THE CHEMICAL BASIS OF MORPHOGENESIS

By A. M. TURING, F.R.S. *University of Manchester*

It is suggested that a system of chemical substances, called morphogens, reacting together and diffusing through a tissue, is adequate to account for the main phenomena of morphogenesis. Such a system, although it may originally be quite homogeneous, may later develop a pattern or structure due to an instability of the homogeneous equilibrium, which is triggered off by random disturbances. Such reaction-diffusion systems are considered in some detail in the case of an isolated ring of cells, a mathematically convenient, though biologically unusual system. The investigation is chiefly concerned with the onset of instability. It is found that there are six essentially different forms which this may take. In the most interesting form stationary waves appear on the ring. It is suggested that this might account, for instance, for the tentacle patterns on *Hydra* and for whorled leaves. A system of reactions and diffusion on a sphere is also considered. Such a system appears to account for gastrulation. Another reaction system in two dimensions gives rise to patterns reminiscent of dappling. It is also suggested that stationary waves in two dimensions could account for the phenomena of phyllotaxis.

The purpose of this paper is to discuss a possible mechanism by which the genes of a zygote may determine the anatomical structure of the resulting organism. The theory does not make any new hypotheses; it merely suggests that certain well-known physical laws are sufficient to account for many of the facts. The full understanding of the paper requires a good knowledge of mathematics, some biology, and some elementary chemistry. Since readers cannot be expected to be experts in all of these subjects, a number of elementary facts are explained, which can be found in text-books, but whose omission would make the paper difficult reading.

1. A MODEL OF THE EMBRYO. MORPHOGENS

In this section a mathematical model of the growing embryo will be described. This model will be a simplification and an idealization, and consequently a falsification. It is to be hoped that the features retained for discussion are those of greatest importance in the present state of knowledge.

The model takes two slightly different forms. In one of them the cell theory is recognized but the cells are idealized into geometrical points. In the other the matter of the organism is imagined as continuously distributed. The cells are not, however, completely ignored, for various physical and physico-chemical characteristics of the matter as a whole are assumed to have values appropriate to the cellular matter.

With either of the models one proceeds as with a physical theory and defines an entity called 'the state of the system'. One then describes how that state is to be determined from the state at a moment very shortly before. With either model the description of the state consists of two parts, the mechanical and the chemical. The mechanical part of the state describes the positions, masses, velocities and elastic properties of the cells, and the forces between them. In the continuous form of the theory essentially the same information is given in the form of the stress, velocity, density and elasticity of the matter. The chemical part of the state is given (in the cell form of theory) as the chemical composition of each separate cell; the diffusibility of each substance between each two adjacent cells must also

be given. In the continuous form of the theory the concentrations and diffusibilities of each substance have to be given at each point. In determining the changes of state one should take into account

(i) The changes of position and velocity as given by Newton's laws of motion.

(ii) The stresses as given by the elasticities and motions, also taking into account the osmotic pressures as given from the chemical data.

(iii) The chemical reactions.

(iv) The diffusion of the chemical substances. The region in which this diffusion is possible is given from the mechanical data.

This account of the problem omits many features, e.g. electrical properties and the internal structure of the cell. But even so it is a problem of formidable mathematical complexity. One cannot at present hope to make any progress with the understanding of such systems except in very simplified cases. The interdependence of the chemical and mechanical data adds enormously to the difficulty, and attention will therefore be confined, so far as is possible, to cases where these can be separated. The mathematics of elastic solids is a well-developed subject, and has often been applied to biological systems. In this paper it is proposed to give attention rather to cases where the mechanical aspect can be ignored and the chemical aspect is the most significant. These cases promise greater interest, for the characteristic action of the genes themselves is presumably chemical. The systems actually to be considered consist therefore of masses of tissues which are not growing, but within which certain substances are reacting chemically, and through which they are diffusing. These substances will be called morphogens, the word being intended to convey the idea of a form producer. It is not intended to have any very exact meaning, but is simply the kind of substance concerned in this theory. The evocators of Waddington provide a good example of morphogens (Waddington 1940). These evocators diffusing into a tissue somehow persuade it to develop along different lines from those which would have been followed in its absence. The genes themselves may also be considered to be morphogens. But they certainly form rather a special class. They are quite indiffusible. Moreover, it is only by courtesy that genes can be regarded as separate molecules. It would be more accurate (at any rate at mitosis) to regard them as radicals of the giant molecules known as chromosomes. But presumably these radicals act almost independently, so that it is unlikely that serious errors will arise through regarding the genes as molecules. Hormones may also be regarded as quite typical morphogens. Skin pigments may be regarded as morphogens if desired. But those whose action is to be considered here do not come squarely within any of these categories.

The function of genes is presumed to be purely catalytic. They catalyze the production of other morphogens, which in turn may only be catalysts. Eventually, presumably, the chain leads to some morphogens whose duties are not purely catalytic. For instance, a substance might break down into a number of smaller molecules, thereby increasing the osmotic pressure in a cell and promoting its growth. The genes might thus be said to influence the anatomical form of the organism by determining the rates of those reactions which they catalyze. If the rates are assumed to be those determined by the genes, and if a comparison of organisms is not in question, the genes themselves may be eliminated from the discussion. Likewise any other catalysts obtained secondarily through the agency of

the genes may equally be ignored, if there is no question of their concentrations varying. There may, however, be some other morphogens, of the nature of evocators, which cannot be altogether forgotten, but whose role may nevertheless be subsidiary, from the point of view of the formation of a particular organ. Suppose, for instance, that a 'leg-evocator' morphogen were being produced in a certain region of an embryo, or perhaps diffusing into it, and that an attempt was being made to explain the mechanism by which the leg was formed in the presence of the evocator. It would then be reasonable to take the distribution of the evocator in space and time as given in advance and to consider the chemical reactions set in train by it. That at any rate is the procedure adopted in the few examples considered here.

2. MATHEMATICAL BACKGROUND REQUIRED

The greater part of this present paper requires only a very moderate knowledge of mathematics. What is chiefly required is an understanding of the solution of linear differential equations with constant coefficients. (This is also what is chiefly required for an understanding of mechanical and electrical oscillations.) The solution of such an equation takes the form of a sum $\Sigma A\,e^{bt}$, where the quantities A, b may be complex, i.e. of the form $\alpha+i\beta$, where α and β are ordinary (real) numbers and $i = \sqrt{-1}$. It is of great importance that the physical significance of the various possible solutions of this kind should be appreciated, for instance, that

(a) Since the solutions will normally be real one can also write them in the form $\mathscr{R}\Sigma A\,e^{bt}$ or $\Sigma\mathscr{R}A\,e^{bt}$ ($\mathscr{R}$ means 'real part of').

(b) That if $A = A'\,e^{i\phi}$ and $b = \alpha+i\beta$, where A', α, β, ϕ are real, then

$$\mathscr{R}A\,e^{bt} = A'\,e^{\alpha t}\cos\,(\beta t+\phi).$$

Thus each such term represents a sinusoidal oscillation if $\alpha = 0$, a damped oscillation if $\alpha < 0$, and an oscillation of ever-increasing amplitude if $\alpha > 0$.

(c) If any one of the numbers b has a positive real part the system in question is unstable.

(d) After a sufficiently great lapse of time all the terms $A\,e^{bt}$ will be negligible in comparison with those for which b has the greatest real part, but unless this greatest real part is itself zero these dominant terms will eventually either tend to zero or to infinite values.

(e) That the indefinite growth mentioned in (b) and (d) will in any physical or biological situation eventually be arrested due to a breakdown of the assumptions under which the solution was valid. Thus, for example, the growth of a colony of bacteria will normally be taken to satisfy the equation $dy/dt = \alpha y$ ($\alpha > 0$), y being the number of organisms at time t, and this has the solution $y = A\,e^{\alpha t}$. When, however, the factor $e^{\alpha t}$ has reached some billions the food supply can no longer be regarded as unlimited and the equation $dy/dt = \alpha y$ will no longer apply.

The following relatively elementary result will be needed, but may not be known to all readers:

$$\sum_{r=1}^{N} \exp\left[\frac{2\pi irs}{N}\right] = 0 \quad \text{if} \quad 0 < s < N,$$

but
$$= N \quad \text{if} \quad s = 0 \text{ or } s = N.$$

The first case can easily be proved when it is noticed that the left-hand side is a geometric progression. In the second case all the terms are equal to 1.

The relative degrees of difficulty of the various sections are believed to be as follows. Those who are unable to follow the points made in this section should only attempt §§ 3, 4, 11, 12, 14 and part of § 13. Those who can just understand this section should profit also from §§ 7, 8, 9. The remainder, §§ 5, 10, 13, will probably only be understood by those definitely trained as mathematicians.

3. Chemical reactions

It has been explained in a preceding section that the system to be considered consists of a number of chemical substances (morphogens) diffusing through a mass of tissue of given geometrical form and reacting together within it. What laws are to control the development of this situation? They are quite simple. The diffusion follows the ordinary laws of diffusion, i.e. each morphogen moves from regions of greater to regions of less concentration, at a rate proportional to the gradient of the concentration, and also proportional to the 'diffusibility' of the substance. This is very like the conduction of heat, diffusibility taking the place of conductivity. If it were not for the walls of the cells the diffusibilities would be inversely proportional to the square roots of the molecular weights. The pores of the cell walls put a further handicap on the movement of the larger molecules in addition to that imposed by their inertia, and most of them are not able to pass through the walls at all.

The reaction rates will be assumed to obey the 'law of mass action'. This states that the rate at which a reaction takes place is proportional to the concentrations of the reacting substances. Thus, for instance, the rate at which silver chloride will be formed and precipitated from a solution of silver nitrate and sodium chloride by the reaction

$$Ag^+ + Cl^- \rightarrow AgCl$$

will be proportional to the product of the concentrations of the silver ion Ag^+ and the chloride ion Cl^-. It should be noticed that the equation

$$AgNO_3 + NaCl \rightarrow AgCl + NaNO_3$$
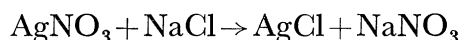
is not used because it does not correspond to an actual reaction but to the final outcome of a number of reactions. The law of mass action must only be applied to the *actual* reactions. Very often certain substances appear in the individual reactions of a group, but not in the final outcome. For instance, a reaction $A \rightarrow B$ may really take the form of two steps $A + G \rightarrow C$ and $C \rightarrow B + G$. In such a case the substance G is described as a catalyst, and as catalyzing the reaction $A \rightarrow B$. (Catalysis according to this plan has been considered in detail by Michaelis & Menten (1913).) The effect of the genes is presumably achieved almost entirely by catalysis. They are certainly not permanently used up in the reactions.

Sometimes one can regard the effect of a catalyst as merely altering a reaction rate. Consider, for example, the case mentioned above, but suppose also that A can become detached from G, i.e. that the reaction $C \rightarrow A + G$ is taken into account. Also suppose that the reactions $A + G \rightleftharpoons C$ both proceed much faster than $C \rightarrow B + G$. Then the concentrations of A, G, C will be related by the condition that there is equilibrium between the reactions $A + G \rightarrow C$ and $C \rightarrow A + G$, so that (denoting concentrations by square brackets) $[A][G] = k[C]$ for some constant k. The reaction $C \rightarrow B + G$ will of course proceed at a rate proportional to $[C]$, i.e. to $[A][G]$. If the amount of C is always small compared with the amount of G one can say that the presence of the catalyst and its amount merely alter the mass action constant

for the reaction $A \to B$, for the whole proceeds at a rate proportional to $[A]$. This situation does not, however, hold invariably. It may well happen that nearly all of G takes the combined form C so long as any of A is left. In this case the reaction proceeds at a rate independent of the concentration of A until A is entirely consumed. In either of these cases the rate of the complete group of reactions depends only on the concentrations of the reagents, although usually not according to the law of mass action applied crudely to the chemical equation for the whole group. The same applies in any case where all reactions of the group with one exception proceed at speeds much greater than that of the exceptional one. In these cases the rate of the reaction is a function of the concentrations of the reagents. More generally again, no such approximation is applicable. One simply has to take all the actual reactions into account.

According to the cell model then, the number and positions of the cells are given in advance, and so are the rates at which the various morphogens diffuse between the cells. Suppose that there are N cells and M morphogens. The state of the whole system is then given by MN numbers, the quantities of the M morphogens in each of N cells. These numbers change with time, partly because of the reactions, partly because of the diffusion. To determine the part of the rate of change of one of these numbers due to diffusion, at any one moment, one only needs to know the amounts of the same morphogen in the cell and its neighbours, and the diffusion coefficient for that morphogen. To find the rate of change due to chemical reaction one only needs to know the concentrations of all morphogens at that moment in the one cell concerned.

This description of the system in terms of the concentrations in the various cells is, of course, only an approximation. It would be justified if, for instance, the contents were perfectly stirred. Alternatively, it may often be justified on the understanding that the 'concentration in the cell' is the concentration at a certain representative point, although the idea of 'concentration at a point' clearly itself raises difficulties. The author believes that the approximation is a good one, whatever argument is used to justify it, and it is certainly a convenient one.

It would be possible to extend much of the theory to the case of organisms immersed in a fluid, considering the diffusion within the fluid as well as from cell to cell. Such problems are not, however, considered here.

4. THE BREAKDOWN OF SYMMETRY AND HOMOGENEITY

There appears superficially to be a difficulty confronting this theory of morphogenesis, or, indeed, almost any other theory of it. An embryo in its spherical blastula stage has spherical symmetry, or if there are any deviations from perfect symmetry, they cannot be regarded as of any particular importance, for the deviations vary greatly from embryo to embryo within a species, though the organisms developed from them are barely distinguishable. One may take it therefore that there is perfect spherical symmetry. But a system which has spherical symmetry, and whose state is changing because of chemical reactions and diffusion, will remain spherically symmetrical for ever. (The same would hold true if the state were changing according to the laws of electricity and magnetism, or of quantum mechanics.) It certainly cannot result in an organism such as a horse, which is not spherically symmetrical.

