THE AMERICAN MINERALOGIST

JOURNAL OF THE MINERALOGICAL SOCIETY OF AMERICA

Vol. 53 MARCH–APRIL, 1968 Nos. 3 and 4

ON LOCATING FIELD BOUNDARIES IN SIMPLE PHASE DIAGRAMS BY MEANS OF DISCRIMINANT FUNCTIONS

FELIX CHAYES, *Geophysical Laboratory, Carnegie Institution of Washington, Washington, D. C. 20008.*

ABSTRACT

The traces of discriminant functions computed from the data are in excellent agreement with field boundaries shown in a number of published phase diagrams. The substitution of a systematic numerical procedure for the conventional inspection technique by which field boundaries are located might reduce the amount of time and effort now sometimes devoted to obtaining points very close to the supposed trace of the field boundary. The principal advantage of the numerical procedure, however, is that it may be applied in multicomponent space, in which the location of field boundaries by graphical inspection is impossible.

INTRODUCTION

Phase-equilibrium studies were among the first systematic experimental researches in petrology, and in most work of this type the actual data gathering is characterized by strong emphasis on quantitative control. In the course of the past half-century there have been vast improvements in the techniques by which temperatures are controlled and recorded, X-ray diffraction has made possible the study of products too fine for effective identification by microscope, and devices permitting generation and measurement of pressure have developed to such an extent that in many laboratories P-T conditions comparable to those at the base of the crust or well into the upper mantle can be reached as a matter of routine. There has been no comparable improvement in the graphical technique by which the experimenter blocks out, from examination of the assemblage of data, the limits of a stability field or the boundary between two such fields. In this crucial operation, in fact, there seems to have been no change at all.

There are many reasons—good and bad—for this conservatism in an otherwise rapidly developing discipline. Probably the most important is just that the location of field boundaries by graphical inspection is simple and usually adequate. Further, although the sentiment may seem heretical to many practitioners of such an eminently quantitative art,

the delineation of field boundaries is essentially qualitative. This does not mean that it is not susceptible to numerical analysis, but it does mean that the curve-fitting techniques familiar to everyone, techniques based on regression statistics of one kind or other, are essentially inapplicable or irrelevant. Whatever its final interpretation, in the actual experimental situation the field boundary is not a line of central tendency. Points lying on opposite sides of a theoretically correct or "true" field boundary differ from each other in an important qualitative sense which finds no expression in regression analysis. In fact, regression analysis presumes that there is *no* qualitative distinction between deviations of opposite sign whereas the existence of just such a distinction is fundamental to the definition of the field boundary.[1]

Numerical construction—more precisely, reconstruction—and analysis of qualitative classifications are two of the objectives of discriminant function analysis, and although the procedure is usually applied to multivariate arrays so that graphical representation is impossible, the principles of the method are unaffected by the number of variables. In all examples so far examined, the traces of discriminant functions computed from the data are in good agreement with boundaries shown graphically in the original publications, and usually located by inspection.

Definition and Calculation of the Discriminant Function

It is presumed that each of the n items in a sample may be unequivocally assigned to one of two mutually exclusive groups by means of some initial set of properties, and that for the ath item, $a = 1, 2, \cdots, n$, observed values of each of a further set of m properties, the variables $(X_{aj}), j = 1, 2, \cdots, m$, are available. In our case the initial set of properties is the phase assemblage, found at the conclusion of the run, which determines whether an item is a member of class 1 or class 2, whereas the second set consists of the temperature and pressure of the run and the composition of the charge. (If composition is the same for all runs it is of course not a variable, and similarly for pressure.)

For each item vector, $\boldsymbol{X}_a = [X_{a1}, X_{a2}, X_{a3}, \cdots, X_{am}]$, we find the scalar, $Z_a = \lambda' \boldsymbol{X}_a$. From the assemblage of Z's we next compute $\bar{z}_1$, $\bar{z}_2$, the mean values of Z for each group, and could then evaluate the ratio

$$G = \frac{(\bar{z}_1 - \bar{z}_2)^2}{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}, \tag{1}$$

[1] For an attempt to escape this dilemma by confining the regression analysis to mean values for pairs of "bracketing points," see Boyd, England, and Davis (1964, pp. 2105–2106).

where n_i is the number of items, and s_i^1 is the observed variance of Z, in group i. In the absence of special definition of Z, the quantity G would be simply the F ratio, of "between"- to "within"-group sums of squares, familiar from ordinary variance analysis. In variance analysis we accept and test whatever value of F emerges either from some one of the raw variables or from a linear transformation of some set of them; for example, the distribution of silica in two groups of rock analyses or the distribution of Q in norms calculated from them.

