On Ternary Continued Fractions,

by

D. N. Lehmer, Berkeley, Cal., U.S.A.

It has been known since Lagrange[1] that the regular continued fraction which represents a quadratic irrationality becomes periodic after a finite number of terms in the expansion, and conversely, that a regular continued fraction which becomes periodic after a finite number of terms is one root of a quadratic equation with rational coefficients. It is therefore useless to look for periodicity in the regular continued fractions which represent cubic and higher irrationalities. To meet this situation Jacobi undertook to extend the continued fraction algorithm as follows:

In the ordinary continued fraction we are concerned with two series of numbers,—the numerators and denominators of the successive convergents—which are given by the recursion formulae:

$$\left.\begin{aligned} A_n &= q_n A_{n-1} + A_{n-2} \\ B_n &= q_n B_{n-1} + B_{n-2} \end{aligned}\right\} \qquad (1)$$

with initial values (1, 0) for A_n and (0, 1) for B_n. In Jacobi's[2] extension of this algorithm we are concerned with three series of numbers given by the recursion formulae:

$$\left.\begin{aligned} A_n &= q_n A_{n-1} + p_n A_{n-2} + A_{n-3} \\ B_n &= q_n B_{n-1} + p_n B_{n-2} + B_{n-3} \\ C_n &= q_n C_{n-1} + p_n C_{n-2} + C_{n-3} \end{aligned}\right\} \qquad (2)$$

with initial values (1, 0, 0) for A_n; (0, 1, 0) for B_n, and (0, 0, 1) for C_n. We shall call the series of numbers (A_n, B_n, C_n) the *convergent sets* and the series of numbers (p_n, q_n) the *partial quotient sets* of a *ternary continued fraction*. Jacobi's problem was to determine if possible an infinite sequence of partial quotient sets which would give a sequence of convergent sets (A_n, B_n, C_n) such that the ratios B_n/A_n and C_n/A_n should converge toward two given irrationalities. He was able to find a periodic series of partial quotient sets which should give certain cubic irrationalities. Bachmann[3] showed later that

(1) Lagrange; Oev. ii, 74.
(2) Jacobi, Werke, vi, 385.
(3) Bachmann, Crelle, Bd. 75, 25 (1873).

for periodicity certain inequality relations must be satisfied by the irrationalities, but whether these are always satisfied is not determined. Berwick[1] has obtained periodic expansions which always avail for cubic irrationalities, but his convergent sets are not given by the recursion formulae indicated above. Others have worked on Jacobi's problem, but the determination of the partial quotient sets which shall ultimately recur, and which shall fit a given pair of cubic irrationalities is a problem which still awaits complete solution.

Some years ago, in 1918 in fact, it occurred to me that the study of the periodic ternary continued fraction apart from the question of fitting it to a given irrationality might yield some results of interest. The study might be called an investigation of the system of difference equations (2) with especial attention to the arithmetical implications. Some of the results were immediately extensible to an n-ary continued fraction easily derived by generalization of the above definition. The chief importance, however, I think of any generalization is its ability to throw light on the ungeneralized field. The extension to quaternary and higher continued fractions is often an amusing game that gives one a virtuous feeling of having done one's whole duty by the subject, but which often adds little more than a complexity difficult to read and of no consequence to the theory of ordinary, or shall we call them, *binary continued fractions*.

The interesting fact developed at once that associated with each *periodic* series of partial quotient sets, there is,—except in certain very interesting special cases,—a definite cubic irrationality. If the p's and q's in the partial quotient sets are related in certain ways the cubic irrationality gives place to a quadratic irrationality or even to a rational field. We shall speak of these cases later.

Using the above recursion formulae with the initial values as indicated one obtains easily $A_1=1$, $A_2=q_2$, $A_3=q_3q_2+p_3$ etc. A convenient computation scheme is:

	A_n	B_n	C_n
	1	0	0
	0	1	0
	0	0	1
p_1q_1	1	p_2	q_1
p_2q_2	q_2	q_2p_1+1	$q_2q_1+p_2$

(1) Berwick, Proc. Lond. Math. Soc. Ser. 2, vol. xii, 1913.

$$\begin{array}{llll} p_3q_3 & q_3q_2+p_3 & q_3q_2p_1+q_3+p_3p_1 & q_3q_2q_1+q_3p_2+q_1p_3+1 \\ \cdots & \cdots & \cdots & \cdots \\ p_nq_n & A_n & B_n & C_n \end{array}$$

