Cortical Scanning: Evidence from Strobe Observations

Stewart W. Wilson
The Rowland Institute for Science
Cambridge, Massachusetts   02142

Abstract

Subjects described imagery seen while viewing a flashing strobe lamp with closed eyes. From the principal images, dual linear cortical scan-like processes are inferred, and then used as the basis of a photo-mechanical simulation of the images. Implications for cortical processing are presented.

Introduction

Experimental (Tootell, et al., 1982) and theoretical (Weiman & Chaikin, 1979; Schwartz, 1977; Wilson, 1983) investigations have suggested that the geometric relation between images on the primate retina and the resultant "images" (patterns of activity) on the primary visual cortex can be represented, approximately, as a complex logarithmic mapping. In particular, Wilson proposed the model shown highly schematically in Fig. 1. The "retina" of Fig. 1a consists of "data fields" whose size and spacing increase linearly with distance from the center of vision. The "cortex" of Fig. 1b is a matrix of identical "message sending units" (MSUs), each of which receives signals from its own retinal data field, processes the signals, and generates a relatively simple output message. The pattern of connections is as suggested by the letters in each figure. Data fields along a ray from center to periphery map into a row of MSUs, and simultaneously, each ring of data fields maps into a column of MSUs. The leftmost column corresponds to the innermost ring, the 12 o'clock ray maps into the top row, and so forth. The mapping illustrated is "complex logarithmic" because retinal azimuth angle and the log of radial distance

from the center of vision map into cartesian row and column coordinates. There is a singularity at the retinal origin which can be dealt with by beginning the mapping at a finite radius. Fig. 2, reproduced from Tootell, et al. and showing the mapping in an actual animal, supports the model to a first approximation.

A significant aspect of the retino-cortical mapping is an apparent size-normalizing property. If a given retinal image is magnified about the center of vision, the cortical "image" does not basically change in size or shape; instead, the image translates uniformly along the cortex. In the case of enlargement, the image slides away from the foveal representation (to the right in Figs.1b and 2b); it slides toward the fovea if the retinal image is reduced in size. This constancy of size and shape arises essentially because, under a logarithmic transformation, a scale factor M is converted to an additive constant log M, which then functions as the translational offset.

Understanding form recognition in animal or machine has traditionally been complicated by the simple fact that the same object can be imaged on the retina at many different sizes (depending, usually, on viewing distance). Recognition of an object is difficult if its scale is unknown. Yet, unless the background is extremely simple, the scale cannot easily be determined until the object is recognized. This circle would be avoided if form recognition occurred in "cortical space." There, for the reasons discussed above, it would be unnecessary to know the scale of an object, and it could be reliably recognized by direct comparison of its cortical image against a "library" of cortical images held in memory.

In such a comparison process, however, matching would have to be possible independent of the unknown input image's translational position on

the cortex. Wilson addressed this question by hypothesizing a process which scans the cortex of Fig. 1b, column by column from left to right, repeatedly. The MSU outputs of each successive column are read out in parallel by the scan and into a sort of multi-channel "shift-register" with as many channels as there are rows in Fig. 1b. Thus the cortical image could be thought of as entering and flowing along this shift register. Clearly, a comparison process set up at a point along the shift register would eventually match any input pattern for which there were a stored equivalent, regardless of the input's original position along the cortex. Thus, patterns could be recognized independent of retinal size and cortical position.

The present article is directed toward the question of whether a scanning process anything like the above hypothetical process actually exists in the primary cortex.

Walter's Scanning Hypothesis

In The Living Brain (1953), Grey Walter devotes a section to discussing reports of imagery seen by subjects while viewing an intense flashing strobe lamp with closed eyes. The primary purpose of this stimulus was to observe the effect of the flashing on the subject's electroencephalogram (EEG). The rate of flashing ranged from a few, to over twenty, flashes per second, depending upon the experimental situation. Often the flashing rate was near 10 fps, so as to observe the effect on the alpha rhythm.

The reports of imagery were unexpected, especially since the imagery was most vivid when the subjects' eyes were closed and the retinas were diffusely illuminated. In the following passage Walter summarizes the

subjects' reports. "The illusion is most marked when the flicker is between 8 and 25 flashes per second and takes a variety of forms. Usually it is a sort of pulsating check or mosaic, often in bright colours. At certain frequencies--around 10 per second--some subjects see whirling spirals, whirlpools, explosions, Catherine wheels."

Walter took the "illusion" seriously and proposed several ingenious arguments why the imagery could be the consequence of a scanning process within the visual system that was engaged in transferring "the spatial image which is received in the projection areas...to the other areas for cognition." An analogy with television is perhaps the most suggestive argument. If the TV studio had perfectly featureless white walls, but they were illuminated by a flashing strobe, then if the strobe rate were near the frame rate of the TV system, definite patterns would nevertheless be seen on the receiver. In particular, the patterns would tend to reveal the path of the scanning beam. As for the brain, Walter suggested "the whirling spiral so many people see under flicker, is it not perhaps an indication of the very path taken by the scanning point in the pattern it makes every tenth of a second?"

In making the scanning hypothesis, Walter did not attach special significance to the particular geometry exhibited by the subjects' imagery. From the perspective of the retino-cortical mapping, however, the geometry is strikingly suggestive: a spiral centered at the retinal origin maps into a straight line lying diagonally across the cortex; a rotating spiral corresponds to moving the line over the cortex (vertically in Fig. 1b). Another pattern from the imagery, the "explosion," could describe, retinally, an expanding circle or set of circles concentric with the center of vision. These would again map into straight lines on the cortex but

lying parallel to the columns and moving orthogonally from left to right. For both the whirling spiral and the explosion, the cortical correspondent is simple and suggests scanning. Particularly because of the geometric relationships, it was decided to repeat Walter's observations both to confirm them and to gain further information.

Observation of Strobe Imagery

Apparatus

A GenRad 1539-A Stroboslave unit was chosen as the stroboscopic light source. At low frequencies it produces a 3 µs pulse (triggered from an external oscillator) at three selectable intensities up to 18×10^6 beam intensity candela. The lamp has a multi-faceted, approximately parabolic, polished metal reflector 10 cm in diameter. The lamp was adjustably mounted so it could be positioned at the exact height of, and as close as desired to, the eyes of subjects who sat in a relaxed position in a comfortable chair. Other than the strobe, the only room illumination was from a small lamp the experimenter used for taking notes, so that the intervals between flashes were quite dark.

Procedure

Subjects were asked to hold their eyes closed firmly enough to be sure they were in fact closed but without strain or squinting. The strobe lamp was positioned within an inch of the subjects' eyes. The subject was shown before starting the knob of an oscillator which controlled the flash rate. He/she was instructed to turn the knob slowly and observe any visual effects which might or might not occur. The frequency was set

initially at about 10 Hz and the subject could go in both directions from there. If a particular frequency seemed to produce interesting effects, the subject was instructed to explore carefully in that vicinity to try to maximize the effect. If, on the other hand, a frequency neighborhood was not very interesting, the subject was instructed to move on. By turning the knob, the subject could vary the rate of the flash between 5 and 25 flashes per second. (This range turned out to contain most of the effects, though the next higher one was occasionally used as well.)

The experimenter asked the subject to describe verbally any interesting observations, particularly of pattern, motion, or color. To reduce potential interference between communicating and observing, subjects were asked to form a fairly complete impression of an image before describing it. Usually the subject would explore for a few moments, say something, and the experimenter would ask a short clarifying question to which the subject would respond; then the cycle would repeat. If the experimenter felt that observations in a particular frequency area were becoming redundant, he might ask the subject to move on. This was sometimes necessary when a subject became fascinated with a particular type of imagery and the experimenter wanted to be sure all frequency neighborhoods were investigated before the subject tired or became visually saturated or, as happened in a few cases, began to experience discomfort.

Besides observations communicated verbally, subjects were provided with pencil and paper for drawing what they saw. Typically, the experimenter might sense that the subject had formed a fairly complete impression of some image and would ask the subject whether he/she could draw it. If so, the flash was stopped while the subject sketched. The drawing intervals also functioned as visual rest periods.

As mentioned, the flash intensity could be set at any of three levels; these were controlled by the experimenter. Sessions always began at the lowest intensity. After the subject had explored the full frequency range and the experimenter felt that any available effects had been reported, the experimenter switched to the middle intensity. The frequency range was then re-explored by the subject, starting usually at 10 Hz. Then, if the subject was willing, the intensity was switched to the highest level, at least for a brief exploration. In general, subjects enjoyed making the observations and wanted to keep going; but if it appeared to the experimenter that a subject was quite sensitive to the light, observations at the high and even middle intensities were curtailed. Fortunately, the more sensitive observers usually made full reports at low strobe intensity.

As a control to be sure that no trace of the lamp's shape or the reflector pattern could be influencing the images seen, a piece of sheet diffusing material was passed between the subject's eyes and the lamp while the subject was reporting some definite image. The experimenter asked "Does what I am doing make any important difference?" The answer was always either no, or else a change in observed colors was noted but no material change in pattern. As a further test, several subjects were asked to view the lamp (always with closed eyes) while it was operated at a frequency well above fusion, such as 100 Hz. In every case, the impression was of a perfectly diffuse, white or "whitish" light. This observation seems conclusively to rule out any imaging of the lamp or reflector on the retinas, and also to show that any coloration introduced by the eyelids was slight. (A diffuser was not employed routinely over the lamp because it reduced brightness.)

Twenty-three male and female subjects, ranging in age from 10 to 74 years, participated in the sessions. They were not prompted to look for any particular type of pattern or geometry or movement. After a session, a subject was asked not to disclose to others what he/she had seen. Sessions typically took an hour, with the strobe observations occupying the first half-hour. The remainder of each session was devoted to an informal examination of the subject's resting alpha rhythm and, separately, to an attempt to trigger the flash from the subject's alpha rhythm signal while the flash was being observed. Walter hypothesized not only that the imagery was due to a scanning process, but also that the scanning might be associated with the alpha rhythm, perhaps even controlled by it. Though our interest was primarily with the images, it was thought worthwhile to look for any alpha rhythm connection. The results were unclear, and to keep this part of the work conceptually distinct from the imagery observations, where the results were much more definite, the apparatus, procedure, and results of the alpha experimentation will be presented elsewhere.

