UCRL-16870

University of California
Ernest O. Lawrence Radiation Laboratory

TWO-WEEK LOAN COPY

This is a Library Circulating Copy which may be borrowed for two weeks. For a personal retention copy, call Tech. Info. Division, Ext. 5545

THE NUMERICAL SOLUTION OF THE PROBLEM OF A SLOW CONDUCTION ELECTRON IN AN IONIC CRYSTAL (POLARON)

Berkeley, California

DISCLAIMER

This document was prepared as an account of work sponsored by the United States Government. While this document is believed to contain correct information, neither the United States Government nor any agency thereof, nor the Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or the Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof or the Regents of the University of California.

UCRL-16870

Research and Development

UNIVERSITY OF CALIFORNIA
Lawrence Radiation Laboratory
Berkeley, California

AEC Contract No. W-7405-eng-48

THE NUMERICAL SOLUTION OF THE PROBLEM OF A SLOW CONDUCTION ELECTRON IN AN IONIC CRYSTAL (POLARON)

Walter Stephen Zimmermann
(Ph. D. thesis)

April 15, 1966

THE NUMERICAL SOLUTION OF THE PROBLEM OF A SLOW CONDUCTION ELECTRON IN AN IONIC CRYSTAL (POLARON)

Walter Stephen Zimmermann

Lawrence Radiation Laboratory
University of California
Berkeley, California

April 15, 1966

ABSTRACT

The object of the computation is to determine the ground-state energy, E_o, and the effective mass, m^*, of the polaron for a range of coupling strengths, g. It is intended to determine these values with greater accuracy than has hitherto been attained, and at the same time to arrive at a quantitative estimate of the accuracy of these results. Because of the second consideration in particular, any ad hoc restrictions or simplifications are studiously avoided.

The procedure is a variational one. The trial state is expanded as a linear combination of basis states chosen in advance. These basis states are carefully chosen to accelerate convergence, and arranged in systematic sequences so that the error from states no included can be estimated.

TABLE OF CONTENTS

	Abstract	i
I.	Introduction	1
II.	The Polaron Problem	6
III.	Previous Solutions and Approximations	16
IV.	Wave Functions and States	26
V.	The Method of Analysis	48
VI.	Numerical Techniques	58
VII.	Results	74
VIII.	Summary and Evaluation	113
	Acknowledgments	122
	Appendix	123
	References	133

The coefficients in the expansion (variational parameters) are determined numerically with the use of advanced computers. The biggest obstacle to achieving accurate results is the limitation imposed by finite computer memory. By choosing numerical procedures carefully adapted to this problem, it is possible to include nearly 1000 states in the fundamental expansion of the trial state.

It was possible to determine the ground-state energy, E_o, for an important but restricted range of coupling strengths, including most of the so-called intermediate-coupling region. The results were generally lower (i.e. better) than the results of other variational computations. The estimated fractional error in E_o ranged from 0.1 per-cent (weaker coupling) to 1.0 per-cent. These estimated errors are usefully small, but at the same time larger than one would have anticipated or hoped from a computation on this scale. At the same time, the possibility of making error-estimates effectively gives also (approximate) lower bounds, an advantage not shared by most previous computations.

The effective mass, m^*, was determined for the same range of coupling strengths. The accuracy of these results is more difficult to estimate, because they do not have the characteristic of being a bound (e.g. upper bound), and because of numerical considerations.

I. INTRODUCTION

The name polaron refers to quantized excitations of the system consisting of a conduction electron interacting with the lattice of an ionic crystal through a polarization field. More precisely, "polaron" refers to a particular simplified model of such an interaction. This model is due to Frohlich[1] and the Hamiltonian corresponding to this model is referred to in the literature as the "Frohlich Hamiltonian."

The Frohlich Hamiltonian describes an interaction whose strength is characterized by a dimensionless coupling constant, g.* For crystals of physical interest, the coupling strength varies from the region of weak coupling ($g^2 < 1$) to coupling strengths of order $g^2 = 10$. As a mathematical problem, the polaron is well-defined for all coupling strengths.

In addition to the coupling constant, g, the "crystal momentum" $\overline{P}$ appears as a parameter when the Hamiltonian is written in a suitable form. The dependence of the ground-state energy of this system on the momentum P, for small P, is conventionally described in terms of the "polaron effective mass" m*, which is defined as

$$m^* = \frac{1}{2} (dE/dP^2)^{-1} \tag{1.1}$$

so the ground state energy E may then be written

---

* Some authors conventionally use α to denote the coupling strength. The correspondence is $\alpha = g^2$.

$$E = E_o + P^2/2m^* \tag{1.2}$$

for small P.

Of the various properties of the system that may be predicted by theory, the effective mass is of particular interest from the physical point of view. This quantity enters into the mobility, μ, through the relation

$$\mu = e\tau/m^* \tag{1.3}$$

and into the cyclotron resonance frequency, ω_c, through the relation

$$\omega_c = eH/m^*c \tag{1.4}$$

Through these observables, m* provides the connection between theory and experiment.

The polaron problem may be viewed apart from its physical motivation as an eigenvalue problem in mathematical physics. The Frohlich Hamiltonian has in the past served as an instructive testing-ground for computational techniques, and it is primarily in this spirit that we shall attack the problem. We shall, nevertheless, avoid introducing any arbitrary or unphysical assumptions, so that our results may be useful to the physicist and the experimentalist.

The polaron problem attracts our attention because it describes a non-trivial field-theoretic problem which is free from divergences and which exhibits some of the chracteristics and difficulties of a wider class of problems. There are several more or less distinct regions of coupling strength, each of which calls for its own methods of attack. Finally, there are an unusually large number of approximations, or modifications, of the polaron problem, each of which is solvable, or has solvable aspects, and which shed some light on the real problem.

If the literature of the polaron problem cannot be described as profound, it can fairly be described as a rich source of original and creative techniques in computational physics. The most notable of these is perhaps Feynman's application of his path-integral formalism for the numerical solution of the polaron problem.[2] This formalism, an ingenious approach with several mathematical and philosophical ramifications, found in the polaron the first test of its powers as a practical computational tool. Our work will provide new confirmation of the surprising accuracy of Feynman's method.

The strong-coupling theory of the polaron has been described[3] as being "...interesting in its own right, presenting us with a number of fascinating mathematical problems." The strong-coupling theory is an exceedingly complicated field which is a subject in itself and which we will not discuss here.

Between the strong-coupling region, which is a very difficult area, and the weak-coupling region, which is subject to treatment by perturbation theory, there lies the so-called intermediate-coupling region, which is not subject to treatment by the techniques suitable for the two extreme regions. This region can be defined in terms of the conventional dimensionless coupling constant, g, as the region roughly corresponding to

$$1 < g^2 < 9 \tag{1.5}$$

The coupling constant is a property of the crystal, and a great majority of the crystals of interest lie in this region.

It is in the intermediate coupling region that the <u>variational techniques</u> have played the most important role. The simplest of these is motivated by the Tomonaga approximation of meson field theory. This approach was developed concurrently by several authors[4,5] but is now popularly referred to as due to Lee, Low, and Pines.[6] The Lee-Low-Pines calculation has implications for our own study, and will be discussed in some detail in a later section.

A generalization of the Lee-Low-Pines approach is given by Lee and Pines[7], and this study will also be reviewed. The Lee-Pines paper is particularly instructive as a model calculation. Judged as a numerical computation it is limited in its scope and does not begin to draw on the full power of modern computational techniques.

Our work is a variational calculation that begins somwhat in the same spirit as the work of Lee-Pines mentioned above. However, we shall proceed in a more systematic manner, avoiding any unnecessary restrictive assumptions, and preparing to draw on the full power of modern computing techniques if that is where the problem leads us (and we shall soon see that it does!)

After giving a specific definition of the polaron problem (Chapter II), we will review those attempts at solutions that are known and that are relevant to our work here. (Chapter III). In Chapter IV we discuss in some detail the choice of a basis of vectors, or states, from which to construct the trial state for the variational computation. In Chapter V we discuss some principles that underly our analysis, and in Chapter VI we outline the numerical techniques that are needed for a problem such as we will encounter. Finally, the results are presented and discussed in Chapter VII, and a brief summary follows.

II. THE POLARON PROBLEM

In this chapter we summarize some of the physical principles and mathematical manipulations that lead one to the Frohlich Hamiltonian, which will then be the starting-point of our calculation. These remarks are intended neither as original nor as rigorous, but are included to lend perspective to the main body of this paper. In what follows we are guided principally by what Frohlich[8] has written in a recent review.

At the outset several assumptions are made with regard to the nature of the oscillations of the crystal lattice. It is assumed the crystal can be treated as a macroscopic dielectric, characterized by a complex dielectric constant, $\mathcal{E}_\omega$. It is further assumed that the dielectric constant $\mathcal{E}_\omega$ can be regarded as independent of wave-vector, and that is has two characteristic absorption frequencies.

From these assumptions, elementary considerations lead to the concept of an "effective dielectric constant" $\bar{\mathcal{E}}$, which relates the polarization field to the electric displacement D according to

$$P(r) = (1/4\pi\bar{\mathcal{E}})D(r) \tag{2.1}$$

where is defined as

$$1/\bar{\mathcal{E}} = 1/\mathcal{E}_\infty - 1/\mathcal{E} \tag{2.2}$$

Here $\mathcal{E}$ is the static dielectric constant and $\mathcal{E}_\infty$ is the dielectric constant for high frequencies. Frohlich makes the simplifying assumption that the frequency of lattice vibrations, ω, is a constant (independent of wavelength). It is also assumed that in the absence of the polaron interaction (i.e. the interaction of the electron with the polarization field) the electron can be described by an "effective mass" m, that will now be referred to as the bare mass, or simply as the electron mass to distinguish this concept from the polaron effective mass m*. For the purpose of computing the electric field, the electron is regarded as a point charge.

The Frohlich Hamiltonian is derived from a Lagrangian of the form

$$\mathcal{L} = \frac{\mu}{2}(\dot{P}^2(r) - \omega^2 P^2(r)) + D(r)\cdot P(r) \qquad (2.3)$$

Here P(r) is the polarization field, and D(r) is the electric field due to the electron. The parameter μ will be determined shortly. Taking P(r) as the generalized field coordinate q(r), the conjugate momentum is found to be

$$p(r) = \delta\mathcal{L}/\delta\dot{P}(r) = \mu\dot{P}(r) \qquad (2.4)$$

and the Hamiltonian

$$H' = \int d^3r \, \frac{\mu}{2}(\dot{P}^2(r) + \omega^2 P^2(r)) - D(r)\cdot P(r) \qquad (2.5)$$

The equations of motion for the field

$$\dot{q}(r) = \delta H'/\delta p(r) = p(r)/\mu \tag{2.6}$$

$$\dot{p}(r) = -\delta H'/\delta q(r) = -\mu\omega^2 q(r) + D(r) \tag{2.7}$$

lead to the relation

$$\ddot{P}(r) + \omega^2 P(r) = (1/\mu)D(r) \tag{2.8}$$

This form shows that in the absence of the electron (D=0) the polarization field is described by the equations of simple harmonic motion, as we would want to require. Consider (2.8) in the static limit $\dot{P} = 0$, and referring to (2.1) we arrive at the expression for the constant μ , hitherto undefined:

$$\mu = 4\pi\bar{\epsilon}/\omega^2 \tag{2.9}$$

Considering the electron as a point charge

$$D(r,r_{el}) = -\nabla_r \frac{e}{|r-r_{el}|} \tag{2.10}$$

The polarization field satisfies curl P = 0 and so may be derived

from a polarization potential, $\phi(r)$:

$$P(r) = (1/4\pi)\nabla_r \phi(r) \tag{2.11}$$

After an integration by parts, the interaction term of the Hamiltonian,

$$H_{int} = -\int d^3r D(r)\cdot P(r) \tag{2.12}$$

can how be written

$$H_{int} = e\phi(r) \tag{2.13}$$

To take into account the electron kinetic energy we modify the Hamiltonian H' (2.5) by the addition of a term so that our Hamiltonian H is now given by

$$H = H' + p_{el}^2/2m \tag{2.14}$$

The electron equations of motion are then easily found:

$$\dot{r}_{el} = p_{el}/m \tag{2.15}$$

$$\dot{p}_{el} = -e\nabla_{el}\ \phi(r_{el}) \tag{2.16}$$

Before proceeding with quantization, it is convenient to introduce temporarily a cubic volume V, and to apply periodic boundary conditions with respect to this volume.

At this point it is conventional in similar problems to Fourier-analyze the fields and to express the commutation relations of the operators in terms of these Fourier components. In our case, the condition that P(r) be real introduces an auxiliary relation which has the effect that the Fourier components are no longer independent. The resulting difficulties are circumvented by the introduction of a complex field, B(r), which has independent Fourier components. Following Frohlich, we write

$$B(r) = \sqrt{\frac{\mu\omega}{2\hbar}}\left(P(r) + \frac{i}{\omega}\dot{P}(r)\right) \tag{2.17}$$

As curl(P) = 0, B(r) has a Fourier-composition that can be written in the form

$$B(r,t) = \frac{1}{\sqrt{V}}\sum \frac{\bar{k}}{|k|} b_k \, e^{i\bar{k}\cdot\bar{r}} \tag{2.18}$$

One can then show straightforwardly that the commutation relations of the B fields

$$\left[B_j(r),\ B^{\dagger}_{j'}(r')\right] = \delta_{jj'}\,\delta(r-r') \tag{2.19}$$

lead in turn to the relations

$$\left[b_k, b_{k'}^{\dagger} \right] = \delta_{kk'} \tag{2.200}$$

$$\left[b_k^{\dagger}, b_{k'}^{\dagger} \right] = \left[b_k, b_{k'} \right] = 0 \tag{2.21}$$

The Hamiltonian (2.14) can then be written

$$H = H_F + H_{el} + H_{int} \tag{2.22}$$

where

$$H_F = \hbar \sum \omega \, b_k^{\dagger} b_k \tag{2.23}$$

$$H_{el} = -(\hbar^2/2m) \nabla_{r_{el}}^2 \tag{2.24}$$

$$H_{int} = 4\, e(\hbar/2V\omega\mu)^{1/2} i \sum \frac{1}{k} \left(b_k e^{-ik\cdot r} - b_k e^{ik\cdot r} \right) \tag{2.25}$$

The factor $4\pi e(\hbar/2V\omega\mu)^{1/2}$ is conventionally written

$$\hbar\omega(h/2m\omega)^{1/4} \, (4\pi g^2/V)^{1/2} \tag{2.26}$$

where g is the dimensionless coupling constant referred to earlier. The constant g is then given by

$$g^2 = \frac{2\pi e^2}{\mu \hbar \omega^3} \sqrt{\frac{m}{2h\omega}} \tag{2.27}$$

Using (2.2) and (2.9) we may then write

$$g^2 = \frac{e^2}{\hbar}\left(\frac{1}{\varepsilon_\infty} - \frac{1}{\varepsilon}\right)\sqrt{\frac{m}{2\hbar\omega}} \tag{2.28}$$

Here

e = charge of the electron

ε_∞ = high-frequency dielectric constant

ε = static dielectric constant

$\hbar$ = Planck's constant / 2π

m = "bare mass" of the electron

ω = frequency of lattice oscillations (longitudinal modes)

It will be more convenient for our purposes to consider the Hamiltonian in the form it takes in the limit V . As periodic boundary conditions lead to the summation over $(2\quad)^3$ points in k-space per unit volume, the replacement

$$\frac{1}{V}\sum \rightarrow \frac{1}{(2\pi)^3}\int d^3k \tag{2.29}$$

gives the properly normalized form in this limit. We will also want to remove the awkward complex factor in the interaction by making the simple unitary transformation

$$a_k = ib_k\,, \qquad a_k^\dagger = -ib_k^\dagger \tag{2.30}$$

With these changes the three terms of the Hamiltonian (2.22) can be written

$$H_F = \int d^3k \; a_k^\dagger a_k \tag{2.31}$$

$$H_{el} = -(1/2m)\nabla^2_{r_{el}} \tag{2.32}$$

$$H_{int} = \int d^3k (V^*(k) a_k^\dagger e^{ik\cdot r} + V(k) a(k) e^{-ik\cdot r}) \tag{2.33}$$

where we have chosen units in which $\hbar = 1$. We have now

$$V(k) = -\frac{\omega}{k}(2m\omega)^{-1/4}\left(\frac{4\pi g^2}{(2\pi)^3}\right)^{1/2} \tag{2.34}$$

For computational purposes it is convenient to transform the Hamiltonian to a new form in which the coordinates of the electron, $\bar{r}_{el}$, do not appear. Consider the unitary operator

$$U = \exp(-i\int d^3k \; a_k^\dagger a_k \; k\cdot r) \tag{2.35}$$

The transformed operators

$$\bar{a}_k = U^{-1} a_k U = a_k e^{ik\cdot r} \tag{2.36}$$

$$\bar{p} = U^{-1} p U = p + \int d^3k \; a_k a_k \; \bar{k}$$

lead to the transformed Hamiltonian

$$H = \left(p - \int d^3k \, a_k^\dagger a_k \, k \right)^2/2m + \int d^3k \, \omega \, a_k a_k$$
$$+ \int d^3k \left(V(k)a(k) + V^*(k)a_k^\dagger \right) \qquad (2.38)$$

where we have dropped the "bars" indicating the operators are transformed. One can verify that $\bar{p}$ (now written p) as defined by (2.38) commutes with the Hamiltonian. One may then choose the eigenstate of the Hamiltonian to be an eigenstate of p, with eigenvalue **P**. Making this choice, and expanding the first term of (2.38), one can write

$$H = P^2/2m - \int d^3k \, a_k^\dagger a_k \, k$$
$$+ \int d^3k \int d^3k' \, a_k^\dagger a_{k'}^\dagger a_k a_{k'} \, (k \cdot k') \qquad (2.39)$$
$$+ d^3k \, (\omega + k^2/2m) \, a_k^\dagger a_k + \int d^3k \left(V(k)a_k + V^*(k)a_k^\dagger \right)$$

We have used the commutation relations

$$\left[a_k, a_{k'} \right] = \delta_{kk'} \qquad (2.40)$$

to write the creation operators to the left of the destruction operators in the third term of (2.39).

The Hamiltonian (2.39) is the particular form of the Frohlich Hamiltonian with which we will begin our calculation.

III. PREVIOUS SOLUTIONS AND APPROXIMATIONS

In this chapter we will review some previously-published attacks on the polaron problem which are precedents for our work, or which are of fundamental importance.

The simplest approach to the polaron treats the interaction term

$$H_{int} = \int d^3k \, (V(k)a_k + V^*(k)a_k^\dagger) \tag{3.1}$$

as a perturbation. In the absence of the interaction (3.1) the ground state of the system is the vacuum state:

$$/\phi_o) = /0) \tag{3.2}$$

The interaction (3.1) allows transitions to one-phonon states of all momenta, k:

$$/\phi_k) = a_k^\dagger /0) \tag{3.3}$$

The first non-vanishing contribution to the energy is the second-order perturbation theory term:

$$\Delta E = -\int d^3k \frac{V^2(k)}{(\omega + k^2/2m - k \cdot P/m)} \tag{3.4}$$

This form can be easily integrated, and one finds for small P

$$E = -\omega g^2 - (g^2/6)(P^2/2m) \tag{3.5}$$

According to (1.1), this corresponds to an effective mass

$$m^*/m = \frac{1}{(1 - g^2/6)} \tag{3.6}$$

A second approach, based on the intermediate-coupling approximation of Tomonaga, is the calculation of Lee-Low-Pines[6], to which we have referred. The Tomonaga approximation is characterized by the assumption that the eigenstate $/\emptyset)$ has the property

$$(k_1, k_2, \ldots k_n \,/\, \emptyset\,) = c_n \prod_1^n f(k_j) \tag{3.7}$$

This theory assumes there is a phonon mode, defined by $f(k)$, which can usefully describe all phonons, and that correlations between phonons may be neglected. The object of the calculation is then to find the functional form $f(k)$ that gives the lowest ground-state energy.

