The Morphogenetic Models of René Thom

Jean Petitot
CAMS, EHESS

2005

At the end of the 1960s, René Thom was the first scientist to develop a general mathematical theory of morphogenetic processes. This chapter presents the fundamental principles of that theory.

1 General content of the model

Let S be a system satisfying the following hypotheses:

(a) Tthere exists an *internal process* (usually unobservable) X which defines the internal states that system S can occupy in a stable manner, and the number of these states is finite;

(b) the internal states of S are in competition with each other and mutually determine each other, and the choice of one of them as the actual state makes the others virtual;

(c) there therefore exists an *instance of selection* I which, on the basis of criteria specific to the system, selects the actual state from among the possible internal states;

(d) the system S is controlled by a certain number of control parameters varying within a space W which we call, to distinguish it from the internal process X, the *external space* (or *control space* or *substrate space*) of S.

We also assume that the control is *continuous*, in the sense that the internal process X is a process X_w which depends continuously on the value w of the control. This process varies continuously when w varies continuously in W and when it is deformed it also deforms the structure of the internal states and their relations of mutual determination. We denote $\mathscr{X}$ the space of possible internal processes X. If the above hypotheses are verified, the system S will be described firstly by the (continuous) field $\sigma : W \to \mathscr{X}$ associating $w \in W$ with the process X_w and then by the instance of selection I.

The standard example is that of the thermodynamic phenomena of phase transitions. In this case, the system S is the thermodynamic system considered, the internal states are the thermodynamic phases (solid, liquid, gas), the instance of selection I is provided by the principle of free energy minimisation and the control parameters are, for example, pressure and temperature. As for the internal process X_w, indescribable because of its complexity, it is the process of molecular dynamics. When the control parameters cross certain specific values, known as critical values[1], they present *phase transitions*, i.e. discontinuities in their observable qualities and abrupt transformations of their internal state. The critical values constitute a subset K of W, determining the *phase diagram*: K partitions W into domains corresponding to the different phases that S can present. In other words, it categorises it and externalises therein the competition between internal states, in the form of a system of discontinuities.

This is a direct consequence of hypotheses (a)-(d). A system $S = (W, \mathscr{X}, \sigma, I)$ manifests itself phenomenologically by the observable qualities $q^1, \dots, q^n$ expressing its internal state. In other words, the internal process X_w is externalized in perceptible qualities q^i_w. When the control w varies continuously, the actual internal state varies continuously (hypothesis (d)) and the qualities q^i_w therefore vary equally. But phenomenologically speaking, a continuous variation is no more than a form of qualitative invariance. It is therefore not significant. So René Thom denoted by *regular point* of W a value w of the control such that the observable qualities q^i_w vary continuously — and therefore remain stable — throughout a neighbourhood U of w (this obviously presupposes that we have defined the concept of neighbourhood on W, i.e. a topology). By definition, the regular points constitute an open set R_W of W, the open set of quality stability. Then let K_W be the closed set complementary to R_W in W. By definition, the points of K_W are the values w of the control such that at least one observable quality q^i_w suffers a discontinuity. These are critical values, crossing which the system S presents critical behaviour. They are also called *catastrophic values*, the closed set K_W being called the *catastrophic set* of S. The K_W are also called *external morphologies*.

As René Thom often stressed, this concept of morphology is purely phenomenological. But it is closely connected to the mathematical concept of *bifurcation*. Let us suppose that the control w follows a path γ in W. Let A_w be the actual internal state initially selected by I. During the deformation of X_w along γ — and therefore, under hypothesis (d), of the structure of A_w and under hypothesis (b), of the relations of mutual determination it has with the virtual states B_w, C_w, etc. — when A_w crosses a critical value, it no longer satisfies the criteria of selection imposed by I under hypothesis (c). The system therefore spontaneously bifurcates from A_w towards another actual (but hitherto virtual) state B_w. This catastrophic transition of internal state manifests itself by a discontinuity in some of the observable qualities q^i_w. In other words, it is the destabilisation (relative to the instance I) of actual internal states under the variation of the control which induces, in the external space W, a set of qualitative discontinuities K_W. In the right cases, the set K_W will constitute a system of interfaces, analogous to a phase diagram, partitioning the external space W into domains, each of which corresponds to a zone in W where one of the internal states is dominant.

[1] Here, the term *critical value* is related to bifurcation theory (and, in the rest of this chapter, to catastrophe theory), and not to the language of thermodynamics. As a general rule, these critical values do not correspond to a critical point in the thermodynamic sense (a particular point where the distinction between the different phases disappears and the phase transition becomes continuous, the singularity manifesting itself in thermodynamic derivatives, of free energy, for example).