Results

The results largely confirmed Walter's summary statement, quoted earlier. Several types of patterns were seen, but most fell into either of two classes: (1) "cross-like," or (2) spiroform. There were many kinds of motion, but, when it had large-scale organization, the motion was usually describable as either radial or rotational. Colors, seen by all but one subject, were vivid and varied over a gamut which, for the subject group as a whole, included all the principal color names. For most subjects, at one time or another, the visual field took on a mosaic or tiled

appearance: the tiles were sometimes of constant size; at other times they increased in size from center to periphery. For every subject there was a definite center to the field at least part of the time. In general, higher strobe intensities produced more centrally symmetric patterns and motions, as well as stronger colors. The frequency range of greatest interest was between about 8 and about 20 flashes per second. Below this range, the stimulus seemed mainly just a flashing whitish light; above it, patterns got more fine-grained and "less interesting" and similarly lost color.

The class of cross-like patterns ("Crosses") included (1) crossing vertical and horizontal members (+ sign), and (2) crossing diagonals (x), (3) patterns of radial vanes, like spokes of a wheel, and (4) a flower or rosette-like pattern of four orthogonal "petals." Fig. 3 shows examples of the Cross class taken from subjects' drawings. Crosses were always seen at the center of the field. The class of spiroform patterns ("Spirals") included principally (1) single-arm spirals having multiple turns (like a snail shell), but also (2) multi-arm spirals not tightly wrapped (something like a turbine fan). Fig. 4 shows examples. The Spirals were nearly always seen at field center. Of the 23 subjects, 15 reported at least one Cross pattern and 13 reported at least one Spiral pattern. Ten subjects saw both Cross and Spiral patterns.

The Radial motion category included (1) expansion or contraction of Spiral forms, (2) "flow," inward or outward, along Cross arms, (3) expansion from the origin of sets of concentric circles, and (4) radial "flow" of the mosaic. The Rotational motion category included (1) rotation of Spiral forms, (2) rotation of Cross arms, and (3) rotation of the whole mosaic, all about the field center. Nineteen subjects saw Radial motion, 17 saw Rotation, and 15 saw both.

Out of the full 23 subjects, 5 saw neither Spirals nor Crosses. Some of what they did see follows, by person. (1) A colorful, mottled field in irregular movement. (2) Similar to (1), with a small whirlpool seen occasionally at field center. (3) Violent, swirling, colorful motion; a close-packed matrix of balls, with whole rows or columns in a sudden linear motion; the field seen as though looking down a tunnel, with the walls sometimes in rotary motion, and sometimes with objects coming from the far end and shooting past the subject. (4) A persistent impression of "beads" rotating together in "galaxies" (without, however, spiral form); the beads moving like many independent "bicycle chains." (5) Beaded chains moving like snakes all over the field, sometimes in loops; at higher intensity, the chains flowing into the center or about the center.

Two subjects saw neither radial or rotational movement. One was (1), above. The other, (6), saw a clear mosaic of hexagons which increased in size from near an unstructured center to the periphery. Ten lines radiated from the center. Structure was visible within the hexagons, but they did not rotate or flow; instead, they seemed to pulsate.

Many of the images just noted were also reported by other subjects, often at lower strobe intensity than the intensity at which those subjects began to see Spirals or Crosses. In fact, Spirals and Crosses were generally experienced at the higher intensities, with the less organized or less centered images or motions occurring at lower intensities. The principal non-Spiral, non-Cross forms were (a) "S-curves," wavelike horizontal or vertical lines seen by four subjects; (b) crossing arcs, like ripples on a quiet pond, especially as though coming from two separate centers, seen by nine subjects; (c) vertical and horizontal straight lines, reported by seven; (d) snakes or chains, reported by four. Examples are

shown in Fig. 5. The S-curves and vertical and horizontal lines were often seen to shift parallel to themselves.

A three-dimensional or depth sensation was reported by nine subjects. Usually this was associated with the "tunnel" image mentioned above, or with an expanding spiral which seemed to come toward and pass the subject. Flow along the arms of a Cross figure was seen by two people to occur in depth. Three subjects reported "edge-circles": a circular motion with the plane of the circle almost parallel to the line of sight.

Patterns having central symmetry were almost always seen centered on the apparent line of sight. But subjects were also facing the lamp. To determine whether the location of the pattern's center was due to the lamp direction or was due to the visual axis direction, twelve subjects while seeing definitely central symmetric patterns were asked to try moving their eyes around under their eyelids. Then they were asked whether the pattern moved too, as though carried by the line of vision, or instead stayed where it was. For one subject the pattern broke up whenever he tried moving his eyes. For another, the pattern center did not move. For a third subject, one pattern moved with the eyes, another did not. For nine subjects, the pattern moved reliably with the eyes. Thus the patterns appeared in general to be centered in the eye (about the fovea), not externally.

There were a few reports of Spiral or Rotating patterns not centered on the apparent line of vision (yet not, from the results above, controlled by the position of the lamp). For instance, two subjects saw brief images of two or more Spirals, present simultaneously, but located in different parts of the field. Another subject saw "edge-circles" away from the line

of sight, though sometimes arranged symmetrically around the apparent center of vision. A fourth subject reported "galaxies" in various parts of the field. Usually, however, if a centrally symmetric pattern or motion was present, there was only one such and it was in the center.

At low strobe intensity, the mosaic pattern mentioned by nearly all subjects was generally made up of constant sized elements. At higher intensity it acquired a center and a radial structure; the elements then got bigger with distance from the center, though for some subjects this was more of a "feeling" than an observation. The experimenter attempted to understand the relation between the Spiral, Cross, or other forms seen, and the mosaic. Were the forms made of mosaic elements or was it a background? It seemed that the mosaic pattern was like a multi-colored weaving whose elements were in constant motion or vibration. The forms resulted when certain elements took on one color and thus stood out from the other colors, or when certain elements all became "dark," or when they contrasted in some other way.

The foregoing descriptions of subjects' observations greatly simplify the phenomena in order to communicate and get a conceptual hold on them. The experimenter's notes indicate the images were often complex, multi-layered, kaleidoscopic, and transitory. Sometimes, several types of general motion, or motion and non-motion, were reported to be simultaneously present, as though superposed. Well-organized patterns such as Spirals or Crosses emerged unpredictably (with no clear relation to strobe frequency, other than being in the range noted earlier), and would dissolve or "switch" just as unpredictably--though a person could sometimes recover a pattern by returning to the same frequency later, and people who got patterns easily could hold them and sometimes even change,

say, a Spiral's direction of rotation "with their mind." The present account has attempted to delineate, within a rather complicated thicket of phenomena, those that seemed clearest and most frequent.

A Basic Scanning Model (M0)

It is striking that the imagery is seen at all. The retinal stimulus is evidently without significant spatial structure. Only temporally is the stimulus patterned; i.e., it is intermittent or periodic. Thus how, of itself, such a stimulus could yield moving geometric forms is not readily apparent. Rather, it seems necessary to postulate an interacting internal process having both space and time dependence. The observed images would then arise stroboscopically from the interaction of the periodic stimulus and the internal process. As suggested by Walter, the latter could be like a scan in which, at any moment, only a discrete portion of the visual field information was being "read." Periodic illumination might then result in perceived patterns formed out of the positions read at the moments of illumination.

This proposal is attractive except for the fact that the patterns to be explained are varied and relatively complex, while, ideally, a plausible scan-like process ought to be simple. Also, from the theoretical point of view of this paper, the same process desirably should permit exploitation of the retino-cortical mapping's potential for size-independent form recognition. In this section we present a primitive scanning model which, by variation of its rate parameters, appears to produce most of the centrally symmetric patterns reported by the observers. At the same time, the model's scanning behavior suits the mapping's normalization property.

In arriving at the model we have disregarded reported imagery which is not centrally symmetric, the assumption being that distinct other processes are responsible. Observers saw centrally symmetric patterns, non-centrally-symmetric patterns and, often, both as aspects of the same complex image. Clearly, there is much to explain and interpret. However, the greatest commonality among reports had to do with centrally symmetric patterns, and it seemed reasonable to try to model them first.

The model has two main components:

(1)   A retinal array (RA) of "data fields" (Wilson, 1983) resembling Fig. 1a and shown in Fig. 6a. RA differs from Fig. 1a in that interstitial data fields have been added. The data fields retain the basic property that their size and spacing increase linearly with distance from the center.

(2)   A pair of fan-like sets of "scanning arms," shown in Fig. 6b. The first fan, "L," rotates counterclockwise and its arms curve clockwise outward. The second fan, "D," is the mirror image of L and it rotates clockwise.

The components of the model are "assembled" by placing L and D on top of RA with all three centers coincident (Fig. 6c).

In the simplest version of the model, M0, there is a strict relationship between the form of L and D and the form of RA. Let 2N denote the number of "rays" (rows leading out from the center) of data fields in RA. Then the number of arms in L (and D) is made N/3 (a whole number), and

they are equally spaced. In addition, each arm is spiroform and falls across the centers of RA data fields such that if an arm crosses the data field at ring i, ray j, the adjacent data fields crossed by that arm have indices i-1, j-1 and i+1, j+1 in the case of L, and i-1, j+1 and i+1, j-1 in the case of D. (A ring of data fields consists of those lying at equal distance from the center; rings are numbered outward.)

Figure 6c shows the model from the point of view of retinal space. However, the scanning action is actually assumed to take place cortically. The corresponding cortical picture is shown in Fig. 7. The arms of L and D are transformed by the mapping into two sets of diagonal straight lines moving upwards and downwards, respectively, across the MSUs of the cortex. We assume for the model that whenever an intersection of an L arm and a D arm falls near the center of an MSU, that MSU's message is "read out" to higher cortical processing levels. The content of the message depends on the stimulus falling at that moment on the corresponding retinal data field. Thus Fig. 6c also represents the scanning process. As L and D rotate, their intersections move generally outward and can be thought of as sampling the data fields which lie under the intersections. Though the sampling is actually occurring cortically in the model, we shall usually find it convenient to take this retinal point of view. Referring again to Fig. 6c, it will be noted that at any moment, the arms of L (and D) cover only every third tract or spiroform sector of RA's data fields. Similarly, the sampling intersections of L and D sample only every third data field. The number three is in fact central to the model, as will be seen.

