REMARKS ON SINGULAR POINTS OF FUNCTIONAL EQUATIONS[1]

BY

LAWRENCE M. GRAVES

Certain equations in function space, including boundary value problems for nonlinear differential equations involving a parameter, may have multiple solutions near a given one. The problem of counting the number of solutions near a given solution has been treated by several writers (see, for example, Schmidt [7], Iglisch [5; 6], Cronin [2; 3; 4,], Bartle [1]), but no method has been devised which will give a definite count in all cases.

The problems we are discussing may be subsumed under the general one of determining the number of solutions $x = x(y)$ near $x = 0$, $y = 0$, of an equation

$$(1) \qquad G(x, y) = 0,$$

where x and y are variables in two Banach spaces $\mathfrak{X}$ and $\mathfrak{Y}$, when the differential $L(dx) = d_x G(0, 0; dx)$ fails to have an inverse. Cronin's methods use the notion of topological degree, and do not usually give the exact number of distinct solutions. Bartle was able to count the exact number of real solutions in cases when the null space of L is one-dimensional, with the help of Newton's polygon method. By extending the notion of Newton's polygon to equations involving several variables we are able to count the exact number of real solutions when the null space of L is two-dimensional. An example involving an elliptic partial differential equation is actually computed in the final section. The theory is applicable to cases when the null space has dimension number greater than two, but the computations become unmanageable.

1. Summary of basic results. We shall first outline the basic theory, essentially as presented by Bartle, but with special hypotheses and notations adapted to the cases we are able to discuss explicitly later. We write the basic equation (1) in the form

$$(2) \qquad L(x) = F(x, y)$$

and make the following assumptions on L and F:

(H₁) L is a linear continuous transformation of the Banach space $\mathfrak{X}$ into itself, having a closed range;

(H₂) L and its adjoint L^* have n-dimensional null spaces, where $n > 0$;

(H₃) $F(x, y)$ maps a neighborhood of the origin in $\mathfrak{X} \times \mathfrak{Y}$ (where $\mathfrak{Y}$ is also a Banach space) into $\mathfrak{X}$;

(H$_4$) $F(x, y) = A(y) + B(y) + C(x, y) + D(x) + \Gamma(x, y)$, where $A(y)$ is linear and continuous, $B(y)$ is a homogeneous polynomial of the second degree, $C(x, y)$ is bilinear and continuous, $D(x) = D(x_1, x_2)$ where $D(x_1, x_2)$ is bilinear and continuous, and $\Gamma(x, y)$ is continuous and satisfies

$$\|\Gamma(x, y)\| \leqq \omega(\|x\| + \|y\|)^3,$$
$$\|\Gamma(x_1, y) - \Gamma(x_2, y)\| \leqq \omega\|x_1 - x_2\|(\|x_1\| + \|x_2\| + \|y\|),$$

with some constant ω.

If $(u_1, \cdots, u_n)$ and $(f_1, \cdots, f_n)$ are bases for the null spaces of L and L^* respectively, there exist elements $g_1, \cdots, g_n$ in $\mathfrak{X}^*$ and $z_1, \cdots, z_n$ in $\mathfrak{X}$ such that

$$(3) \qquad g_i(u_j) = f_i(z_j) = \delta_{ij},$$

and then the operator

$$(4) \qquad H(x) = L(x) + \sum_1^n g_i(x)z_i$$

is easily seen to map $\mathfrak{X}$ onto $\mathfrak{X}$ one-to-one. If $R = H^{-1}$, then (2) is equivalent to

$$(5) \qquad x = RF(x, y) + \sum g_i(x)u_i,$$

since $H(u_i) = z_i$. This leads us to consider the equation

$$(6) \qquad x = RF(x, y) + u.$$

Under the assumptions made on F, the method of successive substitutions shows that equation (6) has a unique solution $x = V(u, y)$ for (x, u, y) near $(0, 0, 0)$, where u is in the null space of L, and that the function V can be written in the form

$$(7) \quad V(u, y) = u + RA(y) + RB(y) + RC(u, y) + RC(RAy, y) + RD(u)$$
$$+ RD(RA(y)) + 2RD(u, RA(y)) + \Gamma_1(u, y),$$

where Γ_1 has properties like those of Γ. With the help of (3) it is seen that the formula $x = V(u, y)$ gives a solution of (5) if and only if

$$g_i(u) = g_i(V(u, y)),$$

and if we write $u = \sum_1^n b_i u_i$, we find from (7) and (3) that these conditions may be written as follows:

$$g_i RA(y) + g_i[RB(y) + RC(RAy, y) + RD(RA(y))]$$
$$+ g_i \sum_j b_j[RC(u_j, y) + 2RD(u_j, RA(y))]$$
$$(8)$$
$$+ g_i \sum_{j,k} b_j b_k RD(u_j, u_k) + g_i \Gamma_1\left(\sum_j b_j u_j, y\right) = 0.$$