There is a fallacy in this argument. It was assumed that the deviations from spherical symmetry in the blastula could be ignored because it makes no particular difference what form of asymmetry there is. It is, however, important that there are *some* deviations, for the system may reach a state of instability in which these irregularities, or certain components of them, tend to grow. If this happens a new and stable equilibrium is usually reached, with the symmetry entirely gone. The variety of such new equilibria will normally not be so great as the variety of irregularities giving rise to them. In the case, for instance, of the gastrulating sphere, discussed at the end of this paper, the direction of the axis of the gastrula can vary, but nothing else.

The situation is very similar to that which arises in connexion with electrical oscillators. It is usually easy to understand how an oscillator keeps going when once it has started, but on a first acquaintance it is not obvious how the oscillation begins. The explanation is that there are random disturbances always present in the circuit. Any disturbance whose frequency is the natural frequency of the oscillator will tend to set it going. The ultimate fate of the system will be a state of oscillation at its appropriate frequency, and with an amplitude (and a wave form) which are also determined by the circuit. The phase of the oscillation alone is determined by the disturbance.

If chemical reactions and diffusion are the only forms of physical change which are taken into account the argument above can take a slightly different form. For if the system originally has no sort of geometrical symmetry but is a perfectly homogeneous and possibly irregularly shaped mass of tissue, it will continue indefinitely to be homogeneous. In practice, however, the presence of irregularities, including statistical fluctuations in the numbers of molecules undergoing the various reactions, will, if the system has an appropriate kind of instability, result in this homogeneity disappearing.

This breakdown of symmetry or homogeneity may be illustrated by the case of a pair of cells originally having the same, or very nearly the same, contents. The system is homogeneous: it is also symmetrical with respect to the operation of interchanging the cells. The contents of either cell will be supposed describable by giving the concentrations X and Y of two morphogens. The chemical reactions will be supposed such that, on balance, the first morphogen (X) is produced at the rate $5X - 6Y + 1$ and the second (Y) at the rate $6X - 7Y + 1$. When, however, the strict application of these formulae would involve the concentration of a morphogen in a cell becoming negative, it is understood that it is instead destroyed only at the rate at which it is reaching that cell by diffusion. The first morphogen will be supposed to duffuse at the rate 0·5 for unit difference of concentration between the cells, the second, for the same difference, at the rate 4·5. Now if both morphogens have unit concentration in both cells there is equilibrium. There is no resultant passage of either morphogen across the cell walls, since there is no concentration difference, and there is no resultant production (or destruction) of either morphogen in either cell since $5X - 6Y + 1$ and $6X - 7Y + 1$ both have the value zero for $X = 1$, $Y = 1$. But suppose the values are $X_1 = 1·06$, $Y_1 = 1·02$ for the first cell and $X_2 = 0·94$, $Y_2 = 0·98$ for the second. Then the two morphogens will be being produced by chemical action at the rates 0·18, 0·22 respectively in the first cell and destroyed at the same rates in the second. At the same time there is a flow due to diffusion from the first cell to the second at the rate 0·06 for the first morphogen and 0·18 for the second. In sum the effect is a flow from the second cell to the first at the

rates 0·12, 0·04 for the two morphogens respectively. This flow tends to accentuate the already existing differences between the two cells. More generally, if

$$X_1 = 1+3\xi, \quad X_2 = 1-3\xi, \quad Y_1 = 1+\xi, \quad Y_2 = 1-\xi,$$

at some moment the four concentrations continue afterwards to be expressible in this form, and ξ increases at the rate 2ξ. Thus there is an exponential drift away from the equilibrium condition. It will be appreciated that a drift away from the equilibrium occurs with almost any small displacement from the equilibrium condition, though not normally according to an exact exponential curve. A particular choice was made in the above argument in order to exhibit the drift with only very simple mathematics.

Before it can be said to follow that a two-cell system can be unstable, with inhomogeneity succeeding homogeneity, it is necessary to show that the reaction rate functions postulated really can occur. To specify actual substances, concentrations and temperatures giving rise to these functions would settle the matter finally, but would be difficult and somewhat out of the spirit of the present inquiry. Instead, it is proposed merely to mention imaginary reactions which give rise to the required functions by the law of mass action, if suitable reaction constants are assumed. It will be sufficient to describe

(i) A set of reactions producing the first morphogen at the constant rate 1, and a similar set forming the second morphogen at the same rate.

(ii) A set destroying the second morphogen (Y) at the rate $7Y$.

(iii) A set converting the first morphogen (X) into the second (Y) at the rate $6X$.

(iv) A set producing the first morphogen (X) at the rate $11X$.

(v) A set destroying the first morphogen (X) at the rate $6Y$, so long as any of it is present.

The conditions of (i) can be fulfilled by reactions of the type $A \to X$, $B \to Y$, where A and B are substances continually present in large and invariable concentrations. The conditions of (ii) are satisfied by a reaction of the form $Y \to D$, D being an inert substance and (iii) by the reaction $X \to Y$ or $X \to Y+E$. The remaining two sets are rather more difficult. To satisfy the conditions of (iv) one may suppose that X is a catalyst for its own formation from A. The actual reactions could be the formation of an unstable compound U by the reaction $A+X \to U$, and the subsequent almost instantaneous breakdown $U \to 2X$. To destroy X at a rate proportional to Y as required in (v) one may suppose that a catalyst C is present in small but constant concentration and immediately combines with X, $X+C \to V$. The modified catalyst reacting with Y, at a rate proportional to Y, restores the catalyst but not the morphogen X, by the reactions $V+Y \to W$, $W \to C+H$, of which the latter is assumed instantaneous.

It should be emphasized that the reactions here described are by no means those which are most likely to give rise to instability in nature. The choice of the reactions to be discussed was dictated entirely by the fact that it was desirable that the argument be easy to follow. More plausible reaction systems are described in § 10.

Unstable equilibrium is not, of course, a condition which occurs very naturally. It usually requires some rather artificial interference, such as placing a marble on the top of a dome. Since systems tend to leave unstable equilibria they cannot often be in them. Such equilibria can, however, occur naturally through a stable equilibrium changing into an unstable one. For example, if a rod is hanging from a point a little above its centre of gravity

it will be in stable equilibrium. If, however, a mouse climbs up the rod the equilibrium eventually becomes unstable and the rod starts to swing. A chemical analogue of this mouse-and-pendulum system would be that described above with the same diffusibilities but with the two morphogens produced at the rates

$$(3+I)\,X-6Y+I-1 \quad \text{and} \quad 6X-(9+I)\,Y-I+1.$$

This system is stable if $I<0$ but unstable if $I>0$. If I is allowed to increase, corresponding to the mouse running up the pendulum, it will eventually become positive and the equilibrium will collapse. The system which was originally discussed was the case $I=2$, and might be supposed to correspond to the mouse somehow reaching the top of the pendulum without disaster, perhaps by falling vertically on to it.

5. Left-handed and right-handed organisms

The object of this section is to discuss a certain difficulty which might be thought to show that the morphogen theory of morphogenesis cannot be right. The difficulty is mainly concerned with organisms which have not got bilateral symmetry. The argument, although carried through here without the use of mathematical formulae, may be found difficult by non-mathematicians, and these are therefore recommended to ignore it unless they are already troubled by such a difficulty.

An organism is said to have 'bilateral symmetry' if it is identical with its own reflexion in some plane. This plane of course always has to pass through some part of the organism, in particular through its centre of gravity. For the purpose of this argument it is more general to consider what may be called 'left-right symmetry'. An organism has left-right symmetry if its description in any right-handed set of rectangular Cartesian co-ordinates is identical with its description in some set of left-handed axes. An example of a body with left-right symmetry, but not bilateral symmetry, is a cylinder with the letter P printed on one end, and with the mirror image of a P on the other end, but with the two upright strokes of the two letters not parallel. The distinction may possibly be without a difference so far as the biological world is concerned, but mathematically it should not be ignored.

If the organisms of a species are sufficiently alike, and the absence of left-right symmetry sufficiently pronounced, it is possible to describe each individual as either right-handed or left-handed without there being difficulty in classifying any particular specimen. In man, for instance, one could take the X-axis in the forward direction, the Y-axis at right angles to it in the direction towards the side on which the heart is felt, and the Z-axis upwards. The specimen is classed as left-handed or right-handed according as the axes so chosen are left-handed or right-handed. A new classification has of course to be defined for each species.

The fact that there exist organisms which do not have left-right symmetry does not in itself cause any difficulty. It has already been explained how various kinds of symmetry can be lost in the development of the embryo, due to the particular disturbances (or 'noise') influencing the particular specimen not having that kind of symmetry, taken in conjunction with appropriate kinds of instability. The difficulty lies in the fact that there are species in which the proportions of left-handed and right-handed types are very unequal. It will be as well to describe first an argument which appears to show that this should not happen.

The argument is very general, and might be applied to a very wide class of theories of morphogenesis.

An entity may be described as 'P-symmetrical' if its description in terms of one set of right-handed axes is identical with its description in terms of any other set of right-handed axes with the same origin. Thus, for instance, the totality of positions that a corkscrew would take up when rotated in all possible ways about the origin has P-symmetry. The entity will be said to be 'F-symmetrical' when changes from right-handed axes to left-handed may also be made. This would apply if the corkscrew were replaced by a bilaterally symmetrical object such as a coal scuttle, or a left-right symmetrical object. In these terms one may say that there are species such that the totality of specimens from that species, together with the rotated specimens, is P-symmetrical, but very far from F-symmetrical. On the other hand, it is reasonable to suppose that

(i) The laws of physics are F-symmetrical.

(ii) The initial totality of zygotes for the species is F-symmetrical.

(iii) The statistical distribution of disturbances is F-symmetrical. The individual disturbances of course will in general have neither F-symmetry nor P-symmetry.

It should be noticed that the ideas of P-symmetry and F-symmetry as defined above apply even to so elaborate an entity as 'the laws of physics'. It should also be understood that the laws are to be the laws taken into account in the theory in question rather than some ideal as yet undiscovered laws.

Now it follows from these assumptions that the statistical distribution of resulting organisms will have F-symmetry, or more strictly that the distribution deduced as the result of working out such a theory will have such symmetry. The distribution of observed mature organisms, however, has no such symmetry. In the first place, for instance, men are more often found standing on their feet than their heads. This may be corrected by taking gravity into account in the laws, together with an appropriate change of definition of the two kinds of symmetry. But it will be more convenient if, for the sake of argument, it is imagined that some species has been reared in the absence of gravity, and that the resulting distribution of mature organisms is found to be P-symmetrical but to yield more right-handed specimens than left-handed and so not to have F-symmetry. It remains therefore to explain this absence of F-symmetry.

Evidently one or other of the assumptions (i) to (iii) must be wrong, i.e. in a correct theory one of them would not apply. In the morphogen theory already described these three assumptions do all apply, and it must therefore be regarded as defective to some extent. The theory may be corrected by taking into account the fact that the morphogens do not always have an equal number of left- and right-handed molecules. According to one's point of view one may regard this as invalidating either (i), (ii) or even (iii). Simplest perhaps is to say that the totality of zygotes just is not F-symmetrical, and that this could be seen if one looked at the molecules. This is, however, not very satisfactory from the point of view of this paper, as it would not be consistent with describing states in terms of concentrations only. It would be preferable if it was found possible to find more accurate laws concerning reactions and diffusion. For the purpose of accounting for unequal numbers of left- and right-handed organisms it is unnecessary to do more than show that there are corrections which would not be F-symmetrical when there are laevo- or dextrorotatory

morphogens, and which would be large enough to account for the effects observed. It is not very difficult to think of such effects. They do not have to be very large, but must, of course, be larger than the purely statistical effects, such as thermal noise or Brownian movement.

There may also be other reasons why the totality of zygotes is not F-symmetrical, e.g. an asymmetry of the chromosomes themselves. If these also produce a sufficiently large effect, so much the better.

Though these effects may be large compared with the statistical disturbances they are almost certainly small compared with the ordinary diffusion and reaction effects. This will mean that they only have an appreciable effect during a short period in which the breakdown of left-right symmetry is occurring. Once their existence is admitted, whether on a theoretical or experimental basis, it is probably most convenient to give them mathematical expression by regarding them as P-symmetrically (but not F-symmetrically) distributed disturbances. However, they will not be considered further in this paper.

6. Reactions and diffusion in a ring of cells

The original reason for considering the breakdown of homogeneity was an apparent difficulty in the diffusion-reaction theory of morphogenesis. Now that the difficulty is resolved it might be supposed that there is no reason for pursuing this aspect of the problem further, and that it would be best to proceed to consider what occurs when the system is very far from homogeneous. A great deal more attention will nevertheless be given to the breakdown of homogeneity. This is largely because the assumption that the system is still nearly homogeneous brings the problem within the range of what is capable of being treated mathematically. Even so many further simplifying assumptions have to be made. Another reason for giving this phase such attention is that it is in a sense the most critical period. That is to say, that if there is any doubt as to how the organism is going to develop it is conceivable that a minute examination of it just after instability has set in might settle the matter, but an examination of it at any earlier time could never do so.

There is a great variety of geometrical arrangement of cells which might be considered, but one particular type of configuration stands out as being particularly simple in its theory, and also illustrates the general principles very well. This configuration is a ring of similar cells. One may suppose that there are N such cells. It must be admitted that there is no biological example to which the theory of the ring can be immediately applied, though it is not difficult to find ones in which the principles illustrated by the ring apply.

It will be assumed at first that there are only two morphogens, or rather only two interesting morphogens. There may be others whose concentration does not vary either in space or time, or which can be eliminated from the discussion for one reason or another. These other morphogens may, for instanse, be catalysts involved in the reactions between the interesting morphogens. An example of a complete system of reactions is given in § 10. Some consideration will also be given in §§ 8, 9 to the case of three morphogens. The reader should have no difficulty in extending the results to any number of morphogens, but no essentially new features appear when the number is increased beyond three.

