CONJECTURES ON THE PERCEPTION OF ELONGATION

by

E. S. Deutsch

Computer Science Center
University of Maryland
Colleqe Park, Maryland
U.S.A.

ABSTRACT

The papers start by discussinq some of the background to the problem of the perception of length together with some known illusions in this field. The difference is then drawn between image contraction and image thinning, both processes of which serve to introduce the proposed conjecture. The suggested conjecture is that lenqth is perceived by a mechanism which first forms a minimal closed path (hull) around the object and then performs an isotropic image reduction operation on the hull of the image. It is conjectured that the elongation of the image is exemplified in the resulting path so obtained; no path should result if the original image possesses no elongation. The conjecture is used to explain the known illusions discussed and a few new ones are suggested.

An isotropic image reduction technique is described and experimental results testing the conjecture, obtained using a variety of imaqes, are presented.

INTRODUCTION

In this paper we will address ourselves to the phenomenon of length perception in qeneral. In the first instance an attempt will be made at distinguishing between two types of image reduction operations, namely that of image contraction and that of image thinning or skeletonization. While both of these operations appear to be similar, the distinction we propose to make between their end results leads to the possibility of a conjecture upon how the human eye perceives whether a given image, whatever its shape, possesses the property of elongation or not. The author is mindful of the fact that his proposed conjecture falls short of explaininq all the different types of experienced phenomena existinq in the field of lenqth perception. However, some interesting results have been obtained on applying the ideas to images in general and to some of the optical illusions in particular.

It seems clear that for an image reduction process to be meaninqful it must exhibit isotropic properties. In other words unless one is specifically interested in the reduction operations exercising a certain bias in a particular direction (1) (see Fiqure 4, it will be explained in more detail later on), the reduction must take place uniformly throughout the entire imaqe. We shall indeed describe such an isotropic process in terms of a computer algorithm; it is the *Very* existence of such a process that gives substance to our conjecture.

Previous work of the author indicates that there are considerable advantages in working with images displayed on hexaqonal arrays (2). Accordinqly, all processes to be described here will utilize such arrays. This apparent restriction does not limit the generality of the main discussion in any way; it may well emerge that what follows lends additional weiqht to the arqument that hexaqonal arrays are the most suitable for imaqe processing in qeneral.

The layout of the remainder of this paper is as follows. We shall first arque the distinction between the two types of image reduction operations mentioned above. A further related discussion then leads to the introduction of the conjecture. The implementation of an isotropic imaqe reduction technique, the use of which has already been implied, is presented next. The compatibility of the conjecture with some of the observed optical illusions is then discussed. An attempt at introducinq new optical illusions in light of the conjecture is also made.

IMAGE REDUCTION: IMAGE CONTRACTION AND IMAGE THINNING

We introduce the distinction between image contraction and imaae thinning by means of an illustration. Assume that the imaqe to be contracted or thinned is a solid disc of uniform qray level distribution. For the present arqument we shall assume that the center of the disc coincides with a point on the display matrix. It will become apparent at once that if reduction takes place *uniformly* throuqhout, and, that provided continuity is maintained, the result of this imaqe contraction operation should reduce the disc to a single dot. An imaqe thinning operation should qive the same result under similar conditions. However, it is quite unnatural to suggest that a thinned disc is a dot. A similar arqument applies to the case when we consider the imaqe to be a uniform square. If the point of intersection of the diagonals coincides with a point on the matrix then the result of both a contraction and a thinninq or skeletonizing operation will be that single dot on the matrix. Aqain, it is quite in order to call that point, in the limit, the contracted form of the original square, but that point is not the skeleton of the square. It is in fact questionable whether the term thinninq or skeletonizing operation is applied to two dimensional imaqes similar to those just described is at all applicable.

Let us now apply a deformation to the disc so that for a qiven set of perpendicular diameters, one diameter is very sliqhtly shorter than the other, such that the end points of this diameter are equi-distant from the center of the disc. The disc deforms into an ellipse. Consider the result of contraction and thinnina ooerations on the deformed disc, as the ratio of the lenqths of the lonqer to the shorter diameter increases. When the ratio of these principle axes is near one or when ø (ø= eccentric angle) 1s near II/2 both these operations will yield an image in the form of a single point. However, as ø decreases a continuity of points emerges i.e., a skeleton. The orientation of this skeleton will be in the direction of the major axis of the particular ellipse. The number of points emerging will be inversely proportional to the value of p. In these circumstances 1t becomes possible to speak more specifically of a thinning or skeletonizing operation rather than a mere contracting operation, for the contraction operation may or may not yield a skeleton dependinq upon its design.

A similar argument holds with reference to the square being deformed, as one of its perpendicular axes is curtailed. A skeleton will emerge as the ratio of the longer to the smaller axis increases, the number of points in the skeleton being proportional to this ratio. Note that in both of the examples just given, the skeleton emerged along the longer center line of the objects, that is, it coincided with a geometric axis of the object.

In both the cases just discussed, the deformation produced a degree of elongation - if only in the intuitive sense - such elongation being registered with the emergence of a skeleton. Our intention, in part, is to generalize this notion together with its converse and suggest that the existence of elongation within an image implies the emergence of a skeleton and vice versa. In other words, given that an image reduction operation is applied to an image; should a skeleton emerge then it can be said that the original image exhibited elongation (i.e., image thinninq has occured). Should, however, only a sinqle point or a clustering collection thereof be obtained then no such property is attributable to the original image (i.e., image contraction has occurred). This would then lead to a qualitative rather than an intuitive definition of elongation.

The distinction we have just proposed is not merely one of semantics. The distinction is between kinds of end results obtained from an image reduction operation; a mere contracted image for some types of images and a skeleton for others.

An additional factor involved in the above speculation would be the information borne by the actual skeleton itself. Intuitively speaking, the degree of elongation exhibited by an ellipse of a given eccentric angle $ø_2$ would exceed that of another ellipse of eccentric angle 0, if $ø_1 > ø_2$. (Assume that their major axes are of equal length.) Application of a reduction operation to two such ellipses should, provided our previous discussion is valid, yield two skeletons (or a point and a skeleton), one of which, that derived from the" ellipse of $ø=ø_2$ being longer in dot count than the other. In Figure 1 are shown two such ellipses their skeletons, superimposed, having been obtained usinq the image reduction mechanism which is described later.

We are thus proposinq both a qualitative and a quantitative measure of elongation within an image. In summary, the production of a skeleton by a uniform reduction operation, siqnifies the existence of the elonqation property within the imaqe, the actual number of points forminq the said skeleton being an estimate of its degree of elonqation.

The discussion so far has been restricted to images of regular shape. What about the elicitation of the elonqation property of irreqular shapes? Consider the images shown in Figure 2. An isotropic imaqe reduction mechanism applied to the riqht-hand imaqe would yield two almost perpendicular skeletons. Yet this imaqe clearly exhibits elongation in the vertical sense. Further remarks concerning this difficulty are deliberately postponed until after the ensuing discussion, whence its solution will become apparent. We now return to our proposition reqarding the qualitative measurement of lenqth.

