

The unity of mathematics: A casebook comprising practical geometry number theory and linear algebra

PETER HILTON and JEAN PEDERSEN

Dedicated to the memory of our colleague and friend Judit Cofman

Abstract. We give a sustained example, drawn largely from earlier publications, of how we may freely pursue a line of mathematical enquiry if we are not constrained, unnaturally, to confine ourselves to a single mathematical subdiscipline; and we draw conclusions from the study of this example which are relevant at many levels of mathematical instruction.

We also include the statement and proof of a new result (Theorem 4.1) in linear algebra which is obviously fundamental to the geometrical investigation which constitutes the leit-motif of the paper.

Key words and phrases: paper-folding, construction of quasi-regular polygons, star polygons, practical geometry, number theory, linear algebra.

ZDM Subject Classification: D20, D30, E10, F60, G20, H60, M10, U60.

1. Introduction

The pedagogical theme of this article is that mathematical reasoning is a unity, not, as the student right through his or her undergraduate years (and even earlier) might think, a collection of very distinct and autonomous disciplines – arithmetic, algebra, geometry, trigonometry, real analysis, complex analysis,

The figures in this article have been reproduced (not always with the same identifying number) from Chapter 4 of [3].

probability theory, statistics, ... The only justification for the separation of the curriculum into these distinct and apparently non-intersecting disciplines is that of administrative convenience – it facilitates the organization and timetabling of classes; and it simplifies enormously the problem of testing whether the students have satisfactorily grasped the material taught. This last consideration seems to be of supreme importance today in the USA and the UK, where regular testing has been mandated at the pre-college level, and is viewed by some very influential people as a certain method of raising standards. Perhaps this is not the place to emphasize our skepticism about the efficacy of constant testing – here we are content to point out that the object of mathematical education is to teach the students to use mathematics effectively, not to devise convenient means of finding out whether the students have acquired certain skills.

For students have surely not learned to use mathematics effectively unless they can decide for themselves what mathematical model will best serve them in solving a problem; and unless they can make intelligent plans as to what questions to ask next once a given problem has been solved. We see no evidence that the influential people referred to above have ever entertained the thought of such objectives in mathematics education.

In this article, we seek to show, by a detailed example, how we think a case study in the learning and using of mathematics should proceed. We have chosen our example because it is one which has occupied a great deal of our attention over the past 20 years. It started, in the hands of one of us (JP) as a piece of practical geometric construction and evolved, under our joint efforts,² into a many-faceted piece of *mathematical* research, involving arithmetic, real analysis,³ linear algebra and number theory. At each stage, the new questions were suggested by the concepts formulated and the results obtained, but those questions – and, of course, the methods used to answer them – might well have had no geometric significance that we could, at the time, see. We were, we would claim, applying mathematics in a very real sense.

Most of our research in this area has been written up; the reader is especially referred to [2], [3] where, in conjunction with our colleague, Derek Holton, we wrote up a number of mathematical topics suitable for study at the undergraduate level, and treated the mathematics arising from paper-folding as Chapter 4 in each of those references. However, we have also included the references [4]–[8], which

²Mathematics is best done by congenial colleagues working together. This is one more valid argument against the rules the authorities impose for testing students' progress.

³This would include the calculus, of course.

deal with various aspects of the mathematics of paper-folding. In fact, reference [8] is somewhat different in purpose, and thus in nature, from the others, since its principal objective is to provide the reader with instructions, and tips, on the actual building of 3-dimensional geometrical models using our paper-folding techniques.

In this article we follow one particular path through the material available in the references, leading to what we refer to as the *Symbol-Creation Theorem*. The preliminary ideas required are all discussed in detail in the references, and are therefore merely summarized in this paper – the summaries constitute Sections 2 and 3 of this paper. On the other hand, the main result of this paper (Theorem 4.1) has not appeared explicitly in the literature, although many of its constituents have been mentioned; and, of course, no explicit proof has been published.⁴ Thus we deal with Theorem 4.1 in complete detail here. In particular, we find ourselves applying a very special method of solving a non-singular system of r linear equations in r unknowns, known as Cramer’s Rule. Interestingly, there are many who today advocate eliminating all discussion of Cramer’s Rule from the curriculum, presumably arguing that the availability of the computer renders the rule superfluous, or inconvenient. This argument would be valid if we were only concerned with *numerical* solutions; however, here the solution given by Cramer’s Rule plays a key role in the *theoretical* development. In any case, we would be uncomfortable if no explanation were given to our students of why, under certain specific conditions, and only under those conditions, a system of linear equations has a unique solution.

The authors would like to express their appreciation of the invitation by Judit Cofman to write this article; and their great sadness at her untimely death.

2. Basic paper-folding procedures

We motivate the paper-folding procedures of this section with a question that has fascinated people at least since the time of the Greeks. The question is whether or not, for a given N , it is possible to construct a regular polygon using Euclidean tools (straight edge and compass)? In fact, Gauss (1777–1855), the Prince of Mathematicians (or Mathematics – it seems that both descriptions appear frequently), completely settled the question by proving that a Euclidean

⁴Aspects of the theory which are irrelevant to the statement and proof of Theorem 4.1 have been entirely omitted from this paper.

construction of a regular N -gon is possible if and only if the number of sides N is a product of the form $N = 2^c \prod \rho_i$, where the numbers ρ_i are distinct Fermat primes – that is, primes of the form $F_n = 2^{2^n} + 1$. Now since F_n is only known to be prime⁵ for

$$F_0 = 3, \quad F_1 = 5, \quad F_2 = 17, \quad F_3 = 257, \quad F_4 = 65537,$$

it is clear that a Euclidean construction of a regular N -gon cannot be known for many values of N ; and, even for these few N , we do not know, to this day, the explicit instructions for their construction in all cases.

Despite Gauss' definitive result we still would like, somehow, to construct all regular polygons. What we describe in this section is a systematic folding procedure that will enable us to approximate, to any degree of accuracy desired, any regular convex N -gon. We will also see, from our examples, that the process will enable us to fold certain regular star $\{\frac{b}{a}\}$ -gons⁶ some of which are shown in Figure 1. For brevity we will refer to the approximations we obtain for both the regular convex N -gons and the regular star $\{\frac{b}{a}\}$ -gons (when $a \geq 2$) as ***quasi-regular polygons***. In most cases the context will make it clear whether or not they are genuine convex polygons. Sometimes we refer to a star b -gon to mean a star $\{\frac{b}{a}\}$ -gon for some a prime to b and satisfying $a < \frac{b}{2}$.

We emphasize that we only allow ourselves to introduce fold lines into a straight strip of paper; these fold lines may have the effect of bisecting angles, but they may not. With these definitions and conditions our approach is to first modify the question so that, instead of asking for an exact construction,⁷ we ask:

*For which $N \geq 3$ is it possible, systematically and explicitly,
to construct quasi-regular (convex) N -gons?*

Surprisingly, as we will show, the answer to this question is: all $N \geq 3$. Furthermore, in showing precisely how this is done we receive a bonus, that is, we see that, by using our paper-folding techniques, we may construct all possible quasi-regular $\{\frac{b}{a}\}$ -gons.

⁵ F_n is now (December 2002) known to be composite for all n in $5 \leq n \leq 30$; but a prime factor of F_n is not known in all these cases.

⁶We will give a more precise definition of these star polygons later in this section. Note that, when we speak of a $\{b/a\}$ -gon, we assume that a, b are coprime; but, if a, b emerge from a calculation they may not, at that stage, be coprime.

⁷Of course, in many cases, such as when $N = 2^c$ (with $c \geq 2$), we can easily give exact constructions, and it would be perverse to construct approximations.

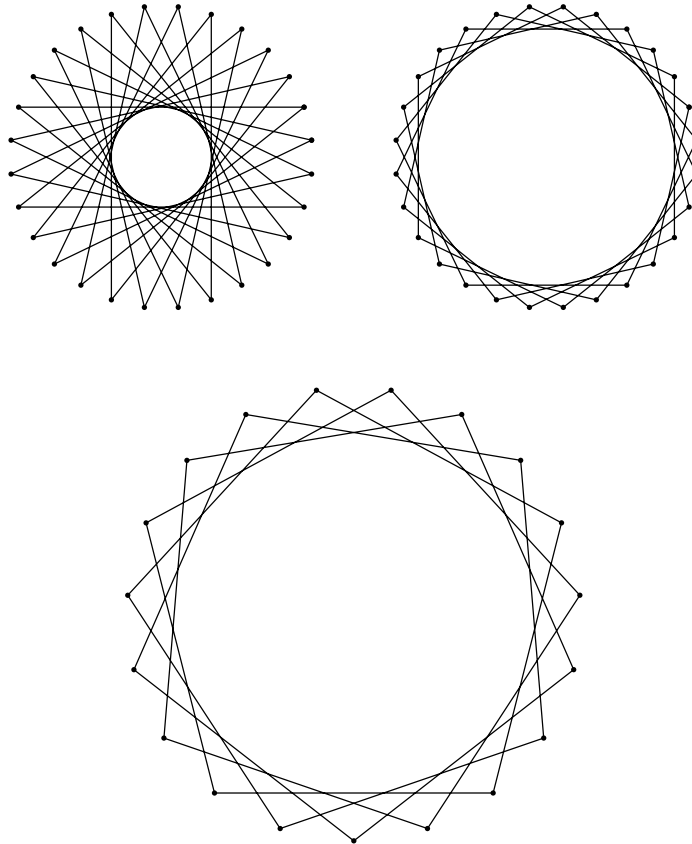


Figure 1. Some star $\{\frac{b}{a}\}$ -gons. (a) $b = 28$, $a = 11$. (b) $b = 27$, $a = 5$.
(c) $b = 19$, $a = 4$.

Let us now begin by describing a precise and fundamental folding procedure, involving a straight strip of paper with parallel edges. We suggest that the reader obtains a long strip of paper and actually carries out the folding procedures described. Adding-machine tape or ordinary unreinforced gummed tape both work well.

