SYMMETRY, BIFURCATIONS AND PATTERN FORMATION (d' apres Satting-
er, Michel, Thom and many others)

Michiel Hazewinkel

Econometric Inst., Erasmus University, Rotterdam and
Centrum voor Wiskunde en Informatica (= Centre for
Mathematics & Computer Science), Amsterdam

1. INTRODUCTION

Nature often seems to like (approximately) symmetric solu-
tions to problems. Mathematically, or more generally, scienti-
fically, it thus becomes our task to understand why, e.g. by
showing that more or less regular patterns are usually the most
stable, or economical, or optimising with respect to a suitable
criterion.

Particular aspects of this theme concern the question of
whether symmetric problems necessarily have symmetric solutions
and how the symmetry of a problem/solution can change as a
parameter varies. This last problem in turn is part of bifur-
cation theory which examines the question of how the set of
solutions of a problem can change in nature as certain parame-
ters vary. Here it turns out to be remarkably fruitful [9, 29,
33] to take the symmetry of a situation into account. Also the
presence of a symmetry group has remarkably strong consequences
for extrema of functions as we shall see.

There are two aspects of symmetry which I consider very
interesting and which perhaps so far have not had all the at-
tention they deserve. One concerns approximate symmetry. Here I
have in mind for instance a symmetric problem which would have
a symmetric solution if the boundary conditions were equally
symmetric. Now suppose the boundary conditions are disturbed,
when will there be an approximatedly symmetric solution, e.g.
when will there be a boundary layer in which the symmetry will
be restored and a fully symmetric solution in the middle or
when will the solution of the new problem be like a crystal
with defects. There seems to be no general theorem concerning
such matters. Other approximate symmetry problems (e.g. how to

recognise them) arise when dealing with a basically symmetric
pattern with a (small) random [3] or systematic disturbance
superimposed. Some preliminary foundational remarks on approxi-
mate symmetry are contained in [17, 28]. Much more remains to
be done and different kinds of approximate symmetry certainly
exist. Still harder to understand approximatedly symmetric en-
tities are the Penrose universes described in section 2 below.
The second aspect concerns the matter that as certain parame-
ters change an object may both lose and gain symmetry, some-
times simultaneously. The systematics of symmetry loss, that is
spontaneous breaking of symmetry, have had considerable atten-
tion over the years, see e.g. [17, 20, 21, 22, 24]. The syste-
matics of gaining symmetry far less. The matter is discussed
below in section 3 in terms of the automorphisms of three di-
mensional algebras.

   Apart from this example there is little new in this paper
and it should basically be seen as a low key introduction to
important work of others, notably the persons mentioned in the
title. I have added a few more references than are strictly
necessary for the purposes of this paper itself, e.g. a few
references to books which treat of bifurcation theory and give
applications [32, 25, 18, 26, 11, 1] and a few (in my view)
closely related matters [2, 10, 41, 23].

2. PATTERN FORMATION AND SYMMETRY AND APPROXIMATE PATTERNS AND SYMMETRIES

   One of the best known examples of pattern formation and
one of the most studied (the two terms are not synonymous) is
the Benard convection. Consider a fluid layer heated from below
as in Figure 1. At a small temperature difference (small tem-
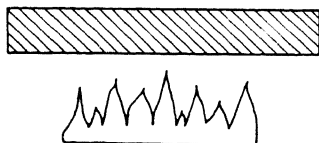perature gradient) heat is transported by conduction and the
solution

Figure 1. Fluid layer heated from below. After [12]

is completely symmetric, that is if we consider only the pos-
sible symmetrics viewable from the top we have E_2 symmetry
where E_2 is the group of rigid motions of the plane (the fluid
layer is assumed to be infinitely extended). At higher tempera-
ture gradients the lighter fluid at the bottom will tend to
rise, cool at the top and return to the bottom. A microscopic
pattern arises which can take the form of rolls, c.f. Figure 2

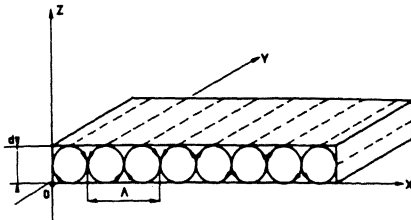or the famous Bénard cells (hexagons), c.f. Figure 3.
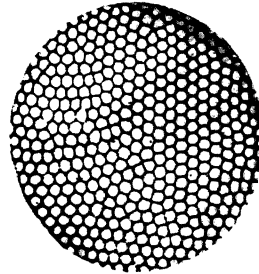

Figure 2. Rolls pattern.
After [12]

Figure 3. Hexagonal
pattern. After [5]

Other less regular patterns develop at still higher temperature
gradients. We will say a little more about Sattinger's bifurca-
tion-in-the-presence-of-symmetry analysis of the Bénard problem
in section 5 below. A highly recommended up to date discussion
of the topic is [9], which includes also a discussion of the
spherical case, a model for convection in the molten layer of
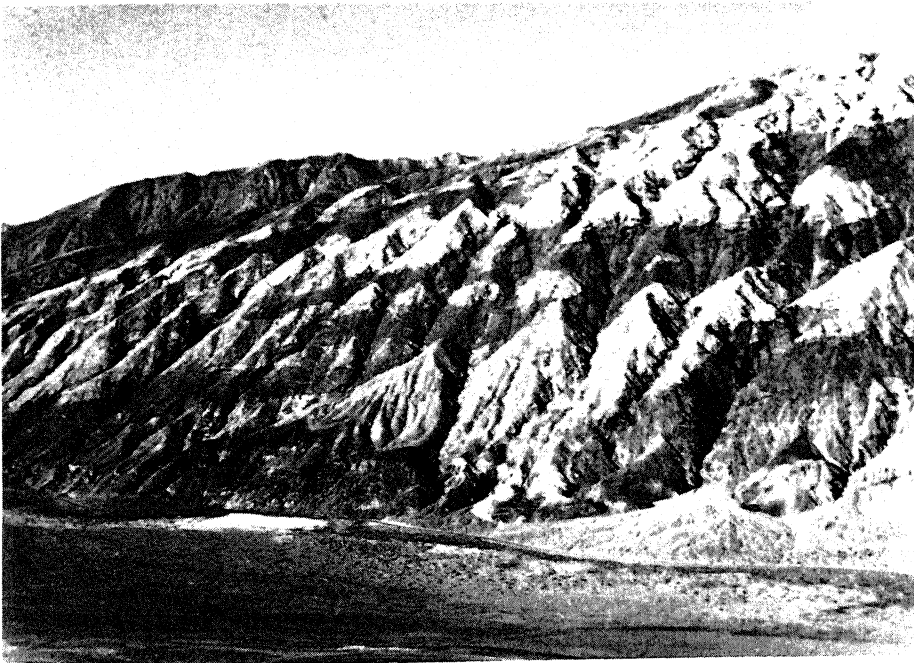the earth between core and mantle.

Figure 4. East wall of Death Valley, California.
After [35].

