INTERACTION OF LIBRARIES AND PUBLISHERS

MANFRED KOCHENT
Mental Health Research Institute, University of Michigan, Ann Arbor, MI 48109, U.S.A.

Abstract—Libraries and publishers have evolved together. Publishers rely on libraries as a minimum market for their scholarly products. Inflationary pressures have caused publishers to increase prices that, in turn, strain library budgets that have not increased as fast, and which, in turn, undermine the minimal demand publishers can count on, adding to inflationary pressure.

A simple mathematical model for the dynamics of the interaction between libraries and publishers is analyzed. It derives a function for the supply curve of scholarly publications, and is used to estimate when an institution will have to spend as much per person on library support as on his or her salary if present trends continue. This is used to argue that present trends are unlikely to continue, but that a discontinuous shift in the production of scholarly output is likely to occur within a decade or two. Likely new forms of communication among scholars in "communicating classes" involving nearly simultaneous communication and a new kind of organized cumulative record are discussed. The implication for institutional changes not only in libraries and publishers and their interrelation but of new kinds of institutions are sketched.

I. INTRODUCTION

Publishers of scholarly journals and books rely on the library market. Current technological developments and economic trends in the processing of the information that has been traditionally recorded in such books and journals may affect both the library community and the publishing industry. The two evolve in a symbiotic relationship. It is interesting and important, not only for those two sectors, but also for those people whose lifestyles are affected by them, to inquire into the consequences of alternative strategies for shaping the future of scholarly books and journals.

To the best of our knowledge, there are few data and no tested models for the demand of scholarly journals. Individual publishers would find it useful to know a formula or table showing the expected number of copies sold as a function of the unit price and the amount per unit spent on marketing for the industry as a whole, as well as the actual amounts (averages, variances) spent. They could then see how much or little they deviate from the average or from how well they might do.

Pricing depends, of course, on costs as well as on estimates of sales. Publishing costs, in turn, depend on technologies, management methods and prevailing prices of production factors. The latter are primarily know-how, energy, space and time. The price of all four factors has been increasing rapidly. As a result, so have the costs of books and journals. That has produced pressures to increase library budgets for purchases of books and journals. The institutions that decide on allocation of library resources may not have been able to respond to these pressures as readily as they have to other inflationary pressures. The resulting spiral may not continue indefinitely without evoking a discontinuous change in the evolution of the libraries-publishers couple.

The aim of this paper is to discuss the occurrence of such a discontinuity, to estimate when it might occur (if it does) and to investigate how to control the course of the publishing-librarianship relation. Our analysis is highly oversimplified in order to emphasize dramatically the authors' points of view.
the likely influence of digital communications technology. Far more sophisticated analytical studies could and should be made that take into account advances made on the economics of scale and scope [1] and that use available data [2].

2. CONCEPTUALIZATION

Researchers add to knowledge by contributing a steady stream of modest increments, punctuated by occasional giant strides or breaks. The units by which knowledge is incremented are represented in what we shall call "claims". Proved theorems and tested hypotheses, expressed and publicly presented so that any expert in the same speciality can verify their validity, are instances of claims. A paper reporting research is likely to be accepted by a reputable scientific journal only if its referees judge it to contain a number of related claims that they understand, believe to be valid, consider to be sufficiently interesting or important as well as novel. Thus a typical article might contain 3-10 claims that add, at least modestly or incrementally, to their speciality. Papers presented at conferences and published in refereed proceedings volumes follow a similar pattern. A section or chapter in a research monograph might be a corresponding package.

We distinguish here between knowledge, know-how and scholarly output of other kinds. We restrict our discussion to knowledge, primarily scientific knowledge, and even more particularly the kind of scientific output that lends itself to representation as claims to discoveries. (Know-how can also be represented as claims about procedures composed of instructions to accomplish a goal; how-to-do it scenarios, computer programs, recipes, blueprints, etc. are examples.)

The aggregation of articles into an issue or into a serial volume does not generally constitute a meaningful unit, as would be the case for an aggregation of sections or chapters into a book. But the journal as a whole is a meaningful aggregate, because it has a known publisher, editor and readership. The largest set of journals and several books can be further aggregated according to the smallest well-defined research community they all serve. In this way, we can characterize the embodiment or record of knowledge by four variables:

1. The number of claims in a given specialty at a specified time.
2. The number of articles (or book sections) in that specialty at that time.
3. The number of journals (or books) in it.
4. The number of distinct research communities in it.