Assume that we have a straight strip of paper that has certain vertices marked on its top and bottom edges, at equally spaced intervals, and further assume that it also has creases or folds along straight lines emanating from the vertices at the top edge of the strip. Further assume that the creases at those vertices labelled A_{nk} , $n = 0, 1, 2, \dots$ (see Figure 2(a)) which are on the top edge, form identical

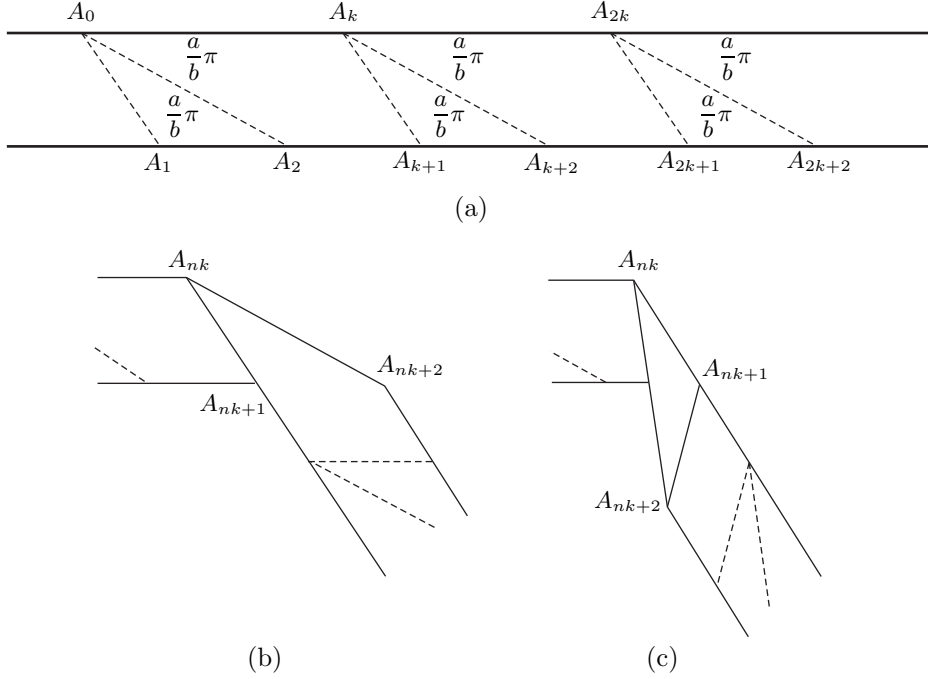


Figure 2

angles of $\frac{a\pi}{b}$ with the top edge, with an identical angle of $\frac{a\pi}{b}$ between the crease along the lines $A_{nk}A_{nk+2}$ and the crease along the lines $A_{nk}A_{nk+1}$ (as shown in Figure 2 (a)). If we fold this strip on $A_{nk}A_{nk+2}$, as shown in Figure 2 (b), and then twist the tape so that it folds on $A_{nk}A_{nk+1}$, as shown in Figure 2 (c), the direction of the top edge of the tape will be rotated through an angle of $2\left(\frac{a\pi}{b}\right)$. We call this process of folding and twisting the FAT-algorithm. Now consider the vertices A_{nk} along the top of the tape, with k fixed and n varying. If the FAT-algorithm is performed on a sequence of angles, each of measure $\frac{a\pi}{b}$, at the vertices given by $n = 0, 1, 2, \dots, b-1$, then the top of the tape will have turned through an angle of $2a\pi$. Thus the vertex A_{bk} will come into coincidence with A_0 ; and the top edge of the tape will have visited every a^{th} vertex of a bounding regular convex b -gon, thus creating a quasi-regular $\left\{\frac{b}{a}\right\}$ -gon. As an example, see Figure 6 (c) where $a = 2$ and $b = 7$. (In order to fit with our usage of “ N -gon” we make a slight adaptation of the Coxeter notation for star polygons (see [1]), so that when we refer to a **quasi-regular** $\left\{\frac{b}{a}\right\}$ -gon we mean a connected sequence

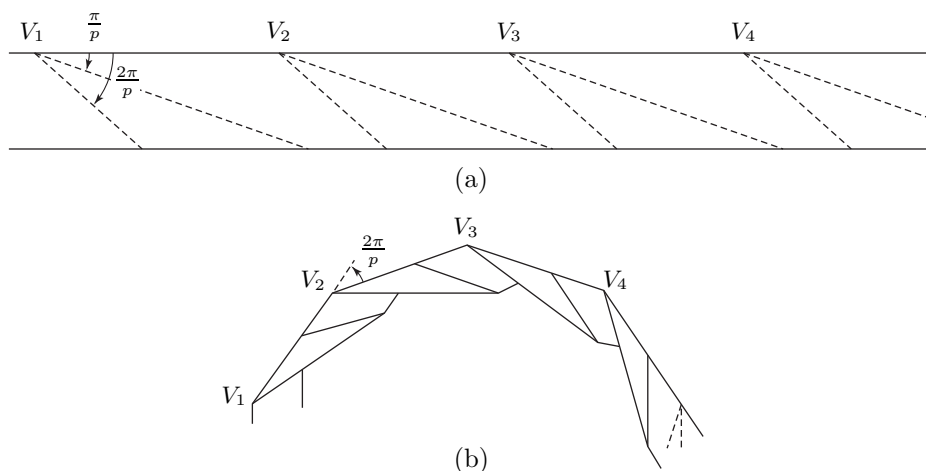


Figure 3

of edges that visits every a^{th} vertex of a quasi-regular b -gon. Thus our N -gon is the special star $\{\frac{N}{1}\}$ -gon. When labeling a convex polygon this way we may well use a lower case letter instead of N .)

Figure 3 illustrates how a suitably creased strip of paper may be folded by the FAT-algorithm to produce a quasi-regular p -gon, (or $\{\frac{p}{1}\}$ -gon). In Figure 3 we have written V_k instead of A_{nk} , since it is more natural in this particular context. Let us now illustrate how the FAT-algorithm may be used to fold a regular convex 8-gon. Figure 4(a) shows a straight strip of paper on which the dotted lines indicate certain special exact crease lines. In fact, these crease lines occur at equally spaced intervals along the top of the tape so that the angles occurring at the top of each vertical line are (from left to right) $\frac{\pi}{2}, \frac{\pi}{4}, \frac{\pi}{8}, \frac{\pi}{8}$. Our immediate interest is focused on the observation that this tape has, at equally spaced intervals along the top edge, adjacent angles each measuring $\frac{\pi}{8}$, and we can therefore execute the FAT-algorithm at 8 consecutive vertices along the top of the tape to produce an exact regular convex 8-gon shown in Figure 4(b). (Of course, in constructing the model one would cut the tape on the first vertical line and glue a section at the end to the beginning so that the model would form a closed polygon.) Notice that the tape shown in Figure 4(a) also has suitable crease lines that make it possible to use the FAT-algorithm to fold a regular convex 4-gon. We leave this as an exercise for the reader and turn to a more challenging construction, the regular convex 7-gon.

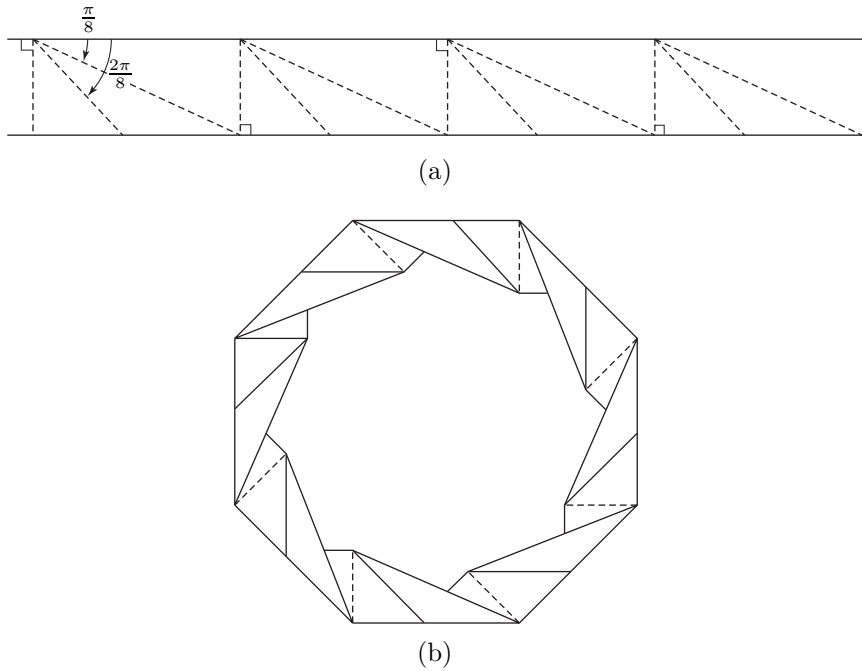


Figure 4

Now, since the 7-gon is the first regular polygon that we encounter for which there does not exist a Euclidean construction, we are faced with a real difficulty in creating a crease line making an angle of $\frac{\pi}{7}$ with the top edge of the tape. We proceed by adopting a general policy we call our *optimistic strategy*. Assume that we *can* crease an angle of $\frac{2\pi}{7}$ (certainly we can come close) as shown in Figure 5(a). Given that we have the angle of $\frac{2\pi}{7}$, it is then a trivial matter to fold the top edge of the strip DOWN to bisect this angle, producing two adjacent angles of $\frac{\pi}{7}$ at the top edge as shown in Figure 5(b). (We say that $\frac{\pi}{7}$ is the *putative* angle on this tape.) Then, since we are content with this arrangement, we go to the bottom of the tape where we observe that the angle to the right of the last crease line is $\frac{6\pi}{7}$ – and we decide, as paper folders, that we will always avoid leaving even multiples of π in the numerator of any angle next to the edge of the tape, so we bisect this angle of $\frac{6\pi}{7}$, by bringing the bottom edge of the tape UP to coincide with the last crease line and creating the new crease line sloping up shown in Figure 5(c). We settle for this (because we are content with an odd multiple of π in the numerator) and go to the top of the tape where

we observe that the angle to the right of the last crease line is $\frac{4\pi}{7}$ – and, since we have decided against leaving an even multiple of π in any angle next to an edge of the tape, we are forced to bisect this angle twice, each time bringing the top edge of the tape DOWN to coincide with the last crease line, obtaining the arrangement of crease lines shown in Figure 5 (d). But now we notice something miraculous has occurred! If we had really started with an angle of exactly $\frac{2\pi}{7}$, and if we now continue introducing crease lines by repeatedly folding the tape DOWN TWICE at the top and UP ONCE at the bottom, we get precisely what we want; namely, pairs of adjacent angles, measuring $\frac{\pi}{7}$, at equally spaced intervals along the top edge of the tape. Let us call this folding procedure the *D^2U^1 -folding procedure* (or, more simply – and especially when we are concerned merely with the related number theory – the *(2, 1)-folding procedure*) and call the strip of creased paper it produces *D^2U^1 -tape* (or, again more simply, *(2, 1)-tape*). The crease lines on this tape are called the *primary crease lines*.