    Another example of a strikingly regular pattern, caused by
erosion in this case, is depicted in Figure 4 and still more
examples are the drainage basin patterns of the Figures 5, 6,
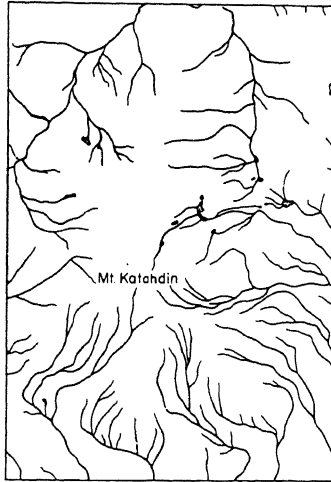and 7 below. These drainage patterns seem perhaps less regular
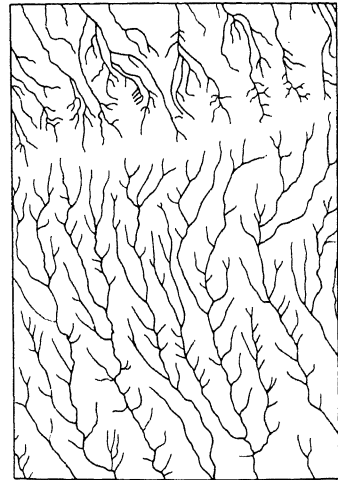

Figure 5. Radial drai-
nage pattern. After [39]
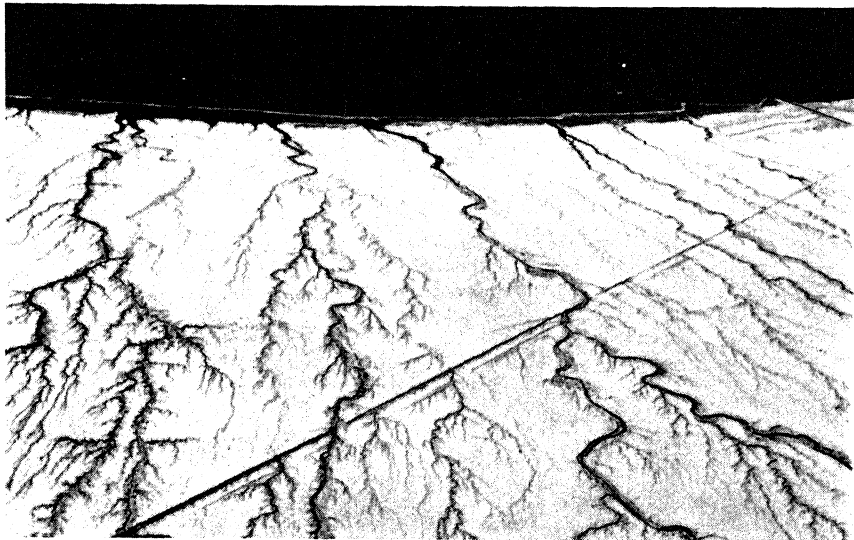

Figure 6. Parallel
drainage pattern. After
[39]

Figure 7. Dendritic channels. After [35].

than the erosion pattern of Figure 4 and the convection pat-
terns of the Figures 2 and 3. Still they are very far from ran-
dom and as such demand an explanation of their (amount of)
regularity. In all these examples the problem is not so much to
understand that something happens, i.e. that some pattern de-
velops, but to explain the striking regularity of the patterns
and in case there are several patterns to understand the selec-
tion mechanisms and the relative stability of these patterns
with respect to each other.

      Still other examples of spontaneously arising regular pat-
terns are the socalled cloud streets, c.f. Figure 8 below, and
the socalled Liesegang rings, which form e.g. when a drop of
silver nitrate is placed on a film of gelatine saturated with
potassum dichromate as in Figure 9.

      However, both nature and man seem to like approximately
symmetric solutions even better. Or solutions whose obvious
regularity is much harder to describe in mathematical terms
then e.g. hexagonal or street patterns. Spiral patterns for
instance, occur very frequently, c.f. Figures 10 and 11 below

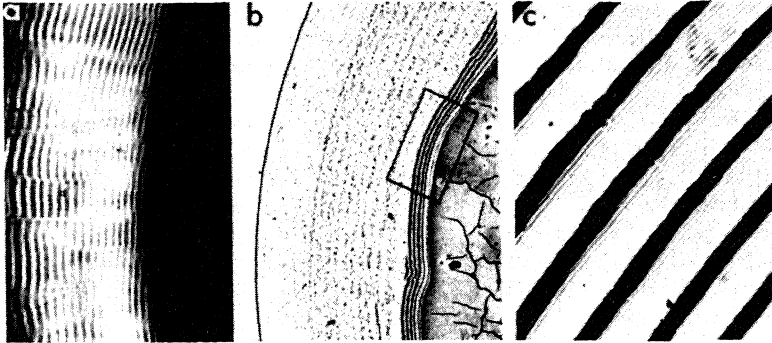Figure 8. Stratocumulus cloud streets. After [37]
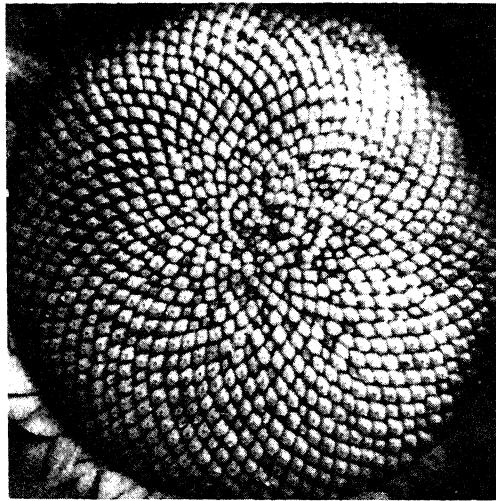
Figure 9. Liesegang rings. After [4]

Figure 10. Sunflower head. Courtesy of Empire magazine

Interesting remarks on the mathematics of spirals can be found in for instance [6]. Still harder to describe kinds of symmetry are those exhibited by various fractal patterns such as the twin dragon pattern of Figures 12 and 13 below which not only have certain more or less obvious symmetries (c.f. Figure 13) and the less obvious rotational symmetry of Figure 12, also has the property that it can be covered with reduced size rep- licas of itself ad infinitum.
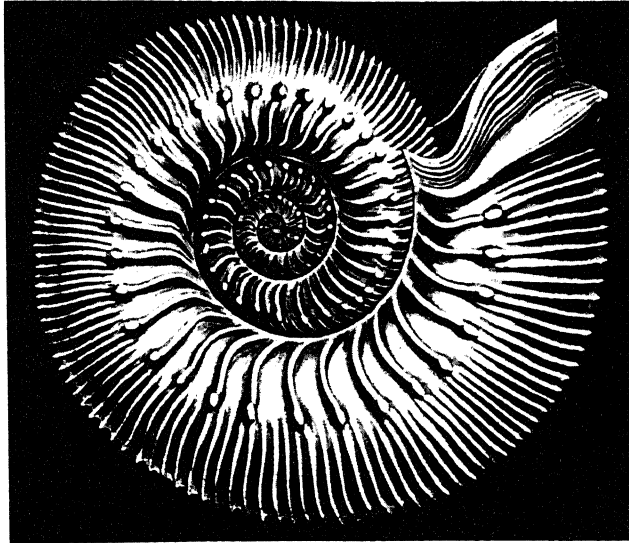
Figure 11. After [36]
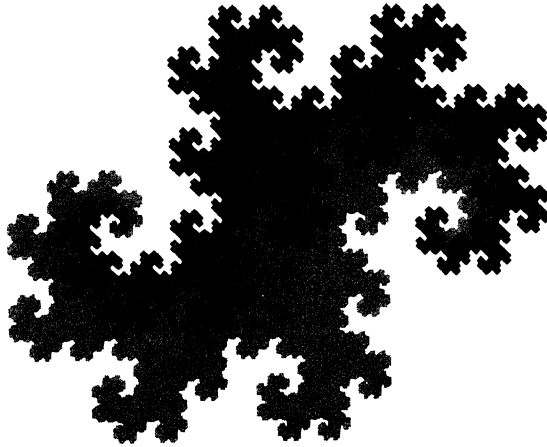
Figure 12. Twin dragon fractal, After B.B. Mandelbrot,
The fractal geometry of nature, Freeman, 1982

    Consider also the spiral tiling of H. Voderberg depicted
in Figure 14. Though it obviously has many regularity proper-
ties and is intuitively very symmetric it is quite difficult to
find symmetries (apart from a 180 degree rotation which also
interchanges colours).