It would be interesting to see how compatible our proposition regarding the attainment of a qualitative estimate of length, is with some of the observed visual phenomena. We have seen Its application in Fiqure I in determining the elongation of two ellipses, and the same argument is applicable to rectanqles of differing dimensions. Yet would the proposition apply to line drawinqs? Consider the Mueller Lyer illusion, Figure 3(a), in which the top vertical line looks lonqer, although the vertical members of each of the drawinqs are of the same length. Furthermore, consider the two rectanqles and the two sections in Figure 3(b) and 3(c) respectively. They are of the same lenqth in each case, yet the upper member appears slightly longer. It was considerations such as these which led one to speculate on how elonqation, in qeneral, is perceived. To assume that there are a number of lenqth perceivinq facilities within the human mind each of which appertains to a different kind of image, say, regular shaped ones, stroke comprising images, etc., is not impossible; it can be neither substantiated nor denied in view of our scanty knowledqe about human perception. It would be exceedingly useful to arrive at a qeneralized mechanism by means of which the elonqation of any image is perceived.

Accordingly, we propose the following conjecture on the overall perception of lenqth. That the elonqation property of an image is perceived by the eye (?) first forminq the (convex) hull around the image and then, treating the hull as an imaqe with uniform qray-level distribution, applyinq isotropic imaqe reduction to this image. By the arguments above the resulting skeleton exemplifies the elonqation both in size and direction.

The conjecture states that the residual skeleton obtained from an isotropic image reduction operation yields a meausre of elonqation. The skeleton itself is derived by drawina the envelope around the object whose length is to be measured. It is the length of the resulting skeleton rather than the extent of the envelope which indicates object elongation. (One can introduce many small undulations in the contour of the envelope thus increasing its extent, without altering the final result.) Once the envelope is formed and the enclosure is 'filled-in', the skeleton should lie along, or *very* close to the longest center line of the 'filled-in' object. This, provided that the reduction takes place isotropically.

The solution to the problem relating to the images in Figure 2 should now be evident. For the formation of the envelope around each of these images yields objects whose longest axis is near vertical. The resulting skeleton accordinq to the conjecture closely approximate this axis.

For the conjecture to be meaningful it must, not only comply with ordinarily observed phenomena but also, cope with, and possibly give rise to the creation of examples of such phenomena. We will devote our attention to both these aspects later. For the moment we digress in order to discuss the concepts of an isotropic image reduction mechanism. We will return to the main problem later on. In the next few paragraphs we review the relevant literature.

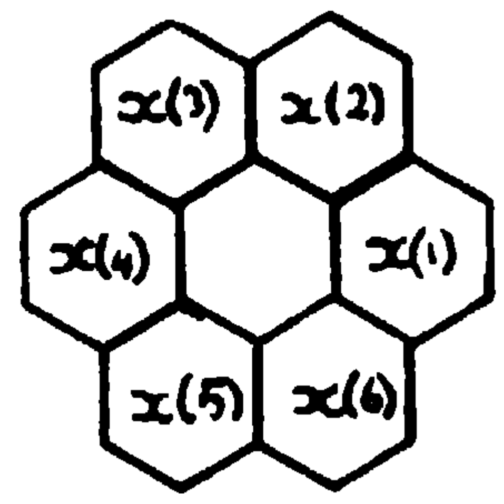
It is pointed out that in no way does our previous discussion imply that because all regular objects are contractable to a dot, that they will be recognized as being identical. The individual shapes of such objects could be identified as being different by means of an edge-detection mechanism, a visual facility for whose existence there 1s considerable evidence. It would Indeed be interesting to evaluate experimentally the scope of an image reduction mechanism together with an edge-detection process.

A very interesting explanation has been proposed by Blum (4) who has used his Medial Axis Transformation to explain some of the optical illusions. The skeletons, termed Medial Axis Functions, produced by this transformation differ from those obtained using the algorithm presented in the next section. His argument, however, is also based upon the determination of the properties of the skeleton, i.e., the Medial Axis Functions. He suggests that the visual process tries to line up these Medial Axis Functions rather than the outline of the objects themselves; it is the speed of formation of these Functions which gives vise to the observed illusions. Figure 3(d) depicts the formation of the Functions for Figure 3(a), the velocity of their generation is represented by, and is proportional to the thickness of the dots. The Functions associated with the converging arrows figure coalesce faster thus producing the illusion that this vertical line is shorter.

TOWARDS AN ISOTROPIC IMAGE REDUCTION PROCESS

We have already mentioned that we will concern ourselves with hexagonal display structures only. Below we Dresent an isotroDic image reduction algorithm appertainlnq to such arrays. For a fuller treatment the reader is referred to (1).

Consider the six peripheral neiqhbors of any element x on the hexaqonal arrav, labelled x(l) throuqh x(6) as shown. Since we are dealIng with images of uniform gray-level distribution, the



value of any element in the array can be either l (qray image) or zero (background). We shall also refer to the value of an element by its name and say that x(l), i.e., element x(1) has the value of either x(l) = 1 or 0, x(1) = 0. This will not lead to a confusion between the name and the value of an element since the meaning will be clear from the context.

Given below are the conditions, all of which must be fulfilled, under which the element x is deletable from the original image durinq the image reduction process.

1. $\sum_{n=1}^{6} x(n) \geq 2$

2. $x(p)\cdot x(p+1)\cdot x(p-1) = 0$

3. $x(p)\cdot x(p+1)\cdot x(p+2) = 0$

4. $\sum_{n=1}^{6} |x(n) - x(n+1)|_x = 2$

5. $x(p) = 1 \ \wedge \ \sum_{n=1}^{6} |x(n) - x(n+1)|_{x(p)} \neq 2$

where p is fixed at 1, or 2 or ....6, and n has a period of 6.

The subscript to the right vertical bar 1n conditions 4 and 5 signifies the name of the x element the peripheral elements of which are involved in the summation.

Each element is scrutinized independently of all the others; the process takes place in parallel. This process, which maintains connectivity, is carried out until no further erasure occurs. Condition 1 preserves end points of the skeleton, conditions 2 and 3 cause the deletion of certain points whose neighbors observe certain connectivity situations, condition 4 preserves the points in on an already derived skeleton and condition 5 preserves intersection points.

The six-way symmetry of the hexagonal array lends generality to the above conditions and renders the latter independent of position within the array of peripheral elements, in the sense that the conditions "look the same" whatever fixed value is assigned to p.

Consider the skeleton of the letter A shown in Figure 4(a). It was derived using one value of p, p = 1, throughout each cycle. The bias of the skeleton towards the lower left portions of the original image is evident. The skeleton will occupy different positions with respect to the original image for different values of p chosen, and it 1s this facet which facilitates isotropic reduction. For if the value of p is continually changed during the reduction process so as to compensate for a bias in a particular direction, then hopefully, a uniform reduction will result.

Whereas the skeleton in Figure 4(a), obtained with p ^{s} 1, is confined to the lower left portions of the image, the skeleton obtained using the value p * 4 was found to occupy the upper right portion of the image (approximately). Indeed, experimental evidence seems to indicate that the bias produced as a result of using the value p = p. during one pass over the image would be neutralized, during a followinq pass, usinq the value p = p_1+3. Thus, in order to attain isotropic reduction, one would simply perform each pass over the image with a different value of p thus making sure that any inbalance caused during one pass is taken care of in the next. More, or less, than six passes may be necessary but the sequences of the values of p would consist of a repetitive group of six differing p values: - P1P2P3P4P5P6.