The Lee-Low-Pines calculation begins with the trial state

$$/\emptyset) = \exp\left[i \int d^3k \, (a_k^+ f(k) + a_k f^*(k) \right] /0) \tag{3.8}$$

which is in fact a special case satisfying (3.7). The function $f(k)$ remains to be determined.

It is suggested that (3.8) be viewed as a unitary transformation from the vacuum state $/0)$ to a new state $/\emptyset)$ with the unitary operator

$$U = \exp\left(i\int d^3k\,\left(a_k^{\dagger} f(k) + a_k f^*(k)\right)\right) \tag{3.9}$$

This transformation yields

$$\bar{a}_k = U^{-1} a_k U = a_k + f^*(k) \tag{3.10}$$

and is referred to as a <u>translation</u> by $f(k)$

The problem may now be viewed as the problem of finding the form $f(k)$ that minimizes the expectation value of the <u>transformed</u> Hamiltonian $\bar{H} = U^{-1}HU$ in the <u>vacuum state</u> $/0)$. This computation is easily carried out, and one finds

$$f(k) = \int \frac{-V^*(k)}{(\omega + k^2/2m + (q-1)k\cdot P/m)} \tag{3.11}$$

If $P=0$ the third term in the denominator does not arise. If $P\neq 0$ q is defined by the relation

$$\int d^3k\; f(k)\; \vec{k} = q\vec{P} \tag{3.12}$$

Considerations of symmetry dictate that a relation of this form should exist. One finds for small P

$$q = (g^2/6)/(1+g^2/6) + O(P^2) + \ldots \tag{3.13}$$

and

$$E = -g^2\omega + (P^2/2m)(1 + g^2/6)^{-1} \tag{3.14}$$

from which the effective mass

$$m^*/m = 1 + g^2/6 \tag{3.15}$$

may be derived.

The accuracy of this calculation may be judged by comparing the results with those of more sophisticated calculations, some of which we shall discuss shortly. Note that E_o, the ground-state energy for P=0, is the same as that given by perturbation theory (3.5).

For the case P=0 the Lee-Low-Pines procedure reduces the polaron problem to one formally equivalent to a fixed-source neutral scalar field model--an elementary problem in field theory. The Hamiltonian defining this problem is of the form

$$H' = \int d^3k\, W(k)\, a_k^\dagger a_k + \int d^3k\, (V(k)a_k + V^*(k)a_k^\dagger) \tag{3.16}$$

If we take $W(k) = \omega + k^2/2m$ and identify $V(k)$ above with our previous

definition (2.34), the Frohlich Hamiltonian (2.39) only by the presence of the additional term

$$H_Q = \int d^3k \int d^3k' \; a_k^\dagger a_{k'}^\dagger a_k a_{k'} \; (\bar{k}\cdot\bar{k}') \tag{3.17}$$

One can attempt to take into account the term (3.17) by considering it as a perturbation, starting from the Lee-Low-Pines state (3.8) as the zero-order solution. This is possible because the state (3.8) (with f(k) given by (3.11)) is in fact an <u>exact</u> solution of the unperturbed "neutral scalar field"-type Hamiltonian (i.e., of the Frohlich Hamiltonian (2.39) minus the "perturbation" (3.17)). The result of this perturbation calculation gives an energy shift

$$E = -0.0159\, g^4 \omega \tag{3.18}$$

so the ground-state energy (for P=0) is given by

$$E_o = -\omega g^2 - 0.0159 \omega g^4 \tag{3.19}$$

We will refer to this computation as the "second" perturbation calculation, to distinguish it from the first calculation we discussed, which treats the interaction term (3.1) as the perturbation.

TABLE 1: Summary of Results

		g^2 =	3	5	7	9
1.	Perturbation Theory	E_o =	-3.00	-5.00	-7.00	-9.00
		m^*/m =	2.00	6.00	-	-
2.	Lee-Low-Pines[6]		-3.00	-5.00	-7.00	-9.00
			1.50	1.83	2.17	2.50
3.	Lee-Low-Pines with correction*		-3.14	-5.40	-7.78	-10.29
4.	Lee-Pines[7]		-3.10	-5.30	-7.58	-9.95
			1.61	2.15	2.82	3.58
5.	Feynman[2]		-3.13	-5.44	-8.11	-11.49
			1.78	3.56	13.2	59.2
6.	Zimmermann**		-3.16	-5.51	-8.04	-10.72
			1.77	2.59	3.24	

* see equation (3.19). Not upper-bound.

** Results of this study. See Chapter VII.

We will see that this result has greater accuracy, and range of validity (in terms of coupling-strength g) than one might be led to conclude from its perturbation-theoretic origin.* Note in this connection, however, that the term (3.17) (the "perturbation") is <u>not</u> characterized by g as a coupling strength.

The point of view discussed here, which treats (3.17) as a perturbation, will be useful to us in our later work. In selecting and ordering the basis of states, we will be guided by the successively higher orders of approximation to the eigenstate, in the sense of this calculation. Later discussions will clarify these remarks.

A more general variational calculation, in the same spirit as that of Lee-Low-Pines, is described in the paper by Lee and Pines[7]. The Lee-Pines approach is a generalization of Lee-Low-Pines in that the phonon may be emitted into either an S-state (as before) or one of three P-states. With certain additional assumptions it is possible to carry through the variational calculation analytically. The numerical results are tabulated in the fourth row of <u>Table 1</u>.

One assumption made by Lee-Pines is that the eigenstate $/\emptyset)$ may be usefully approximated by a state which is a product of an S-wave factor and a P-wave factor. Specifically, the trial function used by Lee-Pines is of the form

---

* See row 3. of <u>Table 1</u>.

$$/\phi) = \exp\left[i \int d^3k \; (f(k) a_k^{+} + f^*(k) a_k) \right] /\phi') \qquad (3.20)$$

where the state $/\phi')$ contains no S-phonons, but only a distribution (to be determined) of P-phonons. Lee-Pines justify this special form on the grounds that the resulting solution can be shown to be <u>exact</u> in the weak-coupling limit ($g^2 \ll 1$), and in the strong-coupling limit for the problem defined with a momentum cut-off, K. The latter observation is an interesting one. However, the introduction of a cut-off in the integrals is a non-trivial modification of the problem, and the significance of the results for the problem without cut-off is not clear.

The introduction of a cut-off in effect creates a new kind of strong-coupling region that has no immediate relation to what is conventionally referred to as the "strong coupling region." Certain integrals that occur in the solution of Lee-Pines, and which are there neglected as being negligible next to $1/g^2$ (in the limit $g^2 \to \infty$) are in fact integrals that diverge (in one case as steeply as K^5) in the limit $K \to \infty$. The lee-Pines strong-coupling solution then describes a region that is "strong" in the sense that

$$K^5/g^2 \ll 1 \qquad (3.21)$$

This is a region that does not exist at all for the problem as we are considering it.

The results of Lee-Pines in the intermediate-coupling region are a considerable improvement over the earlier results of Lee-Low-Pines. In order to judge the accuracy of the Lee-Pines results, we turn to still another approach: the highly original attack on the polaron problem by Feynman.

Feynman approaches the polaron problem by recasting it in terms of his "path integral" formalism. In this formalism, the quantum-mechanical kernel, or Green's function, plays a central role. The kernel is defined by Feynman in terms of a path-integral, or sum over paths, where the integrand is the exponential of the classical action (considered as a functional of the path). This fundamental definition is usually written in the form

$$K(x',t';x,t) = \int \mathcal{D}x(t) \exp\left(\frac{i}{h} \int \mathcal{L}(x,\dot{x},t)\, dt \right) \qquad (3.22)$$

The reader is referred to the original paper of Feynman[9] or to any of several review articles[10,11] for a detailed study of the meaning of this expression.

In the problem at hand one doesn't in fact need the kernel explicitly. The ground-state energy can be shown to be given by the asymptotic decay-rate of the Green's function for large imaginary times:

$$E = -\lim_{t\to\infty} \frac{h}{t} \log K(x', -it; x, 0) \qquad (3.23)$$

Including this limit early in the calculation permits simplifying manipulations.

Unfortunately, the class of path-integrals that can be explicitly computed is very small, and the integrals arising in this problem are not included in that class. Feynman circumvents this difficulty by introducing a novel variational principle. A simplified "model problem" is introduced, which is presumed to represent a decent approximation, in some sense, to the real problem, and for which the corresponding path-integrals can be computed. Feynman shows how the solution of this model problem can be used to get an upper-bound on the ground-state energy of the real problem. The upper-bound characteristic of the approximation is based on the inequality

$$\langle e^{x} \rangle \geq e^{\langle x \rangle} \tag{3.24}$$

where $\langle\ \rangle$ denotes the average value.

The Feynman procedure involves no numerical work until the end of the calculation, where a function (involving one integral that must be solved numerically) of two variables must be minimized with respect to those two variables. This procedure leads to the results tabulated in the fifth row of Table I. The specific figures given are due to Schultz[12].

Keeping in mind the variational character of all three sets of results (i.e. Lee-Low-Pines, Lee-Pines, and Feynman), it is clear that Feynman's are everywhere superior, most markedly in the region of stronger coupling.

Unlike the work of Lee-Low-Pines, Feynman's approach does not seem to admit an obvious generalization that might be used to get an estimate of accurady. If one is inclined to be suspicious, one can point out that the "model problem" introduced by Feynman seems inadequately to reflect the true complexity of the real problem. On the mathematical side, the inequality (3.24) may be too "weak" for accurate numerical results. The weight of the available evidence, including the results of our own calculation, seems to show that these objections carry little force, and attests to the quite remarkable accuracy of Feynman's results over a wide range of coupling constants.

IV. WAVE FUNCTIONS AND STATES

The calculation which we propose to do is, like those of Lee-Low-Pines and Feynman, a variational calculation. Unlike the previous attacks on the polaron problem, we will be prepared to lean heavily on large-scale numerical techniques. We hope in this way to obtain more accurate results and, at the same time, get an estimate of the probable error of these results. Particularly because of the second requirement, we will proceed in as thorough and systematic a manner as possible. We shall avoid introducing arbitrary assumptions or restrictions whose effect on the results cannot be estimated. In particular, we shall introduce no arbitrary restrictions on the size of the basis (such as the restriction to S and P states), except as is determined by limitations of computer space and time.

Accurate results will require a trial state with a very large number of parameters. These parameters will be determined, as always, by the requirement that the Rayleigh quotient $(\phi,H\phi)/(\phi,\phi)$ be stationary. The equations resulting from this condition can, in general, be solved efficiently and accurately only if the variational parameters are <u>linear</u>. That is, our trial state $/\phi)$ should be expressed in terms of a suitable basis of states $/j),(j=1,2,\ldots)$ by a linear expansion

$$/\phi) = \sum_{1}^{n} c_j /j) \tag{4.1}$$

We will refer to equation (4.1) as the "fundamental expansion."

If n terms of this sum are included, one is led to a system of n linear equations. A variety of techniques are known for the numerical solution of such a system, and those appropriate to our problem will be discussed in a later section.

A necessary step in setting up the calculation is then to construct a suitable basis of states $/j)$ $(j=1,2,\ldots)$ to use in the fundamental expansion (4.1). Before constructing such a basis, one needs to choose a representation for the one-phonon wave functions. The many-phonon states $/j)$ are then constructed in the conventional manner by specifying occupations (numbers of phonons) in each of the various states, or modes, each mode being defined in terms of a set of quantum numbers.

The problem as defined by Frohlich is set up in terms of a representation where the phonon wave function u(k) are given by

$$u(k) \simeq e^{ik\cdot r} \tag{4.2}$$

i.e., in the momentum representation. This representation is not useful for our work. What we must do is choose a new complete set of functions u(k) which play the role of (4.2) and which are better suited to our purposes.

In the same spirit as Lee-Pines, we will choose to describe the angular dependence of our phonon wave-functions in terms of an angular-momentum representation. That is, we write

$$u(\vec{k}) = R(k)Y_{LM}(\Omega) \tag{4.3}$$

where Y_{LM} are the spherical harmonics, and Ω represents the pair of angles θ,ϕ. The radial function $R(k)$ is then given by

$$R(k) = \int d\Omega\, Y_{LM}(\Omega)u(\vec{k}) \tag{4.4}$$

R will be independent of M, but will, in general, depend on L, in addition to a "radial quantum number" N. We write then

$$u_{NLM}(\vec{k}) = R_{NL}(k)Y_{LM}(\Omega) \tag{4.5}$$

where the complete set $R_{NL}(k)$ must now be defined.

The calculation of Lee-Low-Pines, which we have discussed in some detail, can be looked at as a special case of the above, with phonons restricted to spherically symmetric states (L=0) and to one radial state (the "radial ground state") N=0. In the same sense, the Lee-Pines calculation corresponds to the restriction N=0, L=0,1 (ignoring for the moment certain other special assumptions made by Lee-Pines).

Because of limitations of computer size, and other factors, it will be necessary to restrict the occupations of states to only a few leading terms from the radial sequence N=0,1,2,.... We will want to choose the ground radial state N=0 to give a respectable approximation to the exact answer, and to choose the remaining ones (N=1,2,...)

if possible, in such a way that only a few terms of the radial series will suffice for an accurate approximation. Before constructing the sequence of radial states R_{NL} in the general case, we turn our attention to the ground radial states R_{OL} for arbitrary L.

It is natural to be guided in this respect by the results of Lee-Pines, and Lee-Low-Pines, because of the close relation between their computation and ours, which we have just pointed out. These calculations determine analytically the optimum functional form of the ground radial states $R_{OL}(k)$ under the restriction that there be only one radial state (Tomonaga approximation), in addition to the restriction on angular momentum.

According to (3.11), the Lee-Low-Pines result (for P=0) suggests that we make the choice

$$R_{00}(k) = \frac{C'_{00}}{k\,(\omega + k^2/2m)} \tag{4.6}$$

It will be contenient to carry out our discussion in terms of the dimensionless unit $x = k/(2m\omega)^{1/2}$, so that (4.6) will be written

$$R_{00}(x) = \frac{C_{00}}{x\,(1 + x^2)} \tag{4.7}$$

The corresponding Lee-Pines results, in contrast, involve three non-linear variational parameters (in addition to the usual normalization

constants). In terms of our present notation, these results suggest

$$R_{00}(x) = \frac{C_{00}\, V(x)\, (x^2+a^2)}{(x^2-b^2)^2 + c^2x^2} \tag{4.8}$$

$$R_{01}(x) = \frac{C_{01}\, x\, V(x)}{(x^2-b^2)^2 + c^2x^2}$$

$$= \frac{C_{01}}{C_{00}} \frac{x}{(x^2+a^2)} R_{00}(x) \tag{4.9}$$

As we have pointed out, we will want to work principally with linear variational parameters. It will therefore not be convenient to use functions with a complicated non-linear parametric dependence such as (4.8) and (4.9). However, the form of these functions will give us some useful clues.

The most important consideration in the choice of phonon wave-functions is to approximate the analytic properties of the solution (insofar as they are known, or can be estimated) as well as possible. In particular, we are concerned with singularities and limiting properties. It is important to study the asymptotic functional forms of $\lim_{x\to 0} R_{0L}(x)$ and $\lim_{x\to\infty} R_{0L}(x)$ and to choose these to match the corresponding properties of the solution. What are the known properties of the solution? On the basis of the Lee-Pines calculation, (4.8) tells us

$$\lim_{x \to 0} R_{00}(x) \simeq V(x) \simeq 1/x \tag{4.10}$$

and from (4.9)

$$\lim_{x \to 0} R_{01}(x) \simeq xV(x) \simeq 1 \tag{4.11}$$

These results suggest

$$\lim_{x \to 0} R_{0L}(x) \simeq x^{L}V(x) \tag{4.12}$$

Similarly,

$$\lim_{x \to \infty} R_{00}(x) \simeq V(x)/x^2 \simeq 1/x^3 \tag{4.13}$$

and

$$\lim_{x \to \infty} R_{01}(x) \simeq R_{00}(x)/x \simeq 1/x^4 \tag{4.14}$$

These results for large x suggest the pattern

$$\lim_{x \to \infty} R_{0L}(x) \simeq x^{-L} V(x) / x^2 \tag{4.15}$$

Conditions (4.12) and (4.15) can be satisfied simultaneously if we assume that the asumptotic behavior of $R_{OL}(x)$ is governed in both extreme regions by a factor of the form

$$R_{OL}(x) \simeq \frac{x^{L-1}}{(1+x^2)^{L+1}} \left(\frac{x}{(1+x^2)} \right)^L \cdot \frac{V(x)}{(1+x^2)} \qquad (4.16)$$

This same factor is suggested by a perturbation-theoretic argument similar to the "second" perturbation-theoretic point of view, discussed earlier. We tentatively accept (4.16) as a condition to be imposed on the radial functions $R_{OL}(x)$.

We are faced, then, with constructing a set of functions $R_{OL}(x)$ with the asymptotic behavior described by (4.16), and which coincide with, or closely approximate, the results of Lee-Pines for the case L=0 and L=1. These and other minor considerations lead us to make the following choice for the leading radial functions $R_{OL}(x)$:

$$R_{OL}(x) = \left(\frac{x}{(1+Z^2x^2)} \right)^L \frac{C_{OL}}{x(1+Z^2x^2)} \qquad (4.17)$$

Here Z is a free parameter, analogous to the parameters a,b,c

appearing in the Lee-Pines solutions (4.18) and (4.19). It is in the nature of a scale factor, and it is intended that Z will initially play a minor role in the calculation. In fact, we begin with Z=1, so that (4.17) coincides with the Lee-Low-Pines result for L=0. After preliminary calculations, we may make adjustments in Z to further optimize our results. The function (4.17) has the simplicity required for the kind of numerical techniques we will need to use, with a small concession (i.e. the scale factor Z) to the flexibility of the three-parameter functions of Lee-Pines.

We have now made a choice for the ground radial functions $R_{OL}(x)$. The next step is to define the complete set $R_{NL}(x)$ for arbitary N. It is convenient to choose the radial functions $R_{NL}(x)$ in the form

$$R_{NL}(x) = C_{NL}\left(\frac{x}{(1 + Z^2x^2)}\right)^L \frac{1}{(1 + Z^2x^2)x} P_{NL} \tag{4.18}$$

where P_{NL} are a set of complete polynomials over the interval ($0 \leq x < \infty$). We intentionally omit the argument of P_{NL} because we will want to make a new choice of independent variable, $v(x)$. We turn now to this question.