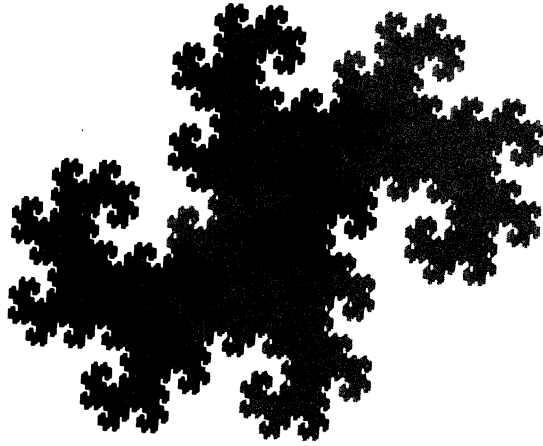Figure 13. Tiling of twin dragon fractal, After B.B. Mandelbrot, The fractal geometry of nature, Freeman, 1982

Figure 14. A non-periodic tiling by H. Voderberg. From [8]

There are patterns whose "regularity" is even more distur-
bing and very hard, possibly impossible to describe in the ma-
thematical framework we usually use in this connection. These
are the Penrose non-periodic tilings, also called the Penrose
universes as described in [8] from which the following is
taken.

The basic tiles are obtained from a diamond with angles of
72° and 108° as drawn in Figure 15. The number t is the golden
ratio $\frac{1}{2}(1 + \sqrt{5})$. The two tiles, called dart and kite, so ob-
tained are marked with a drawn and dashed circle as indicated
and there is an additional tiling rule in that abutting edges
must join circle segments of the same kind so that fitting a dart
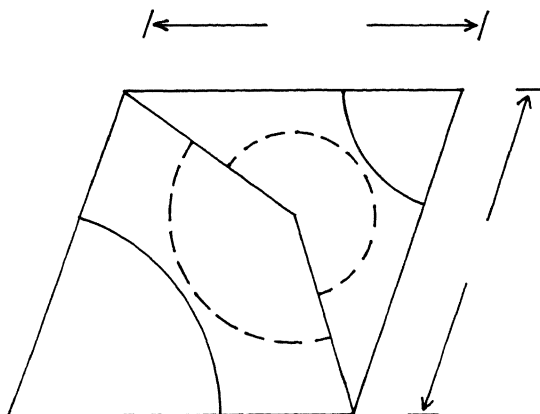and a kite to form a diamond is forbidden. Using these two kinds

Figure 15. The Penrose dart and kite.

of tiles and respecting the additional tiling rules it is pos-
sible to tile the entire plane. Indeed there are innumerable
different ways of doing that. Some of the more striking pat-
terns are depicted below in the Figures 16, 17 and 18. The cen-
tral 10-sided regular polygon consisting of 15 darts and 25
kites in Figure 18 is called a cartwheel. Note also that the
cartwheel pattern of Figure 18 has little symmetry in the ob-
vious sense (only a reflection through a central vertical
line).

Here are some properties of the Penrose tilings (or uni-
verses):
a) all tilings are non-periodic;
b) every point in every tiling is inside a cartwheel;
c) every finite region of diameter ⩽ d of any tiling occurs
   within distance 2d of any point in any other tiling.

In addition there are "fractal properties" in that from a
given tiling another one with larger kites and darts can be
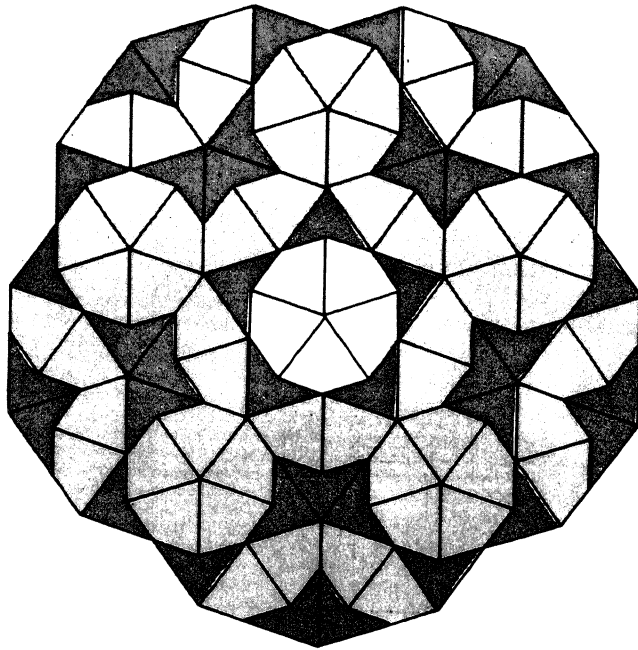constructed in a simple systematic way.

Figure 16
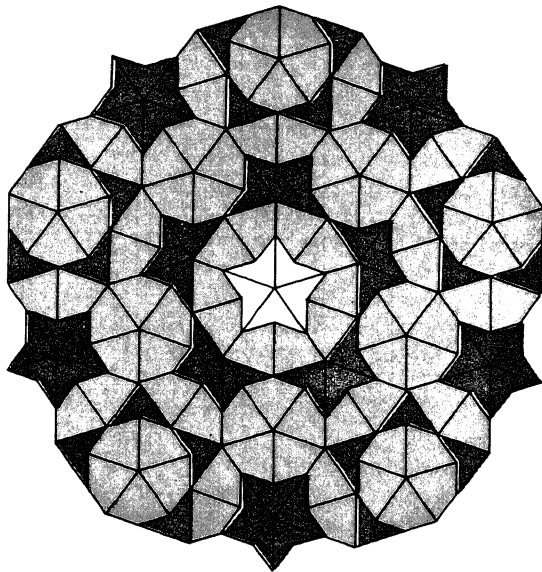
Figure 17
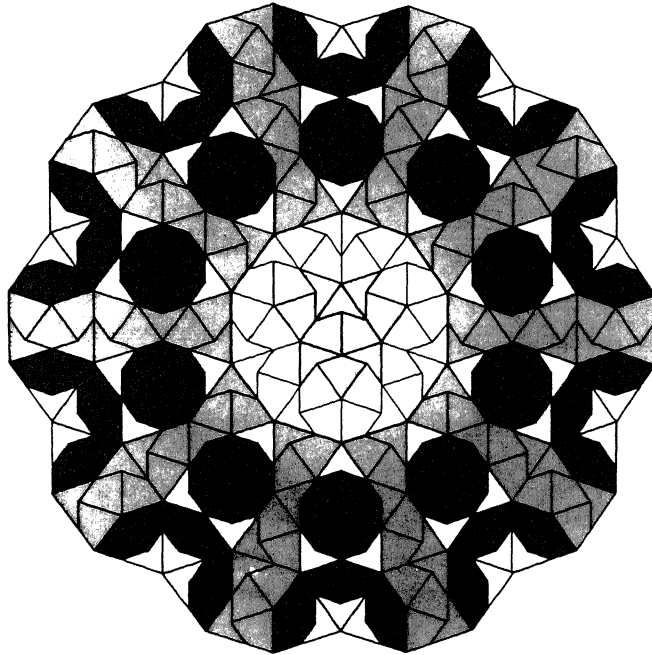
Figure 18

Properties b) and c) above are regularity, indeed symme-
try, properties especially property c) when applied to the same
tiling. The "fractal property" is also a symmetry property of
course. Yet these symmetries are of a different kind then what
we usually understand by the word.

2. CHANGES IN SYMMETRY AS A PARAMETER VARIES

This section contains a number of remarks pertaining to
what can happen to the symmetry group of an object as a para-
meter varies, i.e. during a deformation.