We now place the model in the experimental situation and assume that the retina is uniformly illuminated with a flashing strobe light. Let the

flashing rate be F flashes per second and let the rates of rotation of L and D be V_L and V_D spiroform sectors per second (V_L and V_D are not always equal and they may vary). In each flash, the model will only transmit messages from MSUs whose data fields are sampled by L and D intersections ("LxD") at the moment of the flash. The other MSUs will be silent. Thus the instantaneous perception will be of a spaced array of patches of light.

By the time of the next flash, however, LxD will have moved to new positions on the array, and to a new set of data fields. Thus a different set of patches of light will in general be perceived next. On the following flash, there will be yet another perceived array of patches, and so on. We now make the hypothesis that the forms and motions reported by our observers resulted from perceptual patterns or gestalten arising out of the succession of these instantaneous arrays of patches of light.

Figure 8 shows LxD sampling a certain set of data fields on a particular flash. Let us imagine that on the next flash, LxD has moved in such a way that its intersections again fall precisely on a set of data field centers (integer shift). With this condition, a moment's reflection will show that the second set of sampled data fields will be either (1) the same as the first, or (2) the set formed by shifting the first set exactly one field in one of the directions shown by arrows in Fig. 8.

In case 1, the perception will clearly be of an array which is stationary. In case 2, the perception will be of motion in the direction of the shift, provided the flash frequency is in the range where successive presentations can induce a motion gestalt. Furthermore, the motion, if it is perceived, will be unambiguous in direction. The reason is that each intersection shifts (apparently) by a distance which is only one-third,

i.e., less than half, the distance to any other intersection. Thus it is perceptually clear which way the sampled data fields are "moving." This definiteness of direction results from the 3-sector spacing of the sampling arms; if the arms were spaced every other sector (or every sector), no movement could be perceived.

The arrow directions in Fig. 8 indicate three basic modes of array motion: expansion or contraction; rotation clockwise or counterclockwise; and flow, outwards or inwards, with a spiral curvature like that of the sampling arms themselves. These modes are the same as some of the types of motion reported by our observers.

So far it has been assumed that LxD, on each flash, falls precisely on a set of data field centers. What if this is not true? To find out, and to confirm the previous ideas, a scanning device was built which simulates the components of the model. It consists of three photographic transparencies in superposition, illuminated by a strobe lamp on one side and viewed from the other side. The transparency adjacent the strobe lamp has a pattern of clear circular spots on a black ground and resembles Fig. 6a (with N = 36). The other two transparencies are mounted on clear plastic disks and have spiral arms like those of Fig. 6b, except that the arms have a clear, finite width of approximately one sector with a smoothly changing transparency cross section. In between arms the disk is opaque. Thus, when superposed, the two transparencies yield clear intersections about the size of the adjacent spots in the first transparency. Seeing a spot through an intersection was taken to be analogous to the sampling of a data field in the model.

The L and D disks could be independently rotated by separate motors. When the speeds V_L and V_D and the flashing rate F were adjusted

so that the movements of LxD met the integer shift condition discussed in connection with Fig. 8, it was found that all the predicted motion gestalts occurred. When the integer shift condition was nearly, but not quite, met, the motion was as for the closest integer shift except for an orthogonal drift. For example, with F = 10 hz, V_L = V_D = 10 sectors/sec, the motion is radial outwards. If V_L is changed to 9 sectors/sec, the radial flow takes on a slight clockwise rotation.

To understand all the relationships, combinations of F, V_L, and V_D were tried until a clear pattern emerged. The results are diagrammed in Fig. 9. The axes represent the ratios V_L/F and V_D/F in sectors per flash. The little arrows are located at positions of integer shift and their directions have the same meaning as in Fig. 8. A small circle stands for a stationary pattern. The dashed lines mark loci in which the pattern is without overall motion but, instead of being stationary, it is chaotic. These loci correspond to an apparent shift by L or D or both of exactly 1½ sectors; when that occurs, the motion direction is ambiguous and a definite gestalt does not form.

Using Fig. 9, we can read off non-integer shift motions as follows. Within any square formed by dashed loci of ambiguity, if you go from an arrow toward the central circle, the motion retains its direction but gets slower. If you go from an arrow to an adjacent one, the motion stays as fast but changes direction as indicated. If, finally, you go from an arrow across the locus of ambiguity, the motion changes as shown, but near the locus it will be chaotic--i.e., not organized into an overall flow.

The states of motion described in Fig. 9 are highly varied and change rapidly under change of any of the three variables V_L, V_D, and F. In our experiment, the observers noted that overall patterns of motion

came and went with relatively small changes in strobe flashing rate, roughly 1 hz. Since F is of the order of 10 hz, this is a pattern change within a 10% change in F. The implication, according to Fig. 9, is that V_L and V_D must be of the order of 10 x F, or about 100 sectors/sec.

We assumed in the model, and also for theoretical reasons to be given later, that V_L and V_D can be different in magnitude and are variable by the visual system. In our experiment, observers often noted switches in mode of motion without any adjustment of flash frequency; usually the changes were involuntary, but some observers could make them happen. At first glance, Fig. 9 seems to suggest such changes would require large percentage changes in V_L or V_D in order to make V_L/F or V_D/F change by an integer. But if, as just inferred, V_L and V_D are of the order of 10 x F, only a 10% change is needed, which is quite plausible. Figure 9 consists of an indefinitely repeated pattern. Evidently the "operating point" for the human system, according to our model, is considerably off the graph as shown.

In using the scanning device to construct Fig. 9, we noted an important further gestalt in the regions near horizontal arrows. If the device was set up to give pure rotary motion, the actual figures more closely resembled rotating concentric circles separated by spaces than a uniform rotating array. If now any of the three variables was changed slightly, the circles either began expanding outward or contracting inwards. But surprisingly, the pattern then looked less like circles than like a single, highly wrapped spiral either expanding outward or collapsing inward. Also, there was a feeling of depth, with the center the far point. Thus it may be that the highly wrapped spiral seen by many observers was actually an expanding or contracting set of concentric circles.

Model M0 has suggested a mechanism for a number of the images seen by our observers: radial and rotary flow; multiple-arm spirals; the single, highly wrapped spiral (perhaps); and the intervals of chaotic activity. The model suggests that the observed mosaic pattern is the data field array itself, and shows how the "tiles" of the mosaic may be the individual data fields made visible by the special combination of sampling plus strobe illumination (we return to this point later). But several aspects of the observers' reports are not predicted by M0. In particular, M0 is achromatic whereas the observers reported vivid colors. Furthermore, no understanding is provided of the Cross phenomena. These questions will be addressed in the following two sections.

Color (Model M1)

Under the conditions of the experiment, observers frequently saw vivid colors with a stimulus which is ordinarily regarded as essentially colorless (or bluish-white). Our basic hypothesis in trying to understand this fact is to propose that the scanning process performs a time-dependent trichromatic analysis of the retinal image. Under ordinary visual circumstances this analysis is not evident to us, just as trichromaticity itself is not. But, under the conditions of the experiment, the separate color components tend, stroboscopically, to be revealed.

This hypothesis is incorporated into the model by supposing that L and D, rather than each having N/3 (achromatic) scanning arms as in M0, instead have N/3 scanning arms for each of the three trichromatic components, giving N arms in all per scanner. Figure 10 illustrates this for D; the arms of the three chromatic sets are interlaced: R, G, B, R, G, B, etc. When L and D together scan the image, they rotate oppositely just as in M0. Now, however, the arms sample in R, G, and B. In detail

20

this means that each data field of RA is cyclically sampled by L as to its R, then G, then B, information; and at the same time is being similarly cyclically sampled by D.

We can picture the process from the point of view of the set of intersections LxD. Now, since the scanning arms are sampling separate chromatic components, the intersections can be thought of as sampling pairs of components (where both members of a pair will sometimes be the same, e.g., R and R). Thus LxD is like a mosaic of colors, though of course it is not itself colored. Rather, the waveband pairs sampled at each intersection change, mosaic-like, from one intersection to the next. At the same time, since L and D rotate, the intersections move with respect to the data fields of RA.

The foregoing modifications to M0 make it a new model, M1, incorporating color. Figure 11 shows a part of the LxD of M1. Intersections are labeled as to the color components sampled. The diagram is similar to that of Figure 8 for M0. Moreover, the motion gestalten predicted by Figure 11 for the color case are very similar to Fig. 8's predictions for the achromatic case.

Let us imagine, as we did with Fig. 8, that on one flash LxD superposes exactly on the data fields of RA and on the next flash has moved in such a way as again to superimpose exactly (this is again the "integer shift" situation). It is clear that the effective movement of the RR intersections, for instance, will be exactly one field in the direction shown by one of the little arrows of Fig. 11 (unless LxD duplicates its prior state). The same will be true of the GG and BB intersections. But these are the same kinds of displacements as were discussed in connection with M0. Furthermore, since the RR intersections are spaced

21

by three steps just as were the achromatic intersections (and similarly for GG and BB), these three colors should be observed to display the same set of motion gestalten under the same conditions of V_L, V_D, and F.

What about the "mixed-color" (e.g., RG) intersections? Here, interestingly, the situation is different. Examination of Fig. 11 shows that for any such intersection, the next equivalent intersection is within two steps, not three. As discussed earlier, motion gestalten under integer shift cannot occur in this case, since it is ambiguous which way the underlying data field has moved. Thus if we were to make a working device based on M1 and look at it, we should observe motion gestalts for some of the colors (RR, etc.) and ambiguous or merely pulsating behavior in the others. As noted, subjects sometimes reported interwoven combinations of motion types, or mentioned a feeling of overall motion in a field which was otherwise apparently not moving.