The two morphogens will be called X and Y. These letters will also be used to denote their concentrations. This need not lead to any real confusion. The concentration of X in

cell r may be written X_r, and Y_r has a similar meaning. It is convenient to regard 'cell N' and 'cell O' as synonymous, and likewise 'cell 1' and cell '$N+1$'. One can then say that for each r satisfying $1 \leqslant r \leqslant N$ cell r exchanges material by diffusion with cells $r-1$ and $r+1$. The cell-to-cell diffusion constant for X will be called μ, and that for Y will be called ν. This means that for unit concentration difference of X, this morphogen passes at the rate μ from the cell with the higher concentration to the (neighbouring) cell with the lower concentration. It is also necessary to make assumptions about the rates of chemical reaction. The most general assumption that can be made is that for concentrations X and Y chemical reactions are tending to increase X at the rate $f(X, Y)$ and Y at the rate $g(X, Y)$. When the changes in X and Y due to diffusion are also taken into account the behaviour of the system may be described by the $2N$ differential equations

$$\left. \begin{aligned} \frac{dX_r}{dt} &= f(X_r, Y_r) + \mu(X_{r+1} - 2X_r + X_{r-1}) \\ \frac{dY_r}{dt} &= g(X_r, Y_r) + \nu(Y_{r+1} - 2Y_r + Y_{r-1}) \end{aligned} \right\} \quad (r = 1, ..., N). \tag{6.1}$$

If $f(h, k) : g(h, k) = 0$, then an isolated cell has an equilibrium with concentrations $X = h$, $Y = k$. The ring system also has an equilibrium, stable or unstable, with each X_r equal to h and each Y_r equal to k. Assuming that the system is not very far from this equilibrium it is convenient to put $X_r = h + x_r$, $Y_r = k + y_r$. One may also write $ax + by$ for $f(h+x, y+k)$ and $cx + dy$ for $g(h+x, y+k)$. Since $f(h, k) = g(h, k) = 0$ no constant terms are required, and since x and y are supposed small the terms in higher powers of x and y will have relatively little effect and one is justified in ignoring them. The four quantities a, b, c, d may be called the 'marginal reaction rates'. Collectively they may be described as the 'marginal reaction rate matrix'. When there are M morphogens this matrix consists of M^2 numbers. A marginal reaction rate has the dimensions of the reciprocal of a time, like a radioactive decay rate, which is in fact an example of a marginal (nuclear) reaction rate.

With these assumptions the equations can be rewritten as

$$\left. \begin{aligned} \frac{dx_r}{dt} &= ax_r + by_r + \mu(x_{r+1} - 2x_r + x_{r-1}), \\ \frac{dy_r}{dt} &= cx_r + dy_r + \nu(y_{r+1} - 2y_r + y_{r-1}). \end{aligned} \right\} \tag{6.2}$$

To solve the equations one introduces new co-ordinates $\xi_0, ..., \xi_{N-1}$ and $\eta_0, ..., \eta_{N-1}$ by putting

$$\left. \begin{aligned} x_r &= \sum_{s=0}^{N-1} \exp\left[\frac{2\pi irs}{N}\right] \xi_s, \\ y_r &= \sum_{s=0}^{N-1} \exp\left[\frac{2\pi irs}{N}\right] \eta_s. \end{aligned} \right\} \tag{6.3}$$

These relations can also be written as

$$\left. \begin{aligned} \xi_r &= \frac{1}{N} \sum_{s=1}^{N} \exp\left[-\frac{2\pi irs}{N}\right] x_s, \\ \eta_r &= \frac{1}{N} \sum_{s=1}^{N} \exp\left[-\frac{2\pi irs}{N}\right] y_s. \end{aligned} \right\} \tag{6.4}$$

as may be shown by using the equations

$$\sum_{s=1}^{N} \exp\left[\frac{2\pi i r s}{N}\right] = 0 \quad \text{if} \quad 0 < r < N,$$
$$= N \quad \text{if} \quad r = 0 \text{ or } r = N, \tag{6.5}$$

(referred to in §2). Making this substitution one obtains

$$\frac{d\xi_s}{dt} = \frac{1}{N}\sum_{s=1}^{N} \exp\left[-\frac{2\pi i r s}{N}\right]\left[ax_r + by_r + \mu\left(\exp\left[-\frac{2\pi i s}{N}\right] - 2 + \exp\left[\frac{2\pi i s}{N}\right]\right)\xi_s\right]$$

$$= a\xi_s + b\eta_s + \mu\left(\exp\left[-\frac{2\pi i s}{N}\right] - 2 + \exp\left[\frac{2\pi i s}{N}\right]\right)\xi_s$$

$$= \left(a - 4\mu \sin^2\frac{\pi s}{N}\right)\xi_s + b\eta_s. \tag{6.6}$$

Likewise
$$\frac{d\eta_s}{dt} = c\xi_s + \left(d - 4\nu \sin^2\frac{\pi s}{N}\right)\eta_s. \tag{6.7}$$

The equations have now been converted into a quite manageable form, with the variables separated. There are now two equations concerned with ξ_1 and η_1, two concerned with ξ_2 and η_2, etc. The equations themselves are also of a well-known standard form, being linear with constant coefficients. Let p_s and p_s' be the roots of the equation

$$\left(p - a + 4\mu \sin^2\frac{\pi s}{N}\right)\left(p - d + 4\nu \sin^2\frac{\pi s}{N}\right) = bc \tag{6.8}$$

(with $\mathscr{R}p_s \geqslant \mathscr{R}p_s'$ for definiteness), then the solution of the equations is of the form

$$\left.\begin{aligned}\xi_s &= A_s e^{p_s t} + B_s e^{p_s' t},\\ \eta_s &= C_s e^{p_s t} + D_s e^{p_s' t},\end{aligned}\right\} \tag{6.9}$$

where, however, the coefficients A_s, B_s, C_s, D_s are not independent but are restricted to satisfy

$$\left.\begin{aligned}A_s\left(p_s - a + 4\mu \sin^2\frac{\pi s}{N}\right) &= bC_s,\\[2mm] B_s\left(p_s' - a + 4\mu \sin^2\frac{\pi s}{N}\right) &= bD_s.\end{aligned}\right\} \tag{6.10}$$

If it should happen that $p_s = p_s'$ the equations (6.9) have to be replaced by

$$\left.\begin{aligned}\xi_s &= (A_s + B_s t) e^{p_s t},\\ \eta_s &= (C_s + D_s t) e^{p_s t}.\end{aligned}\right\} \tag{6.9'}$$

and (6.10) remains true. Substituting back into (6.3) and replacing the variables x_r, y_r by X_r, Y_r (the actual concentrations) the solution can be written

$$\left.\begin{aligned}X_r &= h + \sum_{s=1}^{N} (A_s e^{p_s t} + B_s e^{p_s' t}) \exp\left[\frac{2\pi i r s}{N}\right],\\ Y_r &= k + \sum_{s=1}^{N} (C_s e^{p_s t} + D_s e^{p_s' t}) \exp\left[\frac{2\pi i r s}{N}\right].\end{aligned}\right\} \tag{6.11}$$

Here A_s, B_s, C_s, D_s are still related by (6.10), but otherwise are arbitrary complex numbers; p_s and p_s' are the roots of (6.8).

The expression (6·11) gives the general solution of the equations (6·1) when one assumes that departures from homogeneity are sufficiently small that the functions $f(X, Y)$ and $g(X, Y)$ can safely be taken as linear. The form (6·11) given is not very informative. It will be considerably simplified in § 8. Another implicit assumption concerns random disturbing influences. Strictly speaking one should consider such influences to be continuously at work. This would make the mathematical treatment considerably more difficult without substantially altering the conclusions. The assumption which is implicit in the analysis, here and in § 8, is that the state of the system at $t = 0$ is not one of homogeneity, since it has been displaced from such a state by the disturbances; but after $t = 0$ further disturbances are ignored. In § 9 the theory is reconsidered without this latter assumption.

7. Continuous ring of tissue

As an alternative to a ring of separate cells one might prefer to consider a continuous ring of tissue. In this case one can describe the position of a point of the ring by the angle θ which a radius to the point makes with a fixed reference radius. Let the diffusibilities of the two substances be μ' and ν'. These are not quite the same as μ and ν of the last section, since μ and ν are in effect referred to a cell diameter as unit of length, whereas μ' and ν' are referred to a conventional unit, the same unit in which the radius ρ of the ring is measured. Then

$$\mu = \mu'\left(\frac{N}{2\pi\rho}\right)^2, \quad \nu = \nu'\left(\frac{N}{2\pi\rho}\right)^2.$$

The equations are
$$\left.\begin{aligned}
\frac{\partial X}{\partial t} &= a(X-h)+b(Y-k)+\frac{\mu'}{\rho^2}\frac{\partial^2 X}{\partial\theta^2}, \\
\frac{\partial Y}{\partial t} &= c(X-h)+d(Y-k)+\frac{\nu'}{\rho^2}\frac{\partial^2 Y}{\partial\theta^2},
\end{aligned}\right\} \tag{7.1}$$

which will be seen to be the limiting case of (6·2). The marginal reaction rates a, b, c, d are, as before, the values at the equilibrium position of $\partial f/\partial X$, $\partial f/\partial Y$, $\partial g/\partial X$, $\partial g/\partial Y$. The general solution of the equations is

$$\left.\begin{aligned}
X &= h+ \sum_{s=-\infty}^{\infty} (A_s\, \mathrm{e}^{p_s t}+B_s\, \mathrm{e}^{p_s' t})\, \mathrm{e}^{\mathrm{i}s\theta}, \\
Y &= k+ \sum_{s=-\infty}^{\infty} (C_s\, \mathrm{e}^{p_s t}+D_s\, \mathrm{e}^{p_s' t})\, \mathrm{e}^{\mathrm{i}s\theta},
\end{aligned}\right\} \tag{7.2}$$

where p_s, p_s' are now roots of

$$\left(p-a+\frac{\mu' s^2}{\rho^2}\right)\left(p-d+\frac{\nu' s^2}{\rho^2}\right) = bc \tag{7.3}$$

and
$$\left.\begin{aligned}
A_s\left(p_s-a+\frac{\mu' s^2}{\rho^2}\right) &= bC_s, \\
B_s\left(p_s'-a+\frac{\mu' s^2}{\rho^2}\right) &= bD_s.
\end{aligned}\right\} \tag{7.4}$$

This solution may be justified by considering the limiting case of the solution (6·11). Alternatively, one may observe that the formula proposed is a solution, so that it only remains to prove that it is the most general one. This will follow if values of A_s, B_s, C_s, D_s can be found

to fit any given initial conditions. It is well known that any function of an angle (such as X) can be expanded as a 'Fourier series'

$$X(\theta) = \sum_{s=-\infty}^{\infty} G_s e^{is\theta} \quad (X(\theta) \text{ being values of } X \text{ at } t = 0),$$

provided, for instance, that its first derivative is continuous. If also

$$Y(\theta) = \sum_{s=-\infty}^{\infty} H_s e^{is\theta} \quad (Y(\theta) \text{ being values of } Y \text{ at } t = 0),$$

then the required initial conditions are satisfied provided $A_s + B_s = G_s$ and $C_s + D_s = H_s$. Values A_s, B_s, C_s, D_s to satisfy these conditions can be found unless $p_s = p'_s$. This is an exceptional case and its solution if required may be found as the limit of the normal case.

8. Types of asymptotic behaviour in the ring after a lapse of time

As the reader was reminded in § 2, after a lapse of time the behaviour of an expression of the form of (6·11) is eventually dominated by the terms for which the corresponding p_s has the largest real part. There may, however, be several terms for which this real part has the same value, and these terms will together dominate the situation, the other terms being ignored by comparison. There will, in fact, normally be either two or four such 'leading' terms. For if p_{s_0} is one of them then $p_{N-s_0} = p_{s_0}$, since

$$\sin^2 \frac{\pi(N-s_0)}{N} = \sin^2 \frac{\pi s_0}{N},$$

so that p_{s_0} and p_{N-s_0} are roots of the same equation (6·8). If also p_{s_0} is complex then $\mathscr{R}p_{s_0} = \mathscr{R}p'_{s_0}$, and so in all

$$\mathscr{R}p_{s_0} = \mathscr{R}p'_{s_0} = \mathscr{R}p_{N-s_0} = \mathscr{R}p'_{N-s_0}.$$

One need not, however, normally anticipate that any further terms will have to be included. If p_{s_0} and p_{s_1} are to have the same real part, then, unless $s_1 = s_0$ or $s_0 + s_1 = N$ the quantities a, b, c, d, μ, ν will be restricted to satisfy some special condition, which they would be unlikely to satisfy by chance. It is possible to find circumstances in which as many as ten terms have to be included if such special conditions *are* satisfied, but these have no particular physical or biological importance. It is assumed below that none of these chance relations hold.

It has already been seen that it is necessary to distinguish the cases where the value of p_{s_0} for one of the dominant terms is real from those where it is complex. These may be called respectively the *stationary* and the *oscillatory* cases.

Stationary case. After a sufficient lapse of time $X_r - h$ and $Y_r - k$ approach asymptotically to the forms

$$\left. \begin{aligned} X_r - h &= 2\mathscr{R}A_{s_0} \exp\left[\frac{2\pi i s_0 r}{N} + It\right], \\ Y_r - k &= 2\mathscr{R}C_{s_0} \exp\left[\frac{2\pi i s_0 r}{N} + It\right]. \end{aligned} \right\} \tag{8·1}$$

Oscillatory case. After a sufficient lapse of time $X_r - h$ and $Y_r - k$ approach the forms

$$\left. \begin{aligned} X_r - h &= 2 e^{It} \mathscr{R}\left\{A_{s_0} \exp\left[\frac{2\pi i s_0 r}{N} + i\omega t\right] + A_{N-s_0} \exp\left[-\frac{2\pi i s_0 r}{N} - i\omega t\right]\right\}, \\ Y_r - k &= 2 e^{It} \mathscr{R}\left\{C_{s_0} \exp\left[\frac{2\pi i s_0 r}{N} + i\omega t\right] + C_{N-s_0} \exp\left[-\frac{2\pi i s_0 r}{N} - i\omega t\right]\right\}. \end{aligned} \right\} \tag{8·2}$$

The real part of p_{s_0} has been represented by I, standing for 'instability', and in the oscillatory case its imaginary part is ω. By the use of the $\mathscr{R}$ operation (real part of), two terms have in each case been combined in one.