2.1. Example

Consider a rectangle with sides 1 and λ For $\lambda \neq 1$, the symmetry
group is generated by the two reflections across the central
horizontal and vertical axis, so that the symmetry group is the
Klein four group $V_4 = Z/(2) \times Z/(2)$. For $\lambda = 1$ however, there
suddenly appears an additional bit of symmetry, viz a rotation
through $90°$. For this value of λ the symmetry group is suddenly
larger.
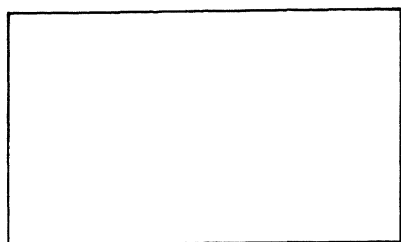
Symmetry group Z/(2) x Z/(2)   Symmetry group D_4
Figure 19
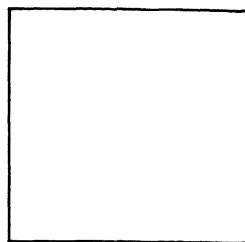
2.2. Example

Consider a small square inside a larger one with the sides
of the small square parallel to those of the big one as in Fig-
ure 20. The parameter which varies is the position of the smal-
ler square

(a)   (b)   (c)   (d)
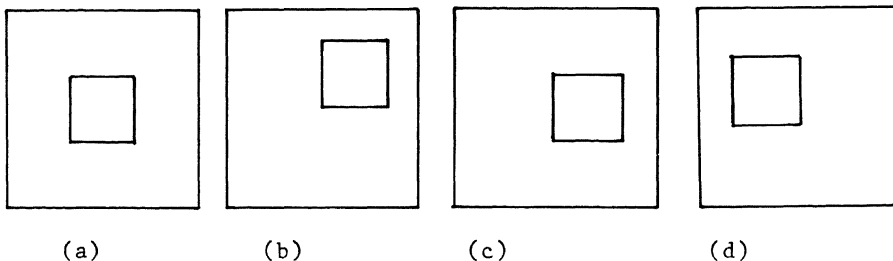
Figure 20

If the centre of the small square is precisely in the middle of
the big one the symmetry group of the whole figure is D_4 (Fig-
ure 20 (a)), if the centre is on a diagonal but not in the cen-
tre the only non-trivial symmetry is a reflection across that
diagonal (Figure 20 (b)). The situation is analogus for the
centre on the horizontal or vertical symmetry axis of the big
square (Figure 20 (c)) and finally if the centre is on none of
these four lines there is no non-trivial symmetry. Thus, in the
parameter square of this example we can describe the symmetry
of the various figures as in Figure 21 below. A similar picture
holds for example 2.1. And indeed in certain reasonably general
situations one can show that this is the general pattern as we
shall see below in section 2.3. Though of course it may happen
that at a certain critical value the symmetry group increases
by an infinite amount as when one considers the rigid
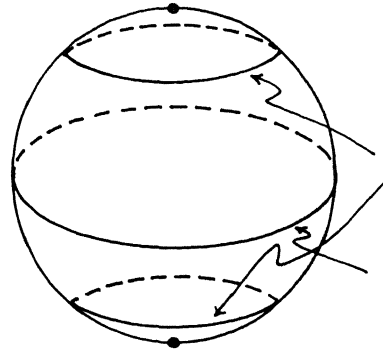
Figure 21   Figure 22

motions of the plane which take a given ellipse with major axis
λ and minor axis 1 into itself. When λ becomes one (circle) the
symmetry group suddenly increases from $Z/(2) \times Z/(2)$ to the
circle group S^1 of all rotations about the centre of the el-
lipse.

2.3. A general mathematical framework for symmetry breaking
     [20, 21]

     We can view the examples above as follows. There is a
group of potential symmetries, in this case the rigid motions
of the plane. For certain isolated parameter values a "large"
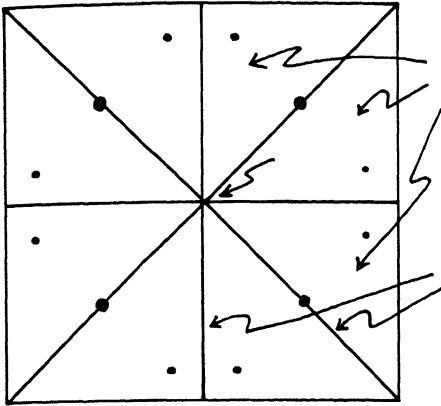subgroup of this group defines actual symmetries and as the
parameter moves away from this value the symmetry is broken to
a smaller subgroup.
     A more precise and mathematical setting for this picture
of symmetry breaking is as follows.
     Let G be a group acting on a set M. (I.e. there is given a
map $G \times M \to M$, written $(g,m) = gm$ such that $g_1 (g_2 m) = (g_1 g_2)m$
$1m = m$; think of gm as the result of applying the transform-
ation g to m). For instance in the case of example 2.2. we
would have $G = E_2$, the group of rigid motions of the plane, and
the set M is the set of all squares of sizes 6 and 2, the lat-
ter inside the other with sides parallel to the co-ordinate
axes of the plane. M can be conveniently described as M
$= \{(P_1, P_2) \quad R^2 \times R^2 : d(P_1, P_2) \leqslant 2\}$ (P_1 is the center of the
first square, P_2 that of the second).

     For each $m \in M$, the _isotropy subgroup_ at m is

$$G_m = \{g \in G : gm = m\} \qquad\qquad (2.1)$$

This is the group of symmetries of the structure represented by m ∈ M. The <u>orbit</u> of m ∈ M is

$$G_m = \{gm : g \in G\} \tag{2.2}$$

G acts transitively on Gm (i.e. for every x,y ∈ Gm there is a g ∈ G such that gx = y). A set with an action of G on it is cal-led a G-set. Two G-sets are isomorphic if there is a bijection φ: M_1 → M_2 such that φ(gm_1) = gφ(m_1) for all $m_1 \epsilon M_1$. If M is a transitive G-set the isotropy subgroups of the points of M are all conjugate and this sets up a bijection between isomorphism classes of transitive G-sets and conjugacy classes of subgroups of G.

A <u>stratum</u> of G is the union of all orbits belonging to one isoclass of G-sets.

Consider for example M = S^2, the sphere, and G = O_2(R) the group of all rotations around the N-S axis and inversion through the origin. An orbit is then the union of two parallel circles at equal northern and southern latitude. There are three strata viz {N} ∪ {S} and the equator, both are strata consisting of a single orbit, and the third stratum is the union of all other orbits. A precisely similar picture is of-fered by the description of example 2.2. by means of Figure 21. By restricting immediately to the subgroup D_4 of E_2 which leaves the larger square invariant we have G = D_4 and M is a square of side 4.

Orbits are e.g. sets of 8 points as indicated by dots in Figure 21, or sets of four points as indicated by crosses, or sets of four points as indicated by small circles and finally the centre is an orbit. There are four strata given precisely by these four types of orbits.

Returning to the general situation. If G is a compact Lie group acting smoothly on a smooth manifold M everything is beautiful: the isotropy subgroups are closed Lie subgroups, or-bits and strata are submanifolds and there exists a G-invariant Riemannian metric on M. A consequence of all this is:

2.4. Theorem

Let G be a compact Lie group acting smoothly on a smooth manifold M. Then for every m ∈ M there is a neighbourhood U of m such that G_m is larger than $G_{m'}$ (up to conjugacy) for all m' ∈ M.

Thus symmetry can suddenly decrease but not suddenly in-crease which is precisely as in examples 2.1. and 2.2. and also as in the example of Figure 22.

Note that in this setting <u>hidden symmetry</u> may occur. To see this consider the following modification of the example of Figure 22. Consider again S^2 and let the group G now consist of rotations around the N-S axis and reflexion through the equator

plane. For M take the space of all unordered pairs of paral-
lels. There is an obvious induced action of G on M by viewing G
as a group of transformations on S^2 so that each element of G
takes an element of M to a possibly different element of M. Let
P be the element of M consisting of twice the equator. The iso-
tropy of P is all of G. The rotations are visible. However, the
reflection (when restricted to the submanifold of S^2 represen-
ted by P) acts just like the identity. That is as a symmetry of
the figure P in S^2 it is a hidden symmetry, which appears as
soon as P is deformed into a pair of close together equal lati-
tude north and south parallels.

2.5. Example

Consider the group of all transformations $R^2 \to R^2$ of the
plane into itself of the form $(f,g) : (x,y) \to (f(x), y + g(x))$
where $f: R \to R$ is a diffeomorphism and $g: R \to R$ is any differ-
entiable map. The inverse of (f,g) is the element $(f^{-1}, -gf^{-1})$
and the composition goes as follows: $(f_1,g_1) \circ (f_1,g_1) = (f_2 f_1,$
$g_1 + g_2 f_1)$. The identity element is (id, o). Thus G is a sub-
group of the group $Diff(R^1)$ of all diffeomorphisms of R^2 into
itself. Now let M be the space of all unordered pairs of ele-
ments of R^2. Consider an element $P = \{x_1,y_1), (x_2,y_2)\} \in M$ and
let us calculate the isotrophy subgroup of P. Suppose $\phi \in G$ is
in G_p. Then we must have

$$\phi(x_1,y_1) = (x_1,y_1) \text{ and } \phi(x_2,y_2) = (x_2,y_2) \qquad (i)$$

or

$$\phi(x_1,y_1) = (x_2,y_2) \text{ and } \phi(x_2,y_2) = (x_1,y_1) \qquad (ii)$$

or both which can only happen if $(x_1,y_1) = (x_2,y_2)$ It is
quite easy to describe the isotropy subgroup for all P. For the
purposes of this example, however, we need only two cases.