In order to preserve conditions (4.12) and (4.15) we will want

$$\lim_{x \to \infty} P_{NL}(v(x)) = \text{constant} \tag{4.19}$$

Also, $v(x)$ should be of such a functional form that the integrals arising in our calculation will not themselves require numerical treatment. We find these requirements are met by the choice

$$v(x) = 1/(1 + Z^2x^2) \tag{4.21}$$

The radial functions $R_{NL}(x)$ then have the form

$$R_{NL}(x) = C_{NL}\left(\frac{x}{(1 + Z^2x^2)}\right)^L \frac{1}{x(1 + Z^2x^2)} P_{NL}(v(x)) \tag{4.22}$$

It remains to determine exactly what polynomials $P_{NL}(v)$ to use. This choice is determined by the condition of orthonormality:

$$\delta_{NN'} = \int_0^\infty x^2\, dx\, R_{NL}(x)\, R_{N'L}(x) \tag{4.23}$$

In terms of the new variable $v(x)$,

$$\delta_{NN'} = \frac{C_{NL}C_{N'L}}{2Z^{L+L'+1}} \int_0^1 dv(1-v)^{L-1/2}\, v^{L+1/2} P_{NL}(v)P_{N'L}(v) \tag{4.24}$$

The "weight factor" $\rho(v)$ of the form

$$\rho(v) = (1-v)^a\, v^b \tag{4.25}$$

on the interval (0,1) defines the <u>Jacobi Polynomial</u> $G_n(p,q,v)$

with

$$p = a + b + 1, \qquad q = b + 1 \tag{4.26}$$

Accordingly

$$P_{NL}(v) = G_N(2L+1,\ L+3/2;\ v) \tag{4.27}$$

For the explicit definition and properties of these polynomials, the reader is referred to Appendix A.

This completes the task of constructing a representation of phonon wave functions. The functions we have chosen are described by three quantum numbers: the radial, or principal quantum number, N; the angular momentum L; and the z-projection of angular momentum, M. We will want to order these states in some specific way so they can be referred to by a single index, i, rather than by the corresponding triplet of indices N_i, L_i, M_i. A particular ordering that will be convenient for our later purposes is the one given in Table 2. The list is constructed for the case where the allowed states correspond to L=0,1,2,3, and N=0,1,2,3,4. Within the limits of computer memory, this description was found by experience to be close to the optimal choice. In any special case where we may want to use another description (such as including a higher range of L but a smaller range of N), we will order the states according to the same general scheme.

TABLE 2: Ordering of Phonon Modes *

		N:	0	1	2	3	4	5
M	L							
0	0		1	17	33	49	65	81
0	1		2	18	34	50	66	82
0	2		3	19	35	51	67	83
0	3		4	20	36	52	68	84
-1	1		5	21	37	53	69	85
-1	2		6	22	38	54	70	86
-1	3		7	23	39	55	71	87
+1	1		8	24	40	56	72	88
+1	2		9	25	41	57	73	89
+1	3		10	26	42	58	74	90
-2	2		11	27	43	59	75	91
-2	3		12	28	44	60	76	92
+2	2		13	29	45	61	77	93
+2	3		14	30	46	62	78	94
-3	3		15	31	47	63	79	95
+3	3		16	32	48	64	80	96

* The figures in the main body of the table give the sequence number of the mode with the particular L,M (left hand columns) and particular N (above). For example, the M=0, L=1, N=1 mode is 18-th mode.

Corresponding to the eigenfunctions we have constructed, we now introduce a new set of creation and destruction operators. Suppose that L_i, M_i, N_i is the i-th triplet of quantum numbers as ordered in Table 2. We then write $u_{N_i L_i M_i}(x)$ simply as $u_i(x)$ and we define

$$A_i = \int u_i^*(x)\, a_x\, d^3x \tag{4.28}$$

From the ortho-normality of the $u_i(x)$ it follows

$$a_x = \sum u_i(x)\, A_i \tag{4.29}$$

Using the commutation relations

$$\left[a_x, a_x^\dagger \right] = \delta(x-x') \tag{4.30}$$

one verifies

$$\left[A_i, A_{i'}^\dagger \right] = \delta_{i,i'} \tag{4.31}$$

Substituting (4.29) in the Hamiltonian (2.39) one finds

$$H = H_0 + H_{int} + H_Q + H_P \tag{4.32}$$

where

$$H_0 = \sum_{i,j} T_{ij} A_i^\dagger A_j \tag{4.33}$$

$$H_{int} = \sum_i (J_i A_i + J_i^* A_i^\dagger) \tag{4.34}$$

$$H_Q = \sum_{i,j,k,l} Q(i,j,k,l) A_i^\dagger A_j^\dagger A_k A_l \tag{4.35}$$

$$H_P = 2P \cdot \sum \overline{K}_{ij} A_i^\dagger A_j \tag{4.36}$$

and where, in turn,

$$T_{ij} = \int d^3x \; u_i^*(\overline{x}) \; (1 + x^2) \; u_j(x) \tag{4.37}$$

$$J_i = \int d^3x \; V(x) \; u_i(\overline{x}) \tag{4.38}$$

$$Q(i,j,k,l) = \int d^3x \int d^3x' \; u_i^*(x) \; u_j^*(x') \; \overline{x} \cdot \overline{x}' u_k(x) \; u_l(x') \tag{4.39}$$

$$\overline{K}_{ij} = \int d^3x \; u_i^*(x) \; \overline{x} \; u_j(x) \tag{4.40}$$

Our next task is to choose a set of states to serve as a basis for the fundamental expansion (4.1) of the trial state. Having decided on a specific representation for the phonon modes, we must prescribe what occupations, or combination of occupations, of these states will be included in the expansion. As even a fairly restricted choice for

the range of N,L leads to a very large number of possible states, we will need to be very selective in choosing which particular states to include.

What states will play a role, in principle, is determined by the selection rules for the various integrals (4.37), (4.38), (4.39) and (4.40). These rules are easily determined by elementary considerations and are summarized in Table 3. We observe that each term of the Hamiltonian separately conserves M, the z-component of L. Also each term except (4.40) (which does not play a role of P=0) conserves parity. From the various approximate solutions we see that states with even parity, and with M=0, contribute the important components to the ground-eigenstate. We conclude that the ground state is characterized, in principle, by even parity if P=0, and in any case by M=0. Accordingly, we need to include only states with these characteristics in the expansion (4.1). (If $P \neq 0$ we will need to allow states with odd parity also).

If one were interested exclusively in the case P=0, one could take advantage of the spherical symmetry of this problem by choosing the trial state explicitly as an eigenstate of L^2. This would effectively reduce the number of variational parameters, but would introduce other complications of a programming and computational nature. Because of these complications, and particularly because we are interested in studying the case $P \neq 0$ also (for which spherical

TABLE 3: Selection Rules

$T_{ij} = 0$ unless
$L_i - L_j = 0$ and $M_i - M_j = 0$ and* $P_i/P_j = 1$
$\overline{K}_{ij} = 0$ unless
$L_i - L_j = \pm 1$ and $M_i - M_j = \pm 1$ or 0 ** and $P_i/P_j = -1$
$Q(i,j,k,l) = 0$ unless
all of the following are true
$L_i - L_k = \pm 1$ $\quad M_i - M_k = \pm 1$ or 0
$L_j - L_l = \pm 1$ $\quad M_j - M_l = \pm 1$ or 0 $\quad P_iP_j = P_kP_l$
$M_i + M_j + M_k - M_l = 0$
or all of the following are true
$L_i - L_l = \pm 1$ $\quad M_i - M_l = \pm 1$ or 0
$L_j - L_k = \pm 1$ $\quad M_j - M_k = \pm 1$ or 0 $\quad P_iP_j = P_kP_l$
$M_i + M_j - M_k - M_l = 0$
$J_i = 0$ unless $L_i = 0$

* P_i = parity of i-th mode = $(-1)^{L_i}$

** z-component of $\overline{K}_{ij} = 0$ unless $M_i - M_j = 0$. This is the compoenent that enters into the computation.

symmetry is violated) we will not choose our trial state explicitly as a function of L^2.

We now make a brief digression to discuss a matter of notation. A state can be defined by giving a sequence of occupation-numbers—i.e. by giving the number of phonons in each of the 96 modes. This occupation-number notation is very awkward in a situation where we must concern ourselves with the details of the states, or actually listing the occupation numbers for each of several hundred states. While we encounter up to 96 modes, we ordinarily will encounter states where the great majority of these modes are unoccupied. That is, the number of phonons in a typical many-phonon state will be three or four, and in the largest case, seven. In this circumstance it is more convenient to use a notation in which the state

$$\frac{1}{\sqrt{N!}} A^{+}_{n_1} A^{+}_{n_2} \ldots A^{+}_{n_N} \; /0) \tag{4.41}$$

is written $/n_1, n_2, \ldots n_N)$. Different states may then have different numbers of entries, depending on the number of phonons. If we restrict ourselves to states with less than seven phonons, for example, no state will require more than six numbers for its complete description (instead of 96 numbers for every state, in the alternative notation).

A special case is represented by the vacuum state, which we will continue to write as $/0)$, without risk of confusion.

If this notation is only a matter of convenience for this text, it is essential for storing the information about the states in the computer. We will see that we will encounter bases with up to 1000 states, so that storing the states alone in the occupation-number notation would require 96,000 words--larger than the memory of most computers! In the program used for this computation, a still more condensed notation, based on the above, was used).

We return now to the problem of explicitly constructing these states. We may begin by considering even-parity states only. These states suffice for the special case P=0. The odd parity states can then be constructed according to the same general principles.

The number of states arising from all possible distributions of a moderate number (e.g. three or four) of phonons among 96 states is, of course, astronomical. Fortunately, we are not faced with this general situation. The restriction $M = \sum M_i = 0$ (the sum being taken over <u>phonons</u>) together with the restriction that parity is even--i.e.

$$\prod_{\text{phonons}} (-1)^{L_i} = (-1)^{\sum_L Ln(L)} = 0 \qquad (4.42)$$

enormously reduces the number of states that must be considered. (In (4.42) n(L) denotes the number of phonons in all state characterized by angular momentum L)

Even with these restrictions, the number of states to be considered results in a problem that exceeds the capacity of any computer. It is essential to introduce at the outset hypothetical notions about the relative "importance " of various basis vectors. A vector is said to be "important" if including it in the fundamental expansion (4.1) results in an energy that is significantly lower than the result when the state is not included. Our intuitive ideas about "importance" are then checked by computation and accepted or modified, as the case may be.

We begin the explict denumeration of states by restricting ourselves to the radial ground-state N=0. All other states can then be though of as arising from these states by "radial excitations" and will be relatively easy to describe.

The one-phonon states can be written down by inspection. The condition on parity dictates the one phonon be in a state of even L: i.e. L=0 or L=2. (We do not allow L=4 states). The other condition dictates M=0. The only possible states are then /1) and /3) (see Table 2 to verify state /1) corresponds to L=0, M=0, N=0, and /3) corresponds to L=2, M=0, N=0).

The two-phonon states are only slightly more complicated. The two occupied modes can have the following pairs of quantum numbers L: L=0, L=0; L=1, L=1; L=2,L=2; L=3, L=3; or L=1, L=3. These "L=profiles" are written S^2, P^2, D^2, F^2, and P^1F^1 respectively. Similarly, the possible pairs of quantum numbers M for the two

phonons are, respectively (by pairs): M=0, M=0; M=+1, M=-1; M=+2, M=-2; M=+3, M=-3. The two-phonon states satisfying these conditions are easily enumerated:

/1,1), /2,2), /5,8), /3,3), /6,9), /11,13),
/4,4), /7,10), /12,14), /15,16), /1,3),
/5,10), /8,9).

The states with three or more phonons require an even more systematic approach. It is, of course, possible to assign the task of constructing these states to the computer. That would be easily programmed, but it would be a short-sighted policy, because we could not possibly do the numerical computation with all the states the computer would give us. The classification schemes that we will discover in the course of constructing the list of states will also be useful for choosing which ones to include, and for ordering them in a suitable sequence.

One scheme that suggests itself is to consider first only S states (a trivial case); then S and P states together, and so on. This will be only temporarily useful because we will still need to face the problem of writing down all six-phonon states, for example, with L = 0,1,2,3 admitted. This will give a very large number of states, so it is necessary to begin thinking about relative orders of importance. Suppose that we have proceeded along such a sequence to the point where we are ready to admit D-phonons. It

is natural to consider first states with one D-phonon, then states with two D-phonons, and so on. But among those states with one D-phonon (for example), we will encounter states with (additionally) two P-phonons, others with four P-phonons, and so on. It would be a plausible conjecture that among those states with one D-phonon, those with successively larger numbers of additional P-phonons would be successively less important. (This conjecture will be confirmed by numerical computations). It is then suggested that we invent an index that describes not only the highest L-mode occupied, but also takes into account the number of phonons in this highest L mode and in the other lower L-modes L=0,1,.... A simple and (as it will turn out) very useful index that has this property is the index we shall refer to as the <u>rank</u>, R. This is defined as

$$R = \sum_{L}^{\infty} L\, n(L) \tag{4.43}$$

where, as in (4.42), n(L) is the number of phonons in all modes characterized by angular momentum L. Note that (4.42) then implies that states with even parity have even rank, states with odd parity odd rank.

The rank has an interesting relation to the second perturbation-theoretic calculation we discussed. In this calculation, the zero-order solution coincides with the Lee-Low-Pines solution, and is described entirely in terms of S-phonons. The first-order theory gives rise to states with rank R=2, and the next order to R=4 states

(in addition, of course, to modifying the distribution of R=2 and R=0 states). Succeeding higher orders give rise to states of rank R=6,8,....

As one considers a sequence of trial functions, containing first states of rank R=0 only, then successively states of R=2, 4,6, the n-th trial state includes all basis vectors that would be included in the n-th order perturbation theory. (The converse is not true: there are a few special cases of states of R=2, for example, that correspond to second-order perturbation, not first). The calculations arranged in this way are not, of course, equivalent to perturbation theory (even ignoring the exception just mentioned). The resulting eigenstates will, in general, not be given by the same set of amplitudes for the various basis vectors, as would be dictated by perturbation theory. The energy as determined by the variational calculation will, in general, be lower than the perturbation theoretic result of corresponding order, in the sense described here.

The usefulness of the rank, R, as defined by (4.43) would at first glance appear to be severely impaired by the fact that the rank is insensitive to the number of S-phonons (L=0). The S-phonons in any case will play a special role in our calculation. For other reasons we will want to consider the number of S-phonons as an independent characteristic, or index, of a state, and we will want to study this characteristic separately. We will see below that the problem can be reformulated in such a way that the S-phonons

(particularly the N=0 S-phonons) play a very minor role in the explicit basis. These circumstances together weaken the force of the objection that the rank R is insensitive to S-phonons.

For the present, we make use of the rank strictly as a "bookkeeping device." For states with a prescribed number of phonons, and a prescribed rank, it is possible to write down by inspection the corresponding allowed L-profiles. For example, three-phonon states of rank R=2 must be either S^1P^2 or S^2D^1. The three-phonon states of rank R=4 will be S^1D^2, $S^1P^1F^1$, or P^2D states. For each of these "L-profiles" it is possible to write down the allowed "M-profiles" that satisfy the requirement M=0. From these distributions it is then an easy matter to write down the states explicitly in terms of the occupation numbers (or rather, in terms of the alternative notation defined by (4.41)).

We have already said that S-phonons will turn out to play a special role and (after some manipulations to be discussed in a later section) in fact are relatively unimportant. This circumstance suggests that we further simplify the labor of constructing long lists of states by writing down only those states with no S-phonons. Each state of this list then leads to a sequence of related admissable states by successive additions of S-phonons. As a shorthand device, each state of our abreviated list can be taken to represent itself together with all other states derived from it in the manner described. In this sense, the vacuum state reminds us of

the whole sequence of states /0), /1), /1,1), /1,1,1),..., formed from the vacuum state by the successive additions of N=0 S-phonons.

With this convention, we can easily write down all states with total phonon occupations up to reasonably high numbers and rank up to R=6, for example. Such a list of states is given in Table 4.

In following discussions we will refer to the states given in Table 4 as "skeleton states" because they form a kind of framework, or skeleton, from which the complete list of states (with S-phonons and radial excitations) can immediately be built. Let us look now at just how this may be done.

A "radial excitation" refers to the transfer of a phonon from a state L,M,N to a state L',M',N' with L'=L, M'=M, N'> N. With the particular ordering of states given in Table 2, this always corresponds to a transition from the i-th state to the (i+16)th state (if N'-N=1), to the (i+32)-nd state of N'-N = 2, and so forth. As an example, consider the "skeleton state" /5,8). From this state we can derive by radial excitations the sequence of states /5,24), /8,21), /21,24), /5,40), /8,37), /21,40),/24,37), /37,40),.... Each of these states is characterized by having two phonons in L=1, M=1 and L=1, M=-1, respectively. They differ in their patterns of radial-mode distributions. We expect that these states, derived by successively higher-order radial excitations will be (at least roughly) in descending order of importance. This expectation will be confirmed by our results.

TABLE 4: List of "Skeleton States" by Rank (R)

R = 2	2, 2	2, 3, 4	2, 2, 5, 10
	5, 8	4, 5, 9	2, 8, 8, 12
		4, 6, 8	2, 5, 5, 14
R = 4	3, 3	3, 7, 8	8, 8, 8, 15
	6, 9	2, 7, 9	5, 5, 5, 16
	11, 13	3, 5, 10	3, 5, 5, 13
	2, 2, 3	2, 6, 10	3, 8, 8, 11
	2, 5, 9	8, 9, 12	2, 2, 2, 2, 3
	2, 6, 8	5, 6, 14	2, 2, 3, 5, 8
	3, 5, 8	8, 13, 15	3, 5, 5, 8, 8
	5, 5, 13	5, 11, 16	2, 2, 8, 8, 11
	8, 8, 11	2, 2, 3, 3	2, 2, 5, 5, 13
	2, 2, 2, 2	2, 2, 6, 9	2, 2, 2, 5, 9
	2, 2, 5, 8	2, 2, 11, 13	2, 2, 2, 6, 8
	5, 5, 8, 8	3, 3, 5, 8	2, 5, 5, 8, 9
		5, 6, 8, 9	2, 5, 6, 8, 8
R = 6	4, 4	5, 8, 11, 13	5, 5, 5, 8, 13
	7, 10	2, 2, 2, 4	5, 8, 8, 8, 11
	12, 14	2, 4, 5, 8	2, 2, 2, 2, 2, 2
	15, 16	2, 2, 7, 8	2, 2, 2, 2, 5, 8
			2, 2, 5, 5, 8, 8
			5, 5, 5, 8, 8, 8

A second consequence of admitting radial excitations is that each skeleton state (or, for that matter, each state with radial excitations) leads to several sequences by the addition of ground-state (N=0) <u>or radially-excited</u> (N > 0) S-phonons. The state /5,8) then reminds us of the existence of the states /1,5,8), /17,5,8), /33,5,8),...; /1,1,5,8), /1,17,5,8), /17,17,5,8), /1,33,5,8),...; /1,1,1,5,8), /1,1,17,5,8),.../17,17,17,5,8),..., /33,33,33,5,8),... Of course, there will be a similar sequence for the state /5,24) and for every other state derived by radial excitations from /5,8).

It is clear from these observations that the moderate number of skeleton states (Table 3) can lead to an enormous number of basis states if more than just a very few terms from each of these sequences are included. Unless the sequences of approximate eigenvalues converge extremely well, the size of the basis will have exceeded the maximum size (as determined by computer characteristics) before satisfactory accuracy is achieved. Having devised a systematic procedure by which we can, more or less easily, write down explicitly those states that enter into the calculation (in principle), we must now turn our attention to determining, systematically by numerical computation, what states must be included, in practice, to achieve a given standard of accuracy. More realistically, we will want to optimize the results for a given size basis (the maximum size we can handle) and then to try to estimate the accuracy of the results.

