ON THE REDUCTION OF ARITHMETICAL BINARY CUBICS WHICH HAVE A NEGATIVE DETERMINANT

By G. B. MATHEWS *and* W. E. H. BERWICK.

[Received February 6th, 1911.—Read February 9th, 1911.]

1. The reduction of a binary cubic with three real roots has been discussed by Arndt, and presents no particular difficulty. In this case the Hessian is a definite form, and the unitary substitution which reduces it is, in general, unique, and converts the cubic into an equivalent form which is reduced by definition. But the case is different when the cubic has two imaginary roots, for the Hessian is then indefinite, and there is no obvious way of deducing from it a unique reduced form of the cubic. The object of this paper is to explain a method of reduction based upon other considerations, and leading, by a finite process of calculation, to a unique result.

All the forms dealt with are supposed written without the attachment of binomial coefficients: thus (a, b, c) means $ax^2+bxy+cy^2$, and so on.

2. Let the cubic considered be given by

$$f = ax^3+bx^2y+cxy^2+dy^3,$$

and let its linear factors be

$$x-\alpha y, \qquad x-\beta y, \qquad x-\gamma y,$$

where α is real, and β, γ complex. It will be supposed, in the first place, that α is irrational.

We have identically

$$f = (x-\alpha y)(ax^2+mxy+ny^2) = (x-\alpha y)\phi,$$

say, where ϕ is a definite form, with irrational coefficients. Now let us, as usual in the geometrical theory of quadratic forms, take a plane cut up into equivalent curvilinear triangles. The fundamental triangle, in the positive half-plane, lies outside the circle $x^2+y^2 = 1$, and between the lines $4x^2 = 1$; the others are derived from it by unitary substitutions. Now, if we plot off the points corresponding to α, β, γ, the first will be on the axis of real quantities, and the others will be within (or on the boundaries of) two conjugate triangles T, T'. To fix the ideas, suppose

that β has a positive imaginary part, and that it is within the triangle T. Then there is a unique substitution S which converts T into the fundamental triangle T_0, and brings β to a position β_0 within T_0. Applying this substitution to f, we obtain an equivalent form

$$f_0 = Sf = (a_0, b_0, c_0, d_0),$$

which is unique, and by definition reduced. In the case when β is on a boundary, there are two reducing substitutions that bring β to the boundary of T_0, but there is only one of them which brings it to the left of the axis of y. Choosing this as the reducing substitution, we again have a unique form f_0. Two reduced cubics cannot be properly equivalent; so that for a given determinant the number of reduced cubics is equal to the number of classes.

3. It remains to discover S by a finite process of calculation. Since

$$f = (x-ay)\{ax^2+(b+a\alpha)xy-d/\alpha.y^2\}$$

identically, it follows that, if t is a sufficiently close rational approximation to α, then the roots of the rational quadratic ψ, given by

$$\psi = ax^2+(b+at)xy-\frac{d}{t}y^2,$$

will be near the roots of ϕ. Now, if β is within T, it will be possible to surround β by a circle C, of radius r, which is also within T, and if we carry the approximation so far as to bring a root of ψ within C, the substitution S which reduces ψ will also reduce ϕ, and therefore f. Practically, there is not much difficulty in a given case, for if, in any way, we get a sequence of rational approximations $t_1, t_2, t_3, \ldots$, to α, and form the corresponding quadratics $\psi_1, \psi_2, \psi_3, \ldots$, the reducing substitutions $S_1, S_2, S_3, \ldots$ must ultimately become the same; and in many cases this repetition shows itself at a very early stage. It is desirable, however, to establish a criterion enabling us to be sure that we have reached a sufficient degree of approximation. One such test will be explained in the next two articles.

