BULLETIN OF
MATHEMATICAL BIOPHYSICS
VOLUME 5, 1943

A MATHEMATICAL THEORY OF THE AFFECTIVE PSYCHOSES

JEROME Y. LETTVIN AND WALTER PITTS

BOSTON CITY HOSPITAL AND THE UNIVERSITY OF CHICAGO

The theory introduces two variables ϕ and ψ. The first represents the intensity of emotion, the second measures the intensity of activity. A set of integrodifferential equations is assumed to govern the variation of ϕ and ψ with respect to time. Since for increasing values of ϕ the conduct of the organism varies from great impassivity through a normal level of feeling to extremes of a circular depression or catatonic excitement; whereas an increase of ψ results in a transition from stupor to manic excitement, the solutions of the equations represent quantitative specifications of different psychotic states.

In the following discussion we shall develop a mathematical theory of the group of mental disorders which may be characterized in the following fashion: That the characteristic course of the disease may be essentially described in terms of the vicissitudes of two variables, the first representing the level of feeling, affect, or emotion in the organism, and the other the level of activity or conation. This group we consider to comprise the circular insanities, the reactive psychoses, and the catatonia of Kahlbaum. It may also be supposed to include affective disorders superimposed upon psychoses of another type, and perhaps also, with a more specific interpretation of the determining variables, certain forms of neurosis; but we shall not consider these latter cases in detail.

We shall denote the fundamental quantities of our theory, regarded as functions of the time t, by the symbols $\phi(t)$ and $\psi(t)$. For increasing values of $\phi(t)$ the conduct of the organism will vary from great impassivity to a normal level of feeling, then to normal strong emotion, and finally to the extremes of a circular depression or catatonic excitement. (We do not here take account of the quality of the emotion). If $\psi(t)$ rises indefinitely, the organism will pass from the greatest stupor and inactivity to normal conation, and ultimately to manic excitement or an idiomotor seizure. We shall measure these variables from an origin placed at the resting values normal for the organism, so that positive values of $\phi(t)$ and $\psi(t)$ correspond to supernormal affective and conative levels, and negative values to subnormal levels.

Consider the change $\Delta \phi$ of $\phi(t)$ during a small interval of time $t, t + \Delta t$. This quantity may be supposed to arise as the sum of three contributions, say $\Delta_1 \phi$, $\Delta_2 \phi$, $\Delta_3 \phi$. The first represents the influence of the contemporary environment at the time in question; for us this can be taken only as given, and specified as an empirical function $M(t)$. The second, $\Delta_2 \phi$, represents the effect of the previous experience of the organism; we may arrive at an expression for it in the following fashion.

The quantity $\Delta_2 \phi(t)$ for any subject is an average over a large number of terms, each of which represents the affectivity of the reaction he habitually makes to some particular kind of situation. Let us denote one of these terms by $\Delta_2 \phi_A(t)$, and consider the contribution to it of the events occurring during some very short interval of time in the past, say that from η until $\eta + d\eta$, where $0 \leqq \eta < t$. Denote this contribution by $\Delta_\eta \Delta_2 \phi_A(t)$.

We observe first that $\Delta_\eta \Delta_2 \phi_A(t)$ depends upon the degree to which the organism is concerned with situations of the given type at the time η; times when he is little or not at all concerned with them will have a negligible direct effect upon his future reactions to them, whereas periods of very great concern will usually have a very large effect. Denoting this concern by a function $Q_A(\eta)$, we shall embody it as a factor in the expression for $\Delta_\eta \Delta_2 \phi_A$, as is simplest to satisfy these conditions. Suppose then that $Q(\eta) > 0$, and let the subject's concern at time η with the given kind of situation have been attended with emotion, so that $\phi(\eta)$ is large. By the principles of learning, this occasion will cause recurrences of this type of matter to be treated with greater emotion than otherwise, or to manifest a higher value of ϕ; and this will be the larger, as $\phi(\eta)$ is greater. But if the subject's reactions at the time η have been of primarily conative type, with a high value for $\psi(\eta)$ and a correspondingly reduced emotion, the recurrences of the same kind of situation will also exhibit a smaller ϕ. This is to say, that $\Delta_\eta \Delta_2 \phi_A(t)$ increases with the value of $\phi(\eta)$ and decreases with $\psi(\eta)$, while the analogously defined $\Delta_\eta \Delta_2 \psi_A(t)$ behaves in an inverse fashion.