To test these ideas, the scanning device built for M0 was modified by replacing the monochromatic L and D transparencies with color transparencies. That latter's arms were not, however, made red, green, and blue, but instead cyan, magenta, and yellow (-R, -G, and -B, respectively). If RGB arms are used, a proper simulation of M1 requires that L and D be superimposed by projection, that is, "additively." This is mechanically a much more complicated set-up than our basic device, which works "subtractively." To make use of the latter, the transparencies were done in -R, -G, -B. The resulting intersections of LxD are "colored" with the complements of the colors of Fig. 11. The behavior should similarly reverse, with -R-R, -G-G, -B-B exhibiting the motion gestalts and the others being ambiguous.

When this device was tested, the motion gestalts predicted in Fig. 11 occurred, under exactly the rate conditions given in Fig. 9. It appeared that the movement was attributable to the "negative" colored intersections (-R-R, etc.) as suggested above, and not to the others. The display would sometimes simultaneously exhibit a static or merely pulsating aspect which seemed to be due to the remaining intersections. Detailed observation of the "M1 machine," though, was harder than for the M0 (achromatic) machine. Scrutiny of individual colors tended to break up the gestalts, which reappeared when attention became more general.

In summary, Model M1 seems consistent, at least as a working hypothesis, with subjects' reports of multiply-moving patchworks or mosaics of vivid colors. In the next section we shall extend the model once more to incorporate the Cross group of phenomena.

Cross Imagery (Model M2)

The Cross imagery introduces axes of symmetry to the field of view. The simplest form of Cross has two straight members orthogonal to each other and intersecting at the center of vision. Similarly, the simplest flower pattern has four major "petals" separated by right angles. The M1 model is, however, rotationally symmetric with no such axial symmetries. We can get the latter by adding to M1 the possibility of a spatial "beat" or moire between the scanning arms and the retinal array.

M1 was formed by having 2N rays of data fields in RA and, in L and D, N/3 (an integer) arms for each of three scanned colors--N arms in all. For example, in the M1 machine, N = 36, with 12 arms of each color in the scanners. A beat can be obtained by changing the rule slightly (giving us model M2). Instead of the same value of N for both RA and LxD,

define N_{RA} such that RA has $2N_{RA}$ rays and define N_{LD} such that L and D each have $N_{LD}/3$ arms in each color; but now let N_{RA} and N_{LD} differ by four. For example, we can keep N_{LD} = 36, but make N_{RA} = 32. LxD will then have 72 rays of intersections as before, but RA will have only 64 rays of data fields.

Figure 12 shows the resulting spatial beat. LxD (in black and white) was superimposed on RA and the combination photographed. A "flower" pattern with fourfold symmetry is apparent. Note that at any constant distance from the center, LxD and RA are in phase at four equally spaced places. Note also that along any radius, the two components also go in and out of phase, so that the beat effect holds both radially and in azimuth.

Figure 12 resembles many subjects' descriptions of the "flower" image. Also, the rule change in going from M1 to M2 is not so ad hoc as might at first appear. In Figure 7 was shown the cortical correspondent of M0: linear scanning lines moving over the cortex. The picture for M1 would simply triple the density of scan lines and assign colors to them. There is an implicit assumption that the scan lines have the same periodicity as the cortical array. Model M2 relaxes this assumption and allows the "wavelength" of the traveling scanning pattern, and that of the cortical array, to differ. No performance disadvantage would seem to stem from this change.

The M1 scanning machine was modified for model M2 by changing to an RA with N_{RA} = 32. It was immediately apparent that besides the motion gestalts for individual intersections of LxD, there were now separate motion gestalts possible for the "petals" of the flower pattern. That is, groups of intersections forming a petal would appear to move as a

unit.   As with M0, the quantitites V_L, V_D, and F were varied until the underlying pattern of relationships presented in Fig. 13 was found (the pattern can be explained in detail analogously to Fig. 9).

The arrows and small circles in Fig. 13 have the same significance as in Fig. 9.   Interestingly, the stationary points of the beat pattern occur for every integer pair of V_L/F and V_D/F, instead of at integer values divisible by three.   Both charts apply, however, so it is quite possible to have an operating point (say, $V_L/F = 1$, $V_D/F = 2$) showing, for example, rotation of the intersections but no movement of the flower pattern.   Thus we have a further way in which perceived motion can be multiple, in correspondence, perhaps, with the multiplicity our subjects reported.

The arrows in Fig. 13 are drawn quite close to the stationary points. It was found that small movement away from the latter would throw the flower pattern into strong radial or rotational movement.   Again, such movement of the beat pattern could be at variance with perceived intersection flow as given by Fig. 9.   All these effects correlate well with our subjects' reports, which included expanding, converging, and rotating motions of the cross images, together with "flow" within the cross arms.

There was a surprise in working with the M2 machine represented by the small letters "p" at half-integer values of V_L/F and V_D/F.   The p stands for "pinwheel" in the fireworks sense, and that describes the effect. Figure 14 shows a diagram: there are four comet-like curving arms; at a point slightly away from the points p, the figure slowly rotates while "light" appears to flow out the arms.   Our subjects made some reports of multi-arm, or fan-like, spiral figures.   Earlier we associated such reports with the states of M0 represented by diagonal arrows in Fig. 9, that is, gestalts of flow along scanner arms.   This explanation may be incomplete,

25

however, at least for some of the reports, because the fan-figures predicted by Fig. 9 are not easily made to appear rotating, while many of the reports mentioned slow rotation. Thus the pinwheels of M2, which do readily rotate, may be more properly associated with what some subjects saw. It is also exciting to think that these may be the "Catherine wheels" (fireworks pinwheels) seen by Walter's subjects.

We have not offered an explanation of members of the Cross class which really do look like crosses and not flowers. What is required, perhaps, is a change in M2 which would permit indefinitely narrow petals in the flowers to make them into crosses. One approach is to increase the pitch of the arms of L and D, that is, make them straighter. This corresponds to a cortical picture in which the scanning lines slide across the cortex at a different angle. The approach looks promising but will not be pursued here.

The Mosaic Pattern

Our hypothesis regarding the mosaic-like pattern reported by most subjects during centrally symmetric imagery is that it represents the "data fields" of Wilson's (1983) MSU model. The "tiles" of the mosaic were described as having various shapes, e.g., triangular, square, with hexagonal or "honeycomb-like" the most frequent. The tiles gave at least "the feeling" of increasing in size from center to periphery and sometimes were reported definitely to do so.

A basic proposition of the MSU model is that the data fields and associated MSUs are discrete. That is, a particular MSU receives information from its data field and generates an output message which is distinct (though not necessarily different) from the output messages of

adjacent MSUs. The data fields of adjacent MSUs may well overlap, but each MSU has its own data field and generates its own message. Hubel & Wiesel (1974) describe the cortical hypercolumn (which we regard as an MSU) as a "2-3 mm region of cortex [which] contain[s] by a comfortable margin the machinery it needs to analyze the region of visual field that it subserves." In their picture, however, the hypercolumn is only discrete in the sense that anatomically and functionally the cortex has a spatial periodicity the unit of which is the hypercolumn. Like a single cycle of a sine wave, the hypercolumn displays a full cycle of properties, but it is not discrete or bounded in any stronger sense. Furthermore, the hypercolumn is thought to generate a plurality of detailed output signals (possibly information-preserving vis-a-vis the input) which merge smoothly with the pluralities of signals from "adjacent" hypercolumns. Thus, from the point of view of computational architecture, the hypercolumn is not a discrete building block, but the dominant regularity in a more finely structured quasi-continuum.

The MSU proposal contrasts with this picture by suggesting that each hypercolumn is in fact a spatially distinct machine. From this viewpoint the retino-cortical mapping represents a rather coarse spatial digitization of the input image. (Note this does not imply that in normal vision the world has to _look_ coarsely digitized; only that peripheral visual knowledge must be very limited.) The strobe experiment bears on the MSU proposal in two ways. First, the observed mosaic tiles are discrete structures which resemble the postulated data fields as to size and spatial arrangement. From psychophysical data Wilson (1983) estimates the angle between rays of data field centers to be approximately 1/8 rad. In the present terms, this is $N_{RA} = 25$, meaning about 50 data fields in a circumferential ring. That

is within the size range suggested by subjects' comments and drawings, though of course the comparison is crude. One might propose, however, that the observed mosaic is only LxD, the scanner intersections; the underlying retinal array might still be essentially continuous. This would seem a strong possibility except that a continuous RA would not permit the interference or "beat" phenomenon we have identified with the Cross and pinwheel imagery. Furthermore, the beat typically is four cycles per revolution, requiring that the resolutions of LxD and RA be very similar. Thus the strobe experiment provides evidence for the discreteness of the data fields and MSUs, and support for the MSU hypothesis.

Size Normalization

This investigation's original motivation stemmed from the retino-cortical mapping's apparent size-normalizing property, and the associated desirability of some form of cortical scanning. We envisaged a process scanning the cortex (Fig. 1b) column by column, from left to right, repeatedly; retinally, the process would be a dilating circle. Our results with the strobe suggested a quite different form of scanning employing (retinally) two multi-armed rotating processes or (cortically) two orthogonal sets of parallel scan lines lying diagonally across the cortex and moving in opposite directions. Do the present results relate at all to the size-normalizing process?

Let us extend the scanning model slightly by giving L and D definable orientations. The simplest way is to choose an arm of each and label it, say, the "zero arm." Now, as a scanner rotates, its orientation is determined by the zero arm's position. The two zero arms will always have an intersection which we can label, say, "x." If L and D

counterrotate at equal speeds, x will move repetitively from center to periphery along some radius. To change the radius we could briefly let L go a little faster, then return its speed to that of D. X would subsequently oscillate on a new radius. Cortically, x is the intersection of the two "zero scan lines." It moves repeatedly left to right along a row of MSUs. To change rows, we briefly change the speed of one of the scan line sets.