A factor influencing the final position and shape of the skeleton within the image is the initial value of p and the remaining values of p in the sequence. We have found that smoother and more centrally disposed skeletons were obtained if the contraction operations were actually performed six times with p. as the initial value of p at the beginning of the ith pass followed by the remaining p values in order. Thus using the basic sequence P1P2p3P4P5P6. P1 would be the first value of p followed by $p_1 \ldots p_6$ 1n the first pass; p_2 would be the first value of p at the beginning of the second round followed by $P_3 \ldots p_6 p$, in that round. The third time round the sequence would be $p_3 \ldots p_6 p_1 p_2$ and so on.

The resulting skeleton is obtained by forming the logical OR of the previous six skeletons and repeating, etc. In practice, the following sequence yielded best results: - p_1 = 1, p_2 = P_1+3, p_3 = 2, p_4 = P_3+3, etc. Figure 4(b) shows the result obtained usinq this sequence; the skeleton is centrally disposed alonq the limbs of the letter. Note that where there is no true center line, that is, where the local part of the limb is an even-number of elements thick, two contiguous skeleton points are generated. Here, the true qeometric center line is situated between these two points. We return now to the main discussion.

THE CONJECTURE - DISCUSSION

In our attempt at establishina credence in our conjecture we shall try to interpret some of the optical illusions usinq the conjecture, and at the same time, suqqest how some new ones might be constructed. We shall be investiqatinq some of the better known illusions, and in doing so, present some of the theories suqqested by others.

The most popular illusion 1n the field of lenqth perception is the Mueller-Lyer lines, or the "arrow" illusion, shown in Fiqure 3(a). Why indeed should the vertical line flanked by the outqoinq arrow heads appear lonqer than an equally long vertical line with inqoinq arrow heads? The most recent attempt at explaining this phenomenon has been made by Greqory (3). It is suqqested that the illusion is produced by our general perception of depth in three dimensions, and by virtue of the fact that there exists an interesting inverse relation - known as Emmert's Law - between the size of an object and its distance from the viewer. According to this theory, the line with the outgoing arrows is beinq associated with the three dimensional representation of a far corner, whereas the line with the ingoing arrows is reminiscent of the representation of a nearer corner. Neglectinq the lengthy discussion, it is finally arqued that the inverse relation causes the image which "looks" as thouqh it were furthest away, (the line representing the far corner) to appear laraer.

The same argument is used to explain why, in the "railway-line" illusion, the upper horizontal rectanqle, Fiqure 3(b), looks longer than the lower one. The two converqinq lines introduce perspective effects and hence the top rectanqle, which appears now further away, is magnified by virtue of the inverse relation. Support for this approach is provided by some experimental evidence. The illusion shown in Figure 3(c), the Jastrow illusion, is not discussed.

We now discuss the arrow illusion in the light of our conjecture. The two lines are displayed as thouqh they were reproduced on an hexagonal array in Figure 5(a). By our conjecture, the lenqth of an object 1s perceived by the eye, in its overall view, forming an imaginary closed hull around the object and then "fillinq in" the entire enclosure, thus treatinq it as a uniform image, Figure 5(b). For the proposed conjecture to be valid, the application of isotropic image reduction to these images should yield two skeletons, the larger one of which should be derived from the imaqe representing the line with the outgoinq arrows. The result of applying the alqorithm discussed in the previous section is shown in Figure 5(c). (The arms of the ingoinq arrows are sliqhtly longer, this is because the difference in shape between the two uniform images in Figure 5(b) would hardly be noticable otherwise.) It is seen that this is in fact the case.

Two arguments in favor of the way our conjecture treats the above illusion are now presented. The first deals with the possibility of accomplishing this very illusion in a different way. Consider the two vertical lines shown in Fiqure 6. Which one appears longer? If these two lines are just another way of representing the arrow illusion then it must also be explicable in terms of our conjecture. Now, it is easy to see that the hulls of both the images of Figure 6 *are* rectangles in which the two shorter sides have been made circular. Isotropic image reduction will yield two skeletons of different lengths and hence the illusion. It would be interesting to see how the above notions of three dimensional depth perception and the inverse size relation would cope with this representation.

The second argument runs as follows. Assuming our conjecture to be valid, is it possible to find an "arrow" whose hull would yield an image which, in turn, would be reducible to a skeleton whose size ranked between those of the other two? Should such an "arrow" be found, then it too must form part of the illusion in that it must appear to be shorter than the line with the outqoing arrows and longer than the line with the ingoing arrows. We suggest that the third "arrow" shown in Figure 7(a) is such a possibility. The author has not as yet performed extensive tests with these drawings, but tentatively, of the numerous persons questioned almost all without exception ranked the arrows 1n the predicted order. The result of applying the reduction operation to these arrows is shown in Figure 7(b).

We now draw our attention to the two vertical lines shown in Figure 8(a). The upper line, with the smaller horizontal line throuqh it, appears slightly longer. This too can be explained by means of our conjecture, as Figure 8(b) indicates. Figure 8(b) also explains why generally, a narrow rectangle of a given length qives the impression of looking slightly longer than a rectangle of the same length but of increased width, (c.f., our early discussion regarding the skeletons obtained from rectangles of varying side ratios.)

Our conjecture does not explain the "railway line" illusion of Figure 3(b) and the Jastrow illusion of Figure 3(c) fully. We feel however, that it has a bearing on the phenomenon. Treatinq the "railway" illusion first; the reason the upper rectangle appears longer is because there is a confusion between the actual boundary of this rectangle and the exceedingly close local boundary dictated by the converging lines. The eye in forming the enclosure, includes the outer lines too, which by our conjecture, results in the upper rectangle looking longer. No such confusion arises in the case of the lower rectangle since the distance between the railway lines and the actual rectangle is fairly large. It is also noteworthy that the illusion only takes place provided one of the objects is very near the converging lines; the illusion will not occur if both objects are drawn far away from them.

A convenient way of testinq this idea would be to ascertain whether an increase 1n the width of the upper rectanqle (or a decrease in the width of the lower rectanqle) causes the illusion either to be less pronounced or to disappear entirely. By a previous arqument, the apparent decrease 1n length caused by increasinq the width of the upper rectangle (or the apparent increase in lenqth caused by decreasing the width of the lower rectanqle) should comoensate for the elonqa tion influence produced by the two bounding lines. Preliminary tests have certainly shown that alterinq the width of one of the rectangles reduces the degree of lenqth discrepancy observed between the rectanqles.

We now discuss the bearing our conjecture has upon the Jastrow illusion. This illusion is reproduced in Figure 9 with the two sections "filled in". The two vertical lines complete the hull of the entire image, the remainder of the hull beinq formed by the edqes of the sections themselves. The bottom section apoears as if 1t were diverqing from the boundary and at the same time converging into the space between the two lines. Note that this effect does not arise in the case of the top section because the corresponding Dortion which gives rise to this impression actually forms part of the enclosure. It is this diverqing effect, interacting between section and boundary which gives rise to the section lookinq smaller, locally.

This arqument does not refute our comments concerning the "railway" illusion, since in the case of the Jastrow illusion, the divergent effect is overwhelming. We have tried the railway illusion using inverted triangles instead of the rectanqles. The trianales were drawn so that their apex pointed away from the point of intersection of the converqinq lines. This was done in order to obtain the diverginq effect of the Jastrow illusion. The illusion seems to disappear under these circumstances (see Fiqure 10). Compare the impression obtained from Fiqure 3(b) with that obtained from Fiqure 10.