We shall suppose that this relation is a simple proportionality, or

$$\begin{aligned} \Delta_\eta \Delta_2 \phi_A(t) &= \beta' Q_A(\eta)\left[\phi(\eta) - \varepsilon_1 \psi(\eta)\right] \\ \Delta_\eta \Delta_2 \psi_A(t) &= \beta' Q_A(\eta)\left[\psi(\eta) - \varepsilon_2 \phi(\eta)\right] \end{aligned} \tag{1}$$

in which β', ε_1, and ε_2 are suitable constants of proportionality.

But the strength of any contribution to $\Delta_2 \phi_A(t)$ from the past will in general be less in an amount dependent upon its remoteness in time, by ordinary forgetting; so that we must multiply the left of equation (1) by some decreasing function of the distance $t - \eta$ into the past. For this we shall select the function $e^{-\mu x}$, which seems to be a fair representation of the "law of forgetting", according to Ebbinghaus and others; the undetermined coefficient μ here allows us to reserve to psychoanalysts and other people who do not consider that an affecting experience ever becomes less efficient, the option of setting $\mu = 0$. We shall see, however, that this assumption requires greatly increased strength in the homeostatic mechanism of the organism to avoid continuously disturbed behavior.

The total magnitude of $\Delta_2\,\phi(t)$ will be obtained by adding together all the quantities $e^{-\mu(t-\eta)}\,\Delta_\eta\,\Delta_2\,\phi_A(t)$ corresponding to a set of past intervals of time which covers $(0\,,\,t)$ completely, and then summing quantities of this kind for each type of situation: Being linear they superpose, and if we allow the intervals into which we have divided a past to increase without limit in number, and their lengths to become uniformly indefinitely small, the error we have made in supposing that ϕ and ψ do not change in the interior of each interval approaches zero, and the sum will be transformed into an integral. We derive

$$\Delta_2\,\phi(t) = \beta' \int_0^t Q(\eta)\, e^{-\mu(t-\eta)} [\phi(\eta) - \varepsilon_1\,\psi(\eta)]\, d\eta\,, \tag{2}$$

together with a corresponding expression for $\Delta_2\,\psi(t)$, where $Q(t)$ is the same kind of average of the Q_A's that $\phi(t)$ is of the ϕ_A's. This $Q(t)$ is in principle determinable empirically, in terms of the intuitive characterization we have given it above, but not easily so in fact. We may approximate it in somewhat the following fashion. On the average, those matters will tend to be of concern to a subject which excite his feelings, emotions, activity; and they will be of the more concern to him, the more they do so. This consideration will perhaps make it plausible to put $Q = \gamma\,\phi$ into equation (2) and its analogue for ψ, where γ is a constant of proportionality.

The last part of $\Delta\,\phi$, the term $\Delta_3\,\phi(t)$, arises out of general homeostatic mechanisms of the organism. These tend to keep the values of ϕ and ψ close to the normal levels: whenever a deviation occurs from these levels, it brings into play restoring forces whose strength increases with the extent of the deviation and tends to reduce it. We may represent this component of $\Delta\,\phi$ and $\Delta\,\psi$ by a function f, so that $\Delta_3\,\phi = -f(\phi)$, $\Delta_3\,\psi = -f(\psi)$, where f is properly monotone and never negative. If we now pass to the limit as $\Delta\,t$ becomes small, and set $\beta = \beta'$, we shall obtain the equations

$$\frac{d\phi}{dt} = \beta \int_0^t \phi(\eta) e^{-\mu(t-\eta)} [\phi(\eta) - \varepsilon_1 \psi(\eta)] \, d\eta + M(t) - f[\phi(t)],$$
$$\frac{d\psi}{dt} = \beta \int_0^t \phi(\eta) e^{-\mu(t-\eta)} [\psi(\eta) - \varepsilon_2 \phi(\eta)] \, d\eta + \bar{M}(t) - f[\phi(t)]. \tag{3}$$