V. THE METHOD OF ANALYSIS

In the previous section we outlined a procedure for the systematic description of a large set of basis vectors. In this connection we found it useful to characterize each state by an index, called the rank. A second characteristic of a given state is the pattern of radial excitations. Still another might be taken to be the number of S-phonons. In order to visualize the process we are about to describe, it may be helpful to think of each of these characteristics (rank, pattern of radial excitations, etc) as a coordinate direction in an abstract vector space, S. The dimensionality of S depends on the number of characteristics we want to focus attention on. Each point in this abstract space, S, then represents a <u>set of states</u>, the characteristics of the set corresponding to the projection of the point onto the various coordinate directions. For example, if the "R-coordinate" is R=2, the set is a set of states with rank R=2. (For most coordinates, or characteristics one may think of, only integer-valued coordinates are permitted).

Practical considerations restrict our basis to a strictly limited number of basis vectors. This situation can be represented, a little loosely, by saying that our basis must correspond to a limited (i.e. not too large) "volume" in our space, S. We will find, for example, that we are able to include states only up to rank R=6 (or, in some circumstances, R=8). That is, we can go only so far out the R-axis. For a given choice in this respect, we will

be limited to states with some restricted pattern of radial excitations. Of course, the restrictions are not independent; if we restrict ourselves more severely in one respect, we can allow more freedom in another. The process of optimizing these choices is a somewhat laborious one, involving repeated computations with different sets of basis vectors, "exploring" each of the various directions in the abstract space S.

In this process we make fundamental use of a notion that is not rigorous but which is extremely useful. This is the notion of the incremental change in the eigenvalue correspoding to a particular state or (more often) set of states. Think of doing a variational calculation with some basis, B_n (consisting of n basis vectors); call the corresponding eigenvalue E_n. Repeating this calculation with some larger basis $B_{n'}$ (which includes B_n), one gets a second (better) approximation $E_{n'}$. The difference $E_{n'}-E_n$ is then thought of as the increment resulting from the set of basis vectors that are included in $B_{n'}$ but not in B_n.

Now strictly speaking, this notion has no meaning, because the increment $E = E_{n'} - E_n$ depends not only on the augmented states $B_{n'}-B_n$ but also on the "background" of states B_n itself. This fact notwithstanding, one finds one can speak as if the increment E were a functional of the set $B_{n'}-B_n$ and that this notion has semiquantitative validity.

We will generally be concerned with sequences of increments $E_1, E_2, \ldots$, and we will find that the properties of these sequences

(such as rate of convergence) will be largely unaffected by the "background" of states which is common to all the bases in the sequence. In this connection, there is one precaution that should be taken. The nested sequence of basis vectors should be arranged so that the corresponding increments in the eigenvalue are monotonically decreasing, or as close as possible to monotonically decreasing. That is to say, we should begin with the most important states, and add sets of states of successively smaller importance. One can be misled if one goes very far down one sequence, studying the role of states of very minor importance, while states of relatively much greater importance (perhaps belonging to some other sequence) have not yet been included. As an extreme example of this situation, one can even reach the apparent conclusion that a particular set of states gives a strictly vanishing contribution to the eigenvalue (when in fact this is not the case). Suppose one excludes from the background of states B_n all states that are connected to the incremental set $B_{n'}-B_n$ by the Hamiltonian. The new set is then "isolated" and the corresponding amplitudes will all vanish. This situation is not likely to arise in practice; it is mentioned to illustrate the role that the "background" of states can play, in principle.

A detailed discussion of numerical results will be the subject for a following chapter. Here we want to anticpate one aspect of the results that was discovered in preliminary computations, and which will require special attention. We refer to the role played by the

phonons in the ground state (L=0, N=0). We enter into this discussion here because the difficulties that arise will require a very important modification of the procedure. This modification will also be reflected in the analysis of numerical results, to be considered in a following chapter.

Consider a particular set of states (e.g. all skeleton states of rank R=4), and then consider the sequence of sets derived from this set by successively adding more ground-state phonons to each state. If there are n states in the original set, there will be n states with one "augmented" S-phonon, n states with two S-phonons, and so on. In the course of studying the convergence of the corresponding sequence of eigenvalues, one finds it necessary to include (for a fairly large class of skeleton states) states with as many as twelve of fourteen S-phonons before even moderate accuracy is achieved. The situation is less severe for weaker coupling ($g^2 \simeq 1$) but becomes intolerable for coupling strengths $g^2 > 5$. As the effect of including (up to) n "augmented" S-phonons is to <u>multiply</u> the size of the basis by n (and the number of matrix elements by n^2), this circumstance could have catastrophic implications for our computation.

Analysis shows that this regretable situation is not really surprising. Recall that the Lee-Low-Pines calculation was concerned with the role of S-phonons, and that our ground state corresponds closely to the (one) state used there. The Lee-Low-Pines solution

gives an amplitude

$$c_n = (n/\emptyset) \simeq \sqrt{\frac{g^{2n}}{n!}} \tag{5.1}$$

for finding exactly n phonons in the ground state. One would not expect to get smooth or rapid convergence until one had included states with 1,2,...n' phonons, where n' is in the asymptotic region (i.e. where $c_{n'}$ is small and remains small for $n > n'$). The close correspondence between the convergence behavior of the numerical results and the predictions of this model suggests that the Lee-Low-Pines calculation is applicable to this analysis.

This insight will also provide us with a means of circumventing the difficulties. Inasmuch as the difficulty is seen to arise from the distribution of ground-state phonons, and inasmuch as the Lee-Low-Pines calculation describes the distribution of phonons in this state (apparently to a good approximation), it is suggested that we incorporate the Lee-Low-Pines work analytically in our procedure before commencing with the numerical computation.

Consider the case P=0, with Z=1, so that we can identify our ground state L=0, N=0 with the Lee-Low-Pines function f(k) as given by (3.11). The Lee-Low-Pines solution (3.8) can be written in our notation

$$/\emptyset) = \exp\left(iW\left(A_1 + A_1^{\dagger}\right)\right) /0) \tag{5.2}$$

A generalization of this result is to write a variational trial state in the form

$$/\phi) = \exp(iW\,(A_1^{\dagger} + A_1)\,)\,/\phi') \tag{5.3}$$

where $/\phi')$ is an arbitrary state to be described by an expansion of the form (4.1). That is, we can regard (5.3) as a unitary transformation from $/\phi)$ to $/\phi')$ and proceed as before.

The factor W arises because the "translation function" f(k) differs from the ground-state $R_{00}(x)$ in that the latter is normalized while the former is not. The Lee-Low-Pines solution corresponds to $W = -(g^2/2)^{1/2}$. One can now regard W as a second non-linear variational parameter (in the same sense as Z).

The unitary operator

$$U = \exp(iW\,(A_1 + A_1)\,) \tag{5.4}$$

generates the transformed operators

$$\bar{A}_i = U^{-1}A_iU = A_i + W\delta_{i,1} \tag{5.5}$$

which equation corresponds to the rule formerly written

$$\bar{a}_x = a_x + f^*(x) \tag{5.6}$$

The transformed Hamiltonian $\bar{H} = U^{-1}HU$ differs from H (in form) by the addition of three sets of terms:

$$\bar{H} = H + H_{Q'} + H_I + H_{int'} \tag{5.7}$$

What is meant by (5.7) is more precisely $\bar{H}(\bar{A}_i, \bar{A}_i) = H(\bar{A}_i, \bar{A}_i) + \ldots$. The term H_I is proportional to the identity operator, I, and is given by

$$H_I = W^2 T_{11} + 2 W J_1 \tag{5.8}$$

The integrals T_{11} and J_1 are defined by (4.37) and (4.38).

The second addition, $H_{Q'}$, has a more complicated form:

$$\begin{aligned} H_{Q'} = W^2 \Big(& \sum_{ij} Q(i,j,1,1) \, (A_i A_j + A_i^\dagger A_j^\dagger) \\ & + Q(i,1,j,1) \, A_i^\dagger A_j \\ & + W \Big(\sum_{ijk} \mathcal{Q}(i,j,k,1) \, (A_i^\dagger A_j^\dagger A_k + A_i^\dagger A_j A_k) \Big) \end{aligned} \tag{5.9}$$

The integrals $\mathcal{Q}(i,j,k,1)$ are defined by (4.39).

Finally, the new interaction-like term $H_{int'}$ is given by

$$H_{int'} = W T_{11} (A_1 + A_1^\dagger) \tag{5.10}$$

An important point is that with a suitable choice of W the term $H_{int'}$ can be made to cancel out the term H_{int} (see (3.1)) exactly. This is the same value of W that minimizes the value H_I , and corresponds to the Lee-Low-Pines solution.

We have restricted our comments to the case P=0 because only in this case does our ground-state $R_{00}(x)$ correspond to the Lee-Low-Pines momentum distribution f(x) (except, as we have said, for normalization). In the general case, P≠0, the operator U (5.4) still generates a unitary transformation that is well defined and that we will find useful. But this is not the same transformation as was used by Lee-Low-Pines for the case P≠0.

For the essentially trivial case $/\emptyset') = /0)$ (in (5.3)) our result coincides with the Lee-Low-Pines calculation. Successive higher approximations (i.e. all non-trivial terms in our computation) can then be looked at as correction to the Lee-Low-Pines solution. In this sense, our computation can be said to begin where the Lee-Low-Pines result stops.

Note that the transformation (5.3) is not equivalent to the Lee-Pines "factored" trial state, given by (3.20), which equation has a similar form. The Lee-Pines state (3.20) includes the implied restriction that the second factor $/\emptyset')$ has no occupations of S-states. We do not make this assumption; our equation (5.3) implies no similar loss of generality.

With the modified trial state (5.3) the Rayleigh quotient $(\emptyset,H\emptyset)/(\emptyset,\emptyset)$ takes the form

$$\frac{(\emptyset U^{-1},\ H\ U\emptyset)}{(\emptyset U^{-1}, U\emptyset)} = \frac{(\emptyset',\ \bar{H}\ \emptyset')}{(\emptyset',\emptyset')} \tag{5.11}$$

which shows that we can use the transformed Hamiltonian $\bar{H}$ and compute matrix elements of $\bar{H}$ between basis states in the same sense as we formerly computed matrix elements of H. Our modified trial state can then be described alternatively as a unitary transformation of the matrix representation of the Hamiltonian. This transformation in some sense brings the matrix closer to diagonal form, simplifying the expansion of the eigenstates.

So far we have described this device as a plausible conjecture. Preliminary calculations (and succeeding ones) confirm that this procedure succeeds almost beyond reasonable expectations. The measure of success in this instance is the question of how many S-phonons must be included in basis states in order to achieve satisfactory numerical convergence. If the factor (5.4) were to describe the distribution of ground-state phonons exactly, then no occupations of this state would have to be included in the explicit basis. Now that is not actually the case, but we will find that we need to include only one or two S-phonons, and these only for the most important "skeleton states."

The usefulness of this transformation would seem to suggest that there may be other similar transformations that further reduce the difficulty of the numerical computation. One such transformation is naturally suggested by the procedure of Lee-Pines. Each additional such transformation, however, introduces additional terms into the Hamiltonian which increase the complexity of the numerical computation in other respects. The additional terms give rise to new matrix elements, complicating the logic of computing these elements, and tending to off-set the advantage of the transformation. While such a transformation would be likely to reduce the size of the basis, the corresponding matrix would be less sparse, introducing compensating difficulties of storage and efficiency. These considerations lead to the question of numerical techniques, which will be the subject of the following chapter.

VI. NUMERICAL TECHNIQUES

In the previous sections we described, in principle, a systematic procedure for determining the ground eigenstate and eigenvalue for the polaron. After transforming the Hamiltonian (2.39) by the unitary operator (5.4) the procedure is to expand the trial state $/\phi)$ as a linear combination of basis states, which are chosen in advance. The Rayleigh Quotient

$$E = \frac{(\phi/\ H\ /\phi)}{(\phi/\phi)} \tag{6.1}$$

then takes the form

$$E = \frac{\sum_{ij} H_{ij} x_i x_j}{\sum_i x_i^2} \tag{6.2}$$

Here x_i are the coefficients in the fundamental expansion (4.1) of the trial state, and the H_{ij} are the matrix elements of the Hamiltonian between the i-th and j-th state of the basis

$$H_{ij} = (i/\ H\ /j) \tag{6.3}$$

As a preliminary step, one must compute these matrix elements. The matrix elements will typically include two factors: One factor is a function of the occupation-numbers of the states (modes) involved in transitions, and arise from the action of the creation and

destruction operators A_i and A_i , respectively, according to the rules

$$A_i \, /n_1, n_2, \ldots) = \sqrt{n_i} \, /n_1, \ldots n_i - 1, \ldots) \tag{6.4}$$

$$A_i \, /n_1, n_2, \ldots) = \sqrt{n_i + 1} \, /n_1, \ldots n_i + 1, \ldots) \tag{6.5}$$

The second factor will be one of the integrals T_{ij}, J_i, $Q(i,j,k,l)$, or K_{ij} given by (4.37), (4.38), (4.39), (4.40). Explicit formulas for these integrals are given in Appendix A.

Assuming now the matrix elements H_{ij} are known, the condition

$$\frac{\partial}{\partial x_i} \left(\frac{\sum_i \sum_j H_{ij} x_i x_j}{\sum_i x_i^2} \right) = 0 \quad (i=1,2,\ldots n) \tag{6.6}$$

leads to the system of equations

$$\sum_j (H_{ij} - E \delta_{ij}) \, x_j = 0 \qquad (i=1,2,\ldots n) \tag{6.7}$$

for the expansion coefficients (variational parameters). Here E means

$$E = \frac{\sum_i \sum_j H_{ij} x_i x_j}{\sum_i x_i^2} \tag{6.8}$$

which, of course, is the energy eigenvalue when the x_i satisfy (6.7).

There are a number of well-known methods for numerically determining the eigenvalue E and the amplitudes x_i. Which method is appropriate depends on the characteristics of the problem at hand. The most important characteristic of our problem in this respect is that the matrix (H_{ij}) is very large, and that it is reasonably sparse. We will need to deal with matrices of dimensionalities ranging from about 100 up to almost 1000. For a typical basis one finds that about 7 per-cent of the matrix elements are non-zero.

The most obvious problem is that of storing the matrix elements themselves. Because these quantities are used repeatedly in the computation, it is almost imperative that they be stored in the "fast memory" of the computer, rather than on tape or disc. Now a matrix of dimensionality 1000 has 10^6 matrix elements, which figure exceeds the capacity of the largest computers by a factor of ten. One must then take advantage of the fact that the matrix is sparse by storing only the non-zero elements (together with a code giving the location of each non-zero element).

These same considerations of space exclude from consideration any numerical methods in the course of which the matrix "fills up" with non-zero elements--i.e. any method that requires storing appreciably more information than is contained in the matrix itself. This effectively excludes computing high powers of the matrix, or its inverse.

These considerations point in the direction of the so-called cyclic iterative methods of numerical analysis. In these schemes the basic operations are matrix multiplication alternated with inversion of a very small, or even trivial (i.e. diagonal) sub-matrix. In terms of our requirements, this approach is suitable because these operations can be carried out very naturally and efficiently for matrices stored in the form we have described. Moreover, the amount of information to be stored remains constant throughout the computation.

Cyclic iterative schemes have another property that will fit very conveniently with our purposes. One ordinarily begins such a computation with a rough (or even arbitrary) approximation, and approaches the solution in a sequence of convergent steps. It is very natural to choose these steps to coincide with the solutions corresponding to successively more general sets of basis vectors. The sequence of intermediate eigenvalues obtained then give us the information we need about the importance of the various classes of states. We have, then, a happy coincidence between our requirements in principle, and what is naturally suggested by practical considerations.

When a sequence of intermediate values is required, the iterative schemes are likely to be more efficient than other procedures. This is because the work is cumulative: each step makes efficient use of the information gained in previous steps. Moreover, the benefits are

greatest where they are needed most. As the basis (and the matrix) increases in size, the solution is presumably converging, so that progressively fewer iterations are required to obtained the next approximation to a desired standard of accuracy. We have here a second important reason for ordering the states so that successive increments in the eigenvalue tend to decrease (in magnitude) monotonically.

The eigenvalue E in (6.7) is determined in principle by the condition

$$\det (H - E) = 0 \tag{6.9}$$

which implies that the system of equations (6.7) is "redundant:" One of the coefficients x_i may be chosen arbitrarily, and the remaining n-1 coefficients are determined by any subset of n-1 equations taken from (6.7). In our case it will be convenient to standardize the solution by taking the first coefficient $x_1 = 1$, and to consider the last n-1 equations as determining the coefficients x_i $(i=2,3,\ldots n)$.

In this problem we begin without precise knowledge of either the eigenvalue E or the eigenvector $X = (x_1, x_2, \ldots)$. From previous calculations (particularly that of Feynman, discussed earlier) we do have an estimate (and upper bound) on the eigenvalue E which may be regarded as a decent approximation. In this special circumstance there is a simple and rapidly convergent iterative scheme for determining the

eigenvalue. Let us assume for the moment that we have a suitable scheme for solving the system

$$\sum_{j=1} (H_{ij} - E\,\delta_{ij})\, x_j = 0 \qquad (i=2,3,\ldots n;\ x_i = 1) \qquad (6.10)$$

for x_i $(i=2,3,\ldots n)$, assuming E is given. (We shall return soon to the question of how this can be done. Suppose that $E = E_o$ is an initial "guess" for the eigenvalue (based, if one chooses, on previous calculations); one can then solve the system (6.10) for the first approximation $X = X^{(1)}$ to the eigenvector. (if the scheme for solving (6.10) is an iterative one, the zero-th order approximation $X=X^{(0)}$ will also be needed). If we can prescribe a rule for determining the next approximation $E = E^{(1)}$ to the eigenvalue using the vector $X^{(1)}$, then by repeated application of this rule one can define a cyclic iterative scheme for determining the eigenvalue E (and the eigenvector X).

We seek a functional of the vector X that will give us the best possible approximation to the eigenvalue for a given approximation to the eigenvector. We recall that the Rayleigh quotient (6.1) itself has the property that it is stationary as a functional of X if X is near the eigenvector. That is to say, an error dX in the eigenvector causes a fractional error of order $(dX)^2$ in the eigenvalue, when the latter is computed from the Rayleigh quotient.

Accordingly, it is suggested that we define our iterative scheme for determining the eigenvalue E by the equations

$$\sum_{j=1}^{n} (H_{ij} - E^{(n)}\delta_{ij})\, x_j^{(n+1)} = 0 \tag{6.11}$$

for i=2,3,...n; and

$$E^{(n)} = \frac{(X^{(n)}, HX^{(n)})}{(X^{(n)}, X^{(n)})} \tag{6.12}$$

To review briefly what this means, we begin with an approximation $E = E^{(0)}$ to the eigenvalue, and solve the system (6.11) for the appoximate eigenvector $X = X^{(1)}$. (We will discuss below just how this may be done). This vector is then used as the argument in (6.12) to compute $E^{(1)}$. The cycle is repeated by again solving (6.11), this time with $E = E^{(1)}$.

It is found in practice that this scheme is both efficient and remarkably insensitive to the zero-th order approximation to the eigenvalue, particularly when the approximate value lies below the true value. If one takes care to slightly under-approximate the eigenvalue, the procedure converges very rapidly, and always to the ground state.

We turn now to the question of solving the system (6.10) of n-1 linear equations, given $E^{(n)}$. It is this step that presents the

greatest difficulties in terms of finding a procedure that is efficient in time and space. The arguments given earlier in favor of the cyclic iterative procedures apply particularly to this problem of determining X for a given E.

For the present discussion we write the equations (6.11) in the form

$$A\,X = Y \tag{6.13}$$

Here A is the matrix H-EI, and Y is the inhomogeneous term (vector) that results from the convention $x_1 = 1$.