The meaning of these formulae may be conveniently described in terms of waves. In the stationary case there are stationary waves on the ring having s_0 lobes or crests. The coefficients A_{s_0} and C_{s_0} are in a definite ratio given by (6·10), so that the pattern for one morphogen determines that for the other. With the lapse of time the waves become more pronounced provided there is genuine instability, i.e. if I is positive. The wave-length of the waves may be obtained by dividing the number of lobes into the circumference of the ring. In the oscillatory case the interpretation is similar, but the waves are now not stationary but travelling. As well as having a wave-length they have a velocity and a frequency. The frequency is $\omega/2\pi$, and the velocity is obtained by multiplying the wave-length by the frequency. There are two wave trains moving round the ring in opposite directions.

The wave-lengths of the patterns on the ring do not depend only on the chemical data a, b, c, d, μ', ν' but on the circumference of the ring, since they must be submultiples of the latter. There is a sense, however, in which there is a 'chemical wave-length' which does not depend on the dimensions of the ring. This may be described as the limit to which the wave-lengths tend when the rings are made successively larger. Alternatively (at any rate in the case of continuous tissue), it may be described as the wave-length when the radius is chosen to give the largest possible instability I. One may picture the situation by supposing that the chemical wave-length is true wave-length which is achieved whenever possible, but that on a ring it is necessary to 'make do' with an approximation which divides exactly into the circumference.

Although all the possibilities are covered by the stationary and oscillatory alternatives there are special cases of them which deserve to be treated separately. One of these occurs when $s_0 = 0$, and may be described as the 'case of extreme long wave-length', though this term may perhaps preferably be reserved to describe the chemical data when they are such that s_0 is zero whatever the dimensions of the ring. There is also the case of 'extreme short wave-length'. This means that $\sin^2(\pi s_0/N)$ is as large as possible, which is achieved by s_0 being either $\frac{1}{2}N$, or $\frac{1}{2}(N-1)$. If the remaining possibilities are regarded as forming the 'case of finite wave-length', there are six subcases altogether. It will be shown that each of these really can occur, although two of them require three or more morphogens for their realization.

(a) *Stationary case with extreme long wave-length.* This occurs for instance if $\mu = \nu = \frac{1}{4}$, $b = c = 1$, $a = d$. Then $p_s = a - \sin^2\dfrac{\pi s}{N} + 1$. This is certainly real and is greatest when $s = 0$.

In this case the contents of all the cells are the same; there is no resultant flow from cell to cell due to diffusion, so that each is behaving as if it were isolated. Each is in unstable equilibrium, and slips out of it in synchronism with the others.

(b) *Oscillatory case with extreme long wave-length.* This occurs, for instance, if $\mu = \nu = \frac{1}{4}$, $b = -c = 1$, $a = d$. Then $p_s = a - \sin^2\dfrac{\pi s}{N} \pm i$. This is complex and its real part is greatest when $s = 0$. As in case (a) each cell behaves as if it were isolated. The difference from case (a) is that the departure from the equilibrium is oscillatory.

(c) *Stationary waves of extreme short wave-length.* This occurs, for instance, if $v = 0$, $\mu = 1$, $d = I$, $a = I - 1$, $b = -c = 1$. p_s is

$$I - \tfrac{1}{2} - 2\sin^2 \frac{\pi s}{N} + \sqrt{\left\{\left(2\sin^2 \frac{\pi s}{N} + \frac{1}{2}\right)^2 - 1\right\}},$$

and is greatest when $\sin^2(\pi s/N)$ is greatest. If N is even the contents of each cell are similar to those of the next but one, but distinctly different from those of its immediate neighbours. If, however, the number of cells is odd this arrangement is impossible, and the magnitude of the difference between neighbouring cells varies round the ring, from zero at one point to a maximum at a point diametrically opposite.
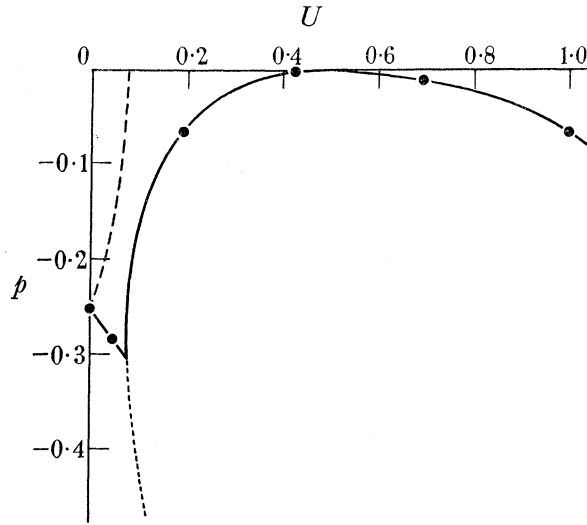

FIGURE 1. Values of $\mathscr{R}p$ (instability or growth rate), and $|\mathscr{I}p|$ (radian frequency of oscillation), related to wave-length $2\pi U^{-\frac{1}{2}}$ as in the relation (8·3) with $I = 0$. This is a case of stationary waves with finite wave-length. Full line, $\mathscr{R}p$; broken line, $-|\mathscr{I}p|$ (zero for $U > 0.071$); dotted line, $\mathscr{R}p'$. The full circles on the curve for $\mathscr{R}p$ indicate the values of U, p actually achievable on the finite ring considered in § 10, with $s = 0$ on the extreme left, $s = 5$ on the right.

(d) *Stationary waves of finite wave-length.* This is the case which is of greatest interest, and has most biological application. It occurs, for instance, if $a = I - 2$, $b = 2.5$, $c = -1.25$, $d = I + 1.5$, $\mu' = 1$, $\nu' = \tfrac{1}{2}$, and $\dfrac{\mu}{\mu'} = \dfrac{\nu}{\nu'} = \left(\dfrac{N}{2\pi\rho}\right)^2$. As before ρ is the radius of the ring, and N the number of cells in it. If one writes U for $\left(\dfrac{N}{\pi\rho}\right)^2 \sin^2 \dfrac{\pi s}{N}$, then equation (6·8) can, with these special values, be written

$$(p - I)^2 + (\tfrac{1}{2} + \tfrac{3}{2}U)(p - I) + \tfrac{1}{2}(U - \tfrac{1}{2})^2 = 0. \qquad (8·3)$$

This has a solution $p = I$ if $U = \tfrac{1}{2}$. On the other hand, it will be shown that if U has any other (positive) value then both roots for $p - I$ have negative real parts. Their product is positive being $\tfrac{1}{2}(U - \tfrac{1}{2})^2$, so that if they are real they both have the same sign. Their sum in this case is $-\tfrac{1}{2} - \tfrac{3}{2}U$ which is negative. Their common sign is therefore negative. If, however, the roots are complex their real parts are both equal to $-\tfrac{1}{4} - \tfrac{3}{4}U$, which is negative.

If the radius ρ of the ring be chosen so that for some integer s_0, $\frac{1}{2} = U = \left(\frac{N}{\pi\rho}\right)^2 \sin^2\frac{\pi s_0}{N}$, there will be stationary waves with s_0 lobes and a wave-length which is also equal to the chemical wave-length, for p_{s_0} will be equal to I, whereas every other p_s will have a real part smaller than I. If, however, the radius is chosen so that $\left(\frac{N}{\pi\rho}\right)^2 \sin^2\frac{\pi s}{N} = \frac{1}{2}$ cannot hold with an integral s, then (in this example) the actual number of lobes will be one of the two integers nearest to the (non-integral) solutions of this equation, and usually *the* nearest. Examples can, however, be constructed where this simple rule does not apply.

Figure 1 shows the relation (8·3) in graphical form. The curved portions of the graphs are hyperbolae.

The two remaining possibilities can only occur with three or more morphogens. With one morphogen the only possibility is (a).

(e) *Oscillatory case with a finite wave-length.* This means that there are genuine travelling waves. Since the example to be given involves three morphogens it is not possible to use the formulae of §6. Instead, one must use the corresponding three morphogen formulae. That which corresponds to (6·8) or (7·3) is most conveniently written as

$$\begin{vmatrix} a_{11}-p-\mu_1 U & a_{12} & a_{13} \\ a_{21} & a_{22}-p-\mu_2 U & a_{23} \\ a_{31} & a_{32} & a_{33}-p-\mu_3 U \end{vmatrix} = 0, \tag{8·4}$$

where again U has been written for $\left(\frac{N}{\pi\rho}\right)^2 \sin^2\frac{\pi s}{N}$. (This means essentially that $U = \left(\frac{2\pi}{\lambda}\right)^2$, where λ is the wave-length.) The four marginal reactivities are superseded by nine $a_{11}, \ldots, a_{33}$, and the three diffusibilities are μ_1, μ_2, μ_3. Special values leading to travelling waves are

$$\left.\begin{array}{lll} \mu_1 = \frac{2}{3}, & \mu_2 = \frac{1}{3}, & \mu_3 = 0 \\ a_{11} = -\frac{10}{3}, & a_{12} = 3, & a_{13} = -1, \\ a_{21} = -2, & a_{22} = \frac{7}{3}, & a_{23} = 0, \\ a_{31} = 3, & a_{32} = -4, & a_{33} = 0, \end{array}\right\} \tag{8·5}$$

and with them (8·4) reduces to

$$p^3 + p^2(U+1) + p(1 + \tfrac{2}{9}(U-1)^2) + U + 1 = 0. \tag{8·6}$$

If $U = 1$ the roots are $\pm i$ and -2. If U is near to I they are approximately $-1-U$ and $\pm i + \frac{(U-1)^2}{18}(\pm i - 1)$, and all have negative real parts. If the greatest real part is not the value zero, achieved with $U = 1$, then the value zero must be reached again at some intermediate value of U. Since P is then pure imaginary the even terms of (8·6) must vanish, i.e. $(p^2+1)(U+1) = 0$. But this can only happen if $p = \pm i$, and the vanishing of the odd terms then shows that $U = 1$. Hence zero is the largest real part for any root p of (8·6). The corresponding p is $\pm i$ and U is 1. This means that there are travelling waves with unit (chemical) radian frequency and unit (chemical) velocity. If I is added to a_{11}, a_{22} and a_{33}, the instability will become I in place of zero.

(*f*) Oscillatory case with extreme short wave-length. This means that there is metabolic oscillation with neighbouring cells nearly 180° out of phase. It can be achieved with three morphogens and the following chemical data:

$$
\left.
\begin{aligned}
\mu &= 1, & \mu_2 &= \mu_3 = 0, \\
a_{11} &= -1, & a_{12} &= -1, & a_{13} &= 0, \\
a_{21} &= 1, & a_{22} &= 0, & a_{23} &= -1, \\
a_{31} &= 0, & a_{32} &= 1, & a_{33} &= 0.
\end{aligned}
\right\} \tag{8.7}
$$

With these values (8·4) reduces to

$$
p^3 + p^2(U+1) + 2p + U + 1 = 0. \tag{8.8}
$$

This may be shown to have all the real parts of its roots negative if $U \geqslant 0$, for if $U = 0$ the roots are near to $-0\cdot6$, $-0\cdot2 \pm 1\cdot3i$, and if U be continuously increased the values of p will alter continuously. If they ever attain values with a positive real part they must pass through pure imaginary values (or zero). But if p is pure imaginary $p^3 + 2p$ and $(p^2 + 1)(U+1)$ must both vanish, which is impossible if $U \geqslant 0$. As U approaches infinity, however, one of the roots approaches i. Thus $\mathscr{R}p = 0$ can be approached as closely as desired by large values of U, but not attained.

9. FURTHER CONSIDERATION OF THE MATHEMATICS OF THE RING

In this section some of the finer points concerning the development of wave patterns are considered. These will be of interest mainly to those who wish to do further research on the subject, and can well be omitted on a first reading.

(1) *General formulae for the two morphogen case.* Taking the limiting case of a ring of large radius (or a filament), one may write $\left(\dfrac{N}{\pi\rho}\right)^2 \sin^2\dfrac{\pi s}{N} = U = \left(\dfrac{2\pi}{\lambda}\right)^2$ in (6·11) or $\dfrac{s^2}{\rho^2} = U = \left(\dfrac{2\pi}{\lambda}\right)^2$ in (7·3) and obtain

$$
(p - a + \mu'U)(p - d + \nu'U) = bc, \tag{9.1}
$$

which has the solution

$$
p = \frac{a+d}{2} - \frac{\mu'+\nu'}{2}U \pm \sqrt{\left\{\left(\frac{\mu'-\nu'}{2}U + \frac{d-a}{2}\right)^2 + bc\right\}}. \tag{9.2}
$$

One may put $I(U)$ for the real part of this, representing the instability for waves of wave-length $\lambda = 2\pi U^{-\frac{1}{2}}$. The dominant waves correspond to the maximum of $I(U)$. This maximum may either be at $U = 0$ or $U = \infty$ or at a stationary point on the part of the curve which is hyperbolic (rather than straight). When this last case occurs the values of p (or I) and U at the maximum are

$$
\left.
\begin{aligned}
p &= I = (d\mu' - a\nu' - 2\sqrt{(\mu'\nu')}\sqrt{(-bc)})(\mu'-\nu')^{-1}, \\
U &= \left(a - d + \frac{\mu'+\nu'}{\sqrt{(\mu'\nu')}}\sqrt{(-bc)}\right)(\mu'-\nu')^{-1}.
\end{aligned}
\right\} \tag{9.3}
$$

The conditions which lead to the four cases (*a*), (*b*), (*c*), (*d*) described in the last section are

(*a*) (Stationary waves of extreme long wave-length.) This occurs if either

(i) $bc > 0$, (ii) $bc < 0$ and $\dfrac{d-a}{\sqrt{(-bc)}} > \dfrac{\mu'+\nu'}{\sqrt{(\mu'\nu')}}$, (iii) $bc < 0$ and $\dfrac{d-a}{\sqrt{(-bc)}} < -2$.

The condition for instability in either case is that either $bc > ad$ or $a + d > 0$.

(b) (Oscillating case with extreme long wave-length, i.e. synchronized oscillations.) This occurs if

$$bc < 0 \quad \text{and} \quad -2 < \frac{d-a}{\sqrt{(-bc)}} < \frac{4\sqrt{(\mu'\nu')}}{\mu'+\nu'}.$$

There is instability if in addition $a+d>0$.

(c) (Stationary waves of extreme short wave-length.) This occurs if $bc < 0$, $\mu' > \nu' = 0$. There is instability if, in addition, $a+d>0$.