These equations form the basis of our theory, and, with the addition of a few subsidiary hypotheses about organic interventions, everything is to be deduced from them. In the present study we shall in general treat M and $\bar{M}$, which represent the external fortunes or misfortunes of the organism, as if they consisted of a series of impulsive shocks between which they are relatively negligible. The "small pains and troubles of daily existence," in Schopenhauer's phrase, which most reasonably go into this term, may be taken account of in the determination of the normal values whence we measure ϕ and ψ.

Our first problem is to determine the simplest form for the restoring function $f(x)$ which will prevent the functions ϕ and ψ from increasing or decreasing without bound and assuming values which are physiologically meaningless. Considering $f(x)$ to be a polynomial, we see that its leading term must be an odd power of x, in order that the homeostasis may work against both positively and negatively abnormal values. If it is in addition not of the first degree, we may demonstrate its adequacy to limit ϕ and ψ in the following way.

Multiply the equations (3) by $e^{\mu t}$ throughout, differentiate, and cancel the exponential factor. We have then

$$\phi'' = -[\mu + f'(\phi)]\,\phi' + \beta\phi^2 - \beta\varepsilon_1\phi\psi - \mu f(\phi)$$
$$\psi'' = -[\mu + f'(\psi)]\,\psi' + \beta\phi\psi - \beta\varepsilon_2\phi^2 - \mu f(\psi), \tag{4}$$

with the initial conditions $\phi'(0) = -f[\phi(0)]$, $\psi'(0) = -f[\psi(0)]$, where we have dropped the terms in M and $\bar{M}$, since the variables cannot become infinite during the integrable impulses of which M and $\bar{M}$ consist, and at other times they obey the equations (4). Now time may be divided into two sorts of interval: those where $|\phi| \geqq |\psi|$, and those where the reverse inequality holds. Considering intervals of the former type, we find, upon rearranging the first of equations (4),

$$\phi' = \frac{\beta\phi^2 - \beta\varepsilon_1\phi\psi}{\mu + f'(\phi)} - \frac{\phi''}{\mu + f'(\phi)} - \frac{\mu f(\phi)}{\mu + f'(\phi)}$$
$$= O(1) - O(\phi) - \frac{\phi''}{\mu + f'(\phi)}, \tag{5}$$

as $\phi \to \infty$ through intervals of the given type, under the given assumptions. Now we may distinguish these kinds of intervals of time. As ϕ increases through intervals of the first kind, we shall have $\phi'' = O[f'(\phi)]$; in this case, by (6) ϕ' is monotonically decreasing without bound for sufficiently large ϕ, and hence will ultimately become negative. As ϕ increases through intervals of the second kind, ϕ'' is negative and of a greater order than $f'(\phi)$: $\phi'' \leqq -\lambda f'(\phi)$, for some $\lambda > 0$. Since $f'(\phi)$ increases without bound with its argument, in intervals of this type ϕ' will also ultimately decrease monotonically and without bound, finally becoming negative. In intervals of the third kind, where $\phi'' \geqq \lambda f'(\phi)$, we discover the same from equation (5), so that we may conclude generally that $\phi(t)$ and $\psi(t)$ are bounded for all time in intervals where $|\phi| \geqq |\psi|$; by applying the same argument to the second equation (4), we establish the conclusion for all time without restriction. This of course does not allow us to exclude singularities in the solution, but since the right of (4) obviously satisfies the Lipshitz condition, that is no problem.

