

A MODEL FOR THE GROWTH OF MATHEMATICAL SPECIALTIES*

M. KOCHEN, A. BLAIVAS

*University of Michigan, Mental Health Research Institute, 205 Washtenaw Place, Ann Arbor,
Michigan 48109 (USA)*

(Received April 1, 1980 in revised form November 18, 1980)

A mathematical model for the growth of two coupled mathematical specialties, differential geometry and topology, is analyzed. The key variable is the number of theorems in use in each specialty. Obsolescences of theorems-in-use due to replacement by more general theorems introduces non-linear terms of the differential equations. The stability of stationary solutions is investigated. The phase portrait shows that the number of theorems in low-dimensional topology relative to those in differential geometry is increasing. The model is qualitatively consistent with the growth of publications in these two specialties, but does not give quantitative predictions, partly because we do not use an explicit solutions as a function of time and partly because only two specialties are used. The methods of analysis and some of the concepts can be extended to the development of more general and realistic models for the growth of specialties.

Introduction

A scientific specialty can be characterized by (*Kochen*¹ 1974): (1) a community of active and supportive contributors (estimated to consist of about 160 scientists at any time, including about 40 transients entering and 40 leaving each year (*Price*² 1976); (2) a cluster of publications linked more to one another by co-citation bonds than to other publications (*Griffith* and *Small*³ 1974); (3) a finite relational structure of assertions representing established knowledge as well as outstanding conjectures (*Kochen*⁴ 1974), (which, in the case of mathematics as a whole, is estimated to grow at about 200 000 theorems per year). One of the remarkable empirical findings in the “mapping of science” (*Price*² 1976; *Griffith* and *Small*³ 1974) is that two dimensions suffice to represent most of the published literature as points in a metric space so that the distances between any two points corresponds to the number of papers citing both.

Suppose that the set of scientists and the set of assertion can also be represented as metric spaces of suitable dimensionalities. We thus have three spaces that

*Supported in part by grant IST 78-16629. The authors would like to express their thanks to Derek de *Solla Price* for very valuable comments that helped us to improve this paper.

we will assume to be continuous, differentiable manifolds. Denote them by K, P, D , for Knowledge, People and Documents, respectively.

At each point $k \in K$ denoting an item of knowledge such as a theorem, let $\underline{K}_t(k, t)$ denote a vector field representing the rate of change in the number of assertions in the neighborhood of k . Here t denotes time, and its use as a subscript means the derivative with respect to t . It is helpful to visualize \underline{K}_t as an arrow at a point k of, say, a plane \underline{K} , which points in the direction in which knowledge in that area grows. This varies both with time and from point to point. Let $\underline{U}_t(k, t)$ denote the vector field corresponding to the rate of change in the number of conjectures at k . Let \underline{K} and \underline{U} denote the integral curves of flow lines corresponding to a system of differential equations, or equivalently, to the action of a continuous 1-parameter group.

Similarly, we assume the existence of vector fields \underline{I}_t and \underline{S}_t and flows $\underline{I}, \underline{S}$ corresponding to the number of actual contributors (analogous to infectives) and of potential contributors (scientists susceptible to "infection") and flows \underline{D} and \underline{P} (documents and proposals).

Differential equations corresponding to those of epidemiology,

$$\underline{I}_t = a\underline{I}\underline{S} + b\underline{I} + c \tag{Eq. (1')}$$

and

$$\underline{S}_t = -a\underline{I}\underline{S} + b'\underline{S} + c \tag{Eq. (1'')}$$

have been analyzed (*Goffman*,⁵ 1966) and found to fit certain kinds of bibliometric data. By analogy, we assume that \underline{K}_t is proportional to both the number of theorems and conjectures as well as the number of contributors. That is, each theorem-in-use helps to transform a conjecture-in-use into a new theorem with probability d per unit time. In addition, each theorem-in-use is used in the proof of e new theorems (e is a fraction). The number of new theorems produced per unit time is also proportional to the number of investigators. Thus,

$$\underline{K}_t = d\underline{K}\underline{U} + e\underline{K} + f + g\underline{I} \tag{Eq. (2)}$$

$$\underline{U}_t = -d\underline{K}\underline{U} + e'\underline{U} + f' + g'\underline{I} \tag{Eq. (3)}$$

Similarly, coupled equations can be written to relate, $\underline{I}, \underline{D}, \underline{S}, \underline{P}$. To account for diffusion effects, we might replace \underline{K} in Eq. (2), by the Laplacian Δ^2 operating on a density function corresponding to \underline{K} . The resulting system will then resemble coupled nonlinear reaction diffusion equations. f is a fixed growth rate.