2. **Newton's polygon for several variables.** Actually we restrict attention to one side of Newton's polygon, and we cannot assert that the method applies to all cases, since a restrictive hypothesis must be made. We may restrict the variables to be real, or allow them all to be complex. We use y to denote a single variable, and $b = (b_1, \cdots, b_n)$ to denote a set of n variables, and consider an equation of the form

$$(9) \qquad \phi(b, y) = \sum_{i+j=0}^{m} \phi_{ij}(b) y^i + \rho(b, y) = 0,$$

where each coefficient ϕ_{ij} is a homogeneous polynomial of degree i mapping the n-dimensional b-space into itself, and ρ is continuous and satisfies

$$(10) \qquad \rho(b, y) = o(\|b\|^m + |y|^m),$$
$$\|\rho(b_1, y) - \rho(b_2, y)\| \leqq \omega \|b_1 - b_2\| (\|b_1\| + \|b_2\| + |y|).$$

We assume also that

$$(11) \qquad \phi_{i0}(b) = 0, \qquad\qquad\qquad i = 0, \cdots, k-1,$$
$$\phi_{k0}(b) \neq 0.$$

At times we shall make the stronger requirement that there is a constant $\alpha > 0$ such that

$$(12) \qquad \|\phi_{k0}(b)\| \geqq \alpha \|b\|^k.$$

When $k = 1$, and (12) holds, it follows from classical existence theorems that (9) has a unique solution $b = b(y)$ near $(0, 0)$. Hence we shall assume $k > 1$. On a Cartesian coordinate system we plot the points (i, j) for which ϕ_{ij} does not vanish identically, and we suppose there is such a point with $i < k$. When $i < k$, the line joining (i, j) to $(k, 0)$ has negative slope. Let $-\mu$ denote the maximum such slope, and let $\mu = p/q$ where p and q are relatively prime. Then substitute

$$(13) \qquad b = \beta \eta^p, \qquad y = \eta^q$$

in (9), and divide by η^{pk} to obtain

$$(14) \qquad \psi(\beta, \eta) = \sum_{(i,j) \in L_0} \phi_{ij}(\beta) + \eta^{r_1} \sum_{(i,j) \in L_1} \phi_{ij}(\beta)$$
$$+ \eta^{r_2} \sum_{(i,j) \in L_2} \phi_{ij}(\beta) + \cdots + \rho(\beta \eta^p, \eta^q)/\eta^{pk} = 0$$

where L_0 is the line through the point $(k, 0)$ with slope $-\mu$, L_1 is the closest parallel line containing plotted points, and L_1 lies at a vertical distance r_1/q above L_0, L_2 is the next closest parallel line containing plotted points, and so on. The remainder term $\rho(\beta \eta^p, \eta^q)/\eta^{pk} = o(\eta^S)$ where S is the lesser of $pm - pk$ and $qm - pk$. We assume $m \geqq k$, $m \geqq \mu k$. The exponents $r_1, r_2, \cdots$, clearly turn out to be positive integers. We have the following theorem.

THEOREM. (a) *Every continuous solution $\beta = \beta(\eta)$ of equation (14) yields a continuous solution $b = b(y)$ of (9) by formulas (13).* (b) *When (12) holds and the equation*

$$(15) \qquad \sum_{(i,j) \in L_0} \phi_{ij}(\beta) = 0$$

has only a finite number of solutions, every solution $b = b(y)$ of (9) continuous near $y = 0$ and vanishing with y is obtained by formulas (13) from a solution $\beta = \beta(\eta)$ of (14) continuous near $\eta = 0$. (When the variables are real, the neighborhoods on the y-axis are to be taken as one-sided neighborhoods with $y > 0$, and the criterion for negative values of y is obtained by first replacing y by $-y$ in (9).)