It is easily shown that for all values of the subscript n we have:

$$\begin{vmatrix} A_{n-2} & B_{n-2} & C_{n-2} \\ A_{n-1} & B_{n-1} & C_{n-1} \\ A_n & B_n & C_n \end{vmatrix} = 1.$$

It appears that if we assume the original set of recursion relations of the form

$$A_n = q_nA_{n-1} + p_nA_{n-2} + r_nA_{n-3}$$

with similar formulae for the B's and C's then the determinant given in the above theorem is equal to the product of the r's, which fact connects this form of ternary continued fraction with ordinary continued fractions of the type:

$$a_1 + \frac{b_2}{a_2+}\,\frac{b_3}{a_3+}\,\cdots\cdot$$

For students of the theory of numbers such fractions are of little interest. The increase in generality is accompanied with a loss of uniqueness which seems to have serious disadvantages in arithmetical work.

For continued fractions of order m the value of the corresponding determinant is found to be:

$$(-1)^{(n-1)(m-1)}.$$

Thus for fractions of odd order the determinant is found to be $+1$ while for those of even order the determinants are alternately $+1$ and -1.

As soon as we have a method of finding a ternary continued fraction whose nth convergent set (A_n, B_n, C_n) is given, then it is clear that we shall have a solution of the indeterminate equation $A_nX+B_nY+C_nZ=1$. In fact, X, Y and Z may be taken as the co-factors in the nth determinant. This problem was solved in general, and the complete solution of the indeterminate equation in any number of unknowns was obtained entirely free from tentative processes, and without the tedious substitutions usually employed. An account of this solution is to be found in the Proceedings of the National Academy of Sciences for April, 1919.

In building up the theory of ternary and higher continued fractions a theorem which is an extension of the well known theorem in ordinary continued fractions was found very useful. It has to do with the continued fraction obtained by writing down the partial quotient sets of two given fractions. Thus if the partial quotient sets of one fraction are $(p_1, q_1; p_2, q_2; \ldots; p_k, q_k)$ and the corresponding convergent set is $(A_kB_kC_k)$, while the partial quotient sets of another fraction are $(p_1', q_1'; \ldots; p'_{k'}, q'_{k'})$ with corresponding convergent set $(A'_{k'}B'_{k'}C'_{k'})$, then the convergent set of order $k+k'$ for the fraction $(p_1q_1; \ldots; p_kq_k; p_1'q_1' \ldots; p'_{k'}q'_{k'})$ is given by the formulae:

$$A''_{k+k'} = A_kC'_{k'} + A_{k-1}B'_{k'} + A_{k-2}A'_{k'},$$
$$B''_{k+k'} = B_kC'_{k'} + B_{k-1}B'_{k'} + B_{k-2}A'_{k'},$$
$$C''_{k+k'} = C_kC'_{k'} + C_{k-1}B'_{k'} + C_{k-2}A'_{k'},$$

and more generally the last three convergent sets are given by the rule for multiplying determinants from the product:

$$\begin{vmatrix} A_k & A_{k-1} & A_{k-2} \\ B_k & B_{k-1} & B_{k-2} \\ C_k & C_{k-1} & C_{k-2} \end{vmatrix} \times \begin{vmatrix} C'_{k'} & C'_{k'-1} & C'_{k'-2} \\ B'_{k'} & B'_{k'-1} & B'_{k'-2} \\ A'_{k'} & A'_{k'-1} & A'_{k'-2} \end{vmatrix} = \begin{vmatrix} A''_{k+k'} & A''_{k+k'-1} & A''_{k+k'-2} \\ B''_{k+k'} & B''_{k+k'-1} & B''_{k+k'-2} \\ C''_{k+k'} & C''_{k+k'-1} & C''_{k+k'-2} \end{vmatrix}.$$

Consider now the *purely periodic continued fraction* of period k:

$$(\overline{p_1, q_1; p_2, q_2; \ldots; p_k, q_k}).$$

Associated with this fraction is the following cubic equation which we call the *characteristic cubic* of the ternary continued fraction:

$$\begin{vmatrix} A_{k-2}-\rho & B_{k-2} & C_{k-2} \\ A_{k-1} & B_{k-1}-\rho & C_{k-1} \\ A_k & B_k & C_k-\rho \end{vmatrix} = 0.$$

Written at length this is

$$\rho^3 - M\rho^2 + N\rho - 1 = 0,$$

where

$$M = A_{k-2} + B_{k-1} + C_k$$

and $N = A_{k-2}B_{k-1} - A_{k-1}B_{k-2} + B_{k-1}C_k - B_kC_{k-1} + A_{k-2}C_k - A_kC_{k-2}.$

We have proved the following theorem concerning this cubic:

The characteristic cubic of any periodic ternary continued fraction remains unaltered by any cyclic permutation of the partial quotient pairs.