Quite a variety of such equations can, of course, be written and analyzed. We do not yet have enough understanding or factual knowledge to select a model that is good for its predictive utility. The ideas and results we present here are a

step beyond that of *Goffman*⁵ in what may become a series of steps toward a satisfactory model. Hopefully, other investigators interested in the growth of specialties will find our considerations a useful basis on which to build or a challenge to the development of more realistic models.

In what follows, we analyze a special model for the growth of two related mathematical specialties. The method of analysis can be extended straightforwardly to more than two dimensions, but the increased computational complexity would obscure the basic features of the model. We may, however, have obtained better predictions. The two specialties, differential geometry and lowdimensional topology, are distinguished by the different journals, by the investigators who consider themselves and are considered by their peers to be experts in these fields, and by characteristic theorems and methods of each.

A model for two subspecialties

Consider the case of two clusters or subspecialties within a field of specialization that are centered at nearby points k_1 and k_2 in knowledge space. For notational convenience let $x' = \underline{K}_t(k_1, t)$ and $y' = \underline{K}_t(k_2, t)$ denote the change rates of "active knowledge" in the mathematical subspecialties: Differential Geometry and Low Dimensional Topology. We interpret the amount of active knowledge as the number of non-obsolete theorems and methods in these fields. While the generation of new knowledge increases the number of such theorems in each pool, it also renders some theorems obsolete in the sense that they now become special cases of the new discoveries. The pool of non-obsolete theorems are those theorems in active use during a given time period.

For one subspecialty, with x non-obsolete theorems, a simple model would be analogous to a tank to which new theorems are added and those becoming obsolete are withdrawn. We assume the process by which new theorems are added to involve the use of non-obsolete theorems in the proof, using $1/a$ theorem-years per new theorem generated. Without obsolescence we would have $x' = ax$. The coefficient, a , can be interpreted as characterizing the creativity of mathematicians. Although the coefficient may actually depend on time and, perhaps on x , we assume it (for simplicity's sake) as a constant, i.e. time independent.

We assume the process by which theorems become obsolete — perhaps because they become special cases of new theorems — to involve an implicit comparison of each non-obsolete theorem with every other one, with $1/2b$ theorem-pair-years per discovery of one that is recognized as a special case. That diminishes the rate at which non-obsolescent theorems grow by bx^2 .

We also assume close contact between the two related subspecialties. This may be interpreted as the existence of side-flows of knowledge from one subspecialty to the other. We suppose, in this case, that the non-obsolete theorems used in proving a new one in a subspecialty x , come not only from that subspecialty but from the neighboring subspecialty y , as well. We assume specifically that the number of theorems in the y -specialty has less influence on the rate of theorem production than the *rate* at which theorems grow in the x -specialty itself. The vigor of the neighboring discipline contributes to growth at a rate of $1/c_{xy}$ new theorems in the y -specialty per new theorem in the x -specialty. That is, encyclopedists, who know a great deal about the known theorems in the neighboring specialty, affect the productivity of theorems in a negligible way compared with the influence of active researchers. Thus, Eq. (2) is replaced by the following pair of differential equations:

$$x' = a_x x + c_{xy} y' - b_x x^2 \tag{Eq. (4)}$$

and

$$y' = a_y y + c_{yx} x' - b_y y^2 \tag{Eq. (5)}$$

where the subscripts refer to the two subspecialties.

A negative value of c can be interpreted as an inhibitory influence of one cluster on another. Note that x' could become negative if existing theorems are discovered to be special cases of other theorems faster than new ones are discovered, but we expect that $x' \geq 0$ and $y' \geq 0$ for all time.

While confining attention to the interaction of but two of many subspecialties is an oversimplification, it is a step beyond the one-dimensional case. Also, there are actual topics in which pairwise interactions are sufficient to account for the major effects.

Assuming that $a_x = b_x r_x$ and $a_y = b_y r_y$ then this pair of equations may be rewritten as:

$$x' = \frac{1}{1 - c_{xy} c_{yx}} [b_x (r_x - x)x + c_{xy} b_y (r_y - y)y] = P(x, y) \tag{Eq. (6)}$$

$$y' = \frac{1}{1 - c_{xy} c_{yx}} [b_y (r_y - y)y + c_{yx} b_x (r_x - x)x] = Q(x, y) \tag{Eq. (7)}$$

We can see that the system of algebraic equations

$$P(x, y) = 0$$

$$Q(x, y) = 0$$

has exactly four roots: $(0,0)$, (r_x, r_y) , $(0, r_y)$ and $(r_x, 0)$. These are stationary points for the system of equations (6) and (7), at which there are maxima, minima, or saddle points.