Suppose that we write

$$A = D + U \tag{6.14}$$

so that if X is a solution of (6.13) it is also a solution of

$$D\,X = -U\,X + Y \tag{6.15}$$

The form of (6.15) then suggests the iterative scheme defined by

$$D\,X^{(n)} = -U\,X^{(n-1)} + Y \tag{6.16}$$

This scheme satisfies all the requirements that arise from limitations of space, if the decomposition A = D + U is suitably chosen. Whether the scheme is also efficient in time will depend on the rate of con-

vergence which depends, in turn, on the decomposition.

If the solution of (6.16), i.e.

$$X^{(n)} = D^{-1}(-U X^{(n-1)} + Y) \tag{6.17}$$

is iterated formally, one arrives at the series

$$X^{(n)} = (-D^{-1}U)^n X^{(0)} + \sum_{j=0}^{n} (-D^{-1}U)^j D^{-1}Y \tag{6.18}$$

As in the analogous algebraic series, to prove convergence it is sufficient to show that the operator $D^{-1}U$ is "small, when smallness is measured in the appropriate way. In this case a suitable measure is the "spectral radius." One can show that the series (6.18) will converge to the solution of (6.13) if the spectral radius of $D^{-1}U$ is less than unity. Here we are not concerned, in fact, with formal questions of convergence, but with practical problems, particular with regard to the rate of convergence. It may happen that a scheme that converges in principle converges so slowly in practice that it is not useful.

One choice for the matrix D corresponds to the well-known Jacobi method.[16] In this procedure, D is chosen to be diagonal. Specifically, D is the matrix derived from the matrix H - $\mathbb{E}$ by setting all off-diagonal matrix elements to 0. In our problem, the Jacobi method does not result in satisfactory convergence, except for the case of very weak coupling. The procedure which we use, and

which we are about to describe, corresponds in a sense to a generalization of the Jacobi method. The matrix D will be chosen to be diagonal, except for a relatively small principal submatrix, D', which will have non-vanishing off-diagonal elements.

Our procedure is motivated by the conjecture that the iterative scheme (6.17) will converge more quickly, generally speaking, if the matrix D is, in some sense, an "important" part of H. It is necessary now to give some kind of operational definition to this notion of "importance."

We have already seen, in the context of our general discussion, that each specific set of basis vectors leads to a matrix H, which matrix in turn defines an eigenvalue problem. This matrix H is referred to as the <u>restriction</u> of the Hamiltonian to the space spanned by a given set of vectors.

In making our choice of the matrix D, let us consider the class of matrices that arise from the Hamiltonian in the manner described (i.e. in the sense of restriction). The eigenvector and eigenvalue of each of these matrices may be regarded as an approximation to the eigenvector and eigenvalue of the Hamiltonian. This approximation will be good or bad according to the particular choice of basis vectors—i.e. according to what restriction of the Hamiltonian is being considered. This suggests that these matrices (i.e. these restrictions of the Hamiltonian) may be regarded as "important" or "unimportant" according to whether the ground eigenvalue and eigenvector give good or bad approximations to the eigenvalue and eigenvector for the polaron. Our conjecture then suggests that D be chosen so that its ground

eigenvalue and eigenvector correspond as closely as possible to the exact eigenvalue and eigenvector of the Hamiltonian.

Practical considerations at the same time put a certain restriction on the choice of D. According to (6.17), the iteration requires solving an equation of the form

$$DX = V \tag{6.19}$$

This step we will want to solve directly (as opposed to iteratively); otherwise we become involved in an absurd regression of iterations within iterations! With limited computer memory, only relatively small linear systems can be solved in this manner. (This remark applies, of course, only to non-trivial linear systems). Suppose that D is of the form

$$D = \left(\begin{array}{c|c} D' & 0 \\ \hline 0 & F \end{array}\right) \tag{6.20}$$

where F is a diagonal matrix, and D' is a matrix that has, in general, non-vanishing off-diagonal elements. The restriction that we should be able to solve (6.19) directly then imposes a restriction on the dimensionality of D'.

Note that the eigenvectors of D' are also eigenvectors of D (insofar as the former may be thought of as being imbedded in the larger space). Our conjecture then implies that D' should be chosen so that its eigenvalue and eigenvector coincide as closely as possible

with the eigenvalue and eigenvector for the polaron. The set of states to which D' corresponds (in the sense of restriction) should then be the set of the most "important" basis vectors--i.e. the basis vectors whose amplitudes in the polaron eigenstate are large. The number of vectors to be included in this set (i.e. the dimensionality of D') is then limited, as we have already said, by practical considerations.

We are now ready to give the prescription according to which the matrix D is chosen in our computation. We state first that D is chosen to be of the special form (6.20) On the basis of intuition, and preliminary computations, a set of about twenty of the most "important" basis vectors is chosen. (In practice this set consisted mainly of the so-called "skeleton states" of rank R less than 6). The matrix D' is then determined by this set. The diagonal elements of the matrix F are chosen so that these elements (as they occur in D) are identical to the corresponding diagonal matrix elements of H - $\mathbb{E}$ (i.e., exactly as they would be given by the Jacobi method).

We emphasize that the choice of D, once made, is not changed in the course of the computation. As new states are added, the diagonal elements corresponding to these states become part of the submatrix F, and the off-diagonal elements become part of the submatrix U.

This procedure, as desribed above, gives satisfactory convergence for a range of coupling constants that is not restricted

to the weak-coupling range. However, the rate of convergence weakens as the coupling strength increases. The convergence becomes too weak to be useful, and then diverges, for coupling strengths within the range we want to study. Fortunately, this situation can be corrected by a simple modification.

The modification is suggested by studying the detailed behavior of the solutions in the case where the convergence is very weak. It is found that, in general, the components x_i of the eigenvector undergo damped oscillations (as a function of the order of approximation, n). With successive approximations, each component oscillates about a mean value which turns out to be very close to the solution when convergence is achieved. This suggests that the vector to be used as the beginning for the (n+1)st approximation should not be simply

$$X^{(n+1)} = D^{-1}(-U X^{(n)} + Y) \tag{6.21}$$

as given by (6.17) but rather by

$$\bar{X}^{(n+1)} = \frac{X^{(n+1)} + X^{(n)}}{2} \tag{6.22}$$

which is the average between the previous value and that suggested by (6.21).

A generalization of this form is to consider the weighted average

$$\bar{X}^{(n+1)} = \frac{X^{(n+1)} + rX^{(n)}}{1 + r} \tag{6.23}$$

where r is a parameter that must be chosen in some way. If the original scheme converges satisfactorily, r may be chosen to be 0. At the other extreme, one can show by simple examples that a choice $r > 1$ may transform a divergent behavior (according to (6.21)) into a convergent one.

It is found that the rate of convergence is not highly sensitive to the value of r, within a certain range. The optimum choices of r were found to increase slightly with coupling constant, ranging from $r = 0.5$ (approximately) for $g^2 = 2$ to $r = 2.0$ for $g^2 = 7$. With these choices, more or less satisfactory rates of convergence were achieved in every case.

Any numerical procedure, and particularly one on the large scale we are considering, is subject to rounding errors. The general problem of determining the eigenvalue and eigenvector of a matrix of dimensionality 1000 would indeed present very severe problems in this respect. In our case the inaccuracies in the result arising from rounding errors are believed to be smaller than errors due to other factors, and will not affect the answers within the number of significant figures we are interested in.

This fortunate circumstance can be attributed to at least three factors. The first factor is that the matrix is sparse. We have mentioned earlier that the matrices we work with generally have about 7 per-cent non-zero elements. This enormously reduces the number of numerical steps in the computation, and greatly reduces the round-off error. A second factor is that the matrix appears to be well conditioned in the context of our numerical procedures. Finally, we are helped by the advance of computer technology. The majority of the computations (and all of the ones with very large bases) were done with the CDC 6600 computer, which has as a standard feature a 60-bit word (which breaks down into a 48-bit fraction, an 11-bit exponent, and a sign). Translated, this means that every number is stored to 14 significant figures, instead of 8 significant figures in the more advanced older machines. The accuracy of our computation then almost equals what was formerly called "double precision."

Various checks are programmed into the procedure to give information about about rounding-errors. The vector X as a solution of (6.13) can easily be checked by direct substitution. We define the vector V by

$$V = A\ X - Z \tag{6.24}$$

If X is an exact solution of (6.13), Z = 0. The norm of Z (e.g. the Euclidean norm) then serves as a useful measure of how well

X approximates the solution of (6.13). As a matter of fact, the vector V can be computed efficiently at every step of the iteration because the non-trivial steps in the right hand side of (6.24) coincide closely with the multiplicative step of our iterative scheme. The procedure followed was to use the norm of V as a test to decide when to stop the iterative cycle (i.e. the cycle for computing X for a given E).

There is an independent check on the accuracy of any approximation $E^{(n)}$ to the eigenvalue E of the matrix H. Recall that the system (6.10) involves the last n-1 equations of the system (6.7). The first equation has then not been explicitly used in the calculation. If $E^{(n)}$ is exactly the eigenvalue E, this first equation will be a linear combination of some set of equations taken from the next n-1 equations. Conversely, if X is a very accurate solution of the last n-1 equations for $E = E^{(n)}$, then if we use

$$(H_{11} - E') + \sum_{j=2}^{n} H_{1j}x_j = 0 \qquad (6.25)$$

(i.e., the first equation of the system (6.7)) to determine E', we will find $E' = E^{(n)}$ if $E^{(n)}$ is the eigenvalue of H. At each order of approximation, E' was computed, and the iterations were continued until E' and $E^{(n)}$ agreed to within one part in 10^5.

VII. RESULTS

In this chapter we shall present an analysis of the numerical results that were obtained according to the procedure outline in earlier sections. For various coupling constants (and for several P for each g) we have a sequence of approximate eigenvalues. These approximate eigenvalues correspond to a sequence of successively more general sets of basis vectors. The difference between successive approximations to the eigenvalue is then referred to as the increment corresponding to states that are included in the larger basis, but not in the smaller one. Our analysis will rely heavily on the information that can be extracted from these sequences of increments.

We restrict ourselves first to the special case $P = 0$. This relatively simpler problem will in fact present enough difficulties to occupy the largest part of our efforts. This special case is at the same time sufficiently general to illustrate many of the questions it is our intention to explore.

We pause to make a few comments about the accuracy of the results. Disregarding for the moment any numerical errors in the computation itself (such as rounding errors), our results give only an approximation to the exact results because our trial state necessarily involves only a finite number of parameters. The main body of this chapter will be devoted to optimizing the result of this finite computation, and to estimating the error due to truncation of the basis.

To carry out this program we need to have at the outset a realistic appraisal of what order of accuracy we hope to achieve. A preliminary survey of results leads to the guess (which will turn out a little optimistic) that it may be possible to reduce the fractional error to about two parts in 10^3 (two-tenths of one per-cent). In any case, it is clear that it will not be possible to reduce the uncertainty to substantially smaller limits (except for a restricted range of coupling strengths). On the basis of this estimate, we make the convention that increments in the eigen-value giving a fraction contribution smaller than two parts in 10^4 may be ragarded as negligible. The smallest increments included will then be one order of magnitude smaller than the projected uncertainty in the result. As limitations of space will present the most serious restriction on our computation, it is absolutely necessary to begin with an estimate of what is important and what is negligible. Only in this way can one systematically and consistently exclude from the basis those states that give unimportant contributions, and makes room for those that may give important ones.

The criteria set down here are consistent with the estimated numerical uncertainties in the intermediate results themselves. These uncertainties are due to round-off error and to a residual error resulting from cutting off the iterations after a finite (in fact, small) number of steps. The second of these two errors is believed

to predominate under the conditions of our computation. The rule used in the computations was to cut off the cyclic iterations when (and only when) the estimated residual error in the eigenvalue was smaller than one part in 10^5. This is better than an order of magnitude smaller than the smallest increments we need to consider (two parts in 10^4). These smallest increments will then have validity to at least 5 per-cent.

We have at our disposal an almost bewildering wealth of numerical information waiting for analysis and interpretation. In order to carry out this analysis, which will involve constructing overlapping sequences of many sorts from out tables of results, it will be helpful to restrict ourselves to one coupling constant. We choose this constant to be $g^2 = 5$. This value is typical of the intermediate coupling region, and will be convenient for demonstrating the convergence patterns that may be found also for weaker and stronger coupling. When it is useful we will give specific examples for other coupling strengths. <u>In what follows, all eigenvalues or increments of eigenvalues refer to the case $g^2 = 5$ unless otherwise stated.</u>

In view of the criteria established earlier, we can say that (for $g^2 = 5$) increments smaller than 1×10^{-3} in magnitude may be regarded as negligible, and all increments are subject to an uncertainty $\pm 5 \times 10^{-5}$.

We are now almost ready to begin our analysis. In an earlier section the concept of rank (R) was introduced. The rank is defined by

$$R = \sum_{L=0} n(L)\, L \tag{7.1}$$

and gives a one-parameter description of the angular-momentum distribution of the phonons. The rank of a state can be related to the order of perturbation-theory in which the state first makes a contribution to the eigenstate (according to a particular perturbation-theoretic scheme).

Preliminary calculations, which we are now ready to present, strongly confirm the usefulness of this notion.

Properties of the Rank (R)

One simple test is to compute a sequence of eigenvalues, using first a basis of R=0 states only (a trivial case); then with a basis of of R=0 and R=2 states; then with R=0, R=2, and R=4 states, and so on. These preliminary calculations were done with the phonons restricted to one radial mode (N=0). In this case the R=0 case (the first in the sequence) is equivalent to the Lee-Low-Pines calculation and the eigenvalue is known to be -5.00 exactly. A sequence of results is given in the following table. The first column describes the basis; the second gives the eigenvalue; the third gives the differences between eigenvalues; and the fourth the number of states in the given basis.

R = 0	-5.0000		5
		-0.2071	
R ≤ 2	-5.2071		7
		-0.0308	
R ≤ 4	-5.2378		21
		-0.0031	
R ≤ 6	-5.2409		67

The convergence of the successive increments is even stronger for weaker coupling, slightly weaker for stronger coupling. In each case the pattern demonstrated here is evident: The sequence of increments between succeeding eigenvalues forms a monotonically decreasing, rapidly convergent (almost geometric) sequence.

One would also like to have it that states with the same rank should be generally of the same order of importance. As an example, consider the states of R=6. Exclusive of S-phonons, the possible distributions in L are D^2; PDF, D^3; P^2D^2, P^3F^1; P^4D; P^6. The distributions are grouped by sets according to whether there are two, three, four, five, or six phonons, respectively. The increments corresponding to each of these classes were found and are given here:

two phonons	-0.0014	4 states
three phonons	-0.0042	11
four phonons	-0.0067	16
five phonons	-0.0038	11
six phonons	-0.0007	4

The close agreement between the second, third, and fourth terms in particular shows that these sets of states can properly be thought of as belonging to the same order of importance. (These figures were not taken from the computation referred to earlier, where phonons were restricted to one radial mode. The sum of these increments will not agree with the increment for all R=6 states quoted before, because the earlier computation was done with a very restricted basis).

The sequence shown here is not a special case, but is typical of many cases that were examined. We conclude that it is useful to group states together by rank, R, with the expectation that states of the same rank will give roughly similar contributions, and that the sequence of eigenvalues will converge as states of successively higher rank are included in the basis.

Role of Ground-State (L=0, N=0) Phonons

In an earlier section we made a unitary transformation of the Hamiltonian, and we asserted that with the transformation the ground-state phonons would play a very minor role in the explicit basis. As an illustration, consider the sequence of states /1), /1,1), /1,1,1),.... These are the states with one, two, three,... phonons in the ground state L=0, N=0, and they are the states from which the Lee-Low-Pines solution was constructed. With the Hamiltonian in its conventional form (untransformed), these states played a dominant role.

The first state /1) was included in the block of states giving the zero-th order approximation, and the corresponding increment is not available. The increments corresponding to the following states are, respectively

/1,1)	-0.00270
/1,1,1)	-0.00014
/1,1,1,1)	-0.00001

The first increment is non-negligible but small, the second is negligible, and the third has no numerical significance. The very rapid convergence of these increments support our conjecture that the factor (5.4) which was suggested by the Lee-Low-Pines calculation, also gives a good approximation to the distribution of ground-state phonons in the more general case. In particular, it is suggested that this factor gives the correct asymptotic distribution--i.e. for large numbers of phonons.

For completeness we should look at the ground-state phonons as they occur in states that also have non-S phonons (i.e. P-phonons, D-phonons, etc.) We refer to these states as "augmented skeleton states." It turns out that the increments corresponding to these sets of states are, in general, smaller than our minimum "important" increment 1×10^{-3}; moreover, the terms of the sequences grow small so rapidly that the values are obscured by numerical uncertainties. While it would be nice to be able to study the convergence of these sequences in detail, we are pleased that the contributions are in fact small, so that the large numbers of states arising in this way

do not have to be included in the basis. This situation is in marked contrast to that of earlier calculations, where these states gave very important contributions (up to those with ten, twelve, or even fourteen phonons).

The very minor role played by the ground state phonons (particularly in "augmented states" can be illustrated again in the context of another aspect of our results, to which we now turn our attention.

The Role of S-Phonons in the Radial Modes N=0,1,2,...

In the previous discussion we were concerned with ground state phonons--i.e. S-phonons in the radial mode N=0. We now look at the more general case of S phonons in arbitrary radial modes. If $N > 0$ we refer to a "radially excited" S-phonon.

Let us restrict our attention first to states with rank R=0. That is to say, we are concerned with states that have only S-phonons. To proceed systematically, we examine first the case of such states with one phonon. The natural sequence of states is then /1), /17), /33), /49),..., corresponding to one S-phonon (L=0) with N=0,1,2,3,.. respectively. The increments for the third, fourth, and fifth states are available, and are given below.

/33)	-0.01944
/49)	-0.00159
/65)	-0.00008

The first two states again had to be included in the initial block of states, and the increments for these are not available.

This sequence typifies what we ideally hope to find throughout this analysis: The sequence of increments converges monotonically and smoothly to 0, and converges rapidly enough that a few terms suffice to reduce the probable residual error to within the established limits.

The very satisfactory convergence of this particular sequence gives encouraging evidence that the choice of radial modes was a suitable one. The conclusion suggested by this simple example is supported by other evidence. For example, one can look at the <u>two phonon</u> states of the same general type.

We arrange these two phonon states into subsets, corresponding to having the phonons restricted to N=0; N=0,1; N=0,1,2; and so forth. The first set contains only /1,1); the second contains /1,17), /17,17); the third /1,33), /17,33), /33,33). Following sets are constructed according to the same pattern. The increments corresponding to these sets are

$N \leq 2$	-0.00621
$N \leq 3$	-0.00105
$N \leq 4$	-0.00012

A similar sequence for <u>three-phonon</u> states is found to be

$N \leq 2$	-0.00063
$N \leq 3$	-0.00015
$N \leq 4$	-0.00001

Each of these cases (the sequence for two-phonon states and the sequence for three-phonon states) demonstrates the same impressive convergence patterns that were found in the one-phonon sequence. Furthermore, a comparison of the total contribution of one-phonon states, the total contribution of two-phonon states, and the total contribution of three-phonon states leads to the conclusion that states of this type with four or more phonons make a contribution that is negligible by our standards. The sequences then converge satisfactorily both in the direction of higher radial excitation, and in the direction of larger phonon occupations.

The S-phonons we have been studying can, of course, occur in states that also have occupations of higher L modes. These states will not have R=0, but rather R=2,4,.... An example of such a state is /17,2,2). We place this state in our scheme of classification by thinking of it as arising from the "skeleton state" /2,2) (two phonons in the mode L=1,M=0) by the addition of one S-phonon in the first excited radial mode (N=1).