Among the forms of f which this result leaves open to us, we shall select the simplest, a cubic polynomial in which for the sake of symmetry the square term is omitted; we put

$$f(x) = \kappa_1 x + \kappa_3 x^3, \quad \kappa_1, \kappa_3 > 0.$$

We now discuss the character of the history determined by the equations 4. This may be represented by the motion of a particle in a plane, whose abscissa and ordinate at a time are respectively the values of ϕ and of ψ which hold at that time; the problem then becomes formally a dynamical system with two degrees of freedom to determine the path of such a particle. Since the coefficients of ϕ' and ψ' in equations (4) are always positive, the system is dissipative; since it is also bounded, a well-known theorem of dynamics (Birkhoff, 1927, Ch. I) allows us to conclude that it will asymptotically approach some stable point of equilibrium. The equilibria will be the real solution of the algebraic equations obtained by setting ϕ', ψ', ϕ'', ψ'' equal to zero in (4), which are

$$\begin{aligned} \beta\,\phi^2 - \beta\,\varepsilon_1\,\phi\,\psi &= \mu(\kappa_1\phi + \kappa_3\phi^3) \\ \beta\,\phi\,\psi - \beta\,\varepsilon_2\,\phi^2 &= \mu(\kappa_1\psi + \kappa_3\psi^3). \end{aligned} \tag{7}$$

We shall solve them under the assumption that $\varepsilon_1 = \varepsilon_2 = \varepsilon$: the motion does not change its character sharply if this is not precisely true, as an examination of the perturbation of first order will readily convince us; and in any case we expect these parameters to be of the same order of magnitude. For numerical applications, a closer approximation can always be found.

The solutions are readily obtained, and, if we exclude those which are certainly imaginary, they are in general five in number:

1. A root $\phi_1 = \psi_1 = 0$, representing the normal values of the variables;
2. A pair of large equal roots
$$\phi_{21} = \frac{\beta(1-\varepsilon)}{2\,\mu\,\kappa_3} + \frac{1}{2\,\mu\,\kappa_3}\,[\beta^2(1-\varepsilon)^2 - 4\,\mu^2\,\kappa_1\,\kappa_3]^{\frac{1}{2}},$$
$$\psi_{21} = \phi_{21};$$
3. A second such pair:
$$\phi_{22} = \frac{\beta(1-\varepsilon)}{2\,\mu\,\kappa_3} - \frac{1}{2\mu\,\kappa_3}\,[\beta^2(1-\varepsilon)^2 - 4\,\mu^2\,\kappa_1\,\kappa_3]^{\frac{1}{2}},$$
$$\psi_{22} = \phi_{22};$$
4. One pair of large roots, negatives of one another:
$$\phi_{31} = \frac{\beta(1+\varepsilon)}{2\,\mu\,\kappa_3} + \frac{1}{2\mu\,\kappa_3}\,[\beta^2(1+\varepsilon)^2 - 4\,\mu^2\,\kappa_1\,\kappa_3]^{\frac{1}{2}},$$
$$\psi_{31} = -\,\phi_{31};$$
5. A smaller pair of such roots
$$\phi_{32} = \frac{\beta(1+\varepsilon)}{2\,\mu\,\kappa_3} - \frac{1}{2\mu\,\kappa_3}\,[\beta^2(1+\varepsilon)^2 - 4\,\mu^2\,\kappa_1\,\kappa_3]^{\frac{1}{2}},$$
$$\psi_{32} = -\,\phi_{32}.$$

Some or all of these roots, except of course ϕ_1, ψ_1, may happen to be imaginary, so that they do not in fact represent equilibria. This will be the case for the pairs (2) and (3) if the radical in them is negative, a condition which holds if and only if

$$1 - 2\,\sigma \leqq \varepsilon \leqq 1 + 2\,\sigma\,; \qquad (8)$$

where $\sigma = \mu\sqrt{\kappa_1\,\kappa_3}/\beta$.

