LAS ECUACIONES DIFERENCIALES COMO MODELOS DETERMINISTICOS EN CIENCIA E INGENIERIA*

48

por

DIEGO BRICIO HERNANDEZ

1987

* Articulo aceptado para su publicaciòn (en dos partes) en otros tantos nÚmeros consecutivos de la revista International Journal of Mathematical Education in Science and Technology (1987).

DIFFERENTIAL EQUATIONS AS DETERMINISTIC MODELS IN SCIENCE AND TECHNOLOGY. Part I: Modelling

by

DIEGO BRICIO HERNANDEZ*

Centro de Investigación en Matemáticas
CIMAT, Apartado Postal 402
36000 Guanajuato, Gto. (MEXICO).

* On leave from Departamento de Matemáticas, Universidad Autónoma Metropolitana-Iztapalapa, Apdo. 55-534, 09340 México D.F. (MEXICO).

ABSTRACT

This paper consists of two parts, appearing in consecutive issues of this Journal. It deals with some of the many aspects of modelling physical systems in terms of differential equations (DE), ordinary and partial. Among the issues touched upon figure a) the important role played by the assumptions underlying a DE based mathematical model (Part I), and b) the interplay between modelling and simulation when using differential equation models (Part II). In a more philosophical vein, there is a section on determinism as well as one on model validation; while fairly general in scope, the concepts examined therein are introduced in direct relation to DE models.

From an educational viewpoint, the present paper contains computational and methodological approaches which, suitably combined, should contribute to a more successful teaching of the subject to engineering and science students. Indeed, and for the sake of both understanding and motivation, prospective teachers of DE's should have a good command of these aspects.

CONTENTS.

1. INTRODUCTION.

This paper draws attention to certain principles governing the application of mathematics to engineering and the physical sciences. Broadly speaking, these applications require

a) producing mathematical descriptions of physical phenomena (MODELLING), and

b) drawing conclusions about those and related phenomena with the help of such models (SIMULATION).

Not surprisingly considering its stated goals, this paper is concerned with both modelling and simulation. Special emphasis is laid, however, upon relating mathematical constructions to their empirical counterparts, keeping faithful to the belief that knowledge is rooted in experiment. In short, the modelling process is viewed here as an instance of translation, from an informal description of a given real phenomena into a (formalized) mathematical construction, the model. This translation is

guided by various assumptions made by the modeller about the behaviour of the real world. The appropriateness of these assumptions is best tested against experimental results, by comparing some of their consequences with the manifestations of reality as observed by the experimenter.

There is a model corresponding to each set of assumptions, and drawing conclusions from a given model is either eased or rendered impossible depending on the set of assumptions it rests upon. No wonder, then, that modelling itself is intimately mingled with simulation, each in turn being related to model complexity in its own particular way. This remark opens up the possibility of achieving unified, systematic treatments of modelling and simulation based upon the idea of model complexity.

Unfortunately, this last notion does not easily lend itself to a precise mathematical treatment; a possibility of doing this, however, is to refer model complexity to algorithm complexity, which can be given a precise meaning, see for instance [1], [2], [3], also [4]. In terms of this notion, one can say that

THE COMPLEXITY OF A GIVEN MODEL IS THAT OF THE LEAST COMPLEX SIMULATION ALGORITHM BASED UPON THAT MODEL.

However this concept of model complexity is rendered operative, some remarks about systematizing modelling/simulation work can be found at the end of section 6 below. See [5], [6] for somewhat more detailed accounts.

In the last resort, however, all modelling and simulation work depends upon the expertise and good sense of the modeller; any systematic theory should be aimed at supplementing rather

than replacing the modeller's craft.

This being said, the reader must be warned that far less ambitious goals are aimed at in most of the paper. For instance, sections 4 and 6 deal with setting up and solving differential equations, respectively. The choice of modelling methodology adopted in section 4 owes a great deal to treatments now classical in the chemical engineering literature, such as [7] and [8]. As to section 6, it is a brief review of numerical methods for differential equations, ordinary and partial.

Determinism no doubt is a fundamental notion underlying all classical physics, hence also its engineering applications. A critical analysis of this concept was done in [9] in relation to modelling; section 3 is concerned with exploring further the experimental origin of this notion from a classical perspective. Determinism is seen to be formalizable into the notion of a state space, as understood in the mathematical system theoretic literature [10].

Together, determinism and the choice of continuous time or space representative set the ground for differential equation models; all extras required are physical laws expressed in differential form plus the necessary differentiability assumptions. The differential equations are ordinary when time is continuous and space is discrete, also when space is one dimensional and time is discrete; in all remaining cases there result partial differential equations. To keep matters within bounds, the main ideas developed in both sections 3 and 5 are illustrated in terms of the differential equation models developed in the elementary examples of section 2. Section 5 is concerned with model validation criteria, and relies heavily

upon concepts describing empirical work. It is seen therein that confrontation with experiment may be critical, and require a careful reexamination of the fundamental assumptions underlying the given model.

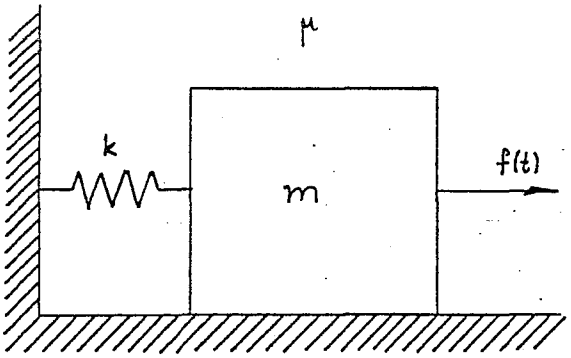
Finally, a few remarks on the numbering of equations, which is internal to each section. If equation (k) of section m is called from section n, it is referred to as eq. (k) if m=n, else as eq. (m.k).

2. SOME MODELLING EXAMPLES.

The following examples illustrate three situations in which dynamical systems are modelled from physical principles, with varying degrees of detail. Though rather elementary in character, they are reckoned to illustrate the various considerations involved in mathematical modelling.

Example 1.

"Consider a mechanical system formed by a block which slides along a horizontal plane while acted upon by an external force. Besides, the mass is attached to a fixed spring, the whole system being submerged in a material fluid. It is desired to model the ensuing motion of this mechanical system."



The assumptions usually made when dealing with this situation are:

A1. The block is a material particle of known mass m,

A2. The spring is linear, i.e. it obeys Hooke's Law, with a known force constant k,

A3. There is no friction between the block and the horizontal surface,

A4. The material fluid helps to dissipate mechanical energy by exerting a retarding force upon the block which varies linearly with its velocity (constant of proportionality $=\mu$),

A5. The external force varies with time as given by a known function $f:[0,\infty)\to\mathbb{R}$.

Let $x:[0,\infty)\to\mathbb{R}$ define the trajectory of the block, with $x(t)$:= position of the block at time t. At this point one would like to invoke

A6. Newton's Second Law,

which requires us to express all the forces as well as the acceleration of the particle in terms of the dynamical variables. It is therefore convenient to introduce some extra smoothness assumptions on x, say

A7. $x\in C^2$.

This new assumption allows us to give all the external forces acting on the particle at time t, namely

$-kx(t)$ [the spring force],

$-\mu\frac{dx}{dt}(t)$ [the viscous force],

$f(t)$ [the external action], plus

$m\frac{d^2x}{dt^2}$ [the inertial force].

Newton's Second Law makes us restrict our search for trajectories to those C^2 functions satisfying the differential equation

$$\frac{d^2x}{dt^2}=-kx(t)-\mu\frac{dx}{dt}+f(t), \tag{1a}$$

over $[0,\infty)$, where we have set $m=1$ for the sake of notational simplicity.

Now, it is quite easy to see that a solution $x\in C^2$ can exist only if

A8. f is continuous,

an assumption that will be readily adopted since we know that indeed there is a trajectory. Moreover, it is a well established experimental fact that fixing both position and velocity at the start of the motion does in fact determine the trajectory, a fact embodying the deterministic character of this analysis. We would like our mathematical description to be as faithful as possible, thus imposing upon the model the basic requirement of reflecting this fundamental fact. In turn, this leads us to supplementing

the DE (1a) by adding the initial condition

$$x(0)=x_0,\ \frac{dx}{dt}(0)=v_0, \tag{1b}$$

hoping thereby to determine one solution of the differential equation and thus reflect the determinism of the situation.

It is a well known fact in the theory of linear differential equations that the initial value problem (IVP) given by conditions (1) does have a unique solution defined on the whole nonnegative time axis. This validates the resulting model at least from the requirement of reflecting determinism, additional validation tests being required by the need to achieve agreement with quantitative experiments.●

Example 2.