Using this theorem we obtain the fundamental recursion formulae for the convergent sets:

$$A_n = MA_{n-k} - NA_{n-2k} + A_{n-3k},$$
$$B_n = MB_{n-k} - NB_{n-2k} + B_{n-3k},$$
$$C_n = MC_{n-k} - NC_{n-2k} + C_{n-3k}.$$

It thus appears that the A's, B's and C's are solutions of the following linear difference equation with constant coefficients:

$$U_{x-3k} - MU_{x-2k} + NU_{x-k} - U_x = 0.$$

The theory of such equations is well understood. (Boole, Finite Differences, p. 208). By referring to this theory, we may write

$$A_n = \sum k_i x_i^n \quad (i = 1, 2, 3, \ldots, sk),$$

where $k_1 k_2, \ldots, k_n$ are independent of n and $x_1 x_2, \ldots, x_n$ are the roots of

$$x^{3k} - Mx^{2k} + Nx^k - 1 = 0,$$

and so are the kth roots of the roots of the characteristic cubic. This result leads to the equation:

$$A_n = \rho_1^{n/k} \sum_{v=1}^{k} P_v \omega^{vn} + \rho_2^{n/k} \sum_{v=1}^{k} Q_v \omega^{vn} + \rho_3^{n/k} \sum_{v=1}^{k} R_v \omega^{vn}, \quad \omega^k = 1,$$

where P_v, Q_v, R_v are independent of n, and ρ_1, ρ_2, ρ_3 are the roots of the characteristic cubic. Similar equations hold for B_n and C_n.

From this last result, we obtain the remarkable theorem: *If the characteristic cubic has one root ρ whose modulus is greater than the modulus of either of the other two roots, then*

$$\lim_{n\to\infty}\left(\frac{A_{n+k}}{A_n}\right) = \lim_{n\to\infty}\left(\frac{B_{n+k}}{B_n}\right) = \lim_{n\to\infty}\left(\frac{C_{n+k}}{C_n}\right) = \rho.$$

If the characteristic cubic has two imaginary roots whose common modulus is greater than the absolute value of the real root, then the fractions A_{n+k}/A_n, B_{n+k}/B_n, C_{n+k}/C_n do not approach any limit as n increases beyond limit.

As an example consider the fraction (3, 1; 2, 2) of period 2.

962 963 964

	A_n	B_n	C_n
	1	0	0
	0	1	0
	0	0	1
3,1	1	3	1
2,2	2	7	4
3,1	5	16	8
2,2	15	49	25
3,1	32	104	53
2,2	99	322	164
3,1	210	683	348
2,2	650	2,114	1,077
3,1	1,379	4 485	2,285
2,2	4,268	13,881	7,072
3,1	9,055	29,450	40,051
2 2	28,025	91,147	46,437
3,1	59,458	193.378	98,521
2,2	184 021	598,500	304,920
3,1	390,420	1,269 781	646,920
2,2	1,208,340	3,929,940	2,002,201
3,1	2,563,621	8,337,783	4,247,881
2,2	7,934,342	25,805,227	13,147,084
3,1	16,833,545	54,748,516	27,892,928
2,2	52,099.395	169,415,269	86,327,905

The characteristic cubic in this example is found to be

$$\rho^3 - 7\rho^2 + 3\rho - 1 = 0.$$

An approximation to the largest root of this cubic is given by the ratio of any number in any one of these columns to the number preceding it by two in that column. Thus

52,099.395/7,934,342 6.5663157701 . . .
169,415,269/25,805,227 6.56631577005 . . . .
86,327,905/13,147,084 6.56631577009 . . . .