Continuing cortically, note that at any moment the position of x defines both a row and a column, the "x-row" and "x-column." The x-column moves from left to right repetitively just as is required in our original size-normalizing scan. Furthermore, the labeled x-row associates an azimuth with this scan. Questions of actual physiology aside, it is not difficult to visualize a comparison scheme in which a 2-D patch of cortex defined in relation to the instantaneous position of x gets matched, as x cycles through its values, against a repertoire of stored 2-D patterns. The scheme will recognize patterns independent of retinal size, provided the azimuth (orientation of the unknown) is known or correctly guessed. How that may occur takes us beyond the scope of this paper, but it is noteworthy that the two-scanner model permits the azimuth of the comparison scheme to be readily changed, via a brief interval of speed difference between the scanners.

Concluding Questions

We have inferred much about the normal visual system from observations made in highly unusual circumstances. How do we know that the observations are not peculiar to the circumstances, and that the inferred scanning mechanism actually exists and operates in normal vision?

29

Of course at this point we do not know, and support from independent directions is needed.

There is little hint of scanning or related strobe effects in ordinary eyes-open vision, but then perception nearly always hides its routine mechanics from us. A more puzzling question is why the mechanics are partially revealed under intense strobe light with eyes closed. Perhaps the disguise is just switched off when those muscles are activated: there is, after all, nothing to see, and strobes are certainly not anticipated. But if that is not the case, then there must be something about unstructured, diffuse illumination (in brief intense flashes) which unmasks the scanning. If so, one should see the imagery with eyes open if one looks at a strobe-illuminated featureless screen larger than the field of view. We have some evidence that effects are seen, but the experiment needs to be done more carefully. If the same effects are indeed produced, it may mean that the seeing of any definite form--that is, in a certain sense, vision itself--causes the system to turn on fully, including all the self-disguising aspects. Such a result would help explain why a strobe-lit ordinary room, though it produces certain odd effects, does not show the scanning imagery.

Turning these questions around, we may ask why in our experiment the imagery was not more prominent than it was. That is, why were there usually only intervals of highly organized images, and why did subjects see the images in different degrees? What was responsible for the image types we have not analyzed? The model shows how the nature of the principal imagery depends on the flash rate and the two scanner rates. A sensible design of the scanning system would have it avoid scanner rates which would risk imagery-producing relationships with any external

30

pulsation of the illumination. But also it may be that the scanning rates just naturally become irregular with closed eyes as the system, getting no structured input, begins to "hunt," as outlined in the previous section, for a productive azimuth setting. Questions such as these, which require more detailed imagery information, suggest the value of experimenting with observers who both see the imagery easily and have the artistic talent to render it very accurately.

References

HUBEL, D. H. & WEISEL, T. N. (1974). Uniformity of monkey striate cortex: a parallel relationship between field size, scatter, and magnification factor. Journal of Comparative Neurology, 158 (3), 295-305.

SCHWARTZ, E. L. (1977). Spatial mapping in the primate sensory projection: analytic structure and relevance to perception. Biological Cybernetics, 25, 181-194.

TOOTELL, R. B. H., SILVERMAN, M. S., SWITKES, E., & DE VALOIS, R. L. (1982). Deoxyglucose analysis of retinotopic organization in primate striate cortex. Science, 218, 902-904.

WALTER, W. G. (1953). The Living Brain. New York: W. W. Norton & Company, Inc.

WEIMAN, C. F. R. & CHAIKIN, G. (1979). Logarithmic spiral grids for image processing and display. Computer Graphics and Image Processing, 11, 197-226.

WILSON, S. W. (1983). On the retino-cortical mapping. Int. J. Man-Machine Studies, 18, 361-389.

Figure Captions

<u>Fig. 1</u>   (a) Schematic retina showing "data fields," each of which sends signals to its corresponding "message sending unit" (MSU) in the schematic cortex of (b). The MSU labeled F, for example, receives information about the stimulus falling on data field F, processes the information and generates a simple output message summarizing the overall form of that stimulus.

<u>Fig. 2</u>   (a) Visual stimulus imaged on retina of macaque monkey. Rings are spaced in logarithmically equal steps. Solid black rectangle encloses that portion of the stimulus that stimulated the region of striate cortex shown in (b). (b) Pattern of brain activation produced by the visual stimulus of (a), as revealed by 2DG autoradiograph from a flat-mounted tissue section. (The striate region shown would correspond to the upper half of Fig. 1a.) Mapping similarity to Fig. 1 is evidenced by strong tendency of rings and rays of (a) to produce equally spaced bands in (b), by thickening of the bands toward the foveal representation, together with the known (Hubel & Weisel, 1974) functional uniformity of the striate cortex. [From Tootell, <u>et</u> <u>al.</u>, (1982). Reprinted by permission of AAAS.]

<u>Fig. 3</u>   Examples of cross-like patterns taken from subjects' drawings. (Actual sizes in this and following two figures typically 5-10 cm in diameter.)

<u>Fig. 4</u>   Examples of spiroform patterns taken from subjects' drawings.

Fig. 5   Examples of principal non-Spiral, non-Cross forms, from subjects' drawings.

Fig. 6   Components of basic scanning model (M0). (a) Data fields of a fragment of the retinal array RA. (b) The scanners L and D. (c) Arms from L and D shown in relation to the RA fragment of (a) at some instant.

Fig. 7   Cortical correspondent of Fig. 6. Solid scan lines correspond to D; dashed lines to L. Circles indicate MSUs connected to the data fields of RA. Scanning actually occurs cortically, on the outputs of the MSUs, but retinal viewpoint (Fig. 6) is usually taken in text.

Fig. 8   LxD sampling a certain set of data fields (dots indicate data field centers). Arrows: possible perceived directions of array movement in case LxD again falls on data field centers at next flash.

Fig. 9   Diagram representing perceived array motion for Model M0 as a function of V_L/F and V_D/F in sectors per flash. Up vertical arrow stands for expanding radial flow; down for contracting. Left-pointing horizontal arrow stands for counterclockwise rotary flow; right-pointing for clockwise. Diagonal arrows indicate flow along spiral arms. See text for further explanation.

Fig. 10   Some scanning arms of D according to Model M1, with each arm sampling either R or G or B (red or green or blue chromatic components).

Fig. 11   LxD for Model M1 with intersections labeled as to pairs of chromatic components sampled.

Fig. 12   A spatial beat pattern obtained by superposing LxD with N_{LD}=36 upon RA with N_{RA}=32, in accordance with Model M2.

Fig. 13   Diagram representing perceived motion of beat pattern in Model M2 as a function of V_L/F and V_D/F in sectors per flash.   Arrows have same directional significance as in Fig. 9, but here refer to perceived motion of the beat pattern.   See text for further explanation.

Fig. 14   Drawing of "pinwheel" pattern seen in the M2 machine at points p in Fig. 13 for which V_L/F takes integer values.

a

b

Fig. 1

a

b

Fig. 2

Close

$>$: FLOW

Far

Slow

Fig. 3

Fig. 4

CENTER

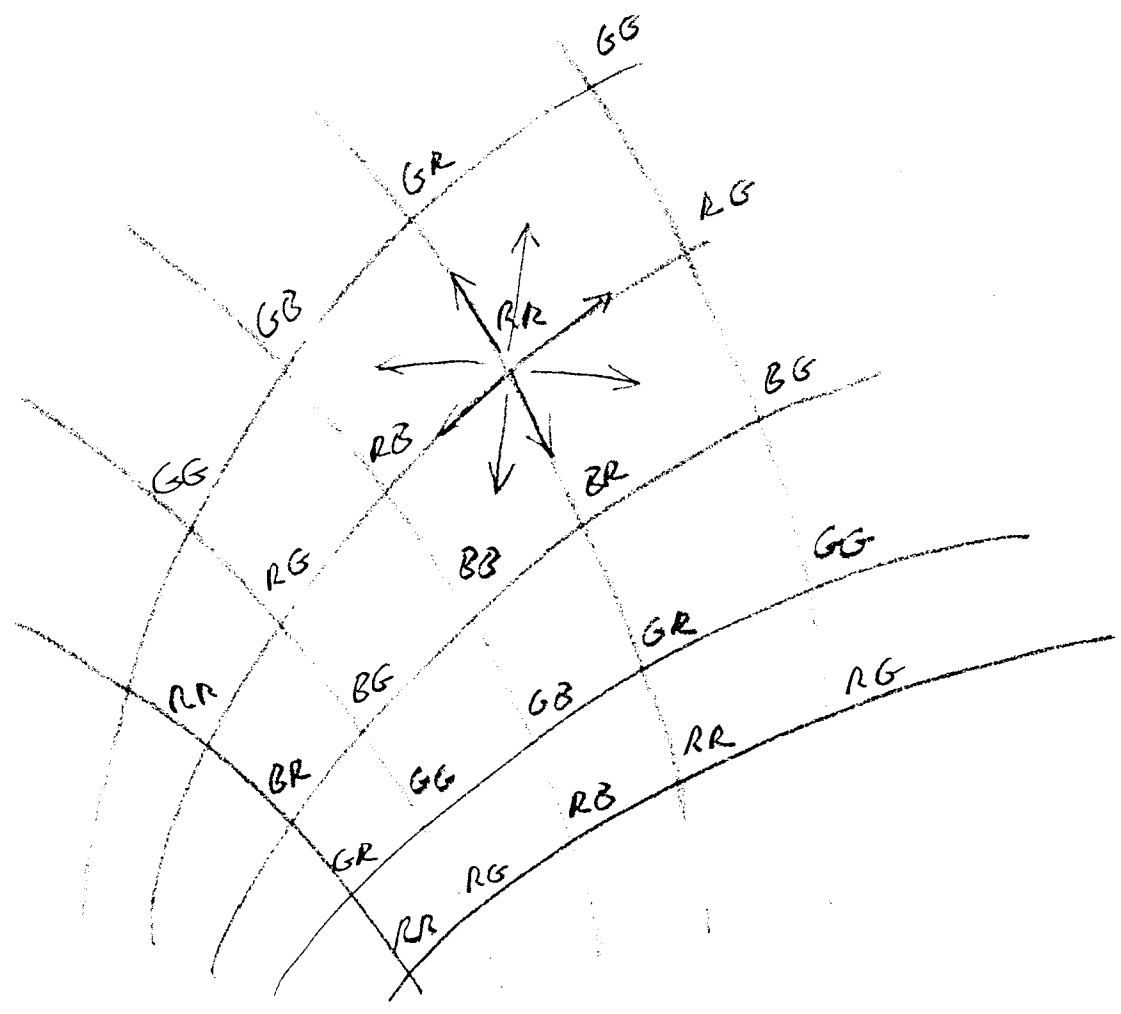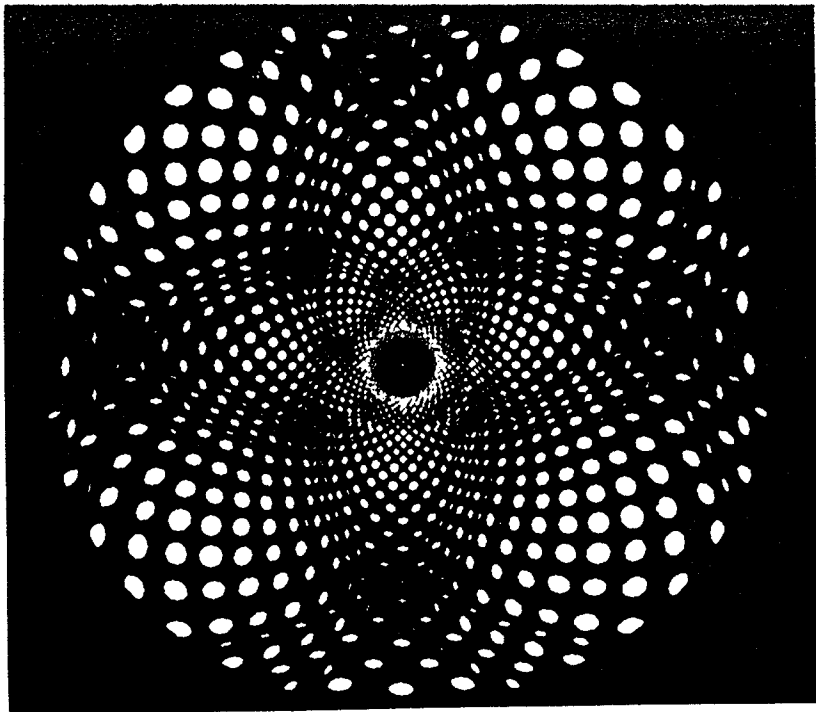moving toward
me

