The Existence Diagram — A Useful Theoretical Tool in Applied Physics

G. ECKER

Institut für Theoretische Physik, Ruhr-Universität Bochum

(Z. Naturforsch. **26 a**, 935—939 [1971]; received 5 March 1971)

The complexity of theoretical work in physics applications often forces severe neglections and crude approximations. Insufficient experimental knowledge of important parameters increases the uncertainty. In this situation the E-diagram method still allows to arrive at useful results through controlled choice of the approximations and sensible limitation of the aims. Choosing "unidirected" deviations only, existence limits can be calculated which restrict the existence of the system to a narrow region — the existence area — of the coordinate plane of variables under consideration. The example of the vacuum arc cathode demonstrates the procedure and its suitablility.

Introduction

One can distinguish two basicely different types of theoretical work:

The one concentrates on finding new concepts to describe the results of experimental frontier research.

The other considers within the range of well-established physical laws phenomena of applied physics which due to their complexity could not be analysed yet.

The second type — the theory in the range of applied physics — is our topic here. It confronts the theoretician primarily with two tasks:

a) The comprehension and principal formulation of the problem in *full*.

b) The construction and evaluation of a workable model from this general description.

The solution of this task is frequently precluded by the fact that the simplest realistic model is still too complex for analytical or even numerical methods and/or the various parameter entering the description are insufficiently known from experiment.

Under these circumstances it has become widespread practice to disect the problem into independent parts to use crude approximations and to take a chance in picking values for the parameters from the total information available.

This procedure has the drawback that the treatment in parts looses all information which follows from the interaction of the parts and that the uncertainty of the crude approximations and the parameters is reflected in an uncertainty of the results.

Reprints request to Prof. Dr. G. ECKER, Ruhr-Universität Bochum, Theoretische Physik, Lehrstuhl I, *D-4630 Bochum-Querenburg*, Buscheystraße, Gebäude N B.

In this situation which occurs frequently in the theory of applications the use of the *E-diagram* can still yield useful and reliable information.

Concept of the *E*-Diagram

If one is forced to neglect terms in the model which are relevant then one causes deviations from the exact solution. If experimental parameters are inaccurately known this too will result in deviations. The effect of the superposition of these deviations on the final result is usually as difficult to trace as the solution of the rigorous model. Therefore it is generally not considered. Consequently the results present themselves with an unspecified uncertainty which yields their value questionable.

Of course, a method — like the existence diagram method — cannot get rid of these deviations. But it can improve the situation by using the deviations in a controlled form and restricting the evaluation to sensibly limited aspects.

In this spirit the *E*-diagram method suggests:

a) For each study we select from the set of variables $\{\zeta_\nu\}$ which describe the system two of particular interest (ζ_a, ζ_b).

b) We do not aim to find the rigorous solution ζ_{as}, ζ_{bs} (point of solution) but will be satisfied if we can reliably restrict existence of our system to a reasonbaly small region within the (ζ_a, ζ_b)-plane. This region will be called existence region (E-region) or existence area (E-area).

c) Such an E-region can be limited by displaced solution characteristics $\overline{\zeta}_{a\nu}(\zeta_b)$ which are displaced due to neglections and parameter approximations chosen such that the deviations from the true characteristics $\Delta\zeta_{a\nu} = \overline{\zeta}_{a\nu} - \zeta_{a\nu}$ are positive or negative semi definite. We call such deviations "unidirected".

d) The "existence area" provides us with a limited but reliable information even for the case of crude approximations. It presents this information in a transparent graphical form. The size of the E-area even for severe approximations can be small enough to draw useful conclusions.

The next section describes these concepts in detail.

Method of the *E*-Diagram

1. The Exact Solution

Suppose our problem can be described by a set of N dependent variables $\{\zeta_i\}$, a set of independent variables $\{\xi_i\}$ and a set of parameters $\{p_i\}$ through a system of general operator equations

$$\Omega_j(\zeta_i, \xi_i, p_i) = 0 \quad (1)$$

with a set of boundary conditions

$$B_j(\zeta_i, \xi_i, p_i) = 0 . \quad (2)$$

If the set of operator Eqs. (1) and boundary conditions (2) is appropriate to describe our system then it defines N solution functions

$$S_j(\zeta_i, \xi_i, p_i) = 0 \quad (3)$$

which define the values of the dependent variables $\{\zeta_i\}$ as functions, of the independent variables $\{\xi_i\}$ and the parameters $\{p_i\}$.