4. For convenience we put $x/y = \theta$, and consider the cubic equation

$$a\theta^3+b\theta^2+c\theta+d = 0$$

with roots α, β, γ. Let u_1, u_2 be two rational approximations to $\beta+\gamma$, such that $u_1 < \beta+\gamma < u_2$, and let v_1, v_2 be two rational approximations to $\beta\gamma$, such that $v_1 < \beta\gamma < v_2$. Further, let z_1, z_2, z_3, z_4 be the roots,

with positive imaginary parts, of the equations

$$\theta^2-u_1\theta+v_1=0, \qquad \theta^2-u_2\theta+v_1=0,$$
$$\theta^2-u_1\theta+v_2=0, \qquad \theta^2-u_2\theta+v_2=0.$$

Then the points z_i are the vertices of a curvilinear parallelogram bounded by parts of the loci

$$2x=u_1, u_2, \qquad x^2+y^2=v_1, v_2$$

respectively. Suppose, now, that the four points z_i are reduced by the same substitution S, defined by

$$S=\begin{pmatrix} p, & q \\ r, & s \end{pmatrix};$$

then the circle $x^2+y^2=v_1$ is transformed by S into

$$(r^2v_1-p^2)(x^2+y^2)+2(rsv_1-pq)x+s^2v_1-q^2=0,$$

which has its centre on the real axis. It follows from this that S transforms the circular arc z_1z_3 into an equivalent arc which lies wholly within the fundamental triangle T_0; for, by supposition, $S(z_1)$ and $S(z_3)$ are reduced points, and if any two reduced points A, B are taken, they can be joined in one, and only one, way by a minor arc of a circle which cuts the real axis orthogonally, and this arc lies wholly within T_0. Similarly with respect to each of the other sides of the curvilinear parallelogram. Hence the transformed quadrilateral lies wholly within T_0, and the point corresponding to β is therefore reduced.

It should be observed that it is assumed here that the approximation has been carried so far that each of the quadratics in θ has complex roots: this must ultimately be the case. Moreover, if we take for granted the known properties of linear substitutions of a complex variable, it is unnecessary to find the equation of the circle into which $x^2+y^2=v_1$ is transformed by S.

5. In order to obtain a set of four auxiliary quadratics, it is convenient to approximate to α by means of an ordinary continued fraction. Thus, suppose that

$$\alpha=\eta_1+\frac{1}{\eta_2}+\frac{1}{\eta_3}+\ldots,$$

where $\eta_2, \eta_3, \ldots$ are all positive integers, and let $\alpha_1, \alpha_2, \alpha_3, \ldots$ be the successive convergents. Then the numbers $\alpha_1-\alpha, \alpha_2-\alpha, \alpha_3-\alpha, \ldots$ are alternately negative and positive, and $|\alpha_i-\alpha|$ is a monotone sequence with limit zero.

Further, if we write

$$-\frac{b}{a}-a_r=u_r, \qquad -\frac{d}{aa_r}=v_r,$$

the numbers $\beta+\gamma-u_1$, $\beta+\gamma-u_2$, ... have signs alternately positive and negative, and so have the numbers $\beta\gamma-v_1$, $\beta\gamma-v_2$, ..., and the limits of the sequences (u_r), (v_r) are respectively $\beta+\gamma$ and $\beta\gamma$. Hence we obtain a sequence of sets of quadratics

$$\theta^2-u_r\theta+v_r=0, \qquad \theta^2-u_{r+1}\theta+v_r=0,$$

$$\theta^2-u_r\theta+v_{r+1}=0, \qquad \theta^2-u_{r+1}\theta+v_{r+1}=0,$$

and there will be a finite index r which will give us a set such as we require. It will generally happen that the first set of quadratics, or possibly more, obviously do not satisfy the proper conditions; but, as soon as this ceases to be the case, we reduce the four quadratics and examine the reducing substitutions to see whether they are identical. If so, we have reached the proper substitution for reducing f; if not, we proceed to the next set of auxiliary quadratics, and so on, until we have attained the object in view.

6. The following example illustrates the points that occur in the practical application of the theory. It is required to reduce (20, 113, 218, 146).

Here we have

$$a=-3+\frac{1}{1+}\frac{1}{3+}\frac{1}{2+}\frac{1}{1+}\cdots,$$

with convergents

$$\frac{-3}{1}, \quad \frac{-2}{1}, \quad \frac{-9}{4}, \quad \frac{-20}{9}, \quad \frac{-29}{13}, \quad \ldots.$$

The first set of auxiliaries is

(60, 159, 146), (20, 53, 73), (60, 219, 146), (20, 73, 73),

of which the third is indefinite; so we must go further. The second set is

(20, 73, 73), (180, 657, 584), (20, 68, 73), (45, 153, 146),

of which the first is ambiguous, and has a root on a boundary line. Without, therefore, proceeding to reduce, we go on to the next stages and obtain the auxiliaries

(45, 153, 146), (200, 680, 657). (180, 617, 584), (1800, 6170, 5913).

