A METHOD FOR ACCELERATING THE CONVERGENCE IN THE PROCESS OF ITERATION*

BY C. C. CAMP

1. *Introduction.* The simplicity and directness of the process of iteration coupled with the fact that errors committed along the way do not vitiate the final result have won for the method a certain degree of widespread popularity. This has become imperiled, however, by the extreme slowness of the convergence in many cases. The object of the present paper is to overcome the difficulty by furnishing a powerful method based on the same kind of analysis as Newton's but better adapted to the process of iteration. The common practice of taking half the sum of two successive approximations, one of which is too large and the other too small, is seen to be inadequate by the following example: $x=2+\pi \sin x$. Here one iterates using the formula $\sin x_2=(x_1-2)/\pi$. For x_1 close to the true root $m_1\equiv dx_2/dx_1=1/(\pi \cos x_2)=-.331$ nearly. The half sum is no improvement here and one finds the same value for the ninth application of the formula as by unmodified iteration, provided one starts with x_1 as the equivalent of 164° in radians. This value is obtained by two applications of the new method. Its equivalent in degrees is 164.05131, and is correct to five decimals.

By ordinary iteration it requires about a score of approximations to solve $6k+10e^{-k}=10$† correct to six decimals. The result, $k=1.126261$, is found by two applications of the new method by starting with $k_1=1.1$, obtained graphically. Here $m_1=.548$. One can use the half sum only when m is negative; and if m is near $-1/3$, no improvement results. The present method consists of one iteration and one application of formula (8), which may be regarded as a formula of interpolation

* Presented to the Society, April 2, 1926.

† See Phillips, *Differential Equations*, p. 15, ex. 5.

or of extrapolation, according to the sign of m. In §2 the method is derived analytically and interpreted geometrically. The rapidity of the convergence is considered in §3, and the last three sections give modifications and extensions of the new method. It is obvious that one may also extend it so as to apply to functions which are merely tabulated by replacing m by a properly corrected first difference.

2. *Derivation and Interpretation of the New Method.* Consider the equation

$$f(x) = 0, \tag{1}$$

in which f is a real single-valued function of a real variable x, continuous together with its first and second derivatives in a properly chosen neighborhood of each root of (1). Let f be further restricted so that $f'(x_0) \neq 0$, $f''(x_0) \neq 0$, when $f(x_0) = 0$. This implies that f possesses only a finite number of maxima and minima in the neighborhood of a zero. Let f, or a modification of f possessing the same desired zero, be decomposed into $f_1(x) - f_2(x)$, so that

$$|f_1'(x_i)| > |f_2'(x_j)| \tag{2}$$

in one of these neighborhoods

$$x_0 - \delta \leqq x \leqq x_0 + \delta, \tag{3}$$

where δ is sufficiently small. That this is always possible is obvious from the following choice

$$f_1(x) \equiv x, \qquad f_2(x) \equiv x - \frac{f(x)}{M(x)}, \tag{4}$$

where $M(x)$ is a function which equals $f'(x)$ at a root of (1) and possesses a continuous second derivative. By continuity $M(x)$ will not vanish in some neighborhood (3). If $|f_1'(x_0)| > |f_2'(x_0)|$, then by continuity there will be a neighborhood about x_0 in which the minimum of the first is greater than the maximum of the second. Condition (2) will therefore be satisfied if it holds for the one point $x_0 = x_i = x_j$. This is evidently true for the choice (4).

By Maclaurin's Theorem we have

$$(5)\qquad \begin{cases} f_1(x) = f_1(x_2) + f_1'(x_2)\Delta x_2 + f_1''(x_2+\theta_2\Delta x_2)\dfrac{\Delta x_2^2}{2}, \\[2ex] f_2(x) = f_2(x_1) + f_2'(x_1)\Delta x_1 + f_2''(x_1+\theta_1\Delta x_1)\dfrac{\Delta x_1^2}{2}, \end{cases}$$

where $\Delta x_1 = x - x_1$, $\Delta x_2 = x - x_2$, $0<\theta_1<1$, $0<\theta_2<1$.