Range of Parameters. Initially, we considered all parameters $a_x, a_y, b_x, b_y, r_x,$ and r_y positive. This is due to an assumption that c_{xy} and c_{yx} would be *positive* influences on the two clusters. Similarly we expected that a_x and a_y would be positive because we interpreted them as *positive* influences on the amount of knowledge in each group and the generation of new knowledge. However, measurement of experimental data (see below) and estimates of these parameters gives us some negative values. We will discuss this point in more detail later but we now assume that these parameters have no restrictions.

The actual behavior of trajectories around the stationary points (i.e. the character of these points) depends especially on whether $c_{xy}c_{yx} < 1$ or $c_{xy}c_{yx} > 1$. To determine qualitatively the solutions to eqs. (6) and (7), we needed some estimates — even very approximate ones — of c_{xy} and c_{yx} . To do this we estimated x', y', x and y for 4 years, 1975, 1976, 1977, and 1978 and for two subspecialties: Differential Geometry (labeled 53A in the classification schedule of *Mathematical Reviews*) and Low Dimensional Topology (55A).

Table 1
Number of new theorems per year

Field	Year			
	1975	1976	1977	1978
53A	225	180	175	70 = $x'(t)$ Differential Geometry
55A	67	55	38	40 = $y'(t)$ Low Dimensional Topology

We estimated, from the number of reviews that appeared in *Mathematical Reviews* a year later, the following relative number of articles published per year.

The number of theorems produced per year is very approximately four times the annual production of articles in these fields. To obtain x from x' for a given year, we multiply $x'(t)$ by the average number of different older articles referring to an article published in year t . If R_x is the average number of references cited in a year- t -paper, and neglecting the probability $p(k)$ that a randomly chosen pair among all the possibilities have papers in common, then $x = x'R_x - \binom{x'}{2} \sum_{k=0}^{R_x} kp(k) +$ higher order terms that are taken to be negligible.

A direct estimate of the average number of papers cited by both papers in randomly chosen pairs from 53A and 55A is approximately zero for both fields. Thus

x, y are estimated, with $R_x = 13, 10, 8$ for 1975, 1976, 1977 respectively and $R_y = 9, 13, \text{ and } 12$.

Table 2
Number of theorems

Field	Year			
	1975	1976	1977	
53A	13 × 225	10 × 180	8 × 175	(Differential Geometry)
55A	9 × 67	13 × 55	12 × 38	(Low Dimensional Topology)

We can now estimate a_x, c_{yx}, b_x by substituting the numbers from the above two tables and Eq. (4) to obtain

$$225 = 2925a_x + 67c_{xy} - (2925)^2 b_x$$

$$180 = 1800a_x + 55c_{xy} - (1800)^2 b_x$$

$$175 = 1400a_x + 38c_{xy} - (1400)^2 b_x$$

Solving these three equations simultaneously for a_x, c_{xy}, b_x gives the estimates of 0.25, -2.73, and 0.000039 respectively. We repeat this procedure with Eq. (5) which takes the form:

$$67 = 603a_y + 255c_{yx} - (603)^2 b_y$$

$$55 = 715a_y + 180c_{yx} - (715)^2 b_y$$

$$38 = 456a_y + 175c_{yx} - (456)^2 b_y$$

to obtain the estimates of -0.21, 0.49, and -0.0002 for a_y, c_{yx} and b_y respectively.

Next it is necessary to find the results for r_x and r_y . This can be done by using Eq. (6) and Eq. (7) with Tables 1 and 2. The results are two linear equations:

$$1\ 018\ 206 = -318r_x + 146r_y$$

$$422\ 312 = 58r_x + 126r_y$$

with the results $r_x = 5873$ and $r_y = -5824$

We used derivatives from Table 1 for Differential Geometry and Low Dimensional Topology for the year 1975, but the results were almost the same for the other three years also. This increases the plausibility of our approximate estimates.

Next we must investigate the solution around the stationary points. This system has four stationary points: $(0,0)$, (r_x, r_y) $(r_x, 0)$ and $(0, r_y)$. Standard qualitative analysis shows that point $(0,0)$ has a positive determinant and a negative radicand.

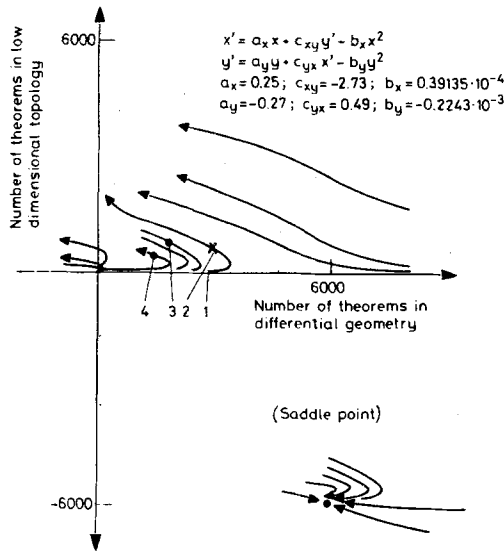


Fig. 1. Direction of change in number of theorems in one specialty relative to another specialty

The system of two specialties can be regarded to start sometime at point 1. In 1975 it was at point 2. In 1976 it moved to point 3, a small deviation from what our model predicts, that it should be on the same curve as points 1 and 2. In 1977 it was at point 4, a further deviation.

