A theorem on Newton's polygon. By R. FRITH, B.A., Trinity College.

[*Received* 26 March, *read* 28 May 1934.]

Castelnuovo* has shown that the maximum freedom of a linear system of curves of genus p is $3p+5$, and that a system with this maximum freedom consists of hyperelliptic curves which can be transformed into a system of curves of order n with an $(n-2)$-ple base point and a certain number of double base points; the only exceptions being that when $p=3$ the system may be transformable into that of all quartics, and when $p=1$ the system may be transformable into that of all cubics. Further, since, in the transformed system, the characteristic series is a g_m^{3p+4} it is non-special and hence the redundancy of the base points is zero, therefore each of the double points of this system reduces the freedom by exactly three; and hence if we remove all the double points we get a system of curves of genus p' and freedom $r'=3p'+5$ with only one base point. If we take this point for origin the system of curves can be represented by a single Newton polygon containing in its interior exactly p points, collinear since the curves are hyperelliptic† $(p \neq 3)$, and containing on its boundary $2p+6$ points. From this we can deduce immediately a theorem concerning convex polygons drawn on squared paper; I shall now give an *a priori* proof of this theorem.

It must, however, be noted that we cannot from this theorem directly deduce Castelnuovo's results, for the Newton polygon for a curve of genus p cannot be said to contain in its interior p points unless the only multiple points of the curve are at the origin and at infinity on the axes. What actually is proved is this (since adding a double point to a system of curves of genus p and freedom r does not affect the difference $r-3p$): *A system of curves of genus p which is such that it can be transformed into a system with not more than three base multiple points of order higher than two, has freedom $r \leqslant 3p+5$, and when $r=3p+5$ the system consists of hyperelliptic curves, the only exceptions being for $p=3$ when the system need not consist of hyperelliptic curves and for $p=1$ when* $r \leqslant 3p+6=9$.

1. On a sheet of squared paper the two sets of perpendicular lines are called unit lines, and the points of intersection of these lines I shall denote by *points*. By a polygon I shall mean a convex polygon having all its vertices at points; the segment AB is said

* G. Castelnuovo, *Annali di Mat.* (2), 18 (1890), 119–128.
† H. F. Baker, *Trans. Camb. Phil. Soc.* 15 (1893), 403–450.

to contain a points when there are a points on the segment AB other than the points A and B, and similarly a region is said to contain a points when there are a points absolutely inside the region.

If k is the number of points on the boundary of a polygon which contains in its interior exactly p points ($p > 0$), then $k \leqslant 2p + 6$ unless $p = 1$; and when $p > 1$ and $k = 2p + 6$ the p points are on a line except when $p = 3$.

When the p points are consecutive points on a unit line a rectangle containing them has $2p + 6$ points on its boundary, hence the maximum value for k when p is given is not less than $2p + 6$.

Given a polygon for which $k \geqslant 2p + 6$, if a portion of it is cut off by a line joining two points on the boundary the resulting polygon is still convex, and if Δ, the resulting increase in the difference $k - 2p$, is not negative then the inequality $k \geqslant 2p + 6$ is still true. For a reason which will later be clear I shall make special note of those cuts for which $\Delta = 0$, or for which $\Delta = 1$, and c, the number of points on the line of section, is two.

I shall make frequent use of the following theorems:

(1) *If $ABCD$ is a parallelogram and A, B, C are points, D is also a point.*

(2) *If in the triangle ABC, BC contains a points, CA contains b points and AB contains c points, then the triangle ABC contains at least $\frac{1}{2}(bc - a)$ points and similarly $\frac{1}{2}(ca - b)$ or $\frac{1}{2}(ab - c)$ points.*

(3) *If in* (2) *a and b are both odd, then $c > 0$. Also, if in* (2) *$a = b$, then $c \geqslant a$.*

(4) *If $a = b = 0$ and there are no points in the triangle ABC, then in the region bounded by the infinite line AB and a parallel line through C there are no points.*

The first of these is obvious. The second is seen to be true by adding a point D to make $ABCD$ a parallelogram with, say, AC as a diagonal. Then on each line through one of the a points on BC and parallel to AB there are c points inside $ABCD$. There are therefore at least ac points inside the parallelogram $ABCD$, of these at most b are on AC, hence there are at least $\frac{1}{2}(ca - b)$ points inside the triangle ABC. The first part of the third follows from the fact that the line joining the centre points on BC and CA is parallel to AB; and the second part follows in a similar manner. The last is seen by taking a point D on the line through C parallel to AB making $ABCD$ a parallelogram which contains no points nor has any on the sides AC and BD.