"A long, thin, uniform metal bar is thermally insulated from its surroundings and has a nonuniform temperature distribution at a certain instant of time. It is desired to model thermal events within the bar; in particular, it is required to predict its future temperature distribution."

As in the previous example, it is imperative to start out by introducing appropriate working assumptions. The first point follows from the fact that the bar is "long and thin", which suggests the appropriateness of considering temperature variations in a longitudinal direction only. In other words, all the points in a given cross section of the bar will be assumed to have the same temperature.

This situation gives rise to the following assumption, in which L denotes the bar length measured in convenient units:

A1. The portion of space occupied by the bar is the closed interval [0,L].

Consequently, a temperature distribution (or temperature "profile") can be described by a function $T:[0,L]\to\mathbb{R}$, where T(z):= temperature at point z of the bar. The only interesting thermal events take place within the bar, in the form of heat conduction; for convenience, this heat transfer mechanism will be assumed to follow a linear law, namely

A2. Fourier's Law [11].

It is necessary to restrict our attention to "sufficiently smooth" temperature profiles if we want to at least express Fourier's Law in this setting. For the time being, we shall be content with stating the assumption that

A3. Each temperature profile is continuous on [0,L] and continuously differentiable on (0,L).

Thus, Fourier's Law of heat conduction states that the rate of heat flow per unit area ("heat flux") along the bar equals

$$-k\frac{dT}{dz}$$

if computed at point z when the temperature profile is T(·). Here k denotes the local heat conductivity.

Therefore, the fact that the metal bar is thermally insulated at the ends translates into the assumption that

A4. Every temperature profile satisfies $T'(0)=T'(L)=0$.

It is an experimental fact that each temperature profile evolves in one and only one way. In other words, we are dealing with a deterministic situation, same as in the previous example. Therefore, each initial temperature profile describes a unique trajectory, which can be most conveniently represented in terms of a function $u:[0,\infty)\times[0,L]\to\mathbb{R}$, with $u(t,\cdot)$ denoting the temperature profile of the bar at time t. Thus the following boundary condition (BC) must hold:

$$\frac{\partial u}{\partial z}(\cdot,0)=0,\ \frac{\partial u}{\partial z}(\cdot,L)=0, \tag{2}$$

together with the initial condition (IC)

$$u(0,\cdot)=f, \tag{3}$$

where f stands for the known non-uniform initial profile.

Because temperature varies in time, thermal energy accumulates at every portion of the bar. From the very definition of heat capacity, the rate of energy accumulation between z_0-h and z_0+h at any instant t is measured by

$$\frac{d}{dt}\left\{\int_{z_o-h}^{z_o+h}\rho C_p u(t,z)\,dz\right\},$$

provided the time derivative exists, where ρ denotes the local linear density of the material.

In turn, the "uniformity" of the bar can be translated into the assumption that

A5. Density, heat capacity and thermal conductivity have constant values, both in time and along the bar.

At this point it is convenient to introduce the purely mathematical assumption that

A6. $\frac{\partial u}{\partial t}$ is continuous,

to obtain the rate of energy accumulation in the section of the bar under analysis in the form of

$$\int_{z_o-h}^{z_o+h}\rho C_p\frac{\partial u}{\partial t}dz.$$

On the other hand, the net heat flow through the element is

given by Fourier's Law as

$$-k\frac{\partial u}{\partial z}(t,z_o-h)-[-k\frac{\partial u}{\partial z}(t,z_o+h)],$$

so that

$$\int_{z_o-h}^{z_o+h}\rho C_p\frac{\partial u}{\partial t}dz=-k\left[\frac{\partial u}{\partial z}(t,z_o-h)\ -\ \frac{\partial u}{\partial z}(t,z_o+h)\right].$$

This last equality merely expresses the validity of

A7. The Law of Conservation of Energy,

which we take for granted here.

At this point it is convenient to even further restrict the class of allowable profiles by introducing the additional assumption that

A8. $u(t,\cdot)\in C^2$ for each $t>0$.

Having done so, it suffices to divide both sides of the previous equality by h and let $h\to 0$ to obtain the DE

$$\frac{\partial u}{\partial t}=\alpha\frac{\partial^2 u}{\partial z^2},\quad t>0,\quad 0<z<L, \tag{4}$$

a necessary condition for the validity of assumptions A1 to A8. Here $\alpha := k/\rho Cp$ denotes the thermal diffusivity of the material.

It is a well known fact in the elementary theory of partial differential equations that the initial boundary value problem (IBVP) defined by (2), (3) and (4) admits a unique solution. It is true, then, that this model does reflect the determinism encountered in the physical situation.■

Example 3.

"A thin, uniform membrane is tightened by forcing its edge to adopt the shape of a nonplanar simple closed curve. Its constituent material is very light and is subject to no other external actions, all internal tensions being uniformly distributed over the membrane. It is required to describe the resulting shape."

This situation can be analyzed along lines very similar to those of the two foregoing examples. A number of hypotheses will have to be put forward in order to carry out that analysis: some of them are of a geometrical nature, some others concern the physics of the situation and the rest are purely mathematical.

The basic treatment is rather well known and can be found in many standard references on mechanics of a continuum [12] or in mathematical treatises like [13]. For this reason this example shall be dealt with in considerably less detail than the previous two.

Broadly speaking, the analysis is normally based upon the following set of assumptions:

A1. The membrane projects onto a plane region identifiable with a closed, bounded region $G \subset \mathbb{R}^2$.

A2. Every perpendicular to G intersects the membrane only once.

A3 The membrane is so thin that such intersection consists of a single point.

From these last two assumptions it follows that the shape of the tightened membrane is the graph of a function $u: G \to \mathbb{R}$, with $u(x,y) :=$ deformation amplitude at $(x,y) \in G$. For mathematical convenience, let's assume that

A4. u is twice continuously differentiable in the interior of G,

A5. The restriction of u to the boundary of G is a known continuous function, say $f: \partial G \to \mathbb{R}$.

Given that the membrane material is very light, it makes sense to neglect the action of gravity. To be precise, we shall asssume that

A6. The only external action on each element of membrane is the tension exerted upon it by the neighboring elements.

Upon letting $a,b \to 0$, the condition for static equilibrium of that portion of the membrane directly above the square $[x_0, x_0+a] \times [y_0, y_0+b] \subset \mathrm{Int}(G)$ transforms into the relation

$$(\sigma u_x)_x + (\sigma u_y)_y = 0,$$

with

$$\sigma := \left[1+(u_x)^2+(u_y)^2\right]^{-1/2}. \tag{5}$$

where T stands for the local tension in the membrane.

If we now introduce the assumption that

A7. The tension is uniformly distributed over the membrane (i.e. T is constant),

plus the additional assumption that

A8. The maximum deformation is small compared with the total extension of the membrane,

then this last condition reduces to

$$\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}=0 \text{ in } G, \tag{6a}$$

which must be supplemented by the BC

$$u\big|_{\partial G}=f. \tag{6b}$$

This boundary value problem (BVP) for Laplace's equation must be satisfied by u if all the previously stated assumptions are to hold. Well known results of Classical Potential Theory, (see e.g. [14]) guarantee that u is fully determined by the various assumptions underlying this mathematical model, thus validating it at least from a qualitative point of view. A

slight extension of this construction yields an IBVP characterizing the dynamic behaviour of this membrane under deterministic conditions. See [13].•

3. DETERMINISM IN MATHEMATICAL MODELS.

The foregoing examples illustrate the construction of mathematical models . In each case a phenomenon was chosen for study, and a relevant model was built through a process akin to translation. Most importantly, this translation process was accomplished by introducing a set of assumptions, of which a most consequential one implied adopting a deterministic viewpoint. Let us now view these concepts from a more general perspective.

For, let us begin by adopting the classical standpoint that physical phenomena refer to changes of observable attributes which take place in space and time. Typical attributes of interest are temperature, position, velocity, electric field, color, concentration, etc. In any case, a mathematical description of a phenomenon must include representatives of these three ingredients, namely TIME, SPACE and ATTRIBUTES. In each case, the mathematical representative chosen must certainly be a mathematical object, i. e. either a set or a concept derived therefrom.

Take TIME, for instance, to be represented by a nonempty set T. Whatever our choice, T must be ordered if we are to distinguish between past, present and future. It must not have "voids" either, hence selecting a set of real numbers for T seems appropriate; its elements will be referred to as instants. We often formulate physical laws in terms of instantaneous rates of change, which in turn must figure in our descriptions; in such cases the choice of an interval for T seems mandatory,

in order to take limits with ease. In our examples and in all that follows, $T=[0,\infty)$: $\{0\}$ stands for the present while $(0,\infty)$ represents the future, many situations requiring "to predict the future in terms of whichever present information is available".