In discriminant function analysis, however, we purposely choose the linear transformation that *maximizes* F. The first problem is to find, for any set of data, the λ that will have this effect, and in this connection it may be shown that $\partial G/\partial\lambda=0$ is equivalent to the condition that

$$c\lambda = (\mathbf{x}'\mathbf{x})^{-1}\mathbf{d}, \tag{2}$$

where the ajth element of $\mathbf{x}$ is the deviation of the ath observation of the jth variable from its group mean, i.e., $x_{aj}=(X_{aj}-\bar{x}_{1j})$ or $(X_{aj}-\bar{x}_{2j})$ depending on whether the ath item is a member of group 1 or group 2, $[d_j]=(\bar{x}_{1j}-\bar{x}_{2j})$, and c is a constant which can be taken as unity since only the relative sizes of the elements of λ are of interest. No other set of weighting coefficients applied to the data can yield a larger G than that defined by (2).[1]

Once λ has been found, each item vector is transformed to the appropriate Z, i.e., $Z_a=\lambda'\mathbf{X}_a$. From the resulting set of Z's *the* discriminant, $\hat{z}$, is computed; this may be taken as the unweighted average, $(\bar{z}_1+\bar{z}_2)/2$, if the parent variance of Z is known or thought to be the same in both groups, or the weighted average, $(s_2\bar{z}_1+s_1\bar{z}_2)/(s_1+s_2)$, if, as here, it is not (Fisher, 1936). If $\bar{z}_1<\bar{z}_2$, the ath item is assigned to class 1 if $Z_a<\hat{z}$ and to class 2 if $Z_a>\hat{z}$. More generally, if the underlying assumptions—that the sampling is random and the distribution of Z is normal—are satisfied and the samples are large, the probability of misclassification in *future* random samples is minimized, and is the same for items drawn from either class. In the work discussed here the sampling is not random, the Z's are based on so few variables that they cannot be even approximately normal unless the X's are normal, very little is known about the distribution of the X's, most of the samples are small, and it is not at all likely that we shall have any considerable number of future samples, random or not, with which to compare present results. As so often happens in petrology, we would like to make probability statements but the situation is such that we must rest content with the best possible sample

[1] The matrix formulation has the usual advantage of compactness, and (2) is obviously ideal for programmed calculation. An excellent derivation free of matrix notation is given by Hoel (1962).

description; we can and indeed will make predictions, but will be unable to attach probabilities to them. All we can hope is that we have done as well as the data permit. In this connection, it is to be noted that no linear transformation will recapture the initial data classification more efficiently than the one with coefficients calculated from (2).

Turning now to the practical problem announced as the subject of this lecture, in a discriminant function based, for instance, on first-order terms in temperature and pressure,

$$\hat{z} = \lambda_p P + \lambda_t T. \tag{3}$$

$\hat{z}$, λ_p, and λ_t are computed from the data, so that the trace of the curve for assigned values of P (or T) can be found directly. The equation may of course be written

$$T = \frac{\hat{z}}{\lambda_t} - \frac{\lambda_p}{\lambda_t} P, \tag{4}$$

in which $(\hat{z}/\lambda_t)$ and $(-\lambda_p/\lambda_t)$ are, respectively, the P intercept and the slope of a straight line in the P-T plane. The resemblance to an ordinary regression line is obvious but only superficial. There is only one discriminant function for any set of data, its trace need not include the mean of either variable, there is no underlying assumption that either variable is dependent or independent in the regression sense, and the scatter of data points from the line is not indicative of the goodness or badness of the fit. The purpose of the line, as a sample description, is to recapture the original partition of the data into two qualitatively distinct classes, and its success or failure is to be judged only by the number of correct assignments it makes.