If we neglect infinitesimals of the first order we obtain

$$(6)\qquad f_1(x_2) = f_2(x_1),$$

the approximate relation used in the process of iteration. If instead of (5) we use the Law of the Mean in connection with (6), we get

$$(7)\qquad x - x_2 = \frac{f_2'(x_1+\theta_1'\Delta x_1)}{f_1'(x_2+\theta_2'\Delta x_2)}(x - x_1),$$

where $0<\theta_1'<1$, $0<\theta_2'<1$. By assumption (2) this shows that the error in the second approximation is less than that in the first, provided x_1 lies in the neighborhood (3). Clearly x_2 will then lie in the same interval.

If we neglect only infinitesimals of order two we have from (5) and (6), upon solving for x and denoting this new approximation by $\bar{x}_2$,

$$(8)\qquad \bar{x}_2 = x_1 + \frac{x_2 - x_1}{1 - m},$$

where

$$(9)\qquad m = f_2'(x_1)/f_1'(x_2).$$

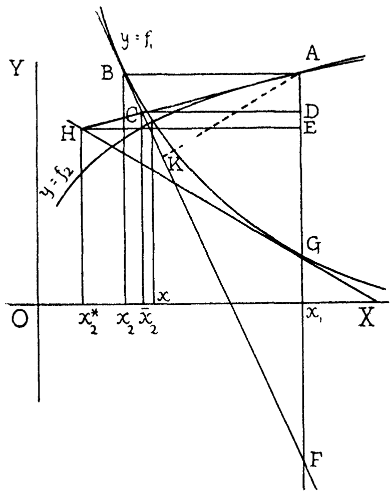
If we use (5) unchanged in connection with (6) and (8) we get

$$(10)\quad x - \bar{x}_2 = \frac{\frac{1}{2}f_1''(x_2+\theta_2\Delta x_2)\Delta x_2^2 - \frac{1}{2}f_2''(x_1+\theta_1\Delta x_1)\Delta x_1^2}{f_1'(x_2) - f_2'(x_1)}.$$

By calling $\bar{m}$ the coefficient of $x - x_1$ in (7) one may put this in the form

$$(11)\qquad x - \bar{x}_2 = \tfrac{1}{2}\Delta x_1^2\left[\frac{\bar{m}^2 f_1''(x_2') - f_2''(x_1')}{f_1'(x_2) - f_2'(x_1)}\right],$$

where x_2' and x_1' are arguments in the numerator of (10). Since the quantities in brackets are bounded above and below, it is clear that the approximation given by formula (8) will be closer than both x_1 and x_2 when Δx_1 is sufficiently small.



If we clear (8) of fractions and simplify we obtain

$$(x_1 - \bar{x}_2)[f_1'(x_2) - f_2'(x_1)] = f_1'(x_2)(x_1 - x_2),$$

or, in the figure, $DF - DA = AF$.

Newton's method may be illustrated by using the tangents at A and G. His form gives

$$(x_1 - x_2^{*})[f_1'(x_1) - f_2'(x_1)] = f_1(x_1) - f_2(x_1),$$

or $EG - EA = AG$.

The new method may be modified by drawing at A a secant line with slope equal to $f_2'(x_2)$ intersecting BF at K. The value of m is thus replaced by $f_2'(x_2)/f_1'(x_2)$ and the interpretation gives $(x_1 - \underline{x})[f_{21}'(x_2) - f_2'(x_2)] = f_1'(x_2)(x_1 - x_2)$ or $NF - NA = AF$, where $\underline{x}_2$ is the abscissa of K, and N is the intersection of AF and a horizontal line through K.

3. *Rapidity of the Convergence of Ordinary Iteration and of the New Method.* For ordinary iteration one has from (7), $\Delta x_2 = \bar{m}\Delta x_1$. An upper bound for $|\bar{m}|$ may be found by replacing the numerator by an upper bound $|\bar{f}_2'|$ and the denominator by a lower bound $|\underline{f}_1'|$ in the interval $x_0 - |\Delta x_1| \leqq x \leqq x_0 + |\Delta x_1|$. If the resulting quotient is called M, one has a method of determining an upper bound for the number of successive approximations required for attaining a given degree of accuracy by pure iteration. For instance, if the error allowed is 10^{-k}, since obviously $|\Delta x_3| \leqq M^2 |\Delta x_1|$ and in general $|\Delta x_i| \leqq M^{i-1} |\Delta x_1|$, we need simply to determine $i-1$

from the relation $M^{i-1}|\Delta x_1| = 10^{-k}$. Hence the number sought is

$$(12) \qquad n = i - 1 = (-k - \log_{10}|\Delta x_1|)/\log_{10} M.$$

Likewise one may obtain a lower bound for the number by replacing M by a lower bound for $|\bar{m}|$. $|\Delta x_1|$ may be replaced by a lower bound in the latter case and by an upper bound in (12).