As to the space representative (a nonempty set S), two typical choices come to mind: a finite set, as in Example 1 and a set $S\subset\mathbb{R}^d$ $(d=1,2,3)$ with a nonempty interior there, as in Examples 2 and 3. Following the now standard terminology, the first choice gives rise to lumped parameter models; the second one, to distributed parameter models. In more classical terms, we deal with particle mechanics and with mechanics of a continuum, respectively. The actual choice of S results from a number of factors, among which are

* symmetry and other geometrical considerations,

and often is dictated by the very practical need of keeping models "as simple as possible".

Thus, in Example 2 the choice of a cylinder of height L and radius R for S does seem natural. Symmetry considerations usually made (e.g. temperature is independent of orientation within any cross section of the cylinder) lead to considering only two space variables . If, moreover, the material is assumed to be isotropic (heat conduction is independent of orientation within the cylinder), then the resulting model would consist of the DE

$$\frac{\partial u}{\partial t} = \alpha\left(\frac{\partial^2 u}{\partial z^2} + \frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r}\right); \quad t\in(0,\infty), \quad 0<z<L, \quad 0<r<R$$

supplemented by the BC's

$$\frac{\partial u}{\partial r}(t,z,0)=0,\ \frac{\partial u}{\partial r}(t,z,R)=0;\ t\in[0,\infty),\ z\in[0,L],$$

$$\frac{\partial u}{\partial z}(t,0,r)=0,\ \frac{\partial u}{\partial z}(t,L,r)=0;\ t\in[0,\infty),\ r\in[0,R].$$

Should any of these simplifying assumptions be dropped, an even more complex model would result. There certainly are situations in which it is mandatory to consider all sort of features in a model, but it is nevertheless advisable to simplify whenever it is possible to do so without losing precision. In any case, experiment and the knowledge derived therefrom remain valuable indicators of just what simplifying assumptions it is apt to consider.

As to attributes, typically only finitely many will be considered (say n), and its representative set A is often chosen to be a subset of $\mathbb{R}^n$ for reasons akin to those leading to the current choice of T. There are instances, however, in which attributes are considered which vary in a discrete fashion (consider, for instance, a phenomenon involving the number of alpha particles detected by a Geiger counter); in such cases, A must be factorizable into its "discrete" and "continuous" factors. These situations are out of the scope of this presentation, though, and A will in fact be assumed to have a nonempty interior in $\mathbb{R}^n$.

Whatever our choices of T, S and A, our description of the phenomenon must provide the following information:

a) At every instant, the spatial distribution of attributes or attribute profile, given by a map e:S→A; let E denote the set of all admissible profiles.

b) The time-space description of the phenomenon, given by the set Ω consisting of all possible trajectories ω:T->E.

The three examples given in section 1 illustrate the type of considerations the modeller has to make in order to choose both E and Ω. Thus, in Example 2 the admissible profiles were real, C^1 functions on [0,L] whose derivatives vanished at each end of the interval and which were indeed C^2 functions on (0,L). As the reader may recall,

* the differentiability requirement was dictated by the need to express heat flux at every point of the bar, while

* the vanishing of the endpoint derivatives expresses thermal insulation;

* the existence of the second derivative at interior points was required in order to measure the local energy accumulation due to heat conduction,

* the continuity of the second derivative there was demanded by the need to interchange derivative and integral and thus express the energy balance.

Finally, the choice of Ω was dictated precisely by the fact that the law of energy conservation must hold, thus ruling out of consideration any trajectory not satisfying the heat equation.

Analogous considerations can be made concerning the choice of the various elements making up the models arrived at in

Examples 1 and 3. For instance, Ω in Example 3 consists of all constant trajectories satisfying the wave equation in two dimensions, with $T=[0,\infty)$.

There is, however, a feature common to all three examples and which so far has remained untouched in this discussion: DETERMINISM. On the other hand, there also is a feature that distinguishes the model in Example 3 from the previous two: time is apparently absent from it. The following remarks touch upon these issues and their origin in experiment.

It is a fundamental assumption in all this that time, space and attributes can be measured, and thereby the observer becomes aware of the phenomenon. Experiment is a source of data, often finitely many in number, in which attribute values are recorded at various instants and points in space. For instance,

* a video camera can be attached to the mechanical system of Example 1 in order to record the motion: in turn, the experimenter can monitor the block's position as a function of time from the resulting shots;

* thermocouples can be inserted at various positions within the metal bar, thus providing continuous monitoring of temperature there, and

* an arrangement similar to that devised for Example 1 can be used to monitor the position of selected points of the membrane.

In addition, any experiment assumes certain fixed "external conditions", which can be as disparate as ambient temperature and pressure, initial position and velocity, initial temperature

profile, etc. This is a very important part of the experimental information, and it provides a full characterization for the experiment. It is so important a piece of information that failing to provide it may lead to serious inconsistencies when mixing data belonging to different experiments.

This being said, the possible experimental conditions will be modelled by the elements of a nonempty set X, this set being referred to as an experimental frame [15]; given an experimental frame X, the data constitute a set $D \subset X \times T \times S \times A$, which often contains all information guiding the analyst in the formulation of simplifying assumptions and the like.

For instance, a phenomenon occurs in steady state under given experimental conditions if "measured attribute values remain constant in time"; this notion can be formalized by observing that it coincides with

$$(x,t',p,a'),(x,t'',p,a'') \in D \Rightarrow a'=a'' \text{ for some } x \in X.$$

More fundamentally, a phenomenon will be said to be deterministic if there is a choice of experimental frame for which the trajectory is determined given each particular set of conditions:

$$\text{For some } X,\ (x,t,p,a'),(x,t,p,a'') \in D \Rightarrow a'=a'' \text{ for every } x \in X.$$

Analogous definitions apply to symmetry considerations and the like.

An important clarification is in order, though, concerning the meaning of "=" in statements like a'=a" above. This is so because of the acknowledged variability of experimental results,

which rules out any contention of ever achieving equality in the mathematical sense. In short, one should rather resort to a notion like "indistinguishability under the conditions of the experiment"; in turn, such vague notion could be made precise if understood to mean

$$|a'-a''|<\varepsilon,$$

with >0 a given tolerance. Indeed a valid alternative could be

$$\frac{|a'-a''|}{|\alpha|}<\varepsilon,$$

where α stands for a suitable average of

$$\{a\in A: (x,t,p,a)\in D\},$$

for each fixed (x,t,p).

The important fact remains that, under determinism, there is an experimental frame that determines a trajectory for every set of conditions; such an experimental frame will be called a STATE SPACE for the corresponding physical system. For instance, a state space for the mechanical system of Example 1 is the set of all ordered pairs (position, velocity); a state space for the bar of Example 2 is the set of all temperature profiles, etc. Thus, the existence of a state space is at the root of determinism.

Summing up, a deterministic model is an ordered collection

$$<T,S,A,E,\Omega,X,f>,$$

in which the first five objects have the meaning already explained, X is a state space and $f:X\to\Omega$ assigns a unique trajectory to every value of the state.

Differential equations provide an adequate framework for the study of determinism (in continuous time) in terms of conditions for the existence and uniqueness of solutions to IVP's.

4. ON SETTING UP DIFFERENTIAL EQUATIONS

A short reflection on the three examples of section 2 suffices to detect a great many similarities between them. Many of these common features are of a structural nature and they have already been considered in section 3, but there is also a number of methodological similarities which we want to focus on in this section.

Let us begin by pointing out that each of these considers a quantity that is conserved: linear momentum in Examples 1 and 3, Energy in Example 2. In each case, the relevant conservation law was expressed in the basic form

$$\text{ACCUMULATION RATE} = \tag{1}$$

$$\text{INPUT RATE} - \text{OUTPUT RATE},$$

which when expressed locally gives rise to a differential equation.

In fact, it seldom occurs that only one source contributes to the build-up of whatever is conserved; also, there may be more than one sink. Thus, in general both input and output terms consist of sums of individual items. Take, for instance Example

1, in which two sinks of linear momentum can be identified, namely the spring and the viscous medium.

In general, it will be necessary to express the local values of the dynamical variable being conserved in terms of the system attributes there as well as in neighbouring points, and perhaps their time derivatives. In turn, the various input and output terms must be expressed in terms of the same quantities by recourse to physical laws. Then an "elementary portion" of the system is selected and the basic conservation equation is written down for it.