Some Practical Illustrations

There is no *a priori* reason for accepting the trace of the discriminant function as an estimator of the field boundary. The argument is completely empirical, and must be made by example rather than precept. In any such situation a single failure outweighs many successes, and although to date no clear counter example has been found, it is unrealistic to suppose that this happy state of affairs will persist indefinitely. It is therefore incumbent upon me to present a considerable number of examples in which the trace of the discriminant function does indeed closely approximate the field boundary proposed by the original investigator. The most convenient way to do this is by a series of diagrams, and I shall, therefore, take advantage of the prerogatives accorded a presidential address to labor this note with an outrageous number of illustrations. In all these illustrations the following conventions apply:

1. When only two phase assemblages are recorded they are represented by solid and open circles (or squares). In two of the diagrams there are three phase assemblages, and here additional self-explanatory symbols are used.

2. The trace of the discriminant function computed from the data plotted in the diagram is shown by a solid line.

3. The path of the published field boundary, traced from the original investigator's working drawing, is shown by a dashed line.

4. When no dashed line is shown the curves lie too close for separate plotting.

5. The discriminant function, stated for T as a function of P in all but one example, is given in Table 1, which is essentially a set of figure captions.

The variables in the discriminant functions shown by the solid lines in Figures 1a to 2b are all of first order only. Higher order terms may of course be included in the calculation of a discriminant function, exactly as in ordinary regression computations. Ideally, one would wish to have some clear statistical basis for deciding whether or not to include them but no such basis is available at present, and most of the samples are too small to support much statistical manipulation. Since the immediate objective is merely to compare the trace of a relatively simple discriminant

TABLE 1. PHASE ASSEMBLAGES AND CALCULATED DISCRIMINANT FUNCTIONS FOR SOME PUBLISHED FIELD BOUNDARIES

Fig.	Phase assemblages	Discriminant function	Reference
1a	rhombic enstatite; clinoenstatite	$T=649.0+1.95P$	Boyd and England, 1965
1b	proto?-ferrosilite; ortho-ferrosilite	$T=909.3+11.38P$	Lindsley, 1965
1c	liquid+quartz; proto?-ferrosilite	$T=1124.2+8.64P$	Lindsley, 1965
1d	quartz; coesite	$T=9.355P-188.8$	Boyd and England, 1960
2a	anorthite+gehlenite+corundum; Ca-Al-pyroxene	$T=149.2P-606.7$	Hays, 1966
2a	Ca-Al-pyroxene; grossularite+corundum	$T=16.31P+962.9$	Hays, 1966
2b	wollastonite+monticellite+gas; akermanite+gas	$T=683.8+10.83P$	Yoder, 1968*a*
2c	pargasite; almandine+diopside+forsterite+nepheline+spinel+anorthite+vapor	$T=688.4+75.85P-4.15P^2$	Boyd, 1959
2d	chlorite+clinopyroxene+anorthite+gas; anorthite+forsterite+gas	$T=613.7+46.39P-2.52P^2$	Yoder, 1966
3a	hematite+magnetite+liquid; acmite+hematite	$T=1031.2+17.85P-0.11P^2$	Gilbert, 1966
3b	liquid; liquid+plagioclase	$T=1248.9+6.84\mathrm{An}-0.039\mathrm{An}^2$	Lindsley, 1966
3b	liquid+plagioclase; plagioclase	$T=1225.0+1.61\mathrm{An}+0.022\mathrm{An}^2$	Lindsley, 1966
3c	diopside; liquid	$T=649.0+1.95P$	Boyd and England, 1963
3d	enstatite; liquid	$T=1560.9+11.53P-0.071P^2$	Boyd, England, and Davis, 1964
4a	anorthite+vapor; vapor+liquid	$T=1499.5-77.31P+4.02P^2$	Yoder, 1968*b*
4b	anorthite+vapor; vapor+liquid	$P=134.6-17.61T+0.58T^2$	Yoder, 1968*b*

