overlap between neighboring $3p$ orbitals are included, the band width is decreased by an estimated 30%. It is emphasized that the estimation procedure does not account completely for overlap since the effects of nonorthogonality to core states have not been considered. The interaction between the $3s$ and $3p$ bands was estimated in the approximation that the Bloch sums are assumed orthonormal. Its effect on the width of the $3p$ band is relatively small, of order 5%. The interaction is zero for some of the eigensolutions by reasons of symmetry. Finally, the inclusion of second-nearest neighbors in the two-center approximation produces an effect of order 10%.

IV. CONCLUSION

The analysis results in a lower bound of 1.0 ev for the width of the halogen band in NaCl to within about 30%. Owing to the extreme nature of the assumed model of a pseudo-argon lattice, the actual value of the width is believed to lie much closer to Shockley's result, which probably represents an upper bound since, as has been mentioned, his calculation is based upon Cl^- orbitals determined without exchange.

V. ACKNOWLEDGMENTS

The author wishes to express his deepest appreciation to Professor Frederick Seitz, who suggested the problem, for his stimulating advice and encouragement, and for his reading of the manuscript. He also wishes to thank Dr. Elihu Abrahams and Dr. Truman Woodruff for helpful discussions and suggestions. A portion of the work was carried out while the author held a Horizons, Inc. fellowship. He is pleased to acknowledge this assistance.

PHYSICAL REVIEW VOLUME 104, NUMBER 5 DECEMBER 1, 1956

Critical Points and Lattice Vibration Spectra*

JAMES C. PHILLIPS†
Department of Physics and Institute for the Study of Metals, University of Chicago, Chicago, Illinois
(Received June 11, 1956)

A method is developed for obtaining directly from group-theoretical and topological considerations a consistent set of critical points belonging to a given secular equation. The treatment includes a detailed discussion of nonanalytic behavior near degeneracies. Comparison with many specific calculations shows that the minimal set so obtained is often the set actually present in the case of short-range forces. The contributions of nonanalytic critical points to the frequency distribution are analyzed. Behavior near critical points is made the basis of a scheme for calculating the main features of frequency distributions from the corresponding minimal sets. An approximate frequency distribution is calculated for aluminum. Possible applications to energy bands are discussed.

INTRODUCTION

THE elastic frequency distribution of a crystal, which determines many of the thermodynamic and optical properties of the crystal, may be defined as $g(\nu)$, where $g(\nu)$ is the fractional number of frequencies between ν and $\nu+d\nu$. The frequency distribution is given by the usual Born-von Kármán theory of crystal dynamics, which treats the lattice as a set of coupled harmonic oscillators. The resulting secular equation is generally too involved to make it practicable to obtain the frequency distribution in closed form; it has proved necessary to resort to various approximate methods. Many investigators have obtained approximate frequency distributions by taking finite samples of the elastic vibrations throughout the Brillouin zone.[1] The objection to this method is that one must take a large sample to obtain good results, which makes the method quite laborious. Some time ago Houston[2] and Montroll[3] proposed analytic methods of approximation which met with fair success. Recently interest in analytic approximations has revived as a result of papers by Montroll,[4] Smollett,[5] and Van Hove.[6] Montroll succeeded in making an exact calculation of the frequency distribution of a special square lattice, and he obtained a distribution containing logarithmic singularities. Smollett showed that such singularities would arise whenever the frequency surfaces in k space had analytic saddle points. Van Hove then demonstrated, by using a topological existence theorem due to Morse,[7] that certain critical points (c.p., points at which $\nabla_k\nu=0$)

* Based on a thesis submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Physics Department of the University of Chicago.

† Now at Bell Telephone Laboratories, Murray Hill, New Jersey.

[1] See reference 1 of L. Van Hove, Phys. Rev. 89, 1189 (1953).

[2] W. V. Houston, Revs. Modern Phys. 20, 161 (1948).

[3] E. W. Montroll, J. Chem. Phys. 10, 219 (1942).

[4] E. W. Montroll, J. Chem. Phys. 15, 575 (1947).

[5] M. Smollett, Proc. Phys. Soc. (London) A65, 109 (1952).

[6] L. Van Hove, Phys. Rev. 89, 1189 (1953).

[7] M. Morse, *Functional Topology and Abstract Variational Theory*, Memorial Sciences Mathematiques (Gauthier-Villars, Paris, 1938), Fascicule 92.

must exist generally, together with the consequent analytic singularities in the frequency distribution.

Van Hove's results predict the existence of *some* c.p. If one knew *all* the c.p., then an interpolation scheme, such as the moment-singularity method proposed by Lax and Lebowitz[8] and by Rosenstock,[9] might suffice to give a good approximation to the distribution. Rosenstock[9] has tried to obtain most of the c.p. for a given secular equation by drawing frequency contours on all the symmetry planes. The drawing of contours is a laborious procedure. Accordingly, we address ourselves in this paper to the problem of determining the most complete statement that can be made about c.p. from general mathematical considerations alone.

In Secs. I, II, and III we present the mathematical tools required in our solution of this problem. We start with the observation that the power of group theory for the location of c.p. required by symmetry alone[10] has not yet been exploited for lattice vibrations. In Sec. I we give a brief description of these group-theoretical techniques. Having located the c.p. required by symmetry, one can make little further progress until these c.p. have been classified. In Sec. II we present perturbation techniques and schematic representations suitable for such classification. We emphasize particularly the existence and classification of nonanalytic behavior near c.p. We then observe that Morse's results actually contain more detailed and powerful topological relations than those used by Van Hove. Moreover, although applied by Morse explicitly only to analytic functions, his general topological methods permit us to derive similar relations for nonanalytic functions. Because of the highly mathematical character of the arguments involved, we have relegated an outline of the proofs of the necessary theorems to the appendix. Section III, on the other hand, contains a description of these theorems which in itself should be sufficient for an understanding of their use. In particular we discuss there the topological classifications of all c.p. discussed in Sec. II.

We present in Sec. IV an explicit method for constructing the so-called minimal set of c.p. and in Sec. V use this method to construct these sets in a number of examples. In Sec. VI we discuss the contribution of nonanalytic c.p. to the frequency distribution. Finally, in Sec. VII the contributions of c.p. are made the basis of a scheme for approximating the frequency distribution. We summarize the method and give a brief discussion of possible applications to energy bands in Sec. VIII.

Before we proceed to the actual solution of the problem, let us review the results of the Born-von Kármán theory. If there are N unit cells in the lattice labeled by n and n', with Z atoms per unit cell, labeled by α, β, and the space has l dimensions, the NlZ equations of motion may be written

$$m_\alpha \frac{d^2\mathbf{x}_\alpha(n)}{dt^2} = \sum_{\beta,n'} \Phi_{\alpha\beta}(n,n')[\mathbf{x}_\beta(n') - \mathbf{x}_\beta(n)]. \quad (1)$$

If we assume

$$\mathbf{x}_\alpha(n) = \mathrm{Re}\left[\boldsymbol{\varepsilon}_\alpha \frac{b}{\sqrt{N}} \exp[i(\mathbf{k}\cdot\mathbf{n} - 2\pi\nu t)]\right], \quad (2)$$

then the translational symmetry of the lattice enables us to reduce Eqs. (1) to an lZ-dimensional secular equation for ν^2. Here b represents the complex amplitude of the wave, $\boldsymbol{\varepsilon}_\alpha$ is a polarization vector, and $\mathbf{k}$ is the wave vector of the vibration. The lZ-dimensional matrix $M(\mathbf{k})$, whose eigenvalues $\nu^2(\mathbf{k})$ provide the lZ branches of the frequency surfaces, has the symmetry and periodicity of the reciprocal lattice. Independent vibrations are obtained by restricting $\mathbf{k}$ to the first Brillouin zone (B.Z.) of the reciprocal lattice. By using the uniform spacing of $\mathbf{k}$ over one zone, it is now easy to show that the normalized frequency distribution[11] $G(\nu^2)$ is given by

$$G(\nu^2) = \frac{V_0}{Zl} \int_{S(\nu^2)} \frac{dS}{|\nabla_k \nu^2|}, \quad (3)$$

where V_0 is the volume of the crystal cell and $S(\nu^2)$ is the surface in the first zone for which $\nu^2(\mathbf{k}) = \nu^2$. The importance of the c.p. at which $\nabla_k \nu^2 = 0$ is evident. We now turn to the methods with which we locate the c.p. necessitated by symmetry considerations alone.

I. GROUP THEORY

The group theory of functions derived from a secular equation having the symmetry and periodicity of the reciprocal lattice has been formulated implicitly by Bouckaert, Smoluchowski, and Wigner[12] in their discussion of energy bands. We shall use their terminology and their abbreviations for the symmetry elements of the B.Z. One of their results is that a solution having wave vector $\mathbf{k}$ must transform as a small representation of $\mathbf{k}$, that is, as an irreducible representation of the group of $\mathbf{k}$. Our solutions (2) are labeled by $\mathbf{k}$ and the polarization set $\{\boldsymbol{\varepsilon}_\alpha\}$. Consequently the set $\{\boldsymbol{\varepsilon}_\alpha\}$ must transform as a small representation. In particular, for one atom per unit cell, we deal with those small representations transforming as one or more components of a vector. One may identify the small representations immediately from their character tables; the latter are given for the cubic lattices in reference 12.