(d) (Stationary waves of finite wave-length.) This occurs if

$$bc < 0 \quad \text{and} \quad \frac{4\sqrt{(\mu'\nu')}}{\mu'+\nu'} < \frac{d-a}{\sqrt{(-bc)}} < \frac{\mu'+\nu'}{(\sqrt{\mu'\nu'})}, \tag{9.4a}$$

and there is instability if also

$$\frac{d}{\sqrt{(-bc)}}\sqrt{\frac{\mu'}{\nu'}} - \frac{a}{\sqrt{(-bc)}}\sqrt{\frac{\nu'}{\mu'}} > 2. \tag{9.4b}$$

It has been assumed that $\nu' \leqslant \mu' > 0$. The case where $\mu' \leqslant \nu' > 0$ can be obtained by interchanging the two morphogens. In the case $\mu' = \nu' = 0$ there is no co-operation between the cells whatever.

Some additional formulae will be given for the case of stationary waves of finite wavelength. The marginal reaction rates may be expressed parametrically in terms of the diffusibilities, the wave-length, the instability, and two other parameters α and χ. Of these α may be described as the ratio of $X-h$ to $Y-k$ in the waves. The expressions for the marginal reaction rates in terms of these parameters are

$$\left.\begin{aligned}
a &= \mu'(\nu'-\mu')^{-1}(2\nu'U_0+\chi)+I, \\
b &= \mu'(\nu'-\mu')^{-1}((\mu'+\nu')U_0+\chi)\alpha, \\
c &= \nu'(\mu'-\nu')^{-1}((\mu'+\nu')U_0+\chi)\alpha^{-1}, \\
d &= \nu'(\mu'-\nu')^{-1}(2\mu'U_0+\chi)+I,
\end{aligned}\right\} \tag{9.5}$$

and when these are substituted into (9.2) it becomes

$$p = I - \tfrac{1}{2}\chi - \frac{\mu'+\nu'}{2}U + \sqrt{\left\{\left(\frac{\mu'+\nu'}{2}U+\tfrac{1}{2}\chi\right)^2 - \mu'\nu'(U-U_0)^2\right\}}. \tag{9.6}$$

Here $2\pi U_0^{-\frac{1}{2}}$ is the chemical wave-length and $2\pi U^{-\frac{1}{2}}$ the wave-length of the Fourier component under consideration. χ must be positive for case (d) to apply.

If s be regarded as a continuous variable one can consider (9.2) or (9.6) as relating s to p, and dp/ds and d^2p/ds^2 have meaning. The value of d^2p/ds^2 at the maximum is of some interest, and will be used below in this section. Its value is

$$\frac{d^2p}{ds^2} = -\frac{\sqrt{(\mu'\nu')}}{\rho^2} \cdot \frac{8\sqrt{(\mu'\nu')}}{\mu'+\nu'}\cos^2\frac{\pi s}{N}(1+\chi U_0^{-1}(\mu'+\nu')^{-1})^{-1}. \tag{9.7}$$

(2) In §§ 6, 7, 8 it was supposed that the disturbances were not continuously operative, and that the marginal reaction rates did not change with the passage of time. These assumptions will now be dropped, though it will be necessary to make some other, less drastic,

approximations to replace them. The (statistical) amplitude of the 'noise' disturbances will be assumed constant in time. Instead of (6·6), (6·7), one then has

$$\left.\begin{aligned}\frac{d\xi}{dt} &= a'\xi + b\eta + R_1(t),\\[1mm]\frac{d\eta}{dt} &= c\xi + d'\eta + R_2(t),\end{aligned}\right\}\qquad(9\cdot8)$$

where ξ, η have been written for ξ_s, η_s since s may now be supposed fixed. For the same reason $a - 4\mu\sin^2\frac{\pi s}{N}$ has been replaced by a' and $d - 4\nu\sin^2\frac{\pi s}{N}$ by d'. The noise disturbances may be supposed to constitute white noise, i.e. if (t_1, t_2) and (t_3, t_4) are two non-overlapping intervals then $\int_{t_1}^{t_2} R_1(t)\,dt$ and $\int_{t_3}^{t_4} R_2(t)\,dt$ are statistically independent and each is normally distributed with variances $\beta_1(t_2 - t_1)$ and $\beta_1(t_4 - t_3)$ respectively, β_1 being a constant describing the amplitude of the noise. Likewise for $R_2(t)$, the constant β_1 being replaced by β_2. If p and p' are the roots of $(p - a')(p - d') = bc$ and p is the greater (both being real), one can make the substitution

$$\left.\begin{aligned}\xi &= b(u + v),\\[1mm]\eta &= (p - a')\,u + (p' - a')\,v,\end{aligned}\right\}\qquad(9\cdot9)$$

which transforms (9·8) into

$$\frac{du}{dt} = pu + \frac{p' - a'}{(p' - p)\,b}R_1(t) - \frac{R_2(t)}{p' - p} + \xi\frac{d}{dt}\left(\frac{p' - a'}{(p' - p)\,b}\right) - \eta\frac{d}{dt}\left(\frac{1}{p' - p}\right),\qquad(9\cdot11)$$

with a similar equation for v, of which the leading terms are $dv/dt = p'v$. This indicates that v will be small, or at least small in comparison with u after a lapse of time. If it is assumed that $v = 0$ holds (9·11) may be written

$$\frac{du}{dt} = qu + L_1(t)\,R_1(t) + L_2(t)\,R_2(t),\qquad(9\cdot12)$$

where $\qquad L_1(t) = \dfrac{p' - a'}{(p' - p)\,b},\quad L_2(t) = \dfrac{1}{p' - p},\quad q = p + bL_1'(t).\qquad(9\cdot13)$

The solution of this equation is

$$u = \int_{-\infty}^{t}(L_1(w)\,R_1(w) + L_2(w)\,R_2(w))\exp\left[\int_{w}^{t} q(z)\,dz\right]dw.\qquad(9\cdot14)$$

One is, however, not so much interested in such a solution in terms of the statistical disturbances as in the consequent statistical distribution of values of u, ξ and η at various times after instability has set in. In view of the properties of 'white noise' assumed above, the values of u at time t will be distributed according to the normal error law, with the variance

$$\int_{-\infty}^{t}[\beta_1(L_1(w))^2 + \beta_2(L_2(w))^2]\exp\left[2\int_{w}^{t} q(z)\,dz\right]dw.\qquad(9\cdot15)$$

There are two commonly occurring cases in which one can simplify this expression considerably without great loss of accuracy. If the system is in a distinctly stable state, then $q(t)$,

which is near to $p(t)$, will be distinctly negative, and $\exp\left[\int_w^t q(z)\,dz\right]$ will be small unless

w is near to t. But then $L_1(w)$ and $L_2(w)$ may be replaced by $L_1(t)$ and $L_2(t)$ in the integral, and also $q(z)$ may be replaced by $q(t)$. With these approximations the variance is

$$(-2q(t))^{-1}\left[\beta_1(L_1(t))^2+\beta_2(L_2(t))^2\right]. \tag{9.16}$$

A second case where there is a convenient approximation concerns times when the system is

unstable, so that $q(t)>0$. For the approximation concerned to apply $2\int_w^t q(z)\,dz$ must have

its maximum at the last moment $w(=t_0)$ when $q(t_0)=0$, and it must be the maximum by a considerable margin (e.g. at least 5) over all other local maxima. These conditions would apply for instance if $q(z)$ were always increasing and had negative values at a sufficiently early time. One also requires $q'(t_0)$ (the rate of increase of q at time t_0) to be reasonably large; it must at least be so large that over a period of time of length $(q'(t_0))^{-\frac{1}{2}}$ near to t_0 the changes in $L_1(t)$ and $L_2(t)$ are small, and $q'(t)$ itself must not appreciably alter in this period. Under these circumstances the integrand is negligible when w is considerably different from t_0, in comparison with its values at that time, and therefore one may replace $L_1(w)$ and $L_2(w)$ by $L_1(t_0)$ and $L_2(t_0)$, and $q'(w)$ by $q'(t_0)$. This gives the value

$$\sqrt{\pi}\,(q'(t_0))^{-\frac{1}{2}}\left[\beta_1(L_1(t_0))^2+\beta_2(L_2(t_0))^2\right]\exp\left[2\int_{t_0}^t q(z)\,dz\right], \tag{9.17}$$

for the variance of u.

The physical significance of this latter approximation is that the disturbances near the time when the instability is zero are the only ones which have any appreciable ultimate effect. Those which occur earlier are damped out by the subsequent period of stability. Those which occur later have a shorter period of instability within which to develop to greater amplitude. This principle is familiar in radio, and is fundamental to the theory of the superregenerative receiver.

Naturally one does not often wish to calculate the expression (9.17), but it is valuable as

justifying a common-sense point of view of the matter. The factor $\exp\left[\int_{t_0}^t q(z)\,dz\right]$ is

essentially the integrated instability and describes the extent to which one would expect disturbances of appropriate wave-length to grow between times t_0 and t. Taking the terms in β_1, β_2 into consideration separately, the factor $\sqrt{\pi}\beta_1(q'(t_0))^{-\frac{1}{2}}(L_1(t_0))^2$ indicates that the disturbances on the first morphogen should be regarded as lasting for a time

$$\sqrt{\pi}\,(q_1(t_0))^{-\frac{1}{2}}(bL_1(t_0))^2.$$

The dimensionless quantities $bL_1(t_0)$, $bL_2(t_0)$ will not usually be sufficiently large or small to justify their detailed calculation.

(3) The extent to which the component for which p_s is greatest may be expected to outdistance the others will now be considered in case (d). The greatest of the p_s will be called p_{s_0}. The two closest competitors to s_0 will be s_0-1 and s_0+1; it is required to determine how close the competition is. If the variation in the chemical data is sufficiently small it may be assumed that, although the exponents p_{s_0-1}, p_{s_0}, p_{s_0+1} may themselves vary appreciably in time, the differences $p_{s_0}-p_{s_0-1}$ and $p_{s_0}-p_{s_0+1}$ are constant. It certainly can happen that

one of these differences is zero or nearly zero, and there is then 'neck and neck' competition. The weakest competition occurs when $p_{s_0-1} = p_{s_0+1}$. In this case

$$p_{s_0} - p_{s_0-1} = p_{s_0} - p_{s_0+1} = -\tfrac{1}{2}(p_{s_0+1} - 2p_{s_0} + p_{s_0-1}).$$

But if s_0 is reasonably large $p_{s_0+1} - 2p_{s_0} + p_{s_0-1}$ can be set equal to $(d^2p/ds^2)_{s=s_0}$. It may be concluded that the rate at which the most quickly growing component grows cannot exceed the rate for its closest competitor by more than about $\tfrac{1}{2}(d^2p/ds^2)_{s=s_0}$. The formula (9·7), by which d^2p/ds^2 can be estimated, may be regarded as the product of two factors. The dimensionless factor never exceeds 4. The factor $\sqrt{(\mu'\nu')}/\rho^2$ may be described in very rough terms as 'the reciprocal of the time for the morphogens to diffuse a length equal to a radius'. In equally rough terms one may say that a time of this order of magnitude is required for the most quickly growing component to get a lead, amounting to a factor whose logarithm is of the order of unity, over its closest competitors, in the favourable case where $p_{s_0-1} = p_{s_0+1}$.

(4) Very little has yet been said about the effect of considering non-linear reaction rate functions when far from homogeneity. Any treatment so systematic as that given for the linear case seems to be out of the question. It is possible, however, to reach some qualitative conclusions about the effects of non-linear terms. Suppose that z_1 is the amplitude of the Fourier component which is most unstable (on a basis of the linear terms), and which may be supposed to have wave-length λ. The non-linear terms will cause components with wave-lengths $\tfrac{1}{2}\lambda, \tfrac{1}{3}\lambda, \tfrac{1}{4}\lambda, \ldots$ to appear as well as a space-independent component. If only quadratic terms are taken into account and if these are somewhat small, then the component of wave-length $\tfrac{1}{2}\lambda$ and the space-independent component will be the strongest. Suppose these have amplitudes z_2 and z_1. The state of the system is thus being described by the numbers z_0, z_1, z_2. In the absence of non-linear terms they would satisfy equations

$$\frac{dz_0}{dt} = p_0 z_0, \quad \frac{dz_1}{dt} = p_1 z_1, \quad \frac{dz_2}{dt} = p_2 z_2,$$

and if there is slight instability p_1 would be a small positive number, but p_0 and p_2 distinctly negative. The effect of the non-linear terms is to replace these equations by ones of the form

$$\frac{dz_0}{dt} = p_0 z_0 + A z_1^2 + B z_2^2,$$

$$\frac{dz_1}{dt} = p_1 z_1 + C z_2 z_1 + D z_0 z_1,$$

$$\frac{dz_2}{dt} = p_2 z_2 + E z_1^2 + F z_0 z_2.$$

As a first approximation one may put $dz_0/dt = dz_2/dt = 0$ and ignore z_1^4 and higher powers; z_0 and z_1 are then found to be proportional to z_1^2, and the equation for z_1 can be written $dz_1/dt = p_0 z_1 - k z_1^3$. The sign of k in this differential equation is of great importance. If it is positive, then the effect of the term $k z_1^3$ is to arrest the exponential growth of z_1 at the value $\sqrt{(p_1/k)}$. The 'instability' is then very confined in its effect, for the waves can only reach a finite amplitude, and this amplitude tends to zero as the instability (p_1) tends to zero. If, however, k is negative the growth becomes something even faster than exponential, and, if the equation $dz_1/dt = p_1 z_1 - k z_1^3$ held universally, it would result in the amplitude becoming

infinite in a finite time. This phenomenon may be called 'catastrophic instability'. In the case of two-dimensional systems catastrophic instability is almost universal, and the corresponding equation takes the form $dz_1/dt = p_1 z_1 + k z_1^2$. Naturally enough in the case of catastrophic instability the amplitude does not really reach infinity, but when it is sufficiently large some effect previously ignored becomes large enough to halt the growth.