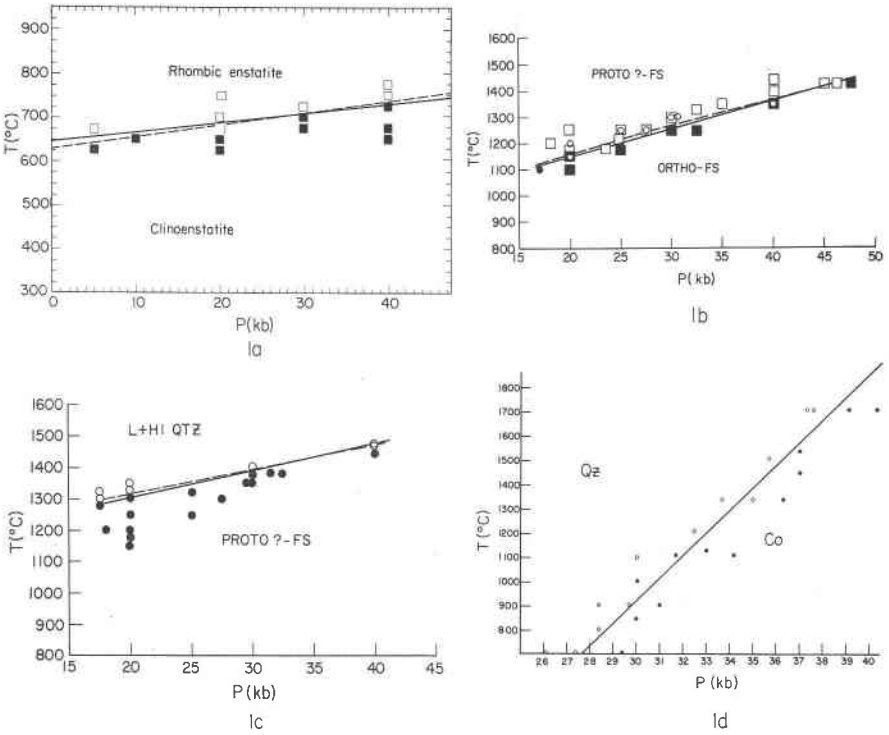


FIG. 1. Boundaries between the fields of (a) rhombic and clino-enstatite, (b) proto?- and ortho-ferrosilite, (c) proto?-ferrosilite and liquid+high quartz, (d) quartz and coesite. (See also Table 1. Symbols as defined in text.)

function with the field boundary located by inspection, the burden of the decision has been placed on the original author(s). A second-order term in one of the variables has been included in the calculation whenever the field boundary shown in the published diagram is curved. Examples are given in Figures 2c to 4b.

There is, incidentally, no requirement that the observed variables be used as such. Although the experimental petrologist measures and records P in bars and T in °C, for instance, the theorist may prefer to think in terms of the log of pressure and the reciprocal of the absolute temperature. These conversions can be made prior to the computation, so that the resulting equation is of the form

$$\hat{z} = \lambda_p \log P + \lambda_t (273.1 + T)^{-1}. \tag{5}$$

Although $\hat{z}$, λ_p, and λ_t will of course have different values in (5) and (3), (5) may be solved explicitly for T, viz.,

$$T = \frac{\lambda_t}{\hat{z} - \lambda_p \log P} - 273.1, \tag{6}$$

and plotted in a P-T coordinate system. In general, however, the trace of (6) will not coincide with the trace of (4), and there seems no a priori