Consider the skeleton states with R=2, and then consider the sequence of states derived from these by adding <u>one</u> S-phonon in successively higher radial modes N=0,1,2,.... The first set of the sequence is then /1,2,2), /1,5,8); the second set is /17,2,2), /17,2,2), /17,5,8). This sequence parallels the sequence /1), /17), /33),...studied earlier.

The sequence of increments for these sets is given below

N = 0	-0.00001
N = 1	-0.03512
N = 2	-0.00261
N = 3	-0.00036

With the very dramatic exception of the first term, the sequence demonstrates the same sharp convergence that was evident in earlier examples. The very small term corresponding to N = 0 illustrates the special role played by the ground state, which is singled out by the unitary transformation (5.4) as a special case. We had stated that the anomolously small role played by the ground state could be illustrated in the context of the whole sequence of radially-excited S-phonons. This example confirms our expectations and supports our earlier conclusions about the usefulness of the unitary transformation. The terms for N=1, N=2, N=3 again show that the radial modes were well chosen, and that terms for N=4 or higher are not expected to give important contributions.

The same principles can be illustrated by a sequence derived from R=2 skeleton states by the addition of <u>two</u> S-phonons. Classifying the paris of S-phonons (as before) by the highest radial mode occupied, we find

N = 0	-0.00039
N ≤ 1	-0.00348
N ≤ 2	-0.00060

A few terms from the sequence for <u>three</u> added S-phonons are known:

$N = 0$	-0.00001
$N \leq 1$	-0.00025

The contribution from all two-phonon increments is about an order of magnitude smaller than for the one-phonon case; the same comparison holds for the three- and two-phonon cases. Moreover, the three-phonon case (above) already gives a negligible contribution. We may confidently conclude that the sequences converge also in the direction of higher occupations (as well as in the direction of higher-order radial excitations), and that similar terms with four or more phonons do not need to be included.

The results quoted were primarily sequences beginning from R=2 skeleton states. The R=2 states were used for this study because they were expected to give the leading contributions, and because there are very few R=2 skeleton states, so it is possible to study several terms of a sequence before running into enormous numbers of states. Below we give some similar examples taken from R=4 and R=6 sequences. With progressively higher rank, we will find limitations of space prevent us from including as many terms of each sequence as in preceeding examples. If the magnitude of the contributions grows smaller with increasing rank (as we hope they will), not as many terms will be needed, for a given standard of accuracy. At the same time, our confidence in the extrapolations may suffer.

We present the sequence of results for adding one S-phonon in successively higher radial modes to the skeleton states of rank R=4.

N = 0	-0.00048
N = 1	-0.01195
N = 2	-0.00018

We see again the phenomenon of a very small leading term (N=0) followed by a relatively larger N=1 term, and a sharply convergent sequence for higher modes.

The terms for adding <u>two</u> phonons to R=4 skeleton states are

N = 0	-0.00011
N = 1	-0.00114

Higher terms of this sequence were not studied.

One term derived from R=6 states was studied. This term was for the addition of one S-phonon in the <u>N=1</u> mode. (We have learned the N=0 case gives negligible contributions). This term is

N = 1	-0.00302

It is instructive to compare the contributions from corresponding terms of R=2, R=4, and R=6. The increments (quoted above) for adding one N=1 S-phonon to each of these sets are

R = 2	-0.03592
R = 4	-0.01195
R = 6	-0.00302

These increments show an almost geometric convergence, with ratio about 1/4. It is strongly suggested that higher terms (e.g. R=8, 10,...) give unimportant contributions

While it is impossible to give rigorous estimates of the cumulative contribution of all terms not included in our basis, the convergence of all sequences studied so far has been so satisfactory that one feels confident these contributions are within the established bounds. Moreover, the results invite the interpretation that the basis used is close to the optimal basis for a given size (with respect to the characteristics we have studied).

The Role of Radially-Excited Non-S Phonons ($N \geqslant 0$, $L = 1,2,\ldots$)

We have taken the point of view that the S-phonons are in a category apart from non-S phonons (L=1,2,,..). In the context of our computation, this is reasonable because the S-phonons (in the dominant radial mode N=0) are described analytically (to a good approximation) by the unitary transformation to which we have referred. We have studied the role of these "ground state" phonons (N=0, L=0) in specific examples and have verified that they now play a minor role. The S-phonons in higher radial modes (N= 1,2,...) have also been studied, and it was found that the importance of these states dropped off rapidly with increasing N.

In these studies, we have considered principally states in which the non-S phonons were in the dominant radial mode N=0. It is now time to study the role of these non-S phonons when they are allowed also in the higher radial modes.

The sequence of states that will interest us may be thought of as being derived from skeleton states (all phonons in N=0 mode) by successive radial excitations (i.e. transitions of a phonon from the state L,M,N to the state L,M, N+1). In our particular scheme of numbering states, this corresponds to a transition from the i-th state to the (i+16)th state. We think of the excited radial states (N=1,2,...) in a category apart from the ground radial state (N=0), and separate the states under consideration into sets according to whether there one, two, three, (or more) phonons in the excited radial states.

The first category of states we examine is the set of states derived from skeleton states with rank R=2 by the excitation of <u>one phonon</u> to successively higher radial modes. The increments are labelled by the radial mode of the phonon, and are given below:

N = 1	-0.02143
N = 2	-0.00071
N = 3	-0.00047

A similar sequence can be studied for the case of states derived from R=2 skeleton states by the excitation of two phonons. The increments are labelled by the highest occupied radial mode.

$N \leq 1$	-0.00921
$N \leq 2$	-0.00091
$N \leq 3$	-0.00057

We find that the magnitude of the increments drops off very sharply with increasingly higher radial modes, and also that the contributions from states with two excited phonons is less than that from states with one exctied phonon. In this connection, it is encouraging to compare the results quoted with the increment for R=2 states with <u>no</u> excited phonons. If we begin with a basis of only R=0 states (E=-5.000) and think of adding the R=2 skeleton states, the increment corresponding to these states is -0.207. The contribution of states with excited phonons is then small by comparison to the contribution from the skeleton states (i.e. the states with all phonons in the ground radial mode). This justifies the special role these skeleton states play in our analysis.

Similar sequences can be studied starting from the R=4 skeleton states or the R=6 skeleton states. Consider the R=4 states. The sequence for raising one phonon to successively higher radial modes is

N = 1	-0.02466
N = 2	-0.00333

Higher terms of this sequence were excluded for lack of space, but their contribution clearly may be expected to be negligible. For R=6 states, the corresponding sequence (for raising one phonon to successively higher radial modes) is

N = 1	-0.01109
N = 2	-0.00254

For the R=4 and R=6 cases, the possibility of raising two or more phonons to excited radial modes was also studied, and the results were qualitatively similar to the R=2 case, and confirmed the general conclusions reached there.

One rather unfortunate result that emerges from these results and from others is that the notion of "rank" which has served so usefully until now is of more limited usefulness in connection with the higher radial modes. Consider the increments (quoted above) for raising one phonon to the first excited radial mode (N=1) for the case R=2, R=4, and R=6 skeleton states, respectively:

R = 2	-0.02143
R = 4	-0.02466
R = 6	-0.01109

Similar sequences quoted earlier, with all phonons restricted to the ground radial mode, gave monotonic and, in fact, very sharp convergence with increasing rank. We find here that the contribution of the R=4 term in this sequence is in fact slightly larger than the R=2 term. The most optimistid guess for the next term (R=8) in the sequence would be based on an extrapolation from R=4 and R=6, and would lead to something of the order -0.00500 (which is not negligible according to our criterion). It is probably useless to speculate on the sum of all succeeding higher terms. This example is the first case where the estimated residual error from states not included exceeded our tentative standard of -0.00100. The number of states of the kind involved (i.e. R=8 states with radial exci-

tatins) is so large that it was not possible to extend the computation to include these terms.

States with Radially-Excited Non-S Phonons and Added S-Phonons.

In the preceeding discussion we considered states derived from skeleton states by successive radial excitations. Inasmuch as skeleton states by definition have no S-phonons, these states will also be without S-phonons, and the radially-excited phonons referred to will be non-S phonons. From these states one can now derive still other sequences of states by successive additions of S-phonons (in the ground or radially-excited states). In an earlier section we discussed such sequences derived from the skeleton states themselves.

The sequences we have just described were studied in accordance with the same principles that have been illustrated in our previous examples. In order to shorten the discussion we omit specific numerical results, and summarize the conclusions derived from these results.

Beginning with a given set of states without S-phonons (but with radially-excited non-S phonons), the sequences derived by addition of one (or more) S-phonons in successively higher radial modes displayed the same smooth convergence that was found in similar sequences derived from skeleton states. Comparison of the contributions from states with one, two, or more augmented phonons showed that the sequences converged (as before) in the direction of larger occupation of S-states.

The one characteristic that differentiates these results from other analogous sequences is that the term corresponding to putting the S-phonons in the ground state (N=0, L=0) is no longer exceptionally small, as it was for sequences derived directly from skeleton states. In the earlier examples the exceptionally small contribution from this term was attributed to the unitary transformation (5.4) which apparently described the distribution of these phonons with great accuracy. In the present case the term corresponding to N=0 plays the role one would normally expect it to play in the sequence N=0,1,2,...; i.e. it gives the leading contribution. As was just mentioned, the succeeding terms display smooth and satisfactory convergence. The <u>normal</u> role of the N=0 term in this case (as compared to the <u>anomolously</u> <u>small</u> role in earlier cases) can only be interpreted by saying that the unitary factor (5.4) does not correctly describe the distribution of ground-state phonons in states with radially-excited phonons present.

In our study of these sequences we encounter the same difficulty that was encountered in our study of states without added S-phonons (but with radially-excited non-S phonons): the sequences of <u>given</u> <u>rank</u> <u>R</u> converge well in all respects, but comparison of contributions from different rank leads to ambiguous or disturbing results. There are a few examples where the contribution from R=4 states is larger than the contribution from R=2 states. Limitations of space in each case prevented including the R=6 term.

Summary

We pause now to review what has been said.

Preliminary calculations (particularly with phonons restricted to the first radial mode) tended to support our conjecture that the rank classification gives a useful hierarchy of approximations. More detailed computations showed that not only do states of higher rank give successively smaller contributions, but different kinds of states of the same rank gave roughly similar contributions.

The next observations confirmed that (after the unitary transformation) the ground-state (L=0, N=0) phonons played a very minor role. This fortunate circumstance was attributed to the unitary transformation (5.4), which in effect gives an analytic description of the distribution of these phonons which closely approximates the correct distribution.

The discussion was extended to S-phonons in the higher radial modes. Sequences of R=0 states (states with only S-phonons) were studied, and the corresponding sequences of increments were found to converge smoothly and rapidly to 0 with increasing occupation of phonons, and with higher patterns of radial excitations. The radially-excited S-phonons were studied also in their role in "augmented states," and similar conclusions were reached.

Next we examined states derived from skeleton states by exciting non-S phonons to higher radial modes. For states of a given rank, these sequences again converged smoothly. Difficulties arose when one compared corresponding contributions from states of rank R=2,

R=4, and R=6. Unlike similar sequences studied earlier, these rank-sequences did not converge smoothly enough to allow inferences about higher terms.

A similar situation was encountered when we examined the states with non-S radial excitations and with added S-phonons. Here again, the various sequences within a block of states of a given rank were found to converge more or less satisfactorily. Comparison between blocks of different rank were inconclusive or even disturbing.

We had pointed out earlier that our analysis could be thought of as "exploring" a region in an abstract space, S, in which each coordinate direction corresponds to one characteristic that may describe a set of states. In the discussion we have just completed, the principal characteristics--i.e. coordinate directions, or dimensions, in S--were rank (R), humber of added S-phonons, and radial excitations. As we explored further along the various directions, we sought to make our basis as efficient as possible by cutting off the sequences of states at the point where the residual error was reasonably expected to be negligible according to our criterion. In many cases the sequences of increments converged so smoothly that we were able to make these estimates with great confidence. In a few other cases the results were ambiguous and no confident estimate of the residual error could be made. This circumstance, of course, does not invalidate the calculation in the sense that our results continue to give upper bounds on the ground-state energy of the polaron; it does somewhat complicate the task of estimating the uncertainty in our results.

Estimates of Error

Let us now face the question of estimating the uncertainty in our results. In making such an estimate we are, of course, faced with several obstacles, both in principle and in practice. The difficulties that arise in principle involve the question of how the various residual errors that we have (perhaps accurately) estimated are to be combined. It is a fundamental characteristic of this kind of problem that the contributions from various sets of states do not combine linearly and independently. Secondly, there is the question of the validity of the extrapolation process that led us to these estimates. We are inclined to have stronger confidence in the extrapolation if we have a larger number of terms in the sequence, less confidence if we have fewer. There is then the practical problem that in some cases it was impossible to compute as many terms of a given sequence as one would have liked.

These difficulties notwithstanding, it is possible to make some meaningful observations. To the extent that the following discussion is valid, our computation offers one advantage over previous computations in that the numerical results also give information that can lead to a semi-quantitative estimate of the accuracy of the results themselves. Given a confidence in certain estimates that cannot be rigorously justified, one then has not only an upper bound but also a rough lower bound on the energy of the polaron.

Recall that earlier in the discussion we compared the eigenvalues corresponding to bases with states $R \leq 2$, $R \leq 4$, and $R \leq 6$, respectively. These comparisons were made with a basis with phonons restricted to the ground radial mode (N=0), and showed very impressive convergence. One is naturally led to look at the same sequence for a general basis, in which radial excitations are allowed. In particular, we may select from the "optimal" basis we have been studying all states with $R \leq 2$, then all states with $R \leq 4$, and so on. The eigenvalues for this sequence of bases are given below:

$R = 0$	-5.00000
$R \leq 2$	-5.34304
$R \leq 4$	-5.46897
$R \leq 6$	-5.50264

The increments corresponding to each of these sets are given in the next table:

$R = 2$	-0.343
$R = 4$	-0.126
$R = 6$	-0.034

Extrapolating these results, one is led to something of the order of -0.010 for the increment corresponding to the R=8 states. The validity of this estimate may be open to question on the ground that the sequence of states arranged by rank may seem to have been biased in favor of the lower-rank states. In studying the various sequences that begin with skeleton states and proceed by successive

excitations, additions of S-phonons, etc., it was not possible to include nearly as many terms for the higher-rank states than for the lower-rank states, due to lack of space. (For example, if we consider the states derived by exciting one non-S phonon to the N=1 mode, there are three such R=2 states, twenty-six R=4 states, and one hundred twenty-three R=6 states). This procedure, of including fewer terms for the higher-rank states, can be justified on the ground that the leading terms for the lower-rank states are larger (in magnitude), so that it is necessary to include more terms to achieve a given standard of accuracy. It is almost certainly true that the fractional error in the R=6 increment (-0.034) is greater than the fractional error in the increments for R=4 or R=2. That is, we might guess that we know the R=2 increment to within one part in one hundred, but the R=6 increment to only one part in ten. This is acceptable, because the R=6 increment is an order of magnitude smaller, and the terms enter into the results additively. This having been said, the fact remains that we are less confident in the higher-rank increments than the lower-rank increments, because as the rank (and the number of skeleton states) increases we are able to include fewer terms of all sequences, and our estimates and extrapolations are less well-founded. On top of this, there were a few sequences studied that did not exhibit convincing convergence in the direction of higher rank, and a few cases in which the terms even appeared to grow with increasing rank. One must say that it is possible (or probable, depending on how one interprets several

results) that the absolute error in the R=6 increment, for example, is larger than the error for R=4 or R=2. Making a rather conservative (i.e. pessimistic) view of this effect, one might be led to modify the estimate of the R=8 increment from -0.010 to -0.020.

Whatever residual errors remain in each of the sets of states (for a given rank) also affect us in another respect: In addition to affecting our extrapolation to the R=8 term and succeeding terms, these residues accumulate in some way (i.e. not linearly or independently) to give an error in the approximate eigenvalue (even for a basis cut off with states of rank R=6). It is difficult to make quantitative estimates of the cumulative effect of these errors. There were a number of residual errors that were estimated (more or less realistically) to be each of order -0.001 or less. We estimate that these small increments accumulate to give something an order of magnitude larger--i.e. of order -0.010. That is to say, we anticipate that the energy would be lower by -0.010 (approximately) if all these next higher order terms were included in the basis. Making the plausible assumption that the extrapolated estimate of the contribution from R=8 states will be unaffected by this perturbation, we are led to estimate an error of order -0.030 in our result.

If these estimates are not grossly over-optimistic, the eigenvalue for $g^2 = 5$ is accurate to within about <u>one-half of one percent</u>. Making a generous allowance for the accumulation of higher terms not studied, and for overly-optimistic estimates, we can say that <u>the eigenvalue -5.50 is almost certainly not in error by more than 1.0 percent</u>.

Another measure of what has been accomplished is, of course, the comparison of our ground-state energy E_o to that obtained by other variational calculations. A survey of such a comparison was given by Table 1, for a range of coupling strengths. For $g^2 = 5$, which is the case we are studying, the "next-best" result (Feynman's) is $E_o = -5.44$. Next is the figure $E_o = -5.40$, according to the Lee-Low-Pines result corrected by perturbation theory. The comparison is slightly misleading, because this result (-5.40) is not an upper-bound. Our result is seen to be a modest but significant improvement over Feynman's result, and a considerable improvement over Lee-Pines result $E_o = -5.30$.

In one sense a more suitable standard for comparison is not E_o itself, but that part of E_o that we are computing numerically in a non-trivial way. If we separate out the Lee-Low-Pines term $-g^2\omega$ from the eigenvalue, and write

$$E = -g^2\omega - E' \tag{7.2}$$

then our computation is directed toward determining E'. Our result for $g^2 = 5$ is $E' = 0.50$; the others are $E' = 0.44$ (Feynman), and $E' = 0.30$ (Lee-Pines).

This separation is also useful in relation to the discussion of errors. One finds that many important sets of states give contributions (for different g) more closely proportional to E' (which is in turn roughly proportional to g^4) than to E_o.

The analysis we have made for $g^2 = 5$ is summarized in Table 5, together with a corresponding summary for $g^2 = 1$, $g^2 = 3$, and $g^2 = 7$. The results are, of course, in units of ω. If the estimated error (second row from bottom) is computed as a fraction of E_o, the results are (approximately)

g^2	fractional (probable) error (in per-cent)
1.0	0.10 per-cent
3.0	0.20 per-cent
5.0	0.50 per-cent
7.0	1.10 per-cent

Judged in relation to E' (see (7.2)) the error does not grow so steeply as a function of g; on the other hand, the errors are then much larger in absolute value. For the experimentalist, the form given in the table above is the meaningful one, and our results accurate enough to be useful. At the same time, the error is larger than one would hope to achieve from a computation on such a large scale. We return to this point in the next chapter.

The Case $P \neq 0$, and the Effective Mass m^*

Previous discussions of the general case have been restricted in most cases to determining the energy for small P. In this range one conventionally writes

$$E_P = E_o + P^2/2m^* \qquad (7.3)$$

which defines the effective mass m^*

TABLE 5: Results and Error Estimates

Basis (by Rank)	$g^2 = 1$	$g^2 = 3$	$g^2 = 5$	$g^2 = 7$
	E_o = Energy Eigenvalue (Upper Bounds)			
R = 0	-1.0000	-3.000	-5.000	-7.000
R ≤ 2	-1.0108	-3.127	-5.343	-7.641
R ≤ 4	-1.0149	-3.158	-5.469	-7.944
R ≤ 6 *	-1.0162*	-3.164*	-5.503*	-8.038*
	Sequences of Increments in E_o			
R = 2	-0.0108	-0.127	-0.343	-0.641
R = 4	-0.0041	-0.031	-0.126	-0.303
R = 6	-0.0013	-0.008	-0.034	-0.098
R = 8 (Est).	(-0.0004)	(-0.002)	(-0.010)	-0.030)
Estimated Error	-0.0012	-0.006	-0.030	-0.090
Estimated Energy**	-1.0174	-3.170	-5.533	-8.128

* These figures represent the "best" upper bounds.