This point would have complex roots of a secular equation with a positive real part, which results in an unstable focus here.

Point (r_x, r_y) has the same positive determinant and negative radicand as point $(0,0)$. The sign of the real part of the root is negative, making this a stable focus.

At point $(r_x, 0)$ we have a negative determinant and a positive radicand. This point will have two real roots, one will be positive and one will be negative; this is a saddle point.

The next point $(0, r_y)$ gives the same solution as the previous one (negative determinant, positive radicand and two real roots with different signs). This will also be a saddle point.

Interpretation. The actual characteristics of the equation pair (5) (6) are shown in Fig. 1. This figure shows only a portion of the entire map primarily in the positive ($x > 0, y > 0$) quadrant of the x, y -plane. The solutions were obtained using a Runge-Kutta procedure. In 1975 our system was at the point marked 2 (2925603) on a curve with an origin in the focus point (0,0). Presumably our system began its movement near point 1 on the same curve. The direction of motion is indicated by an arrow. This map may be used for prediction of further behavior of our system. We can see that Low Dimensional Topology would probably slowly increase the active theorems in use while Differential Geometry would lose them. This process cannot be scaled in terms of time. It may take hundreds of years to approach the y -axis: nothing can be predicted exactly.

However, we also have to take into account that the actual behavior of the system may not follow the obtained curve exactly. The reason is obvious: we compared the system of two equations, neglecting to connect the clusters with other mathematical specialties. We forcibly reduced a multidimensional case to a two-dimensional case. Although we tried to select two subspecialties with strong enough connections between each other and comparatively weak links to other subfields, such strongly dominating links may occur in the future. Thus, what we just stated may divorce the actual movement of the system from the predicted curve.

This did happen when we tried to trace the actual behavior of the two clusters over time and to compare the results with data for the following years (see Table 2). In 1976 the system was at point (1 800 715) theorems, point 3, on Fig. 1, that is a deviation of predicted behavior. The next year the system was at point (1 400 456), point 4 on Fig. 1. Thus the *actual* behavior of the two clusters does not exactly follow the predicted curves. However, the tendency is almost the same. To appreciate the deviations, one should take into account the roughness of our experimental estimates and the possibility of outside influences that were mentioned before.

Other parts of the plane with negative x and y values should only be considered as a mathematical offspring of our assumptions and should not be taken into account in a practical sense. However, the behavior of solutions around other non-stationary points could influence the trajectories in the working area of the x, y -plane and for this purpose the analysis should be made.

Our estimates gave us two striking results: $c_{xy} < 0$ and $r_y < 0$. The first result seems to indicate that Differential Geometry is under the negative influence of the neighboring field, Low Dimensional Topology (possibly due to people outflow) and is actually declining. This is also visible in Table 1. The second result, which was negative ($r_y = -5824$) does not fall into the positive range for theorems on the x, y -plane. This is in direct contrast to what we would normally expect to occur.

Conclusion

It appears fruitful to describe the dynamics of knowledge in a cluster of related mathematical specialties by means of coupled non-linear differential or difference equations for the number of theorems-in-use, the number of published papers reporting them and the number of active authors. We tried the simplest set of non-linear equations. Their solutions in phase-space gave a qualitative description of how the number of knowledge items in one specialty, or theorems in the specialties we studied, change relative to the other. The fit, with observed changes, though fair, is not good enough to validate the model based on the particular equations we used, but is evidence that the method is useful. The results are interesting and may contribute to our qualitative understanding of the content of knowledge.

References

1. M. KOCHEN, *Principles of Information Retrieval*, Melville/Wiley, Los Angeles 1974.
2. D. de SOLLA, PRICE, Studies in Scientometrics. Part I. Transience and Continuance in Scientific Authorship. *Int. Forum on Inf. and Document*, 1976. FID, Moscow, 1, 2.
3. H. SMALL, B. C. GRIFFITH, The Structure of Scientific Literatures, I: Identifying and Graphing Specialties. *Science Studies*, 4 (1974) 17–40.
4. M. KOCHEN, *Integrative Mechanisms in Literature Growth*, Westpoint, Conn., Greenwood, 1974.
5. W. GOFFMAN, Mathematical Approach to the Spread of Scientific Ideas – The History of Mast of Mast Cell Research. *Nature*, 212 (1966) October 5061.