How do we prove that this evident convergence actually takes place? A very direct approach is to admit that the first angle folded down from the top of the tape in Figure 5 (a) might not have been precisely $\frac{2\pi}{7}$. Then the bisection forming the next crease would make the two acute angles nearest the top edge in Figure 5 (b) only approximately $\frac{\pi}{7}$; let us call them $\frac{\pi}{7} + \epsilon$ (where ϵ may be either positive or negative). Consequently the angle to the right of this crease, at the bottom of the tape, would measure $\frac{6\pi}{7} - \epsilon$. When this angle is bisected, by folding up, the resulting acute angles nearest the bottom of the tape, labeled $\frac{3\pi}{7}$ in Figure 5 (c), would in fact measure $\frac{3\pi}{7} - \frac{\epsilon}{2}$, forcing the angle to the right of this crease line at the top of the tape to have measure $\frac{4\pi}{7} + \frac{\epsilon}{2}$. When this last angle is bisected twice by folding the tape down, the two acute angles nearest the top edge of the tape will measure $\frac{\pi}{7} + \frac{\epsilon}{2^3}$. This makes it clear that every time we repeat a D^2U^1 -folding on the tape the error is reduced by a factor of 2^3 .

We see that our *optimistic strategy* has paid off – by blandly assuming we have an angle of $\frac{\pi}{7}$ at the top of the tape to begin with, and folding accordingly, we get what we want – successive angles at the top of the tape which, as we fold, rapidly get closer and closer to $\frac{\pi}{7}$, whatever angle we had, in fact, started with!

We confidently expect that the reader, furnished with this particular, but not special, example of our general procedure for constructing quasi-regular polygons, will now know how to construct any other example, at least in the case of a convex polygon. We further expect the reader to be able, easily, to pass from the

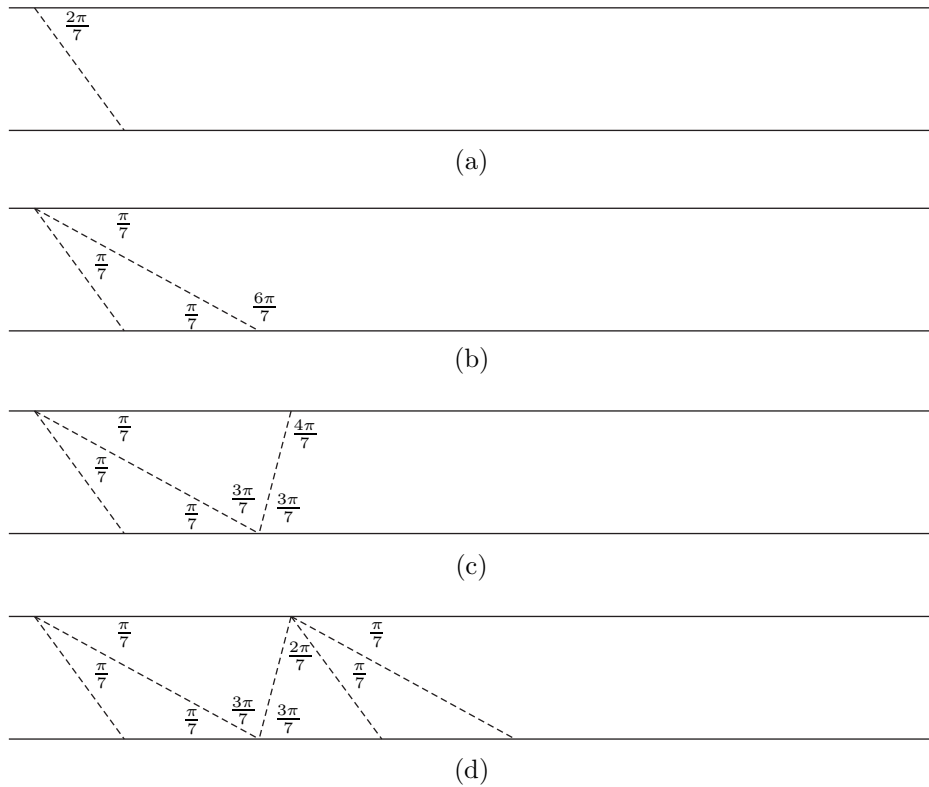


Figure 5

construction of a quasi-regular convex N -gon, with N odd, to the construction of a quasi-regular convex⁸ N -gon for any N . See Figure 9 for an example.

In practice the approximations we obtain by folding paper are quite as accurate as the *real world* constructions with a straight edge and compass – for the latter are only perfect in the mind. In both cases the real world result is a function of human skill, but our procedure, unlike the Euclidean procedure, is very forgiving in that it tends to reduce the effects of human error – and, for many people (even the not so young), it is far easier to bisect an angle by folding paper than it is with a straight edge and compass.

⁸For the construction of the general star polygon, see Chapter 5 of [6].

Figure 6 (c, d) shows the regular $\{\frac{7}{2}\}$ - and $\{\frac{7}{3}\}$ -gons that are produced from the D^2U^1 -tape by executing the FAT-algorithm on the crease lines that make angles of $\frac{2\pi}{7}$ and $\frac{3\pi}{7}$, respectively, with an edge of the tape (if the angle needed is at the bottom of the tape, as with $\frac{3\pi}{7}$, simply turn the tape over so that the required angle appears on the top). In Figure 6 (c, d) the FAT-algorithm was executed on every other suitable vertex along the edge of the tape so that, in (c), the resulting figure, or its flipped version, could be woven together in a more symmetric way and, in (d), the excess could be folded neatly around the points.

It is now natural to ask the following two questions.

Question 1. Can we use the same general approach used for folding a convex 7-gon to fold a convex N -gon with N odd, at least for certain specified values of N ? If so, can we always prove that the actual angles on the tape really converge to the putative angle we originally sought?

Question 2. Do we always get a quasi-regular $\{\frac{b}{a}\}$ -gon with any general iterative folding procedure, perhaps with other periods, such as those represented by

$$D^3U^3, D^4U^2, \text{ or } D^3U^1D^1U^3D^1U^1?$$

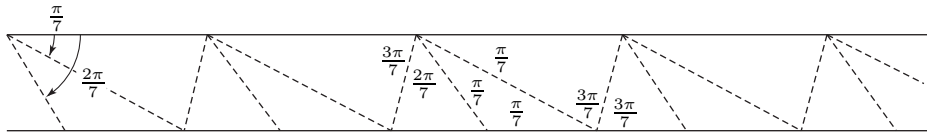
How does the folding procedure determine $\frac{b}{a}$? (The *period* is determined by the repeat of the *exponents*, so the above examples have periods 1, 2 and 3 respectively.)

The answer to Question 1 is yes and even more is true. We will soon show an algorithm for determining the folding procedure that produces tape from which you can construct any given quasi-regular $\{\frac{b}{a}\}$ -gon, if a, b are odd with $a < \frac{b}{2}$. Because discussing how to construct quasi-regular $\{\frac{b}{a}\}$ -gons when b is odd and a is even would distract us from getting to our main result we refer the reader to Chapter 4 of [2] for this result. Question 2 will be answered in Section 4.

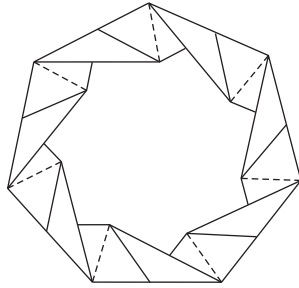
However before we begin to answer Question 1 let us just look at the general **1-period** folding procedure D^nU^n (because the result has such interesting historical implications). A typical portion of the tape would appear as illustrated in Figure 7 (a).

It turns out that the smallest angle u_k at the top, and bottom, of this tape approaches $\frac{\pi}{2^n+1}$; that is,

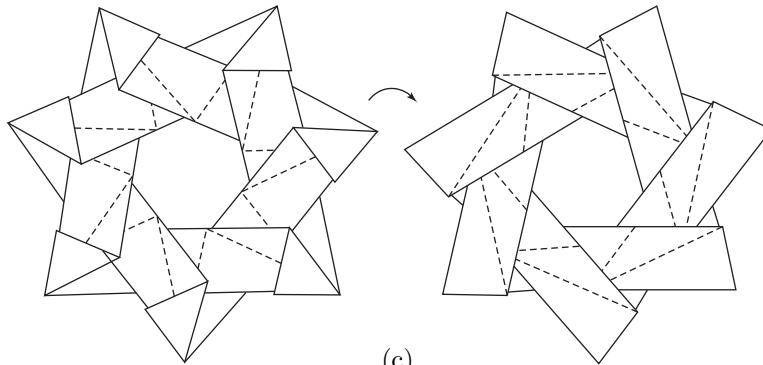
$$u_k \rightarrow \frac{\pi}{2^n+1} \text{ as } k \rightarrow \infty. \quad (2.1)$$



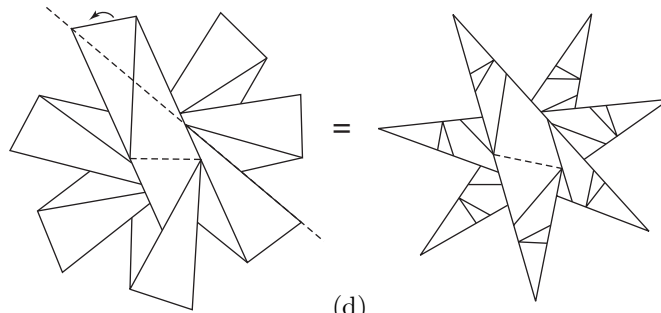
(a)



(b)



(c)



(d)

Figure 6

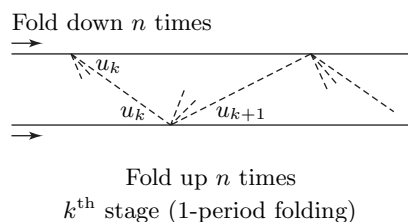


Figure 7

A proof of (2.1) similar to the one provided above for the tape whose smallest angle approached $\frac{\pi}{7}$ may be given. In fact, we can see that, if the original fold down made an angle differing by ϵ_0 from the putative angle of $\frac{\pi}{2^n+1}$, then the error at the k^{th} stage of the $D^n U^n$ -folding procedure would be given by

$$|\epsilon_k| = \frac{|\epsilon_0|}{2^{nk}}. \quad (2.2)$$

Hence we see that the $D^n U^n$ -folding procedure produces tape from which we may construct quasi-regular $(2^n + 1)$ -gons – and, of course, these include those N -gons for which N is a Fermat number, prime or not. We would like to believe that the ancient Greeks and Gauss would have appreciated the fact that, when $n = 1, 2, 4, 8$ and 16 , the $D^n U^n$ -folding procedure produces tape from which we can obtain, by means of the FAT-algorithm, a quasi-regular 3-, 5-, 17-, 257- and 65537-gon, respectively. What's more, if $n = 3$, we approximate the regular 9-gon whose non-constructibility by Euclidean tools is very closely related to the non-trisectibility of an arbitrary angle.