** Derived by adding probable error to upper bound. This figure does not give an upper bound.

Heretofore we included in our basis only states with even parity. In the general case $P \neq 0$ the term (4.36) in the Hamiltonian connects states with even parity to states of odd parity. The states with odd parity that must now be added to the basis are states of odd rank; i.e. R=1, R=3, R=5,... The basis of odd-parity states is constructed according to the same principles as in the even-parity case. In particular, one finds that the sequence of bases including states of successively higher rank gives a smoothly convergent sequence of approximate eigenvalues. The convergence of this sequence is extremely strong for very small P, and moderate for P of order unity. (No computations were made for very large P).

At this juncture we are faced with a severe problem arising from limitations of space. In principle, one would want to treat the even and odd-parity states on an even footing, so that the basis should include roughly as many states of one category as of the other. Now our previous results were based on a set of states with more than 900 basis vectors; this problem itself stretches the capacity of the available computers nearly to the limit. It is not possible to include in addition a similar set of odd-parity states.

One answer to this difficulty is to cut back the even-parity states, selecting for example, the 500 most important states, allowing room for a comparable number of odd-parity states. A second alternative, which is the one we choose, is to restrict the computation (as is customary) to small momenta, so that the convergence with increasing rank of the odd parity-states is very strong, and rela-

tively few odd-parity states need to be included. These computations suffice to determine the effective mass, m*, which is in any case the quantity of greatest interest. We find that we are able to carry the computation out far enough to get some information about how the energy deviates from the quadratic dependence on P that is observed for small P.

An analysis was carried out, similar to the one we have discussed, (but in slightly abbreviated form) to determine the optimal basis of odd-parity states within the given restrictions on the size of this basis. These exploratory computations resulted in a basis of 750 of the even-parity states previously discussed, and 200 odd-parity states of rank R=1, R=3, and R=5. With this basis, the estimated error in the eigenvalue (for non-vanishing but small P) was found to be comparable to the error estimated in the special case P=0.

If we now take P to be be a dimensionless unit (corresponding to what was formerly written $P/\sqrt{2m\omega}$), we can write (7.3) as

$$E_P = E_o + P^2(m/m^*)\omega \tag{7.4}$$

from which

$$m^*/m = P^2\omega/(E_p - E_o) \tag{7.5}$$

which, in principle, should be independent of P for small P.

The basic soundness of our effective-mass computation is supported by the fact that our results coincide exactly with the perturbation-theoretic result (3.6) for weak coupling and closely approximate the Lee-Low-Pines intermediate-coupling result (3.15) for $g^2 = 1$. One side-comment is appropriate in this connection. For the case $P=0$ we pointed out that the Lee-Low-Pines computation corresponds to a trivial limit in the context of our work (i.e. to the case where the basis set includes only the vacuum state). The same correspondence does not hold in the case $P\neq 0$, because our unitary transformation (5.4) is no longer equivalent to the Lee-Low-Pines transformation (3.9). In our case the perturbation arising from non-vanishing P must be taken into account explicitly, rather than being described to the first approximation by the unitary transformation. The close correspondence between the two results (in the region where Lee-Low-Pines is expected to give a good approximation) is then not the trivial observation it would be in the special case $P=0$.

Feynman[2] points out that in his approach the effective-mass computation is attended by difficulties not present in the special case $P=0$. In particular, the Feynman <u>energies</u> for $P\neq 0$ do <u>not</u> represent upper bounds. In any case, the effective mass, even when computed from upper-bound approximations to the energy (such as Lee-Pines' results, or ours), is not itself an upper (or lower) bound on the exact result. This circumstance complicates the interpretation and comparison of results.

The error-analysis of the effective-mass results involves an additional degree of speculation that was not present in the analysis of the P=0 results. Consider an elementary study of the propagation-of-error from the energies E_P and E_o to the effective mass ratio m^*/m. From (7.5) one finds

$$\frac{d(m^*/m)}{(m^*/m)} = \frac{d(E - E_o)}{(E - E_o)} \tag{7.6}$$

$$\simeq \frac{E}{E - E_o}\left(\frac{dE}{E} - \frac{dE_o}{E_o}\right) \tag{7.7}$$

To minimize the error in m^*/m one would then want to choose a basis that gives E_P and E_o with comparable accuracy (so that the second factor in (7.7) is small; and one would want to compute m^*/m from (7.5) with P as large as possible (but within the range where E_P depends quadratically on P) in order to maximize $E-E_o$.

Because the second factor in (7.7) may have either sign, the fractional error in m^*/m may be positive or negative; this shows that our approximation does not have an upper- or lower-bound characteristic. As this factor may also vanish, it is possible to arrive at the <u>exact</u> result for m^*/m, even when E and E_o are known only crudely. We can in no way derive (or estimate) a <u>lower bound</u>

on the magnitude of the fractional error in m^*/m. An upper-bound estimate on the error in m^*/m must be based on the possibility that there is (effectively) no cancellation between the two terms in the factor $dE/E - dE_o/E_o$. The factor $E/(E-E_o)$ is quite large (typically of order 50 or larger). If the error in the energy is comparable to that estimated earlier (about 1.0 per-cent), then in the worst of circumstances the effective mass as given by our computation has a fractional error of order 50 per-cent. This is certainly an overly-pessimistic view, particularly considering the close agreement between our results and other reliable computations for weak and moderate coupling. For the basis described earlier and used in these computations, the fractional errors dE/E and dE_o/E_o are probably close in magnitude; (because of the variational character of the energy, they must be of the same sign). The difference $dE/E - dE_o/E_o$ is probably then an order of magnitude smaller than either of these terms: i.e. the difference is perhaps of order 0.1 per-cent. Our value m^*/m may then have a fractional error of order 5.0 per-cent. These estimates do not carry the same confidence as our earlier estimates of the error in E_o.

We are now ready to give a few numerical examples to illustrate the above discussion.

Consider first the weak-coupling limit (e.g. $g^2 = 0.2$). The energy E_o (i.e. for P=0) in this case is found by our computation to be $E_0 = -0.20062$, in very close agreement with the "second" perturbation-theoretic computation (3.19) which predicts $E_o =$ -

-0.20064. In this weak-coupling limit we computed E_P for a range of P as given below:

P	m*/m
0.10	1.0344
0.20	1.0346
0.40	1.0357
0.60	1.0376
0.80	1.0408
1.00	1.0458
1.20	1.0539
1.40	1.0686
1.60	1.0987
1.80	1.1613
2.00	1.2564

The second column gives m*/m for various P under the assumption that the energy depends strictly quadratically on P (i.e. m*/m as given by (7.5)). One sees that m*/m remains nearly constant for a conveniently wide range of P. To say something specific, m*/m remains constant to within 1.0 per-cent for P as large as 0.80. The small-P limit m*/m = 1.034 is in close agreement with the weak-coupling result (3.6). The trend of the figures in the second column above for larger P shows that the energy drops below the curve given by the strict quadratic dependence (7.4), if P is not small.

A corresponding set of results for $g^2 = 1.00$ shows the same qualitative behavior, and gives an effective mass ratio $m^*/m = 1.19$. This is in close agreement with the perturbation-theoretic result $m^*/m = 1.20$ and the Lee-Low-Pines approximation $m^*/m = 1.165$. The close agreement between our result and the others in this case and in the previous case gives a good indication that the numerical errors in our computation are much smaller than a very conservative error-analysis might lead one to believe.

For $g^2 = 3$, we have the sequence of results as follows:

P	m^*/m
0.10	1.778
0.20	1.783
0.40	1.800
0.60	1.829
0.80	1.871

The limiting value for small P, $m^*/m = 1.77$ is in remarkably close agreement with Feynman's $m^*/m = 1.78$, and in reasonable agreement with Lee-Pines' $m^*/m = 1.61$. Lee-Low-Pines predict $m^*/m = 1.50$.

For $g^2 = 5$ (which is the case we have studied in some detail) a similar sequence yields the result $m^*/m = 2.59$, which is intermediate between Feynman's $m^*/m = 3.56$ and Lee-Pines' $m^*/m = 2.15$. If Feynman's result is correct, our result is in error by 40 per-cent. Our error-analysis indicates that this is not inconceivable, but is at the same time not likely.

For $g^2 = 7$, our result is $m^*/m = 3.24$, which is much closer to Lee-Pines' $m^*/m = 2.82$ than to Feynman's result $m^*/m = 13.2$. For this case, intermediate-coupling theory predicts 2.16. The wide discrepancy between these various results attests to the special difficulties attending the effective-mass computation, as compared to the simpler computation for determing E_o. The result of our analysis suggests that Feynman's result for the effective mass is in this instance quite likely to be in error by at least 50 percent.

The results discussed here are summarized in Table 6.

Computations with Alternative Bases

We conclude this chapter with a brief discussion of some side-computations that shed light on the validity of our assumptions and the accuracy of our computation.

The radial functions $R_{NL}(x)$ as given by (4.22) include a non-linear variational parameter Z. Although Z enters into the radial functions in an apparently very complicated way, the resulting matrix elements (see Appendix 1) depend on Z in a fairly simple way (i.e. they turn out to be proportional to various powers of Z). We indicated that our computations would be made initially with Z=1, so that our ground radial state N=0 corresponds exactly to that of Lee-Low-Pines.

TABLE 6: Polaron Effective Mass

$g^2 =$	0.20	1.00	3.00	5.00	7.00
Z.*	1.034	1.19	1.77	2.59	3.24
LLP.*	1.033	1.17	1.50	1.83	2.17
LP.*			1.61	2.15	2.82
F.*			1.78	3.56	13.2

Z. = Zimmermann (this study).

LLP = Lee-Low-Pines[6]

LP = Lee-Pines[7]

F = Feynman[2]

The question arises whether our results could be improved (or, in fact, whether they change at all) by a different choice of Z. Because of the non-linear character of Z, it is not possible to compute the "best" value of Z in the same way it is possible to compute the optimal value of the other 900 or so linear coefficients. One must resort to doing a series of repeated computations, each independent of the others and for different Z. One can then plot out the dependence of the energy on Z and roughly determine the optimum point.

If it were to turn out that the energy depends very sensitively on Z, this would be a very inefficient procedure. Now in principle the energy should be independent of Z if sufficiently many terms from the radial sequence N=0,1,2,... are included in the basis. This is true because the radial states (4.22) form a complete set for any Z. (The case Z=0 must be excluded because then certain integrals diverge). Inasmuch as our basis is restricted in certain respects with regard to radial excitations, one expects to find a weak dependence of the energy on Z. This expectation is confirmed by an examination of the results.

The nature of the dependence of E on Z will be illustrated by a sequence of results for $g^2 = 4$. These results are summarized in the table below:

Z	E_o
0.65	-4.29408
0.75	-4.30177
0.85	-4.30075
0.95	-4.29490
1.00	-4.28765

One finds that the optimal Z in this case in near Z = 0.75. These results show the weak but non-trivial dependence of E_o on Z that we expected to find. Results for other coupling strengths are qualitatively similar. Because of the weak character of this dependence, it is not necessary to choose a finer mesh of points on the Z-scale than is indicated by the tabulation above. Any change in E_o that may result from a more accurate determination of Z would be small compared to the errors from other sources. All results quoted in this chapter, including the detailed study for the case $g^2 = 5$, were derived from a basis with the optimal value of Z for that coupling constant, as determined by a tabulation similar to the one given above.

A second study was made which was directed toward the same question of the suitability of the radial basis. In this study the phonons were restricted to one radial mode, but the functional form of this radial mode $R_{OL}(x)$ was taken in a very general form. Consider a function $r_{OL}(x)$ constructed from our radial basis $R_{NL}(x)$

by a linear expansion:

$$r_{OL}(x) = \sum_{N=0}^{N'} c_N R_{NL}(x) \tag{7.8}$$

where the maximum index N' is something like N'=3 or N'=4. Using $r_{OL}(x)$ in the role formerly played by $R_{OL}(x)$, one can, for a given set of coefficients c_N, compute the matrix elements of the Hamiltonian. For a given set of coefficients one can then determine the ground state energy E_o. By varying the coefficients c_N one can optimize the energy by the same procedure that we optimized the energy as a function of Z. This computation is a little more complicated because we must search out a space of three of four parameters instead of a one-parameter space. Unlike the previous case, one finds the dependence of E_o on the coefficients c_N is in this case very sharp. In particular, one finds for all coupling-strengths that <u>the functional form of the ground state chosen</u> (i.e. $R_{OL}(x)$) <u>is almost exactly the optimal choice</u>. To put it otherwise, the optimal choice of coefficients is $c_1 = 1$; $c_2 = c_3 = \ldots = 0$. A relatively small departure from this optimal condition results in significantly inferior approximations to E_o.

Still another study was directed toward checking the validity of the assumption described by (4.16). We concluded there that it is suitable to choose the radial modes $R_{OL}(x)$ for different L to be proprtional to the factor $(x/(1+x^2))^L$. This choice was

based on a heuristic argument that may not seem entirely convincing. Using a moderate-sized basis, a number of computations were made, for various coupling strengths, in which the L-dependence we are discussing was based on simple alternatives to the form (4.16). The result of these studies was again that the choice originally made gave superior results to other simple alternatives.

For a given choice of phonon modes, and for a given basis, it is possible to arrive at different conclusions with regard to the probable error, according to what particular systematic sequence of results is studied. Our results were arrived at by using a sequence that made use of the notion of "rank." Another alternative is to arrange the states in sequences according to the highest angular momentum mode occupied: i.e., after states with only S-phonons, we have states with S and P phonons, then S,P, and D, phonons, and so forth. For $g^2 = 6$, a typical set of results is quoted here.

S only	-6.000	
		-0.5145
S,P	-6.5145	
		-0.1321
S,P,D	-6.6466	
		-0.0199
S,P,D,F	-6.6665	

The projected probable error derived from these sequences is typically a little smaller, but of the same order, as when derived from the "rank" sequences.

A review and discussion of these results will be included in the following chapter.

VIII. SUMMARY AND EVALUATION

The concept with which we have approached the polaron problem is, in essence, a very simple one. The basic notion is known as the Rayleigh-Ritz variational principle, which refers to the well-known fact that the expectation value of the Hamiltonian of a quantum-mechanical system in a "trial state" $/\phi)$ is stationary as a functional of $/\phi)$ when the trial state is exactly the eigenstate. Moreover, the stationary point is also a <u>minimum</u> point, so that the energy in any approximate eigenstate gives an upper bound on the energy eigenvalue of the system.

To apply this principle, it is necessary to vary the trial state $/\phi)$ in some systematic and numerically accessible way. The simplest procedure, which is the one used here, is to expand the trial state as a linear combination of basis states, which have been chosen earlier. The coefficients of this linear expansion then become the variational parameters, and the energy of the system is a quadratic function of these parameters. The variation of the energy with respect to these coefficients leads to a system of linear equations. The eigenvalue E is determined by the condition that this linear, homogeneous system has a non-trivial solution.

A very straightforward application of this procedure to the polaron problem leads immediately to an astronomically large number of basis states, and to a numerical computation several orders of magnitude beyond the scope of available computers and known techniques.

The difficulties presented by the very large scale of the problem are attacked simultaneously from two directions. On the one hand, it is possible to find a sequence of unitary transformations, which may be carried out explicitly, and which bring the Hamiltonian (in a sense) closer to diagonal form. This procedure may also be looked at as incorporating certain factors explicitly in the trial state, analogously to well-known procedures in the solution of elementary quantum-mechanical problems. With these devices the problem may be reduced considerably from its original size. It is still necessary to pay careful attention to the numerical techniques used for the solution of this problem, in order to make efficient use of the capacity of modern computers. By reducing the scale of the problem as much as possible by analytic transformations, and by stretching the practical techniques for the numerical solution to their limits, it is possible to proceed systematically with the computation.

The choice of the basis for the fundamental expansion of the trial state is a crucial step. Underlying this basis is a representation of phonon wave functions, or modes. The states then describe the pattern of occupations of these modes. In the selection of these modes we were guided principally by the results of previous computations, particularly those of Lee, Low, and Pines[6], and Lee and Pines.[7] In an important sense our computation may be

thought of as a natural extension of these calculations. In these computations, the optimal functional form of the phonon modes are determined explicitly under certain very restricted conditions.

Having determined the phonon modes, one must select the pattern of allowed occupations in some systematic and orderly way. We were guided in this task by a scheme for the classification of states. An important role in this scheme was played by an index, R, called the rank, which describes the pattern of the distribution of phonons among the modes with various angular momenta. This concept allows us to give practical meaning to the intuitive notion that the lower angular momenta are most important.

The selection of basis states was considerably simplified by a unitary transformation referred to earlier. Considered as a transformation of the <u>states</u>, the effect of this unitary operator was to distribute phonons in a particular way into the "ground state" N=0, L=0. Whereas this particular mode had played a predominant role, the unitary transformation had the effect of reducing the mode N=0, L=0 to very minor importance. That is to say, the predominantly important basis vectors were now states with no phonon occupations of this mode. Of these states, the most important states are those with all phonons in the leading radial mode N=0. States of this type were used as a point of reference

in our discussion, and were referred to as "skeleton states." Other states were thought of as being derived from these states by "radial excitations," or by additions of S-phonons, or both.

The numerical techniques suitable to the solution of the problem naturally involve a sequence of approximations, each utilizing a basis generalized from the previous instance by the inclusion of an additional set of states. The corresponding sequence of approximate eigenvalues, and the increments in the eigenvalues corresponding to the various sets of states, served as the basis for our analysis. This analysis was in the first instance directed toward selecting the optimal basis within the given restrictions on the size of the basis. Having determined such a basis, the analysis was directed toward estimating the residual error in the results arising from the finite nature of the computation. This analysis was carried out by arranging the sets of states in a systematic way, and by extrapolating the sequences of increments corresponding to these sequences of states.

By a progressive sophistication of techniques it was possible to extend the problem to a basis with more than 900 states. That is to say, it was possible to determine accurately the ground state energy eigenvalue of a Hamiltonian matrix of dimensionality 900. With this basis, we were able to obtain approximations (and upper bounds) to the ground-state energy E_o (for $P=0$) of the polaron for an important range of coupling strengths. The results obtained

were superior to (i.e. lower than) the results of other methods for coupling stregths up to $g^2 = 7$, approximately. The method is applicable in principle to stronger coupling, but within the limitations imposed by finite computer memory it gives inferior results.

It was possible to determine approximations to the effective mass, m^*, for the same range of coupling strengths. These results, however, do not have an upper- or lower-bound characteristic. This defect is shared by most other approaches to the polaron effective mass.

The error analysis was carried out according to the procedure outlined. The results for weak and moderate coupling were relatively more accurate than for stronger coupling, and were in agreement with the results of other computations (where the other computations were valid). For example, our result for $g^2 = 1$ is in close agreement with the "corrected" Lee-Low-Pines computation (see equation (3.19)). The estimated error grew sharply with increasing coupling strength; the fractional error (estimated) varied from 0.1 per-cent for $g^2 = 1$ to about 1.0 per-cent for $g^2 = 7$.