Finally we present an instance which seems to contradict the conjecture. Consider the two images with their unequal vertical lines shown in Figure 11. The extremities of the squares flankinq both of these lines are the same. The envelope of each of the images 1s therefore the same; a rectangle whose shorter sides corresponds to the outer side of the small squares and whose larqer sides are equal to the distance between these squares' outer side. The con lecture implies that since the resultinq skeletons for each of the hull is of the same length, the elonqation, or rather the extent of the vertical lines should appear similar. Yet this is not what is observed.

CONCLUSION

As usual, the kind of material described in this paper is particularly vulnerable to criticism and counter arguments. The author is fully aware of this and particularly so on the following two counts. First, because most of the material discussed is, to the author's knowledge, new, and second, because the author 1s knowingly encroaching upon a domain which is a little alien to his own.

Chronologically, an isotropic image reduction technique was all that we were after originally. This then gave rise to the speculations we described above. The fact that we could actually simulate some of the ideas on a computer seemed particularly attractive. It is of course realized that the investigations have been limited in their scope - for example, we have not mentioned the Herring illusion at all - but they seem interesting enough to warrant further attention.

REFERENCES

(1) Deutsch, E. S., Towards Isotropic Image Reduction, University of Maryland, Technical Report, No. TR-70-139, October 1970.

(2) Deutsch, E. S., Thinning Algorithms for Rectangular, Hexagonal and Triangular Arrays, University of Maryland, Technical Report, No. TR-70-115, June 1970.

(3) Gregory, R. L., Eye and Brain, World University Library, McGraw Hill Book Co.: New York 1966.

(4) Blum, H., A Transformation for Extracting New Descriptors of Shape in Models for the Perception of Speech and Visual Form, Proceedings of a Symposium, Weiant Wathen-Dunn, Editor, M.I.T. Press, Cambridge, 1967.

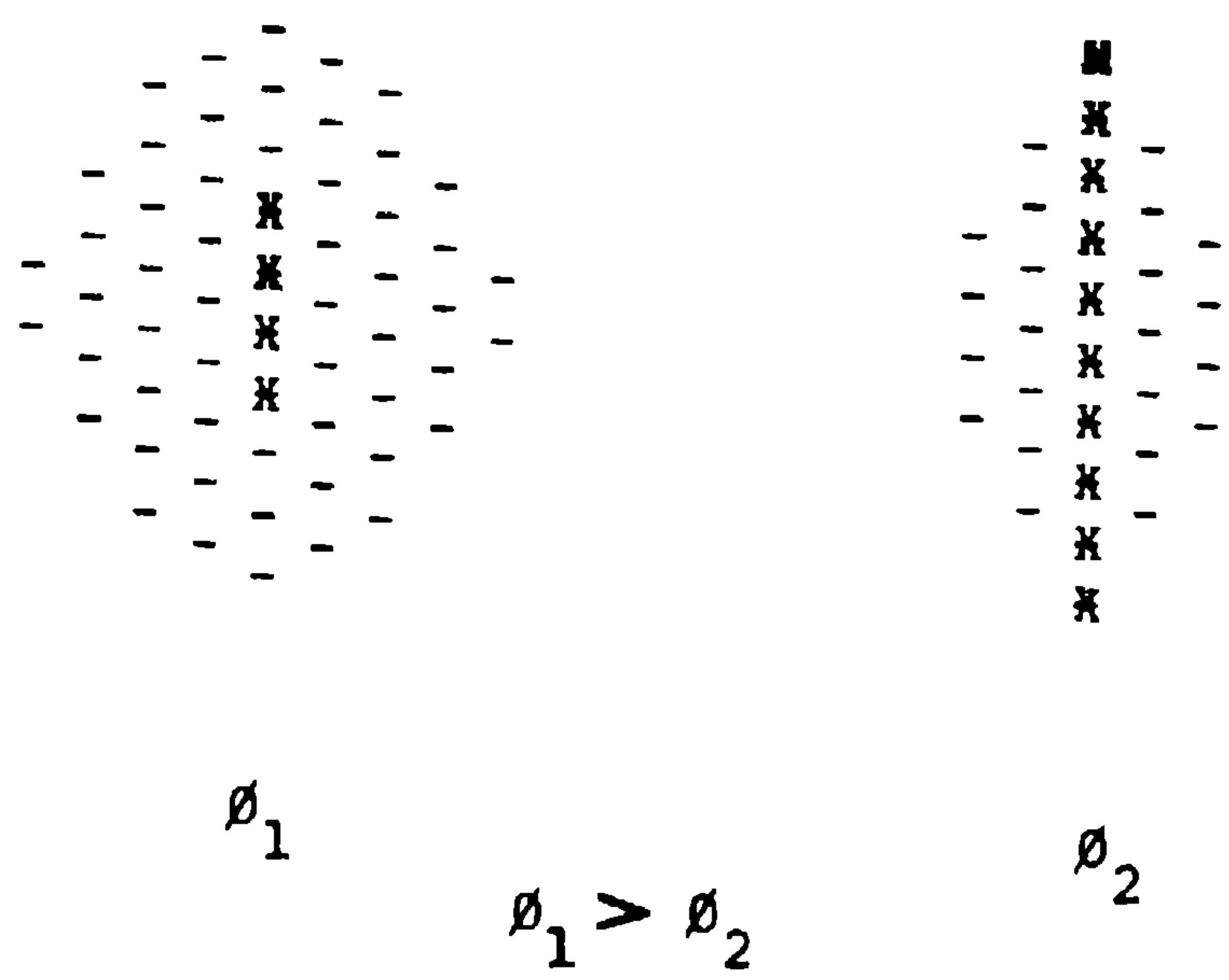


Figure 1.