The case $N = 2^n + 1$ is untypical of the general case in an important respect, since we may construct $(2^n + 1)$ -gons from our folded tape by special methods (not involving the FAT-algorithm), in which, however, the top edge does not describe the polygon, as it does in the FAT-algorithm. Figure 8 shows how the $D^2 U^2$ -tape shown in part (a) may be folded along just the short lines of the creased tape to form the *outline* of the quasi-regular pentagon shown in (b), and along just the long lines of the creased tape to form the *outline* of the slightly larger quasi-regular pentagon shown in (c); and, finally, we show in (d) the quasi-regular pentagon formed by an edge of the tape when the FAT-algorithm is executed.

It is easy to see how to construct quasi-regular polygons with $2^c N$ sides, N odd, if we already know how to construct quasi-regular N -gons. If, for example, we wished to construct a quasi-regular 10-gon then we take the $D^2 U^2$ -tape (which you may recall produced FAT 5-gons) and introduce a **secondary crease line** by

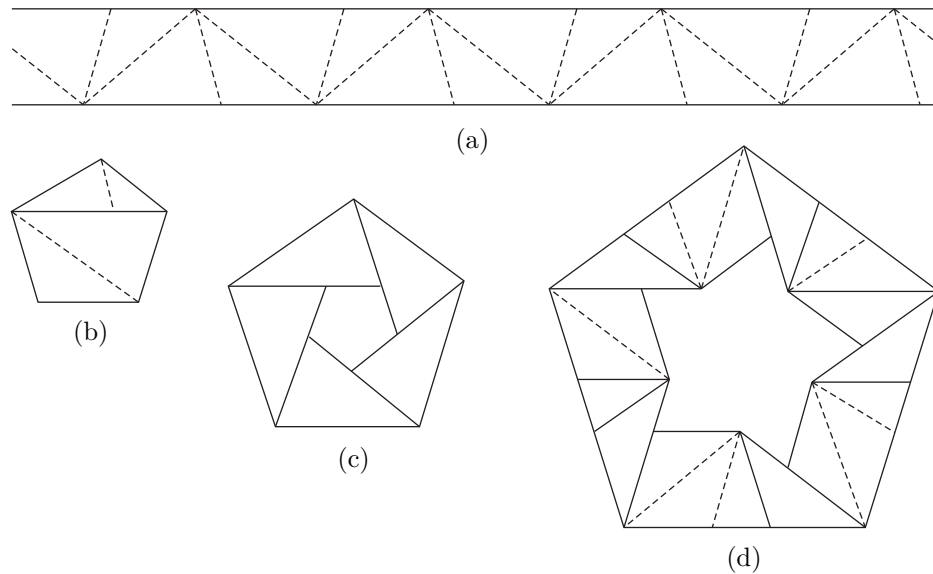


Figure 8

bisecting each of the angles of $\frac{\pi}{5}$ next to the top (or bottom) edge of the tape. The FAT-algorithm may be used on the resulting tape to produce the quasi-regular convex FAT 10-gon, as illustrated in Figure 9. It should now be clear how to construct a quasi-regular 20-gon, 40-gon, 80-gon, . . .

It should be clear from this example that if we can describe how to construct quasi-regular N -gons for any odd number $N \geq 3$, then we will, in fact, also be able to construct any quasi-regular polygon having an even number of sides.

There is a great deal of fascinating mathematics, of a number-theoretic nature, arising from the 2-period folding procedures $D^m U^n$. However, this has already been published in detail elsewhere (see [2]–[4], [6]–[8]), so we will omit it from this summary, and move on, at last, to answering our first question.

Turning, then, to Question 1, we now show, by considering a particular but not special case, how to determine the folding instructions for producing tape from which we can construct any given quasi-regular $\{\frac{b}{a}\}$ -gon, with b, a odd and $a < \frac{b}{2}$.

Thus, suppose we want to construct a quasi-regular $\{\frac{11}{3}\}$ -gon. Then, of course, $b = 11$, $a = 3$ and we proceed precisely as we did when we wished to construct the regular convex 7-gon; that is, we adopt our *optimistic strategy*

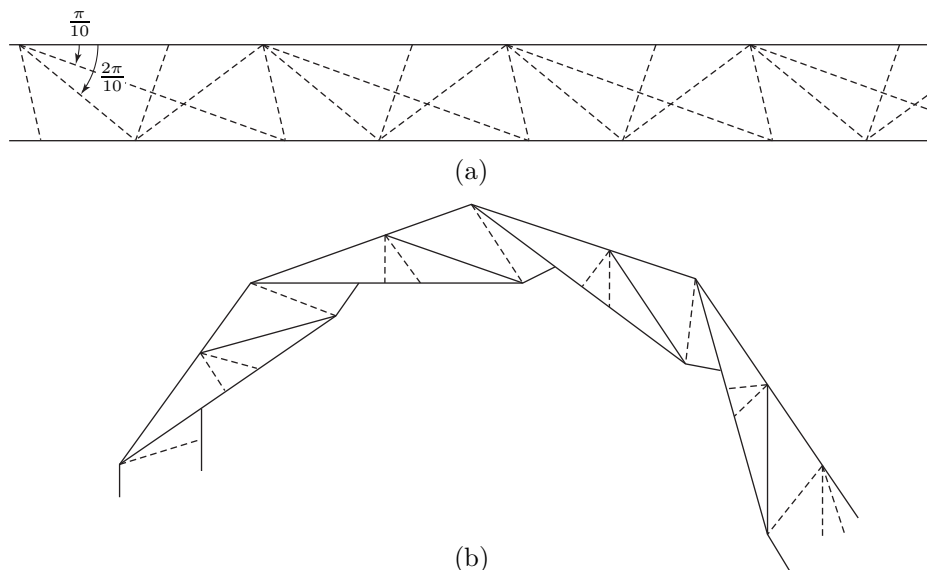


Figure 9

which, as you recall, means that we *assume* we've got what we want and, as we will show, we then actually get an arbitrarily good approximation to what we want! This time we assume we can fold the desired putative angle of $\frac{3\pi}{11}$ at A_0 (see Figure 10 (a)) and we adhere to the same rules that we used in constructing the quasi-regular 7-gon, namely, we adopt the following principles:

- (1) Each new crease line goes in the forward (left to right) direction along the strip of paper.
- (2) Each new crease line always bisects the angle between the last crease line and the edge of the tape from which it emanates.
- (3) The bisection of angles at any vertex continues until a crease line produces an angle of the form $\frac{a'\pi}{b}$ where a' is an *odd* number; then the folding stops at that vertex and commences at the intersection point of the last crease line with the other edge of the tape.

Once again the *optimistic strategy* works; and following this procedure results in tape whose angles converge to those shown in Figure 10 (b). We could denote this folding procedure by the expression $D^1U^3D^1U^1D^3U^1$, interpreted in the obvious way on the tape – that is, the first exponent “1” refers to the one bisection (producing a line in a downward direction) at the vertices A_{6n} (for $n = 0, 1, 2, \dots$)

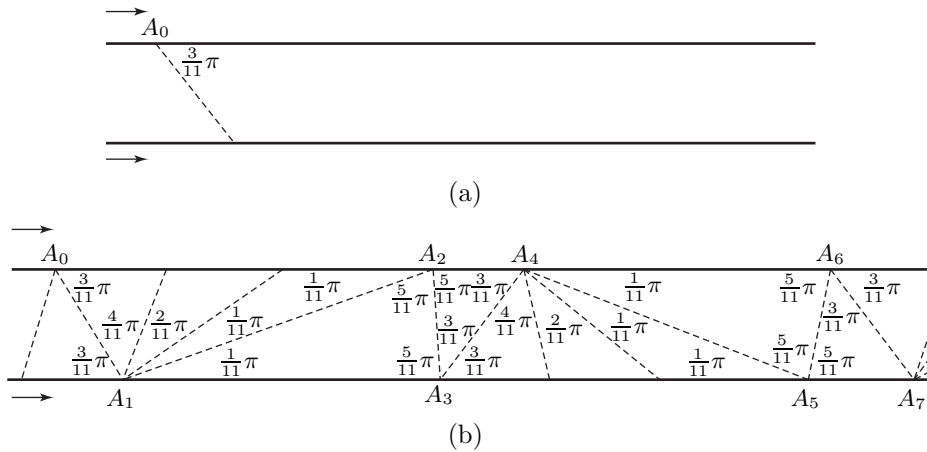


Figure 10

on the top of the tape; similarly, the “3” refers to the 3 bisections (producing creases in an upward direction) made at the bottom of the tape through the vertices A_{6n+1} ; etc. However, since the folding procedure is *duplicated* halfway through, we can abbreviate the notation and write it simply as $\{1, 3, 1\}$, with the understanding that we alternately fold from the top and bottom of the tape as described, with the number of bisections at each vertex running, in order, through the values $1, 3, 1, \dots$. We call this a *primary folding procedure of period 3* or a *3-period folding*.