$x_1 = x_2$ and $y_1 \neq y_2$.
$$\qquad (a)$$
Then $G_p = \{(f,g): f(x_1) = x_1, g(x_1) = 0\}$

$x_1 \neq x_2$ and $y_1 \neq y_2$

Then $G_p = \{(f,g) : f(x_1) = x_1, f(x_2) = x_2, g(x_1) = \qquad (b)$

$\qquad = g(x_2) = 0\} \cup$

$\quad \{(f,g): f(x_1) = x_2, f(x_2) = x_1, g(x_1) = y_2 - y_1,$

$\quad g(x_2) = y_1 - y_2\}$

In case (b) the set

$N_p = \{(f,g) : f(x_1) = x_1, \ f(x_2) = x_2, \ g(x_1) = g(x_2) = 0\}$ is a

normal subgroup of G_p and G_p is in fact the disjoint union N_p

αN_p where $\alpha = (f,g)$ is any element of G such that

  $f(x_1) = x_2, \ f(x_2) = x_1$, $g(x_1) = y_2 - y_1$, $g(x_2) = y_1 - y_2$.

    Now let $P = ((x_1,y_1), (x_2,y_2))$ approach a point Q

$=((x,y_1), (x,y_2))$ with $y_1 \neq y_2$. Then we see that as the limit
point is reached there is both symmetry gain in that the N_p
part of G_p becomes bigger and sudden symmetry loss in that the
αN_p part of G_p disappears. It is in fact easy to show that
there are no inclusion up to conjugacy relations between G_p and
G_Q.
    Thus theorem 2.6. does not hold in more general situ-
ations. It seems that this kind of phenomenon cannot happen
when considering the symmetry group (of motions) of figures or
patterns in Euclidean space as these figures change. But I do
not know of a general theorem to this effect.

3. DESIGN SYMMETRY VERSUS GENERIC SYMMETRY [14]

    In example 2.7. we saw that in more general cases a state-
ment like that of theorem 2.6 does not hold. A more complicated
but also more suggestive example of the same phenomenon of si-
multaneous symmetry loss and symmetry gain during a deformation
is obtained by considering the automorphisms of three (or
higher) dimensional algebras over R. That is the subject of
this section.

3.1. Algebra structures

    Let $V = R^3$ be the vectorspace of all 3-tuples of real num-
bers. An associative algebra structure with unit on V is given
by a bilinear map (the multiplication).
    m: $V \times V \to V$, $(x,y) \to xy$

such that $(xy)z = x(yz)$ and such that there exists a $1 \in V$ with
$1x = x1 = x$ for all $x \in V$. By choosing a basis in V suitably we
can assume that $1 = e_1 = (1,0,0)$ and we shall do so. Let e_2, e_3
be the other basis elements. Then because of the bilinearity
the multiplication is specified by 12 constants (the so-called
structure constants)

$$e_2 e_2 = \gamma_{22}^1 e_1 + \gamma_{22}^2 e_2 + \gamma_{22}^3 e_3 \qquad e_2 e_3 = \gamma_{23}^1 e_1 + \gamma_{23}^2 e_2 + \gamma_{23}^3 e_3$$

$$e_3 e_2 = \gamma_{32}^1 e_1 + \gamma_{32}^2 e_2 + \gamma_{32}^3 e_3 \qquad e_3 e_3 = \gamma_{33}^1 e_1 + \gamma_{33}^2 e_2 + \gamma_{33}^3 e_3$$

In order that the algebra be associative these γ_{ij}^k have to satisfy certain relations. E.g. from $e_2(e_2 e_3) = (e_2 e_2)e_3$ one obtains

$$\gamma_{23}^2 \gamma_{22}^1 + \gamma_{23}^3 \gamma_{23}^1 = \gamma_{22}^2 \gamma_{23}^1 + \gamma_{22}^3 \gamma_{33}^1$$
$$\gamma_{23}^1 + \gamma_{23}^3 \gamma_{23}^2 = \gamma_{22}^3 \gamma_{33}^2 \qquad\qquad (3.1)$$
$$\gamma_{23}^2 \gamma_{22}^3 + \gamma_{23}^3 \gamma_{23}^3 = \gamma_{22}^1 + \gamma_{22}^1 \gamma_{23}^3 + \gamma_{22}^3 \gamma_{33}^3$$

The precise form of these conditions will not be important for us.

3.2. Isomorphisms and automorphisms

A map $\phi: A \to B$ from an algebra A to an algebra B is an isomorphism if it is an isomorphism of vectorspaces and if moreover $\phi(1_A) = 1_B$ and $\phi(xy) = \phi(x)\,\phi(y)$ for all $x,y \in A$. Here 1_A and 1_B denote the unit elements of A and B. An isomorphism $\phi: A \to A$ is called an automorphism and Aut(A) denotes the group of automorphisms of A. This group should also be considered as the group of symmetries of the algebra A. Indeed the situation fits the general framework described above. Let G be the group of all vector space isomorphisms $\phi: V \to V$ such that $\phi(e_1) = e_1$ (because we only consider algebra structures on V for which e_1 is the unit element). Let M be the space of all 12-tuples $(\gamma_{22}^1, \ldots, \gamma_{33}^3)$ such that all associativity relations like

(3.1) hold. For $\phi: V \to V$, $\phi(e_1) = e_1$, let $\bar{e}_2 = \phi(e_2)$, $\bar{e}_3 = \phi(e_3)$ and let $\bar{\gamma}_{22}^1, \ldots, \bar{\gamma}_{33}^3$ be defined by

$$\bar{e}_i \cdot \bar{e}_j = \bar{\gamma}_{ij}^1 e_1 + \bar{\gamma}_{ij}^2 \bar{e}_2 + \bar{\gamma}_{ij}^3 \bar{e}_3 \quad , \; i,j = 2,3$$

The group element ϕ now acts on M by

$$(\gamma_{22}^1, \ldots, \gamma_{33}^3) \to (\bar{\gamma}_{22}^1, \ldots, \bar{\gamma}_{33}^3).$$

One sees immediately that if A is an associative algebra with

structure constants $\gamma = (\overset{1}{\gamma_{22}}, \ldots, \overset{3}{\gamma_{33}})$ then $\mathrm{Aut}(A) = G_\phi$, the

isotropy subgroup of γ M.

Two algebras A and B are isomorphic if and only if they
(or more precisely the corresponding 12-tuples of structure
constants) are in the same orbit.

3.3. The isomorphism classes of three dimensional algebras

It turns out that up to isomorphism there are six differ-
ent algebra structures on $V = R^3$. They are

$A_1 \simeq R[X]/X(X-1)(X-2)$

$A_2 \simeq R[X]/X(X^2 + 1)$

$A_3 \simeq R[X]/X^3$

$A_4 \simeq R[X]/X^2(X-1)$

$A_5 \simeq R[X,Y]/(X^2,Y^2,XY)$

A_6 with basis 1, e_2, e_3 and the multiplication rules
  $e_2^2 = 1$, $e_3^3 = 0$, $e_2e_3 = e_3$, $e_3e_2 = -e_3$

Here if $f(X) = X^3 + a_2X^2 + a_1X + a_o$ is a polynomial $R[X]/f(X)$

denotes the 3 dimensional algebra with basis 1, X, X^2 and multi-

plication rules $XX = X^2$, $XX^2 = X^2X = -a_2X^2 - a_1X - a_o$. The al-

gebra A_5 is defined similarly.