We now consider the presentation of the set of ζ-functions in a N-dimensional ζ-space.

In this space a solution function S_i determines a $(N-1)$-dimensional hyperspace.

The intersection of any combination of $(N-1)$ of these hyperspaces determines a one dimensional line (characteristic) in the ζ-space which passes through the point of solution. According to the total number of possible combinations of the N solution functions in groups to $(N-1)$ we expect to find N of these characteristics.

For practical applications it is of particular interest to consider the projection of these characteristics into a coordinate plan (ζ_a, ζ_b). Of course, these projections will also be a system of N characteristics * $\zeta_{a\nu}(\zeta_b)$ intersecting in one point — the projection of the point of solution (PS).

* The choice of the representation $\zeta_{a\nu}(\zeta_b) = \zeta_a$ of course has only reasons of convenience and means no preference for ζ_a in any way.

As a basis of the following discussion this is schematically demonstrated in Fig. 1 for the case $N=6$ where — for the sake of simplicity — we have chosen the $\zeta_{a\nu}(\zeta_b)$-characteristics as straight lines.

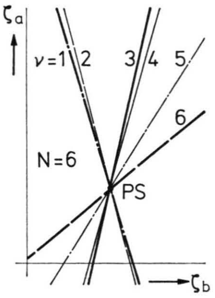

Fig. 1. Schematic presentation of linear characteristics $\zeta_a = \zeta_{a\nu}(\zeta_b)$ for the case of $N=6$ variables. PS projection of the point of solution into the (ζ_a, ζ_b) plane.

The analytical presentation of these $\zeta_{a\nu}(\zeta_b)$ characteristics follows through elimination of all dependent variables except ζ_a, ζ_b from a combination of $(N-1)$ S-functions.

2. The Approximate Solution

Deviations of the $\zeta_{a\nu}(\zeta_b)$-characteristics are caused by

a) Insufficient knowledge of the experimentally given parameters: Δp_j.

b) Model assumptions in the operator equations and boundary conditions (1, 2): $\Delta\Omega_j$; ΔB_j.

A given set of Δp_j and $\Delta\Omega_j$, ΔB_j produces corrections which displace the characteristic (ν) from $\zeta_{a\nu}(\zeta_b)$ to $\overline{\zeta}_{a\nu}(\zeta_b)$. The displacement $\Delta\zeta_{a\nu} = \overline{\zeta}_{a\nu}(\zeta_b) - \zeta_{a\nu}(\zeta_b)$ generally depends in sign and magnitude on Ω, B, $\Delta\Omega$ and ΔB.

However, in many practical cases it is not difficult to choose on grounds of qualitative reasoning all $\Delta\Omega_j$, ΔB_j and Δp_j so that for a given characteristic (ν) the deviations $\Delta\zeta_{a\nu}(\zeta_b)$ are semi definite, — positive or negative.

In other words we can choose for a given characteristic (ν) our neglections and parameters such that we can reliably state on what side of the displaced characteristic the point of solution will be found. Therefore such a displaced characteristic $\overline{\zeta}_{a\nu}$ with "unidirected deviations" can be considered as "existence limit".

Now it is easy to construct existence areas:

If we consider only one characteristic $\nu = 1$ and choose e. g.

$$\Delta\zeta_{a1}(\zeta_b) > 0 \quad (4)$$

then te toal half plane below the existence limit $\bar{\zeta}_{a1}(\zeta_b)$ is the existence area.

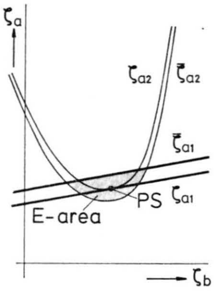

Fig. 2. Example for a finite E-area in the case of two characteristics (schematic).

A better result is already achieved if one considers two characteristics ($\nu = 1, 2$), e. g.

$$\Delta\zeta_{a1}(\zeta_b) > 0; \quad \Delta\zeta_{a2}(\zeta_b) < 0 \quad (5)$$

The existence area

$$\bar{\zeta}_{a1} > \zeta_a > \bar{\zeta}_{a2} \quad (6)$$

is for linear characteristics still an infinite sector. For other characteristics it may be finite (see e. g. the schematical presentation of Fig. 2 where for the sake of simplicity we used

$$\Delta\zeta_{a1} = +\text{const}; \quad \Delta\zeta_{a2} = -\text{const}$$

so that the characteristics are shifted only but not distorted).