From (11) one can estimate the errors made by using (8) in place of (5). Since the quantities in brackets are bounded above and below, evidently

$$(13) \qquad |x - \bar{x}_2| < k\Delta x_1^2,$$

where k is a constant.

By repeated applications the sequence of values obtained from

$$(14) \qquad \begin{cases} \bar{x}_k = x_{k-1} + \dfrac{x_k - x_{k-1}}{1 - m}, \\ f_1(x_k) = f_2(x_{k-1}), \qquad (k = 2, 3, 4, \cdots) \end{cases}$$

where $m = f_2'(x_{k-1})/f_1'(x_k)$, will have errors which are not greater than

$$\rho\epsilon^2, \ \rho\epsilon^4, \ \rho\epsilon^8, \ \cdots, \quad \text{where} \quad \rho\epsilon = |\Delta x_1|, \quad \rho = \frac{1}{k}.$$

In Newton's method one has similarly

$$(15) \quad x - x_2^{+} = \tfrac{1}{2}\Delta x_1^2 \left[\frac{f_1''(x_1 + \theta_3 \Delta x_1) - f_2''(x_1 + \theta_4 \Delta x_1)}{f_1'(x_1) - f_2'(x_1)} \right],$$

$$0 < \theta_3 < 1, \qquad 0 < \theta_4 < 1.$$

The two methods coincide when $f_1(x)$ is a linear function. But the new method is more powerful than Newton's in the following extensive cases:

Case I. When $f_1(x)$ is non-linear, $f_1''(x)$ is of constant sign and $f_2(x)$ is linear in the interval

$$(16) \qquad x_0 - |\Delta x_1| \leqq x \leqq x_0 + |\Delta x_1|;$$

Case II. When $f_1''(x)$, $f_2''(x)$ are of opposite sign and $f_1'(x), f_2'(x)$ are each of constant sign in some interval (16).

We choose Δx_1 small enough so that an application of Newton's method will give $\overset{+}{x_2}$ a better value than x_1. It is sufficient* if x_1 is so chosen that

$$\left|\frac{f_1(x_1) - f_2(x_1)}{f_1'(x_1) - f_2'(x_1)}\right| < \frac{A}{2B},$$

where A is the minimum of $|f_1'(x)-f_2'(x)|$ and B is the maximum of $|f_1''(x)-f_2''(x)|$ in the interval between x_1 and

$$x_1 - \frac{2[f_1(x_1) - f_2(x_1)]}{f_1'(x_1) - f_2'(x_1)}.$$

Both cases may be proved geometrically. In Case I, since $|f_1'(x)| > |f_2'(x)|$, evidently f_1 will be monotonic and have a continuously turning tangent in view of the fact that f_1'' is of constant sign. The same is true for $f_1(x)$ and $f_2(x)$ in Case II. In either case let these curves intersect at C whose abscissa is x_0 in the restricted interval. Draw tangents at A on f_1 and at B on f_2 for $x=x_1$. Call the intersection of the tangents D. A horizontal line through B cuts f_1 at E, at which a tangent is drawn intersecting BD at F. The tangent at any point of arc AC will cut BD in a point of the line segment DG where G is the point on BD with the same abscissa as C. The abscissa $\bar{x}_2$ of F will therefore be between $\overset{+}{x}$ and x_0 the true value of the root.

As an illustration of Case I let us take the famous Wallis equation, namely $x^3=2x+5$. Here if $x_1=2$, $m_1=.1541$. By Barlow's Tables $1/(1-m_1)=1.182$, $x_2=2.08$, $\bar{x}_2=2.09456$. By iteration, $x_3=2.094552776$, and $\bar{x}_3=2.094551483$, where $1/(1-m_2)=1.179$. The result of two applications of the new method thus furnishes nine digits correct.† The same number is furnished by Newton's method with three applications. Iteration requires six operations to get six digits correct.

* Cf. Cauchy, Œuvres Complètes, (2), vol 4 (1899), p. 576, Theorem III.

† Cf. Whittaker and Robinson, *The Calculus of Observations*, p. 86.