We now consider the behavior of $\nu^2(\mathbf{k})$ in the neigh-

[8] M. Lax and J. L. Lebowitz, Phys. Rev. **96**, 594 (1954).
[9] H. B. Rosenstock, Phys. Rev. **97**, 290 (1955).
[10] See the considerations of G. Dresselhaus, Phys. Rev. **100**, 580 (1955), and R. Parmenter, Phys. Rev. **100**, 573 (1955), on energy bands in the zincblende structure.
[11] We calculate $G(\nu^2)$ since ν^2 is the eigenvalue of the secular equation. One obtains $g(\nu)$ by observing that $g(\nu) = 2\nu G(\nu^2)$.
[12] Bouckaert, Smoluchowski, and Wigner, Phys. Rev. **50**, 58 (1936).

borhood of $\mathbf{k}_0$. Let $u_j(\mathbf{k}_0)$, $j=1, \cdots, n$, be a set of n-fold degenerate solutions ($1\leq n\leq Zl$) belonging to the representation Γ_α of the group of $\mathbf{k}_0$. Then to second degree in $\xi=\mathbf{k}-\mathbf{k}_0$, the $\nu_j^2(\mathbf{k})$ are the eigenvalues of

$$W_{ij}(\xi)=\nu_{i0}^2\delta_{ij}+(i|\,\xi\cdot\nabla_k\nu^2|\,j)+\tfrac{1}{2}(i|\,\xi\xi\colon\nabla_k\nabla_k\nu^2|\,j)+\sum_m\frac{(i|\,\xi\cdot\nabla_k\nu^2|\,m)(m|\,\xi\cdot\nabla_k\nu^2|\,j)}{\nu_{i0}^2-\nu_{m0}^2}. \quad (4)$$

The matrix W is obtained from second-order degenerate perturbation theory.[13] The index m labels modes not belonging to the original set j and $\nu_{l0}^2=\nu_l^2(\mathbf{k}_0)$, $l=i$ or m. We shall call the point $\mathbf{k}_0$ an *ordinary* c.p. if

$$(i|\,\xi\cdot\nabla_k\nu^2|\,j)=0, \quad i,j=1, \cdots, n. \quad (5)$$

The operator $\nabla_k\nu^2$ transforms like a vector. It will generate a representation of the group of $\mathbf{k}$ which, when reduced, will be found to contain a set of small representations Γ_β. In the absence of screw axes and glide planes, the Γ_β will be those small representations labeled by components of vectors.[14] In any case all the matrix elements will vanish unless one of the direct products $\Gamma_\alpha\times\Gamma_\beta\times\Gamma_\alpha$ contains the unit representation. The second order terms include an operator transforming like the trace of a tensor. The diagonal matrix elements of this term will not vanish except for special choices of the force constants. Therefore terms to second order suffice to characterize the frequency surfaces in small neighborhoods.

At a general point in the zone, the group of $\mathbf{k}_0$ contains only the identity. The only small representation of the group of $\mathbf{k}_0$ is the one-dimensional unit representation; there is only accidental degeneracy. Further, we have $(i|\,\xi\cdot\nabla_k\nu^2|\,i)\neq 0$ except by accident, for $\Gamma_\alpha\times\Gamma_\beta\times\Gamma_\alpha$ *is* the unit representation. Thus c.p. are not required to occur at arbitrary $\mathbf{k}_0$ by symmetry conditions. Finally, inspection of Eq. (4) shows that $\nu^2(\mathbf{k})$ is analytic about nondegenerate points.

At points of higher symmetry, representations of higher dimensionality may occur. A systematic test at all symmetry points of the condition that $\Gamma_\alpha\times\Gamma_\beta\times\Gamma_\alpha$ does not contain the unit yields all ordinary c.p. required by symmetry. Further, it yields those points for which some components of the gradient vanish. In applying the topological considerations one may regard these points as potential c.p. Finally, when $n>1$, the $\nu^2(\xi)$ are solutions of a secular equation and hence are analytic only by accident. We must therefore distinguish two kinds of ordinary c.p., the *analytic* and the *fluted* (the fluted points being produced by degeneracies at symmetry points).

There are lZ branches of the frequency surfaces in $\mathbf{k}$ space. The topological considerations to be developed are applicable only to single branches forming closed manifolds. Difficulty in identifying suitable separate branches can arise only at those points of degeneracy where branches cross. This situation is illustrated by the solid lines in Fig. 1. One way to label branches in the vicinity of crossover points, which we shall call the *ordered* labeling, defines the ith branch at $\mathbf{k}$ to be the one of frequency $\nu_i^2(\mathbf{k})$ where $i<j$ implies $\nu_i^2(\mathbf{k})\leq\nu_j^2(\mathbf{k})$. This definition of branches suffers from the difficulty that it can introduce generalized nonanalytic c.p. at crossover points. We call such points *singular* c.p. and define them as points where one or more components of the gradient change sign discontinuously while the remaining components vanish. The topological characterization of c.p. given in the appendix requires that such points be considered c.p.[15] Ordinary c.p. produce infinities in the frequency distribution, or in its first derivative, for $l=2$ or 3, respectively, but singular c.p. produce at most discontinuities in the first derivative of the frequency distribution. As an example of a singular c.p. introduced by the ordered labeling, we note that the branch (1)–(3) in Fig. 1 might have a generalized minimum at $\mathbf{k}_0$.

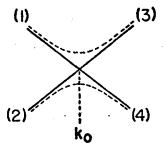

FIG. 1. Frequency surfaces near a crossover point $\mathbf{k}_0$.

Crossover points can occur where degeneracy is necessitated by symmetry but where all components of the gradient are not required to vanish. At symmetry points where several one-dimensional representations exist, accidental degeneracies of the kind pictured in Fig. 1 also can occur.[16] The group-theoretical considerations described above enable one to locate the singular c.p. required by symmetry and to locate within a family of "accidental" crossovers the singular c.p.

It might be thought that it would be more convenient to label the branches in Fig. 1 by (1)–(4) and (2)–(3) so as to avoid introducing singular c.p. However, such a labeling is inconvenient for several reasons. The most important of these is that the degeneracy at $\mathbf{k}_0$ is removed as soon as one leaves the symmetry plane, as is indicated by the dotted lines in Fig. 1. Therefore the alternate labeling would make the surfaces discontinuous. Henceforth we will always define our branches according to the ordered labeling.

II. BEHAVIOR NEAR C.P.

We now turn our attention to a study of the detailed behavior of $\nu^2(\mathbf{k})$ in the neighborhood of ordinary and singular c.p. If the c.p. is analytic, we may expand

[13] P. O. Löwdin, J. Chem. Phys. **19**, 1396 (1951).

[14] For the case of screw axes and glide planes, see G. Dresselhaus, thesis, University of California, 1955 (unpublished).

[15] Our definition of a generalized c.p. differs from Van Hove's (reference 6) because he includes what we have called fluted points among his generalized c.p.

[16] Accidental degeneracies have been discussed in detail by C. Herring, Phys. Rev. **52**, 365 (1937).

$\nu^2(\mathbf{k})$ to second degree in $\xi=\mathbf{k}-\mathbf{k}_0$. In the normal form, the expansion becomes:

$$\nu^2=\nu_0^2-\sum_{i=1}^{j} b_i\xi_i^2+\sum_{i=j+1}^{l} b_i\xi_i^2;\quad b_i>0. \tag{6}$$

Here j is the index of the form; the topological features of an analytic c.p. are completely specified by its index.

No such expansions are available for the fluted points. The geometrical classification which we shall now introduce does not require the existence of such expansions. It is based on the number of separate *sectors* in the neighborhood of the c.p. in which $\nu^2>\nu_0^2$ or $\nu^2<\nu_0^2$. A sector is an angle, or solid angle, taken with the c.p. as apex, within which $\nu^2-\nu_0^2$ is everywhere of the same sign. If $\nu^2>\nu_0^2$, we have a *positive* sector; whereas if $\nu^2<\nu_0^2$, we have a *negative* sector.

The following procedure is a graphic way of counting the numbers of positive and negative sectors. In polar coordinates we have $\nu^2(\xi)-\nu_0^2=\lambda(\xi,\theta,\cdots)$. Consider ξ fixed and sufficiently small that the perturbation expansion in Eq. (5) is valid. Let $\lambda=0$ for all angles be represented by a *reference circle* or *sphere*. For fixed angle let values of λ be plotted radially, positive values outside and negative values inside the reference surface. The radial scale is chosen in any convenient manner. The numbers (P,N) of positive and negative sectors around $\mathbf{k}_0$ may now be read off directly from this sketch.

Let us consider as an example the simplest case, a two-fold degeneracy in two dimensions. We simplify the discussion by assuming here, and throughout the paper, that a center of inversion is present. Since the secular equation must then have the full symmetry of the square, it may be written in the form

$$\begin{vmatrix} b\xi_x^2+c\xi_y^2-\lambda & d\xi_x\xi_y \\ d\xi_x\xi_y & c\xi_x^2+b\xi_y^2-\lambda \end{vmatrix}=|c_{ij}-\lambda\delta_{ij}|=0 \tag{7}$$

in the neighborhood of the c.p. with b, c, and d as general parameters. In Fig. 2 we plot two cases; the λ lines are solid, the reference circle dotted, and all the positive sectors are shaded.

In Fig. 2(a) the case $-c>b>0$ and $d=0$ leads to degenerate saddle points, while in 2(b) the case $-c>b>0$ and $-c-b>d$ produces fluted points. Each of the saddle points in 2(a) has sector numbers (2, 2), but the outer λ line in 2(b) has the numbers (4, 4) while the inner has (0, 1). Evidently the (4, 4) point is intrinsically fluted, for the only analytic combinations in two dimensions are (1, 0), (2, 2), and (0, 1). On the other hand, it is not surprising that the inner line is topologically equivalent to an analytic maximum and that it makes a similar contribution to the frequency distribution.

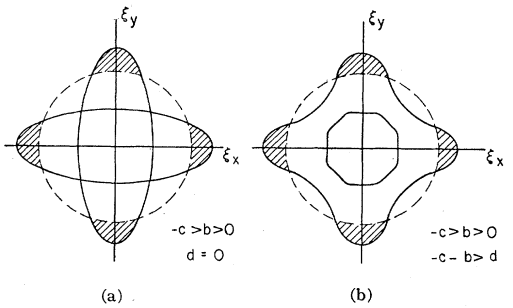

FIG. 2. λ lines about a two-fold degenerate c.p. in two dimensions. (a) Degenerate saddle points. (b) Fluted points.