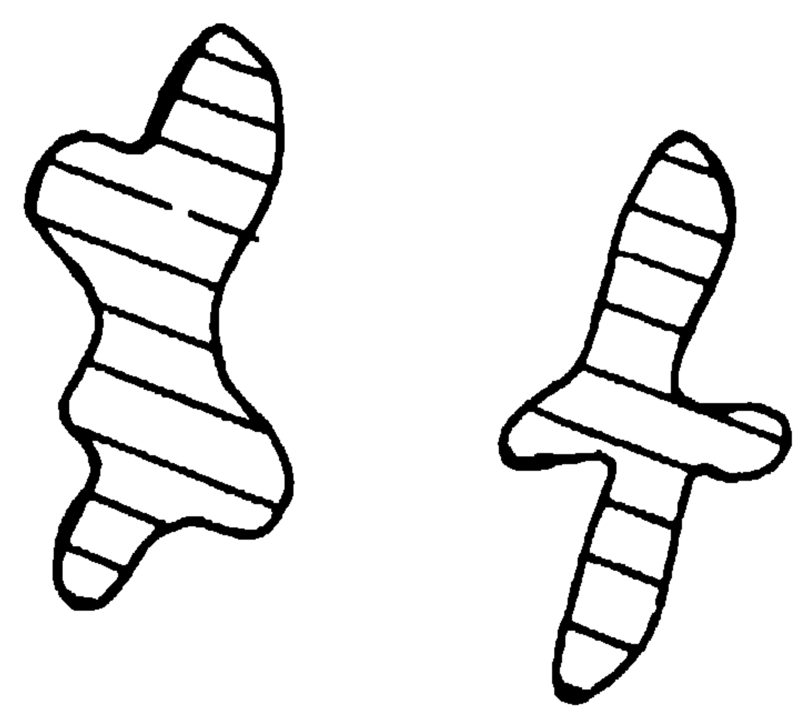
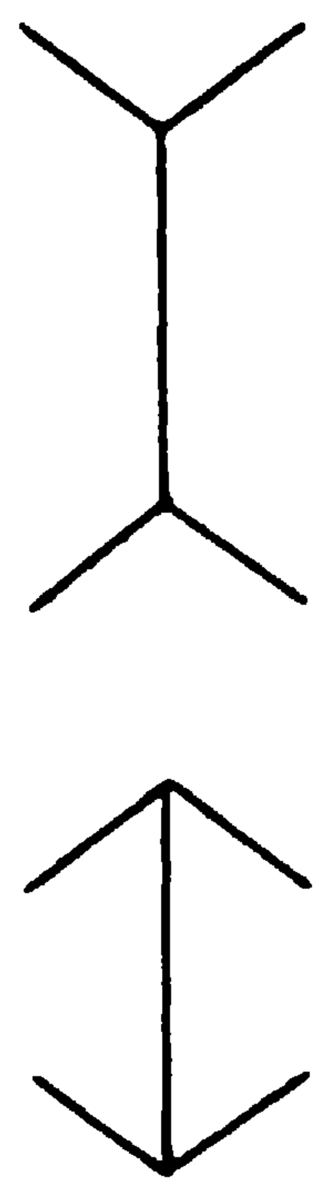
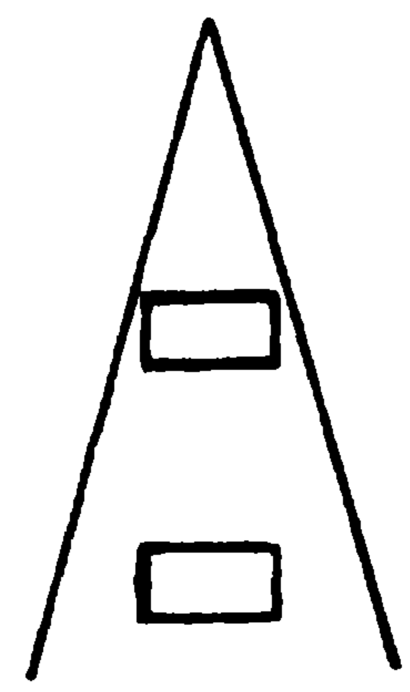


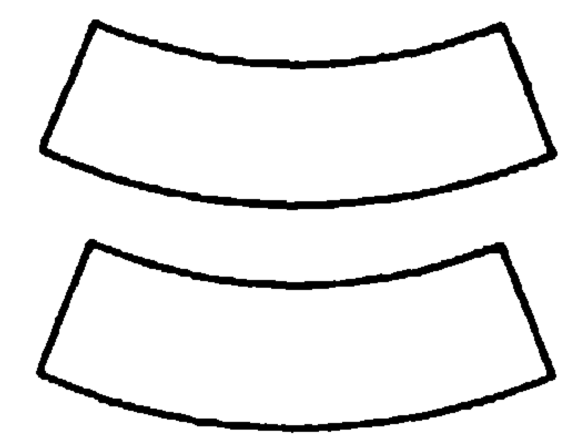
Figure 2.



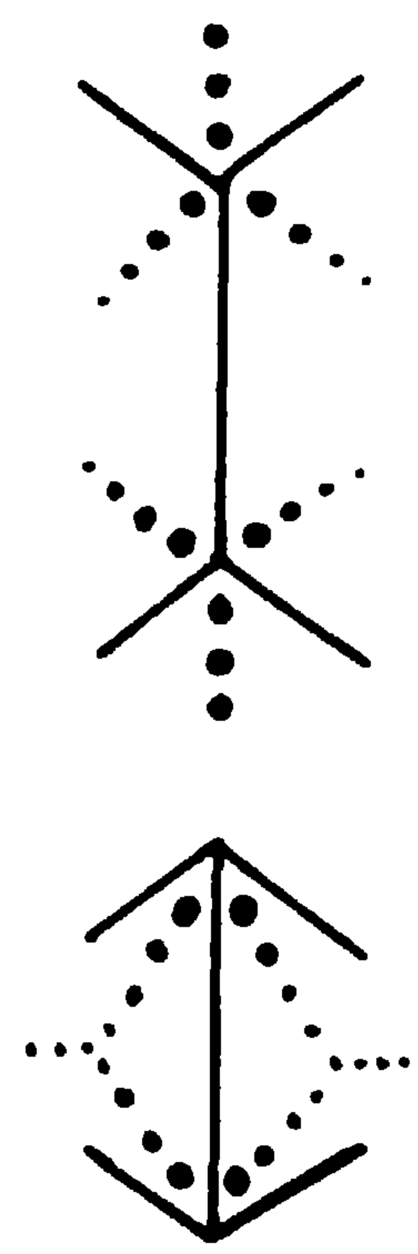
a



b



c



d

Figure 3.