2 Morphodynamics and structural stability

Thomian morphodynamics is based on the possibility of specifying the general model in mathematical terms. The first specification consists in assuming that, with regard to their nature, the internal processes X_w constitute a space $\mathscr{X}$ equipped with a natural *topology* $\mathscr{T}$ significant for the type of process studied. This means that we can tell when two internal processes X and Y are neighbours and we can therefore rigorously define the continuity of the field $\sigma : W \to \mathscr{X}$. By moving in $\mathscr{X}$ we can therefore deform its elements X.

We then assume that we can define the *qualitative type* of the processes X. The qualitative type is a relation of equivalence (generally defined by the action of a group G on $\mathscr{X}$), which is a weak, qualitative identity. Let $\tilde{X}$ be the class of equivalence of X for the qualitative type (i.e. the orbit of X under the action of G). We seek to characterise that which remains invariant when X varies in $\tilde{X}$ (i.e. varies for a constant qualitative type) by means of discrete information, for example the values of a finite number of invariants. At the level of the invariants, the variation in a class of equivalence $\tilde{X}$ is reduced to the identity. Therefore, there is only qualitative variation when a deformation in $\mathscr{X}$ causes a change in the class of equivalence. The variation is manifested by a discontinuity in the value of certain invariants and we find the "right" situation of the general model.

Compared to a standard approach, which would consist in studying, for each physical system of type $S = (W, \mathscr{X}, \sigma, I)$, the internal processes X_w as isolated entities, morphodynamics introduces a double shift in perspective. Firstly, it takes as its object of study not only the processes X_w but also the parameterized families $(X_w)_{w \in W}$, by focusing on the geometry of the catastrophic sets K_W induced in the external spaces W by the destabilisation of the internal states defined by X_w. Secondly, and above all, it considers these families as the image of fields $\sigma : W \to \mathscr{X}$ sending the external space W (which is generally a part of the standard space $\mathbf{R}^n$ with n dimensions) into the generalised space $\mathscr{X}$.

Now, from the moment that we possess, for a space $\mathscr{X}$, a topology $\mathscr{T}$ and a relation of equivalence defining the qualitative type, we can define a concept of structural stability. Let $X \in \mathscr{X}$. We say that X is *structurally stable* if all Y close enough to X (in the sense of $\mathscr{T}$) are equivalent to X. The process X is therefore structurally stable if its qualitative type resists small perturbations, or if the class $\tilde{X}$ is open (in the sense of $\mathscr{T}$) in X. Let $K_{\mathscr{X}}$ then be the closed subset of $\mathscr{X}$ composed of the structurally unstable $X \in \mathscr{X}$. $K_{\mathscr{X}}$ is an intrinsic catastrophic set, canonically associated with $\mathscr{X}$. It is categorised by a discriminating morphology, which classifies the qualitative types of its structurally stable elements.

Let $\sigma : W \to \mathscr{X}$ be the characteristic field of a system $S = (W, \mathscr{X}, \sigma, I)$. Let $K'_W = \sigma^{-1}(K_{\mathscr{X}} \cap \sigma(W))$ be the trace of $K_{\mathscr{X}}$ on W through the intermediary of σ. The hypothesis of morphodynamic modelling is that the catastrophic set K_W of S can be deduced from K'_W on the basis of the instance of selection I. This means that a value w of the control belongs to K_W (i.e. is a critical value) if and only if the situation in w is correlated in the manner regulated by I with a situation belonging to K'_W. The external morphology is essentially the apparent outline on the substrate space of the internal dynamics. It is therefore the analysis (at the same time local and global) of the intrinsic catastrophic sets $K_{\mathscr{X}}$ that lies at the heart of this theory.

If we introduce the additional hypothesis that a field σ can only concretely exist if it is itself structurally stable, we see that such a constraint drastically limits the complexity that K'_W can present. In the right cases, we can even obtain a classification of the local structures of K'_W and therefore of the local external morphologies. The theory thus brings to light purely mathematical constraints acting on the domain of morphogenetic phenomena.