The equilibria corresponding to the roots (4) and (5) will exist under an inequality for the radical occurring there which may easily be transformed into

$$\begin{aligned} &\varepsilon > 2\,\sigma - 1 \\ &\varepsilon \leqq -\tfrac{1}{3} + \tfrac{2}{3}\sqrt{1 - 3\,\sigma^2}\,. \end{aligned} \qquad (9)$$

For all sufficiently small σ, the inequalities (8) and (9) will obviously leave a range for ε in which all five equilibria exist. In perhaps the most common case, μ and κ_3 will be quite small, and β and κ_1, of the order of unity, so that σ is close to zero, the condition (9) is satisfied by virture of the sign of ε; and (8) requires only that ε should not be in the immediate neighborhood of unity. Still, if μ, κ_1, or κ_3 is quite large or β is small, both inequalities may fail, and all equilibria save the origin cease to exist; this latter is then necessarily stable. An organism whose parameters are so related cannot possibly exhibit disturbed behavior under any circumstances.

It is important to investigate the behavior of the solution in the neighborhood of the points of equilibrium, first, to determine their stability or instability, and second, to throw light on the fluctuations of the organism about the normal level which are insufficient to pass into permanent disturbance. If $\phi = \nu_1$, $\psi = \nu_2$ be a point of equilibrium, and if we set

$$\bar{\phi} = \phi - \nu_1, \quad \bar{\psi} = \psi - \nu_2,$$

we shall derive expressions for $\bar{\phi}$ and $\bar{\psi}$ of the form

$$\begin{aligned} \bar{\phi}(t) &= \sum_{i=1}^{4} \bar{A}_i \, e^{\lambda_i t} \\ \bar{\psi}(t) &= \sum_{i=1}^{4} \bar{A}_i \, e^{\lambda_i t} \end{aligned} \tag{10}$$

where the $\bar{A}_i$ are certain linear combinations of the A_i, these depend on the boundary conditions, and, of the system (4) for this equilibrium, the λ_i are the roots of the characteristic equation which is

$$\begin{vmatrix} \alpha_1 - \lambda & 0 & \gamma_1 & -\beta\,\varepsilon_1\,\nu_1 \\ 0 & \alpha_2 - \lambda & \beta\,\nu_2 - 2\,\beta\,\varepsilon_2\,\nu_1 & \gamma_2 \\ 1 & 0 & -\lambda & 0 \\ 0 & 1 & 0 & -\lambda \end{vmatrix} = 0$$

where

$$\begin{aligned} \alpha_1 &= -\,[\mu + f'(\nu_1)] \\ \alpha_2 &= -\,[\mu + f'(\nu_2)] \\ \gamma_1 &= 2\,\beta\,\nu_1 - \beta\,\varepsilon_1\,\nu_2 - \mu\,f'(\nu_1) \\ \gamma_2 &= \beta\,\nu_1 - \mu\,f'(\nu_2), \end{aligned}$$

provided that all these roots are distinct and that none of them vanishes. This equation may be written in the form

$$\lambda^2(\lambda - \alpha_1)(\lambda - \alpha_2) - \lambda[\gamma_1(\lambda - \alpha_1) + \gamma_2(\lambda - \alpha_2)] + \gamma_1 \gamma_2 + \beta^2 \varepsilon \nu_2(\nu_2 - 2 \varepsilon \nu_1) = 0 . \tag{11}$$