Proof. Part (a) of the theorem is obvious. Now suppose that $b = b(y)$ is a solution of (9) which vanishes with y. We shall show that the corresponding $\beta = \beta(\eta)$ tends to a unique finite limit when η approaches zero. Suppose first that $\beta(\eta)$ is unbounded near $\eta = 0$. Then on a sequence of values of y tending to zero, $\eta^p / \|b\|$ tends to zero, $y^p / \|b\|^q$ tends to zero, and

$$\phi_{ij}(b) y^j = O(\|b\|^i \, |y|^j)$$
$$= o(\|b\|^k)$$

except for $i = k$, $j = 0$, and $\rho(b, y) = o(\|b\|^k)$ by (10) and the assumption $m > k$, $m \geq \mu k$. This leads to a contradiction with (9) and (12), so the set S of limiting values of $\beta(\eta)$ as η approaches zero is compact. It is readily seen that S is connected, and by (14) the points β in S satisfy (15). Hence S reduces to a single point β_0, and if we set $\beta(0) = \beta_0$, $\beta(\eta)$ is continuous near $\eta = 0$. This completes the proof.

We note that when $\psi(\beta_0, 0) = 0$ and the differential $d_\beta \psi(\beta_0, 0; d\beta)$ has an inverse, the implicit function theorem yields a unique solution $\beta(\eta)$ of (14) near β_0.

We now examine in detail the special case in which $n = k = 2$, and the variables are real. We shall denote the two components of ϕ, ψ, ϕ_{ij}, etc., by (ϕ', ϕ''), (ψ', ψ''), etc., but those of b and β will be written (b, c) and (β, γ) respectively. There are two subcases.

CASE I. μ is an integer, so (13) becomes $b = \beta \eta^\mu$, $y = \eta$. The line L_0 contains the point $(2,0)$ and one or both of $(0, 2\mu)$ and $(1, \mu)$, so equation (15) has the form

$$(15') \qquad \phi'_{20}(\beta) + d' + h'\beta + l'\gamma = 0,$$

$$(15'') \qquad \phi''_{20}(\beta) + d'' + h''\beta + l''\gamma = 0,$$

where

$$\phi_{20}'(\beta) = A'\beta^2 + B'\beta\gamma + C'\gamma^2,$$
$$\phi_{20}''(\beta) = A''\beta^2 + B''\beta\gamma + C''\gamma^2,$$
$$(d', d'', h', h'', l', l'') \neq (0, 0, \cdots, 0).$$

CASE II. $\mu = p/2$, where p is odd, so (13) becomes $b = \beta\eta^p$, $y = \eta^2$. The line L_0 contains the points $(2, 0)$ and $(0, p)$ but no others. Equation (15) reduces to (15'), (15''), with $h' = h'' = l' = l'' = 0$, $(d', d'') \neq (0, 0)$.

The graphs of (15') and (15'') are conics in the β-plane, and the assumption (12) is certainly fulfilled if either of these is elliptic. The only case when (12) can fail is when an asymptote of (15'), if (15') is hyperbolic (or an axis of (15'), if (15') is parabolic) is parallel to a corresponding line for (15''). Furthermore, the Jacobian of (15'), (15'') vanishes at a common solution if and only if the graphs are tangent there. Thus by the theorem above and the implicit function theorem, a point β_0 at which (15'), (15'') intersect without being tangent gives a unique neighboring solution $\beta(\eta)$ of (14), to which corresponds a solution $b = b(y)$ of (9) defined at least for small positive y. Also when the condition (12) holds, and (15'), (15'') are not identical ellipses, every solution $b = b(y)$ of (9) determines a solution β_0 of (15'), (15''). Hence the method will give the exact number of solutions of (9) except for certain special cases.

3. An example. Consider the equation

$$(16) \qquad \Delta x + \lambda x + y A_1 + y^2 B_1 + y C_1 x + D_1 x^2 + \Gamma_1(x, y) = 0$$

where A_1, B_1, C_1, and D_1 are continuous functions of (s, t) on the unit square Q, λ and y are real parameters, Δ is the Laplacian, and the solution x is to be continuous on Q, and of class C'' on the interior of Q, and is to vanish on the boundary of Q. We use the norm appropriate to the space of continuous functions on Q, and assume that there is a constant ω such that

$$\|\Gamma_1(x, y)\| \leqq \omega(\|x\| + |y|)^3,$$
$$\|\Gamma_1(x_1, y) - \Gamma_1(x_2, y)\| \leqq \omega\|x_1 - x_2\|(\|x_1\| + \|x_2\| + |y|).$$

If we set

$$u_{ij}(s, t) = 2 \sin \pi \, is \, \sin \pi it,$$
$$N_{ij}(s, t, \sigma, \tau) = u_{ij}(s, t) u_{ij}(\sigma, \tau),$$

the Green's function for the square Q is given by the formula

$$G = \frac{1}{\pi^2} \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \frac{N_{ij}}{i^2 + j^2},$$

and with its help we find that (16) is equivalent to

$$(17) \qquad x - \lambda G(x) = y G(A_1) + y^2 G(B_1) + y G(C_1 x) + G(D_1 x^2) + G(\Gamma_1),$$

where we have written $G(x)$ for $\int_Q G(s, t, \sigma, \tau) x(\sigma, \tau) d\sigma d\tau$, etc. The equation (17) is in the form considered in §1, with $L(x) = x - \lambda G(x)$.