We can obtain all possible fluted points in two dimensions by observing that each sector is bounded on the reference circle by the intersections of the λ line with the reference circle, points where $\lambda=0$. The angles for which $\lambda=0$ for any mode are obtained by setting $|c_{ij}|=0$ and are independent of ξ. Let us therefore define the line $\lambda'=|c_{ij}|$. Every sector of the reference circle bounded by the points $\lambda=0$ for a single mode will then occur among the set of sectors bounded by the points $\lambda'=0$ on the λ' line. Thus we need only determine the possible sectors of the λ' line. The most general form of λ' is $r(\xi_1^4+\xi_2^4)+s\xi_1^2\xi_2^2$ since it is a fourth degree polynomial having the symmetry of the square. In each sector the λ' line must have a radial maximum or minimum. Thus one can locate all possible sectors by finding the solutions to $\delta\lambda'=0$ subject to $\delta\xi=0$. The only solutions are found to be the $\langle 10\rangle$ and $\langle 11\rangle$ directions. The combinations of sectors in these directions having the symmetry of the square are analytic, warped minima or maxima, and the (4, 4) combination which is illustrated in Fig. 2(b). Thus we have found all fluted points in two dimensions.

We can generalize these ideas to three dimensions by considering a λ surface and its accompanying reference sphere. Consider the two-fold degeneracies first. If the eigenvectors of the two modes lie in a plane of symmetry of the zone, their frequency contours behave in that plane much as do those of the twofold degeneracies in two dimensions. Normal to the plane the behavior is analytic, since in the secular determinant there are no cross terms between the coordinates in the plane and that normal to the plane. Thus, as will be shown later, nothing new is introduced from the topological standpoint. If the eigenvectors do not lie in a plane of symmetry, a separate analysis of the sectors on the reference sphere should be made in each case.

A three-fold degeneracy involves three equivalent modes, and since inversion is present, the secular equation must possess cubic symmetry. Thus it must have the form $|D_{ij}-\lambda\delta_{ij}|=0$, with $D_{ii}=b\xi_i^2+c(\xi_j^2+\xi_k^2)$ and $D_{ij}=d\xi_i\xi_j$. The possible sectors of the λ surfaces may now be determined in the same way as in the two-dimensional case, with the appropriate modifications of terminology: e.g., the sectors are now bounded on the reference sphere by closed lines. The result is that separate sectors are possible only in the directions $\Delta_\xi=\langle 100\rangle$, $\Sigma_\xi=\langle 110\rangle$, $\Lambda_\xi=\langle 111\rangle$ and along a general direction for which $\xi_i=\xi_j$. Here the subscript ξ empha-

sizes that we are dealing with directions in ξ space away from the c.p. in question. The fluted points resulting from the various combinations of these sectors are listed in the next section.

Singular c.p. can also be classified according to their sector numbers. As Van Hove remarks,[6,15] the only singular c.p. for $l=2$ are generalized maxima and minima; thus we need only analyze singular points for $l=3$. Consider first the singular c.p. arising from accidental degeneracies. According to Herring,[16] accidental degeneracies can occur at isolated points in the zone or along closed curves. A simple example is a closed curve of contact C which lies in a symmetry plane. Then the relative maxima and minima along C produce singular c.p., as was first noticed by Van Hove. Herring shows that the components of $\nabla_k\nu^2$ do not vanish at these points except along C. Thus for the lower or upper branch, respectively, the λ surface will be entirely inside or outside the reference sphere except near the two directions where C intersects the reference sphere. Here the λ surface may change sign. In tabular form, the results for the sector numbers are

	Lower branch	Upper branch
Minimum along C	(2, 1)	(1, 0)
Maximum along C	(0, 1)	(1, 2)

Near an isolated point of contact, Herring has also shown that the first-order terms will vanish only in one direction. Thus the sector numbers in the table represent all the possibilities that can occur for both kinds of accidental degeneracy.

More complicated behavior can occur at the singular points necessitated by symmetry. This is because symmetry may often require that $\nabla_k\nu^2$ vanish in a plane. In this case fluted behavior is possible in this plane, so that it is also possible to obtain the combinations (1, 4) or (4, 1) which we call F_1 or F_2 points, respectively.

This completes the discussion of the sector numbers of all c.p. that will occur without special choices of the force constants.

III. TOPOLOGICAL RESULTS

If we analyze the c.p. required by symmetry according to the methods of the last section, we obtain in each branch a set of points which we call the *symmetry set* $\mathcal{S}$. The important idea in Morse's topological work[7] is that the existence of some c.p. necessitates the existence of others. Morse finds that the total numbers of the different kinds of c.p. are related. The symmetry set generally does not satisfy these relations. The smallest set satisfying these relations and containing the symmetry set is called the *minimal set* $\mathfrak{M}$. The chief purpose of this paper is to devise a method for constructing the minimal set.

The topological relations can be stated quite simply for a branch containing only analytic c.p. In two dimensions let p_j denote an analytic c.p. of index j and n_j denote the total number of these in the manifold. Then the following relations hold:

$$\begin{aligned} n_0 &\geqslant 1, \\ n_1-n_0 &\geqslant 1, \\ n_2-n_1+n_0 &= 0. \end{aligned} \qquad (M_2)$$

Similarly, in three dimensions let P_j denote an analytic c.p. of index j and N_j denote the total number of these. Then the relations that must hold are:

$$\begin{aligned} N_0 &\geqslant 1, \\ N_1-N_0 &\geqslant 2, \\ N_2-N_1+N_0 &\geqslant 1, \\ N_3-N_2+N_1-N_0 &= 0. \end{aligned} \qquad (M_3)$$

In both cases the equality sign in the last relation makes it the most useful one for constructing the minimal set. A much weaker set (M') of inequalities only can be obtained from these equations by adding consecutive relations. Neither set can be applied to most branches because many degenerate c.p. will be intrinsically nonanalytic. Van Hove[6] has already applied the relations (M') to branches containing only analytic c.p. and c.p. topologically equivalent to them. In the Appendix we prove by Morse's general methods that the relations (M_l) can be used in the presence of the fluted and singular c.p. that occur in our problem. In doing so, each nonanalytic point may be assigned an index j and a topological weight $q\geqslant 1$.[17] By definition, q is the number of times a c.p. of index j is to be counted in computing the numbers n_j or N_j which enter the relations (M_l). For example, the point in two dimensions with sector numbers (4, 4) has $j=1$ and $q=3$; it is to be counted as three saddle points in the relations (M_2).

We can assign the index j and the number q to the given nonanalytic point once we know the sector numbers (P,N) described in the last section. The topological methods for doing this are presented in the appendix; here we shall only give the results in terms of a few rules.

1. If the sector numbers are (1, 0) or (0, 1), then $q=1$ and $j=0$ or l, respectively. That is, a warped minimum or maximum is topologically equivalent to an analytic minimum or maximum, respectively.

2. In two dimensions, an (n,n) point has $j=1$ and $q=n-1$.

3. In three dimensions, usually only one of P or N will be greater than 1. In the former case $j=2$ and $q=P-1$, and in the latter case $j=1$ and $q=N-1$. In general a point must be counted *both* as a $j=1$ point with $q_1=N-1$ *and* as a $j=2$ point with $q_2=P-1$.

These rules cover all possibilities that can occur for general choices of the force constants.

[17] Van Hove's generalized c.p. are obtained in the special case $q=1$.

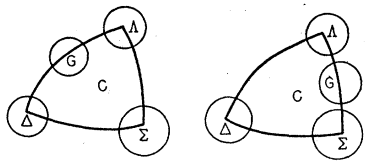

FIG. 3. Sectors around a threefold degenerate point.

A convenient notation is obtained by writing the index j as a subscript to the symbol for the non-analytic point. Thus the (4, 4) point in two dimensions is called an f_1 point. All other c.p. occurring in two dimensions, both fluted and singular, are equivalent to p_j's. Henceforth we use the symbols p_j and P_j for all c.p. topologically equivalent to analytic c.p. of index j. We have completed the task of listing q and j in two dimensions; we now do the same for three dimensions.

We begin with ordinary c.p. where complications will arise only at degeneracies necessitated by symmetry. At a twofold degeneracy lying in a plane of symmetry, the possibilities are determined by the behavior in the plane. If the behavior in the plane is that of a p_j point, only a P point can result. But if we have an f_1 point in the plane, an F_1 or F_2 point of weight 3 will result. Other twofold degeneracies can be handled by our general methods. The possibilities arising from a threefold degeneracy are somewhat more involved. To classify such points, we recall the statement made at the end of the last section to the effect that possible sectors can occur only in the directions Δ_ξ, Σ_ξ, Λ_ξ, and a general direction in the $\xi_i=\xi_j$ plane. The situation for a cubic element $0\leq z\leq y\leq x$ of the reference sphere is illustrated in Fig. 3. The situations arising when three or four of the possible sectors have the same sign and hence coalesce are easy to treat. When the signs along the four special directions indicate the possibility of two positive and two negative sectors, one must distinguish two cases, each of which is illustrated in Fig. 3. In Fig. 3(a) the possible sector in the $\xi_i=\xi_j$ plane, which we label the G_ξ sector, lies between Δ_ξ and Λ_ξ. In Fig. 3(b), on the other hand, the G_ξ sector lies between Λ_ξ and Σ_ξ. The case of two positive and two negative will arise only when alternant sectors in the sequence $(\Delta G\Lambda\Sigma\Delta)_\xi$ or $(\Delta\Lambda G\Sigma\Delta)_\xi$ have opposite sign. Further, the two sectors of the same sign as the central area, C_ξ, must coalesce.

From these arguments all possibilities for ordinary c.p. arising from a threefold degeneracy have been obtained; they are listed in Table I. The left column contains the symbol T_j for the point, the first five columns give the "sign" of the λ surface in the five important directions, the next column gives the sector numbers (P,N), and the last column gives the topological weight q. To use the results of Table I for a particular c.p., one notices that the secular equation factors in the Δ_ξ, Σ_ξ, and Λ_ξ directions, so that the information required by the first three columns is easily obtained. That required in the next two is usually best obtained graphically.

The situation is much simpler for the singular points. If they are necessitated by symmetry, then $q=3$ at most (at the F_1 or F_2 points discussed in Sec. II). If they arise from accidental degeneracies, we know from Sec. II that they have sector numbers (1, 0), (0, 1), (2, 1), or (1, 2), so that $q=1$ in each case. By combining this fact with the result obtained earlier that ordinary points will have $q>1$ only at symmetry points, we obtain the important conclusion that all c.p. not in the symmetry set have $q=1$. This result simplifies the problem of constructing the minimal set, which is what we discuss next.