3 The theory of dynamical systems

The main mathematical specification of the general model consists in postulating that the internal process X is a *differentiable dynamical system* on a differentiable manifold M of internal parameters characteristic of the system S considered. We call the space M the internal space (to distinguish it from the external space W). A dynamical system X on M consists in associating with each point x of M a tangent vector $X(x)$ of M at x, a vector varying differentiably with x. X is therefore a vector field differentiable on M, in other words, in terms of local coordinates $x_1,\ldots,x_n$, a system of ordinary differential equations:

$$\frac{dx_i}{dt} = f_i(x_1,\ldots,x_n)$$

where the f_i (components of the field) are differentiable functions of x_j and where t is the time parameter. Given such a field, integrating it consists in finding, in M, differentiable curves parameterised by time (i.e. differentiable applications $\gamma : \mathbf{R} \to M$, $t \mapsto \gamma(t) = (x_1(t),...,x_n(t))$ which admit at each point for velocity vector $dx/dt = d\gamma(t)/dt$ the field vector $X(x) = X(\gamma(t))$. We say that the field is a dynamical system if we can integrate the trajectories over an infinite time (i.e. if the trajectories do not leave M); if one and only one trajectory passes through each point (the principle of determinism: the initial condition $x(0)$ at time $t = 0$ univocally determines the future evolution $x(t)$ for $t > 0$ and the past evolution $x(t)$ for $t < 0$); and if the trajectories vary differentiably according to the initial conditions.

Then let $f_t : M \to M$ be the application which assigns to every point x of M the point at time t of the trajectory emanating from x at time $t = 0$. It is easy to see that f_t is a diffeomorphism of M (an automorphism of its differentiable manifold structure) and that the application $t \mapsto f_t$ of the additive group $\mathbf{R}$ in the group of diffeomorphisms of M is a morphism of the group. f is called the *flow* of the dynamical system X. It is the integral version of the vector field X.

René Thom suggested that the models $S = (W, \mathscr{X}, \sigma, I)$ where the internal processes X_w are dynamical systems should be called *metabolic models*. Their internal states need to be defined. The basic idea is to introduce a difference between fast and slow dynamics, in other words between two timescales, one internal and fast, the other external and slow. In the internal space, the fast internal dynamics creates attractors that specify the local phenomenological quality of the substrate. The slow external dynamics operates in the substrate space W. As we assume that the internal dynamics of the evolution of instantaneous states is infinitely fast compared to the external dynamics of evolution in the external spaces W (a condition known as *adiabaticity*), the only significant states are the asymptotic states (for $t \to +\infty$) defined by the X_w, i.e. the limit regimes.

Now, the analysis of these asymptotic states has proved to be unexpectedly and formidably difficult. The complexity of a general dynamical system is prodigious. Firstly, the ideal determinism — which is mathematical — does not in any way entail determinism in the physical sense of the term (in the sense of "predictability"). An initial condition can only be defined approximately. It is not represented by a point x_0 in M but by a small domain U that "thickens" x_0. For determinism to be physical, the trajectories emanating from U must form a tube that "thickens" the trajectory of γ emanating from x_0. This means that the trajectory γ is stable with respect to small perturbations of its initial condition. A physically deterministic dynamical system is therefore a dynamical system (by definition ideally deterministic) which has stable trajectories.

There is no reason why this should generally be the case. There are even dynamical systems (for example geodesic systems in Riemann manifolds with negative curvature) presenting the property that all their trajectories are unstable, and presenting it in a structurally stable way. As Vladimir Arnold observed ([1], p. 314–315, our translation):

> The possibility of structurally stable systems with complicated movements, each of which is in itself exponentially unstable, is one of the most important discoveries to be made in differential equation theory in recent years. [...] In the past, it was assumed that systems of generic differential equations could only contain simple, stable limit regimes: positions of equilibrium and cycles. If the system was more complicated (conservative, for example), it was accepted that under the effect of a weak modification in the equations (for example by taking into account small, non-conservative perturbations) the complicated movements "break down" into simple movements. Now we know that this is not the case, and that in the functional space of vector fields, there exist domains composed of fields where the phase curves [the trajectories] are more complex. The conclusions to be drawn from this affect a large number of phenomena in which deterministic objects have "stochastic" behaviour.

Physical indeterminism (chaos, chance, randomness, etc.) is therefore perfectly compatible with mathematical determinism. As Thom pointed out ([4], p. 124):

> What we call "laws of chance" are in fact no more than properties of the most general deterministic system.

Let us return to the specification of the general model. In terms of dynamical systems, the internal states of S are the attractors of X_w. The very tricky concept of *attractor* generalises the concept of stable equilibrium point. Intuitively, it is a stable asymptotic regime, a closed set A, X — invariant and indecomposable for these two properties (i.e. minimal) — which attracts (i.e. captures asymptotically) all the trajectories emanating from the points of one of its neighbourhoods. The largest neighbourhood of A having this property is called the *basin of attraction* of A, denoted $B(A)$. In simple cases, the attractors have a simple topological structure (attractive point or attractive cycle), they are finite in number and their basins of attraction are "good" domains (of simple form) separated by separatrices. But this description is too optimistic, because the attractors may be infinite in number, their basins of attraction may be inextricable intermingled, and the attractors may have a very complicated topology (strange attractors).