In the lumped parameter case S is a finite set, (say $S=\{p_1,\ldots,p_m\}$), and the natural choices for elementary sets are $\{p_1\},\ldots,\{p_m\}$. Thus the above operation will be repeated a total of m times and a system of m ordinary differential equations (ODE) will result.

In contrast, in the distributed parameter case (say $S\subset R^d$, with a nonempty interior there), for every interior point $p\in S$ there is a positive number r_o such that the open ball of radius r_o centered in p [denoted $V(p,r_o)$] consists of interior points of S. Take the sets in the collection

$$\{V(p,r):p\in Int(S),0<r<r_o\}$$

to be elementary in this case. The basic conservation equation will be written down for each such $V(p,r)$, and a differential relation valid at p will result upon letting $r\to 0$. The arbitrariness of p implies then that such relation holds in the whole interior of S and hence results in a partial differential equation (PDE) valid there.

Besides supplying the all important conservation principle, Physics enters into the above by providing expressions for

a) the basic variable being conserved, in terms of the attributes and their derivatives,

b) each input and output term, as a function of the same variables [the "constitutive equations", Bird, op. cit.].

Item a) is usually "forced" upon us by Physics itself and there is not much we can do about it: the mechanical energy of the block in Example 1 is given by

$$\frac{1}{2}kx^2+\frac{1}{2}m\left(\frac{dx}{dt}\right)^2,$$

its linear momentum by

$$m\frac{dx}{dt},$$

etc. These are the very definitions of the concepts we need and we have no choice but to take them as given. On the contrary, there usually is a whole spectrum of physical laws to choose from when it comes to selecting a constitutive equation. The choice is dictated by our knowledge of the reliability of each such physical law over the whole range of variation of the attributes and their derivatives, but no less so by the basic "practical" constraint of avoiding unnecessarily complex models.

Take, for instance, the constitutive equations for the

viscous and spring forces acting on the block of Example 1. Both intuition and physical measurements suggest that viscous action is absent if the block is not moving; they also suggest that the corresponding force is a function of velocity, directed in the oposite direction. Similar remarks apply to the spring restoring action as related to deformation.

Hence it is natural to expect constitutive equations such as

$$\text{spring force}=\varphi(x),$$

$$\text{viscous force}=\psi\left(\frac{dx}{dt}\right),$$

with φ and ψ satisfying

$$\varphi(0)=0;\ x\varphi(x)<0 \text{ if } x\neq 0 \tag{2a}$$

$$\psi(0)=0;\ v\psi(v)<0 \text{ if } v\neq 0 \tag{2b}$$

In this more general setting, instead of equation (1.1a) one obtains

$$\frac{d^2x}{dt^2}=\varphi(x)+\psi\left(\frac{dx}{dt}\right)+f(t), \tag{2c}$$

which is nonlinear in general. Moreover, if the corresponding model is to reflect the underlying determinism, existence and uniqueness of solutions to the IVP associated with this DE must be secured. This imposes stringent requirements upon φ and ψ: for instance

allowed by both conservation laws and constitutive equations is

* both φ and ψ are continuously differentiable,

although milder smoothness conditions do indeed suffice. In practice, such functions result from experiment and are often given in graphical form, labelled the "characteristic response curve" of the corresponding element. It can be safely said, then, that these additional smoothness assumptions arise from purely mathematical considerations and are not supplied by the physics of the situation. Their importance should not be minimized, though, given their function in the model.

What we did in Example 1 was to use our knowledge about φ and ψ to replace $\varphi(x)$ and $\psi(dx/dt)$ by $-kx$ and $-\mu dx/dt$, respectively, with $k:=-\varphi'(0)$, $\mu:=-\psi'(0)$, two positive numbers. The resulting simplification procedure is nothing but linearization, a welcome byproduct of differentiability.

Quite analogous remarks can be made concerning our use of Fourier's Law in Example 2, which implies the constitutive equation

$$\text{heat flux}=-k\frac{\partial u}{\partial z};$$

in general one should expect

$$\text{heat flux}=\varphi\left(\frac{\partial u}{\partial z}\right),$$

with φ very much like in (2a). A more refined expression would in fact be

$$\text{heat flux} = F\left(u, \frac{\partial u}{\partial z}\right),$$

with $F(u,\cdot)$ as φ in (2a); it would yield the nonlinear PDE

$$\rho C_p \frac{\partial u}{\partial t} = \frac{\partial}{\partial z} F\left(u, \frac{\partial u}{\partial z}\right), \tag{4}$$

instead of (1.4). It was through the same linearization process based upon assuming F to be differentiable and $\partial F/\partial u$ to be negligible that the linear PDE in Example 2 was derived.

It is difficult to make apparent and even more so to render precise the criteria guiding the modeller (scientist, engineer or whatever) in making the right choice of constitutive equation. They ultimately rely on the very basic requirement of "making model predictions agree with experiment", of course, but the task is never accomplished via a blind "search in model space". On the contrary, those guiding criteria constitute a good part of the modeller's expertise , whose build-up takes years of higher learning and subsequent practice.

Often, however, a scientist gets involved in work of a really innovative nature, for which earlier experiences and training do not suffice. In such cases further experimentation is mandatory in order to test the model resulting from a given set of assumptions. We shall be coming back to these considerations in the following section.

In the meantime, let us insist on the all important role

played by both conservation laws and constitutive equations in setting up differential equations describing physical phenomena. These two elements do not have the same rank, though, and the choice of conservation law seems to be the more fundamental one; once that choice has been made, however, the choice of constitutive equations determines a differential equation.

We shall conclude by remarking that these choices are far from exhausting all the possible ways of influencing the outcome when setting up a differential equation. To illustrate, mention must be made of our "last minute assumption" that T was constant in Example 3, one that resulted in the (linear) Dirichlet problem for Laplace's equation. In general, tension can be expected to depend on both deformation and rate of strain, say

$$T=T(u,u_x,u_y);$$

the BVP to solve in this more general case would be

$$(Tu_x)_x+(Tu_y)_y=0, \tag{5a}$$

$$u|_{\partial G}=f \tag{5b}$$

instead of (2.6). Moreover, dropping the assumption also made that "the maximum deformation is small compared with the total extension of the membrane" would result in the even more complex nonlinear BVP

$$(\sigma u_x)_x+(\sigma u_y)_y=0, \tag{6a}$$

$$u|_{\partial G}=f, \tag{6b}$$

where σ has been defined in (2.5).

The reader should reflect upon the importance of introducing the right assumptions when formulating a DE based model: too relaxed assumptions result in defective models; too stringent ones give rise to unmanageable models. The right balance lies somewhere in the middle, as one should expect.

PARTIAL CONCLUSIONS.

The main goal of this first part has been to bring to focus the very important role played by the assumptions (implicit or otherwise) underlying a given mathematical model. This has been done in the specific context of differential equation based models, but the corresponding concepts remain applicable to any other class of mathematical models.

In particular, it has been found that DE models can be characterised in term of the assumptions of

a) determinism and
b) continuous time,

plus the more specific ones concerning

c) the physical laws governing the phenomenon.

In addition, the differential equations will be ordinary if the space is assumed to be discrete while the time representation is continuous, or if space is assumed to be one dimensional if only the steady state is to be described. In every other case the resulting DE's will be partial.

A methodology has been given for the construction of DE based mathematical models, following the established procedure in the physical and engineering sciences. It is based on the idea of coupling a choice of conservation law with appropriate choices of constitutive equations specifying the various sources (energy, matter, momentum, electrical charge or whatever) with the appropriate attributes (temperature, concentration, velocity, electric potential, etc.).

The choice of physical law in c) is most consequential with regards to the complexity of the resulting model. In this connection, a simulation based methodology is outlined in Part II, whereby the "right" model complexity can be obtained: a model should be complex enough that no significant features of the real system are overlooked, while still allowing for model manipulations with the computational tools available.

BIBLIOGRAPHY.

[1] CHAITIN, G.J.,"On the length of programs for computing finite binary sequences". JACM 16(1969) 145-159.

[2] CHAITIN, G.J., "Information theoretic limitations of formal systems". Symposium on Computational Complexity, Courant Institute, 1971.

[3] CHAITIN, G.J.,"A theory of program size formally identical to information theory". JACM 22 (1975), 329-340.

[4] TRAUB, J.F. & H. WOZNIAKOWSKI,"A general theory of optimal algorithms". Academic Press, New York,,1980.

[5] HERNANDEZ, D.B.,"On partial discretization and preservation relations", Math. Modelling (1983),223-244.

[6] HERNANDEZ, D.B.,"Structure simplification of models", to appear in the International Encyclopedia on Systems and Control (M. Singh, ed.), Pergamon Press, Oxford, to appear.