2. Let A, C, B be three consecutive vertices of the polygon, there being a points on BC, b points on CA and c points on AB.

There are then five possible cases: when the resulting polygon (by cutting along AB) contains points and

(i) the triangle ABC contains points and $c \neq 0$;

(ii) the triangle ABC contains no points and $c \neq 0$;

(iii) the triangle ABC contains points and $c = 0$;

(iv) the triangle ABC contains no points and $c = 0$; lastly

(v) the resulting polygon contains no points.

First, if $a = b = 0$ and $c \neq 0$, cutting along AB increases k by $c - 1$ and decreases p by at least c, therefore $\Delta \geqslant c - 1 + 2c > 1$. If one of a and b, (say a), is not zero we consider the cases separately.

(i) By cutting along AB k is increased by $c - a - b - 1$ and (α) when $b \leqslant c$, p is decreased by at least $\frac{1}{2}(ca - b) + c$, then

$$\Delta \geqslant c - a - b - 1 + ca - b + 2c$$
$$= 2(c - b) + (a + 1)(c - 1) \geqslant 0,$$

and (β) when $b > c$, p is decreased by at least $\frac{1}{2}(ab - c) + b$, then

$$\Delta \geqslant c - a - b - 1 + ab - c + 2c$$
$$= (a - 1)(b - 1) + 2(c - 1) \geqslant 0.$$

(A) *Equality* ($\Delta = 0$) *only occurs when* $c = 1$ *and one of* a *or* b *is one;* $c = 2$ *does not give* $\Delta = 1$.

(ii) Since the triangle contains at least $\frac{1}{2}(ac - b)$ points we must have $ac \leqslant b$, therefore $b \neq 0$ and we obtain $bc \leqslant a$ and $ab \leqslant c$. Hence $a = b = c = 1$ and we get $\Delta = 0$.

(B) *When* $a = b = c = 1$ *and there are no points in the triangle* ABC *we have* $\Delta = 0$.

(iii) If $a = b = 0$ and ABC contains $d\,(\neq 0)$ points,

$$\Delta = -1 + 2d > 0.$$

If $a = 0$ and $b \neq 0$, then since the triangle ABC contains at least one point it contains at least $\frac{1}{2}(b + 1)$ points, hence

$$\Delta \geqslant -1 - b + (b + 1) = 0.$$

Finally if both a and b are not zero the triangle ABC contains at least $\frac{1}{2}ab$ points and $\Delta \geqslant -a - b - 1 + ab$. This expression is only negative for (a, b) or (b, a) equal to $(0, h)$, $(1, h)$ or $(2, 2)$, where h is arbitrary. The case $(0, h)$ is not allowed, for neither a nor b is zero; the case $(2, 2)$ is not possible, since from (3) c would then be at least two. The case $(1, h)$ is, also from (3), only possible when h is even and therefore at least two. In this case suppose that $a = 1$ and b is even, then cut out C by the line BX, where X is the point on

https://doi.org/10.1017/S0305004100012470 Published online by Cambridge University Press

AC next to A. Since there is an odd number of points on CX there is, by (3), certainly one point on BX, and hence we have case (i) or (ii).

(C) *The case* $a=c=0$, $b\neq 0$, *when there are just* $\frac{1}{2}(b+1)$ *points inside the triangle, gives* $\Delta=0$; $(a, b)=(2, 3)$, $c=0$ *also gives* $\Delta=0$.

3. Hence all vertices of types (i), (ii) and (iii) can be removed and we reach eventually either a triangle, or a polygon whose vertices are either all of type (iv), or some of type (v) and the rest of type (iv).