The other two roots of this equation are imaginary. It is not true, however, that the cubic always has imaginary roots, as the example of the fraction (5,2) will show. The cubic in this case is $\rho^3 - 2\rho^2 - 5\rho - 1 = 0$ the roots of which are all real. Practically all of the work done so far in connection with Jacobi's extension of the continued fraction algorithm has been concerned with cubic irrationalities with negative discriminant. The ternary continued fraction applied without change to all cubic fields alike. It also reaches down into the realm of quadratic irrationalities. In this connection we have the curious theorem that if p is always greater by 2 than q in a periodic ternary continued fraction one root of the characteristic cubic is unity and the other roots will give quadratic irrationalities. This is a special condition, sufficient, but not necessary. One of my students, Dr. Coleman([1]), has worked out in detail the necessary and sufficient condition. Another of my students is at present working on the problem of finding a ternary continued fraction to fit any given quadratic irrationality. This seems to yield results with little difficulty.

The characteristic cubic has been defined by means of the three convergent sets at the end of the first period. This may be called the *first* characteristic cubic, the *second* being formed in the same way from the three convergent sets at the end of the second period, and so in general. Connected with these different cubics we have the interesting theorem: The roots of the characteristic cubic of order k are the kth powers of the roots of the first characteristic cubic. This theorem is true for quaternary and higher continued fractions.

We have seen that the largest root, if any, of the characteristic cubic is approximated by the ratio of A_{n+k}/A_n or B_{n+k}/B_n or C_{n+k}/C_n. One asks at once if anything can be said of the ratios B_n/A_n or C_n/A_n or C_n/B_n. It is not difficult, using the solutions of the difference equation given above to show that these ratios also approximate with increasing n to cubic irrationalities connected by linear fractional relations with the largest root, if any, of the characteristic cubic. We have in fact the theorem: *If the characteristic cubic have one root ρ whose modulus is greater than the modulus of either of the other two roots, then the ratio B_n/A_n approaches a limit σ_1 related to ρ by the equation*:

$$\sigma_1=\frac{\rho B_k+A_kB_{k-2}-A_{k-2}B_k}{\rho A_k+A_{k-1}B_k-A_kB_{k-1}}=\lim\left(\frac{B_n}{A_n}\right).$$

Similarly $$\sigma_2=\frac{\rho C_{k-1}+C_{k-2}A_{k-1}-C_{k-1}A_{k-2}}{\rho A_{k-1}+A_kC_{k-1}-A_{k-1}C_k}=\lim\left(\frac{C_n}{A_n}\right)$$

([1]) Coleman, American Journal of Mathematics, Vol. LII, No. 4, Oct. 1930.

and $$\sigma_3=\frac{\rho C_{k-2}+C_{k-1}B_{k-2}-C_{k-2}B_{k-1}}{\rho B_{k-2}+B_kC_{k-2}-B_{k-2}C_k}=\lim\left(\frac{C_n}{B_n}\right).$$

The actual cubic equations satisfied by these three ratios are obtainable with some little difficulty from these three equations. The coefficients of these cubics are however too complicated to be of much use. The case where the ternary continued fraction has a finite number of non-recurring partial quotient sets is not difficult to treat. The results are entirely similar to the results in the purely periodic case.

It will be remembered that in the case of the expansion of a pure quadratic surd $\sqrt{R}$ in a continued fraction the denominators of the complete quotients in the expansion recur periodically. These denominators are the values of the Pellian quadratic x^2-Ry^2, the values of x and y being the numerators and denominators of the successive convergents. A parallel state of affairs is found for ternary fractions which represent the simple cubic irrationality $\theta=R^{\frac{1}{3}}$. In fact another student of mine has shown that if l, θ, θ^2 expand into a ternary continued fraction which ultimately becomes periodic and if m be the value of the "Pellian cubic"

$$x^3+Ry^3+R^2z^3-3Rxyz$$

for $x=A_n$, $y=B_n$, $z=C_n$, then the series m is periodic([1]). Mr. Daus has also found many other interesting analogies for ternary continued fractions of the above type with those for the quadratic field. The great difficulty in the whole subject for the student interested in the numbertheoretic implications is the lack of uniqueness in the expansions. It is quite possible to get entirely different expansions which will give the same ultimate ratios for the convergent sets. Thus the purely periodic fraction (4,2; 3,5) will give the same values of the irrationalities $\sigma_1, \sigma_2, \sigma_3$ as the mixed fraction (4,2; 3,6; 1,2; 0,1; 0,1; 0,1; 1,2; 4,5; 1,1; 1,4; 0,5; 1,3; 0,1; 0,1; 0,1; 1,2; 1,6). This field of inquiry I have not yet thoroughly explored, but a number of interesting results have already come out of it. It will be observed that the second of the above fractions is characterized by the peculiarity that all of the partial quotient sets with the exception of the first are such that the p less than or equal to the q. Such fractions I have called proper fractions and it would seem likely from experimental evidence that a proper fraction is obtainable equivalent to