As long as claims or articles are not circulated as separate items, the journal or book is the physical unit which presently characterizes much of the traffic of recorded knowledge. It is these units which, at present, researchers actually buy and handle (carry, read, shelf, etc.), which publishers produce and sell. Research libraries serve as intermediaries between researchers and publishers.

A scientific research library is generally part of an institution that adds to research or uses new knowledge. While most of these are special libraries, affiliated with an industrial laboratory, university or independent research institute, some are large governmental or private libraries (e.g. the U.S. Library of Congress, British Library, N.Y. Public Library, Lenin Library, etc.). Special libraries concerned with a specialty service the needs of researchers who are employed by the parent institution and who belong to a research community in that specialty. Thus, the number of libraries in a specialty is at present proportional to the number of institutions concerned with that specialty. Insofar as journals and books are still physical objects to be "handled by unaided human hands and eyes, they should ideally be within "arms's reach", and every institutional unit should have its own library that provides its users with the objects they need most urgently within minutes of need recognition.

The number of primary research journals published in the U.S. [King 3] varies with time, as in Table 1.

This looks like a reasonably continuous and monotonically increasing growth curve, with a current growth rate of almost 3% per yr. (Indeed, both a linear regression and the equation $e^{0.028(1-1704)}$ fit these data well.) The growth rate of the pool of educated people who need and use the information in such journals is currently growing at more than 3% worldwide [4]. But does
this translate into economic demand that continued growth in the number of journals will supply?

The scientific article or scholarly paper rather than the journal is probably a more basic unit for formal communication within a research community. In some journals, the number of articles published per yr has increased. It is therefore the number of articles existing at time \( t \) that may be a more appropriate measure of scientific output than \( J(t) \). This, labeled \( N(t) \) and measured in thousands, has grown as in Table 2 [5]:

Table 2

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>( N(t) )</td>
<td>208</td>
<td>217</td>
<td>252</td>
<td>305</td>
<td>354</td>
<td>382</td>
</tr>
</tbody>
</table>

(The average price of an individual, institutional and foreign subscription respectively went up from $8.00, 9.50, 9.50 in 1960 to $32., 36., and 43. in 1977.)

We observe first that the growth of knowledge does not depend entirely on need or demand. When a researcher makes a discovery he needs to communicate it, to record it. A work of scientific (or other) research is complete only when it is reported in public documented form for citation and criticism by the community of inquiry concerned with that work. Discoveries are more often the result of a process that moves by its own momentum—the next experiment or conjecture dictated by logic of the field of inquiry and by what techniques have become feasible than by the result of support for or utilitarian need for such discoveries[6]. (The latter are necessary but not sufficient.) Thus, a discoverer will publish his findings even if only one or two peers will read it with appreciation or criticism. But will authors forever find it convenient to express themselves through the scientific journals as we know them and will the needs of readers be met better by these forms than by alternatives to present book/journal technologies that have begun to appear[7]? Eventually, individual subscribers are likely to look to professional middle men services such as libraries (or their replacements in the future), such as Lockheed, SDC, etc. Let us examine the role of libraries in greater detail.

The average number of journals subscribed to by a library at time \( t \) should be roughly proportional to that library's journal budget in year \( t \). In many U.S. special libraries, the fraction of their total book/journal budget devoted to journals has increased from 25-40% to 60-75% and 90% in some cases[8]. Non public library budgets tend to be a fixed percentage of the total indirect cost budgets of the institution they serve, and there are limits to the growth of the latter. As we approach zero population growth, which is an eventual necessity for most countries, the size of library budgets, in constant monetary units, is expected to eventually stabilize.

As Rescher[9] points out, “The resource requirements for continuing the accustomedly smooth linear course of scientific progress increase in geometric proportion. But the resources made available in the zero-growth world which now lies around the corner will cumulate only linearly.” Insofar as library budgets for books and journals are included in the resources required for scientific progress, we can expect that requirements will begin to exceed what is available in the near future if current trends continue. Of course, we shape trends by our interventions, and we can neither extrapolate nor predict them.

We might expect that journals with large circulation or size, which are also journals that are highly cited, will be subscribed to by most libraries, so that the probability of the most popular and most cited journal being acquired by a library is high. For the less popular, less cited or more specialized journals, it is less likely that very many libraries hold the same journal.