(5) Case (a) as described in § 8 represents a most extremely featureless form of pattern development. This may be remedied quite simply by making less drastic simplifying assumptions, so that a less gross account of the pattern can be given by the theory. It was assumed in § 9 that only the most unstable Fourier components would contribute appreciably to the pattern, though it was seen above (heading (3) of this section) that (in case (d)) this will only apply if the period of time involved is adequate to permit the morphogens, supposed for this purpose to be chemically inactive, to diffuse over the whole ring or organ concerned. The same may be shown to apply for case (a). If this assumption is dropped a much more interesting form of pattern can be accounted for. To do this it is necessary to consider not merely the components with $U = 0$ but some others with small positive values of U. One may assume the form $At - BU$ for p. Linearity in U is assumed because only small values of U are concerned, and the term At is included to represent the steady increase in instability. By measuring time from the moment of zero instability the necessity for a constant term is avoided. The formula (9·17) may be applied to estimate the statistical distribution of the amplitudes of the components. Only the factor $\exp\left[2\int_{t_0}^{t} q(z)\,dz\right]$ will depend very much on U, and taking $q(t) = p(t) = At - BU$, t_0 must be BU/A and the factor is

$$\exp\left[A(t - BU/A)^2\right].$$

The term in U^2 can be ignored if At^2 is fairly large, for then either $B^2 U^2/A^2$ is small or the factor e^{-BUt} is. But At^2 certainly is large if the factor e^{At^2}, applying when $U = 0$, is large. With this approximation the variance takes the form $C e^{-\frac{1}{2}k^2 U}$, with only the two parameters C, k to distinguish the pattern populations. By choosing appropriate units of concentration and length these pattern populations may all be reduced to a standard one, e.g. with $C = k = 1$. Random members of this population may be produced by considering any one of the type (a) systems to which the approximations used above apply. They are also produced, but with only a very small amplitude scale, if a homogeneous one-morphogen system undergoes random disturbances without diffusion for a period, and then diffusion without disturbance. This process is very convenient for computation, and can also be applied to two dimensions. Figure 2 shows such a pattern, obtained in a few hours by a manual computation.

To be more definite a set of numbers $u_{r,s}$ was chosen, each being ± 1, and taking the two values with equal probability. A function $f(x, y)$ is related to these numbers by the formula

$$f(x, y) = \Sigma u_{r,s} \exp\left[-\tfrac{1}{2}((x - hr)^2 + (y - hs)^2)\right].$$

In the actual computation a somewhat crude approximation to the function

$$\exp\left[-\tfrac{1}{2}(x^2 + y^2)\right]$$

was used and h was about 0·7. In the figure the set of points where $f(x,y)$ is positive is shown black. The outlines of the black patches are somewhat less irregular than they should be due to an inadequacy in the computation procedure.
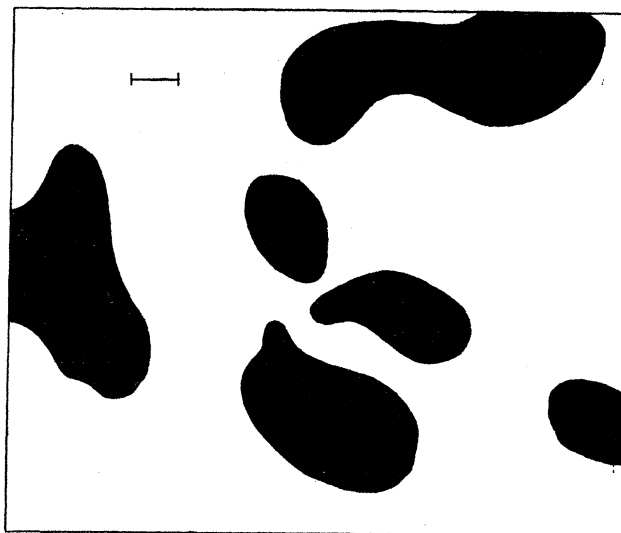

FIGURE 2. An example of a 'dappled' pattern as resulting from a type (a) morphogen system. A marker of unit length is shown. See text, §9, 11.

10. A NUMERICAL EXAMPLE

The numerous approximations and assumptions that have been made in the foregoing analysis may be rather confusing to many readers. In the present section it is proposed to consider in detail a single example of the case of most interest, (d). This will be made as specific as possible. It is unfortunately not possible to specify actual chemical reactions with the required properties, but it is thought that the reaction rates associated with the imagined reactions are not unreasonable.

The detail to be specified includes

(i) The number and dimensions of the cells of the ring.

(ii) The diffusibilities of the morphogens.

(iii) The reactions concerned.

(iv) The rates at which the reactions occur.

(v) Information about random disturbances.

(vi) Information about the distribution, in space and time, of those morphogens which are of the nature of evocators.

These will be taken in order.

(i) It will be assumed that there are twenty cells in the ring, and that they have a diameter of 0·1 mm each. These cells are certainly on the large rather than the small side, but by no means impossibly so. The number of cells in the ring has been chosen rather small in order that it should not be necessary to make the approximation of continuous tissue.

(ii) Two morphogens are considered. They will be called X and Y, and the same letters will be used for their concentrations. This will not lead to any real confusion. The diffusion constant for X will be assumed to be $5 \times 10^{-8}\,\mathrm{cm^2\,s^{-1}}$ and that for Y to be $2\cdot5 \times 10^{-8}\,\mathrm{cm^2\,s^{-1}}$. With cells of diameter 0·01 cm this means that X flows between neighbouring cells at the
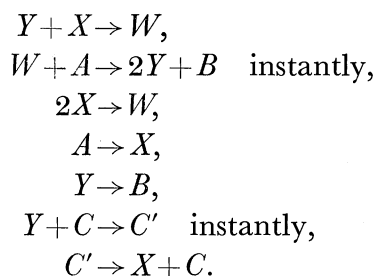
rate 5×10^{-4} of the difference of X-content of the two cells per second. In other words, if there is nothing altering the concentrations but diffusion the difference of concentrations suffers an exponential decay with time constant 1000 s, or 'half-period' of 700 s. These times are doubled for Y.
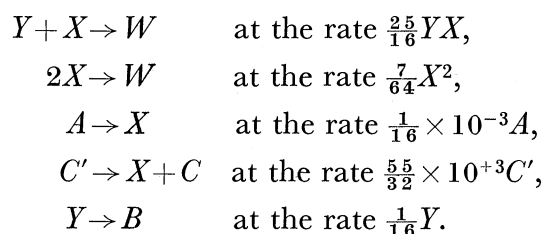
If the cell membrane is regarded as the only obstacle to diffusion the permeability of the membranes to the morphogen is 5×10^{-6} cm/s or 0.018 cm/h. Values as large as 0.1 cm/h have been observed (Davson & Danielli 1943, figure 28).

(iii) The reactions are the most important part of the assumptions. Four substances A, X, Y, B are involved; these are isomeric, i.e. the molecules of the four substances are all rearrangements of the same atoms. Substances C, C', W will also be concerned. The thermodynamics of the problem will not be discussed except to say that it is contemplated that of the substances A, X, Y, B the one with the greatest free energy is A, and that with the least is B. Energy for the whole process is obtained by the degradation of A into B. The substance C is in effect a catalyst for the reaction $Y \rightarrow X$, and may also be regarded as an evocator, the system being unstable if there is a sufficient concentration of C.

The reactions postulated are

$$Y + X \rightarrow W,$$
$$W + A \rightarrow 2Y + B \quad \text{instantly,}$$
$$2X \rightarrow W,$$
$$A \rightarrow X,$$
$$Y \rightarrow B,$$
$$Y + C \rightarrow C' \quad \text{instantly,}$$
$$C' \rightarrow X + C.$$

(iv) For the purpose of stating the reaction rates special units will be introduced (for the purpose of this section only). They will be based on a period of 1000 s as units of time, and 10^{-11} mole/cm^3 as concentration unit*. There will be little occasion to use any but these special units (s.u.). The concentration of A will be assumed to have the large value of 1000 s.u. and the catalyst C, together with its combined form C' the concentration $10^{-3}(1 + \gamma)$ s.u., the dimensionless quantity γ being often supposed somewhat small, though values over as large a range as from -0.5 to 0.5 may be considered. The rates assumed will be

$$\begin{aligned}
Y + X \rightarrow W & \quad \text{at the rate } \tfrac{25}{16} YX, \\
2X \rightarrow W & \quad \text{at the rate } \tfrac{7}{64} X^2, \\
A \rightarrow X & \quad \text{at the rate } \tfrac{1}{16} \times 10^{-3} A, \\
C' \rightarrow X + C & \quad \text{at the rate } \tfrac{55}{32} \times 10^{+3} C', \\
Y \rightarrow B & \quad \text{at the rate } \tfrac{1}{16} Y.
\end{aligned}$$

With the values assumed for A and C' the net effect of these reactions is to convert X into Y at the rate $\tfrac{1}{32}[50XY + 7X^2 - 55(1 + \gamma)]$ at the same time producing X at the constant rate $\tfrac{1}{16}$, and destroying Y at the rate $Y/16$. If, however, the concentration of Y is zero and the rate of increase of Y required by these formulae is negative, the rate of conversion of Y into X is reduced sufficiently to permit Y to remain zero.

* A somewhat larger value of concentration unit (e.g. 10^{-9} mole/cm^3) is probably more suitable. The choice of unit only affects the calculations through the amplitude of the random disturbances.

In the special units $\mu = \frac{1}{2}$, $\nu = \frac{1}{4}$.

(v) Statistical theory describes in detail what irregularities arise from the molecular nature of matter. In a period in which, on the average, one should expect a reaction to occur between n pairs (or other combinations) of molecules, the actual number will differ from the mean by an amount whose mean square is also n, and is distributed according to the normal error law. Applying this to a reaction proceeding at a rate F (s.u.) and taking the volume of the cell as 10^{-8} cm^3 (assuming some elongation tangentially to the ring) it will be found that the root mean square irregularity of the quantity reacting in a period τ of time (s.u.) is $0\cdot004\sqrt{(F\tau)}$.

TABLE 1. SOME STATIONARY-WAVE PATTERNS

cell number	first specimen				second specimen: incipient	'slow cooking': incipient	four-lobed equilibrium	
	incipient pattern		final pattern					
	X	Y	X	Y	Y	Y	X	Y
0	1·130	0·929	0·741	1·463	0·834	1·057	1·747	0·000
1	1·123	0·940	0·761	1·469	0·833	0·903	1·685	0·000
2	1·154	0·885	0·954	1·255	0·766	0·813	1·445	2·500
3	1·215	0·810	1·711	0·000	0·836	0·882	0·445	2·500
4	1·249	0·753	1·707	0·000	0·930	1·088	1·685	0·000
5	1·158	0·873	0·875	1·385	0·898	1·222	1·747	0·000
6	1·074	1·003	0·700	1·622	0·770	1·173	1·685	0·000
7	1·078	1·000	0·699	1·615	0·740	0·956	0·445	2·500
8	1·148	0·896	0·885	1·382	0·846	0·775	0·445	2·500
9	1·231	0·775	1·704	0·000	0·937	0·775	1·685	0·000
10	1·204	0·820	1·708	0·000	0·986	0·969	1·747	0·000
11	1·149	0·907	0·944	1·273	1·019	1·170	1·685	0·000
12	1·156	0·886	0·766	1·451	0·899	1·203	0·445	2·500
13	1·170	0·854	0·744	1·442	0·431	1·048	0·445	2·500
14	1·131	0·904	0·756	1·478	0·485	0·868	1·685	0·000
15	1·090	0·976	0·935	1·308	0·919	0·813	1·747	0·000
16	1·109	0·957	1·711	0·000	1·035	0·910	1·685	0·000
17	1·201	0·820	1·706	0·000	1·003	1·050	0·445	2·500
18	1·306	0·675	0·927	1·309	0·899	1·175	0·445	2·500
19	1·217	0·811	0·746	1·487	0·820	1·181	1·685	0·000

The diffusion of a morphogen from a cell to a neighbour may be treated as if the passage of a molecule from one cell to another were a monomolecular reaction; a molecule must be imagined to change its form slightly as it passes the cell wall. If the diffusion constant for a wall is μ, and quantities M_1, M_2 of the relevant morphogen lie on the two sides of it, the root-mean-square irregularity in the amount passing the wall in a period τ is

$$0\cdot004\sqrt{\{(M_1+M_2)\,\mu\tau\}}.$$

These two sources of irregularity are the most significant of those which arise from truly statistical cause, and are the only ones which are taken into account in the calculations whose results are given below. There may also be disturbances due to the presence of neighbouring anatomical structures, and other similar causes. These are of great importance, but of too great variety and complexity to be suitable for consideration here.

(vi) The only morphogen which is being treated as an evocator is C. Changes in the concentration of A might have similar effects, but the change would have to be rather great. It is preferable to assume that A is a 'fuel substance' (e.g. glucose) whose concentration does

not change. The concentration of C, together with its combined form C', will be supposed the same in all cells, but it changes with the passage of time. Two different varieties of the problem will be considered, with slightly different assumptions.