I shall consider first the case in which all the vertices are of type (iv). There are in the triangle ABC at least $\frac{1}{2}ab$ points, since

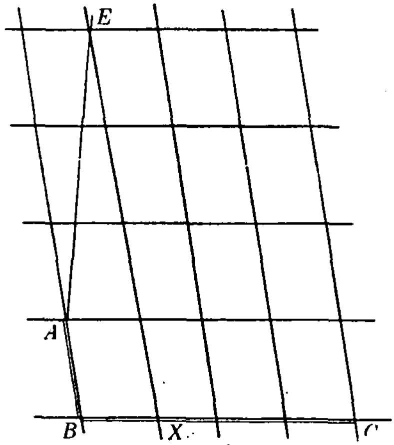

Fig. 1.

c is zero. But there are no points in the triangle, therefore one of a or b is zero. That is any side containing at least one point has adjacent sides with no point on them. If there is a side with points on it let it be BC, if there is not choose BC arbitrarily. Let the adjacent sides be AB and CD and suppose D to be further than A from BC. Then since the vertex C is of type (iv) the lines BD and CD and the triangle BDC contain no point. Hence, from (4), the point A must lie on the line through D parallel to BC; and the infinite line BC, the parallel line through D, i.e. the line AD, and the set of equidistant parallels contain on them all the points. Let the first point on the other side of the polygon through the vertex A be E (Fig. 1); and through E draw the line parallel to AB. Since A is a vertex of type (iv) there are no points on BE or in the triangle BAE, hence, by (4), there are no points in the region bounded by the infinite line AB and the parallel line through E, which line meets BC in the point X, the first point after B on BC

(this is clear, since the lines AB and EX give a second and different set of parallel lines containing all the points, and since AB meets two consecutive lines of the other system in points it follows that every meet of lines of different systems is a point). If there are r points on EX ($r \neq 0$ unless AE is parallel to BC which would mean that $p = 0$), then cutting along EX gives $\Delta = r - 2 + 2r > 0$, and when the line of section has two points on it, i.e. $r = 2$, $\Delta \neq 1$. Hence this case can now be neglected.

4. I consider the next polygon which has some vertices of type (v). First of all I show that a polygon having no point in its interior is either a triangle or a quadrangle with two of its sides parallel. For if it is not a triangle take three sides AB, BC, CD and let A be nearer than D to BC. On the line through A parallel to BC take the point E to make $ABCE$ a parallelogram. The point E is not in the polygon and there are two alternatives. (α) D is on AE. In this case either AD is a side of the polygon which is then a quadrangle with AD parallel to BC, and there can be no points on AB or DC unless D is at E, in which case there may be points on AB and DC, but if there are there are no points on AD or BC. Or else AD is not a side of the polygon; in this case there can be no points on AD and there are therefore no points in the region, in which the rest of the polygon must lie, between the infinite line AB and its parallel through D. This case therefore cannot arise. (β) D is not on AE. Then CD must pass through E or between A and E. When it passes through E we have AB parallel to CD as in (α). When it passes between A and E, meeting AE in λ, not necessarily a point, there are no points on $A\lambda$ and hence no points between the infinite line AB and its parallel through λ (in which region the rest of the polygon must lie), and thus this case cannot arise. So that if the polygon is not a triangle it is a quadrangle $ABCD$ with two sides, say AB and CD, parallel and with no point on the other two sides, in this case BC and DA.

When the polygon is a triangle ABC, then if none of a, b, c are zero the triangle contains at least $\frac{1}{2}(bc - a)$ points, so that $bc \leqslant a$, and similarly $ca \leqslant b$ and $ab \leqslant c$ giving $a = b = c = 1$. If one, say a, is zero, then $bc \leqslant a$ shows that one of b and c is zero also.