On an attractor, the trajectories of a dynamical system present *recurrence*. Intuitively, the recurrence of a trajectory γ means that if $x \in \gamma$, then γ passes arbitrarily close to x again after an arbitrarily long time and so γ returns infinitely often close to its initial position. The trivial cases of recurrence are the fixed points of X (the points where X equals zero, i.e. the trajectories reduced to a point point) and the cycles of X (closed trajectories). But there generally exists non-trivial recurrence. If γ is a complicated recurrent trajectory and A is its topological closure, then A is a whole domain of M (a non-empty closed set) where γ is dense.

Whatever we may make of these difficulties, Thom assumed in his morphodynamic models that for almost all initial conditions $x_0 \in M$ ("almost all" and not all because we must take into account the separatrices between basins), the trajectory emanating from x_0 is captured asymptotically by an attractor A_w of the internal dynamics X_w. This corresponds to a hypothesis of local equilibrium: the fast internal dynamics drives the system towards a stable asymptotic regime corresponding to an internal state.

Once these various hypotheses have been established, the general model becomes a mathematical programme: general structure of dynamical systems (qualitative dynamics or global analysis); geometric characterisation of structurally stable dynamical systems and their attractors; analysis of the ergodic properties of strange attractors; analysis of possible causes of instability; analysis of the deformations (perturbations) of structurally unstable systems; study of the geometry (which can be extremely complex) of catastrophic sets $K_{\mathscr{X}}$; etc. This programme, which we could call the Thom-Smale programme, is an extension of that of Poincaré and Birkhoff. It is in fact the programme of modern qualitative dynamics.

But if the Thom-Smale programme is of immense scope, it is also of immense difficulty. That is why Thom proposed to simplify it.

4 The theory of singularities and "elementary" morphogenetic models

As the complexity of general dynamical systems is too great to master, we can start by carrying out a rough study, of a thermodynamic nature. This consists in ignoring the fine structure (the complicated topology) of attractors. This step is all the more necessary since the empirical catastrophic sets K_W are usually much simpler than those induced by the bifurcations of general dynamical systems. The aim is therefore to understand how systems can be internally chaotic (stochasticity of the attractors defining internal states) and externally ordered (simplicity of observable morphologies).

The idea is to apply to general systems that which can be observed to occur in the case of gradient systems, namely the existence of gradient lines and level manifolds. To do so, we use the fact that, if A is an attractor of a dynamical system X on a manifold M, we can build, on the basin $B(A)$ of A, a positive function f (called a *Lyapunov function*) which decreases strictly along the trajectories in $B(A)-A$ and vanishes on A. This function is a sort of local entropy, expressing the fact that over time, $B(A)$ contracts on A analogously to a gradient system. But it does not allow us to say anything about the internal structure of the attractor.

The next step is to retain, out of all the bifurcations of the attractors, only those that are associated with their Lyapunov function. This reduction is similar to thermodynamic averaging. It corresponds to a change in the level of observation, from the fine level described by X_w to the rougher level described by f_w. It is analogous to the reduction that is performed in Landau's mean-field theory of phase transitions. Thom gives the following justification for it ([5], p. 521, our translation):

> Personally, I like to think that what plays a role, is not the concept — too fine — of attractor, but a class of equivalence of attractors, equivalent because encapsulated in the level manifold of a Lyapunov function (a quasi-potential), provided that the attractor avoids implosions of an exceptional nature. I believe that this may be the path to follow to find a satisfactory mathematical definition of the concept of stationary asymptotic regime of a dynamics. From this perspective, the fine internal structure of the attractor is of little importance: the only thing that matters is the Lyapunov function that encloses it in one of its level manifolds. But we can consider that only the structure of the tube enclosing the attractor is phenomenologically important, and we thus obtain a problem that is similar to elementary catastrophe theory.

"Elementary" morphodynamic models consist in applying quasi-potentials — the Lyapunov functions — to the gradient system derived from a potential. We assume that the internal dynamics X_w is in fact the gradient dynamics associated with a differentiable potential function $f_w : M \to \mathbf{R}$. The internal states determined by f_w are then its minima (if f is equated with an energy, this principle is that of the energy minimisation of the system). In Thomian terminology, this sort of system is called a *static model*. Mathematically, the theory of static models is therefore an integral part of the bifurcation theory of potential functions. Now, for potentials, there exists a simple characterisation of structural stability (*Morse theorem*). Under the hypothesis that the manifold M is compact, $f : M \to \mathbf{R}$ is stable if and only if:
(i) its critical points (i.e. its minima, maxima and saddle points) are non-degenerate, in other words they are not fusions of several minima, maxima or saddle points; and
(ii) its critical values (i.e. the values $f(x)$ for critical x) are distinct.