IV. CONSTRUCTION OF THE MINIMAL SET

Using the rules given in the last section, we can now see if the symmetry set satisfies Morse's relations. If it does, then the symmetry set $\mathfrak{S}$ is the minimal set $\mathfrak{M}(\mathfrak{S}=\mathfrak{M})$; if it does not, then we must search for the additional c.p. in $\mathfrak{M}$. We shall indicate explicitly how this is done in three dimensions. Let us define the multiplicity of a point in the B.Z. as the number of points symmetrically equivalent to it. Then the extra c.p. may be classified in order of decreasing symmetry or increasing multiplicity according to whether they occur in (i) symmetry lines, (ii) symmetry planes, or (iii) general points of the zone. We shall now show how the Morse relations can be used to locate extra c.p. in each category. From the results of the last section, we know that the extra c.p. all have $q=1$, which simplifies our task considerably.

A c.p. will have nonvanishing probability of occurring on a symmetry line only if symmetry requires the normal components of $\nabla_k\nu^2$ to vanish. A c.p. will then occur if ν^2 has a maximum or a minimum along the line. The one-dimensional form (M_1) of Morse's equations is that the number of maxima and minima are equal; further, maxima and minima must alternate alone a line. In this way we obtain all c.p. required on

TABLE I. Possible intrinsically fluted c.p. arising from a threefold degeneracy.

T_j	Δ_ξ	Σ_ξ	Λ_ξ	G_ξ	C_ξ	(P,N)	q
δ_1	−	+	+	+	+	(1, 6)	5
δ_2	+	−	−	−	−	(6, 1)	5
σ_1	+	−	+	+	+	(1, 12)	11
σ_2	−	+	−	−	−	(12, 1)	11
λ_1	+	+	−	+	+	(1, 8)	7
λ_2	−	−	+	−	−	(8, 1)	7
S_1	+	+	+	−	+	(1, 24)	23
S_2	−	−	−	+	−	(24, 1)	23
A_1	+	−	−	+	+	(1, 20)	19
A_2	−	+	+	−	−	(20, 1)	19
B_1	−	+	−	+	+	(1, 14)	13
B_2	+	−	+	−	−	(14, 1)	13
C_1	−	+	+	−	+	(1, 30)	29
C_2	+	−	−	+	−	(30, 1)	29
D_1	+	−	+	−	+	(1, 36)	35
D_2	−	+	−	+	−	(36, 1)	35

symmetry lines by the one-dimensional topological relations (M_1).

The remaining points of $\mathfrak{M}$ which occur in symmetry planes can usually be located by using only the relations (M_2). In so doing, we must be careful to consider enough zones so that the symmetry plane is periodic. For example, in the f.c.c. lattice the plane $k_z=0$ is periodic only if one considers the first two B.Z. The relations (M_2) may require c.p. along the symmetry lines in the plane which were not required by the one-dimensional relations.

Next, the c.p. of $\mathfrak{M}$ occurring at general points in the zone may be located by using the relations (M_3). As before, these relations may require new c.p. in the elements of higher symmetry previously studied.

In both the two- and three-dimensional analyses, the multiplicities of symmetry elements can be used as a guide in locating additional c.p. For example, if one requires four saddle points in a cubic (100) plane, one locates them along the ⟨010⟩ or ⟨011⟩ directions. These in turn must be made consistent with the previous solution of M_1 along the symmetry line. If one location requires further c.p. while the other does not, the latter is chosen.

Finally, we must search explicitly for singular points arising from accidental degeneracies. Some of these can be found by observing that two symmetry points are often connected by a symmetry line and that along this line each solution of the secular equation will belong to a small representation. For each small representation, the frequency must be a smooth function of wave number along this line.[16] From the analysis of the symmetry set, the ordering in frequency of the representations is known near the two symmetry points. If the ordering of two representations near one symmetry point is reversed near the other point, there must be a crossover, or accidental degeneracy, along the symmetry line. In branches defined by the ordered labeling, such crossovers can give rise to singular c.p. This happens if the discontinuities in slope along and normal to the symmetry line occur with a reversal in sign. The same arguments can be used to locate singular c.p. on closed curves of contact lying in symmetry planes. An example of such c.p. is given in the next section.

We have now constructed the minimal set. This set contains the symmetry set, the singular points due to crossovers required by the symmetry set, and the smallest number of additional c.p. consistent with the Morse relations (M_i) for all i. Let us designate by

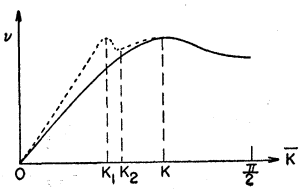

FIG. 4. A schematic dispersion curve illustrating kinks.

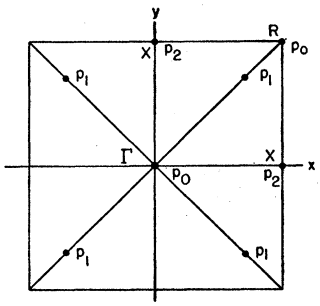

FIG. 5. C.p. in the upper branch of the square lattice, $1/2>\tau>1/5$.

$\mathfrak{M}'(\neq\mathfrak{M})$ the smallest set properly containing $\mathfrak{M}$ and consistent with the requirements of symmetry and connectivity. The extra c.p. in $\mathfrak{M}'$ which do not occur in $\mathfrak{M}$ must have high multiplicity and simultaneously be consistent with all the Morse relations. Therefore we expect $\mathfrak{M}'$ to be considerably larger than $\mathfrak{M}$, which suggests that $\mathfrak{M}$ may often exhaust all the c.p. present in a given branch. We can obtain a stronger criterion for the completeness of the minimal set from the following more detailed considerations of the nature of the extra c.p.

The situation in one dimension is illustrated by the dispersion curve of Fig. 4. The solid curve is a typical dispersion curve for short-range forces. The minima at 0 and $\pi/2$ are obtained as part of the symmetry set and the maximum at K is a consequence of these minima. If there had been a kink in the dispersion curve, as indicated by the dotted portion in the figure, there would have been an additional pair of c.p. at K_1 and K_2. Such kinks can always be added to the minimal set and a set satisfying the relations (M_i) will still be obtained. More explicitly, a kink consists of two c.p. in the same symmetry element of the zone. If the symmetry element is an h-dimensional manifold, the two c.p. must have adjacent indexes j and $j+1$ when considered in h, $h+1$, $\cdots$ dimensions. For example, if the symmetry element is a line, one extra c.p. must be a maximum and the other a minimum along the line. The line will lie in a symmetry plane, and in the plane the pair must have indexes j' and $j'+1$. Finally, for $l=3$ the points must have three-dimensional indexes j'' and $j''+1$. It is now clear that such kinks will be produced only by terms in the secular equation that vary rapidly with k, and these arise only from long-range forces.

The above reasoning breaks down if there are crossovers caused by accidental degeneracies. In effect, these produce kinks, as can be seen in Fig. 6 of the next section. However, we assume that in constructing the minimal set we have located all accidental degeneracies and have included in $\mathfrak{M}$ the c.p. arising from the corresponding kinks.

It is also possible that the true set of c.p., $\mathcal{T}$, does not contain $\mathfrak{M}$, i.e., that some of the c.p. that we have added to $\mathcal{S}$ to form $\mathfrak{M}$ are not in $\mathcal{T}$. This can be checked directly from the secular equation. In such a case $\mathfrak{M}$ can

be modified to obtain a set $\mathfrak{M}_1$ consistent with $\mathcal{T}$. Then the arguments advanced above regarding completeness should still be valid for $\mathfrak{M}_1$. Calculations in such cases may be quite complicated, however, and we restrict ourselves in this paper to examples where $\mathcal{T}\supset\mathfrak{M}$.

We conclude that a consistent minimal set is usually exhaustive for the case of short-range forces only. These plausibility arguments are reinforced by the results of a number of calculations, some examples of which are presented in the next section.

V. EXAMPLES OF THE MINIMAL SET

We now present a few examples to illustrate the method for constructing the minimal set. The simplest one is Montroll's square lattice with central first and second neighbor forces.[4] Group theory shows that both branches must have c.p. at the origin Γ, at the corner R, and at the centers of the sides X. Moreover, Γ and R are twofold degenerate. The symmetry set for the lower branch with Montroll's parameter $\tau<1/2$ consists of a fluted minimum at Γ, p_1's at X, and a fluted maximum at R. Since these c.p. already satisfy the relations (M_2) in this case the minimal set is identical with the symmetry set. In the upper branch there are two cases, depending on whether $\tau>1/5$ or $\tau<1/5$. For $\tau<1/5$ the upper branch has a fluted minimum at Γ, p_2's at X, and an f_1 at R. Again the minimal set is identical with the symmetry set. Finally, for $1/2>\tau>1/5$, the upper branch has fluted minima at Γ and R, and p_2's at X. In this case, which is illustrated in Fig. 5, the minimal set is obtained from the symmetry set by the addition of p_1's along the four diagonals from Γ to R. In all cases the minimal sets include all the c.p. found by Montroll.

We now analyze the more complicated case of the simple cubic lattice. The symmetry points labeled by Γ, R, M, and X[18] must be ordinary c.p. in all three modes; Γ and R are threefold degenerate, M and X are twofold degenerate and lie in planes of symmetry. We use Blackman's[19] secular equation in which τ measures the ratio of second to first neighbor forces. We study the c.p. for $0<\tau<[2(10)^{\frac{1}{2}}-5]/6\simeq 0.221$ and present the results in Table II. The first three lines concern $0<\tau<1/10$. In this case the symmetry sets form minimal sets in all three modes. For $1/10<\tau<1/7$ the λ_1 point in the high mode becomes a P_0 point. From Morse's equations we see immediately that this implies the appearance of a P_1 point along Λ. Similarly as τ increases further one may say that further changes in the symmetry sets produce new c.p. along the symmetry elements of the zone. The nature and location of these c.p. can usually be deduced immediately from the three-dimensional form of Morse's equations. In some cases it is necessary to find points in planes of symmetry, e.g., the plane $k_i=0$ (which we call an s_{I} plane) or $k_i=k_j$ (which we call an s_{II} plane). This is easily done by utilizing the two-dimensional form of Morse's formulas in the way which we discussed earlier. The results agree with those of Rosenstock, who calculated frequency contours up to $\tau=0.221$.