4. *Modifications of the Method.* Just as one may modify Newton's method by various substitutions for the derivative, so it is expedient at times to change the value of m. One may determine approximately by sketching the curves whether a secant line drawn at $x=x_1$ on f_2 with a slope $f_2'(x_2)$ will tend to accelerate the convergence still further in the early terms of the sequence. With certain combinations of f_1 and f_2 this also simplifies the form of m and is very effective in diminishing the arithmetic. Formula (9) then becomes

$$m = f_2'(x_2)/f_1'(x_2), \tag{17}$$

and both advantages are shown in the following example: $4x^2=x^3+5$. Here $m=3x_2/8$. If we take $x_1=1.4$ then $x_2=1.391402$, and by the modified form $\underline{x}_2=1.38202$ whereas $\bar{x}_2=1.3817745$ by (8), (9). Then by iteration from $\underline{x}_2$ we have $x_3=1.381994$ and by (17) $\underline{x}_3=1.38196603$. This is correct to eight figures.

One may make small errors either accidentally, by choosing simple values for $1/(1-m)$ or by using the same value of m more than once. The final result is correct as illustrated by the problem, $y=.5-\log_{10} y$. By inspection the initial value found from a table was $y_1=.6675$; m was taken as $-.65$ and by error y_2 was found to be .6756. Accordingly $\bar{y}_2=y_1+.606\times.0081=.6724$. By (6) $y_3=.6723723$ and by using 606 again $\bar{y}_3=.6723832$ which checks to seven digits.

If the f's have simple derivatives up to the third order, it may be desirable to use a formula which neglects infinitesimals of only the third order or higher. One needs merely to extend equations (5) to one more term each and then keep terms of the second order. It is easy to derive the new approximation by using (6) and solving a quadratic in x. If we designate the new approximation by $\tilde{x}_2$, we have

$$\begin{aligned}\tilde{x}_2(f_1'' - f_2'') = f_2' - f_1' + x_2 f_1'' - x_1 f_2'' \\ + [(f_1' - f_2')^2 + 2(x_2 - x_1)Q]^{1/2},\end{aligned} \tag{18}$$

where $Q\equiv(f_2'f_1''-f_1'f_2'')+\frac{1}{2}(x_2-x_1)f_1''f_2''$, where f_1', f_1'' have the argument x_2, and where f_2', f_2'' have the argument x_1. One application each solves $4x^2 = x^3 + 5$ and the Wallis equation with errors beyond the sixth decimal place, provided the first approximations are 1.4 and 2, respectively.

5. *Extensions to Systems of Equations.* The method may be extended to a system of n equations in n unknowns. Let us consider for simplicity the case $n = 2$, and let the equations to be solved be

$$f(x, y) = 0, \qquad g(x, y) = 0. \tag{19}$$

Assume that the slopes of these curves at a common point (x_0, y_0) are different and that in the neighborhood of each intersection the ordinate of each curve represents a single-valued function of x possessing continuous derivatives up to the second order. At a solution (x_0, y_0), the Jacobian does not vanish, i. e.,

$$J = \begin{vmatrix} f_x & g_x \\ f_y & g_y \end{vmatrix} \neq 0. \tag{20}$$

Let f be broken up into $f_1(x, y) - f_2(x, y)$ and g similarly into $g_1(x, y) - g_2(x, y)$ so that at the solution sought

$$\left|\frac{\partial f_1}{\partial y}\right| > \left|\frac{\partial f_2}{\partial y}\right| \quad \text{and} \quad \left|\frac{\partial g_1}{\partial x}\right| > \left|\frac{\partial g_2}{\partial x}\right|, \tag{21}$$

also, if possible so that

$$\begin{cases} |f_x| + |f_{2y}| < r|f_{1y}| \\ |g_y| + |g_{2x}| < r|g_{1x}| \end{cases}, \tag{22}$$

where $0 < r < 1$, and the subscripts x and y denote partial differentiations. Condition (22) is sufficient to insure the convergence of the sequence of iterated values obtained from the system

$$f_1(x_1, y_2) = f_2(x_1, y_1), \tag{23}$$

$$g_1(x_2, y_1) = g_2(x_1, y_1), \tag{24}$$

since on account of continuity these inequalities will still hold if we keep x and y in sufficiently small intervals about x_0 and

y_0, respectively. To prove this it is sufficient to expand the f's and g's of x and y by Taylor's theorem about the points indicated by the arguments in (23), (24) and then employ these equations. The errors in the sequences of values for x and y will then have zero for their limit.