The error-analysis for the effective-mass computation was more problematical. Because m^* is not a "vatiational" quantity, the error may be of either sign, or, in fact, may vanish altogether. It is therefore impossible to get a lower-bound estimate on the magnitude of the error. It is possible to get an upper-bound

estimate on the error in m*, but this estimate is very weak and almost certainly leads to a grossly pessimistic result.

In comparison to previous approaches to the polaron problem, the biggest advantage of the procedure used here is that, in principle, it gives us, along with the results, a possibility for estimating the accuracy of these results. Whereas other approaches have involved certain restrictive assumptions whose effect on the results cannot be estimated, the only restrictions in our procedure are the restrictions on the size of the basis, made necessary by practical considerations. A large part of our effort was devoted to estimating the effect of this truncation of the basis. While our estimates, which were based on the sequence of intermediate results, were not rigorous, they were at the same time meaningful.

The conclusion of these studies gives an error estimate that is generally small enough to make the results useful, but larger than one would hope or expect from a computation on such a very large scale. At the outset it was certainly intended that a computation with a basis of many hundreds of states (and in particular with nearly 1000 states) would yield results with probable errors smaller than the 1.0 per-cent estimated in a typical case. It must be regarded as a disappointment that the scale of the problem is so out of proportion to the accuracy of the results.

This conclusion is underlined particularly by a comparison with Feynman's work. The path-integral formalism, conceptually more sophisticated and more complicated than our approach, is almost trivial from the point of view of the numerical computations involved, and gives results close to ours for $g^2 \lesssim 7$ and actually lower than ours for stronger coupling. This comparison, and the light it sheds on Feynman's approach, is in itself one of the interesting results of our study. It would not have been antidpated that the Feynman method, involving as it does some assumptions that are difficult to evaluate, would in fact be competetive with a variational computation involving many hundreds of parameters.

It is possible that some modifications of our approach, but within the present framework, would result in a more favorable comparison. One alternative is to make further unitary transformations, of the same general character as those already discussed, in order to bring the matrix still closer to diagonal form. One such transformation is suggested by the Lee-Pines calculation. Given their restriction to S- and P-phonons, and certain additional assumptions about the form of the trial state, the resulting problem can be solved explicitly. (This explicit solution still contains such parameters as may have been included in the functions describing the radial modes). It is suggested that the unitary transformation that affects this solution might be

incorporated into our computation in the same sense as the Lee-Low-Pines transformation was incorporated. In this case the trivial version of our computation would be equivalent to the Lee-Pines calculation, rather than to the less accurate one of Lee-Low-Pines. While this would undoubtedly be of some usefulness, it would also carry with it certain practical difficulties. Each such unitary transformation has the effect of adding certain new classes of terms to the Hamiltonian operator; the effect of these terms from a computational point of view is to increase considerably the density of the Hamiltonian matrix (i.e. the number of non-zero elements for a matrix of given size). Unless the transformation is of significant usefulness in itself, the loss may be greater than the gain.

Another possibility is to look still more carefully at the choice of radial modes. It is regarded as likely that some improvement could be made in this respect, but it is also unlikely that the improvement would be very great.

As our interest was primarily methodological, we would point out that many of the difficulties in this computation may be regarded as arising from specific characteristics of this particular problem which may not be present in other important problems. It is possible to invent a number of simple modifications of the polaron problem, for which results of comparable accuracy could be achieved with a much smaller basis. Alternatively, for these model problems much more accurate results could be obtained for a basis as large as the

one used here. One example of such a model is that used by Devreese and Evrard[13,14] to study the possible existence of excited bound states of the polaron. In this model, phonons may be emitted with only one momentum k, or its opposite, -k. An analogous model in angular-momentum space may be constructed by restricting the angular-momentum vector to a plane (M=0 for all phonons). For any problems with such characteristics, or for problems for which these models are meaningful approximations, the procedure studied here may be of very great usefulness.

We have already said that one of the interesting results of our work is the light it sheds on Feynman's approach. This suggests it may be possible to find an approach that combines the conceptual eleagance and apparent power of the path-integral formalism with a method for evaluating the probable error of the results. One possible approach may be based on the numerical evaluation of the path-integrals instead of resorting to the "model problem." The numerical evaluation of such path integrals has already been studied by several authors, but not with particular reference to this problem.

ACKNOWLEDGMENTS

I am grateful to Professor Charles Schwartz for drawing my attention to this problem, for fundamental suggestions with regard to the solution, and for patient assistance with the computation.

I also want to express my thanks to Dr. David L. Judd, Director of the Theoretical Group at the Lawrence Radiation Laboratory (Berkeley), for allowing me to complete this research at the Laboratory. In this connection, I am also thankful for the support of the Computer Center at the Laboratory, without which this work could not have been carried out. In particular, I want to thank Mr. Chuck Stevenson and Mr. Dave Stevens for valuable assistance with programming.

This research was carried out with the financial support of the Atomic Energy Commission and the United States Air Force Office of Scientific Research.*

* Grant No. AF-AFOSR-130-63

APPENDIX A

In this Appendix we compute the various integrals that arise when the Hamiltonian is expressed in terms of the discrete modes defined by the creation and destruction operators (4.28). These integrals are given by (4.37), (4.38), (4.39), and (4.40). They are

$$T_{ij} = \int d^3x \, u_i^*(\bar{x}) \, (1 + x^2) \, u_j(\bar{x}) \tag{A.1}$$

$$J_i = \int d^3x \, u_i(\bar{x}) \, V(x) \tag{A.2}$$

$$Q(i,j,k,l) = \int d^3x \int d^3x' \, u_i^*(\bar{x}) \, u_j^*(x') \, \bar{x}\cdot\bar{x}' \, u_k(\bar{x}) \, u_l(\bar{x}') \tag{A.3}$$

$$K_{ij} = \int d^3x \, u_i(x) \, \bar{x} \, u_j(x) \tag{A.4}$$

With the aid of (4.5), i.e.

$$u_i(\bar{x}) = u_{L_i M_i N_i}(\bar{x}) = R_{N_i L_i}(x) \, Y_{L_i M_i}(\Omega) \tag{A.5}$$

one can write

$$T_{ij} = \delta_{ij} + \delta_{L_i L_j} \delta_{M_i M_j} \int_0^\infty x^2 dx \, R_{N_i L_i}(x) \, x^2 \, R_{N_j L_j}(x) \tag{A.6}$$

Similarly

$$\bar{K}_{ij} = \delta_{L_i, L_j \pm 1} \int_0^\infty x^2 dx \, R_{N_i L_i}(x) \, \bar{x} \, R_{N_j L_j}(x) \tag{A.7}$$

Making use of

$$\bar{x} \cdot \bar{x}' = x \, x' \, P_1^0(\cos \phi_{xx'}) \tag{A.8}$$

$$= (4\pi/3) \, x \, x' \sum_m{}' \, Y^*_{1m}((\Omega)) \, Y_{1m}(\Omega) \tag{A.9}$$

we can write

$$Q(i,j,k,l) = (4\pi/3) \left(\int_0^\infty x^2 dx \, R_{N_i L_i}(x) \, x \, R_{N_k L_k}(x) \right) \left(\int_0^\infty x^2 dx \, R_{N_j L_j}(x) \right.$$

$$\left. x \, R_{N_l L_l}(x) \right) \sum_m{}' \, S_{ik}(m) \, S_{jl}(-m) \, (-1)^m \tag{A.10}$$

where in turn

$$S_{ij}(M) = \int d\Omega \, Y^*_{L_i M_i}(\Omega) \, Y^*_{1,M}(\Omega) \, Y_{L_j M_j}(\Omega) \tag{A.11}$$

These integrals over three spherical harmonics may be expressed in terms of Clebsch-Gordon coefficients, and are well-known. If we define

$$I(N,L;\, N',L',\, s) = \int_0^\infty x^2 dx \, R_{NL}(x) \, x^s \, R_{N'L'}(x) \tag{A.12}$$

then the integral in (A.10) is $I(N,L;N',L';\ 2)$ and the integral in (A.11) is $I(N,L;\ N',L',\ 1)$. The normalization for the radial functions $R_{NL}(x)$ may then be written

$$I(N,L;\ N,L,\ 0) = 1 \tag{A.13}$$

The integrals we need to compute are then J_i (A.2) and $I(N,L;\ N',L',\ s)$ for s=0, s=1, and s=2. In working out these integrals we will repeatedly encounter a particular definite integral. For convenience we define

$$B(J,K) = \int_0^1 (1 - v)^J\, v^K\, dv = \Gamma(J+1)\,\Gamma(K+1)/\Gamma(J+K+2) \tag{A.14}$$

provided $J > -1$ and $K > -1$.

Using the definition (4.22) of the radial functions

$$R_{NL}(x) = C_{NL}\left(\frac{x}{1 + Z^2x^2}\right)^L \frac{1}{x\,(1 + Z^2x^2)}\, P_{NL}(v(x)) \tag{A.15}$$

one can write

$$I(N,L;\ N',L';\ s) = \frac{C_{NL}\, C_{N'L'}}{2\, Z^{L+L'+s+1}} \int_0^1 (1 - v)^{\frac{L+L'+1+s}{2}} (v)^{\frac{L+L'+s-1}{2}}\, P_{NL}(v)\, P_{N'L'}(v)\, dv \tag{A.16}$$

In the ortho-normality condition (A.13) this integral occurs with $L=L'$, $s=0$. We are therefore concerned with polynomials ortho-normal on the interval (0,1) with respect to the weight function

$$w(v) = (1 - v)^{L-1/2} v^{L+1/2} \tag{A.17}$$

These polynomials are defined by the "Rodrigues' Formula"

$$P_{NL}(v) = \frac{1}{w(v)} \frac{d^N}{dv^N} \left(w(v)\, v^N (1 - v)^N \right) \tag{A.18}$$

where $w(v)$ is as in (A.17). These polynomials are known in the literature as the <u>Jacobi Polynomials of the Second Kind</u>, and are conventionally denoted by $G_N(2L+1, L+3/2; v)$. (See, for example, reference 15).

We wish to define these polynomials strictly as given by (A.18): i.e., without a normalization constant. The constant C_{NL} is determined so that the whole radial function $R_{NL}(x)$ is normalized.

If

$$P_{NL}(v) = \sum_{j=0}^{N} a_j^{(N,L)} v^j \tag{A.19}$$

one can easily show

$$a_0^{(N,L)} = \Gamma(N + L + 3/2) / \Gamma(L + 3/2) \tag{A.20}$$

and

$$a_N^{(N,L)} = (-1)^N \, \Gamma(2N + 2L + 1)/\Gamma(N + 2L + 1) \tag{A.21}$$

The normalization integral I(N,L; N,L, 0) may be computed by substituting (A.18) into (A.16) and differentiating by parts N times to obtain

$$I(N,L;\ N,L;\ 0) = \frac{(-1)^N C_{NL}^2}{2Z^{2L+1}} \int_0^1 \frac{d^N}{dv^N}\left[\frac{1}{w(v)} \frac{d^N}{dv^N} \left(w(v)\ v^N\ (1 - v)^N \right)\right] w(v)\ (1 - v)^N\ v^N\ dv \tag{A.22}$$

According to (A.18) the factor in brackets $\left[\quad\right]$ is again $P_{NL}(v)$. Then from (A.19) and (A.21)

$$\frac{d^N}{dv^N} P_{NL}(v) = N!\ a_N^{(N,L)} \tag{A.23}$$

One finds

$$I(N,L;\ N,L,\ 0) = \frac{(-1)^N C_{NL}^2}{2Z^{2L+1}}\ N!\ a_N^{(N,L)}\ B(N+L+1/2,\ N+L-1/2) \tag{A.24}$$

where the last factor is an abbreviation for an integral, as given by (A.14).

After some simplification, the normalization condition (A.13) then leads to

$$C_{NL} = 2\, Z^{L+1/2} \sqrt{\frac{(N + 2L)!}{N!}} \frac{1}{\Gamma(N + L + 1/2)} \tag{A.25}$$

The same general procedure may be followed to compute the integral $I(N,L;\ N',L';\ 2)$ that arises in the term T_{ij}. We can assume without loss of generality $N' > N$. Also, $L=L'$ (otherwise the term does not enter into (A.10)). After substitution and an N'-fold integration by parts

$$I(N,L;\ N',L;\ 2) = \frac{(-1)^{N'}\, C_{NL} C_{N'L}}{2\, Z^{2L+2}} \int \frac{d^{N'}}{dv^{N'}} \frac{1}{v} P_{NL}(v)$$

$$w(v)\ (1 - v)^{N'}\ v^{N'}\ dv \quad - \quad \frac{\delta_{NN'}}{Z^2} \tag{A.26}$$

If $N' > N$,

$$\frac{d^{N'}}{dv^{N'}}\left(\frac{1}{v} P_{NL}(v)\right) = (-1)^{N'}\ N'!\ a_o^{(N',L)}\ v^{-1-N'} \tag{A.27}$$

We have then

$$I(N,L;\ N',L;\ 2) = \frac{C_{NL} C_{N'L}}{2\, Z^{2L+1}}\ N'!\ a_o^{(N',L)}\ B(N' + L - 1/2,\ L - 1/2)$$

$$- \frac{\delta_{NN'}}{Z^2} \tag{A.28}$$

After substitution of C_{NL} (A.25); $a_o^{(N,L)}$ (A.20); and $B(N' + L - 1/2, L - 1/2)$ (A.14); and after considerable simplification, one finds

$$I(N,L;\ N',L,\ 2) = \frac{2}{Z^2}\sqrt{\frac{(N + 2L)!\ N'!}{(N' + 2L)!\ N!}}\ \frac{(2N + 2L + 1)}{(2L + 1)} - \frac{1}{Z^2}\,\delta_{NN'} \tag{A.29}$$

The integrals $I(N,L;\ N',L',\ 1)$ involve only a few additional complications. We are interested in the cases $L' = L + 1$ and $L' = L - 1$. (We assume as before $N' > N$). We denote $I(N,L;\ N',L';\ 1)$ as I_+ and I_- in these two cases, respectively. It is convenient also to write $w_L(v)$ where formerly we wrote $w(v)$, in order to distinguish between the cases for L and L'. We note

$$w_{L'}(v) = w_{L+1}(v) = v\,(1 - v)\,w_L(v) \tag{A.30}$$

Using (A.30) and making the usual substitutions, one finds

$$I_+ = \frac{(-1)^{N'} C_{NL} C_{N'L'}}{2\,Z^{2L+3}} \int_0^1 \frac{d^{N'}}{dv^{N'}}\left(\frac{1}{v}\,P_{NL}(v)\right) (1 - v)^{N'+L+1/2}\ v^{N'+L+3/2}\ dv \tag{A.31}$$

Again after substitution and simplification, one arrives at the result

$$I_+ = I(N,L;\ N',L+1;\ 1) = \sqrt{\frac{N'!\ (N + 2L)!}{N!\ (N' + 2l + 2)!}}\ \frac{(2N + 2L + 1)}{Z} \qquad (A.32)$$

By an analogous procedure, one may show

$$I_- = I(N,L;\ N',L-1,\ 1) =$$

$$= \frac{C_{NL}\ C_{N',L-1}}{2\ Z^{2L+1}}\ (-1)^N\ N'!\ a_N^{(N,L)}\ B(N'+L'-1/2,\ N'+L'+1/2) \qquad (A.33)$$

if $N' = N+1$, and $I_- = 0$ otherwise. With this stipulation

$$I_- = \sqrt{\frac{N + 1}{N + 2L}} \qquad (A.34)$$

It remains only to compute the integral J_i (A.2). Using (A.9) and the definition

$$V(x) = \sqrt{\frac{4\pi g^2}{(2\pi)^3}}\ \frac{\omega}{(2m\omega)^{3/4}}\ \frac{1}{x} \qquad (A.35)$$

one can show

$$J_i = \sqrt{\frac{1}{4\pi}} \int_0^\infty R_{NO}(x)\, V(x)\, x^2 dx \tag{A.36}$$

$$= C_{NO} \sqrt{\frac{2g^2}{\pi}} \int \frac{1}{1 + Z^2 x^2} P_{NO}(v(x))\, dx \tag{A.37}$$

if $L = L_i = 0$. Otherwise, $J_i = 0$. We have written N for N_i in (A.37). Changing variables from x to v, and substituting (A.18) for $P_{NO}(v)$, one finds

$$J_i = D \int_0^1 v^{-1/2} (1 - v)^{-1/2} P_{NO}(v)\, dv \tag{A.38}$$

$$= D \int \frac{1}{v} \frac{d^N}{dv^N} \left(v^{N+1/2} (1 - v)^{N-1/2} \right) dv \tag{A.39}$$

where

$$D = \frac{C_{NO}}{2Z} \sqrt{\frac{2g^2}{\pi}} = \sqrt{\frac{2g^2}{\pi Z}} \frac{1}{\Gamma(N + 1/2)} \tag{A.40}$$

After and N-fold integration by parts,

$$J_i = D\, N!\, B(N - 1/2, -1/2) \tag{A.41}$$

which, after simplification, turns out to be independent of N:

$$J_i = \sqrt{\frac{2g^2}{Z}} \tag{A.42}$$

This completes the list of integrals that are needed for the matrix elements.

REFERENCES

1. H. Frohlich, H. Pelzer, and D Zienau, Phil. Mag. 41, 221 (1950).
2. R. P. Feynman, Phys. Rev. 97, 660 (1955).
3. G. R. Allcock, Strong-Coupling Theory of the Polaron, in Polarons and Excitons, edited by C. G. Kuper and G. D. Whitfield, (Plenum Press, New York, 1963). p. 46.
4. M. Gurari, Phil Mag. 44, 329 (1953).
5. S. V. Tiablikov, Zh. Eksp. i. Teor. Fiz., 26 (1954)
6. T. D. Lee, F. Low, and D. Pines, Phys. Rev. 90, 297 (1953)
7. T. D. Lee and D. Pines, Phys. Rev. 92, 883 (1953).
8. H. Frohlich, Introduction to the Theory of the Polaron, in Polarons and Excitons (see Ref. 3).
9. R. P. Feynman, Revs. Mod. Phys. 20, 367 (1948).
10. I. M. Gel'fand and A. M. Yaglom, J. Math. Phys. 1, 48 (1960).
11. W. T. Martin and I. Segal (editors), Analysis in Function Space, (The M.I.T. Press, Cambridge, 1964).
12. T. D. Schultz, Phys. Rev. 116, 526 (1959).
13. J. Devreese and R. Evrard, Physics Letters 11, 278 (1964).
14. J. Devreese and R. Evrard, Phys. Stat. Sol. 9, 403 (1965).
15. U. W. Hochstrasser, Orthogonal Polynomials, in Handbook of Mathematical Functions, edited by Milton Abramowitz and Irene A. Stegun, (National Bureau of Standards, 1964).
16. Richard S. Varga, Matrix Iterative Analysis (Prentice Hall, Inc., Englewood Cliffs, 1962).

This report was prepared as an account of Government sponsored work. Neither the United States, nor the Commission, nor any person acting on behalf of the Commission:

A. Makes any warranty or representation, expressed or implied, with respect to the accuracy, completeness, or usefulness of the information contained in this report, or that the use of any information, apparatus, method, or process disclosed in this report may not infringe privately owned rights; or

B. Assumes any liabilities with respect to the use of, or for damages resulting from the use of any information, apparatus, method, or process disclosed in this report.

As used in the above, "person acting on behalf of the Commission" includes any employee or contractor of the Commission, or employee of such contractor, to the extent that such employee or contractor of the Commission, or employee of such contractor prepares, disseminates, or provides access to, any information pursuant to his employment or contract with the Commission, or his employment with such contractor.