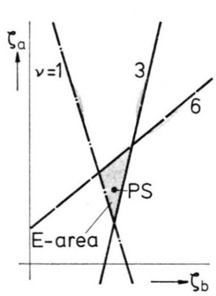

Fig. 3. E-area limited by three characteristics only in accord with the example of Eq. (7).

For linear characteristics three limits are necessary to produce a finite existence region. Figure 3 shows the case

$$\Delta\zeta_{a1} < 0; \quad \Delta\zeta_{a3} < 0; \quad \Delta\zeta_{a6} > 0 \quad (7)$$

for the example of Fig. 1.

If we take more than three limits into account, e. g. additionally

$$\Delta\zeta_{a2} > 0; \quad \Delta\zeta_{a4} > 0; \quad \Delta\zeta_{a5} > 0 \quad (8)$$

the existence area may be unaffected or further reduced depending on the circumstances (see e. g. Figs. 4 and 5).

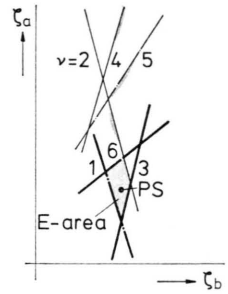

Fig. 4.

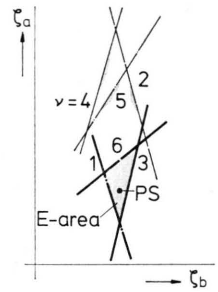

Fig. 5.

Figs. 4 and 5. Demonstration of the effect of the additional characteristics on the limitation of the E-area of Fig. 3.

Trivially additional characteristics can never cause an increase of the existence area.

Example of the Vacuum Arc

As an example for practical demonstration we choose the theory of the *vacuum arc cathode*.

The theory of the arc cathode is in general extremely complicated (Ref. [1]). There are at least three model regions in strong interaction:

a) The electrode, b) The sheath, c) The plasma body.

In actual fact the electrode should be further subdivided into a molten and a solid region. There are different types of sheathes, depending on whether one studies space charge effects or kinetic models. The plasma body in its full extension is composed of different models depending on whether the Saha equation is valid, the assumption of quasi neutrality holds, etc.

Moreover the whole phenomenon is non-stationary.

Essential simplification is achieved by the assumption that the system of the electrode and the plasma immediately in front of the cathode spot is self-contained and stationary. This assumption is corroborated by experimental evidence. The main advantages are that we can use for the neutral par-

[1] G. Ecker, Ergebn. exakt. Naturwiss. **33**, 1 [1961].

ticle density in the plasma sphere the density given by the vapour pressure of the electrode material at the temperature of the electrode spot and that we can calculate the conditions in the plasma from the interaction within itself and with the electrode surface only.

But even so the problem is not tractable. The obstructing problems are:

α) The ion current from the plasma to the surface of the electrode is uncertain due to the lack of knowledge about the radial losses and the uncertainty of the electron temperature which in turn is a consequence of the uncertainty of the energy balance in the plasma.

β) The extension of the electrode sheath is an unknown quantity. It may be replaced by the value of the potential fall across the sheath (U_c), but this too is an uncertain figure.

γ) Due to the assumption of self-containment it is very difficult to find information about the evaporation rate of the neutrals from the electrode surface since this would require knowlede of the distribution of the particle density throughout the whole plasma body.

δ) The energy balance in the cathode, particularly in the cathode spot, is the worst problem of them all. Even with modern computer technique it is impossible to solve for the temperature distribution in the electrode under the influence of the ion bombardment, the evaporation from the surface, the heat conduction from a circular area into the electrode, the heat loss due to melting at the surface, etc.

Nevertheless, in this desperate situation the existence diagram yields useful information.

From the large number of dependent variables describing the vacuum cathode, the temperature in the electrode cathode spot (T) and the current density (j) are those of predominating experimental and theoretical interest. We therefore choose a presentation of our problem in the $(T-j)$-plane.

We further choose three characteristics. The first is determined through the law of charge conservation, the second is related to the law of neutral particle conservation and the third formulates the total energy balance.

We can here only scetch the analysis:

In the case of the charge conservation we use the current continuity relation

$$j = j_e(E_c, T) + j_+ \tag{10}$$

where j_+ and j_e designate respectively the ion current density from the plasma and the electron emission from the cathode surface. E_c is the electric field at the cathode surface. T is the temperature of the cathode spot at the electrode surface.

The emission current density j_e can be evaluated with the knowledge of T, E_c and the work function φ via the emission probability and the supply function (Ref. [2]).

