SOLUTION OF THE EQUATION $ze^z = a$

BY E. M. WRIGHT

Communicated by Richard Bellman, December 15, 1958

The roots of the equation $ze^z = a$ $(a \neq 0)$ play a role in the iteration of the exponential function [2; 3; 11] and in the solution and application of certain difference-differential equations [1; 9; 10; 12]. For this reason, several authors [4; 5; 7; 8; 9; 12] have found various properties of some or all of the roots. Here we "solve" the equation in the following sense. We list the roots Z_n, where n takes all integral values, and define Z_n precisely for each n. We give a rapidly convergent series for Z_n for all n such that $|n| > n_0(a)$; the first few terms provide a very good approximation to Z_n. In general, n_0 is fairly small. Finally we show how to calculate each of the remaining Z_n $(-n_0 \leqq n \leqq n_0)$ numerically by giving a variety of methods to find a first approximation to Z_n and showing how to improve this to any required degree of accuracy.

We cut the complex z-plane along the negative half of the real axis and take $|\arg z| \leqq \pi$ in the cut-plane. If we put $w = z + \log z$, we have $dw/dz = (z+1)/z$ and there is a branch-point at $z = -1$. The cuts in the w-plane are the two semi-infinite lines on which $w = u \pm \pi i$, $u \leqq -1$. It can be proved that there is a one-to-one correspondence between the points of the z-plane and those of the w-plane, excluding the cuts in each case, so that the function $z(w)$ is uniquely defined in the cut w-plane.

We write $A = |a|$, take $\log A$ real and $\log a = \log A + i\alpha$, where $-\pi < \alpha \leqq i\pi$. All the roots of our equation are given by $Z_n = z(\log a + 2n\pi i)$, where n takes all integral values. Z_n is thus precisely defined except when $\alpha = \pi$ and $\log A \leqq -1$, (i.e. when a is real and $-e^{-1} \leqq a < 0$). In this one case, $\log a$ and $\log a - 2\pi i$ lie one on each of the two cuts in the w-plane; $z(\log a)$ has two real values, one less than -1 and one between -1 and 0, while $z(\log a - 2\pi i)$ has the same two values. If $-e^{-1} < a < 0$, we define Z_{-1} and Z_0 to be these two real values, distinguishing them arbitrarily by $Z_{-1} < -1 < Z_0 < 0$. If $a = -e^{-1}$, the equation (1) has a double root at $z = -1$ and we put $Z_{-1} = Z_0 = -1$. In addition, when a is real and positive, Z_0 is real. There are no other real roots for any a.

For every nonreal root Z_n, we write $Z_n = X_n + iY_n$. It is easily proved that Y_0 lies between 0 and α, that

$$(2n - 1)\pi + \alpha < Y_n < 2n\pi + \alpha \qquad (n \geqq 1)$$

and that

$$2n\pi + \alpha < Y_n < (2n+1)\pi + \alpha \qquad (n \leqq -1).$$

We define the sequence of polynomials $P_m(t)$ by

$$P_1(t) = t, \qquad P_{m+1}(t) = P_m(t) + m\int_0^t P_m(\sigma)d\sigma.$$

In particular,

$$P_2 = t + \frac{1}{2}t^2, \qquad P_3 = t + \frac{3}{2}t^2 + \frac{1}{3}t^3,$$

$$P_4 = t + 3t^2 + \frac{11}{6}t^3 + \frac{1}{4}t^4, \quad P_5 = t + 5t^2 + \frac{25}{6}t^3 + \frac{25}{12}t^4 + \frac{1}{5}t^5.$$