3. DYNAMICS OF THE LIBRARIES-PUBLISHERS COUPLE

Suppose the price of publications is determined by the equilibrium of a free market, with libraries continuing to be a major part of the demand. If production costs increase faster than
the rate of inflation, while library budgets increase at a lower rate, it is a question of time before some growth discontinuities occur.

To estimate that time, we analyze a simple model characterized by the following variables:
1. $C_s(p)$: number of copies of a book or journal supplied at price $p$.
2. $C_l(p)$: number of libraries that buy such a publication at price $p$.
   The equilibrium price, $p$, is obtained by setting $C_l(p) = C_s(p)$.
3. $q$: quality of a publication.
4. $p(q)$: price as a function of quality.
5. $N(q)$: number of publications of quality $q$.
6. $B$: total budget of the library community.

Then $p(q)C_l(p)N(q)$ is the total amount spent by libraries on purchasing publications. This should be approximately equal to $B$.

Let us examine these functions in turn. Let $k$ be the unit (operating) cost of producing and marketing a publication as well as royalties to the author and $K$ the fixed (overhead) cost. For now assume no economies of scale or scope, subsidies, etc. Recall that $p$ is the price set by the publisher and $C_s$ the number produced. If the publisher expects to sell $\hat{C}_l$ copies at price $p$, then his net revenue is $\hat{C}_l p - C_s k - K$. (We ignore the inventory costs of storing an excess of unsold copies for now.) If the publisher seeks net revenues that are a fixed percentage $P$ of the total cost $C_s k + K$ as profit, then it follows that $C_s(p) = \left[\frac{\hat{C}_l p}{(1 + P)} - K\right]/k$.

Realistically, publishers do not produce a book unless they are quite sure of selling at east 1000 copies, usually to libraries; and about 300 subscriptions for a new journal. (The first press run of a book is usually less, perhaps 1000 copies.) If $M$ denotes this minimum, then we expect that $M p_{\min} = (M k + K) = P(M k + K)$, and

$$P_{\min} = (k + K/M)(1 + P).$$

A plausible supply curve then resembles that shown by the step-like solid line in Fig. 1.

With the equation

$$C_s(p) = \begin{cases} 0 & \text{if } p < (k + K/M)(1 + P) = p_{\min} \\ M + 1/k[\hat{C}_l p/(1 + P) - K], & \text{if } p > p_{\min}. \end{cases}$$

Neither publishers nor librarians know the demand function for publications. Careful studies of library subscriptions [10] showed that libraries proceed in several stages when they must reduce expenses. First, they eliminate duplicate holdings. Secondly, they seek to reduce costs of bindings, etc. Then they cancel journals that its patrons judge to be of marginal utility or quality. Only in the third or fourth stage will price begin to affect the choice of what to cancel or what not to buy.

A plausible family of demand curves may be expressed by $C_l(p) = C(1 + p)^{-a}$. The exponent is very close to 0, making the curve nearly horizontal, but it is closer to 0 for high quality than for lower quality publications.

![Fig. 1.](image-url)
We can reasonably expect costs $K$ and $k$ to increase exponentially over the next few years, at rates exceeding 12% or so unless new technologies displace human labor costs in the publishing industry sooner than expected or unless inflation decreases. Hence, $p_{\text{min}}$ is likely to increase exponentially too, with the slope of the supply curve flattening. If the demand curve stays fixed, (as shown in Fig. 1 by the dashed line) it will before long intersect the supply curve making the market price $p_{\text{min}}$.

Suppose now that the quality of a publication is measured by its “impact factor”, the average number of citations per article in a journal. Data from ISI [11] suggests that $N(q)$ may decrease with $q$ as $N_d(1 + q)^{-b}$. The total expenditures by all libraries would then be

$$N_d(1 + q)^{-b}p_{\text{min}}(q)M(q) = B.$$ 

When will the amount an institution spends per library user on its library services equal that user’s salary, if present trends continue? If $I$ is an individual’s present income, and it is raised at $r\%$ per yr, his income after $t$ yr from now will be $I(1 + r)^t$. If $C$ is the amount presently spent per user on library services, and that had to increase (to meet inflated prices) at $R\%$ per yr, this will be $C(1 + R)^t$ in $t$ yr. The value of $t$ at which these are equal is

$$t = (\log I/C)/(\log [(1 + R)/(1 + r)]).$$

Thus, if $I = 30,000$ and $C = 3000$, while $R = 0.40$ and $r = 0.06$, then $t$ is about 8.28 yr. If $I = 40,000$, $C = 3000$, $R = 0.20$, and $r = 0.10$, then $t = 30$. By early next century, an institution may have to spend as much per person on library support as on his salary, if present trends and practices continue. That, however, is not likely.