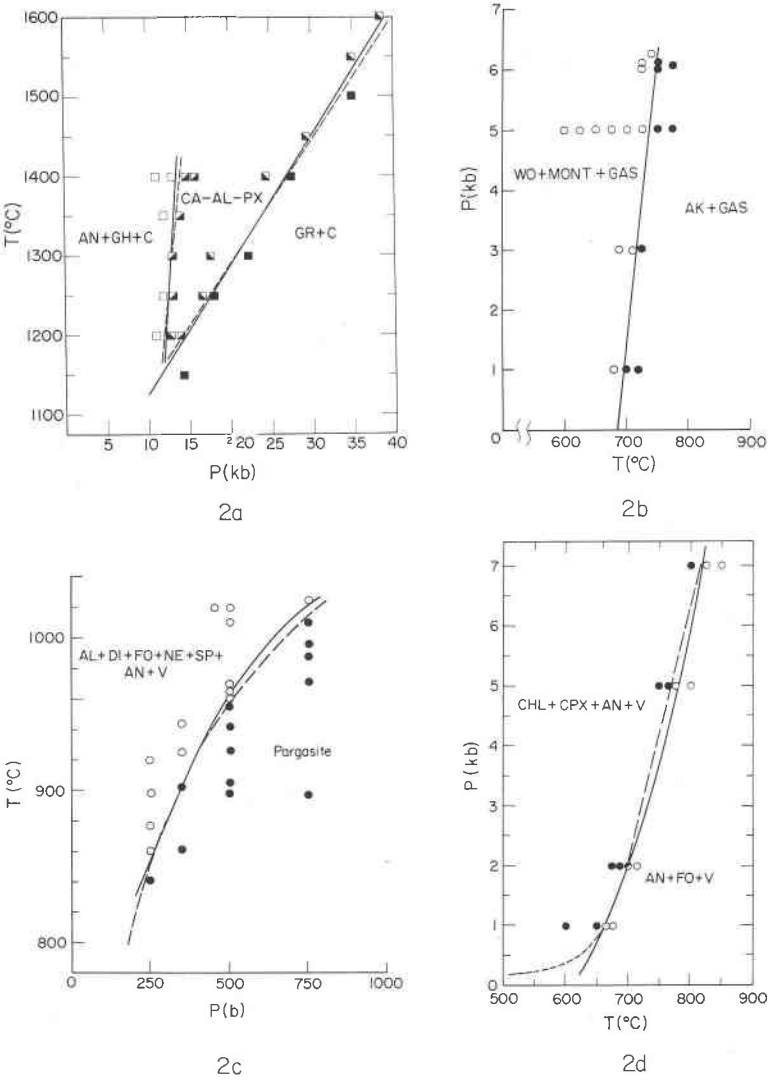


FIG. 2. Boundaries of the fields of (a) Ca-Al-pyroxene, (b) wollastonite+monticellite+gas, (c) pargasite, (d) anorthite+forsterite+vapor. (See also Table 1. Symbols as defined in text.)

basis for deciding which will either partition the data more efficiently or agree better with the field boundary located graphically by the experimentalist.

Even from cursory inspection of Figures 1–3 it is evident that the areal partitions effected by the computed discriminant functions are very like those established by the previously published field boundaries. In four of the fourteen partitions shown in these figures the agreement is so close that the two lines cannot be shown separately, and in five they are virtually coincident over large portions of the experimental range. Even in the five examples in which the boundaries do indeed differ except for a point of intersection, the divergence is usually not large in relation to experimental uncertainty. In detail, however, these graphs show the usual differences between lines fitted by calculation and inspection. Chief among these is increasing discrepancy at and beyond the limits of the experimental range. Much of this is inevitable—alternative lines fitted by inspection alone will usually behave in the same fashion—but some of it probably stems from the circumstance that in fitting the data by eye the investigator may be influenced by information not part of his investigation and hence not shown on his diagram. Curves showing the effect of pressure on the melting temperatures of diopside, enstatite, anorthite, or acmite, for instance, ought to extrapolate close to the melting points of the first three or the melting range of the fourth, as determined at atmospheric pressure. This information does not appear in either the scatter diagrams or the accompanying tables in the references on which Figures 3a, 3c, 3d, and 4[1] are based, for its determination is not part of the work reported in these references.

In each calculation the data shown on the appropriate diagram have been given unit weight, and differences between the dashed and solid lines may sometimes reflect the original investigator's decision that some of his results were more reliable than others. Something of this sort may be responsible for the discrepancy between the plagioclase liquidus curves shown in Figure 3b; the dashed line lies above the bracket at An_{20} but intersects the bracket at An_{10}, whereas the solid line intersects the bracket at An_{20} but lies below the bracket at An_{10}. Finally, there are examples of conflict for which no reasonable explanation is available, as, for instance, in the classification of points on the 500-bar section in Figure 2c or the 5-kilobar section in Figure 2d.

[1] The original working diagram from which the dashed line in Figure 4a was traced contained no data for the melting point of anorthite at atmospheric pressure. The bracket shown in Figures 4a and 4b, and included in the computation, is taken from Day and Sosman (1911). This is the only instance in which data points other than those shown on a published diagram have been used.