For every sufficiently large positive n, we write $H = 2n\pi + \alpha - \pi/2$, $\beta = \log(A/H)$ and

$$\eta = \sum_{j=0}^{\infty} (-1)^j P_{2j+1}(\beta) H^{-2j-1}. \tag{1}$$

We can show then that

$$Y_n = H + \eta, \qquad X_n = (H+\eta)\tan\eta$$

or, if we wish to calculate X_n only without first calculating η, we may use the series

$$X_n = \beta + \sum_{j=1}^{\infty} (-1)^j P_{2j}(\beta) H^{-2j}. \tag{2}$$

To obtain these expansions we take iH as a first approximation to Z_n and note that $\beta = w(Z_n) - w(iH)$. Hence, by the Taylor's series for $z(w)$, we have

$$Z_n = iH + \sum_{m=1}^{\infty} \beta^m [d^m z/dw^m]_{z=iH}. \tag{3}$$

Some manipulation enables us to deduce (1) and (2). We can show that the series in (1) and (2) are convergent and the results valid if

$$2H|\beta| < (H-1)^2,$$

$$(\log A)^2 < \left(H - \frac{1}{2}\pi\right)^2 + 2(1 + \log A)\log H + 1$$

are both satisfied, which they clearly are for large enough n.

To calculate η from a reasonable number of terms of (1), we must have β/H fairly small. We observe that the series in (1) and (2) have real terms, a matter of importance for numerical calculation.

For n negative, we write $H=-2n\pi-\pi/2-\alpha>0, \beta=\log(A/H)$ and define η by (1). We have then

$$Y_n = -H-\eta, \qquad X_n = (H+\eta)\tan\eta$$

and (2) is still true.

There will remain a few values of n for which the series (1) and (2) diverge or converge too slowly to provide a convenient means of calculating Z_n. For such an n, we have to calculate $z(w)$, where $w=\log A+(2n\pi+\alpha)i$. Now

$$\frac{dw}{dz}=\frac{z+1}{z},$$

$$\frac{dz}{dw}=\frac{z}{z+1},$$

$$\frac{d^2z}{dw^2}=\frac{dz}{dw}\,\frac{d}{dz}\left(\frac{z}{z+1}\right)=\frac{z}{(z+1)^3}.$$

Hence, if δz, δw denote corresponding small changes in z and w, we have

$$\delta z = z(z+1)^{-1}\delta w + O\{z(z+1)^{-3}(\delta w)^2\}. \tag{4}$$

Thus, if we have a first approximation z_0 to z, we calculate $w_0=w(z_0)$ and take $\delta w=w-w_0$. We then apply the correction $\delta z=z_0(z_0+1)^{-1}\delta w$ to z_0 to obtain z_1 (say). If we write $w=u+iv$ and $z=x+iy$, we may calculate $\lambda=\{(x_0+1)^2+y_0^2\}^{-1}$ and use the correction in the real form

$$\begin{aligned} \delta x &= \{1-\lambda(x_0+1)\}\delta u - y_0\lambda\delta v,\\ \delta y &= y_0\lambda\delta u + \{1-\lambda(x_0+1)\}\delta v. \end{aligned} \tag{5}$$

Next we calculate $w_1=w(z_1)$ and, if this still differs appreciably from w, we repeat the process. It is usually possible to use the same coefficients of δu, δv in (5) at each step. Provided z_0 is not near -1, the process converges fairly rapidly by (4).

But z_0 is near -1 if and only if w is near $-1\pm\pi i$. Let us suppose, for example, that w lies near $-1+\pi i$, so that z must be near -1. We can show that

$$z = -1+\sum_{m=1}^{\infty} c_m\omega^m, \tag{6}$$

where

$$c_1 = -3c_2 = 36c_3 = 270c_4 = 4320c_5 = -17010c_6 = 1,$$

$$c_7 = -\frac{139}{5443200}, \qquad c_8 = -\frac{1}{204120}, \qquad c_9 = -\frac{571}{2351462400},$$

$$c_m = -c_{m-1}(m+1)^{-1} - \frac{1}{2}\sum_{h=2}^{m-1} c_h c_{m-h+1} \qquad (m \geqq 3)$$

and $\omega = i2^{1/2}(w+1-\pi i)^{1/2}$. If w lies on the lower edge of the cut in the w-plane ending at $-1+\pi i$, we take ω real and positive; if w does not lie on this cut, we take $\mathcal{I}(\omega)>0$. The radius of convergence of the series in (6) is $2\pi^{\frac{1}{2}}$. If w lies near $-1-\pi i$, the same series gives us $z(w)$, but $\omega = i2^{1/2}(w+1+\pi i)^{1/2}$ and $\mathcal{I}(\omega)<0$, unless ω is real. Thus if (say) $\log a$ is near $-1+\pi i$, (6) enables us to calculate Z_0 and Z_{-1}.