[7] HIMMELBLAU, D. & K.B. BISCHOFF, "Process Analysis and Simulation". John Wiley and Sons, New York, 1968.

[8] ARIS, R., "Mathematical modelling techniques". Pitman, London, 1978.

[9] HERNANDEZ, D. B., "On some guiding principles in mathematical modelling". Math. Modelling 2 (1981),179-190.

[10] KALMAN, R., P.L. FALB & M.A. ARBIB, "Topics in Mathematical System Theory", Mac Graw Hill, New York, 1969.

[11] BIRD, R.B., W.E. STEWART & E.N. LIGHTFOOT,"Transport Phenomena". John Wiley & Sons, New York, 1961.

[12] SEGEL, L.A., "Mathematics applied to continuum Mechanics". Macmillan, New York, 1977.

[13] COURANT, R. & D. HILBERT,"Methods of mathematical physics" (2 vols.), Interscience Publishers, New York,1962.

[14] KELLOGG, O.D., "Foundations of Potential Theory". Springer, Berlin, 1929 (reprinted by Dover, Inc.).

[15] ZEIGLER, B.P., "Theory of modelling and simulation", John Wiley and Sons, New York,1976.

DIFFERENTIAL EQUATIONS AS DETERMINISTIC MODELS IN SCIENCE AND TECHNOLOGY. Part II: Simulation

by

DIEGO BRICIO HERNANDEZ*

Centro de Investigación en Matemáticas
CIMAT, Apartado Postal 402
36000 Guanajuato, Gto. (MEXICO).

* On leave from Departamento de Matemáticas, Universidad Autónoma Metropolitana-Iztapalapa, Apdo. 55-534, 09340 México D.F. (MEXICO).

ABSTRACT

This paper consists of two parts, the first of which appeared in the previous issue of this Journal. It deals with some of the many aspects of modelling physical systems in terms of differential equations (DE), ordinary and partial. Among the issues touched upon figure a) the important role played by the assumptions underlying a DE based mathematical model (Part I), and b) the interplay between modelling and simulation when using differential equation models (Part II). In a more philosophical vein, there is a section on determinism as well as one on model validation; while fairly general in scope, the concepts examined therein are introduced in direct relation to DE models.

From an educational viewpoint, the present paper contains computational and methodological approaches which, suitably combined, should contribute to a more successful teaching of the subject to engineering and science students. Indeed, and for the sake of both understanding and motivation, prospective teachers of DE's should have a good command of these aspects.

CONTENTS.

SUMMARY OF PART I.

The main goal of the first part was to bring to focus the very important role played by the assumptions (implicit or otherwise) underlying a given mathematical model. This was done in the specific context of differential equation based models, but the corresponding concepts remain applicable to any other class of mathematical models.

In particular, it was found that DE models can be characterised in terms of the assumptions of

a) determinism, and
b) continuous time,

plus the more specific ones concerning

c) the physical laws governing the phenomenon.

In addition, the differential equations are ordinary if the space is assumed to be discrete while the time representation is

continuous, or if space is assumed to be one dimensional if only the steady state is to be described. In every other case the resulting DE's are partial.

In addition, a methodology was given for the construction of DE based mathematical models, following the established procedure in the physical and engineering sciences. Such methodology is based on the idea of coupling a choice of conservation law with appropriate choices of constitutive equations specifying the various sources (energy, matter, momentum, electrical charge or whatever) with the appropriate attributes (temperature, concentration, velocity, electric potential, etc.).

It was pointed out there that the choice of physical law in c) is most consequential with regards to the complexity of the resulting model. In this connection, a simulation based methodology is outlined in this second part, whereby the "right" model complexity can be obtained: a model should be complex enough that no significant features of the real system are overlooked, while still allowing for model manipulations with the computational tools available.

The list of references given at the end has been numbered consecutively from the last item in the corresponding list provided in Part I. Any reference to item [i], with i<16, points to the first part of the Bibliography.

5. MODEL VALIDATION.

In rather coarse terms, model validation refers to making sure that model predictions agree with actual observations, hence recurring to experiment is conceptually mandatory in order to validate a given model. Situations arise, however, in which the

modeller does not actually carry out any new experiments to that effect, relying instead on his/her own or other people's accumulated expertise in order to judge on the goodness of such model. In any case, experience has the final word in these matters.

To be precise, one may adopt the preferential ordering given in [16] for scientific theories and adapt it to ordering models. Then, given two models M_1 and M_2, it is said that M_1 is preferrable to M_2 if (M_1,M_2) satisfies at least one of the following criteria, while (M_2,M_1) satisfies none of the remaining ones. In the following statements, (M,N) is an ordered pair of models, and it is assumed that if one of them can be applied to a given set of data then so can the other:

i) The scope of M is greater than the scope of N, i.e. when considered as formal systems, M has more interpretations than N.

ii) M has greater accuracy of prediction than N, and the predictions are not error free.

iii) M is less complex than N.

In this context, and from an epistemological viewpoint, it does make sense to give greater weight to point ii) , and regard iii) as a (most important) requirement of a practical nature. That is, provided the empirical contexts we are concerned with do lie within the scope of both models. Accordingly, this section and the next will be centered upon requirements ii) and iii), respectively.

Now, it does make sense to ask ourselves just what we mean

by "predictions agreeing with observations". For purposes of illustration, consider the experimental arrangement suggested in section 4 for monitoring the motion of the block in Example 1. Experiment yields data which can be presented as

$$(t_1,x_1),\ldots,(t_N,x_N), \tag{1}$$

i. e. a table of instants and corresponding measured positions, assuming a common underlying initial state.

On the other hand, the model predicts a set of values

$$x(t_1),\ldots,x(t_N) \tag{2}$$

for that particular initial state, hence full agreement is achieved if and only if

$$x(t_i)=x_i, i=1,\ldots,N, \tag{3}$$

whatever the choice of a positive integer N and of time instants $t_1,\ldots,t_N$. In other words, we are asking for a perfect match between theoretical predictions and actual observations for all possible realizations of all conceivable experiments in order to proclaim the model "good".

Now, this may be highly desirable and logically sound, but it is also highly unrealistic. In the first place, it must be observed that only finitely many sets of data such as (1) can ever be collected, so that these stringent requirements cannot even be verified in practice. Second, it frequently happens that r such tables differ in their second row even if

a) corresponding pairs have the same time components, and

b) the data were obtained for identical initial state.

This last statement is rather serious, insofar as (if taken at face value) it bluntly contradicts the determinism that lead to the model in the first place. This difficulty resolves, however, if one recalls that "=" in the operational definition of determinism actually meant "indistinguishability under the conditions of the experiment" and we ought to keep this meaning in (3) in order to preserve consistency.

This being said, one still has to cope with the practical limitation of never being able to perform "all conceivable experiments". Even if data accumulate and the collected information actually grows without bounds, it is still finite at every instant.

A situation results in which the goodness of the model can only be tested up to the available information, giving up all claims of ever achieving definitiveness when giving a positive evaluation. In strict terms, one cannot rule out the possibility of ever coming across new experimental evidence forcing us to reject a formerly acceptable model. Rejection, on the other hand, is definitive.

In real life, model validation may take a slightly different turn, particularly in those situations where a reasonable knowledge of the physical system is available. In engineering work, for instance, it is often known that a constitutive equation must have a certain form, in which only finitely many parameters remain to be specified; in such cases, modelling requires estimating those parameters from experimental data.

To be specific, let us illustrate this in terms of Example 1. Suppose it is known that dissipation is negligible ($\mu=0$), while

$$\text{spring force}=-k\,\text{sign}(x)\,|x|^{\theta},$$

where θ is a real, positive unknown parameter.

We are actually dealing with a whole family of models, one for each value of the parameter. The question arises of choosing the best such model.

Observe that each model in the family produces one and only one trajectory for eac choice of initial state

$$s:=\left(x(0),\frac{dx}{dt}(0)\right),$$

thereby giving rise to a family

$$\{x(t,\theta,s):\theta>0,s\in\mathbb{R}^2\}$$

of predicted values for the block's position at time t. This in turn results in an overall measure of discrepancy with the data which depends on the parameter value as well as on the initial state s. In general, any measure of overall discrepancy with the data is a function of two variables, say $Q:(0,\infty)\times\mathbb{R}^2\to\mathbb{R}$, such that

$Q(\theta,s):=$ overall discrepancy with data corresponding to initial state s and parameter value θ.

A frequent choice for Q is

$$Q(\theta,s) := \sum_{i=1}^{N} [x(t_i,\theta,s)-x_i(s)]^2,$$

but some other definitions of Q do occur in the engineering literature [17],[18].