By extending the expansions to one more term each and then retaining only infinitesimals of order two, one obtains readily the analog of (8), namely

$$
(25) \qquad \begin{cases} \bar{x}_2 = x_1 + \dfrac{1}{D} \begin{vmatrix} (y_2 - y_1)f_{1y}, & f_{1y} - f_{2y} \\ (x_2 - x_1)g_{1x}, & g_{1y} - g_{2y} \end{vmatrix} \\[2ex] \bar{y}_2 = y_1 + \dfrac{1}{D} \begin{vmatrix} f_{1x} - f_{2x}, & (y_2 - y_1)f_{1y} \\ g_{1x} - g_{2x}, & (x_2 - x_1)g_{1x} \end{vmatrix} \end{cases}
$$

where D is the form taken by J when f, g are separated into parts and the arguments are changed to agree with (23), (24). These same arguments are to be understood in the partial derivatives of (25). It is clear that D will also be different from zero for a sufficiently small neighborhood of (x_0, y_0). It can be seen that if (x_1, y_1) is in such a neighborhood then (x_2, y_2) will be also, on account of (22). When $|x_0 - x_1|$, $|y_0 - y_1|$ are sufficiently small, $(\bar{x}_2, \bar{y}_2)$ will also be inside this neighborhood and the sequence of values (x_k, y_k), $(\bar{x}_k, \bar{y}_k)$ found by repeating the iterations

$$
(26) \qquad \begin{cases} f_1(x_k, y_{k+1}) = f_2(x_k, y_k), \\ g_1(x_{k+1}, y_k) = g_2(x_k, y_k), \end{cases}
$$

and applying (25) successively will converge to the solution (x_0, y_0).

It is obvious how the formulas would be written for the case of a greater number of variables n.

In case the inequalities (22) do not hold for a given system it is easy to show that (19) can always be replaced by a new system

$$
(19') \qquad \begin{cases} F(x, y) = 0, \\ G(x, y) = 0, \end{cases}
$$

which will satisfy (21), (22), when rewritten in the form $F \equiv F_1 - F_2$, $G \equiv G_1 - G_2$. The new system will have the same solution as (19) if (20) is satisfied, provided (x, y) is confined to a sufficiently small region about (x_0, y_0). It is sufficient to choose

$$G_1 \equiv x, \quad G_2 \equiv x - \frac{1}{J}\begin{vmatrix} f(x,y), & g(x,y) \\ f_y(x,y), & g_y(x,y) \end{vmatrix}, \tag{27}$$

$$F_1 \equiv y, \quad F_2 \equiv y - \frac{1}{J}\begin{vmatrix} f_x(x,y), & g_x(x,y) \\ f(x,y), & g(x,y) \end{vmatrix}. \tag{28}$$

The amount of arithmetic work may be shorter for this change in the equations to be solved than for more direct methods by which (19) can still be solved without (22) being satisfied.

Geometrically one may solve a set (19) by the theory of the method for one variable, provided $J \neq 0$, i. e., the curves (19) intersect without being tangent to each other. We assume as before that the functions defined by the graphs of (19) possess continuous second derivatives. By a proper choice of f's and g's suppose (21) is satisfied and that at (x_0, y_0) the slope of f is numerically less than that of the curve g. One may proceed as follows:

First Method. After solving by a table, by interpolation, or graphically for an approximate solution (x_1, y_1) of (19), one tests the value of y_1 by trial in (23) and then solves (24) for x_2. Next comes an application of the formula

$$\bar{x}_2 = x_1 + \frac{x_2 - x_1}{1 - m_x}, \tag{29}$$

where $m_x \equiv g_{2x}(x_1, y_1)/g_{1x}(x_2, y_1)$. Equation (23) is then solved in which x_1 is replaced by $\bar{x}_2$ and y_1 by y_1 as corrected above. One then employs the formula

$$\bar{y}_2 = y_1 + \frac{y_2 - y_1}{1 - m_y}, \tag{30}$$

where $m_y \equiv f_{2y}(\bar{x}_2, y_1)/f_{1y}(\bar{x}_2, y_2)$. The process is repeated until a set (x_k, y_k) is the same as (x_{k-1}, y_{k-1}).