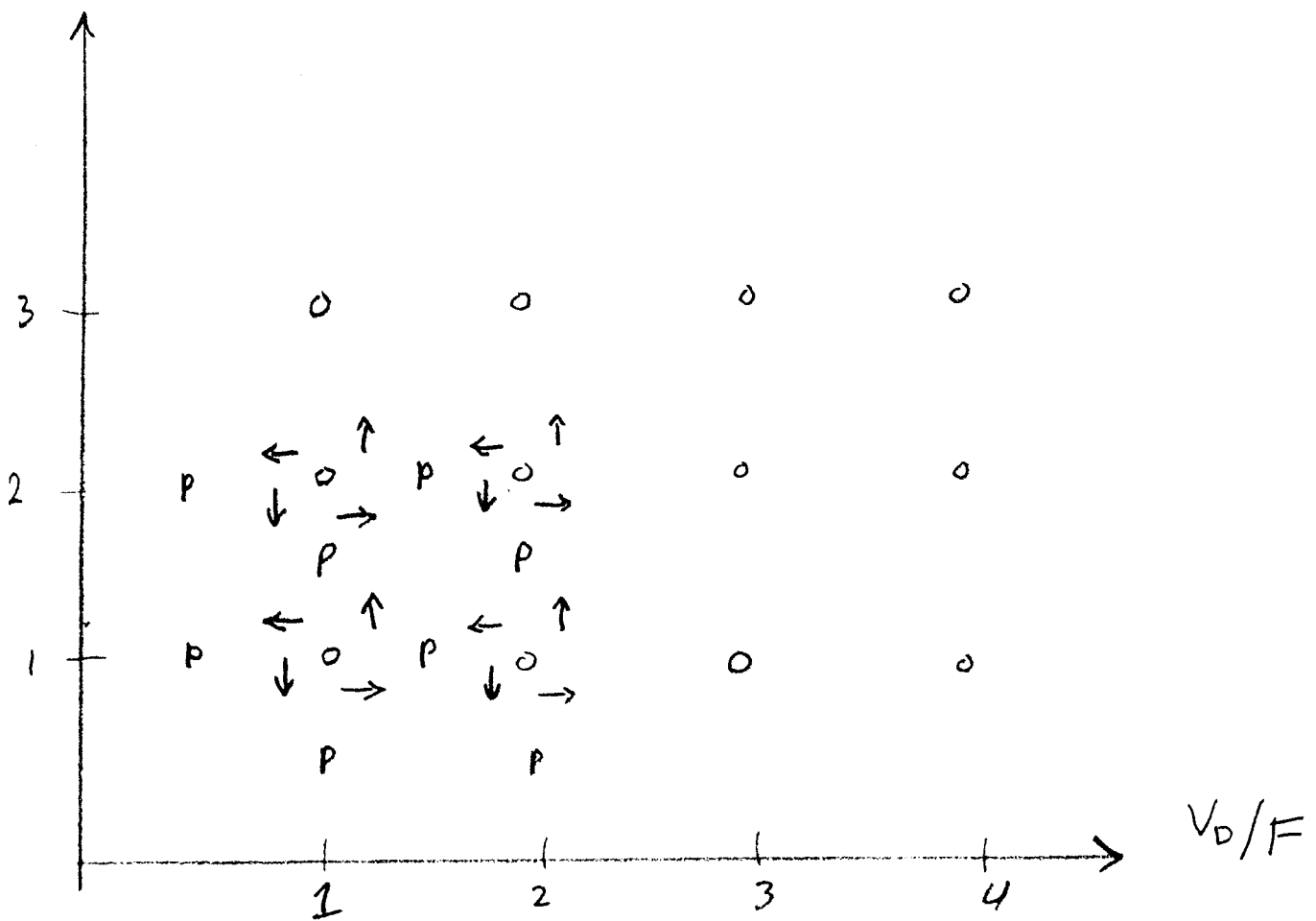C.V.