The chain of reduction for the first is

$$(45, 153, 146) \sim (146, 139, 38) \sim (38, 13, 20) \sim (20, -13, 38),$$

and for the second

$$(200, 680, 657) \sim (657, 634, 177) \sim (177, 74, 297),$$

with reducing substitutions

$$\begin{pmatrix} 0, & 1 \\ -1, & -1 \end{pmatrix}\begin{pmatrix} 0, & 1 \\ -1, & -2 \end{pmatrix}\begin{pmatrix} 0, & 1 \\ -1, & 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0, & 1 \\ -1, & -1 \end{pmatrix}\begin{pmatrix} 0, & 1 \\ -1, & -2 \end{pmatrix},$$

which are different; so we must still go on. The next set of auxiliaries is

(1800, 6170, 5913), (5220, 17893, 17082),

(2600, 8890, 8541), (7540, 25781, 24674),

and it is found that these are all reduced by the same substitution

$$\begin{pmatrix} 0, & 1 \\ -1, & -1 \end{pmatrix}\begin{pmatrix} 0, & 1 \\ -1, & -2 \end{pmatrix}\begin{pmatrix} 0, & 1 \\ -1, & 0 \end{pmatrix} = \begin{pmatrix} 2, & -1 \\ -1, & 1 \end{pmatrix}.$$

Finally, applying this to the cubic, we obtain the reduced form

$$f_0 = S(20, 113, 218, 146) = (-2, 12, -11, 21).$$

7. For the sake of completeness, we will consider the case when f has a rational linear factor, say,

$$f = (lx+my)(px^2+qxy+ry^2),$$

where l, m are integers prime to each other. Then p, q, r must also be integers, and the reducing substitution is simply that which reduces (p, q, r). This covers all those cases where β is on a boundary of a triangle T: for, in such a case, the reduced form must be either

$$f_0 = a(x-ay)(x^2 \pm xy+ny^2)$$

or

$$f_0 = a(x-ay)(x^2+mxy+y^2),$$

and in each case a must be rational in order that the coefficients of f_0 may be rational. It follows incidentally that every reduced cubic which has no rational root must have one complex root *within* T_0.

8. Without attempting at present to discuss the necessary or sufficient conditions to be satisfied by the coefficients of a given cubic in order that it may be reduced, we will prove one interesting theorem about the leading coefficient a of a reduced cubic.

Let A, B, C be the points corresponding to the roots α, β, γ, and let D be the discriminant of f. Then, since

$$D = a^4(\beta-\gamma)^2(\gamma-\alpha)^2(\alpha-\beta)^2,$$

it follows that

$$|D| = a^4\overline{BC}^2.\overline{CA}^2.\overline{AB}^2 = a^4\Delta,$$

where $\overline{BC}$ is the length of the straight line BC. Now, since B, C are not outside of T_0 and its conjugate, it is obvious from a figure that we obtain a minimum value of Δ by supposing B, C at the points corresponding to $(-1\pm i\sqrt{3})/2$, and A half-way between them. In this limiting case we have

$$\overline{BC} = \sqrt{3}, \qquad \overline{CA} = \overline{AB} = \frac{\sqrt{3}}{2}, \qquad \Delta = \frac{27}{16}.$$

Hence, in general, for a reduced form

$$|D| > 27a^4/16$$

or

$$|a| < \sqrt[4]{\frac{16\Delta}{27}}.$$

In particular, so long as $|D| < 27$, $|a| = 1$; and the only case of $|D| = 27$ with $|a| > 1$ is the form

$$(2x-y)(x^2-xy+y^2).$$

9. It will be seen that the principles of this paper apply not only to cubics, but also to higher binary forms. For instance, if we have a quintic with three real and two complex roots, there will be in general just one reducing substitution which will bring the complex roots to T_0, and this substitution can actually be discovered. If the quintic has two pairs of complex roots, there will be, in general, two distinct reducing substitutions, and in order to find them it will be necessary to approximate to each of the real irrational quadratic factors of the quintic. If the quintic has all its roots real, it will have a definite quadratic covariant, and this may be used for purposes of reduction. However, the discussion of the further applications of the theory may be reserved for another occasion.