3.4. The deformation/contraction relations between the six iso-
morphism classes

By the symbol $A \Longrightarrow B$ we understand that there is a family
of algebra structures $A(t) = (\overset{1}{\gamma_{22}}(t), \ldots, \overset{3}{\gamma_{33}}(t))$ isomorphic
for small $t \neq o$ to A and such that $A(o)$ is isomorphic to B. In
other words $A \Rightarrow B$ means that the orbit corresponding to the
isomorphism class B is in the closure of the orbit correspond-
ing to isomophism class A.

With this notation the pattern of contraction/deformation
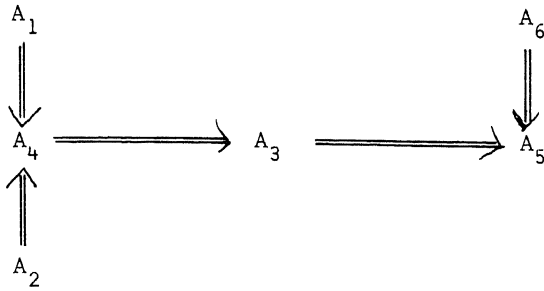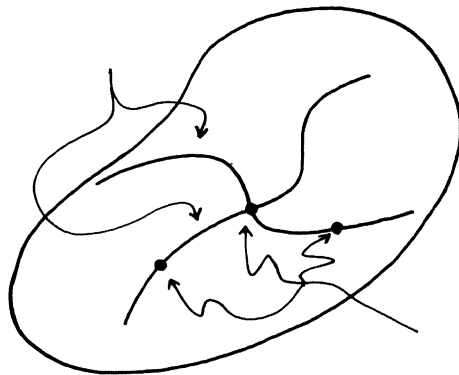relations between $A_1, \ldots, A_6$ is as follows

$$A_1$$

$$\Downarrow$$

$$A_4 \xrightarrow{\hspace{3cm}} A_3 \xrightarrow{\hspace{3cm}} A_5$$

$$\Uparrow$$

$$A_2$$

$$A_6$$

$$\Downarrow$$

Figure 23

Figure 24

Concentrating on the A_1, A_3, A_4 part of Figure 23 the situation in γ-space is somewhat like depicted in Figure 24. Compare this e.g. with Figure 22 and Figure 21.

The corresponding pattern of automorphism groups is as follows

$$\text{Aut}(A_4) = S_3 \qquad\qquad\qquad\qquad \text{Aut}(_6) = \{(a,b) : b \neq o\}$$

$$\Downarrow \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \Downarrow$$

$$\text{Aut}(A_4) = R \setminus \{o\} \xRightarrow{\hspace{1cm}} \overset{\text{Aut}(A_3) =}{\underset{\{(a,b) \ : \ a \neq o\}}{}} \xRightarrow{\hspace{1cm}} \overset{\text{Aut}(A_5) =}{\underset{GL_2(R)}{}}$$

$$\Uparrow$$

$$\text{Aut}(A_2) = S_2$$
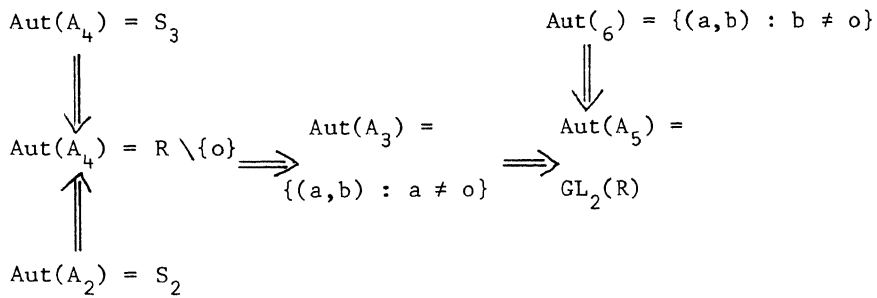
Figure 25

Here S_n is the permutation group on n letters and $GL_m(R)$ is the
group of real invertible m x m matrices. The multiplication
rules of $Aut(A_3)$ and $(Aut(A_6)$ are respectively (a,b)(c,d) =

(ac, ad + bc^2), (a,b)(c,d) = (c+ad, bd).

We see that as a rule during a contraction (⟶)
(i)   symmetrics are both gained and lost
(ii) the dimension of the symmetry group does not become
      less

3.5. Design versus generic symmetry

At least in a large number of cases. there seem to be two
sources of symmetry. The first is what I like to call "generic
symmetry" it is the symmetry which is possessed by almost all
of the structures under consideration. As an example consider
algebras of the form $C[X]/f(X)$ where C denotes the complex num-
bers and $f(X)$ is a polynomial of degree n. Almost all $f(X)$ have
n distinct roots and as a consequence almost all of these alge-
bras have S_n as their automorphism group. The second source is
what I like to call design symmetry which arises e.g. when
parts of the structures under consideration are very carefully
arranged in such a way that a large symmetry group arises. For
instance if precisely two roots of $f(X)$ are equal and all
others are different from each other and from this double root,
then the automorphism group picks up a factor R-{0} but not
all roots are of the same kind anymore and the generic symmetry
group drops to S_{n-2} (the permutations of the (n−2) unequal
single roots). It seems to me that the way the symmetry groups
can change often can be understood systematically in these
terms. During contraction (⟹) the generic symmetry group tends
to become smaller and the design symmetry group larger or,
equivalently, during a deformation (⟸) the generic symmetry
group becomes larger and the design symmetry group smaller
(i.e. gets broken). This last phenomenon was of course the sub-
ject matter of section 2 above.

4. CONSEQUENCES OF THE PRESENCE OF SYMMETRY

Quite generally the presence of symmetry in a (mathemati-
cal) problem can have enormous influence and it can greatly
facilitate solving a problem. We shall see this when examining
bifurcation phenomena in the presence of symmetry in the next
section and we have already seen examples in the previous sec-
tion. Here we describe some more material around this theme.

4.1. Symmetric problems and their solutions

A first question to examine is whether symmetric problems necessarily have symmetric solutions. This is discussed in considerable detail by W.C. Waterhouse in [40] who calls the principle that this be the case the Purkiss principle. Here are some of his examples where the principle holds.
- of all rectangles with a given perimeter the square has the largest area.
- Take four positive numbers whose product is 16. Then their sum is least when all numbers are equal.
- For a given mean $\overline{x} = n^{-1} (x_1 + \ldots + x_n)$ the value of $x_1^2 + \ldots + x_n^2$ is least when all x_i are equal.
    It is clear that statements to the effect that under certain circumstances the Purkiss principle holds are precisely the desired sort of mathematical explanations of why nature likes symmetric solutions (Cf. the introduction).
    There are also quite simple counter examples to the Purkiss principle. For instance one from Bpniakovsky: find the minimum of $f(x,y) = (x^2 + (y-1)^2)((x-1)^2 + y^2)$. This is symmetric in x and y. The two solutions, however, are (1,0) and (0,1). Another example (from [20]) which I like very much is the following. Consider four towns located on the corners of a square. What is the shortest road system that joins these four towns. It is not very difficult to see that there are two solutions which are depicted in Figure 26. The angle between the horizontal and the top left oblique segment in Figure 26 is 30°. This solution is better than the other obviously possible candidates: three edges of the square or the two diagonals.
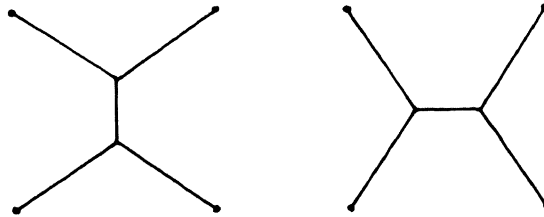

Figure 26

Note that in both counter examples the total set of solutions is symmetric under the full symmetry group of the problem but that the individual solutions (if there is more than one) are only invariant under an isotropy subgroup. That is, there is symmetry breaking in precisely the sense of section 2 above with as the manifold M the space of all solutions.