To prove the convergence we can use an error-correction type of proof like that given earlier in this section for the 7-gon. We leave the details to the reader, and explore here what we can do with this $(1, 3, 1)$ -tape. First, note that, starting with the putative angle $\frac{3\pi}{11}$ at the top of the tape, we produce a putative angle of $\frac{\pi}{11}$ at the bottom of the tape, then a putative angle of $\frac{5\pi}{11}$ at the top of the tape, then a putative angle of $\frac{3\pi}{11}$ at the bottom of the tape, and so on. A careful inspection of this tape shows that we could use the FAT algorithm on it to fold quasi-regular $\{\frac{11}{a}\}$ -gons, when $a = 1, 2, 3, 4, 5$. To put the result in a form that suggests the desired generalization, we may say that, if there are crease lines enabling us to fold a star $\{\frac{11}{a}\}$ -gon, there will be crease lines enabling us to fold star $\{\frac{11}{2^k a}\}$ -gons, where $k \geq 0$ takes any value such that $2^{k+1}a < 11$. These features, described for $b = 11$, would be found with any *odd* number b . However, this tape has a special symmetry as a consequence of its odd period; namely, if it is “flipped” about the horizontal line half way between its parallel edges,

the result is a *translate* of the original tape. As a practical matter this special symmetry of the tape means that we can use either the top edge or the bottom edge of the tape to construct our polygons. On tapes with an *even* period the top edge and the bottom edge of the tape are not translates of each other (under the horizontal flip), which simply means that care must be taken in choosing the edge of the tape used to construct a specific polygon. Figures 11 (a, b) show the completed $\{\frac{11}{3}\}$ -, $\{\frac{11}{4}\}$ -gons, respectively.

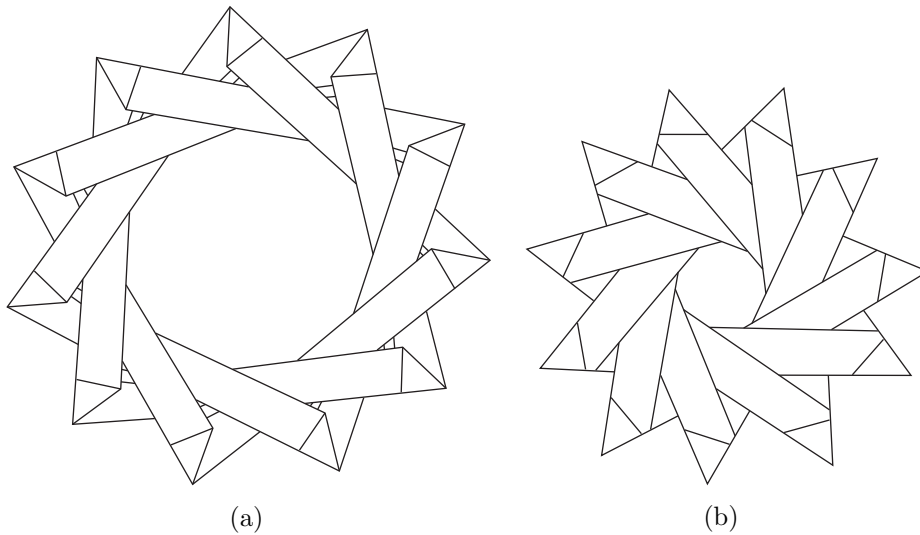


Figure 11

Now, to set the scene for the number theory of Section 3, and to enable us systematically to determine the folding procedure for any given a and b , let us look at the patterns in the *arithmetic* of the computations when $a = 3$ and $b = 11$. Referring to Figure 10 (b) we observe that

the smallest angle to the right of A_n where	is of the form $\frac{a}{11}\pi$ where	and the number of bisections at the next vertex ⁹
$n = 0$	$a = 3$	$= 3$
1	1	1
2	5	1
3	3	3
4	1	1
5	5	1

We could write this in shorthand form as follows:

$$(b =)11 \left| \begin{array}{ccc} (a =)3 & 1 & 5 \\ & 3 & 1 & 1 \end{array} \right| \quad (2.3)$$

Observe that, had we started with the putative angle of $\frac{\pi}{11}$, then the symbol (2.3) would have taken the form

$$(b =)11 \left| \begin{array}{ccc} (a =)1 & 5 & 3 \\ & 1 & 1 & 3 \end{array} \right| \quad (2.3')$$

In fact, it should be clear that we can *start anywhere* (with $a = 1, 3$ or 5) and the resulting symbol, analogous to (2.3), will be obtained by cyclic permutation of the matrix component of the symbol, placing our choice of a in the first position along the top row.

In general, suppose we wish to fold a $\{\frac{b}{a}\}$ -gon, with b, a odd and $a < \frac{b}{2}$.

Then we may construct a symbol¹⁰ as follows. Let us write

$$b \left| \begin{array}{cccc} a_1 & a_2 & \dots & a_r \\ k_1 & k_2 & \dots & k_r \end{array} \right| \quad (2.4)$$

where b, a_i ($a_1 = a$) are odd, $a_i < \frac{b}{2}$, and

$$b - a_i = 2^{k_i} a_{i+1}, \quad i = 1, 2, \dots, r, \quad a_{r+1} = a_1. \quad (2.5)$$

At this stage, we do not assume that $\gcd(b, a) = 1$ but we have assumed that the list a_1, a_2, \dots, a_r is without repeats. Indeed, if $\gcd(b, a) = 1$ we say that the symbol (2.4) is *reduced*, and if there are no repeats among the a_i 's we say that the symbol (2.4) is *contracted*. (It is, of course, theoretically possible to consider symbols (2.4) in which repetitions among the a_i are allowed.) We regard (2.4) as encoding the general folding procedure to which we have referred. The key fact, answering Question 1, is that, with the given data (involving just b and $a = a_1$), there is always a number r , the *period* of the folding procedure, such that $a_{r+1} = a_1$, but there are no repeats among a_1, a_2, \dots, a_r .

⁹Notice that, referring to Figure 10 (b), to obtain an angle of $\frac{3\pi}{11}$ at A_0, A_6, A_{12}, \dots , the folding instructions would more precisely be $U^3 D^1 U^1 D^3 U^1 D^1 \dots$. But we don't have to worry about this distinction.

¹⁰More exactly, a 2-symbol, since there is a notion of a t -symbol for any integer $t \geq 2$. We will not discuss this here, but refer the interested reader to Chapter 4 of [2] and Chapter 4 of [3].

Example 1. If we wish to fold a 31-gon we may start with $b = 31$, $a = 1$ and construct the symbol

$$(b=)31 \left| \begin{array}{cc} (a=)1 & 15 \\ 1 & 4 \end{array} \right|$$

which tells us that folding D^4U^1 will produce tape (usually denoted (4,1)-tape) that can be used to construct a FAT 31-gon. In fact, this tape can also be used to construct FAT

$$\left\{ \frac{31}{2} \right\}^-, \left\{ \frac{31}{4} \right\}^-, \left\{ \frac{31}{8} \right\}^- \quad \text{and} \quad \left\{ \frac{31}{15} \right\}^- \text{-gons.}$$

However, if we wish to fold a $\left\{ \frac{31}{3} \right\}$ -gon we start with $b = 31$, $a = 3$ and construct the symbol

$$(b=)31 \left| \begin{array}{cc} (a=)3 & 7 \\ 2 & 3 \end{array} \right|$$

which tells us to fold D^2U^3 – or, more simply, to use the (2,3)-folding procedure – to produce (2,3)-tape from which we can fold the FAT $\left\{ \frac{31}{3} \right\}$ -gon. Again, we get more than we initially sought, since we can also use the (2,3)-tape to construct

$$\text{FAT } \left\{ \frac{31}{6} \right\}^-, \left\{ \frac{31}{12} \right\}^-, \left\{ \frac{31}{7} \right\}^- \quad \text{and} \quad \left\{ \frac{31}{14} \right\}^- \text{-gons.}$$

However, we don't have a folding procedure that produces the $\left\{ \frac{31}{5} \right\}$ -gon. Thus we construct another symbol, this time with $b = 31$, $a = 5$.

$$(b=)31 \left| \begin{array}{cccc} (a=)5 & 13 & 9 & 11 \\ 1 & 1 & 1 & 2 \end{array} \right|$$

which tells us to fold $D^1U^1D^1U^2$ – or, more simply, to use the 4-period (1,1,1,2)-folding procedure – to produce (1,1,1,2)-tape from which we can fold the FAT $\left\{ \frac{31}{5} \right\}$ -gon. Once again, we get more than we asked for, we can also use the (1,1,1,2)-tape to construct

$$\text{FAT } \left\{ \frac{31}{10} \right\}^-, \left\{ \frac{31}{13} \right\}^-, \left\{ \frac{31}{9} \right\}^- \quad \text{and} \quad \left\{ \frac{31}{11} \right\}^- \text{-gons.}$$

We can combine all the possible symbols for $b = 31$ into one *complete symbol*, adopting the notation

$$31 \left| \begin{array}{cccc|cccc} 1 & 15 & 3 & 7 & 5 & 13 & 9 & 11 \\ 1 & 4 & 2 & 3 & 1 & 1 & 1 & 2 \end{array} \right|. \quad (2.6)$$

Notice in (2.6) that the total amount of folding would be the same to produce any quasi-regular (convex or star) 31-gon. Since it is very difficult to bisect an angle 4 times you may wish to use the second or third parts of this symbol to produce the tape. Even if you really want a convex 31-gon it may be easier, in practice, to produce the star polygon first and then use the vertices of that polygon to determine the convex polygon.

Example 2. Suppose we wish to fold a $\{\frac{93}{3}\}$ -gon.¹¹ Constructing the symbol we obtain

$$93 \left| \begin{array}{cc} 3 & 45 \\ 1 & 4 \end{array} \right| \quad (2.7)$$

which tells us that folding D^4U^1 will produce tape that can be used to construct a FAT $\{\frac{93}{3}\}$ -gon. But, recall, from Example 1, this is precisely how we folded to get a FAT 31-gon. In fact, this must surely be one of the most difficult methods for finding out that $\frac{93}{3} = 31!$

Example 2 shows why we will, in general, prefer to have $\gcd(b, a) = 1$.

In the next section we formalize these ideas and prove that there is always a unique contracted symbol for given odd numbers b, a with $a < \frac{b}{2}$.