Since $|\nu_1| = |\nu_2|$ for all the points of equilibrium, we have $\alpha_1 = \alpha_2 = \alpha$, and equation (11) is easily soluble, with roots

$$\lambda = \frac{\alpha}{2} \pm \tfrac{1}{2}[\alpha^2 + 2(\gamma_1 + \gamma_2) \pm 2\sqrt{(\gamma_1 + \gamma_2)^2 - 4\chi}]^{\frac{1}{2}}, \tag{12}$$

in which $\chi = \gamma_1 \gamma_2 + \beta^2 \varepsilon \nu_1(\nu_2 - 2 \varepsilon \nu_1)$ and each of the four λ_i is obtained by a different choice for the two doubtful signs in expression (12). Now for the equilibria we are considering to be of stable type, in the sense that no sufficiently small deviation from that point will generate forces which tend to augment it, the condition is that none of the λ_i, or their real parts, if some of them be complex, shall exceed zero. This will be the case here if and only if

$$\alpha = -[\mu + f'(\nu)] < 0 ,$$

which is always true,

$$\gamma_1 + \gamma_2 = 3\beta \nu_1 - 2\mu \kappa_1 - 6\mu \kappa_3 \nu^2_1 < \beta \varepsilon \nu_2 , \tag{13}$$

and

$$\chi = (3\mu \kappa_3 \nu^2_1 + \beta \varepsilon \nu_2 + \mu \kappa_1 - 2\beta \nu_1)(3\mu \kappa_3 \nu^2_2 + \mu \kappa_1 - \beta \nu_2) + \beta^2 \varepsilon \nu_1(\nu_2 - 2\nu_1 \varepsilon) > 0 . \tag{14}$$

For the equilibrium at the origin, these reduce to

$$-2\mu \kappa_1 < 0 , \quad \mu^2 \kappa^2_1 > 0 ,$$

so that the origin is always a stable point, as we should expect. For the other equilibria, the conditions become extremely cumbersome when handled directly; we shall therefore approximate them in the following manner. The quantity $\sigma^2 = \mu^2 \kappa_1 \kappa_3/\beta^2$ may in general be expected to be quite small in comparison with unity; we may therefore expand in the expressions for the points of equilibrium in powers of σ^2, substitute in (13) and (14), and terminate the expansion at the first term that gives a determinate form to the conditions. We discover that for the points ϕ_{21}, ψ_{21} and ϕ_{31}, ψ_{31} we may neglect σ^2 itself, whereas for the smaller equilibria at ϕ_{22}, ψ_{22} and ϕ_{32}, ψ_{32} it is necessary to take account of terms in σ^2; and for one initial value, namely $\varepsilon = \frac{1}{2}$, we must proceed to terms in σ^4. The results are as follows: the condition (13) is satisfied

1) By ϕ_{21} , ψ_{21} unless $3/5 \leqq \varepsilon < 1$;
2) By ϕ_{22} , ψ_{22} unless $3 < \varepsilon \leqq 2 + \sqrt{10}$;
3) By ϕ_{31} , ψ_{31} if $\varepsilon \geqq 0$, i.e., always;
4) By ϕ_{32} , ψ_{32} if $\varepsilon \geqq 0$.