The values of λ for which the null space of L is nontrivial are $\lambda = \pi^2 k$, where k is a sum of two squares, and then an orthonormal basis for the null space is given by the functions u_{ij} with $i^2 + j^2 = k$. Since G is symmetric we may take $f_{ij} = u_{ij}$ as a basis for the null space of L^*, and so $g_{ij} = f_{ij}$, $z_{ij} = u_{ij}$.

Also if we denote the identity by I, we have, corresponding to formula (4) of §1,

$$H = L + \sum_{i^2+j^2=k} N_{ij}$$
$$= I - k \sum_{i^2+j^2 \neq k} N_{ij}/(i^2 + j^2),$$

and the inverse R of H is given by

$$R = I + k \sum_{i^2+j^2 \neq k} N_{ij}/(i^2 + j^2 - k).$$

Since $g_{ij}RG = g_{ij}G = g_{ij}/\pi^2 k$, $i^2+j^2 = k$, the equations (8) become in this case (after multiplying by $\pi^2 k$),

$$y g_{ij}(A_1) + y^2[g_{ij}(B_1) + g_{ij}(C_1 RG(A_1)) + g_{ij}(D_1(RG(A_1))^2)]$$
$$+ y \sum_{\alpha^2+\beta^2=k} b_{\alpha\beta}[g_{ij}(C_1 u_{\alpha\beta}) + 2g_{ij}(D_1 u_{\alpha\beta} RG(A_1))]$$
$$+ \sum_{\alpha^2+\beta^2=k} \sum_{\gamma^2+\epsilon^2=k} b_{\alpha\beta} b_{\gamma\epsilon} g_{ij}(D_1 u_{\alpha\beta} u_{\gamma\epsilon}) + \rho = 0,$$

or, in the notation of equation (9) of §2,

(18) $y\phi_{01} + y^2\phi_{02} + y\phi_{11}(b) + \phi_{20}(b) + \rho = 0.$

By direct computation we obtain the following:

$$g_{ij}(1) = \begin{cases} 8/\pi^2 ij & \text{for } i, j \text{ odd,} \\ 0 & \text{for } i, j \text{ even;} \end{cases}$$

$$g_{ij}(u_{\alpha\beta} u_{\gamma\epsilon}) = \begin{cases} \dfrac{128 ij\alpha\beta\gamma\epsilon}{\pi^2[(i^2 - \alpha^2 - \gamma^2)^2 - 4\alpha^2\gamma^2][(j^2 - \beta^2 - \epsilon^2)^2 - 4\beta^2\epsilon^2]} \\ \qquad\qquad\qquad\qquad \text{for } (i + \alpha + \gamma)(j + \beta + \epsilon) \text{ odd,} \\ 0 \qquad\qquad\qquad\quad \text{for } (i + \alpha + \gamma)(j + \beta + \epsilon) \text{ even;} \end{cases}$$

$$RG(u_{\alpha\beta}) = \begin{cases} \dfrac{u_{\alpha\beta}}{\pi^2(\alpha^2 + \beta^2 - k)} & \text{if } \alpha^2 + \beta^2 \neq k, \\ \dfrac{u_{\alpha\beta}}{\pi^2 k} & \text{if } \alpha^2 + \beta^2 = k. \end{cases}$$

We now proceed to calculate explicitly the form which equation (15)

takes in a few special cases of (16), in order to illustrate the results. In the first two we take $\lambda = 10\pi^2$, so $(i, j) = (1, 3)$ or $(3, 1)$.

CASE I. A_1 and D_1 are constants, with $A_1 D_1 \neq 0$. Then both components of Q_{01} reduce to $8A_1/3\pi^2$, and one component of Q_{20} is

$$\frac{128 D_1}{27\pi^2}\left(X - \frac{9Y}{7}\right)\left(X + \frac{9Y}{25}\right)$$

while the other is

$$\frac{128 D_1}{27\pi^2}\left(Y - \frac{9X}{7}\right)\left(Y + \frac{9X}{25}\right),$$

where $X = b_{13}$, $Y = b_{31}$. Thus $p = 1$, $q = 2$ in (13), and equation (15) is equivalent to

$$\left(X - \frac{9Y}{7}\right)\left(X + \frac{9Y}{25}\right) = -\frac{9A_1}{16D_1},$$

$$\left(Y - \frac{9X}{7}\right)\left(Y + \frac{9X}{25}\right) = -\frac{9A_1}{16D_1}.$$

The asymptotes of these hyperbolas have different directions, so the hypothesis (12) holds, and since the asymptotes separate each other, there are always exactly two intersections. Hence for $\lambda = 10\pi^2$, and y sufficiently small (positive or negative), equation (16) has exactly two real solutions x_1 and x_2 which are near the initial solution $x = 0$.