This example illustrates the general principle that forces of longer range produce more c.p. Thus for τ small, and hence essentially only first neighbor forces, the symmetry set contains all the c.p. As τ increases, the second neighbor forces introduce more c.p. The important point is that the minimal set still contains all the additional c.p. introduced by these forces. We may expect that as forces of still longer range become significant, this will no longer be the case. Kinks and consequently c.p. not in the minimal set may then occur. Furthermore in this case for $\tau>[2(10)^{\frac{1}{2}}-5]/6$ the actual set $\mathcal{T}$ of c.p. in the high mode no longer contains the minimal set for that mode.[20] This occurs because the P_1 point along Λ spontaneously becomes a P_3 point at $\tau=[2(10)^{\frac{1}{2}}-5]/6$ without an accompanying change in the symmetry set. At the same time additional p_1 points are introduced into the s_{II} planes. Even with the aid of the Morse relations we now find the situation sufficiently more involved that we restrict ourselves to $\tau<[2(10)^{\frac{1}{2}}-5]/6$.

The final example is f.c.c. aluminum, using Walker's secular equation.[21] Several aspects are now important that were not present in the first two examples. The first is that there is a twofold degeneracy at L which does not lie in a plane of symmetry. The general behavior of frequency surfaces at L in f.c.c. lattices can be taken from the work of Dresselhaus.[14] In general the expressions are complicated, but for one atom per unit cell they simplify to analytic expressions. The next point is that there must be a crossover along Σ. This

TABLE II. Critical points in the simple cubic lattice, using central first and second neighbor forces. The parameter τ is a measure of the ratio of second-neighbor forces to first-neighbor forces.

Branch	Γ	R	X	M	Λ	S	Σ
$0<\tau<1/10$							
High	P_0	λ_1	P_3	F_2			
Middle	P_0	δ_2	P_1	P_3			
Low	P_0	P_3	P_1	P_2			
$1/10<\tau<1/7$							
High	P_0	P_0	P_3	F_2	P_1		
Middle	P_0	δ_2	P_1	P_3			
Low	P_0	P_3	P_1	P_2			
$1/7<\tau<1/5$							
High	P_0	P_0	P_3	F_2	P_1		
Middle	P_0	λ_1	P_1	P_3		P_2	
Low	P_0	P_3	P_1	P_2			
$1/5<\tau<(2\sqrt{10}-5)/6$							
High	P_0	P_0	P_3	P_1	P_1		P_2
Middle	P_0	λ_1	P_1	P_3		P_2	
Low	P_0	P_3	P_1	P_2			

[18] We use the notation of reference 12 for the cubic lattices.
[19] M. Blackman, Repts. Progr. Phys. **8**, 11 (1941).
[20] I am greatly indebted to Dr. Rosenstock for directing my attention to this case.
[21] C. B. Walker, Phys. Rev. **103**, 558 (1956).

can be deduced by noticing that Σ is continued beyond $K=2\pi a^{-1}(1/2, 1/2, 0)$ into S which ends at X. Thus the high (longitudinal) mode along Σ becomes the middle (transverse) mode at X, so that there must be a crossover somewhere along $\Sigma-S$. Actually the crossover must occur along a curve of contact in the s_I plane. The curves for the middle and high modes along Σ are sketched in Fig. 6, and detailed frequency contours showing the curve of contact in the s_I plane are given in Walker's paper. This accidental degeneracy produces a singular maximum in the middle mode and a singular P_1 in the high mode along Σ at the point marked k_1 in the sketch. The second aspect is that symmetry requires that the high and middle modes be degenerate at $W=2\pi a^{-1}(1, 1/2, 0)$ but with $\partial\nu^2/\partial k_y\neq 0$. The middle mode has an f_1 point in the k_xk_z symmetry plane at W; consequently W is a singular F_2 point for that mode. The high mode has a fluted p_0 in the symmetry plane, so that there is a singular P_0 in the high branch at W. The minimal sets for each branch, including the singular points produced by the kink along Σ, are presented in

TABLE III. Critical points in the frequency surfaces of aluminum, using Walker's secular equation.

Branch	Γ	L	X	W	Σ	SI	SII	Σ	Σ
High	P_0	P_3	P_3	P_0		P_1	P_2	P_1	P_2
Middle	P_0	P_1	P_1	F_2	P_3				
Low	P_0	P_1	F_2	P_3					

Table III. Comparison with Walker's results shows that the minimal sets are exhaustive for all branches.

VI. FREQUENCY DISTRIBUTIONS NEAR C.P.

Having located the c.p., we proceed to assess the singular contributions of their neighborhoods in $\mathbf{k}$ space to the frequency distribution $G(\nu^2)$. We denote the lowest order singular contribution of each c.p. by ΔG^+ for $\nu^2>\nu_0^2$ and ΔG^- for $\nu^2<\nu_0^2$. The ΔG's of analytic c.p. have already been evaluated by Van Hove,[6] so that we need consider only the fluted points and the singular points. Both kinds of c.p. generally have such involved frequency surfaces that the resulting integrals cannot be evaluated in closed form. We shall present simple approximations that should be adequate for most purposes.

The first step is to put Eq. (3) in a simpler form. If $\nu^2=\nu_0^2+\phi(\eta_1, \cdots, \eta_l)$ then it is easy to show[22] that the contribution to the frequency distribution of the neighborhood of the c.p. is given by

$$\frac{V_0}{Zl}\int_S \frac{W(\eta_1, \cdots, \eta_l)d\eta_1\cdots d\eta_{l-1}}{|\partial\phi/\partial\eta_l|}. \tag{8}$$

Here $W(\eta_1, \cdots, \eta_l)$ is the appropriate weighting function that makes $W(\eta_1, \cdots, \eta_l)d\eta_1\cdots d\eta_{l-1}=dS_l$,[23] and η_l is determined implicitly by $\phi(\eta_1, \cdots, \eta_l)=\nu^2-\nu_0^2$ on S.

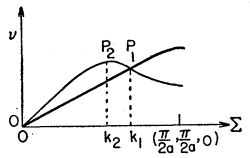

FIG. 6. Sketch of dispersion curves in the high and middle branches along $\Sigma=(110)$ direction in aluminum.

The simplest example to consider is a minimum in two dimensions. If the minimum is analytic, we have ($\epsilon=\nu^2-\nu_0^2$, $\xi=\mathbf{k}-\mathbf{k}_0$)

$$\epsilon=b\xi_x^2+c\xi_y^2=r^2(b\cos^2\theta+c\sin^2\theta). \tag{9}$$

By using the second half of (9) in (8) with $\eta_1=\theta$ and $\eta_2=r$, we find that $\Delta G^-=0$ and $\Delta G^+=\pi V_0/Zl(bc)^{\frac{1}{2}}$ in agreement with Van Hove's results. We observe that a circular minimum with $\epsilon=ar^2$ would give the same result if $a=(bc)^{\frac{1}{2}}$. Thus a is the geometric mean between the extreme values b and c. On the other hand, if one sets $b>c>0$ and $b+c>d$ in Eq. (7), both surfaces have fluted minima at $\mathbf{k}_0$. We may approximate the outer λ line, which has extreme values b and $(b+c+d)/2$, by a circle of radius $a_0=[b(b+c+d)/2]^{\frac{1}{2}}$, and the inner line, of extreme values c and $(b+c-d)/2$, by $a_i=[c(b+c-d)/2]^{\frac{1}{2}}$. The accuracy of this approximation can be checked in two cases. If $d=b-c$, both curves become circles and the approximation is exact. If $d=0$, we have two noninteracting analytic minima which can be treated exactly, or viewed as a limiting case of upper and lower fluted minima. Then the fractional error in the total jump for both surfaces computed from the approximations a_0 and a_i turns out to be $1-(2bc)^{\frac{1}{2}}(b^{\frac{1}{2}}+c^{\frac{1}{2}})(b+c)^{-\frac{3}{2}}$, which is about 0.08 for $b=2c$. Mean radii may also be used to compute approximate jump decreases in the frequency distribution arising from maxima.

The only remaining fluted point for $l=2$ is the f_1 point which can be obtained by choosing $-c>b>0$ and $b+c+d<0$ in Eq. (7). The frequency contours of the higher surface, which is an f_1 point, are shown in Fig. 7. We shall calculate ΔG^+ as an example of the approximations that can be employed. Because of symmetry, we need consider only the upper halves of the hyperbolas $\nu^2>\nu_0^2$ in the first octant. We see, from solving the secular equation, that these curves are given by

$$\epsilon=\frac{b+c}{2}(\xi_1^2+\xi_2^2)+\left[\left(\frac{b-c}{2}\right)^2(\xi_1^2-\xi_2^2)^2+(d\xi_1\xi_2)^2\right]^{\frac{1}{2}}, \tag{10}$$

and G is given by

$$G(\nu^2)=\frac{8\pi V_0}{2Z}\int_{(\epsilon/b)^{1/2}}^{L}\frac{d\xi_1}{|\partial\epsilon/\partial\xi_2|}. \tag{11}$$

[22] R. Courant, *Differential and Integral Calculus II* (Interscience Publishers, Inc., New York, 1936).

[23] H. Margenau and G. Murphy, *The Mathematics of Physics and Chemistry* (D. Van Nostrand Company, Inc., New York, 1943), p. 190.

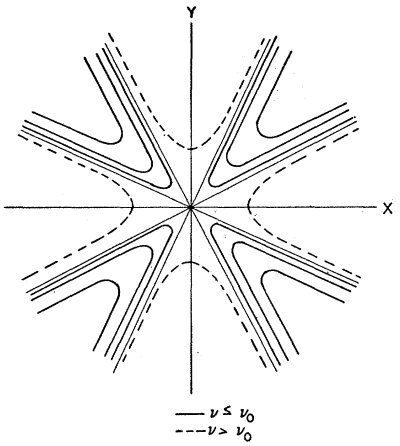

FIG. 7. Frequency contours of an f_1 point.