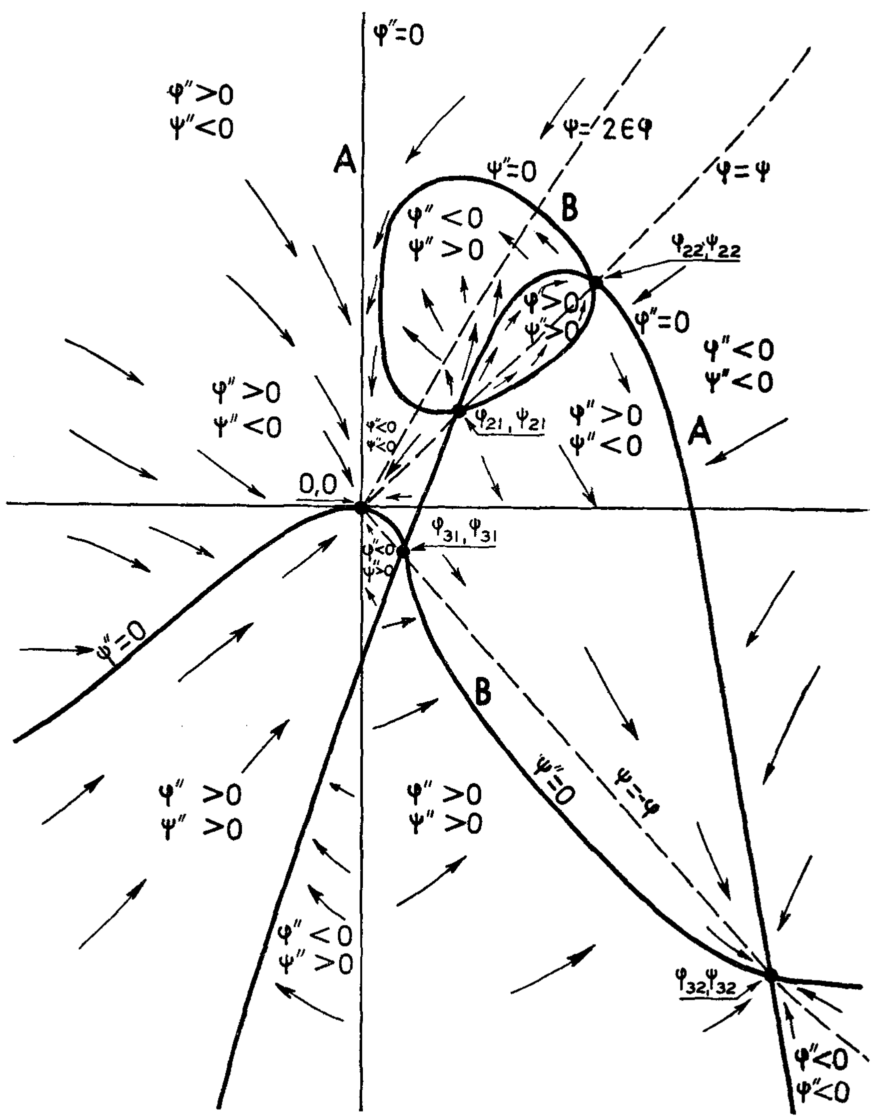

FIGURE 1

The inequality (14) is fulfilled

1) By ϕ_{21}, ψ_{21} unless $\frac{1}{2} < \varepsilon \leqq 1$;
2) By ϕ_{22}, ψ_{22} if $\varepsilon < \frac{1}{2}$;
3) By ϕ_{31}, ψ_{31} if $\varepsilon \geqq 0$;
4) By ϕ_{32}, ψ_{32} unless $\sqrt{7} - 2 < \varepsilon < 1$.

We may obtain a general insight into the qualitative character of the motion determined by our equations from the first figure. The degenerate algebraic curve consisting of a parabola A together with the axis of ψ divides the plane into the regions for which ϕ'' is positive and those for which it is negative, while the closed curve B together with the cubic parabola separate the regions of positive from those of negative ψ''. Clearly, the intersections of these curves will constitute the points of equilibrium. In each of these regions, we now have arrows, whose slope is the value of ψ''/ϕ'', which represent roughly the direction of motion of a particle placed there with no kinetic energy. The presence of kinetic energy and the dissipative forces will modify these considerations somewhat, and the initial conditions in (4) will superpose upon these forces an initial velocity with components $-f(\phi)$, $-f(\psi)$ toward the origin.

The difference in character between the stable and the unstable equilibria becomes very clear in the diagram. As remarked above, these are no periodic orbits in the large, and unless continually disturbed, the particle will ultimately settle toward one of the equilibria.

This work was aided in part by a grant from the Dr. Wallace C. and Clara A. Abbott Memorial Fund of the University of Chicago.

LITERATURE

Birkhoff, G. 1927. *Dynamical Systems.* New York: American Mathematical Society Colloquium Publication No. 9.