([1]) Daus, Am. Jour. Math. xliv, No. 4, Oct. 1922.

any improper one. The most general results which I have been able to prove in this connection however are the following equivalences:

The purely periodic improper fraction $(\overline{2p, 2p-3})$ is equivalent to the mixed fraction

$$(2p, 2p-2; \overline{1, 2p-2; 0,1; 0, p-2; 1,2; 0,1; p-4, p+2; 1, 2p-1}),$$

and the purely periodic improper fraction $(2p+1; 2p-2)$ is equivalent to the mixed proper fraction;

$$(2p+1; 2p-1; \overline{1, 2p-1; 0,1; 0, p-1; 0,2; 0,1; p-3, p-2; 1, 2p-1}).$$

It is easy to show that the characteristic cubic of these fractions has three real roots.

It should be noted in this connection that Jacobi's fractions, from the way the partial quotients are obtained are proper fractions. On that account they cover a very restricted part of the field. Dr. Coleman has, indeed, just obtained a proof that Jacobi's fraction can never represent a quadratic irrationality.

It is well known that the ordinary periodic continued fraction represents one root of a quadratic equation with rational coefficients the other root of which equation may be obtained by inverting the order of the partial quotients. In other words, inverting the order of the partial quotients does not change the field of the irrationality involved. The same is not true of ternary and higher continued fractions, and it is easy to find examples in which the ternary continued fraction and its *inverse* represent irrationalties belonging to different fields. It becomes an interesting question to find what fractions may be inverted without changing their discriminating cubic. This question I have not answered completely, but I have been able to prove the extraordinary theorem that if the p's and q's satisfy a linear relation $Ap_i+Bq_i+C=0$, then the discriminating cubic is the same for the fraction and its inverse(¹). This theorem holds also for quaternary and higher fractions, so that if the partial quotient sets are interpreted as coordinates in space, then if they represent points on a straight line the corresponding fractions may be inverted without changing the discriminating equation; that is without changing the field of the irrationality involved. The proof of this I have carried out in detail for the ternary and quaternary cases, but the method employed is too full of algebraic difficulties for the general case. In the ternary and quaternary cases the proof has been made by induction, and is characterized by a curious feature that the induction requires several simultaneous assumptions which are then extended from n to $n+1$. Thus to prove the M of the first cubic equal to the M' of the second we assume it true for n and proceed to $n+1$ partial quotient pairs. But this extension to $n+1$ requires the assumption of the truth of another equation. This in turn requires the truth of a third and so on till after six such assumptions we return to the original assumption $M=M'$. All of these assumed equations are verified for $n=1, 2$ etc. and must then hold for all values of n.

(¹) Bull. Amer. Math. Soc., Aug. 1931.

Besides such continued fractions which I have called *linear* ternary continued fractions there are of course the fractions which are palindromic, that is, whose partial quotient sets read the same backwards as forwards. I have not been able to show that these exhaust all the cases where the irrationality is unchanged by inverting the pairs, but I am assured by experimental evidence that it is so.

From a given ternary continued fraction another which we shall call the reciprocal may be obtained by replacing (A_n, B_n, C_n) by (A'_n, B'_n, C'_n) where $A'_n=B_{n-2}C_{n-1}-B_{n-1}C_{n-2}$, etc. This new fraction is obtainable in an interesting way from a set of partial quotient pairs as follows: The fraction reciprocal to the fraction

$$(p_1, q_1; p_2, q_2; p_3, q_3, \ldots)$$

is

$$(0, 0; 0, -p_1-q_1, -p_2; -q_2, -p_3; \ldots).$$

If we have a purely periodic fraction, its reciprocal is periodic and if the characteristic cubic of the first is $\rho^3-M\rho^2+N-1=0$, then the characteristic cubic of the second reciprocal to it is $\rho^3-N\rho^2+M\rho-1=0$, so that the equations are reciprocal and the roots of one are the reciprocals of the roots of the other. If the first cubic has three real roots, an approximation to the largest will be given by the first equation and an approximation to the smallest by the second. If, however, the first equation has two imaginary roots with modulus less than the real root, then the first fraction will give an approximation to the real root, while the second will give no approximation at all.