3. The Quasi-order Theorem

Thus we claim that, given positive odd integers b, a with $a < \frac{b}{2}$, there is always a unique contracted symbol

$$b \left| \begin{array}{cccc} a_1 & a_2 & \dots & a_r \\ k-1 & k_2 & \dots & k_r \end{array} \right|, \quad a_1 = a, \quad a_i \neq a_j \quad \text{if } i \neq j. \quad (3.1)$$

where each a_i is odd, $a_i < \frac{b}{2}$, and

$$b - a_i = 2^{k_i} a_{i+1} \quad i = 1, 2, \dots, r, \quad a_{r+1} = a_1. \quad (3.2)$$

We argue as follows. We fix b and let S be the set of positive odd numbers $a < \frac{b}{2}$. Given $a \in S$, define a' by the rule

$$b - a = 2^k a', \quad k \text{ maximal}; \quad (3.3)$$

¹¹You might have noticed that this should turn out to be a $\{\frac{31}{1}\}$ -gon, but suppose we just proceed (as some of our less sophisticated students might do these days) without making this observation.

that is, we take as many factors of 2 as we can out of $b - a$. Notice that $k \geq 1$, since $b - a$ is certainly even. We claim that $a' \in S$. First, a' is obviously odd. Second, $2a' \leq 2^k a' = b - a < b$, so $a' < \frac{b}{2}$. Thus (3.3) describes a function $\psi : S \rightarrow S$, such that $\psi(a) = a'$. We will show that ψ is a **permutation** of the finite set S ; it is sufficient to show that ψ maps S onto itself; and, to show this, it is certainly enough to exhibit a function $\varphi : S \rightarrow S$ such that $\psi\varphi(a') = a'$. We define φ as follows: given $a' \in S$, let k be *minimal* such that $2^k a' > \frac{b}{2}$ and set $\varphi(a') = a$, where

$$a = b - 2^k a'. \quad (3.4)$$

Notice that $k \geq 1$, since $a' < \frac{b}{2}$, so that a is odd; that $a < \frac{b}{2}$ since $2^k a' > \frac{b}{2}$; and that $b > 2^k a'$, since $2^{k-1} a' < \frac{b}{2}$, so that a is positive. Thus $a \in S$; and comparison of (3.3), (3.4) shows that, as claimed, $\psi\varphi(a') = a'$. Thus ψ is a permutation and φ is the inverse permutation.

The permutation ψ has one more important property. We write $\psi(a) = a'$, as above, and claim that

$$\gcd(b, a) = \gcd(b, a'). \quad (3.5)$$

For it is clear from (3.3) that if $d \mid b$ and $d \mid a'$ then $d \mid b$ and $d \mid a$. Conversely, if $d \mid b$ and $d \mid a$, d is odd and $d \mid 2^k a'$ so $d \mid b$ and $d \mid a$. Thus if a_1 in (3.1) is coprime to b , so are a_2, a_3, \dots, a_r , and we may, if we wish, confine ψ and φ to the subset S_0 of S consisting of those $a \in S$ which are coprime to b ; that is, we may confine ourselves to reduced symbols.

We now use the fact that, given any permutation ψ of a finite set S_0 and any $a \in S_0$, then a must generate a cycle, in the sense that, if we iterate ψ , getting

$$a, \psi(a), \psi^2(a), \psi^3(a), \dots$$

(here $\psi^2(a) = \psi(\psi(a))$, etc. and we may write $\psi^0(a)$ for a)

we must eventually repeat, that is, we will find $m > 0$ such that

$$a, \psi(a), \dots, \psi^{m-1}(a)$$

are all different but $\psi^m(a) = a$. In case this is not clear to you, we give the easy proof. Certainly, since S_0 is finite, we must eventually repeat in the weaker sense that we find $s \geq 0$, $m > 0$ such that $\psi^s(a) = \psi^{s+m}(a)$. Suppose this is the *first* time we get a repeat. We claim that $s = 0$; for, if not, we have

$$\psi(\psi^{s-1}(a)) = \psi(\psi^{s-1+m}(a)).$$

But ψ is one-one, so $\psi^{s-1}(a) = \psi^{s-1+m}(a)$ and the given repeat wasn't our first repeat. Thus $s = 0$ and $a = \psi^m(a)$. Of course, with regard to (3.1), $m = r + 1$. This completes the proof of our claim.

So we have a universal algorithm for folding a $\{\frac{b}{a}\}$ -gon, where a, b are coprime odd integers with $a < \frac{b}{2}$. But, from the number-theoretic point of view, it turns out that we have much more. For, reverting to (3.1), let

$$k = \sum_{i=1}^r k_i \quad (\text{we may call this the } \mathbf{fold-total}).$$

Then, as in Chapter 4 of [2], we prove

THEOREM 3.1 (*The Quasi-order Theorem*). *Suppose that (3.1) is not only contracted but also reduced. Then the quasi-order of 2 mod b is k . That is, k is the smallest positive integer such that*

$$2^k \equiv \pm 1 \pmod{b}.$$

In fact, $2^k \equiv (-1)^r \pmod{b}$.

PROOF. The proof is really a triumph of technique! First, we find it convenient to think in terms of the φ -function rather than the ψ -function. Thus we work *backwards* in constructing our symbol (3.1). Also we will find it convenient to repeat the initial number a_1 . Precisely, we write our *modified symbol* as

$$b \begin{pmatrix} c_1 & c_2 & c_3 & \cdot & \cdot & \cdot & c_r & c_1 \\ \ell_1 & \ell_2 & \ell_3 & \cdot & \cdot & \cdot & \ell_r & \end{pmatrix} \quad (3.6)$$

where¹² (compare (3.2) and (3.4)) ℓ_i is minimal such that $2^{\ell_i} c_i > \frac{b}{2}$ and

$$b - c_{i+1} = 2^{\ell_i} c_i, \quad i = 1, 2, \dots, r \quad (c_{r+1} = c_1). \quad (3.7)$$

We set $\ell = \sum_{i=1}^r \ell_i$, and our first task will be to prove that

$$2^\ell \equiv (-1)^r \pmod{b}. \quad (3.8)$$

To this end, consider the $(\ell + 1)$ numbers, all less than $\frac{b}{2}$,

$$\begin{aligned} & c_1, 2c_1, \dots, 2^{\ell_1-1} c_1, c_2, 2c_2, \dots, 2^{\ell_2-1} c_2, c_3, \\ & \dots, c_r, \dots, 2^{\ell_r-1} c_r, c_1. \end{aligned} \quad (3.9)$$

¹²In fact, if we compare (3.1) and (3.6) $k_i = \ell_{r+1-i}$, $a_i = c_{r+2-i}$.

In this sequence there are r places where we switch from c_i to c_{i+1} . If we rewrite the sequence (3.9) as

$$n_1, n_2, n_3, \dots, n_{\ell+1},$$

then

$$n_{j+1} \begin{cases} = 2n_j & \text{if there is no switch,} \\ \equiv -2n_j \pmod{b} & \text{at a switch, by (3.7).} \end{cases} \quad (3.10)$$

Since there are r switches, we conclude from (3.10) that

$$n_{\ell+1} \equiv (-1)^r 2^\ell n_1 \pmod{b}. \quad (3.11)$$

But $n_{\ell+1} = n_1 = c_1$, and c_1 is coprime to b . Thus (3.11) implies that

$$2^\ell \equiv (-1)^r \pmod{b},$$

which is (3.8).

To show that ℓ is the quasi-order of $2 \pmod{b}$, we must show that, for every positive $m < \ell$, the congruence

$$2^m \equiv \pm 1 \pmod{b} \quad (3.12)$$

is false. Now (3.12) implies, in the light of (3.10), that $n_{m+1} \equiv \pm c_1 \pmod{b}$, with $m+1 < \ell+1$. We first show that

$$n_{m+1} \equiv c_1 \pmod{b}$$

is impossible. Now, as we have remarked, it follows from the definition of the φ -function that all the numbers n_j in the sequence (3.9) satisfy $n_j < \frac{b}{2}$. Thus

$$n_{m+1} \equiv c_1 \pmod{b} \text{ implies } n_{m+1} = c_1.$$

But either n_{m+1} is even or it is some c_i different from c_1 . Thus, since the symbol (3.6) is contracted and, of course, c_1 is odd, $n_{m+1} = c_1$ is impossible.

Finally we show that $n_{m+1} \equiv -c_1 \pmod{b}$ is impossible. For $n_{m+1} + c_1$ is a positive integer less than b , so it is not divisible by b .

This completes the proof of the theorem. \square

The Quasi-order Theorem is striking in that, given b , we compute k starting with any a which is odd, less than $\frac{b}{2}$, and coprime to b . Of course, the choice $a = 1$ is always available and is the one to make when we seek folding instructions for producing a regular convex b -gon.

There are two rather remarkable examples of our symbol (3.1). One is

$$23 \left| \begin{array}{cccccc} 1 & 11 & 3 & 5 & 9 & 7 \\ 1 & 2 & 2 & 1 & 1 & 4 \end{array} \right|$$

telling us that

$$2^{11} \equiv 1 \pmod{23}, \quad \text{or} \quad 23 \mid 2^{11} - 1. \quad (3.13)$$

This is remarkable because $2^{11} - 1$ is a *Mersenne number*, that is, a number of the form $2^p - 1$, where p is prime. Abbé Mersenne hoped that all these numbers would be prime; but (3.13) shows that this is not so (of course, this was already known long before the invention of the symbol (3.1)).

The second example is even more remarkable; it is

$$641 \left| \begin{array}{ccccccccc} 1 & 5 & 159 & 241 & 25 & 77 & 141 & 125 & 129 \\ 7 & 2 & 1 & 4 & 3 & 2 & 2 & 2 & 9 \end{array} \right| \quad (3.14)$$

telling us that

$$2^{32} \equiv -1 \pmod{641}, \quad \text{or} \quad 641 \mid 2^{32} + 1. \quad (3.15)$$

This is striking result because $2^{32} + 1$ is a *Fermat number*, that is, a number of the form

$$2^{2^n} + 1.$$

The French mathematician Pierre Fermat hoped that all these numbers would be prime; but (3.15) shows that this is not so (the factorizability of $2^{32} + 1$ was, in fact, first noticed by the Swiss mathematician Leonhard Euler).

Our symbol seems only to give one factor of $2^{11} - 1$ or $2^{32} + 1$, and not the complementary factor. In Section 6 we will show that this is not so – we can get both factors from our symbol.

4. The Symbol-Creation Theorem

We now introduce a theorem that answers, very precisely, Question 2 of Section 2. This theorem has been strongly hinted at, for example, in Chapter 4 of [2] and Chapter 4 of [3], but it has not previously been explicitly stated, nor has the proof been given in detail.

Recall that, in Section 2, we introduced the idea of a *contracted* symbol. We will say that a folding instruction vector (k_1, k_2, \dots, k_r) is *contracted* if there

exists no proper factor s of r such that our vector arises by iterating the vector (k_1, k_2, \dots, k_s) .