4.2. Extrema of symmetric functions

    Here is a result that shows how strong the influence of
the presence of non-trivial symmetry can be. The setting is
that of section 2.3 above, i.e. a compact Lie group G (for in-
stance a finite one) acting smoothly on a smooth manifold M.
Let F be the set of all functions f on M which are invariant
under G, i.e. such that f(gm) = F(m) for all m ε M, g ε G

    Theorem (cf. e.g. [21]). If an orbit is isolated in its
stratum it is critical for all f Є F (i.e. df = 0 at all points
of that orbit) and inversely if an orbit is critical for all
f Є F it is isolated in its stratum.

5. BIFURCATION IN THE PRESENCE OF SYMMETRY

5.1. General remarks and first examples

    Bifurcation theory is concerned with how the set of solu-
tions of a problem can change as a parameter varies. For a
first introdution to bifurcation theory I refer the reader to
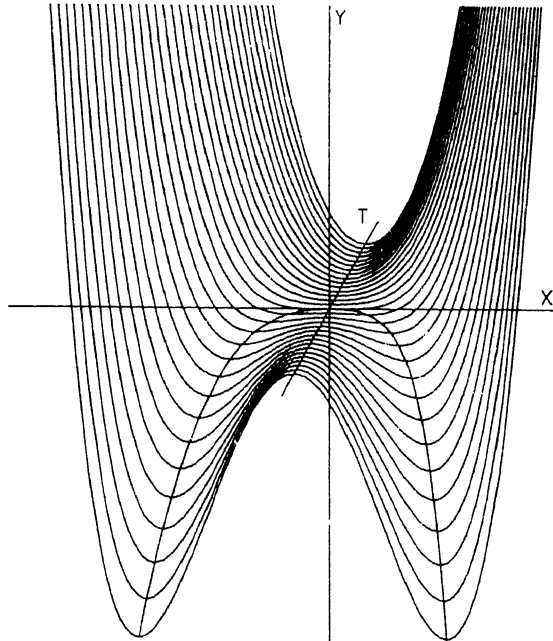my chapter "Bifurcation phenomena. A short introductory tu-

Figure 27

torial with examples" in this volume [15]. Two of the simplest
examples of bifurcations are the pitchfork bifurcation depicted
in Figure 27 where a minimum bifurcates into two minima and a
maximum, and the Hopf bifurcation where a stable equilibrium
point bifurcates into an oscillatory cycle. In the case of the
bifurcating valley there is also obviously a kind of symmetry
breaking involved and in the case of the Hopf bifurcation also
seems to involve symmetry loss when viewed in space-time space,
cf. fig 28 below. The continuous translations symmetry gets
broken into a discrete group of translation symmetries. Quite
generally it seems clear from the setting of sections 2 and 4
above that if a solution of a problem is symmetric with
symmetry group G, and for parameter values $\lambda < \lambda_0$ there is
(locally) a single solution, and if at λ_0 this solution
bifurcates into several, then the new solutions will have as
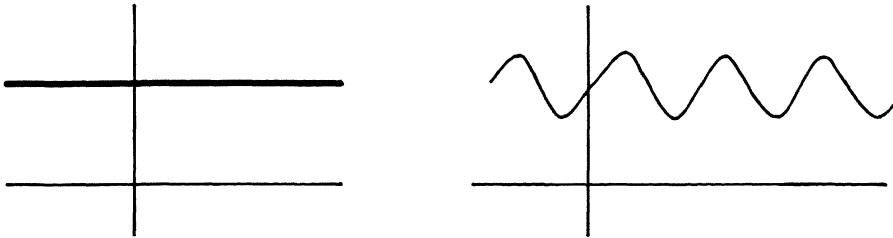symmetry group isotropy subgroups of G. This should severely

Figure 28

restrict the possible bifurcations. To see how this works, how
the results of section 4 apply, and how representation theory
can be used to advantage we first need a slightly more precise
description of a little bit of bifurcation theory.

5.2. Microsynopsis of some bifurcation theory

Consider a map $G(\lambda,.) : B_1 \to B_2$ depending on a parameter λ
(where λ may also denote a vector of parameters). Here the B_i
are Banach spaces. At this stage it does not hurt to think of
the B_i as finite dimensional spaces, say $B_1 = B_2 = R^2$ and to
think of $(\lambda,.)$ as a map given by an expression like $(x^3 + y^3 + \lambda xy^2, x^2 + \lambda xy)$. We are interested in studying $G(\lambda,u) = 0$, $u \in B_1$, and in how the solution set of this equation changes as λ
varies. In most of the interesting examples the B_i are suitable
spaces of functions and $G(\lambda,.)$ is e.g. a differential operator
depending smoothly on λ. An explicit example occurs below.

Suppose $G(\lambda_0, u_0) = 0$; consider the partial derivative
$G_u(\lambda_0,u_0)$ at (λ_0, u_0). This is a linear map $B_1 \to B_2$. If this
linear map is invertible then the implicit function theorem
says that there is a differentiable mapping $\lambda \to u(\lambda)$ for λ near

λ_0 such that $G(\lambda, u(\lambda)) = 0$ and that locally (i.e. for (λ, u) close to (λ_0, u_0) this is the only solution. Thus for bifurcation phenomena we are interested in points where $G_u(\lambda_0, u_0)$ is not invertible. In the finite dimensional case of $B_1 = B_2 = R^n$ we are thus interested in cases where the n × n matrix of partial derivatives $G_u(\lambda_0, u_0)$ is not of full rank. If there is a symmetry group involved it is obvious how (generalised) results like those of section 4.2 above could, indeed will be important.

Assume that $G_u(\lambda_0, u_0)$ is a Fredholm operator of index zero, so that the kernel is finite dimensional and the range is closed of co-dimension equal to the dimension of the kernel (automatically the case if $B_1 = B_2 = R^n$.). In the simplest case dim N = 1, where N = Ker $G_u(\lambda_0, u_0)$. Let $B_1 \subset B_2$ and let P : $B_2 \rightarrow$ N be the projection on N and write Q = Id − P. Then we can rewrite the equation $G(\lambda, u) = 0$ as

$$QG(\lambda, v + \psi) = 0 \quad \text{and} \quad PG(\lambda, \psi) = 0$$

where v = Pu, ψ = Qu. Fix v, then $QG_u(\lambda_0, u_0)$ is an isomorphism, so by the same implicit function theorem used above we can find $\psi(\lambda, v)$ as a function of (v, λ) so that $Q(\lambda, v + \psi(\lambda, v)) = 0$ near (λ_0, v_0). Thus it remains to solve the so-called bifurcation equations

$$F(\lambda, v) \equiv PG(\lambda, v + \psi(\lambda, v)) : R \times N \rightarrow N, \quad F(\lambda, v) = 0$$

In the simplest case (dim N = 1) it readily follows that $F(\lambda, v)$ is of the form

$$F(\lambda, v) = a(\lambda - \lambda_0) v + \dots.$$

so that in the non-degenerate case, the solution set near $(\lambda_0, 0)$ looks like two lines crossing each other vertically, a pitchfork bifurcation.

5.3. Equivariant bifurcation theory

Now suppose that there is a group of symmetries involved. I.e. there is a group H acting linearly on B_1 and B_2 and $G(\lambda, u)$ is equivariant which means that $G(\lambda, gu) = g G(\lambda, u)$ for all g G. Here is a general result [33, theorem 13].

Theorem. Let $G(\lambda, u) : B_1 \rightarrow B_2$ be analytic and equivariant w.r.t a compact group H. Let $G(\lambda_0, u_0) = 0$, $gu_0 = u_0$ for all g H, and let $G_u(\lambda_0, u_0)$ be Fredholm of index zero with kernel N_0. Then N_0 is invariant under H (i.e. gN_0 N_0 for all g H) and the $F(\lambda, v)$ are equivariant.

Now in many interesting bifurcation problems dim N > 1. Then the simple analysis of 5.2 above does not apply. However,

if there is symmetry it may easily happen that N is an irredu-
cible representation of H and if we know which one (i.e. as a
rule, if we know enough of the representation theory of H) this
is just as good as the case dim N = 1.