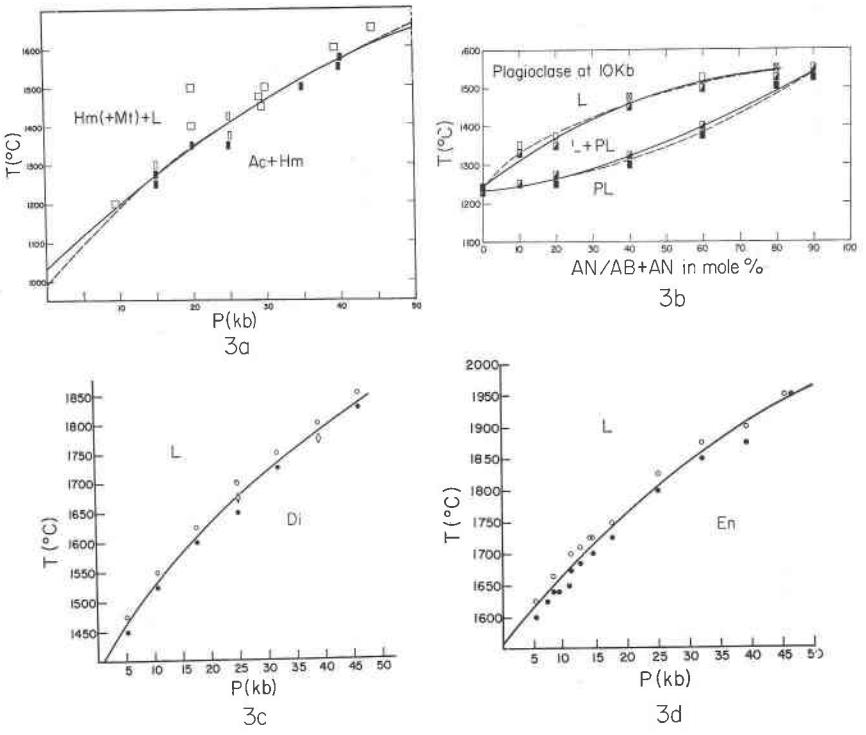


FIG. 3. Upper boundaries of the fields of (a) acmite+hematite, (b) plagioclase and plagioclase+liquid, (c) diopside, (d) enstatite. (See also Table 1. Symbols as defined in text.)

On the whole, however, in these examples the differences between computed discriminant functions and field boundaries fitted by inspection are small in an absolute sense and, for most practical purposes, negligible. It is not to be expected that agreement of this caliber will always be obtained. No unequivocal counter-example has so far been found, but it is perhaps as well to conclude this section with an example in which the quality of agreement evidently depends largely on the choice of variables used in the computation.

In the examples shown in Figures 1–3, inclusive, the computed discriminant functions—see Table 1—are all stated for temperature as an explicit function of pressure or composition. Where only linear variables are employed in the calculation, as for the first seven equations in the table, the distinction between implicit and explicit statements is purely formal. If a single-valued function for T is desired, however, T must occur only in first-order terms in the calculation. In the examples shown in Figures 2c–3d, inclusive, all of the computations involved a second-order term, but T

was not used as the second-order term, for whether the field boundary is linear or curvilinear, we speak of pressure or composition affecting melting or reaction temperatures. Actually, discriminant functions have been calculated for a number of the arrays with a second-order term in T, with results so similar to those shown in the figures that it does not seem worth recording them.

The only exception so far encountered is shown in Figure 4. The solid line in Figure 4a and the next-to-last line of Table 1 show the discriminant calculated with T as a first-order term only, and P as both first and second order. Of the 32 data points, five are intersected and five are clearly misclassified by the computed function. The field boundary provided by the original investigator intersects four points *but misclassifies none.* (As noted above, the bracket at atmospheric pressure is not part of the data in question.) If the discriminant function is computed as $P = f(T, T^2)$ instead of $T = f(P, P^2)$ however, the result is an almost perfect replica of the dashed line in Figure 4a. The trace of this function is shown in Figure 4b, and its equation is given in the last line of Table 1.

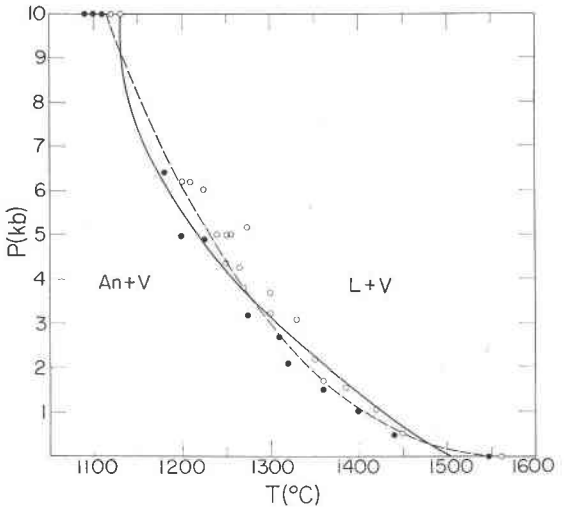
In this example the choice of variables is obviously critical. As already suggested, there is no need to use P and T at all, and there is some reason for preferring log P and $(T+273.1)^{-1}$ on theoretical grounds. A function linear in these variables would be curvilinear in P and T. The curvature of field boundaries in P-T space is ordinarily gentle, and probably could often be accommodated in this fashion. Preliminary calculations indicate, however, that for the data of Figure 4 at least one term of order greater than 1 will be needed even if P and T are replaced by log P and $(T+273.1)^{-1}$.