5. Now consider a polygon $ABCDG$ such that there are no points in the quadrangle $ABCD$. Then AD, BC must be the parallel lines, for there are no points between the infinite lines which form the parallel sides of a quadrangle containing no points. The infinite lines BC, AD and the equidistant parallel set contain all the points. There is a similar set of equidistant parallel lines among which is the infinite line AB, and every intersection of lines of different sets is a point and *vice versa*. Let AD contain r points and BC contain s points, then $s \leqslant 2r$ if a point G is to exist

making $ABCDG$ a polygon which is not a triangle. Let Y be the next point on the line CD (Fig. 2) and X the next point to Y on the line through Y parallel to BC, which line meets AB in Z, the next point to A on this line. Then we may suppose the side DG to meet the line YZ on the same side of X as is Y, for otherwise we could move C to C', the polygon still being convex, containing the same number of points inside it and one more point on its boundary*. Similarly the side AG must pass to the left of W. There are two cases; first when $r < s$. There are on YZ $2r - s\ (\geqslant 0)$ points all of which we have just seen must lie in the polygon. Next we note that the two sides AG and DG cannot have on

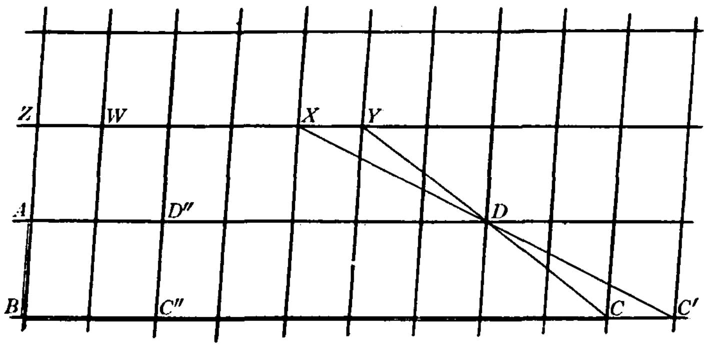

Fig. 2.

them more than $r - (s - r - 1) = 2r - s + 1$ points, including G; for all such points must lie on the r lines parallel to AB which pass through the r points on AD, there cannot be more than one on any of these lines, and on the $s - r - 1$ of them which pass between Y and D there are none of the points. Thus

$$k - 2p \leqslant s + 4 + 2r - s + 1 - 2(r + 2r - s)$$
$$= 2s - 4r + 5 \leqslant 5$$

for $s \leqslant 2r$, so that this case cannot arise.

Secondly when $r \geqslant s$, if we cut BC out by a cut along AD

$$\Delta = r - s - 2 + 2r$$
$$= 3r - s - 2$$
$$\geqslant 2(r - 1) \geqslant 0,$$

when $r > 0$. When $r = 0$, $s = 0$ and there can be no points in the polygon, i.e. $r \neq 0$.

* Such an alteration would make $k \geqslant 2p + 7$, and it will be seen that this is impossible.

https://doi.org/10.1017/S0305004100012470 Published online by Cambridge University Press

(D) *When* $r = s = 1$, $\Delta = 0$; *and* $\Delta = 1$ *for no value of* r.

6. In the next case we have a quadrangle $ABCD$ with no points in the triangle BCD. If there are points on AC then all the interior points lie on this line, for otherwise we could, by (ii), cut off the vertex D or B. In this case when all the points are on AC we see that either there is just one point on AC or else there are no points on the sides AB, BC, CD, DA (from § 4). This case cannot arise, for then $k = 4 < 6$. When there is just one interior point on AC either there is just one point on each of the four sides, then $k = 2p + 6$ and it is easy to see that $ABCD$ is a parallelogram and is in fact a special case of a type dealt with in full later; or else $k < 2p + 6$. There remains the case in which AC, and similarly BD, have no points on them, and all the points lie on one side of each

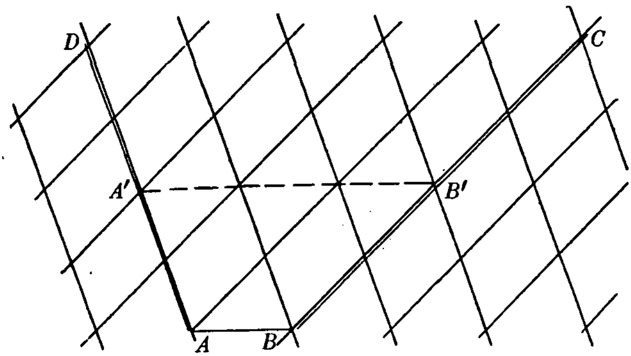

Fig. 3.