There are therefore two causes of structural instability:
(i) the degeneracy of critical points, corresponding to what are called *bifurcation* catastrophes;
(ii) the equality of two critical values, corresponding to what are called *conflict* catastrophes.

Each of these two very distinct types of catastrophe has a corresponding type of instance of selection I, which Thom called conventions:
(i) the *convention of perfect delay*, according to which the system S remains in an internal state (a minimum of f_w) as long as that state exists: there can only be catastrophe when a minimum disappears through fusion with another critical point (bifurcation);
(ii) the *Maxwell convention*, according to which the system S always occupies the absolute minimum of f_w: there can only be catastrophe when another minimum becomes the absolute minimum (conflict).

5 The principles of morphodynamic models

As we have seen, morphodynamic models receive a natural interpretation within the context of systems theory. In this setting, the space W is a control space and the phenomena we seek to account for are of the critical type. Most of the rigorous and accurate physical applications of morphodynamics are of this type: diffraction catastrophes and wave-front dislocations in wave optics (with their consequences for semi-classic approximation in quantum mechanics); the theory of phase transitions and phenomena of spontaneous symmetry breaking in ordered media; stability of defects in mesomorphic phases; bifurcation (buckling) of elastic structures; constrained differential equations, singular perturbations and chaotic solutions (feedback-induced chaos); theory of shock waves; analysis of singularities in variational systems; regime changes in hydrodynamics, chemical kinetics and thermodynamics (dissipative structures and spontaneous self-organisation of matter, etc.); strange attractors, deterministic chaos and routes to turbulence; etc.

In these rigorous and accurate applications, we know the internal dynamics of the system explicitly, one way or another. We therefore seek to derive mathematically the catastrophic sets K_W from our explicit knowledge of the fields $\sigma : W \to \mathscr{X}$ and, quite naturally, we postulate that the internal process X_w causally generates the external morphology K_W. Analysis of various physical examples shows that often, a “flesh” of fine-scale complex processes (renormalisation groups, oscillating integrals, etc.) is grafted onto a “skeleton” of medium-scale singularities. this allows us to speak with a certain degree of precision about the morphological infrastructures of certain classes of physical phenomena. Now, these infrastructures are phenomenologically dominant. We therefore possess - for the first time - a link between the mathematical formalisms of physical objectivity and the phenomenology of the manifestation.

6 The models of morphogenesis

The junction between physical models and morphological schemes is made by considering that the control space W is the spatio-temporal extension of a material substrate. Consequently, the models describe the qualitative variation of perceptible qualities that can be observed in the substrate. This is the case for the models proposed by Thom for embryogenesis. They are based on two guiding ideas.

The first is that the attractors of internal dynamics define local metabolic regimes on the substrate (whence the name of metabolic models) and that, since these regimes are controlled by the spatio-temporal extension of the substrate, their catastrophes (made elementary by thermodynamic averaging) manifest themselves as differentiations of qualities in the substrate, in other words as processes of morphogenesis.

The second idea, more speculative, is that it is possible to interpret the topology of attractors defining local regimes in terms of their functional significance within the global regulation of the organism.

References

[1] Arnold V. (1978) *Mathematical methods of classical mechanics*, Mir (Moscow).

[2] Thom R. (1980) *Modèles mathématiques de la morphogenèse*, Christian Bourgois (Paris), in French. English translation: *Mathematical models of morphogenesis* (1983), John Wiley and sons (Chichester).

[3] Thom R. (1972) *Stabilité structurelle et Morphogenèse*, Benjamin (New York), Édiscience (Paris), in French. English translation: *Structural stability and morphogenesis* (1975), Benjamin (Reading MA).

[4] Thom R. (1980) Halte au hasard, silence au bruit, *Le Débat* **3**, 119–132, Paris (Gallimard), in French. Commented english edition: Stop Chance! Silence Noise! *SubStance* **40** (1983).

[5] Thom R. (1990) *Apologie du Logos*, Hachette (Paris).

Collected works of René Thom are available on CD-Rom at the Institut des Hautes Études Scientifiques, 35 route de Chartres, 91440 Bures-sur-Yvette.