The singularity in G is produced by the tips of the hyperbolas so that the upper limit L need not be specified. Near the tips $\xi_2 \ll \xi_1$ so that, to second order in ξ_2, Eq. (10) becomes

$$\epsilon \simeq b\xi_1^2 + \left(c + \frac{d^2}{b-c}\right)\xi_2^2. \tag{12}$$

We observe that (12), which is exact in the terms that produce the chief singularity in $G(\nu^2)$, could have been obtained more simply from second-order perturbation theory. This remark will be of use when we come to the three-dimensional fluted c.p. For the f_1 point we have

$$\Delta G^+ = -\frac{\pi V_0}{Z a^+} \ln\left|\frac{\nu^2}{\nu_0^2} - 1\right|, \tag{13}$$

with $a^+ = \{-b[c+d^2/(b-c)]\}^{\frac{1}{2}}$. By performing a unitary transformation to the diagonal axes $\xi_1 = \pm\xi_2$, we can derive a similar formula for ΔG^-, with $a^- = \frac{1}{2}(b+c+d)^{\frac{1}{2}} \times [b+c-d+2(b-c)^2/d]^{\frac{1}{2}}$.

At a singular minimum we will have $\epsilon = f(\theta) r$. Then $\Delta G^- = 0$ and

$$\Delta G^+ = \frac{\pi V_0}{2Z} \int_0^{2\pi} \frac{r\,d\theta}{f(\theta)} = \epsilon \left\{ \frac{\pi V_0}{2Z} \int_0^{2\pi} \frac{d\theta}{f^2(\theta)} \right\}, \tag{14}$$

so that a singular minimum produces a discontinuous increase in $G'(\nu^2)$.

In three dimensions there are more possibilities to consider. We begin by listing in Table IV the singular terms that each kind of c.p. may be expected to contribute to the distribution $G(\nu^2)$. We note that the saddle points have a behavior that is qualitatively different from that of the maxima and minima. From the table, we see that only ordinary points and singular maxima and minima having one discontinuous component of $\nabla_k \nu^2$ can produce discontinuities in $G'(\nu^2)$. Thus we need analyze in detail only these points. Before doing so, we outline the arguments used to obtain the results of Table IV. The first and fourth lines are obtained immediately if one assumes $\epsilon \sim r^2$ or r, respectively. The second line follows if one observes that the reference sphere can be divided into two zones. Designate by the equator the intersection of the plane in which $\nabla_k \nu^2 = 0$ with the reference sphere. Then the first zone, in which $\epsilon \sim r^2$, is an equatorial band of width $\sim \epsilon^{\frac{1}{2}}$ while the second zone is the rest of the sphere, for which $\epsilon \sim r$. The second zone contributes terms $\sim \epsilon^2$, which may be neglected in comparison with the contribution of the first zone, which is proportional to $\epsilon^{\frac{1}{2}} \times \epsilon^{\frac{1}{2}} = \epsilon$. A similar argument gives the result of the third line, since the zone where $\epsilon \sim r^2$ then occupies a solid angle $\sim \epsilon$, so that its contribution is $\sim \epsilon^{\frac{1}{2}} \times \epsilon = \epsilon^{\frac{3}{2}}$. The results of the last four lines are obtained by direct calculation.

Let us now consider ordinary c.p. in detail. Van Hove[6] has already indicated that fluted points as well as analytic points will contribute singular terms $\sim \epsilon^{\frac{1}{2}}$, and has calculated the contributions of analytic points explicitly. The contributions of the remaining points can be approximated by the techniques used for $l=2$. Thus, we approximate warped spheres of fluted maxima and minima by spheres of mean radius A. If the minimum or maximum is twofold degenerate, the frequency contours are given by

$$\epsilon = a f(\theta) \rho^2 + g z^2. \tag{15}$$

Here z is the coordinate normal to the symmetry plane and a is the mean radius in the symmetry plane obtained by the methods discussed earlier. The radius of the equivalent sphere is given by $A^3 = a^2 g$. On the other hand, if the minimum or maximum arises from a threefold degeneracy, so that it has cubic symmetry, we can take $A^3 = \alpha\beta\gamma$. Here α, β, and γ are the radii in the (100), (110), and (111) directions, respectively, easily obtained from the secular equation for these directions. The validity of this approximation has been checked in two ways for the special case $b=2c$ and $d=0$ in the secular equation. The first way was to fit the values α, β, γ to a combination of cubic harmonics, and then do the integral numerically, using elliptic integrals. The result agreed to within the limits of error with

TABLE IV. Singularities contributed to $G(\nu^2)$ by various c.p.

Designation of c.p.	No. of discontinuous components of $\nabla_k \nu^2$	ΔG
Maxima and minima		
Ordinary	0	$\epsilon^{\frac{1}{2}}$
Singular (symmetry)	1	ϵ
Singular (curve of contact)	2	$\epsilon^{\frac{3}{2}}$
Singular (isolated contact)	3	ϵ^2
Saddle points		
Ordinary	0	$\epsilon^{\frac{1}{2}}$
Singular, $\nabla_k \nu^2 \neq 0$ along axes of hyperboloids	1, 2, 3	0
Singular	1 ($\perp$ axes of hyperboloids)	$\epsilon^{\frac{1}{2}} \ln\epsilon$
Singular	2 ($\perp$ axes of hyperboloids)	$\epsilon^{\frac{1}{2}}$

that given by the approximate radius A. The second way used the fact that for $d=0$ we have three interpenetrating ellipsoids, so that the total ΔG for all three surfaces can be evaluated exactly. The fractional error of the approximate formula in this case was 0.10.

Fluted saddle points can be treated by perturbation techniques similar to those used for the f_1 point. Consider for example a δ_1 point, whose frequency surfaces for $\nu^2 \simeq \nu_0^2$ are shown in Fig. 8. The contribution of the cup-like surfaces for $\nu^2 < \nu_0^2$, which are shown in Fig. 8(a), can be evaluated by observing that the singular contribution again comes from the tip of the cup. Because of symmetry, we need only consider the cup along the positive ξ_1 axis. Then for the region of interest, $\xi_2^2+\xi_3^2 \ll \xi_1^2$, so from second order perturbation theory

$$\epsilon \simeq b\xi_1^2+\left(c+\frac{d^2}{b-c}\right)(\xi_2^2+\xi_3^2), \tag{16}$$

where $c>-b>0$ assures that we have a δ_1 point. As before, the perturbation expansion is exact in the term that produces the leading singularity in $G(\nu^2)$. In this case,

$$\Delta G^-(\nu^2)=-\frac{6\pi V_0}{Zla^{\frac{3}{2}}}(\nu_0^2-\nu^2)^{\frac{1}{2}}; \quad a^3=-b\left(c+\frac{d^2}{b-c}\right)^2, \tag{17}$$

where the factor 6 occurs instead of Van Hove's 2 for an analytic point, as there are now 6 cups instead of 2. For $\nu^2 > \nu_0^2$, these perturbation methods no longer apply. Although we have not been able to calculate the result explicitly, we believe that the connected, tipless surfaces for $\nu^2 > \nu_0^2$ should make similar contributions whether they are analytic hyperboloids of one sheet or the fluted surfaces of Fig. 8(b).[24] Since the connected analytic surfaces do not contribute a singularity to $G(\nu^2)$, we shall *assume* that the same result holds for the connected surfaces of the δ_1 point for $\nu^2 > \nu_0^2$. All other fluted saddle points may be treated similarly, so that this completes the analysis of the ΔG's of all ordinary c.p.

We consider now singular maxima and minima for which $\nabla_k \nu^2 = 0$ in a plane of symmetry, while there is a crossover normal to the plane. Suppose, e.g., the point is a fluted minimum in the plane and a singular minimum along the normal to the plane, which we take to be the ξ_3 axis. Then if the mean radius in the plane is a, we may approximate the frequency surfaces by two expressions (which are needed because of the crossover; a, g, $h>0$):

$$\begin{aligned} \epsilon &\simeq a\rho^2 f(\theta)+g\xi_3, \quad \xi_3>0 \\ &\simeq a\rho^2 f(\theta)-h\xi_3, \quad \xi_3<0. \end{aligned} \tag{18}$$

We still have $\Delta G^-=0$, and a direct calculation gives

$$\Delta G^+=\frac{\pi V_0}{Zla}\left(\frac{1}{g}+\frac{1}{h}\right)\epsilon. \tag{19}$$

This completes our quantitative analysis of the major singularities in $G(\nu^2)$ produced by nonanalytic c.p.

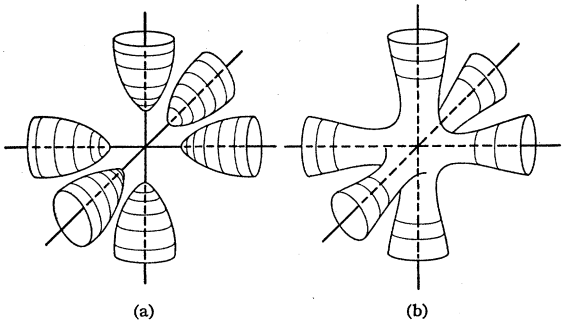

FIG. 8. Frequency surfaces near a δ_1 point. (a) $\nu^2 \lesssim \nu_0^2$; (b) $\nu^2 \gtrsim \nu_0^2$.

[24] Van Hove has arrived at a similar conclusion regarding the similarity of the ΔG's arising from analytic and fluted c.p.

VII. APPROXIMATE CALCULATIONS OF THE FREQUENCY DISTRIBUTION

It can be shown from Eq. (3) that only c.p. contribute singularities to the frequency distribution. Thus it may be hoped that a knowledge of these singularities, together with appropriate smooth curves connecting them, will be enough to give a good approximation to the exact distribution. The moment-singularity method developed by Lax and Lebowitz[8] is a refinement of this idea. They proposed that the strengths of the singularities, together with the coefficients of a first few Legendre polynomials, be left as disposable parameters. The parameters are then determined from a corresponding number of moments of the distribution. By using a method due to Montroll, one can calculate the latter without solving the secular equation.