Second Method. This is the same before the application of (29), in place of which one uses

$$(31)\quad \bar{x}_2 = x_1 + \frac{x_2 - x_1}{1 - m}, \quad \text{where } m \equiv \frac{\text{slope of } f \text{ at } (x_1, y_1)}{\text{slope of } g \text{ at } (x_2, y_1)};$$

$$(32)\qquad m = \frac{(f_{1x} - f_{2x})(g_{1y} - g_{2y})}{(f_{1y} - f_{2y})(g_{1x} - g_{2x})},$$

wherein the arguments are the same as in (23), (24). The process is repeated as in the first method until the required accuracy is reached. The following example illustrates the method:

$$f(x,y) \equiv 5y^3 + x^2 - 2xy - 4 = 0,$$
$$g(x,y) \equiv x^3 + 2y^2 - 1 = 0.$$

Take $f_1 \equiv 5y_2^3$, $f_2 \equiv 2x_1y_1 - x_1^2 + 4$, $g_1 \equiv 1 - x_2^3$, $g_2 \equiv 2y_1^2$. For a first approximation start with $x_1 = -.65$, $y_1 = .8$. The second of conditions (22) is not satisfied and it is easily seen that (25) will give a diverging sequence. Nevertheless we can still obtain a solution as follows:

The value y_1 improved by trial gives .7977; x_2 from (24) $= -.64844$. Now $m = -.675$, hence from (31) $\bar{x}_2 = -.6491$. From (23) $y_2 = .798235$. In (30) $m_y = -.1358$ and $\bar{y}_2 = .79817$; $x_3 = -.649658$ from (24); and from (31) if we use $m = -.6725$ then $\bar{x}_3 = -.649434$; $y_3 = .798070$, $\bar{y}_3 = .798082$ if we take $m_y = -.136$; $x_4 = -.6494036$, $\bar{x}_4 = -.649416$ where m is not recalculated. Finally $y_4 = .7980875$ and $\bar{y}_4 = .798087$, where m_y is kept as $-.136$. The results are correct to six decimals.

6. *Extension to the Determination of Complex Roots.* Just as Newton's method remains valid, so the method of iteration is applicable to the computation of complex roots. Moreover, the latter method is simpler and in case the modulus of m is sufficiently small it has the usual advantages. However, when the convergence is slow it is imperative to have a method which leads more rapidly to the root. That the present method will accelerate the convergence as in the case of real roots is evident since Taylor's Theorem goes over with a slight change to

expansions in complex numbers. The arithmetic of iteration is easier than for Newton's method and the application of formula (8) is easier after one has found a few approximations than a change to Newton's method at that stage of the calculation.

The method will be illustrated by an example already solved by Cauchy* although he has erroneously stated that his final result, (163), p. 490, is correct up to the seventh decimal. If we write his equation, $e^x - x = 0$, in the form $x_2 = \log x_1$ then near the root sought $|m|$ is near .7275. Hence pure iteration gives a slowly converging sequence, which will approach the root spirally from $x_1 = i$, as follows:

i, $1.5708i$, $.45158+1.5708i$, $.49129+1.29086i$, $.32295+1.20713i$, $.22281+1.3094i$, $.28384+1.40225i$, $.35807+1.3715i$, $\cdots$

If after the third approximation is computed formula (8) is employed, the fourth application of the new method gives the root $\bar{x}_7$ as $.3181315+1.3372357i$, which is less than Cauchy's result by 2 in the seventh decimal of the real part. That the value here given is correct and that Cauchy's is incorrect has been verified by the use of Peter's Zehnstellige Logarithmen, by which the calculation has also been extended to nine decimals. Cauchy used five applications of Newton's method starting with $x_1 = 0$, whereas the new method by the same number of applications furnishes nine decimals as follows:

$$x_1 = i;\ x_2 = 1.5708i,\ \bar{x}_2 = .2854 + 1.2854i;$$
$$x_3 = .2751 + 1.3523i,\ \bar{x}_3 = .3310 + 1.3376i;$$
$$x_4 = .3206 + 1.3282i,\ \bar{x}_4 = .318165 + 1.337276i;$$
$$x_5 = .31816565 + 1.3371892i,\ \bar{x}_5 = .3181315 + .3372358i;$$
$$x_6 = .318131574 + 1.3372357215i,\ \bar{x}_6 = .318131505 + 1.337235701i.$$

THE UNIVERSITY OF ILLINOIS

* Œuvres, IIe Série, Tome IV, p. 485.