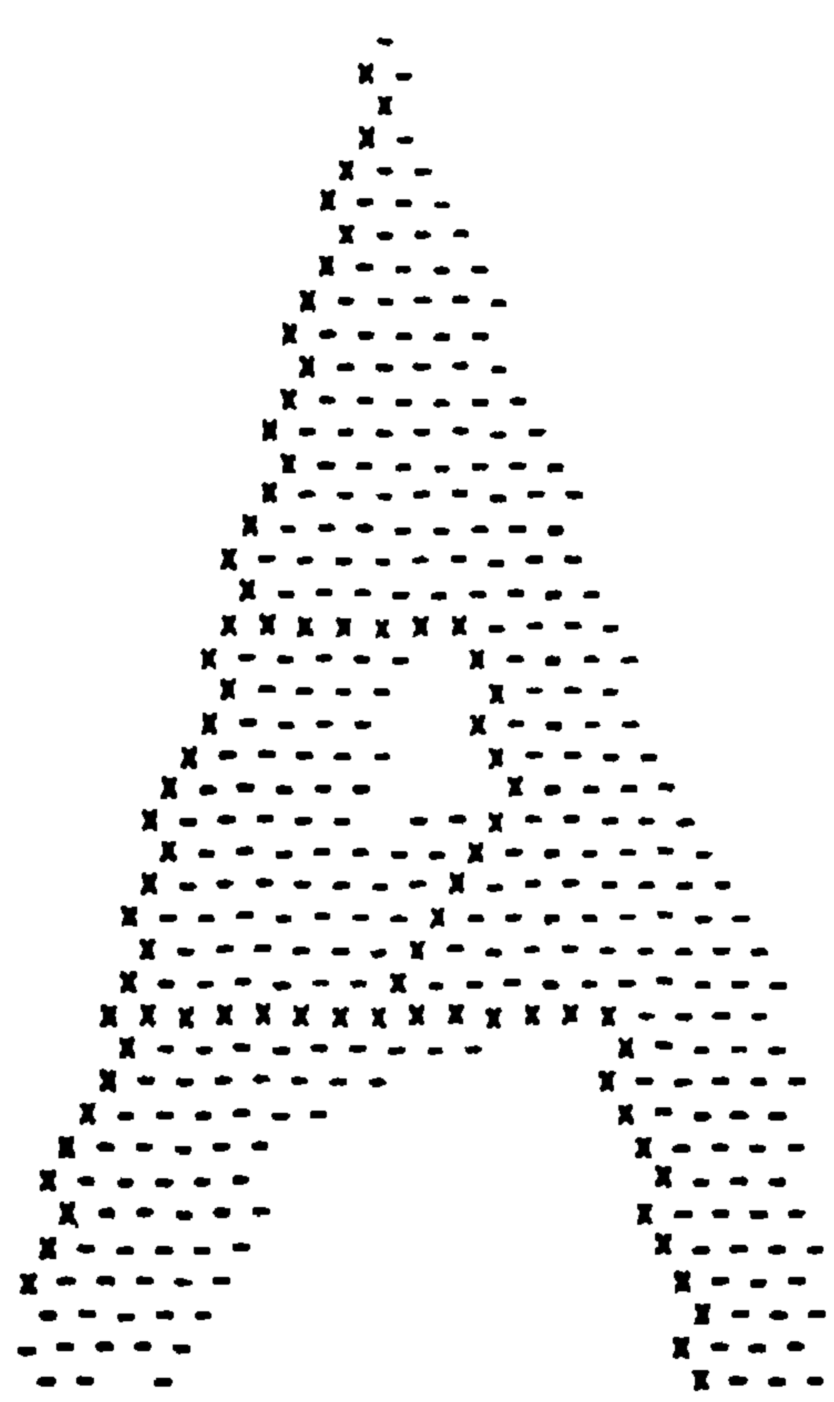


Figure 4(a)

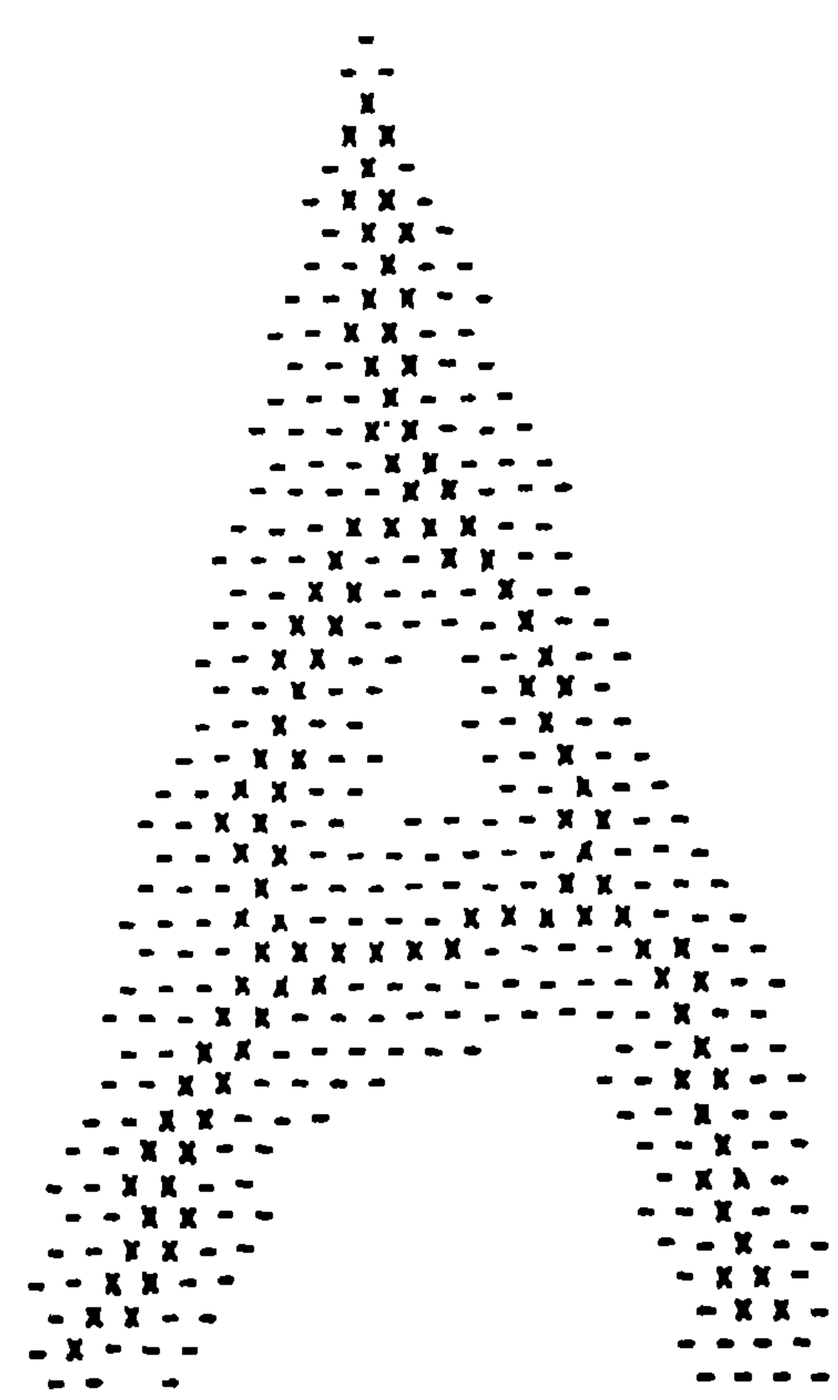
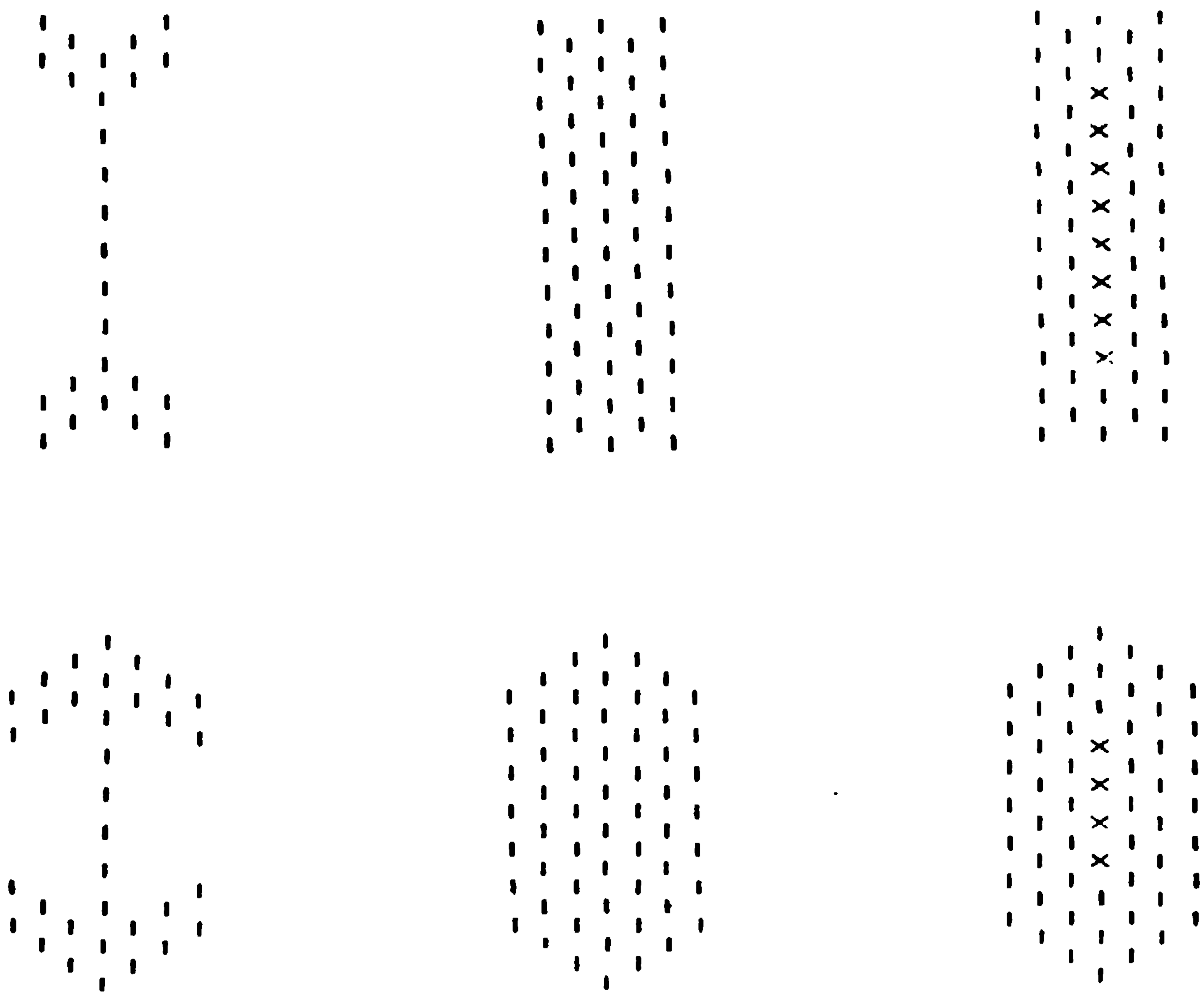