It is the modeller's responsibility to select the best model in the family (i.e. the best parameter value). The obvious rationale for this purpose requires $Q(\theta,s)$ to become as small as possible for each value of s actually represented in the data. For notational convenience, let $s_1,\ldots,s_m$ be precisely those values of the state. This rationale can be implemented as follows:

For each $j\in\{1,\ldots,m\}$, choose θ_j so that

$$Q(\theta_j,s_j)=\min\{Q(\theta,s_j):\theta>0\}$$

then choose the "best" parameter value by letting

$$\theta^* := \theta_k,$$

where k is such that $Q(\theta_j,s_j)$ achieves its minimum value (say Q_{min}) for j=k.

There still remains a fundamental question to be asked, namely

* Is best good enough?

By this we mean, is $Q_{min}<\varepsilon$, the prespecified tolerance? If

it is, fair enough. Let's now concern ourselves with what can be done when "best is not good enough".

For that purpose, one must remember that a particular model is arrived at as a result of a rather long series of decisions: choices had to be made on

* the dimensionality of the space representation,

* whether time was to be discrete or continuous,

* which physical phenomena were to be considered in the description and which were to be ignored

* whether determinism should be assumed or not,

* the conservation principle to be invoked,

* the constitutive equations to be chosen, etc.

It is only natural that a failure to meet the validation specification must lead to a reexamination of all this decision making.

In normal modelling practice, it often occurs that failure to validate the model results from a decision to disregard a certain aspect of the phenomenon, so that taking it into account in a new model solves the problem. As a typical illustration of this, consider Example 2. An important ingredient in the corresponding model was the assumption that

* the bar was insulated along its length,

and not only at its two ends. A more cautious approach would consider the possibility of actually having heat losses along the sides of the bar, say to the surrounding medium. For the sake of simplicity, assume this medium is bulky enough as to maintain a constant temperature in spite of receiving heat from the bar. Assuming heat losses occur through natural convection only, the resulting model would be of the form

$$\frac{\partial u}{\partial t} = \alpha\frac{\partial^2 u}{\partial z^2} - \beta u; \quad t>0, \quad 0<z<L$$

$$u_z(t,0)=0, \quad u_z(t,L)=0; \quad t\in[0,\infty).$$

The new parameter can now be estimated from experimental data along the same lines given earlier. However, a more refined analysis would assess the importance of these heat losses by conducting a sensitivity analysis on this new parameter.

This model refinement type of work is frequent in the modern, mathematically oriented engineering literature [19], [20]. However, some of the decisions made for modelling purposes are of a more fundamental nature, hence they are more consequential than others. Also, and as a byproduct of this, decisions are not all independent, hence a rather complex "decision tree" would be required in order to fully represent the logic involved. Nothing as ambitious as this shall be attempted here; we will content ourselves with the much more modest aim of exploring further the notions of determinism and continuity of the space and/or time representations, concepts which have been seen to lie at the root of all our modelling with differential equations. In practice, however,

these two notions are often taken for granted and seldom if ever is any further attention granted to them.

For instance, take the assumption of time being continuous; by this we mean that T should be an interval of real numbers, not necessarily $T=[0,\infty)$. For lumped parameter models, choosing T otherwise (say T at most countable) leads to difference equation models, such as

$$x_{k+1}=f(k,x_k), \tag{4}$$

in which trajectories are vector sequences. These models yield their predictions by simple iteration, hence they are very manageable from a computational point of view. On the other hand, they always have a state space and are necessarily deterministic. These properties have rendered them attractive and made them the recipients of much attention lately [21].

An important particular case obtains when $T=\{0\}$, a possible choice of time representative for steady state models. Then, instead of a vector difference equation such as (4) one gets a system of algebraic equations such as

$$g(x)=0, \tag{5}$$

with $g:\mathbb{R}^d\to\mathbb{R}^d$. Any solution of (5) is a steady state predicted by the model, of which there can be more then one: some of them may be observable, some of them not. In any case, experimental confirmation must be sought for these possible steady states. The unobservability of some of them is sometimes traceable to unstability, which must in turn be investigated by means of a dynamical model. See [22], [23] for this type of analysis in the field of reaction-diffusion systems.

On the contrary, distributed parameter models still give rise to differential equations (ordinary or partial, see Example 3) for the steady state; other choices of T lead to systems of functional difference equations, with essentially the same properties as (4).

With all its implied consequences as to the type of model it gives rise to, the choice of a continuous representative for either time or space is not really essential. It is an often convenient choice simply because so many physical laws are expressed in differential form, hence in terms of limits. From a more fundamental viewpoint, what makes this choice attractive for modelling purposes is the completeness of the real number system. In a more practical vein, most of Classical Physics is expressed in terms of limiting operations, so differential equations based models may be confidently expected to stay with us for quite a while.

Let us end this section by pondering on the possible means and consequences of discarding the assumption of determinism. In the notation and terminology of section 3, we are forced to do so if no state space exists i.e. if for every choice of experimental frame X there is at least one $x \in X$ and one pair $(x,t,p,a'),(x,t,p,a'') \in D$ with a' and a" distinct. Again, "distinct" must be read as "distinguishable under the conditions of the experiment".

In the notation of this section and making reference to table (1), this would occur if r repetitions of the same experiment (i.e. for the same initial state and the same time instants $t_1,\ldots,t_N$) yield differing second rows. Again, "equal" and "different" must be understood in terms of

"distinguishability under experimental conditions". In an attempt at quantifying this notion one might proceed as follows.

For every $i=1,\ldots,N$, let c_i stand for the average of all the differing x_i's obtained in the r repetitions of the experiment, and let B_i stand for a neighbourhood of c_i. Then it makes sense to ask for each repetition whether the trajectory actually observed belongs to the set

$$\{\omega\in\Omega:\omega(t_1),\ldots,(t_N)\in B_N\}. \qquad (6)$$

Count the number of those repetitions for which an affirmative answer was given and divide it by r, the total number of repetitions of the same experiment. Let the resulting quotient be denoted by f_r, a number between 0 and 1 hereafter referred to as the relative frequency of the event that a trajectory belongs to set (6).

A positive number η may be selected, with the idea of rejecting the assumption of determinism if and only if $f_r<\eta$.

It may be felt that we are making it hard to reject this assumption by proceeding in this fashion; this is indeed the case, simply because we are dealing with a rather fundamental assumption, with deep implications for our modelling.

On the other hand, a variant of this criterion could have been chosen from the very beginning: a deterministic viewpoint is adopted only if $f_r>1-\eta$. Then we say that "the trajectory values at instants $t_1,\ldots,t_N$ are $c_1,\ldots,c_N$, respectively", explaining away all discrepancies from this statement in terms of "experimental error".

One feels justified in this procedure by the commonly made observation that, as r increases without bounds, f_r fluctuates about a fixed number $p \in [0,1]$, with deviation amplitudes decreasing "in the average". This observation embodies the so-called Principle of Statistical Regularity, and p is identified wih "the probability that an observed trajectory belongs to set (6)". In other words, one is simply acknowledging the varability present in all experimental situations while assessing determinism itself.

If determinism is finally rejected, then it is convenient to remind ourselves that a probability can be assigned to all sets of trajectories such as (6)-at least in principle-according to the Principle of Statistical Regularity. This opens up the possibility of giving a probabilistic description of the corresponding physical model in the terms made precise by Kolmogorov's Extension Theorem [24]. Thus we see that probability spaces appear as the natural models to consider in the absence of determinism.

6. SIMULATION AND MODEL COMPLEXITY.

It should be clear by now that conclusions must be drawn from a given model in order to test its validity by comparing them with experiment , hence the advantage of having access to expedient procedures for "solving the model's equations". Simulation is the art of devising those efficent algorithms as well as implementing them in order to obtain the required predictions.

On the other hand, the fundamental role played by the various assumptions underlying a given model must be apparent by now, together with the long series of decisions required in modelling. It is only natural to expect a high degree of

interplay between the corresponding decision making and algorithm design i.e. between modelling and simulation. In this section we shall attempt to illustrate this interaction while giving an overview of algorithm design.

For instance, consider the situation described in Example 1. A deterministic model was arrived at under the stated assumptions, in which the basic problem to be solved for validating it is the IVP (1.1). Now, this IVP is indeed very simple: in the particular case in which $\mu^2<4k$ (underdamped oscillations), its solution can be expressed in closed form as

$$x(t)=Ae^{-t/2}\sin(\omega t+\delta)+\int_0^t h(t-s)f(s)ds \qquad (1)$$

where $\omega^2:=k-\mu^2/4$, both A and δ can be expressed in terms of x_0 and v_0, and

$$h(t):=\frac{A}{\omega}\exp\{-t/2\}\sin\omega t.$$

Analogous expressions hold in the complementary cases.