of the diagonals. Let these diagonals meet in μ, not a point, and suppose all the points to lie in the region $DC\mu$. Since the triangle ABC contains no point and AC contains no point, either AB or BC contains no point. If AB contains a point, then BC, and similarly AD, contain no point. Then if D is further than C from AB an argument frequently used shows that the region between the infinite line AB and its parallel through D contains no point. But the whole of the triangle $CD\mu$ lies in this region and therefore this case never arises. Thus AB must contain no point. Then the infinite line BC, its parallel through A, and the set of equidistant parallels contain all the points. Similarly we get a set parallel to AD, and the two sets, if the quadrangle is to contain points, are different. Now move the infinite line AB parallel to itself until it once again passes through a point. It is then clear that in its new position it meets every line of either set in points and in particular meets AD in the point A' and BC in the point B', which are by construction the points next to A and B on these lines. There are then on $A'B'$ $r\ (\neq 0)$ points all of which lie inside the quadrangle $ABCD$. Then cut AB off by cutting along $A'B'$; $\Delta = r - 2 + 2r > 0$, and is not equal to one when r is two. Thus the quadrangle is either reduced or is such that all its interior points are on the line $A'B'$, and in this case $k - 2p \geqslant 6$ only when D and C are the next points to A' and B' on the lines AD and BC respectively, and then $k - 2p = (2r + 6) - 2r = 6$.

7. The last case to be considered is when the polygon reduces to a triangle ABC for which $k \geqslant 2p + 6$. If $a = b = 0$, then since the triangle contains at least one point it contains at least $\frac{1}{2}(c+1)$ points and then $k - 2p \leqslant c + 3 - (c+1) \leqslant 2$; therefore no two of a, b, c are zero. If $a = 0$ the triangle contains at least $\frac{1}{2}bc$ points and $k - 2p \leqslant b + c + 3 - bc = 4 - (b-1)(c-1) \leqslant 4$. Hence we may suppose none of a, b, c to be zero. In the triangle there are at least $\frac{1}{2}(ab - c)$ points, and therefore if $k - 2p - 6 > 0$ we have

$$a + b + c + 3 - ab + c - 6 \geqslant 0,$$

$$2(c-1) - (a-1)(b-1) \geqslant 0,$$

and similarly
$$2(a-1) - (b-1)(c-1) \geqslant 0,$$

and, adding,
$$(a + c - 2)(b - 3) \leqslant 0.$$

Now suppose that we have so chosen a, b, c that both a and c are not greater than b; then the above inequality gives two cases. First $b - 3 \leqslant 0$ and $a + c - 2 \geqslant 0$. This implies $a \leqslant 3$, $c \leqslant 3$ and then $k = a + b + c + 3 \leqslant 12$, and hence we must have $p \leqslant 3$. For $p = 3$ we must have $a = b = c = 3$ and $k - 2p = 6$, and this is the exceptional case when the points are not in line. It is easy to see that the only other case satisfying these inequalities is $p = 1$, $a = b = c = 2$, $k - 2p = 7$, which is the other exceptional case.

The other pair of inequalities gives $a + c - 2 \leqslant 0$, $b - 3 \geqslant 0$. These lead directly to $a = c = 1$. The triangle contains at least $\frac{1}{2}(b-1)$ points and therefore

$$k - 2p - 6 \leqslant (5 + b) - (b - 1) - 6 = 0.$$

Hence there must be exactly $\frac{1}{2}(b-1)$ points inside the triangle and they must lie on the line joining the unique points on the sides BC and AB.

8. The original polygon has been reduced to one of the four cases:

I. A triangle $p = 1$, $a = b = c = 2$, $k - 2p = 7$.

II. A triangle $p = 3$, $a = b = c = 3$, $k - 2p = 6$.

III. A triangle $a = c = 1$, $k - 2p = 6$.

IV. A quadrangle with two parallel sides and one point on each of the other sides, $k - 2p = 6$.

Referring to the cuts (A), (B), (C), (D) we see that in every case there is either one or no point on the line of section; and further, since $k-2p$ has at no stage been reduced, in the original polygon we must have $k-2p=6$ and every Δ zero, or else, if the polygon reduces to I, $\Delta = 0$ or 1, there being two points on the line of section. Since all such cuts have been noted and in none of them are there two or three points on the line of section, cases I and II are isolated and cannot be built upon.