    We also know a priori that the bifurcating solutions will
have symmetry groups which are isotropy subgroups of H acting
on N. A very simple example of a bifurcation situation with
symmetry (rotational symmetry in this case) is the one of a
stiffish rubber bar with opposite forces acting on the two end
points as shown in Figure 30. The corresponding bifurcation
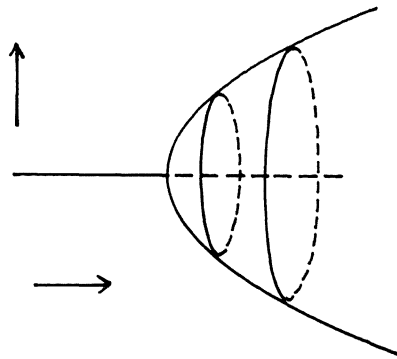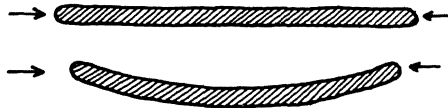diagram is sketched of Figure 29.

Figure 29

Figure 30

    The idea is the following. Just as in case dim N = 1 we
know that $F(\lambda,v)$ must start off with a term λv (if $\lambda_o = 0$), we
know in the equivariant case that $F(\lambda,v)$ must start off with
very specific terms which are determined by the representation
of H, which is involved. In many cases degree considerations
and the value of dim N rule out all but a few possible (known)
representations which may make a case like dim N = t (which is
usually totally intractable in the general case) quite easy to
do in a symmetric case.

5.4. Example: Bénard convection [33, 34]; cf. also [9]

    In case of the Bénard convection such an analysis has ac-
tually been carried out by Sattinger, loc. cit. The equations
involved are the Boussinesq equations which are

$$\Delta u_k + \delta_{k3}\, \theta - \frac{\partial p}{\partial X_k} = P_r^{-1} \sum_j u_j \frac{\partial u_k}{\partial X_j} \qquad , \ k = 1,\ 2,\ 3$$

$$\Delta \theta + Ru_3 = \sum_j u_j \frac{\partial \theta}{\partial X_j}$$

$$\sum_j \frac{\partial u_j}{\partial X_j} = 0$$

Here u_1, u_2, u_3 are the velocity components of a fluid element as functions of the co-ordinates X_1, X_2, X_3. The function θ is the temperature profile and P is the pressure. R denotes the Rayleigh number, P_1 the Prandtl number and Δ denotes the Laplacian

$$\frac{\partial^2}{\partial X_1^2} + \frac{\partial^2}{\partial X_2^2} + \frac{\partial^3}{\partial X_3^2}$$

and δ_{k1} the Kronecker delta. The bifurcation parameter λ is R. Thus in this case B_1 and B_2 are suitable spaces of 5-tuples of functions $(u_1, u_2, u_3, \theta, P)$ of three variables X_1, X_2, X_3, and it is then immediately clear how to write these equations in the form $G(\lambda, u) = 0$.

The group involved is E_2, the group of rigid motions of the plane which consists of all motions

$$g: \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \to \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$$

where the matrix is orthogonal and hence of determinant ± 1.

The action of E_2 on the spaces of 5-tuples of functions B_1, B_2 is now given by

$$gf(X_1, X_2, X_3) = \begin{pmatrix} B & 0 \\ 0 & I \end{pmatrix} f\,(g^{-1}(X_1, X_2), X_3) \quad , \quad f \in B_i,\ g \in E_2$$

Here B is the 2 x 2 matrix which gives the rotation/reflection part of $g \in E_2$ and I is a 3×3 identity matrix. It is a small exercise to check that the Boussinesq equations are indeed equivariant w.r.t. this action. The reason is that the physics is independent of the observer.

In this case it turns out that dim N is infinite. However,
under the assumption that we restrict our attention to solu-
tions which are periodic with respect to a hexagonal lattice,
cf. Figure 31, dim N becomes finite dimensional and an explicit
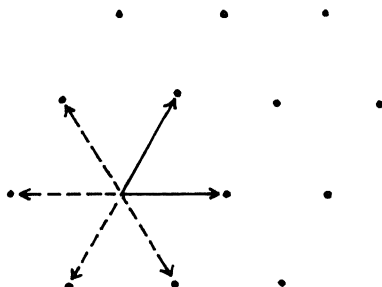bifurcation analysis can be carried out [33, 34, cf. also 9].

Figure 31

It turns out that the stability of the bifurcation solutions
depends on two parameters in the manner depicted in Figure 32
below. This agrees with experimental data. Of course, this does
not give us a complete mathematical description of the Benard
convection. It remains to be shown that the E_2 symmetry has to
break through a lattice. This is still an open problem, though
it is hard to see how E_2 symmetry could get broken otherwise to
a stable solution (cf. also [38, section II.1]). It also re-
mains to analyse the relative stability of various lattice pat-
terns with respect to each other. Cf. in this connection [31,
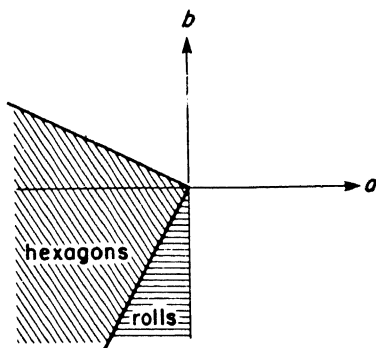32].

Figure 32

5.5. Drainage patterns. An example . Competing singularities

Let me conclude with a possible model for the formation of

drainage patterns due to Thom [38], which I like particularly.
Imagine a sandy slope on which a gentle rain is continuously
falling. At the top small rivulets will form at random, these
will merge and form larger rivulets further down etc., to form
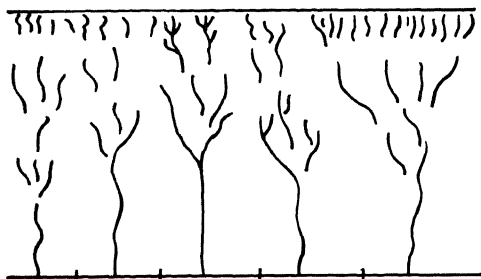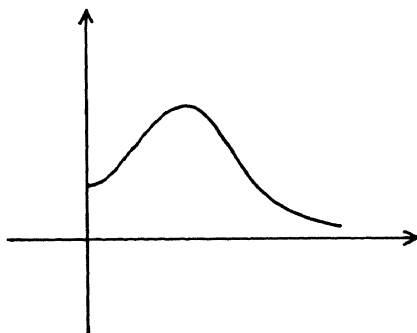a pattern somewhat like the one depicted in Figure 33.

Figure 33

Figure 34

Let X_i denote the position of the i-th watershed near or at the
bottom. Assume that the "erosion power" of a stream is propor-
tial to the width of its basin. Then the position of the i-th
watershed will move according to the differential equation

$$\dot{X}_i = C(X_i - X_{i-1}) - C(X_{i+1} - X_i) = C(2X_i - X_{i-1} - X_{i+1})$$

$$(5.1)$$

Obviously equipartition $(X_i = \frac{1}{2}(X_{i-1} + X_{i+1}))$ is a stationary
solution. For an analysis of the stability of such a solution
consider two streams at positions X and -X on R with the divide
at u near zero. Then assuming that C also depends on X we find

$$\dot{u} = 2\ C(X)\ u + 2\ XC'(X)\ u + u^2\ (.....)$$

for small u. So that we will have stability if $C(X) + XC'(X) <$
0. Assuming that $C(X)$ is somewhat like Figure 34, which is not
unreasonable, one would expect a characteristic wave length for
the pattern at bottom given by L = "smallest" X such that $C(X)$
$+ X C'(X) < 0$. The picture of Death Valley some pages back
(Figure 4) is a nice example of just such a pattern with, appa-
rently, a characteristic wave length. It remains of course to
do a complete (bifurcation) analysis which leads from the ran-
dom pattern at the top to the regular pattern at the bottom.
The governing equation above is a discretisation of a certain
(anti-) diffusion equation. Numerical experiments, [16], with
such equations are encouraging but it seems that we are quite
far from a completely satisfactory theory at the moment.