The results are shown in table 1. There are eight columns, each of which gives the concentration of a morphogen in each of the twenty cells; the circumstances to which these concentrations refer differ from column to column. The first five columns all refer to the same 'variety' of the imaginary organism, but there are two specimens shown. The specimens differ merely in the chance factors which were involved. With this variety the value of γ was allowed to increase at the rate of 2^{-7} s.u. from the value $-\frac{1}{4}$ to $+\frac{1}{16}$. At this point a pattern had definitely begun to appear, and was recorded. The parameter γ was then allowed to decrease at the same rate to zero and then remained there until there was no
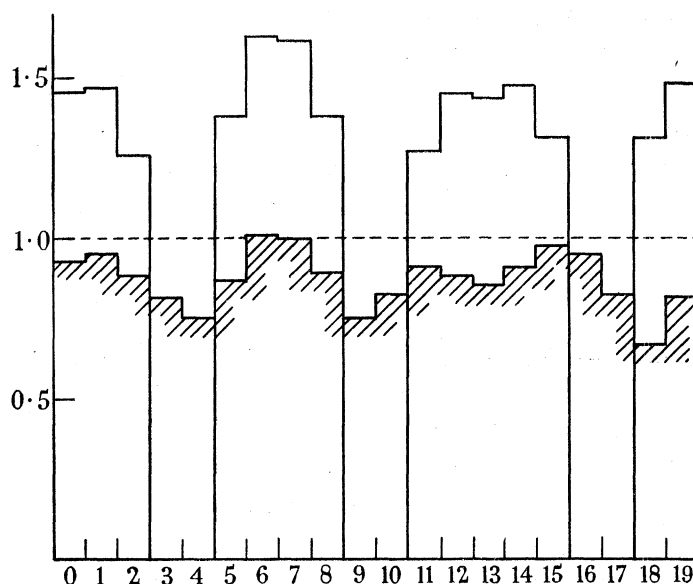

FIGURE 3. Concentrations of Y in the development of the first specimen (taken from table 1).
------- original homogeneous equilibrium; ///// incipient pattern; ——— final equilibrium.

more appreciable change. The pattern was then recorded again. The concentrations of Y in these two recordings are shown in figure 3 as well as in table 1. For the second specimen only one column of figures is given, viz. those for the Y morphogen in the incipient pattern. At this stage the X values are closely related to the Y values, as may be seen from the first specimen (or from theory). The final values can be made almost indistinguishable from those for the first specimen by renumbering the cells and have therefore not been given. These two specimens may be said to belong to the 'variety with quick cooking', because the instability is allowed to increase so quickly that the pattern appears relatively soon. The effect of this haste might be regarded as rather unsatisfactory, as the incipient pattern is very irregular. In both specimens the four-lobed component is present in considerable strength in the incipient pattern. It 'beats' with the three-lobed component producing considerable irregularity. The relative magnitudes of the three- and four-lobed components depend on chance and vary from specimen to specimen. The four-lobed component may often be the stronger, and may occasionally be so strong that the final pattern is four-lobed. How often this happens is not known, but the pattern, when it occurs, is shown in the last

two columns of the table. In this case the disturbances were supposed removed for some time before recording, so as to give a perfectly regular pattern.

The remaining column refers to a second variety, one with 'slow cooking'. In this the value of γ was allowed to increase only at the rate 10^{-5}. Its initial value was -0.010, but is of no significance. The final value was 0.003. With this pattern, when shown graphically, the irregularities are definitely perceptible, but are altogether overshadowed by the three-lobed component. The possibility of the ultimate pattern being four-lobed is not to be taken seriously with this variety.

The set of reactions chosen is such that the instability becomes 'catastrophic' when the second-order terms are taken into account, i.e. the growth of the waves tends to make the whole system more unstable than ever. This effect is finally halted when (in some cells) the concentration of Y has become zero. The constant conversion of Y into X through the agency of the catalyst C can then no longer continue in these cells, and the continued growth of the amplitude of the waves is arrested. When $\gamma = 0$ there is of course an equilibrium with $X = Y = 1$ in all cells, which is very slightly stable. There are, however, also other stable equilibria with $\gamma = 0$, two of which are shown in the table. These final equilibria may, with some trouble but little difficulty, be verified to be solutions of the equations (6·1) with

$$\frac{\mathrm{d}X}{\mathrm{d}t} = \frac{\mathrm{d}Y}{\mathrm{d}t} = 0,$$

and
$$32f(X, Y) = 57 - 50XY - 7Y^2, \quad 32g(X, Y) = 50XY + 7Y^2 - 2Y - 55.$$

The morphogen concentrations recorded at the earlier times connect more directly with the theory given in §§ 6 to 9. The amplitude of the waves was then still sufficiently small for the approximation of linearity to be still appropriate, and consequently the 'catastrophic' growth had not yet set in.

The functions $f(X, Y)$ and $g(X, Y)$ of § 6 depend also on γ and are

$$f(X, Y) = \tfrac{1}{32}[-7X^2 - 50XY + 57 + 55\gamma],$$

$$g(X, Y) = \tfrac{1}{32}[7X^2 + 50XY - 2Y - 55 - 55\gamma].$$

In applying the theory it will be as well to consider principally the behaviour of the system when γ remains permanently zero. Then for equilibrium $f(X, Y) = g(X, Y) = 0$ which means that $X = Y = 1$, i.e. $h = k = 1$. One also finds the following values for various quantities mentioned in §§ 6 to 9:

$$a = -2, \quad b = -1.5625, \quad c = 2, \quad d = 1.500, s = 3.333,$$

$$I = 0, \quad \alpha = 0.625, \quad \chi = 0.500, \quad (d-a)(-bc)^{-\frac{1}{2}} = 1.980,$$

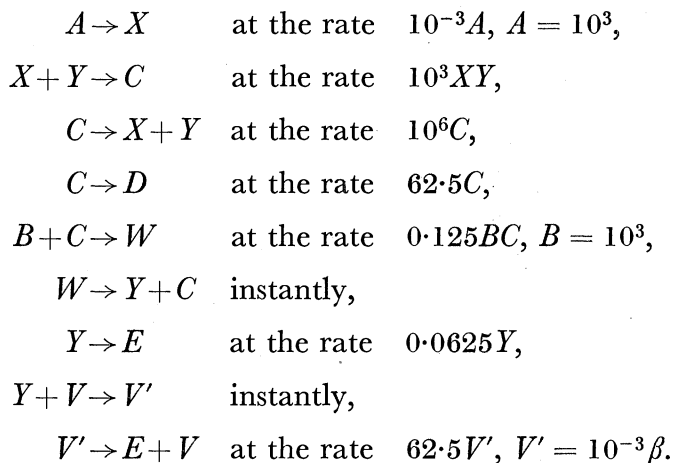$$(\mu+\nu)(\mu\nu)^{-\frac{1}{2}} = 2.121, \quad p_0 = -0.25 \pm 0.25i,$$

$$p_2 = -0.0648, \quad p_3 = -0.0034, \quad p_4 = -0.0118.$$

(The relation between p and U for these chemical data, and the values p_n, can be seen in figure 1, the values being so related as to make the curves apply to this example as well as that in § 8.) The value $s = 3.333$ leads one to expect a three-lobed pattern as the commonest, and this is confirmed by the values p_n. The four-lobed pattern is evidently the closest competitor. The closeness of the competition may be judged from the difference $p_3 - p_4 = 0.0084$,

which suggests that the three-lobed component takes about 120 s.u. or about 33 h to gain an advantage of a neper (i.e. about $2\cdot7:1$) over the four-lobed one. However, the fact that γ is different from 0 and is changing invalidates this calculation to some extent.

The figures in table 1 were mainly obtained with the aid of the Manchester University Computer.

Although the above example is quite adequate to illustrate the mathematical principles involved it may be thought that the chemical reaction system is somewhat artificial. The following example is perhaps less so. The same 'special units' are used. The reactions assumed are

$$A \to X \qquad \text{at the rate} \quad 10^{-3}A, \; A = 10^3,$$

$$X + Y \to C \qquad \text{at the rate} \quad 10^3 XY,$$

$$C \to X + Y \quad \text{at the rate} \quad 10^6 C,$$

$$C \to D \qquad \text{at the rate} \quad 62\cdot5 C,$$

$$B + C \to W \quad \text{at the rate} \quad 0\cdot125 BC, \; B = 10^3,$$

$$W \to Y + C \quad \text{instantly,}$$

$$Y \to E \qquad \text{at the rate} \quad 0\cdot0625 Y,$$

$$Y + V \to V' \qquad \text{instantly,}$$

$$V' \to E + V \quad \text{at the rate} \quad 62\cdot5 V', \; V' = 10^{-3}\beta.$$

The effect of the reactions $X + Y \rightleftarrows C$ is that $C = 10^{-3}XY$. The reaction $C \to D$ destroys C, and therefore in effect both X and Y, at the rate $\frac{1}{16}XY$. The reaction $A \to X$ forms X at the constant rate 1, and the pair $Y + V \to V' \to E + V$ destroys Y at the constant rate $\frac{1}{16}\beta$. The pair $B + C \to W \to Y + C$ forms Y at the rate $\frac{1}{8}XY$, and $Y \to E$ destroys it at the rate $\frac{1}{16}Y$. The total effect therefore is that X is produced at the rate $f(X, Y) = \frac{1}{16}(16 - XY)$, and Y at the rate $g(X, Y) = \frac{1}{16}(XY - Y - \beta)$. However, $g(X, Y) = 0$ if $Y \leqslant 0$. The diffusion constants will be supposed to be $\mu = \frac{1}{4}, \nu = \frac{1}{16}$. The homogeneity condition gives $hk = 16, k = 16 - \beta$.

It will be seen from conditions $(9\cdot4a)$ that case (d) applies if and only if $\frac{4}{k} + \frac{k}{4} < 2\cdot75$, i.e. if k lies between $1\cdot725$ and $9\cdot257$. Condition $(9\cdot4b)$ shows that there will be instability if in addition $\frac{8}{k} + \frac{k}{8} > \sqrt{3} + \frac{1}{2}$, i.e. if k does not lie between $4\cdot98$ and $12\cdot8$. It will also be found that the wave-length corresponding to $k = 4\cdot98$ is $4\cdot86$ cell diameters.

In the case of a ring of six cells with $\beta = 12$ there is a stable equilibrium, as shown in table 2.

TABLE 2

cell	0	1	2	3	4	5
X	7·5	3·5	2·5	2·5	3·5	7·5
Y	0	8	8	8	8	0

It should be recognized that these equilibria are only dynamic equilibria. The molecules which together make up the chemical waves are continually changing, though their concentrations in any particular cell are only undergoing small statistical fluctuations. Moreover,

in order to maintain the wave pattern a continual supply of free energy is required. It is clear that this must be so since there is a continual degradation of energy through diffusion. This energy is supplied through the 'fuel substances' (A, B in the last example), which are degraded into 'waste products' (D, E).

11. RESTATEMENT AND BIOLOGICAL INTERPRETATION OF THE RESULTS

Certain readers may have preferred to omit the detailed mathematical treatment of §§ 6 to 10. For their benefit the assumptions and results will be briefly summarized, with some change of emphasis.

The system considered was either a ring of cells each in contact with its neighbours, or a continuous ring of tissue. The effects are extremely similar in the two cases. For the purposes of this summary it is not necessary to distinguish between them. A system with two or three morphogens only was considered, but the results apply quite generally. The system was supposed to be initially in a stable homogeneous condition, but disturbed slightly from this state by some influences unspecified, such as Brownian movement or the effects of neighbouring structures or slight irregularities of form. It was supposed also that slow changes are taking place in the reaction rates (or, possibly, the diffusibilities) of the two or three morphogens under consideration. These might, for instance, be due to changes of concentration of other morphogens acting in the role of catalyst or of fuel supply, or to a concurrent growth of the cells, or a change of temperature. Such changes are supposed ultimately to bring the system out of the stable state. The phenomena when the system is just unstable were the particular subject of the inquiry. In order to make the problem mathematically tractable it was necessary to assume that the system never deviated very far from the original homogeneous condition. This assumption was called the 'linearity assumption' because it permitted the replacement of the general reaction rate functions by linear ones. This linearity assumption is a serious one. Its justification lies in the fact that the patterns produced in the early stages when it is valid may be expected to have strong qualitative similarity to those prevailing in the later stages when it is not. Other, less important, assumptions were also made at the beginning of the mathematical theory, but the detailed effects of these were mostly considered in § 9, and were qualitatively unimportant.

The conclusions reached were as follows. After the lapse of a certain period of time from the beginning of instability, a pattern of morphogen concentrations appears which can best be described in terms of 'waves'. There are six types of possibility which may arise.

(*a*) The equilibrium concentrations and reaction rates may become such that there would be instability for an isolated cell with the same content as any one of the cells of the ring. If that cell drifts away from the equilibrium position, like an upright stick falling over, then, in the ring, each cell may be expected to do likewise. In neighbouring cells the drift may be expected to be in the same direction, but in distant cells, e.g. at opposite ends of a diameter there is no reason to expect this to be so.

This is the least interesting of the cases. It is possible, however, that it might account for 'dappled' colour patterns, and an example of a pattern in two dimensions produced by this type of process is shown in figure 2 for comparison with 'dappling'. If dappled patterns are to be explained in this way they must be laid down in a latent form when the foetus is only

a few inches long. Later the distances would be greater than the morphogens could travel by diffusion.

(b) This case is similar to (a), except that the departure from equilibrium is not a unidirectional drift, but is oscillatory. As in case (a) there may not be agreement between the contents of cells at great distances.

There are probably many biological examples of this metabolic oscillation, but no really satisfactory one is known to the author.

(c) There may be a drift from equilibrium, which is in opposite directions in contiguous cells.

No biological examples of this are known.

(d) There is a stationary wave pattern on the ring, with no time variation, apart from a slow increase in amplitude, i.e. the pattern is slowly becoming more marked. In the case of a ring of continuous tissue the pattern is sinusoidal, i.e. the concentration of one of the morphogens plotted against position on the ring is a sine curve. The peaks of the waves will be uniformly spaced round the ring. The number of such peaks can be obtained approximately by dividing the so-called 'chemical wave-length' of the system into the circumference of the ring. The chemical wave-length is given for the case of two morphogens by the formula (9·3). This formula for the number of peaks of course does not give a whole number, but the actual number of peaks will always be one of the two whole numbers nearest to it, and will usually be *the* nearest. The degree of instability is also shown in (9·3).

The mathematical conditions under which this case applies are given in equations (9·4a), (9·4b).

Biological examples of this case are discussed at some length below.

(e) For a two-morphogen system only the alternatives (a) to (d) are possible, but with three or more morphogen it is possible to have travelling waves. With a ring there would be two sets of waves, one travelling clockwise and the other anticlockwise. There is a natural chemical wave-length and wave frequency in this case as well as a wave-length; no attempt was made to develop formulae for these.

In looking for biological examples of this there is no need to consider only rings. The waves could arise in a tissue of any anatomical form. It is important to know what wavelengths, velocities and frequencies would be consistent with the theory. These quantities are determined by the rates at which the reactions occur (more accurately by the 'marginal reaction rates', which have the dimensions of the reciprocal of a time), and the diffusibilities of the morphogens. The possible range of values of the reaction rates is so immensely wide that they do not even give an indication of orders of magnitude. The diffusibilities are more helpful. If one were to assume that all the *dimensionless* parameters in a system of travelling waves were the same as in the example given in §8, one could say that the product of the velocity and wave-length of the waves was 3π times the diffusibility of the most diffusible morphogen. But this assumption is certainly false, and it is by no means obvious what is the true range of possible values for the numerical constant (here 3π). The movements of the tail of a spermatozoon suggest themselves as an example of these travelling waves. That the waves are within one cell is no real difficulty. However, the speed of propagation seems to be somewhat greater than can be accounted for except with a rather large numerical constant.