Advantages of a Systematic Numerical Procedure for Locating Field Boundaries

It has now been shown that a numerical procedure strange and perhaps even repulsive to many petrologists will very often yield results closely comparable to those readily obtained by simple inspection. What is the purpose of this exercise?

There is first of all the intrinsic interest of the problem, an interest that will be apparent to readers concerned either with the progressive "quantification" of the more traditional earth sciences or with the computational question of form and pattern recognition. Quite aside from any practical consequences, one would simply like to know if the thing can be done; evidently it can.

In systems simple enough so that field boundaries may be constructed graphically there is no particular advantage in computing them unless reaction rates are such that in their vicinity equilibrium cannot be at-



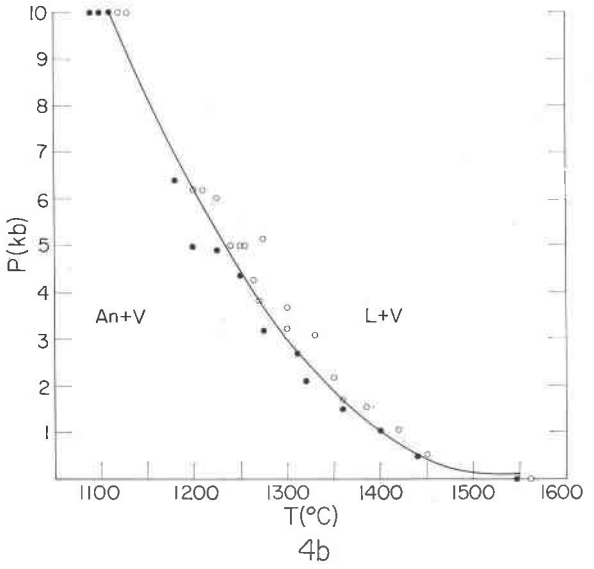


FIG. 4. Boundaries between the fields of anorthite+vapor and liquid+vapor, (a) computed for $T=f(P, P^2)$, (b) computed for $P=f(T, T^2)$. The dashed line in (a) is traced from the investigator's working drawing. (See also Table 1.)

tained in reasonable time. The construction of field boundaries by inspection requires a considerable concentration of data in the immediate vicinity of the boundary, and on both sides of it, for otherwise there will be many possible constructions and no particular reason for preferring one to another. For any given set of variables formed from a given data array, however, there is only one discriminant function, and examination of Figures 1a, c, and d, 2b and c, and 3a suggests that the discriminant function is much less influenced by dispersion than either regression techniques or fitting by inspection. If a boundary were to be located by computation rather than inspection, the requirement that data be concentrated in its immediate vicinity—the definition, that is to say, of "immediateness"—could be materially relaxed. This could be a considerable practical advantage, whether it led simply to an increase in haphazard sampling or to a properly randomized sampling of the broadened critical region.

Finally, and probably of most importance in the near- and mid-future, if the number of variables is such as to preclude graphical representation, so that the location of field boundaries by inspection is impossible, a numerical procedure that will perform the multidimensional equivalent of this operation becomes indispensable. Given the knowledge that results obtained by it are in good agreement with those reached by inspection where inspection is possible, it seems reasonable to suppose that the discriminant function could be used to "locate" phase fields in multicomponent space, where inspection is impossible.

Acknowledgments

All examples used in this paper are of field boundaries determined at the Geophysical Laboratory and already published, mostly in recent Annual Reports of the Laboratory. The analysis requires more information than is usually printed, however, and I am grateful to all the authors whose experimental results are used here for making available to me their working drawings and data tables. I am also indebted to these and other fellows and staff members of the Laboratory for extensive discussion, both of specific examples and of the broad general issues involved.