We know that, given odd numbers b, a with $a < \frac{b}{2}$ there is a unique set of folding instructions to create a quasi-regular $\{\frac{b}{a}\}$ -gon. We now prove the converse.

THEOREM 4.1 (*The Symbol-Creation Theorem*). *Given the folding instruction vector (k_1, k_2, \dots, k_r) , there exists a unique reduced symbol*

$$b \left| \begin{array}{cccc} a_1 & a_2 & \dots & a_r \\ k_1 & k_2 & \dots & k_r \end{array} \right| \quad (4.1)$$

which is contracted if, and only if, (k_1, k_2, \dots, k_r) is contracted.

We first indicate the line of proof. At the outset we deal with the case $r = 1$, and then assume $r \geq 2$ thereafter. Given k , then

$$b \left| \begin{array}{c} a \\ k \end{array} \right|$$

is a reduced symbol if, and only if, b, a are odd, $a < \frac{b}{2}$, $\gcd(b, a) = 1$ and

$$b - a = 2^k a,$$

that is

$$b = (2^k + 1)a. \quad (4.2)$$

but (4.2) shows that $a \mid b$. Thus, since $\gcd(b, a) = 1$, we must have $a = 1$, $b = 2^k + 1$, and thus we have our unique reduced symbol (with $r = 1$)

$$2^k + 1 \left| \begin{array}{c} 1 \\ k \end{array} \right|. \quad (4.3)$$

Of course, since $r = 1$, (4.3) is contracted.

Thus we assume henceforth that $r \geq 2$. The line of proof is then as follows. We first suppose that (4.1) is a reduced symbol. Then we will show, in Section 6, that

$$Ba_i = bA_i, \quad (4.4)$$

where

$$B = 2^k - (-1)^r, \quad k = \sum_{i=1}^r k_i, \quad (4.5_1)$$

and, for each i ,

$$A_i = 2^{k-k_{i-1}} - 2^{k-k_{i-1}-k_{i-2}} + \dots + (-1)^r 2^{k_i} - (-1)^r. \quad (4.5_2)$$

Here, for convenience, we interpret the subscripts on the k 's to be written mod r to allow for uniformity of the formula. It is plain that B, A_i are odd; and it is not difficult to show that $A_i < \frac{B}{2}$. For, since each k_i is strictly positive, $B \geq 2^k - 1$ and $A_i \leq 2^{k-1} - 1$. Now since $b = a_i + 2^{k_i} a_{i+1}$, it follows from (4.4) that

$$B = A_i + 2^{k_i} A_{i+1}, \quad \text{for all } i. \quad (4.6)$$

Hence

$$\gcd(B, A_i) \quad \text{is independent of } i. \quad (4.7)$$

Notice that B, A_i are entirely determined by (k_1, k_2, \dots, k_r) . Thus if $\gcd(B, A_i) = q$, then q is entirely determined by (k_1, k_2, \dots, k_r) , and it follows from (4.4) that, in view of the requirement that b, a_i are coprime positive integers, b, a_i are given by the equations

$$B = qb, \quad A_i = qa_i. \quad (4.8)$$

This establishes the *uniqueness* of the reduced symbol (4.1) for a given (k_1, k_2, \dots, k_r) .

Now (4.6) is also an easy consequence of (4.5₁), (4.5₂). Thus we see that there is a symbol

$$B \left| \begin{array}{cccc} A_1 & A_2 & \dots & A_r \\ k_1 & k_2 & \dots & k_r \end{array} \right|.$$

It now follows from (4.8), regarded as the definitions of b, a_i , that (4.1) is also a symbol, which is reduced since, by (4.8) and $\gcd(B, A_i) = q$, we know that $\gcd(b, a_i) = 1$. We notice that $a_i < \frac{b}{2}$, since $A_i < \frac{B}{2}$. Thus we do have a reduced symbol (4.1) for a given (k_1, k_2, \dots, k_r) , proving the *existence* part of the statement of the theorem.

Of course if the folding instructions vector (k_1, k_2, \dots, k_r) is contracted, so is the symbol (4.1). Conversely, if (k_1, k_2, \dots, k_r) is formed by iterating (k_1, k_2, \dots, k_s) , then we may form the reduced symbol (using the *existence* statement above)

$$b \left| \begin{array}{cccc} a_1 & a_2 & \dots & a_s \\ k_1 & k_2 & \dots & k_s \end{array} \right|; \quad (4.9)$$

and we would obtain a reduced symbol (4.1) by iterating (4.9). By *uniqueness*,

this would have to be the (unique) reduced symbol for the given (k_1, k_2, \dots, k_r) ; and, by construction, it is not contracted. Thus the theorem would be proved, once we have established (4.4).

5. Some matrix calculations

Our plan is to establish (4.4) by solving the equations

$$a_i + 2^{k_i} a_{i+1} = b, \quad 1 \leq i \leq r \quad (5.1)$$

in the unknowns (a_1, a_2, \dots, a_r) , for fixed k_1, k_2, \dots, k_r, b . Our method of solution will be to apply *Cramer's Rule*, named after the Swiss mathematician Gabriel Cramer (1704–1752) who used it in a book on curves in 1750. This will involve calculating the determinants of certain matrices, and we devote this section to those calculations.

We generalize the algebraic problem (slightly!) and consider the $(r \times r)$ matrix¹³

$$M = \begin{pmatrix} 1 & \lambda_1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \lambda_2 & \dots & 0 & 0 \\ \vdots & & & & & \\ 0 & 0 & 0 & \dots & 1 & \lambda_{r-1} \\ \lambda_r & 0 & 0 & \dots & 0 & 1 \end{pmatrix}. \quad (5.2)$$

Our first result will be

$$\det M = 1 - (-1)^r \lambda_1 \lambda_2 \dots \lambda_r. \quad (5.3)$$

This is trivial if $r = 2$. Now, expanding by the first row in (5.2), we see that

$$\det M = 1 - \lambda_1 \det N, \quad (5.4)$$

where

$$N = \begin{pmatrix} 0 & \lambda_2 & 0 & \dots & 0 & 0 \\ 0 & 1 & \lambda_3 & \dots & 0 & 0 \\ \vdots & & & & & \\ 0 & 0 & 0 & \dots & 1 & \lambda_{r-1} \\ \lambda_r & 0 & 0 & \dots & 0 & 1 \end{pmatrix}. \quad (5.5)$$

¹³Recall that $r \geq 2$.

We may now calculate $\det N$ by induction on r , where we may assume $r \geq 3$. For, if $r = 3$, then

$$N = \begin{pmatrix} 0 & \lambda_2 \\ \lambda_3 & 1 \end{pmatrix},$$

so $\det N = -\lambda_2\lambda_3$. Thus our inductive hypothesis is that, in (5.5),

$$\det N = (-1)^r \lambda_2 \dots \lambda_r, \quad (5.6)$$

and this has been verified for $r = 3$.

Let us, temporarily, write $N(r)$ for (5.5), to emphasize that N depends on r . It is then plain that $\det N(r) = -\lambda_2 \det N(r-1)$, where

$$N(r-1) = \begin{pmatrix} 0 & \lambda_3 & 0 & \dots & 0 & 0 \\ 0 & 1 & \lambda_4 & \dots & 0 & 0 \\ \vdots & & & & & \\ 0 & 0 & 0 & \dots & 1 & \lambda_{r-1} \\ \lambda_r & 0 & 0 & \dots & 0 & 1 \end{pmatrix}.$$

By our inductive hypothesis, $\det N(r-1) = (-1)^{r-1} \lambda_3 \dots \lambda_r$, so that

$$\det N(r) = (-1)^r \lambda_2 \lambda_3 \dots \lambda_r$$

and (5.6) is established, whence (5.3) follows from (5.4). Notice that, if $\lambda_i = 2^{k_i}$, then, in (4.5₁),

$$B = (-1)^{r+1} \det M. \quad (5.7)$$

We next calculate $\det M_i$, where M_i is the matrix obtained from M by replacing its i -th column by the vector

$$\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}.$$

Due to the symmetry in the matrix M , it will suffice to calculate $\det M_1$, since we can then infer $\det M_i$ by “rotating the suffixes”.

Now

$$M_1 = \begin{pmatrix} 1 & \lambda_1 & & & & & & \\ 1 & 1 & \lambda_2 & & & & & \\ 1 & 0 & 1 & \lambda_3 & \mathbf{0} & & & \\ \vdots & & & & & & & \\ 1 & 0 & 0 & 0 & \dots & 1 & \lambda_{r-1} & \\ 1 & 0 & 0 & 0 & \dots & 0 & 1 & \end{pmatrix}.$$

As before, we write $M_1(r)$ for M_1 to emphasize the dependence on r and observe that, trivially, $\det M_1(2) = 1 - \lambda_1$. We will prove by induction on r that

$$\det M_1 = \det M_1(r) = 1 - \lambda_1 + \lambda_1 \lambda_2 - \dots + (-1)^{r-1} \lambda_1 \lambda_2 \dots \lambda_{r-1}. \quad (5.8)$$

For, expanding by the first row of M_1 , we see that

$$\det M_1(r) = 1 - \lambda_1 \det M_1(r-1), \quad \text{where}$$

$$M_1(r-1) = \begin{pmatrix} 1 & \lambda_2 & & & & & & \\ 1 & 1 & \lambda_3 & & & & & \\ 1 & 0 & 1 & \lambda_4 & \mathbf{0} & & & \\ \vdots & & & & & & & \\ 1 & 0 & 0 & 0 & \dots & 1 & \lambda_{r-1} & \\ 1 & 0 & 0 & 0 & \dots & 0 & 1 & \end{pmatrix}.$$