(*f*) Metabolic oscillation with neighbouring cells in opposite phases. No biological examples of this are known to the author.

It is difficult also to find cases to which case (*d*) applies directly, but this is simply because isolated rings of tissue are very rare. On the other hand, systems that have the same kind of symmetry as a ring are extremely common, and it is to be expected that under appropriate chemical conditions, stationary waves may develop on these bodies, and that their circular symmetry will be replaced by a polygonal symmetry. Thus, for instance, a plant shoot may at one time have circular symmetry, i.e. appear essentially the same when rotated through any angle about a certain axis; this shoot may later develop a whorl of leaves, and then it will only suffer rotation through the angle which separates the leaves, or any multiple of it. This same example demonstrates the complexity of the situation when more than one dimension is involved. The leaves on the shoots may not appear in whorls, but be imbricated. This possibility is also capable of mathematical analysis, and will be considered in detail in a later paper. The cases which appear to the writer to come closest biologically to the 'isolated ring of cells' are the tentacles of (e.g.) *Hydra*, and the whorls of leaves of certain plants such as Woodruff (*Asperula odorata*).

Hydra is something like a sea-anemone but lives in fresh water and has from about five to ten tentacles. A part of a *Hydra* cut off from the rest will rearrange itself so as to form a complete new organism. At one stage of this proceeding the organism has reached the form of a tube open at the head end and closed at the other end. The external diameter is somewhat greater at the head end than over the rest of the tube. The whole still has circular symmetry. At a somewhat later stage the symmetry has gone to the extent that an appropriate stain will bring out a number of patches on the widened head end. These patches arise at the points where the tentacles are subsequently to appear (Child 1941, p. 101 and figure 30). According to morphogen theory it is natural to suppose that reactions, similar to those which were considered in connection with the ring of tissue, take place in the widened head end, leading to a similar breakdown of symmetry. The situation is more complicated than the case of the thin isolated ring, for the portion of the *Hydra* concerned is neither isolated nor very thin. It is not unreasonable to suppose that this head region is the only one in which the chemical conditions are such as to give instability. But substances produced in this region are still free to diffuse through the surrounding region of lesser activity. There is no great difficulty in extending the mathematics to cover this point in particular cases. But if the active region is too wide the system no longer approximates the behaviour of a thin ring and one can no longer expect the tentacles to form a single whorl. This also cannot be considered in detail in the present paper.

In the case of woodruff the leaves appear in whorls on the stem, the number of leaves in a whorl varying considerably, sometimes being as few as five or as many as nine. The numbers in consecutive whorls on the same stem are often equal, but by no means invariably. It is to be presumed that the whorls originate in rings of active tissue in the meristematic area, and that the rings arise at sufficiently great distance to have little influence on one another. The number of leaves in the whorl will presumably be obtainable by the rule given above, viz. by dividing the chemical wave-length into the circumference, though both these quantities will have to be given some new interpretation more appropriate to woodruff than to the ring. Another important example of a structure with polygonal

symmetry is provided by young root fibres just breaking out from the parent root. Initially these are almost homogeneous in cross-section, but eventually a ring of fairly evenly spaced spots appear, and these later develop into vascular strands. In this case again the full explanation must be in terms of a two-dimensional or even a three-dimensional problem, although the analysis for the ring is still illuminating. When the cross-section is very large the strands may be in more than one ring, or more or less randomly or hexagonally arranged. The two-dimensional theory (not expounded here) also goes a long way to explain this.

Flowers might appear superficially to provide the most obvious examples of polygonal symmetry, and it is probable that there are many species for which this 'waves round a ring' theory is essentially correct. But it is certain that it does not apply for all species. If it did it would follow that, taking flowers as a whole, i.e. mixing up all species, there would be no very markedly preferred petal (or corolla, segment, stamen, etc.) numbers. For when all species are taken into account one must expect that the diameters of the rings concerned will take on nearly all values within a considerable range, and that neighbouring diameters will be almost equally common. There may also be some variation in chemical wave-length. Neighbouring values of the ratio circumferences to wave-length should therefore be more or less equally frequent, and this must mean that neighbouring petal numbers will have much the same frequency. But this is not borne out by the facts. The number five is extremely common, and the number seven rather rare. Such facts are, in the author's opinion, capable of explanation on the basis of morphogen theory, and are closely connected with the theory of phyllotaxis. They cannot be considered in detail here.

The case of a filament of tissue calls for some comment. The equilibrium patterns on such a filament will be the same as on a ring, which has been cut at a point where the concentrations of the morphogens are a maximum or a minimum. This could account for the segmentation of such filaments. It should be noticed, however, that the theory will not apply unmodified for filaments immersed in water.

12. Chemical waves on spheres. Gastrulation

The treatment of homogeneity breakdown on the surface of a sphere is not much more difficult than in the case of the ring. The theory of spherical harmonics, on which it is based, is not, however, known to many that are not mathematical specialists. Although the essential properties of spherical harmonics that are used are stated below, many readers will prefer to proceed directly to the last paragraph of this section.

The anatomical structure concerned in this problem is a hollow sphere of continuous tissue such as a blastula. It is supposed sufficiently thin that one can treat it as a 'spherical shell'. This latter assumption is merely for the purpose of mathematical simplification; the results are almost exactly similar if it is omitted. As in § 7 there are to be two morphogens, and $a, b, c, d, \mu', \nu', h, k$ are also to have the same meaning as they did there. The operator ∇^2 will be used here to mean the superficial part of the Laplacian, i.e. $\nabla^2 V$ will be an abbreviation of

$$\frac{1}{\rho^2}\frac{\partial^2 V}{\partial \phi^2} + \frac{1}{\rho^2 \sin^2 \theta}\frac{\partial}{\partial \theta}\left(\sin \theta \frac{\partial V}{\partial \theta}\right),$$

where θ and ϕ are spherical polar co-ordinates on the surface of the sphere and ρ is its radius. The equations corresponding to (7·1) may then be written

$$\left.\begin{aligned}
\frac{\partial X}{\partial t} &= a(X-h)+b(Y-k)+\mu'\nabla^2 X,\\
\frac{\partial Y}{\partial t} &= c(X-h)+d(Y-k)+\nu'\nabla^2 Y.
\end{aligned}\right\} \tag{12·1}$$

It is well known (e.g. Jeans 1927, chapter 8) that any function on the surface of the sphere, or at least any that is likely to arise in a physical problem, can be 'expanded in spherical surface harmonics'. This means that it can be expressed in the form

$$\sum_{n=0}^{\infty}\left[\sum_{m=-n}^{n}A_n^m P_n^m(\cos\theta)\,\mathrm{e}^{im\phi}\right].$$

The expression in the square bracket is described as a 'surface harmonic of degree n'. Its nearest analogue in the ring theory is a Fourier component. The essential property of a spherical harmonic of degree n is that when the operator ∇^2 is applied to it the effect is the same as multiplication by $-n(n+1)/\rho^2$. In view of this fact it is evident that a solution of (12·1) is

$$\left.\begin{aligned}
X &= h+\sum_{n=0}^{\infty}\sum_{m=-n}^{n}(A_n^m\,\mathrm{e}^{iq_n t}+B_n^m\,\mathrm{e}^{iq_n' t})\,P_n^m(\cos\theta)\,\mathrm{e}^{im\phi},\\
Y &= k+\sum_{n=0}^{\infty}\sum_{m=-n}^{n}(C_n^m\,\mathrm{e}^{iq_n t}+D_n^m\,\mathrm{e}^{iq_n' t})\,P_n^m(\cos\theta)\,\mathrm{e}^{i\phi},
\end{aligned}\right\} \tag{12·2}$$

where q_n and q_n' are the two roots of

$$\left(q-a+\frac{\mu'}{\rho^2}n(n+1)\right)\left(q-d+\frac{\nu'}{\rho^2}n(n+1)\right)=bc \tag{12·3}$$

and

$$A_n^m\left(q_n-a+\frac{\mu'}{\rho^2}n(n+1)\right)=bC_n^m,$$

$$B_n^m\left(q_n'-a+\frac{\mu'}{\rho^2}n(n+1)\right)=cD_n^m. \tag{12·4}$$

This is the most general solution, since the coefficients A_n^m and B_n^m can be chosen to give any required values of X, Y when $t=0$, except when (12·3) has two equal roots, in which case a treatment is required which is similar to that applied in similar circumstances in §7. The analogy with §7 throughout will indeed be obvious, though the summation with respect to m does not appear there. The meaning of this summation is that there are a number of different patterns with the same wave-length, which can be superposed with various amplitude factors. Then supposing that, as in §8, one particular wave-length predominates, (12·2) reduces to

$$\left.\begin{aligned}
X-h &= \mathrm{e}^{iq_{n_0} t}\sum_{m=-n_0}^{n_0}A_{n_0}^m P_{n_0}^m(\cos\theta)\,\mathrm{e}^{im\phi},\\
b(Y-k) &= \left(q_{n_0}-a+\frac{\mu'}{\rho^2}n(n+1)\right)(X-h).
\end{aligned}\right\} \tag{12·5}$$

In other words, the concentrations of the two morphogens are proportional, and both of them are surface harmonics of the same degree n_0, viz. that which makes the greater of the roots q_{n_0}, q_{n_0}' have the greatest value.

It is probable that the forms of various nearly spherical structures, such as radiolarian skeletons, are closely related to these spherical harmonic patterns. The most important application of the theory seems, however, to be to the gastrulation of a blastula. Suppose that the chemical data, including the chemical wave-length, remain constant as the radius of the blastula increases. To be quite specific suppose that

$$\mu' = 2, \quad \nu' = 1, \quad a = -4, \quad b = -8, \quad c = 4, \quad d = 7.$$

With these values the system is quite stable so long as the radius is less than about 2. Near this point, however, the harmonics of degree 1 begin to develop and a pattern of form (12·5) with $n_0 = 1$ makes its appearance. Making use of the facts that

$$P_1^0(\cos\theta) = \cos\theta, \quad P_1^1(\cos\theta) = P_1^{-1}(\cos\theta) = \sin\theta,$$

it is seen that $X-h$ is of the form

$$X-h = A\cos\theta + B\sin\theta\cos\phi + C\sin\theta\sin\phi, \tag{12·6}$$

which may also be interpreted as
$$X-h = A'\cos\theta', \tag{12·7}$$

where θ' is the angle which the radius θ, ϕ makes with the fixed direction having direction cosines proportional to B, C, A and $A' = \sqrt{(A^2+B^2+C^2)}$.

The outcome of the analysis therefore is quite simply this. Under certain not very restrictive conditions (which include a requirement that the sphere be relatively small but increasing in size) the pattern of the breakdown of homogeneity is axially symmetrical, not about the original axis of spherical polar co-ordinates, but about some new axis determined by the disturbing influences. The concentrations of the first morphogen are given by (12·7), where θ' is measured from this new axis; and $Y-k$ is proportional to $X-h$. Supposing that the first morphogen is, or encourages the production of, a growth hormone, one must expect the blastula to grow in an axially symmetric manner, but at a greater rate at one end of the axis than at the other This might under many circumstances lead to gastrulation, though the effects of such growth are not very easily determinable. They depend on the elastic properties of the tissue as well as on the growth rate at each point. This growth will certainly lead to a solid of revolution with a marked difference between the two poles, unless, in addition to the chemical instability, there is a mechanical instability causing the breakdown of axial symmetry. The direction of the axis of gastrulation will be quite random according to this theory. It may be that it is found experimentally that the axis is normally in some definite direction such as that of the animal pole. This is not essentially contradictory to the theory, for any small asymmetry of the zygote may be sufficient to provide the 'disturbance' which determines the axis.

13. Non-linear theory. Use of digital computers

The 'wave' theory which has been developed here depends essentially on the assumption that the reaction rates are linear functions of the concentrations, an assumption which is justifiable in the case of a system just beginning to leave a homogeneous condition. Such systems certainly have a special interest as giving the first appearance of a pattern, but they are the exception rather than the rule. Most of an organism, most of the time, is developing

from one pattern into another, rather than from homogeneity into a pattern. One would like to be able to follow this more general process mathematically also. The difficulties are, however, such that one cannot hope to have any very embracing *theory* of such processes, beyond the statement of the equations. It might be possible, however, to treat a few particular cases in detail with the aid of a digital computer. This method has the advantage that it is not so necessary to make simplifying assumptions as it is when doing a more theoretical type of analysis. It might even be possible to take the mechanical aspects of the problem into account as well as the chemical, when applying this type of method. The essential disadvantage of the method is that one only gets results for particular cases. But this disadvantage is probably of comparatively little importance. Even with the ring problem, considered in this paper, for which a reasonably complete mathematical analysis was possible, the computational treatment of a particular case was most illuminating. The morphogen theory of phyllotaxis, to be described, as already mentioned, in a later paper, will be covered by this computational method. Non-linear equations will be used.

It must be admitted that the biological examples which it has been possible to give in the present paper are very limited. This can be ascribed quite simply to the fact that biological phenomena are usually very complicated. Taking this in combination with the relatively elementary mathematics used in this paper one could hardly expect to find that many observed biological phenomena would be covered. It is thought, however, that the imaginary biological systems which have been treated, and the principles which have been discussed, should be of some help in interpreting real biological forms.

REFERENCES

Child, C. M. 1941 *Patterns and problems of development.* University of Chicago Press.

Davson, H. & Danielli, J. F. 1943 *The permeability of natural membranes.* Cambridge University Press.

Jeans, J. H. 1927 *The mathematical theory of elasticity and magnetism,* 5th ed. Cambridge University Press.

Michaelis, L. & Menten, M. L. 1913 Die Kinetik der Invertinwirkung. *Biochem. Z.* **49**, 333.

Thompson, Sir D'Arcy 1942 *On growth and form,* 2nd ed. Cambridge University Press.

Waddington, C. H. 1940 *Organisers and genes.* Cambridge University Press.