The moment-singularity method was applied specifically to two-dimensional lattices which have infinite singularities in their distributions. We shall develop a somewhat different procedure for sketching $G(\nu^2)$ for a three-dimensional lattice. Our expressions for the frequency distribution near each singularity give us the shape of the distribution there but not its magnitude. It is laborious to calculate the higher moments of the distribution; therefore we reduce the number of moments to be calculated by computing the shape of the distribution near each important c.p. by the methods of the last section. Then we will obtain a good approximation to the total distribution if we can estimate the heights of the chief peaks. The simplest procedure is to leave the height of the main peak in each branch undetermined, and to add enough linear and quadratic terms to the expansions about each singularity to match smoothly the distributions at intermediate points. The constant in each branch is then determined by normalization. The natural generalization of this scheme, in the spirit of the moment-singularity method, is to leave the magnitudes of n_j points in the jth branch to

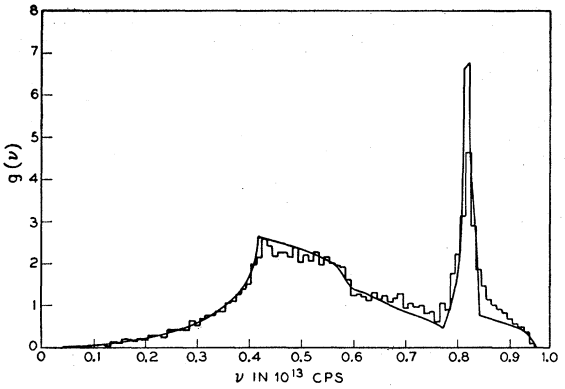

FIG. 9. The frequency distribution of aluminum, using Walker's secular equation. The histogram was obtained by Walker from a sampling of 2791 points in 1/48th of the zone. The smooth curve was obtained by the interpolation scheme described in the text.

be determined. The normalization conditions then leave only n_j-1 undetermined parameters in the jth branch. These can be obtained by computing the first $\sum_{j=1}^{l}(n_j-1)$ moments of the distribution and solving the resulting simultaneous equations.

We have employed this method to calculate the frequency distribution of aluminum from Walker's[20] secular equation. Only the simplest form of the method, with the heights of the main peaks determined by normalization, was used. The result for the total distribution is shown in Fig. 9, together with the histogram obtained by Walker from the solution of the secular equation at 2791 points in 1/48th of the zone. Comparison with Walker's separate histograms for each branch showed that the heights of the main peaks of the lower and middle branches were obtained with about a 10% error. This was not surprising, since both branches had broad maxima, so that the normalization condition could be expected to determine the height with some degree of accuracy. Moreover, the linear and quadratic terms which were added to these branches to produce smooth curves were small compared to the terms due to the singularities. On the other hand, there is a very narrow peak in the high branch. The height of this peak will be hard to determine by any moment method since the moments are not sensitive to a shift of part of the distribution to the wings of the peak. In our case, for example, our estimate of the height of the narrow peak is about 50% too large. It might be thought that this error could be greatly reduced by including fourth order terms about the maximum in the high mode, thus increasing the magnitude of the right wing of the peak. The improvement is slight, however, owing to the slow convergence of the higher terms, which in turn is a consequence of the distortions producing the narrow peak. Thus, in such cases if one wants a more exact distribution, one must simply calculate the frequency surfaces for the critical frequencies on each edge of the peak and estimate the volume of k space between the surfaces. Since only two surfaces are required, the additional labor should not be great.

One may conclude from this calculation that if only broad peaks are present in a distribution, the distribution can be sketched with about 10% accuracy by a simple interpolation between singularities.

VIII. CONCLUDING REMARKS

We now summarize the results of the paper. We have shown that by analyzing the matrix elements of $\nabla_k\nu^2$ by group theory, one can find all c.p. that must occur because of symmetry. The degeneracies required by symmetry at these c.p. are obtained at the same time. Possible accidental degeneracies can also be inferred from the analysis of symmetry points. Next we gave a detailed analysis of possible frequency surfaces near degeneracies and showed how the sector numbers could be used to classify topologically the fluted and singular c.p. which can arise. For each branch all c.p. necessitated by symmetry can be analyzed in this way, and we have called the resulting set of c.p. the symmetry set $\mathcal{S}$ belonging to the branch in question. We next showed that by using the Morse relations (M) one can extend $\mathcal{S}$ to a larger set $\mathfrak{M}$, which we have called the minimal set. $\mathfrak{M}$ is the smallest set of c.p. which is completely consistent with connectivity requirements as well as symmetry requirements, and includes the singular points arising from the cross-overs obtained earlier. Arguments were advanced for the completeness of $\mathfrak{M}$ under certain circumstances. Several examples were given in which the $\mathfrak{M}$'s belonging to particular secular equations were constructed. Most of the minimal sets so constructed agreed with those derived from detailed frequency contours.

We next showed how to estimate the singular contributions to the frequency distribution $G(\nu^2)$ of the neighborhoods of fluted and singular c.p. The frequency distribution of each branch could then be approximated by a smooth curve having the appropriate singularities. A sample calculation showed that the method could be expected to give good results for real three-dimensional lattices.

We may conclude that it is easy to construct the minimal set, that its construction yields information about frequency surfaces throughout the zone from a detailed analysis about symmetry points alone, and that the minimal set forms a suitable basis for approximation to the frequency distribution.

It should be noted that the methods developed in this paper are quite general and can be applied to any system of functions derived as solutions of a secular equation defined on a Brillouin zone. In particular, with minor modifications, the method for constructing the minimal set can be carried over immediately to the energy bands of the Bloch theory of electrons in crystal lattices. The minimal set should be particularly useful in the calculation of complicated energy bands, since it can

be constructed once the secular equation is known only in the neighborhood of each symmetry point.

The modifications that would be required for energy bands depend mainly on the fact that each band need no longer belong to a vector representation. One consequence of this is that fourfold degeneracies can occur. These can be analyzed by the same methods that we have applied to twofold and threefold degeneracies.

We have seen that the minimal set can be constructed from the symmetry set. The latter in turn is obtained once one knows both the order of energy values at the points of the symmetry set and the curvatures in each branch at these points. These values can be obtained from effective mass data and from detailed calculations such as those based on orthogonalized plane waves. Each minimal set will determine the qualitative shape of the energy surfaces for that particular branch, its extrema, and the qualitative shape of the energy spectrum of the band. Moreover, from our calculations on the simple cubic lattice, it is clear that a whole range of the parameters may correspond to the same minimal set. Thus in some cases one need only estimate the parameters, or only determine the signs of the curvatures at symmetry points, to obtain the extrema of the bands.

These arguments based on the minimal set should be compared with the interpolation scheme for constructing energy bands proposed by Slater and Koster.[25] They suggested that a tight-binding approximation would be suitable in many cases, but that instead of computing the required matrix elements, one should regard them as disposable parameters, to be determined from more exact values at the symmetry points. One would introduce as many parameters as could be determined from the available data, and decide which parameters were significant by physical arguments. The method of Slater and Koster establishes a secular equation with fewer data than are required to specify the actual symmetry set and hence the minimal set. Once the secular equation is established, however, its energy contours, extrema, and energy distribution can be studied by our methods as well as by the sampling techniques used by Slater and Koster.

If there are enough data available, the symmetry set can be restricted to a few possibilities. The corresponding minimal sets can be constructed and the common features analyzed. A given interpolation scheme will correspond to one of these sets. Thus one can see how representative this scheme is of the various possibilities consistent with the data.

Finally, we should discuss the completeness of minimal sets for energy bands. If one makes a tight-binding approximation and assumes that only nearer neighbors contribute significantly to the matrix elements, the arguments made earlier in connection with short-range forces should suffice. Additional complications could arise from a large number of bands closely spaced in energy. This could produce many kinks, and the singular points arising from accidental degeneracies might be hard to classify. On the other hand, if the bands are well separated, one can use the f-sum rule to argue that the large curvatures necessary for kinks will probably not occur.[26] More restrictive statements might be obtained as a result of a detailed application of these methods to various energy bands.

The author would like to express his gratitude to Professor M. H. Cohen for suggesting this problem and for his continued suggestions and encouragement throughout the solution of the problem and during the preparation of the manuscript. The author also expresses his appreciation to Professor E. H. Spanier for checking the topological results. He is also indebted to Professor C. B. Walker of the Institute for the Study of Metals, Eugene Blount, Professor G. F. Dresselhaus, and Marshall Osborn for various conversations. This research was performed while the author held a General Electric Fellowship.

APPENDIX

It was stated in Sec. III that the Morse relations, when properly interpreted, are valid in the presence of fluted and singular c.p. We now derive this result by the methods of Morse.[7] A detailed description of these methods can be found in the book by Seifert and Threlfall.[27] We use an outline of their presentation as an introduction to our discussion.

We begin by defining some of the elementary notions of algebraic topology. Let Ω be a topological domain, i.e., a set of points with a metric such that every point has a well-defined neighborhood. Further, let Ω be closed, that is, have no boundaries. Because of the periodic boundary conditions, our Brillouin zones are such domains. They are topologically equivalent to l-dimensional toroids. For example, a square with opposite sides identified may be deformed into a doughnut.

We now introduce the concept of the topological connectivity of a manifold. Define a k chain $\mathfrak{X}^k$ as a k-dimensional subset of Ω. A closed k chain is called an absolute k cycle. If an absolute k cycle $\mathfrak{Z}^k$ bounds a $(k+1)$ chain $\mathfrak{X}^{k+1}$ of Ω, $\mathfrak{Z}^k$ is said to be homologous to zero, $\mathfrak{Z}^k \sim 0$. Two chains $\mathfrak{X}^k$ and $\mathfrak{Y}^k$ are homologous in case their sum is homologous to zero (if one chain is deformable into another, it is homologous to it). The r absolute cycles $\mathfrak{Z}_1^k, \cdots, \mathfrak{Z}_r^k$ are homologously independent in case, with $\epsilon_\rho = 0$ or 1, $\sum_{\rho=1}^{r} \epsilon_\rho \mathfrak{Z}_\rho^k \sim 0$ implies $\epsilon_\rho = 0$. The kth connectivity number R^k of Ω is defined as the maximal number of homologously independent absolute k cycles of Ω. For a two-dimensional torus, the

[25] J. C. Slater and G. F. Koster, Phys. Rev. **94**, 1948 (1954).