Let's now relax the assumption requiring linearity in both spring and viscous medium constitutive equations, just as was done in section 4. There results an extended model embodied in (4.3), with φ and ψ continuously differentiable functions satisfying (4.2).

It is possible to investigate this nonlinear model using qualitative methods [25], but there is no hope of ever explicitly

solving the differential equation as in the linear case. The road is open for methods yielding approximate solutions, though, and we would like to concentrate on them from now on.

For, observe that the IVP for (4.3) can be expediently converted into one for a system of ODE's in the vector variable $z:=(x,v)$, with $v:=dx/dt$, namely

$$\dot{z}=F(t,z),\ z(t_0)=z_0, \tag{2}$$

with $t_0:=0$, $F(t,x,v):=\left(v,\phi(x)+\psi(v)+f(t)\right)^T$.

Let us take this IVP as our starting point. However, we will explicitly deal with the more general situation in which F is an $\mathbb{R}^n$-valued vector function.

From well known results in the theory of ordinary differential equations [25], under the current differentiability assumptions there exists a closed interval $[a,b]$ with $t_0\in(a,b)$, on which a unique solution z of (2) exists. Let d be the uniform convergence metric in $C[a,b]$, namely

$$d(f,g):=\max\{|f(t)-g(t)|:t\in[a,b]\}.$$

An ε-approximate solution of the IVP (2) over $[a,b]$ is any function $w:[a,b]\to\mathbb{R}^n$, continuous there, such that $d(w,z)<\varepsilon$. An approximate solution algorithm for (2) will yield an ε-approximate solution for any positive ε. There exist many kinds of such algorithms, but we will concentrate in what follows upon numerical methods.

For, the IVP (2) is readily seen to be equivalent to the

nonlinear Volterra type integral equation

$$z(t)=z_0+\int_0^t F(s,z(s))ds. \tag{3}$$

Discretize time, replacing the basic interval $[t_0,b]$ by the partition

$$t_0,t_1,\ldots,t_m:=b,$$

with

$$t_{k+1}=t_k+h,\ k=0,1,\ldots,m-1,$$

for a positive stepsize h.

Then, (3) gives rise to

$$z(t_{k+p})=z(t_k)+\int_{t_k}^{t_{k+p}} F(s,z(s))ds,$$

for each positive integer p. Use the approximate integration rule

$$\int_{t_k}^{t_{k+p}} h(s)ds \sim \sum_{j=0}^{p} c_j h(t_{k+j}) \tag{4}$$

upon the right hand side of the last relation to get

$$z(t_{k+p}) \sim z(t_k) + \sum_{j=0}^{p} w_j F(t_{k+j}, z(t_{k+j})). \tag{5}$$

Note that approximate integration rules like (4) can be devised for which the right and left hand sides of (5) differ by less than any prespecified tolerance [26]. This suggests the possibility of generating an ε-approximate solution of (2) by interpolating the finite sequence of points

$$(t_0, z_0), (t_1, z_1), \ldots, (t_m, z_m)$$

generated by the algorithm

$$w_0 := z_0 \tag{6a}$$

$$w_{k+p} := w_k + \sum_{j=0}^{p} c_j F(t_{k+j}, w_{k+j}). \tag{6b}$$

This algorithm is said to be multistep if $p>1$; otherwise it is single step. Note that (6) defines a recurrence relation similar to (5.4) in the single step case. On the contrary, any multistep instance of (6) reduces to a p^{th} order difference equation, which must be solved by previously generating $w_1, \ldots, w_{p-1}$ via an appropriate single step algorithm.

Algorithm (6) is said to be explicit if $c_p=0$; otherwise it is called implicit. Formula (5.4) defines an explicit, single step algorithm and it can be implemented as a simple iteration. On the other hand, implementing an implicit instance of (6) requires solving an equation at every step of the iteration.

Thus, implicit methods are computationally more expensive, their advantage lying in their greater numerical stability [27].

An ε-approximate solution can indeed be obtained by interpolating the points generated by (6), provided h is chosen sufficiently small. Choosing h small enough to comply with given precision specifications frequently requires some trial and error, computational efficiency often demanding variable stepsizes. In some cases, however, there are error estimates that can be used to ease the task of selecting the right stepsize [27].

Discretization can occur under many guises when developing computational simulation algorithms. Take, for instance, the BVP (2.5) arrived at in Example 3. The basic region G being bounded, it can be enclosed in a rectangle R of sufficently large sides, say $R=[a,b]\times[c,d]$. In turn, both sides of the rectangle can be partitioned, say by

$$a=:x_0<x_1<\ldots<x_m:=b,$$
$$c=:y_0<y_1<\ldots<y_n:=d;$$

let P denote the set consisting of the points (x_i,y_j) thus determined which actually lie in G. Discretization is then achieved by substituting P for G. Just for the sake of convenience, assume the planar mesh thus generated consists of squares, each of side h.

It is an easy consequence of Taylor's Theorem that

$$\frac{f(x+\delta x)-2f(x)+f(x-\delta x)}{(\delta x)^2} \to f''(x) \qquad (7)$$

when $\delta x \to 0$, provided f is C^2 at x. This result suggests that the following procedure can be applied to the discretization of the BVP (2.5).

Number the points of P in such a way that the first N of them are interior to G, say $p_1,\ldots,p_N$, while $p_{N+1},\ldots,p_{N+M}$ are not. For any real function v on G, write v_i for $v(p_i)$, $i=1,\ldots,N+M$. Finally, let $p_{N+M+1},\ldots,p_{N+M+K}$ be the points in the mesh which are not in P, and define $v_{N+M+1},\ldots,v_{N+M+K}$ arbitrarily.

Let u be the solution to the BVP being considered; clearly $u_{N+1},\ldots,u_{N+M}$ are known from the data: they coincide with $f(p_{N+1}),\ldots,f(p_{N+M})$. For each index value i not exceeding N, let p_j, p_k, p_m and p_n be the first neighbours of p_i, then approximate each of the partial derivatives making up the Laplacian of u at p_i as suggested by (7).This gives rise to

$$u_j+u_k+u_m+u_n-4u_i \sim 0,$$

a relation involving at most five unknown values. Some of the known values should better be eliminated, though, in case they have indices greater than N+M. To illustrate how, assume $j>N+M$, while $i,k,m,n<N$, so that there necessarily is a boundary point q in the segment between p_i and p_j; denote the abscissae of these three points by x_b, x_i and x_j, respectively. Then,

$$u(q) \sim [(x_j-x_b)u_i+(x_b-xi)u_j]/h,$$

thus making it reasonable to replace u_j by

$$[hf(q)-(x_j-x_b)u_i]/(x_b-x_i)$$

in the corresponding relation. One is then left with the approximate equality

$$u_k+u_m+u_n-cu_i \sim b_i, \tag{8}$$

where $c:=4+(x_j-x_b)/(x_b-x_i)$, $b_i:=-hf(q)/(x_b-x_i)$. Similar relations obtain quite analogously for each of the N values of i.

Let x be the N-vector satisfying the system of linear equations

$$\mathbf{Ax=b} \tag{9}$$

the i^{th} of which is obtained by substituting = for $\sim$ in the corresponding approximate relation (8).

Systems like (9) may be accurately solved by rather well known techniques , using widely available computational packages [28]. Then the components $x_1,\ldots,x_N$ of x approximate the values $u_1,\ldots,u_N$ of the solution at the interior points $p_1,\ldots,p_N$, respectively. A measure of the total error is

$$e_h:=\max\{|x_i-u_i|:i=1,\ldots,N\},$$

and it can be rendered smaller than any $\varepsilon>0$ by choosing h small enough. An ε-approximate solution of the BVP can then be constructed by merely interpolating the points

$$(p_i,x_i),\ i=1,\ldots,N.$$

Note that there are as many linear equations as there are interior points of G in P, each of them with <u>at most</u> 5 unknowns, hence a proportion not greater than 5/N of the elements of A will

be nonzero. This remark says that **A** is sparse, a circumstance with very important practical consequences, since it permits saving resources in terms of both

a) memory requirements for the mere storage of NxN matrix **A**,

b) computational effort required in order to solve (9).

In fact, at most 5N numbers and positions (say 15N memory locations) will be required in order to store the NxN matrix **A**, a substantial economy for large N. Moreover, appropriately labelling the points in P helps to reduce the bandwidth of **A**, thereby simplifying the elimination procedure required for solving (9). See [29] for details.