III. There is no side with no point on it, therefore we must build on to a side containing one point. Let k', p' refer to the polygon to be added, including the line of section, and let k, p refer to the polygon after section, so that $k-2p=6$, and for the polygon before the cut we have (there being one point on the line of section)

$$(k+k'-4)-2(p+p'+1)=6, \text{ whence } k'-2p'=6.$$

If the cut is of type (D) it is clear that the points inside the new polygon will still lie in a line. For, referring to Fig. 2, we add the portion $ABC''D''$, and AD'' is the line of section; the triangle must lie between the infinite lines AB and $C''D''$ and the resulting

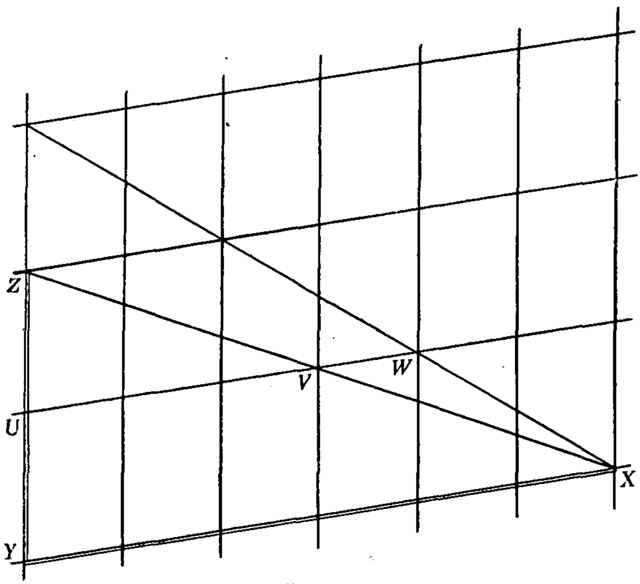

Fig. 4.

polygon is of type IV. In all the other cases we add another triangle of type III. Let the final triangle be XYZ, there being one point U on ZY and one point V on ZX, all the p interior points lying on the line UV. It is required to show that all the

https://doi.org/10.1017/S0305004100012470 Published online by Cambridge University Press

$p + p' + 1$ points inside the new polygon still lie on the line UV. When ZXA, the added triangle, contains no point, the $p + 1$ points in the polygon $XYZA$ are still on the line UV and A must lie on XY so that $XYZA$ is still a triangle of type III. When the triangle does contain points, let the next point to V on UV be W (Fig. 4). Then if the points in the triangle ZXA do not lie on the line UV the side XA must either pass through W or between W and V. (Every meet of the lines of the two parallel sets shown in the figure is a point and all points are accounted for in this way.) The figure shows the case of $p = 2$, and it is clear that this case cannot give a quadrangle $XYZA$ for which $k - 2p = 6$. *A fortiori*, the same is true for $p > 2$. For $p = 1$, however, we are able to obtain (though not by any of the types of cuts which we have made) case II, but that is all. Hence the points lie on the line UV and the resulting polygon is of type IV.

IV. There may be now a side with no point on it. From Fig. 3, the diagram for this case, the line AB being supposed to be the line of section, the whole of the added triangle must lie between the continued lines DA and CB, and comparing with (C), the only case of a section along a line containing no point, we see that this cannot occur. For the rest, the added part must lie between the infinite lines AB and CD, and hence we see that adding to case IV simply gives another polygon of the same type.

9. Concluding: In the original polygon either $k = 2p + 6$ and the p points are in line, and the polygon is a quadrangle with two parallel sides and with one point on each of the other two, or, as a limiting case, a triangle, with the line of points parallel to the parallel sides; or $p = 3$, $k = 2p + 6$ and the polygon is a triangle with three points on each side; or $p = 1$, $k = 7$ and the polygon is a triangle with two points on each side.