Figure 4(b)



a b c

Figure 5

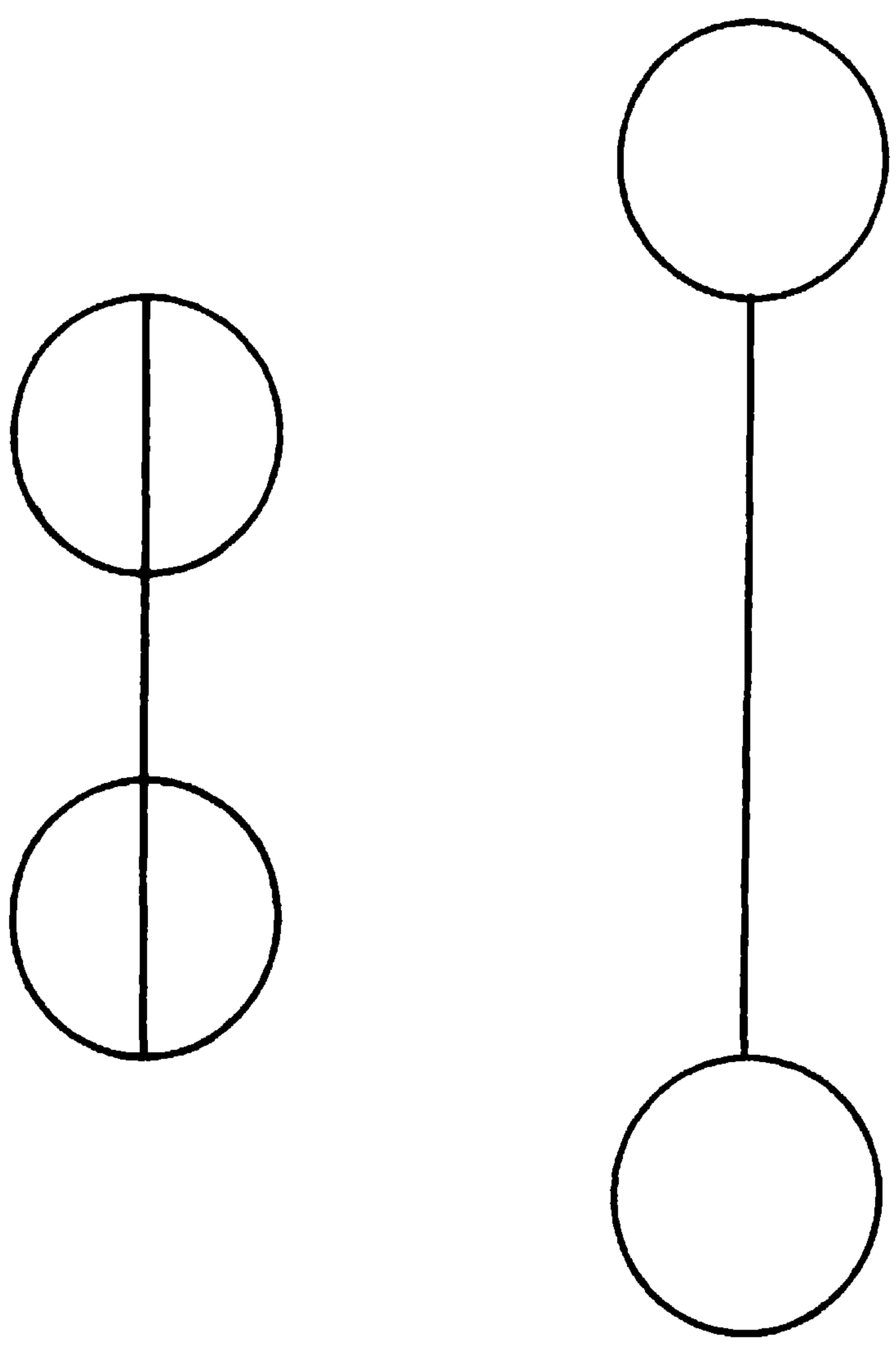


Figure 6

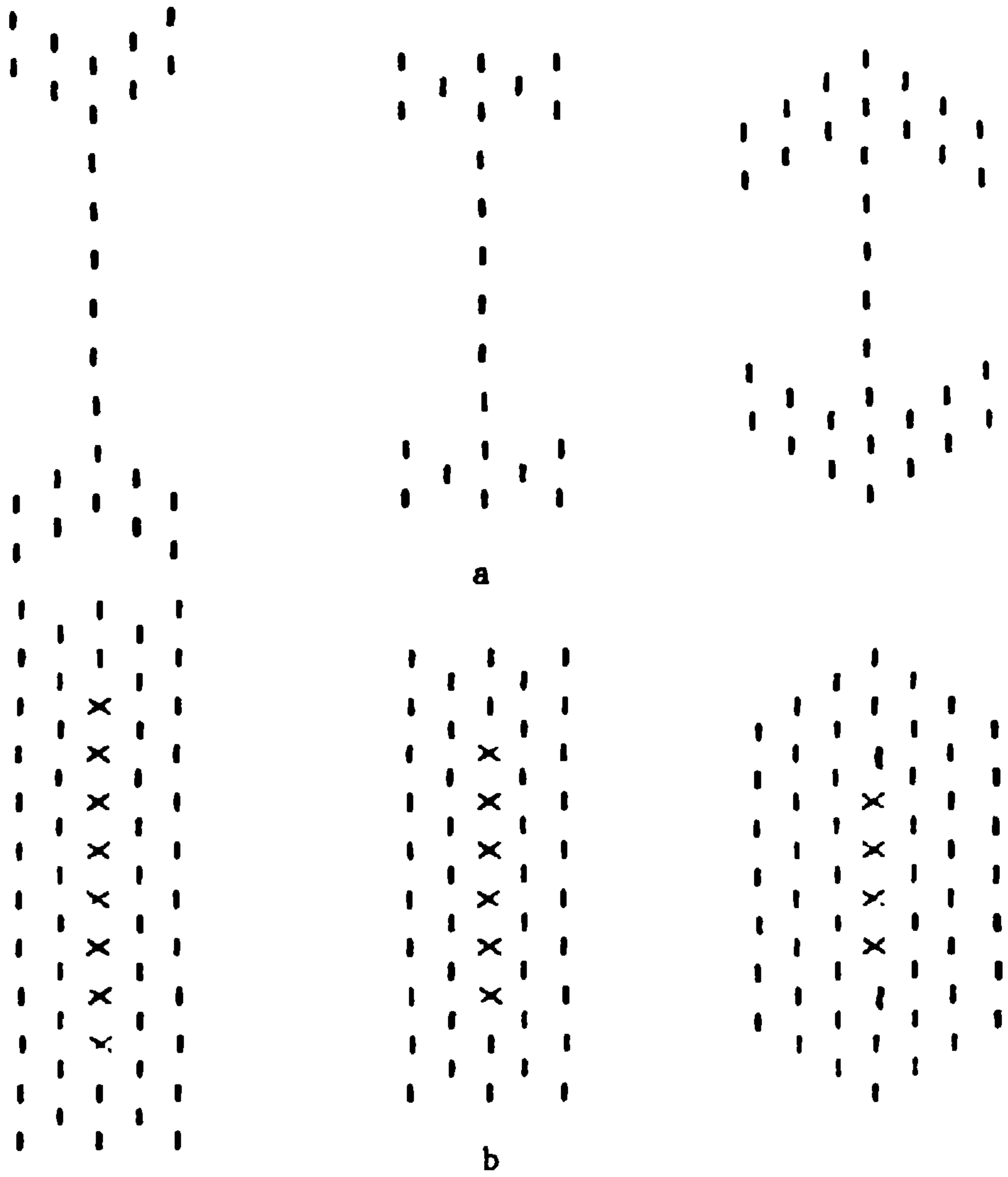