This type of discretization is well documented, see for instance [30]; it can be extended to dynamic models like that resulting from Example 2 or from its extensions, the unbounded character of the t variable giving rise to some interesting mathematical problems. Alternative discretization techniques applicable in all these cases are the so called finite element methods; they are very convenient from the point of view of their greater flexibility, when compared with the finite difference techniques we have just described. Nevertheless, they do employ more sophisticated mathematical machinery to go from the BVP to be solved to the resulting linear system (9). See [31].

Both of these discretization techniques are applicable to the solution of nonlinear BVP's or IBVP's like (4.5) and (4.6), those BVP's resulting from relaxing certain simplifying assumptions in Example 3. When applied to them, both finite difference and finite element techniques give rise to finite

dimensional systems too, but this time they are nonlinear, of the type (5.5).

Now, most nonlinear equation solvers work sequentially: a sequence $\{x_k\}$ of iterates is generated from a given starting point x by resorting to a suitably devised scheme such as (5.4), the sequence converging to the solution as $k \to \infty$. However, the really good ones (like those in the Newton-Raphson family) are implicit, and require solving a linear system like (9) at every iteration, see [32]. Hence dealing with very general nonlinear models may be computationally very costly-and even prohibitive, depending on the current economics of computation.

This leads us straight back to the question of how detailed a differential equation model ought to be. Sure enough, making too many simplifying assumptions may yield simple but very bad models; on the other hand, attempting to "take every aspect into account" may lead to a very faithful model...if and when predictions can indeed be extracted from it. Where is the right balance?

In this connection, it is appropriate to mention some recent developments in Identification Theory [33], [18] dealing with the construction of discrete time models of time series. An interesting notion of model complexity appears there, in the Rissanen approach to structure determination [34], see also [16], and it might be thought feasible to apply it here. It must be borne in mind, however, that the models considered therein are already cast in an algorithmic fashion, whereas most "serious" DE models must be somehow discretized before becoming computationally feasible. In our terminology, Rissanen's information theoretical approach pertain to ALGORITHM rather than MODEL complexity: as such it can certainly constitute

the main ingredient of a formalized notion of model complexity, along the lines suggested in the Introduction. There are, however, other concepts of algorithm complexity which one might also like to consider for these purposes.

Once a satisfactory notion of model complexity is adopted, a reasonably systematic approach would start out by making a list of all the features worthy of consideration in a given modelling situation. Using the modeller's expertise, some of these features may be selected for actual consideration, a finite (and not too large) family of possible models arising from this, each characterized in terms of a set of assumptions. In turn, this family of models can be structured by the relation "being derived from by simplifying", which makes it into a lattice-like ordered structure. See [35] for an enlightening illustration of such procedure, based on the "sound engineering judgement" of model complexity.

In general, the resulting discrete family of models can then be explored from simpler to more complex in a systematic fashion, moving "up" the complexity ladder only if agreement with experiment is not satisfactory, otherwise exploring "sidewise". This procedure would make it possible to attain a convenient combination of model complexity and simulation feasibility.

7. CONCLUSIONS.

As we see, modelling as an activity is highly dependent on the modeller's expertise and good judgement, and no other possibility seems open so far. Nevertheless, a rationale for model selection has been outlined, based upon the notion of model complexity. Moreover, it has been indicated how this last notion can evolve from that of algorithm complexity, at least in

principle: the complexity of a DE model coincides with that of the least complex simulation algorithm based upon that model. The work of Traub et al. on complexity theory seems highly relevant in connection with this idea of model complexity, see e.g. Question 1 on p. 3 of [4].

It is clear from the presentation given above that modelling and simulation are highly interrelated: simulation is carried out in terms of a model and, conversely, the choice of a model requires substantial simulation work. Hence it is only fair that a definition of model complexity takes simulation into account. A promising area for further research is the development of these ideas using an appropriate notion of algorithm complexity, and Rissanen's information theoretical notions seem most relevant. On a more practical vein, it seems quite desirable to develop modelling/simulation packages based upon these ideas of model complexity, along the lines of the last paragraph of section 6 .

It should be apparent that most of the foregoing considerations apply without change to models other than those given in terms of DE's. On the other hand, from pedagogical reasons it seems sound to center the discussion on a particular family of models: the well known interest of DE based models in modern science and engineering explains our choice of family. It is clear that, together, determinism and the choice of continuous time or space representative set the ground for the emergence of DE models; all extras required are physical laws expressed in differential form plus the necesary differentiability assumptions. In turn, the actual setting up of DE models can be systematized to a certain extent via the "conservation law+constitutive equation" approach described above. The DE's are ordinary when time is continuous and space is discrete, also when space is one dimensional and time is

discrete; in all remaining cases there result PDE's.

Space limitations prevented us from considering input-output DE based models, either with regards to their origin in experiment or their peculiarities from the point of view of simulation. We stop short of doing that, however, and indeed the present treatment suffices in order to deal with a fixed input function. In fact, readers conversant with mathematical system theory [10] will undoubtedly note that, given both determinism and the notion of state space, it only remains to relate causality to experiment in order to arrive at the notion of dynamical system as the fundamental input-output model.

8. ACKNOWLEDGEMENTS.

F. Vàzquez made the author aware of [16]. S. Bittanti pointed out the relevance of Rissanen's work to the ideas on model complexity presented in an earlier version of this paper; both he and L. Kolawole made valuable editorial comments therein, many of which have been incorporated into this new version. Earlier stimulating discussions on modelling held with R. Canales, J. Morcos, S.T. Guillen and A. Alonso are also acknowledged. The first draft of this paper was processed on board the M/V "Amparo" (courtesy of Transportacion Maritima Mexicana), en route from Tampico to Genoa. Both the final editing and printing were carried out on the computing facilities of the Dipartimento di Elettronica (Politecnico di Milano). Throughout, support came from the Universidad Autonoma Metropolitana as well as from CONACYT (Mexico), grant 85-01597.

BIBLIOGRAPHY.

[16] CAINES, P.,"On the scientific method and the foundations of system identification". Conference on Information Sciences and Systems, The Johns Hopkins University,1985.

[17] SEINFELD,J.H. & L. LAPIDUS,"Process Modeling, Estimation and Identification". Prentice Hall Inc., Englewood Cliffs NJ, 1974.

[18] BITTANTI, S. (ed.),"Identificazione parametrica". CLUP, Milano, 1981.

[19] ARIS, R., "Mathematical modelling techniques". Pitman, London, 1978.

[20] STEWART, W.E., W.H. RAY & C.C. CONLEY, "Dynamics and modelling of reacting systems". Academic Press, New York, 1980.

[21] GREENSPAN, D., "Discrete models". Applied Mathematics and Computation No. 3, Addison Wesley, Reading MA, 1978.

[22] VARMA, A. & R. ARIS, "Stirred pots and empty tubes", pp. 20-55 of "Chemical Reactor Theory", L. Lapidus & N.R. Amundson, eds. Prentice Hall, Englewood Cliffs NJ, 1977.

[23] LUSS, D., "Steady state multiplicity and uniqueness criteria for chemically reacting mixtures", pp. 131-159 of [20].

[24] LAMPERTI, J., "Probability: A Survey of the Mathematical Theory". Benjamin Co., New York, 1966.

[25] LEFSCHETZ, S., "Differential Equations: Geometric Theory". Interscience, New York, 1964 (reprinted by Dover, Inc.).

[26] DAVIS, P.J. & P. RABINOWITZ,"Methods of Numerical Integration". Academic Press, New York, 1984.

[27] GEAR, G. W., "Numerical initial value problems in ordinary differential equations". Prentice Hall, Englewood Cliffs.

[28] FORSYTHE, G. E. & C. B. MOLER,"Computer solution of linear algebraic systems". Prentice Hall, Englewood Cliffs NJ, 1967.

[29] GEORGE, A. & J.W. LIU, "Computer solution of large sparse positive definite systems". Prentice Hall, Englewood Cliffs NJ, 1981.

[30] FORSYTHE, G. & W. WASOW, "Finite difference methods for partial difference equations". John Wiley and Sons, New York, 1964.

[31] STRANG, G.N. & G.J. FIX, "An analysis of the finite element method". Prentice Hall, Englewood Cliffs NJ,1973.

[32] ORTEGA, J.M. & W.C. RHEINBOLDT, "Iterative solutions of nonlinear equations". Academic Press, New York,1970.

[33] ASTROM, K.J. & P.E. EYCKOFF,"System identification: a survey". Automatica 7, 123-162.

[34] RISSANEN,J.,"Modelling by shortest data description". Automatica 14 (1978),465-471.

[35] ARIS, R."Hierarchies of models in reactive systems", pp. 1-36 of [20].