Fig. 5

RA

a

L

D
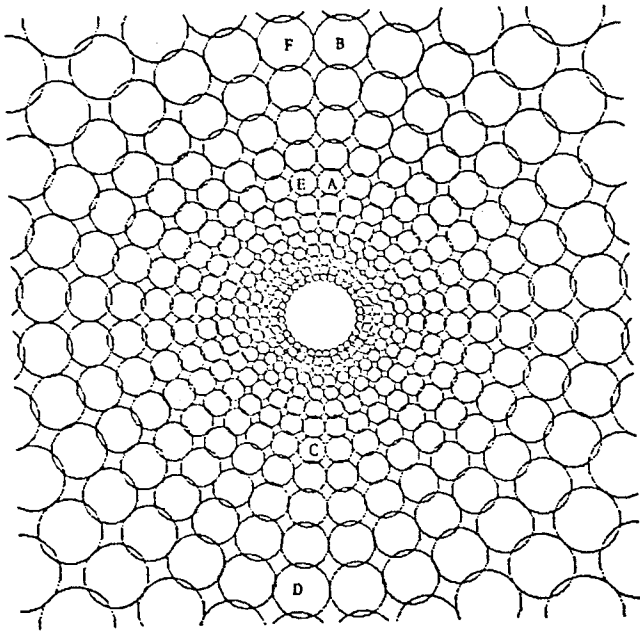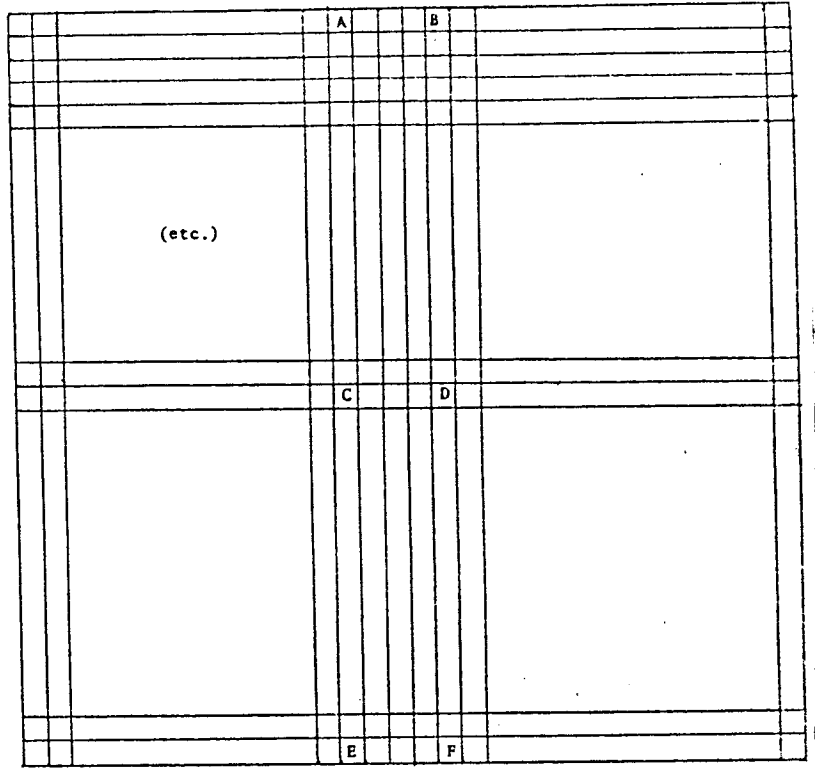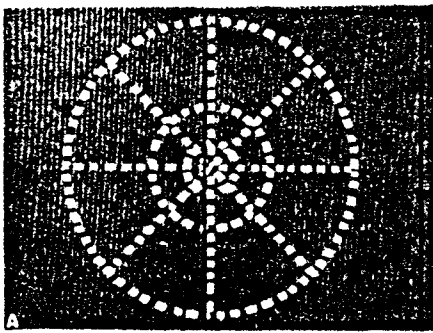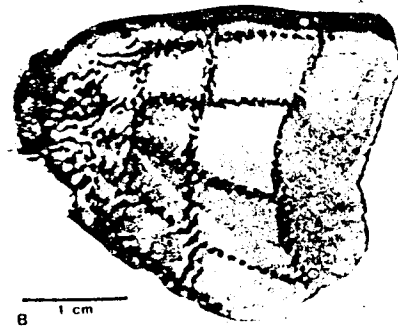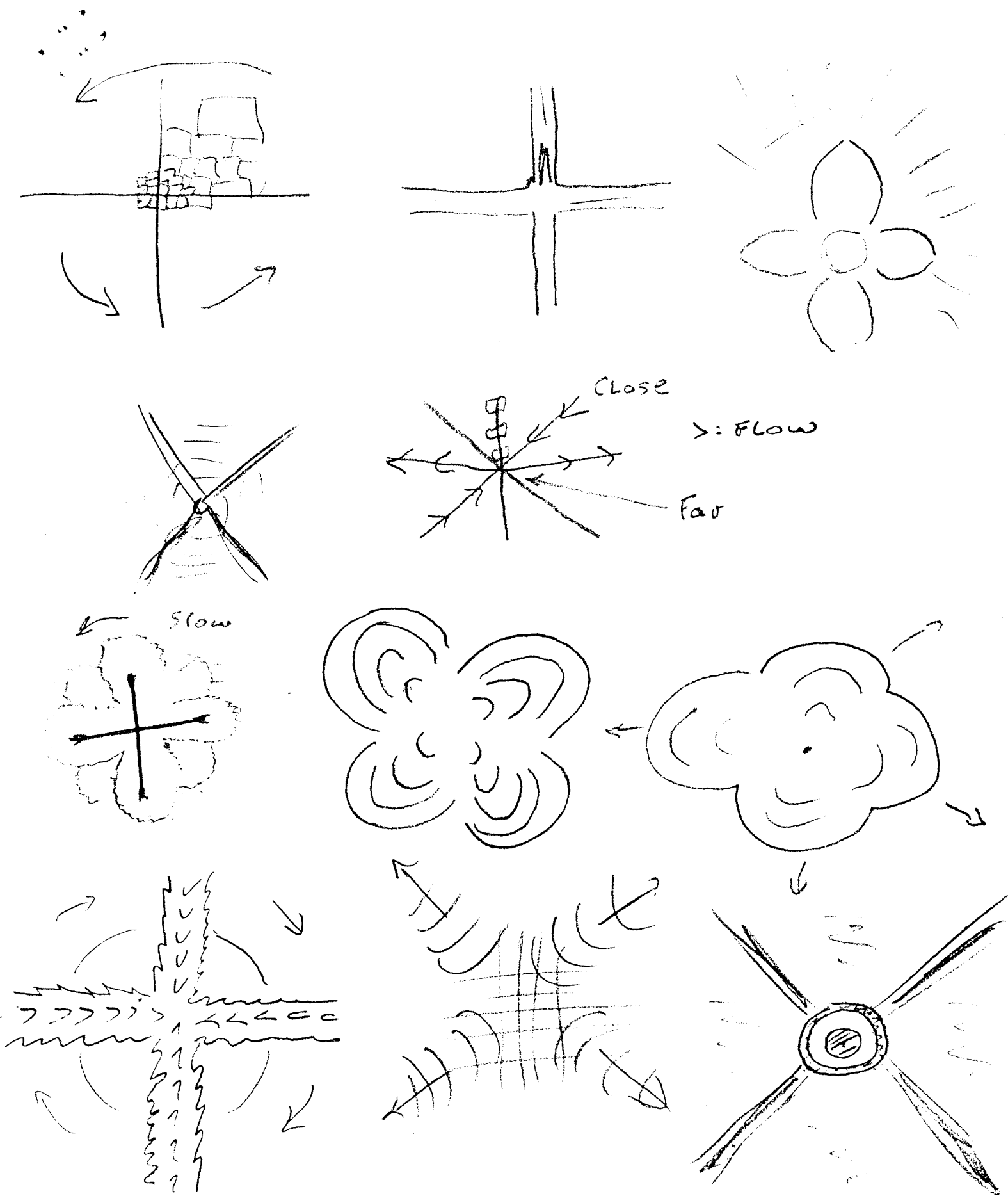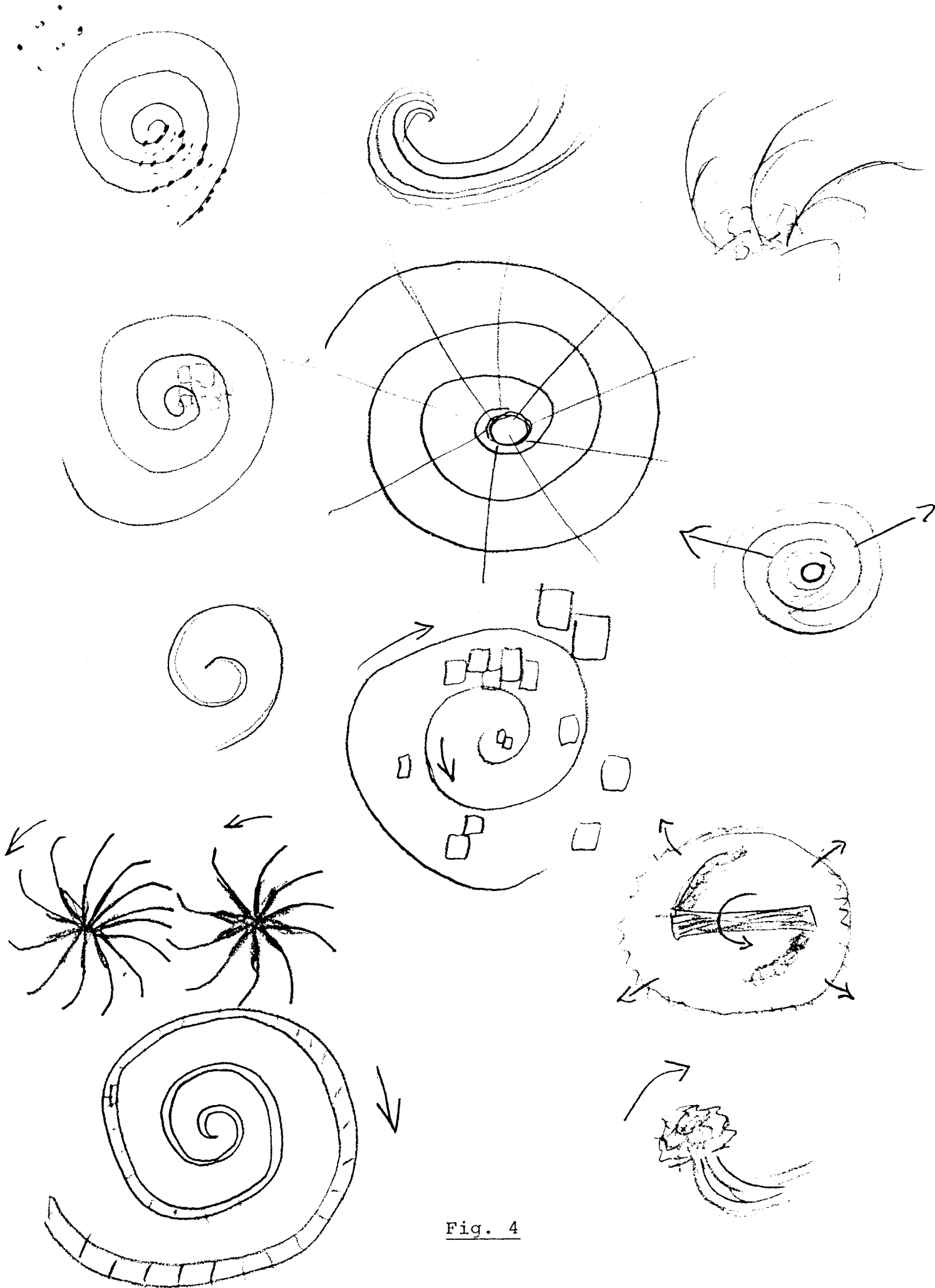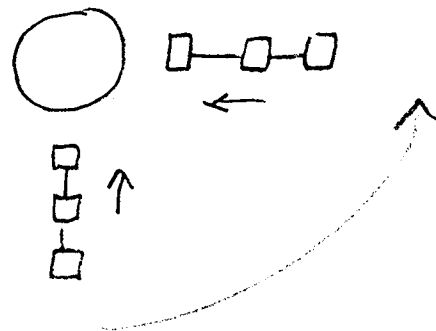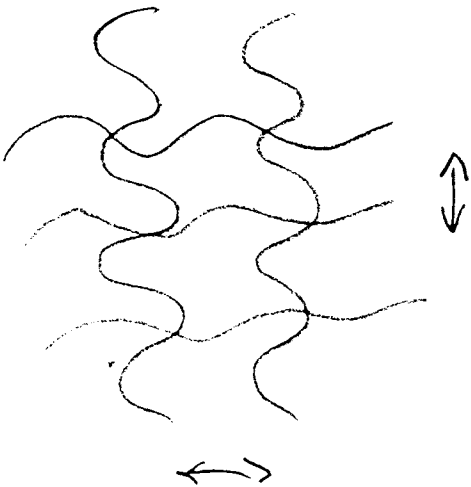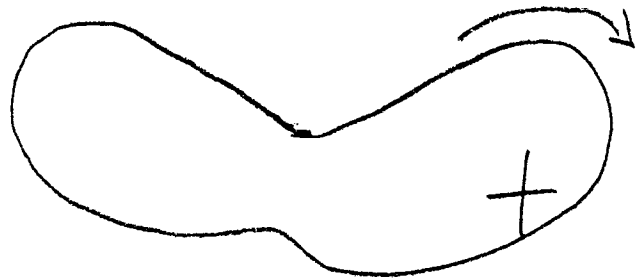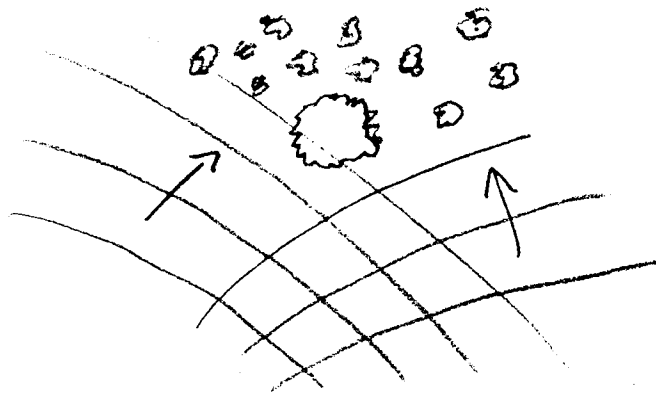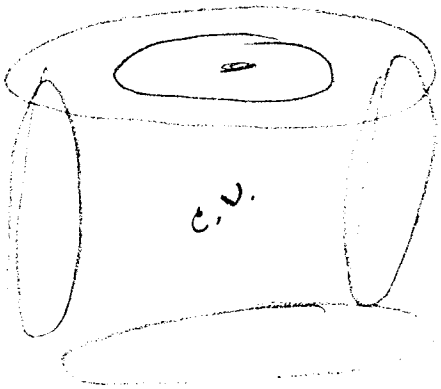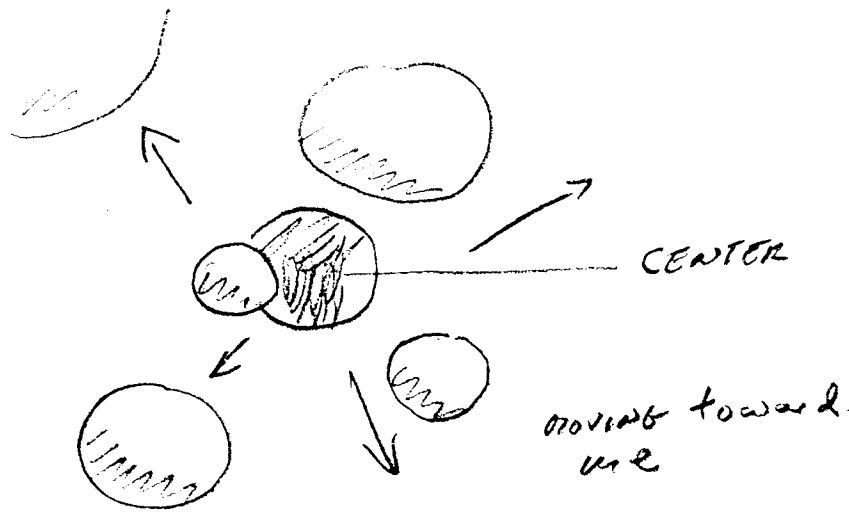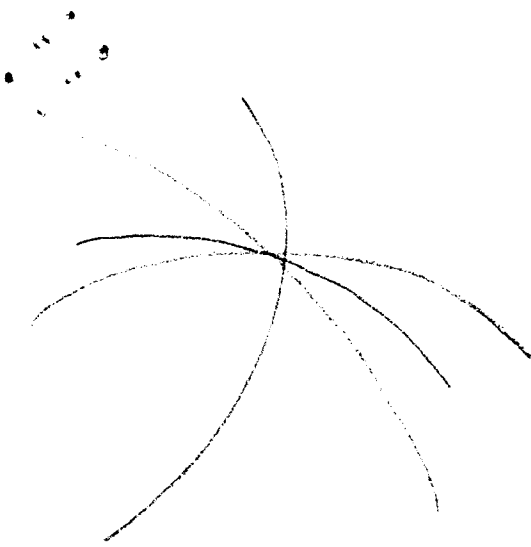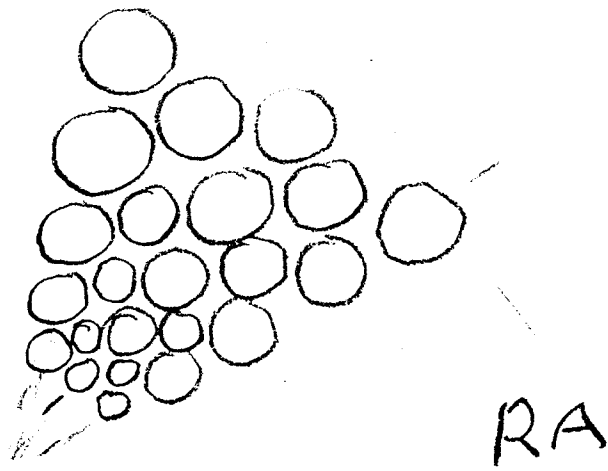
b
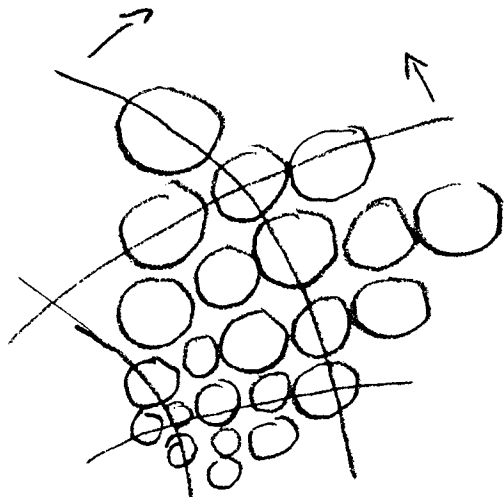
c