The electric field E_c follows from the current density and the cathode drop U_c through the McKeown equation (Ref. [3]).

$$E_c^2 \propto U_c^{1/2} \left\{ j_+ \left(\frac{m_+}{m_-}\right)^{1/2} - j_e \right\} \tag{11}$$

In the spirit of our E-diagram technique we choose the value U_c such that the corresponding characteristic $\overline{T}_{cc}(j)$ is shifted to lower values of T.

The value of the ion current density j_+ is prescribed by the qualities in the plasma which are difficult to calculate correctly. It is however not difficult to make a reliable estimate (Ref. [4]) which shows that

$$j_+/j \leqq U_c/(U_c + U_i) \tag{12}$$

holds. The $(<)$-sign again shifts the characteristic $\overline{T}_{cc}(j)$ to lower values of T.

From (10), (11) and (12) – using the $(=)$-sign in the latter – we calculate $\overline{T}_{cc}$ shown in Fig. 6 for the example of copper.

We know little about the neutral particle balance and in fact within our model cannot calculate this balance. But there is a simple reliable requirement which we can use for the calculation of a corresponding existence limit (Ref. [5]). The requirement reads

$$j_+ \leqq \frac{e}{4} \frac{p_{ev}(T)}{(m_+ \, k\,T/3)^{1/2}} \tag{13}$$

where $p_{ev}(T)$ designates the vapour pressure of the electrode material at the temperature T. It states that the number of ions going back to the cathode surface must be smaller then the number of neutrals

[2] E. L. Murphy and R. H. Good, Jr., Phys. Rev. **102**, 1464 [1956].

[3] S. S. MacKeown, Phys. Rev. **34**, 611 [1929].

[4] G. Ecker, G. E.-R & D-Report No. 70 - C 213 [1970].

[5] T. H. Lee and A. Greenwood, J. Appl. Phys. **32**, 916 [1961].

evaporated from the cathode surface, — a fact which is selfevident for a vacuum arc.

Evaluating this Eq. (13) together with Eqs. (10) and (11) engaging the (=)-sign in Eq. (13) we find another limit $\overline{T}_{nc}(j)$ which according to the (<)-sign in Eq. (13) is shifted to lower values of the temperature. It too is shown in Fig. 6 for copper.

For the third existence limit $\overline{T}_{ec}$ we use the energy balance in a crude approximation. We neglect all loss processes like heat conduction, latent heat of melting, etc. except the energy taken away by evaporation and electron emission from the metal. Furthermore we choose the energy gain provided by the ion bombardment and the energy created by Joule's heating to the larger side. In general we dispose of all approximations and parameter values so that we increase the temperature compared to reality $\overline{T}_{ec}(j) > T_{ec}(j)$. The basic formula is then for the calculation of the corresponding limit

$$j\left[(U_c+U_i)\,\frac{j_+(j,\overline{T}_{ec})}{j}-\varphi+\frac{0.31}{\sqrt{\pi}}\,(j\,I)^{1/2}\,\varrho\right] = \frac{\varepsilon}{4}\,\frac{p_{ev}(\overline{T}_{ec})}{(m_+\,k\,\overline{T}_{ec}/3)^{1/2}} \qquad (14)$$

where ϱ is the specific resistivity, ε the evaporation energy and U_i the ionization potential. The evaluation of this formula simultaneously with Eqs. (10) and (11) provides us with the limit $\overline{T}_{ec}$ of Fig. 6 which according to our assumptions is shifted to higher values of temperature. Due to the resistive heating effects this energy limit depends — in principle — on the total current. However, for copper the effect of resistivity is negligible in the range $I < 200$ A (I total spot current) and we restrict ourselves in Fig. 6 to this low current range.

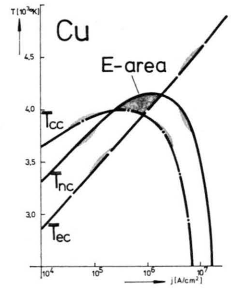

Fig. 6. Practical example of the existence diagram for the vacuum arc copper cathode. T temperature, j current density in the spot at the cathode surface.

In spite of the fact that the very complex situation in the vacuum arc cathode forced us to introduce severe approximations we see from Fig. 6 that the result described by the three existence limits offers an existence area which is quite small. An area which in fact is sufficiently small to allow conclusions with respect to current density and temperature which are theoretically and experimentally valuable and probably not less accurate then the experimental informations available for comparison.