Figure 7

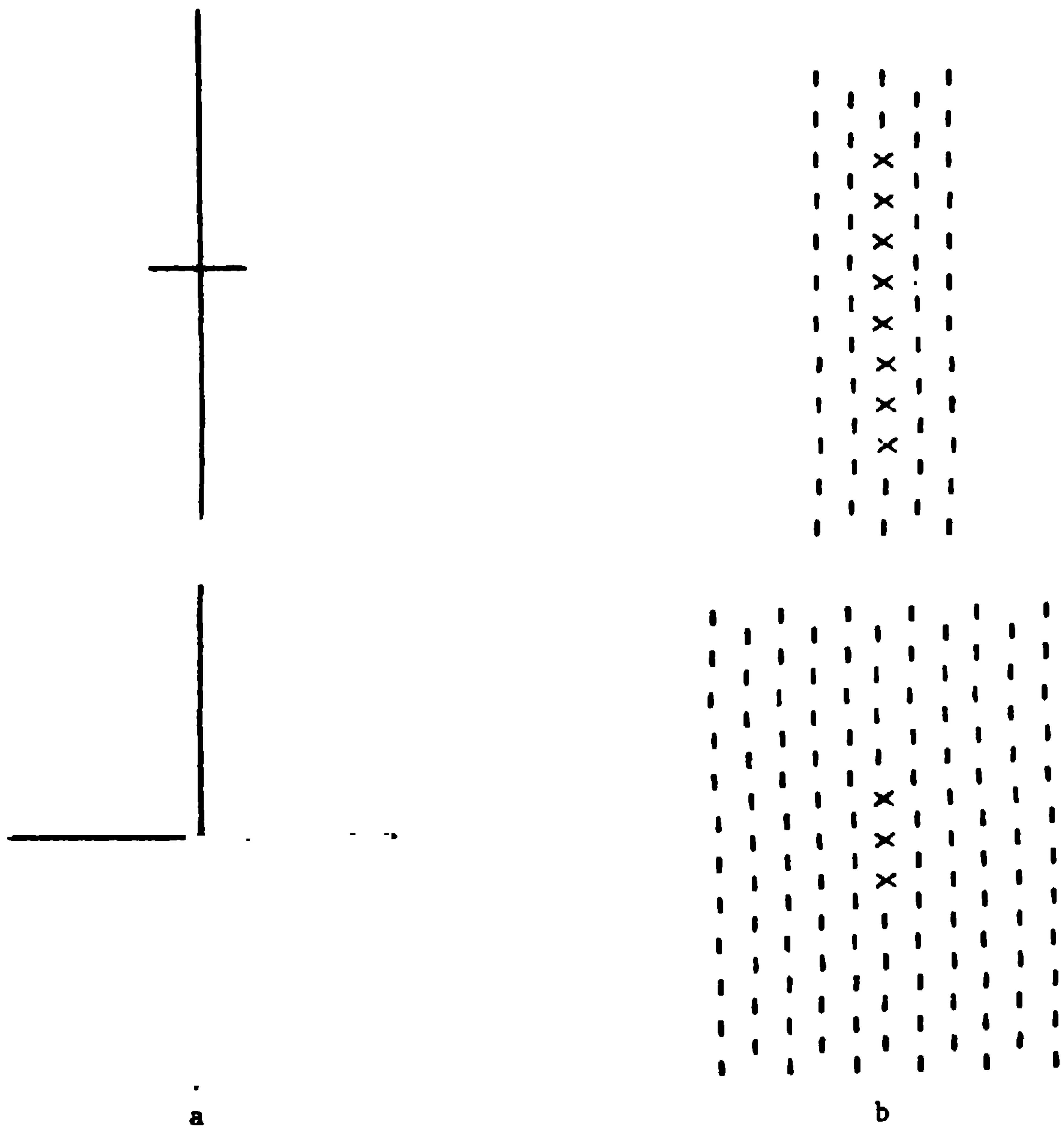


Figure 8

i

Figure 9.

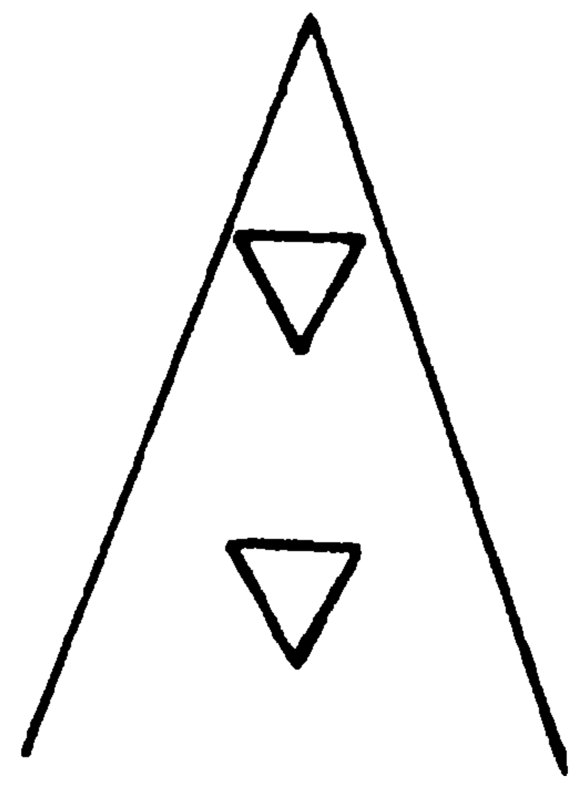


Figure 10.