Thus, by the inductive hypothesis,

$$\det M_1(r-1) = 1 - \lambda_2 + \lambda_2 \lambda_3 - \dots + (-1)^{r-2} \lambda_2 \lambda_3 \dots \lambda_{r-1},$$

so that

$$\begin{aligned} \det M_1 &= 1 - \lambda_1 (1 - \lambda_2 + \lambda_2 \lambda_3 - \dots + (-1)^{r-2} \lambda_2 \lambda_3 \dots \lambda_{r-1}) \\ &= 1 - \lambda_1 + \lambda_1 \lambda_2 - \dots + (-1)^{r-1} \lambda_1 \lambda_2 \dots \lambda_{r-1} \end{aligned}$$

and (5.8) is established. Notice that, if $\lambda_i = 2^{k_i}$, then, in (4.52),

$$A_1 = (-1)^{r+1} \det M_1,$$

and, more generally,

$$A_i = (-1)^{r+1} \det M_i. \quad (5.9)$$

6. Completing the proof of the Symbol-Creation Theorem

It remains only to use our calculations in Section 5 to prove (4.4). We will then have completed the proof of our main result, Theorem 4.1 (The Symbol-creation Theorem). As forecast at the start of Section 5, we proceed to solve the set of simultaneous equations

$$a_i + 2^{k_i} a_{i+1} = b, \quad i = 1, 2, \dots, r \quad (a_{r+1} = a_1) \quad (6.1)$$

using *Cramer's Rule*. This will involve us in, first, constructing the matrix M of (5.2) with $\lambda_i = 2^{k_i}$ and then calculating $\det M$, which we know (5.7) to be given by

$$\det M = 1 - (-1)^r 2^k = (-1)^{r+1} B. \quad (6.2)$$

Since $\det M \neq 0$, we know that the set of simultaneous equations (6.1) has a *unique* solution in the field \mathbb{R} of real numbers. Of course, the solution depends on b ; indeed, to find the solution we must construct, for each i , the matrix $M^{(i)}$ obtained by replacing the i th column vector of M by the vector

$$\begin{pmatrix} b \\ b \\ \cdot \\ \cdot \\ \cdot \\ b \end{pmatrix}.$$

The unique solution is then given, by Cramer's Rule, by the equations

$$\frac{a_1}{\det M^{(1)}} = \frac{a_2}{\det M^{(2)}} = \dots = \frac{a_r}{\det M^{(r)}} = \frac{1}{\det M}. \quad (6.3)$$

But plainly $\det M^{(i)} = b \det M_i = (-1)^{r+1} b A_i$, by (5.9). Thus we may rewrite (6.3) as

$$\frac{a_i}{b A_i} = \frac{1}{B}, \quad 1 \leq i \leq r,$$

or

$$B a_i = b A_i, \quad i = 1, 2, \dots, r. \quad (6.4)$$

This proves (4.4) and completes the proof of Theorem 4.1.

We close this section with an example of the theorem.

Example 6.1. Consider the folding instruction vector $(1, 2, 3, 4)$. Thus $r = 4$, $k_1 = 1$, $k_2 = 2$, $k_3 = 3$, $k_4 = 4$, $k = 10$. Hence $b = 2^{10} - 1 = 1023$. Moreover, $A_1 = 2^6 - 2^3 + 2 - 1 = 57$, so that $q = \gcd(1023, 57) = 3$, whence, by (4.8), $b = 341$, $a_1 = 19$. We may calculate a_2, a_3, a_4 similarly. Thus

$$A_2 = 2^9 - 2^5 + 2^2 - 1 = 483, \quad \text{so} \quad a_2 = 161,$$

$$A_3 = 2^8 - 2^7 + 2^3 - 1 = 135, \quad \text{so} \quad a_3 = 45,$$

$$A_4 = 2^7 - 2^5 + 2^4 - 1 = 111, \quad \text{so} \quad a_4 = 37,$$

and the unique reduced (contracted) symbol with these folding instructions is

$$341 \left| \begin{array}{cccc} 19 & 161 & 45 & 37 \\ 1 & 2 & 3 & 4 \end{array} \right|. \quad (6.5)$$

We draw attention to two important features of our theorem, exemplified in the example.

Remark 1. The method of proof of the theorem tells us just what (star) polygons we obtain with the given folding instructions, namely, regular $\left\{ \frac{b}{2^c a_i} \right\}$ -gons, where a_i , $1 \leq i \leq r$, is calculated as in Section 4 and c is chosen so that $b > 2^{c+1} a_i$. This means, in our example, the folded strip would enable us to construct, using the FAT algorithm, regular

$$\left\{ \frac{341}{19} \right\}^-, \left\{ \frac{341}{38} \right\}^-, \left\{ \frac{341}{76} \right\}^-, \left\{ \frac{341}{152} \right\}^-, \left\{ \frac{341}{161} \right\}^-, \\ \left\{ \frac{341}{45} \right\}^-, \left\{ \frac{341}{90} \right\}^-, \left\{ \frac{341}{37} \right\}^-, \left\{ \frac{341}{74} \right\}^-, \left\{ \frac{341}{148} \right\}^- \text{-gons.}$$

Remark 2. The procedure given for calculating a_1, a_2, \dots, a_r is not the easiest in practice. For, once we have used it to calculate b and a_1 , it is far simpler just to calculate the symbol in the usual way from the data

$$b \left| \begin{array}{cccc} a_1 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{array} \right|. \quad (6.6)$$

This, of course, also provides us with a check on our calculations, since from (6.6) we should recover the given vector (k_1, k_2, \dots, k_r) of folding instructions. The reader may like to try this with $b = 341$, $a_1 = 19$, which should yield the vector $(1, 2, 3, 4)$.

7. An application

Consider our example, in Section 3, of the symbol

$$641 \left| \begin{array}{cccccccc} 1 & 5 & 159 & 241 & 25 & 77 & 141 & 125 & 129 \\ 7 & 2 & 1 & 4 & 3 & 2 & 2 & 2 & 9 \end{array} \right|. \quad (7.1)$$

We saw that this leads, by the Quasi-order Theorem, to the remarkable result that

$$641 \mid 2^{32} + 1,$$

and $2^{32} + 1$ is the Fermat number $F_5 = 2^{2^5} + 1$, which is thus shown not to be prime. The natural question at this point is – what is the complementary factor? This can, of course, in principle be obtained by simply dividing $2^{32} + 1$ by 641 – one might call this the *elementary* or *classical* method, but it is not an attractive method. Can we not use the symbol (7.1) to find a superior method? The answer is that we can.

If we return to the general case of the symbol

$$b \left| \begin{array}{cccc} a_1 & a_2 & \dots & a_r \\ k_1 & k_2 & \dots & k_r \end{array} \right| \quad (7.2)$$

then we know, by (4.8), that

$$B = 2^k - (-1)^r = qb \quad \text{and} \quad A_i = qa_i, \quad i = 1, 2, \dots, r. \quad (7.3)$$

It follows from the first equation in (7.3) that

The factor of $2^k - (-1)^r$ complementary to b is q .

Hence we need to calculate q . As we have said, the elementary method is to divide $2^k - (-1)^r$ by b ; but the second set of equations in (7.3) shows that we may instead divide A_i by a_i for any value of i ! This method is far more attractive since the numbers A_i are much smaller than B and the numbers a_i are less than one half of b .

This new method, which we may describe as the *paper-folder's method*, by contrast with the *elementary method*, is particularly attractive if $a_i = 1$ for some value of i (as in our example). For if $a_i = 1$ then $q = A_i$. Thus, in our example, the complementary factor is

$$A_1 = 2^{23} - 2^{21} + 2^{19} - 2^{17} + 2^{14} - 2^{10} + 2^9 - 2^7 + 1, \quad (7.4)$$

which is fairly easily calculated to be 6, 700, 417, and thus we know that

$$F_5 = 641 \times 6, 700, 417. \quad (7.5)$$

Notice that the biggest number we have to deal with in calculating $q (= A_1)$ in this example is 2^{23} which is smaller than $2^{32} + 1$ by a factor of approximately $2^9 = 512$. Of course, we can use further tricks to simplify the calculation of A_1 in (7.4). For example, it is easy to see from (7.4) that

$$A_1 = 3 \times 2^{21} + 3 \times 2^{17} + 15 \times 2^{10} + 3 \times 2^7 + 1, \quad (7.6)$$

providing an immediate simplification.

Given the symbol (7.2) there are two criteria to adopt in deciding which value of i to choose in order to calculate q :

- (i) choose the smallest a_i , so that the division of A_i by a_i is as easy as possible – in particular, if some $a_i = 1$, as in our example, choose that value of i ;
- (ii) choose i so that k_{i-1} is as big as possible, so that the leading power of 2 occurring in the expression for A_i is as small as possible.

It will often occur, as in our example, that these two criteria lead to the same choice¹⁴ of i .

Implicit in this discussion is, as before, the presence of an effective check on our calculation of the quasi-order of 2 mod b , since we have (at least) two independent procedures for calculating q .

¹⁴We have no counterexample at our fingertips.

References

- [1] H. S. M. Coxeter, *Regular Polytopes*, Macmillan Mathematics Paperbacks, New York, 1963.
- [2] P. Hilton, D. Holton and J. Pedersen, *Mathematical Reflections – In a Room With Many Mirrors*, 2nd printing, Springer, New York, 1998.
- [3] P. Hilton, D. Holton and J. Pedersen, *Mathematical Vistas – From a Room With Many Windows*, Springer, New York, 2002.
- [4] P. Hilton and J. Pedersen, Folding regular star polygons and number theory, *Math. Intelligencer* **7**, no. 1 (1985), 15–26.
- [5] P. Hilton and J. Pedersen, Certain algorithms in the practice of geometry and the theory of numbers, *Publ. Sec. Mat. Univ. Autonoma Barcelona* **29**, no. 1 (1985), 31–64.
- [6] P. Hilton and J. Pedersen, Geometry in Practice and Numbers in Theory, *Monographs in Undergraduate Mathematics* **16** (1987), 37, (Available from Department of Mathematics, Guilford College, Greensboro, North Carolina 27410, USA).
- [7] P. Hilton and J. Pedersen, On the complementary factor in a new congruence algorithm, *Int. Journ. Math. and Math. Sci.* **10**, no. 1 (1987), 113–123.
- [8] P. Hilton and J. Pedersen, *Build Your Own Polyhedra*, Dale Seymour Publications, (reprinted 1994).

PETER HILTON
DEPARTMENT OF MATHEMATICAL SCIENCES
SUNY BINGHAMTON
BINGHAMTON, NEW YORK 13902-6000
USA
AND
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF CENTRAL FLORIDA
ORLANDO, FLORIDA 32816-1364
USA

JEAN PEDERSEN
DEPARTMENT OF MATHEMATICS AND COMPUTER SCIENCE
SANTA CLARA UNIVERSITY
SANTA CLARA, CALIFORNIA 95053-0290
USA

E-mail: jpedersen@scu.edu

(Received January 25, 2002)