Fig. 6
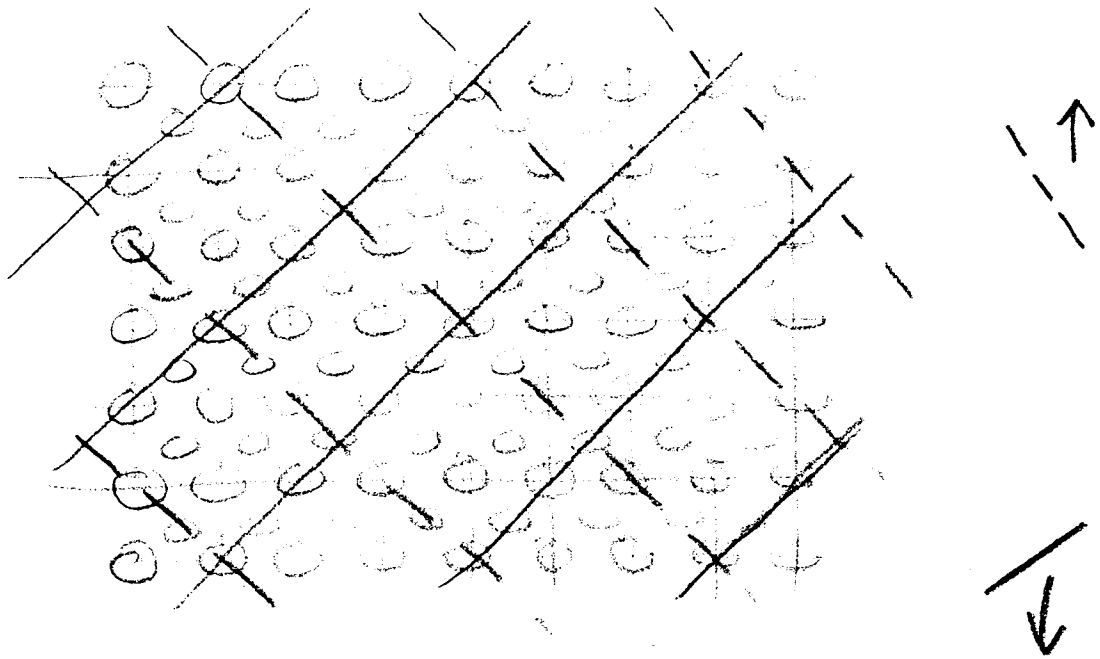
Fig. 7
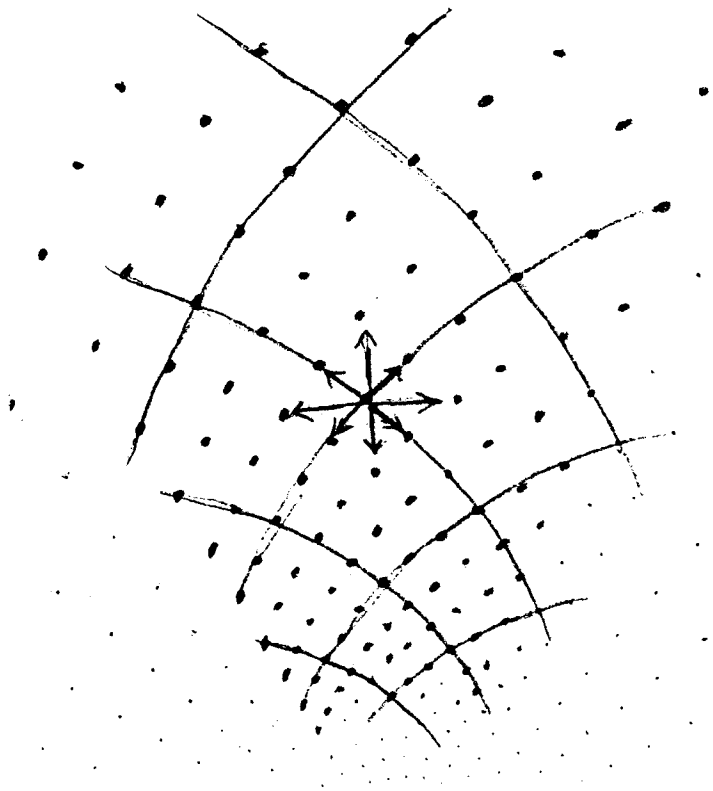
Fig. 8

Fig. 9

Fig. 10

Fig. 11

Fig. 12

Fig. 13

Fig. 14