CASE II. $A_1 = a_1 u_{23}$, and B_1, C_1, D_1, a_1 are constants. Then $\phi_{01} = 0$, the components of ϕ_{20} are the same as in Case I, while the components of ϕ_{02} are

$$8B_1/3\pi^2 + 2^9 a_1^2 D_1/5 \cdot 3^5 \cdot \pi^6,$$

$$8B_1/3\pi^2 + 2^9 a_1^2 D_1/3 \cdot 5 \cdot 7^2 \cdot \pi^6,$$

and those of ϕ_{11} are $C_1 X$ and $C_1 Y$.

If we set $\delta = 3^3 \cdot 5^2 \cdot \pi^2 C_1/2^{12} D_1$, the equations of the hyperbolas may be written

$$\left(X - \frac{9Y}{7} + \delta\right)\left(X + \frac{9Y}{25} + \frac{7\delta}{25}\right) = \frac{7\delta^2}{25} - \frac{9B_1}{16D_1} - \frac{4a_1^2}{45\pi^4},$$

$$\left(Y - \frac{9X}{7} + \delta\right)\left(Y + \frac{9X}{25} + \frac{7\delta}{25}\right) = \frac{7\delta^2}{25} - \frac{9B_1}{16D_1} - \frac{36a_1^2}{245\pi^4}.$$

These hyperbolas may have two, three, or four intersections. In the first case, there are exactly two real solutions of (16) for each small value of y, in the second case there are at least two, and in the third case there are exactly four.

In particular there are four real solutions when both hyperbolas reduce to pairs of lines.

In the next two cases we take $\lambda = 5\pi^2$, so $(i, j) = (1, 2)$ or $(2, 1)$.

CASE III. $A_1 = u_{12}u_{13}$, $D_1 = u_{12}$. Then the components of ϕ_{01} are $2^9/63\pi^2$ and 0, while the components of ϕ_{20} are $9X^2/4 + Y^2$ and $2XY$, with $X = b_{12}$, $Y = b_{21}$, and equation (15) is equivalent to

$$
(19) \qquad
\begin{aligned}
9X^2 + 4Y^2 &= -\, 2^{11}/63\pi^2, \\
2XY &= 0,
\end{aligned}
$$

so (16) has no real solutions for small positive y. But since changing the sign of y changes the sign of the right-hand sides of (19), (16) has four real solutions for small negative y.

CASE IV. $A_1 = u_{21}u_{13}$, $D_1 = u_{12}$. Then the components of ϕ_{01} are 0 and $2^9/225\pi^2$, while the components of ϕ_{20} are the same as before, so (19) is replaced by

$$
\begin{aligned}
9X^2 + 4Y^2 &= 0, \\
2XY &= -\, 2^9/225\pi^2,
\end{aligned}
$$

and there are no real solutions of (16) for y near zero.

BIBLIOGRAPHY

1. R. G. Bartle, *Singular points of functional equations*, Trans. Amer. Math. Soc. vol. 75 (1953) pp. 366–384.

2. Jane Cronin, *The existence of multiple solutions of elliptic differential equations*, Trans. Amer. Math. Soc. vol. 68 (1950) pp. 105–131.

3. ———, *Branch points of solutions of equations in Banach space*, Trans. Amer. Math. Soc. vol. 69 (1950) pp. 208–231.

4. ———, *Analytic functional mappings*, Ann. of Math. vol. 58 (1953) pp. 175–181.

5. R. Iglisch, *Reelle Lösungsfelder der elliptischen Differentialgleichungen $\Delta u = F(u)$ und nichtlinearer Integralgleichungen*, Math. Ann. vol. 101 (1929) pp. 98–119.

6. ———, *Über die Vielfachheit einer Lösung in der nichtlinearer Integralgleichungen von E. Schmidt*, Jber. Deutschen Math. Verein vol. 39 (1930) pp. 65–72.

7. E. Schmidt, *Über die Auflösung der nichtlinearer Integralgleichungen und die Verzweigung ihrer Lösungen*, Math. Ann. vol. 65 (1908) pp. 370–399.

UNIVERSITY OF CHICAGO,
    CHICAGO, ILL.