[26] I am indebted to Dr. C. Herring for this observation.

[27] H. Seifert and W. Threlfall, *Variationsrechnung Im Grossen*, Hamburger Mathematische Einzelschriften (B. G. Teubner, Leipzig und Berlin, 1938), Vol. 24. I am indebted to Professor E. H. Spanier for bringing this book to my attention.

nonzero connectivity numbers are

$$R^0=1, \quad R^1=2, \quad R^2=1,$$

while for a three-dimensional torus the nonzero ones are

$$R^0=1, \quad R^1=3, \quad R^2=3, \quad R^3=1.$$

Morse succeeded in constructing similar connectivity characterizations for neighborhoods of separate c.p. To do so, one must consider parts of cycles, and in considering each neighborhood "discard" the rest of the manifold. This can be done in the following way. If $Z\subseteq H\subseteq\Omega$, we say that a chain contained in H is homologous to zero "mod Z" if it lies entirely in Z. For the chains $\mathfrak{x}^k$ and $\mathfrak{y}^k$ $\mathfrak{x}^k\sim\mathfrak{y}^k$ mod Z is equivalent to $\mathfrak{x}^k+\mathfrak{y}^k\sim 0$ mod Z, since addition of chains is taken mod 2. A chain $\mathfrak{z}^k$ of H is called closed mod Z or a relative cycle mod Z if its boundary is ~ 0 mod Z. Thus, a relative k cycle on H mod Z is a k chain lying in H with its boundary in Z. We can now define linear independence of relative cycles. It is understood that a relative cycle $\mathfrak{z}^k$ mod Z is homologous to zero mod Z when it is homologous to a chain $\mathfrak{x}^{k+1}$ of Z. If $\mathfrak{z}_1{}^k, \cdots, \mathfrak{z}_s{}^k$ are s relative k cycles, they are homologously independent mod Z in case ($\epsilon_\rho=1$ or 0) $\sum_{\rho=1}{}^s \epsilon_\rho\mathfrak{z}_\rho{}^k\sim 0$ implies $\epsilon_1=\cdots=\epsilon_s=0$.

Now suppose J is a continuous function on Ω. If α is an arbitrary value of J, we denote the set of all points of Ω for which $J<\alpha$ by $\{J<\alpha\}$. Let g be a point of Ω and $J(g)=\alpha$. We shall call all the relative k cycles in the set $\{J<\alpha\}+g$ mod$\{J<\alpha\}$ the relative k cycles, or just k cycles, belonging to the point g. The "mod$\{J<\alpha\}$" device used here is the way that we succeed in "throwing away" all but a neighborhood of g in determining the cycles belonging to g. Now we define the maximal number $m^k=m^k(g)$ of homologously independent k cycles belonging to g to be the kth-type number of the point g. The type numbers of most points of Ω are all zero. More generally, the type numbers characterize the topological features of each point. In fact, we now define a c.p. as a point not all of whose type numbers are zero. We show that this definition contains the definition given in Sec. I in terms of components of the gradient. In our manifolds, ∇J is continuous at all except crossover points. Consider a point g where one component of the gradient is continuous and not zero. Then the relative cycles belonging to g end at g. However, they can all be deformed to below the level $\alpha=J(g)$ by slipping them down along the slope of J through g. Thus they are all homologous to zero, and all the type numbers of g are zero. Thus g will be a c.p. only if all components of the gradient which are continuous at g also vanish there. On the other hand, if a component is discontinuous, it must change sign at g, or else the same deformation could still be effected. Hence g can be a c.p. only if one or the other condition holds for all three components of the gradient. Thus the analytical definition of c.p. given in Sec. I is indeed contained in our topological definition.

As an example, we shall determine the type numbers of the saddle point of the function $J=x^2-y^2$ occurring at the origin. Clearly $m^0=0$, since a point on $\{J<0\}+(0,0)$ is deformable into a point on $\{J<0\}$. Also $m^2=0$, since if an area on $\{J<0\}+(0,0)$ contains the origin, the origin is a part of its boundary, and so it is not a relative two-cycle mod$\{J<0\}$; while if it does not contain (0,0), then it lies entirely in $\{J<0\}$, and so is homologous to zero mod$\{J<0\}$. A relative 1 cycle which is not deformable to $\{J<0\}$ must pass through the origin, with its end points (boundaries) on opposite sides of the origin (see Fig. 10). All such 1 cycles are deformable into one another, so that there is exactly one homologously independent relative 1 cycle and $m^1=1$.

We may also consider the n-fold fluted point[28] (the ordinary saddle point of Fig. 10 is the special case

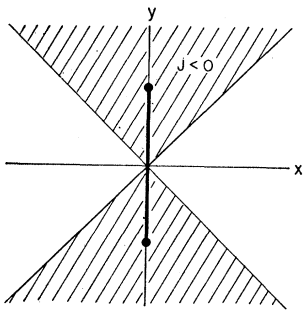

FIG. 10. An analytic saddle point in two dimensions.

$n=2$). This contains n sectors where $J<0$. Here as before $m^0=m^2=0$, but now $m^1=n-1$. Each nonzero one cycle must have its end points lying in two different sectors. If we label a cycle by sectors in which its end points are located, then every nonzero cycle may be expressed as a linear combination of the n cycles (1,2), (2,3), $\cdots$, (n,1). However, the sum of these cycles is zero, since they bound the two chain which is the neighborhood of the origin. Thus one cycle must be removed, so that $m^1=n-1$. A similar argument for the multiple three-dimensional saddle points shows that if the sector numbers are (r,1) or (1,s), the nonzero type numbers are $m^2=r-1$ or $m^1=s-1$, respectively. For example, if the sector numbers are (1,s), the only nontrivial cycles are the one cycles ending in separate sectors, and only $s-1$ of these are homologously independent, or nonbounding. If the sector numbers are (r,1), the result is obtained by applying the same argument to the function $-J$. A more detailed argument can be given to justify the more general result for the (r,s) points quoted in rule 3 of Sec. III.

Consider now a nonzero relative k cycle $\mathfrak{z}^k$ associated with a critical point g; again let $\alpha=J(g)$, and denote the boundaries of $\mathfrak{z}^k$, which lie in $\{J<\alpha\}$, by $R\partial\mathfrak{z}^k$. If $R\partial\mathfrak{z}^k$ is homologous to zero on $\{J<\alpha\}$, it bounds a relative k cycle $\mathfrak{u}^k$ lying entirely below α. Thus $\mathfrak{z}^k+\mathfrak{u}^k=\mathfrak{Z}^k$ is an

[28] See p. 39 of reference 7.

absolute k cycle, and we say that the relative cycle $\mathfrak{z}^k$ is completable below α to an absolute cycle.

We are now ready to state, and outline the proof of, the form of Morse's formulas appropriate to our problem. Let J have only a finite number of critical points on Ω, and let the sum of the type numbers of each critical point be finite. If we denote the sum of the kth type numbers of all critical points by N^k, then the following equations hold between the N^k and the connectivity numbers R^k of Ω:

$$\begin{aligned} &N^0 \geqslant R^0, \\ &N^1-N^0 \geqslant R^1-R^0, \\ &\cdots \\ &\cdots \\ &\cdots \\ &\sum_{i=1}^{l}(-1)^i N^i=\sum_{i=1}^{l}(-1)^i R^i. \end{aligned} \qquad (M_l)$$

Note that in the last relation it is the equality that holds. For now let β be a parameter that increases continually from the minimum of J to its maximum. The connectivity numbers of $\{J\leq\beta\}$ will change only when β is a critical value α. Let $J(g)=\alpha$ and $\mathfrak{z}^k$ be a relative cycle associated with g. If $\mathfrak{z}^k$ is completable below α, the addition of g to $\{J<\alpha\}$ will increase the kth connectivity number of $\{J<\alpha\}$ by one, so that we may call a completable relative k cycle an "increasing" one. If $\mathfrak{z}^k$ is not completable, the absolute $(k-1)$ cycle $R\partial\mathfrak{z}^k$, which is not homologous to zero on $\{J<\alpha\}$, becomes so by the addition of g. Thus the $(k-1)$th connectivity number of $\{J<\alpha\}$ is decreased by one by the addition of g, and $\mathfrak{z}^k$ may be said to be a "decreasing" cycle. If we denote the total number of independent "increasing" relative k cycles by N_+^k, and the total of the decreasing ones similarly by N_-^k, we have, when β has attained its maximum,

$$R^k=N_+^k-N_-^{k+1}, \qquad (20)$$

but also

$$N^k=N_+^k+N_-^k \qquad (21)$$

and $N_-^0=N_-^{l+1}=0$, since there are no decreasing 0 cycles, and there are no $(l+1)$ cycles at all. Subtracting (20) from (21) gives

$$N^k-R^k=N_-^k+N_-^{k+1}, \qquad (22)$$

$$N^0-R^0=N_-^0+N_-^1=N_-^1\geqslant 0. \qquad (23)$$

The rest of Morse's formulas are derived in a similar way; the equality holds in the last equation because $N_-^{l+1}=0$.

The equations (M_l) which we have just stated were obtained by Morse for any manifold for which all sums of type numbers N^k were finite. For our manifolds we have shown that this is the case. In addition, we have shown that the type numbers m^k of each c.p. in our manifolds usually have the form $m^k=q\delta_{jk}$, so that is is sufficient to know q and j for such c.p. Moreover, in the Morse relations (M_l) the sum of the type numbers N^k can be replaced by the total number of c.p. of index k, N_k, where the αth c.p. is to be counted $q_{\alpha k}$ times in computing N_k. Then, using the connectivity numbers stated earlier, we obtain the form of the Morse equations stated in Sec. III. Thus, to make a topological analysis of our manifolds it was only necessary to evaluate the type numbers of the critical points that arise.