If w lies between the cuts, i.e. if $u<-1$, we have (see [6], for example)

$$z = \sum_{m=1}^{\infty} (-1)^{m-1} m^{m-1} (m!)^{-1} e^{mw}. \tag{7}$$

For $u \leqq -2$, the first few terms give the value of z with sufficient accuracy. This gives us Z_0 when $|a| \leqq e^{-2}$.

Even if the series (1) does not converge sufficiently rapidly to be useful to calculate Z_n, the first one or two terms may provide a sufficient approximation to enable us to apply our correction procedure. (A similar remark applies to (7) and even to (6).)

If $|w|>4$ and w does not lie between the cuts in the w-plane, a useful value for z_0 is $w-\log w$, where $\log w$ has its principal value. The next approximation z_1 will be accurate to at least one decimal place and further approximations converge rapidly. For $|w| \leqq 4$, we have constructed a table of $w(z)$, which gives a satisfactory value of z_0 by inspection, except near $z=-1$.

Alternatively drawing can be used to obtain the first approximation. Given u, v, we have to solve

$$x + \log r = u, \qquad y + \theta = v, \tag{8}$$

where $r^2 = x^2+y^2$, $\tan\theta = y/x$. To solve these equations graphically, we use (i) a sheet of paper, the (x, y) plane, carrying circles $r=k$ and radii $\theta=h$ for various values of k and h, and (ii) a sheet of tracing paper, the (X, Y) plane, on which the lines $X=-\log k$ and $Y=-h$ are drawn. We place the origin of the (X, Y) plane at the point (u, v) on the (x, y) plane, make the corresponding axes parallel and then plot on a second sheet of tracing paper (the second (x, y) plane) placed over the first the intersections of $X=-\log k$ with $r=k$ and that of

$Y=-h$ with $\theta=h$. Through these two sets of points can be drawn the two curves (8) and their intersection in the (x, y) plane gives the required solution.

REFERENCES

1. R. Bellman, Ann. of Math. vol. 50 (1949) pp. 347–355.
2. G. Eisenstein, J. Reine Angew. Math. vol. 28 (1844) pp. 49–52.
3. L. Euler, *Opera Omnia* (i) vol. 15, Leipzig and Berne, 1927, pp. 268–297.
4. N. D. Hayes, J. London Math. Soc. vol. 25 (1950) pp. 226–232.
5. ———, Quart. J. Math. Oxford Ser. (2) vol. 3 (1952) pp. 81–90.
6. A. Hurwitz and R. Courant, *Funktionentheorie*, 3d ed., Berlin, 1929, pp. 141–142.
7. E. M. Lémeray, Nouv. Ann. de Math. (3) vol. 15 (1896) pp. 548–556 and vol. 16 (1896) pp. 54–61.
8. ———, Proc. Edinburgh Math. Soc. vol. 16 (1897) pp. 13–35.
9. O. Polossuchin, Thesis, Zurich, 1910.
10. F. Schürer, Ber. Verh. Sachs. Akad. Wiss. Leipzig Math.-phys. Kl. vol. 64 (1912) pp. 167–236 and vol. 65 (1913) pp. 239–246.
11. E. M. Wright, Quart. J. Math. Oxford Ser. (2) vol. 18 (1947) pp. 228–235.
12. ———, J. Reine Angew. Math. vol. 194 (1955) pp. 66–87, esp. pp. 72–74.

UNIVERSITY OF ABERDEEN, SCOTLAND