Most of the calculations were done during the late fall of 1967, when the pressure of Society and other business made it impossible for me to devote full attention to the project. W. Bryan kindly prepared most of the data decks and assisted with the actual machine computation.

A preliminary version of this paper (Chayes, 1968), translated into Russian and based largely on the examples shown here in Figures 1d, 3c, and 3d, is expected to appear shortly in a Festschrift volume, honoring A. B. Vistelius.

References

Boyd, F. R. (1959) Hydrothermal investigation of amphiboles. *In* P. H. Abelson (ed.), *Researches in Geochemistry*. John Wiley & Sons, Inc., New York, p. 377–396.

——— and J. L. England (1960) The quartz-coesite transition. *J. Geophys. Res.* **65**, 749–756.

——— AND ——— (1963) Effect of pressure on the melting of diopside, $CaMgSi_3O_8$, and albite, $NaAlSi_3O_8$, in the range up to 50 kb. *J. Geophys. Res.* **68,** 311–323.

——— AND ——— (1965) The rhombic enstatite-clinoenstatite inversion. *Carnegie Inst. Wash. Year Book.* **64,** 117–119.

———, ———, and B. T. C. Davis (1964) Effect of pressure on the melting and polymorphism of enstatite, $MgSiO_3$. *J. Geophys. Res.* **69,** 2101–2109.

CHAYES, F. (1968) On locating field boundaries in simple phase diagrams by means of discriminant functions (in Russian). In *Problems of Mathematical Geology* (M. A. Romanova and O. V. Sarmanov, editors), Nauka—Russian Scientific Publishing House (in press).

DAY, A. L., AND R. SOSMAN (1911) The melting points of minerals in the light of recent investigations on the gas thermometer. *Amer. J. Sci.*, 4th ser., **31,** 341–349.

FISHER, R. A. (1936) The use of multiple measurements in taxonomic problems. *Ann. Eugenics.* **8,** 179–188.

GILBERT, M. CHARLES (1966) Acmite. *Carnegie Inst. Wash. Year Book.* **65,** 241–244.

HAYS, J. F. (1966) Lime-alumina-silica. *Carnegie Inst. Wash. Year Book.* **65,** 234–239.

HOEL, P. G. (1962) *Introduction to mathematical statistics.* 3d ed. John Wiley & Sons, Inc., New York.

LINDSLEY, D. H. (1965) Ferrosilite. *Carnegie Inst. Wash. Year Book.* **64,** 148–149.

——— (1966) Melting relations of plagioclase at high pressures. *Carnegie Inst. Wash. Year Book.* **65,** 204–205.

YODER, H. S., JR. (1966) Spilites and serpentinites. *Carnegie Inst. Wash. Year Book* **65,** 269–279.

——— (1968a) Akermanite and related melilite-bearing assemblages. *Carnegie Inst. Wash. Year Book.* **66,** 471–477.

——— (1968b) Experimental studies bearing on the origin of anorthosite. In *Origin of anorthosites,* New York State Geological Survey (in press).

CLASSIC MINERAL OCCURRENCES

Several years ago the Council of the MSA discussed the desirability of including in *The American Mineralogist* more material of interest to the semi- and nonprofessional members of the Society. They agreed to revive the old "Famous Mineral Localities" series, in a form that would interest the amateur but retain the editorial standards of the journal. Prof. R. H. Jahns was assigned as editor of the new series, which it was hoped would summarize data on the location, geologic setting, minerals, paragenesis and origin of classic mineral occurrences.

The following paper by John Sinkankas, on Amelia, Virginia, is the first result of this program. While it does not fully review the previous work in this area, and might well have been published as a regular contribution to this journal, it does effectively summarize the more recent finds in this important groups of pegmatites.

Unfortunately about twenty other projected manuscripts for this series have not yet been forthcoming. Hopefully the publication of this first contribution will encourage some of the other authors, previously invited, to help in a successful continuation of this project. And anyone with a suggestion for a suitable contribution, either by himself or someone else, should *outline* his proposal to Prof. R. H. Jahns, Dean, School of Earth Sciences, Stanford University, Stanford, California, with a copy to William T. Holser, Editor of the journal.