For some high quality journals, $M(q)$ is small. Hence $p_{\text{min}}(q)$ is high (e.g. Brain Research, now over $1600 per yr), and for some, $M$ is high (e.g. Science). On the whole, the above total expenditure cannot exceed the libraries’ budgets, for academic libraries, these have been 3–4% of the academic institutions’ overhead budget, but the fraction has been declining. While these budgets have also increased exponentially, the rate has been about 6% per yr, compared with increases in $p_{\text{min}}$ of more than 15% per yr. For the equation to balance, either $N_0$ or $M$ must decrease or $b$ must increase [12].

What is the most likely initial response to increasing $p_{\text{min}}$? Decreasing $M$ will increase $p_{\text{min}}$ even more. While publishers of very high-quality specialized publications might make the decision to publish even if $M(q)$ is lowered, and the specialists might pay the price they charge, that is not true for lower quality publications. Thus $M(q)$, averaged over $q$, might stay the same.

The percentage of publications that is of high quality may increase. We could also extend this model to argue that the total capital investment in the publishing industry is proportional to its average profitability and that this may decline with increasing prices, causing publishers who do not carry primarily high-quality publications to go out of business. Consequently, also, libraries would be stocked primarily with high-quality publications. This is a positive outcome of the present inflationary spiral.

As seems to be the case for libraries who have reduced all costs other than cancelling or not acquiring publications on the basis of price, we will shortly reach the slope where $b$, which measures sensitivity of demand to quality, increases. But $N_0$, the number of “zero-quality publications” is likely to decrease first. Here we face a difficulty. If the quality of a publication is the number of citations per item in it, the acquisitions librarian has no way of knowing this for a newly announced publication other than waiting 3–5 yr, waiting for reviews or relying on the estimates and judgments of potential readers and users. The publisher will, of course, have made such a judgment. Hence we may have at least two differing judgments for many publications, the librarian estimating (conservatively) it to be 0 quality, and the publisher/author estimating it (perhaps wishfully) to be of at least marginal utility. It could be argued that no one could or should prejudge the value of a work that is not yet done. (Yet this is done with proposals all the time, on the basis of the author’s knowledge, plan of attack, ideas, etc.) Risks must be taken, but those taking them would like to be as judiciously informed and as intelligent about them as possible.
The production of new knowledge has always been one of the riskiest of human enterprises. By ironic coincidence, it has also been the most productive and successful. For the modest investments in research and scholarship, by comparison with other expenditures of the global economic system throughout its history, the yield, in terms of discoveries, innovations and their economic impact, has been enormous. Moreover, it is an investment level that we are free to choose, unlike expenditures for defense, health, welfare, etc. that are increasingly forced upon us.

4. EFFECT OF PLAUSIBLE INTERVENTION

Suppose that libraries pool their limited resources. One device is to coordinate acquisitions. In a university library system this might work as follows. Assume, for the sake of argument, a university had a matrix organization in which the rows correspond to projects, tasks and enterprises requiring coordinated interdisciplinary team-work, and in which the columns correspond to classical academic disciplines. A given faculty member may belong to both a disciplinary department and a functional team. Assume that there is a small special library associated with nearly every department and every team. A faculty member upon learning of a new book or journal that he considers important, sends to both his team and his departmental library a notice listing author, title, publisher, date, price, likely readership and why it is important for them. The two small, special libraries decide if they should and could acquire the items; if not they transmit the notice to a library that is a level higher than their own, e.g. a university medical or a law library. They make a similar decision, transmitting the notice to a university-wide library, if appropriate. After that it may go to a state, a regional or a national library. Libraries at the same level may have formulated a collection policy where each library specializes in different areas, so that at least one or two copies are always somewhere available in the region, through interlibrary loans or copies on demand. Of course, every library in the system would be informed of the decision made by every library in response to the notice it receives.

A regional library consortium of this sort then could acquire more publications with a given budget after costs of maintaining the system are deducted. Of course quality of service is lowered, with the cost of delays in getting needed publications, of copying, mailing, etc. passed on to the user. Libraries that do the lending are not willing to absorb the costs, and even if it is charged to the end user, they will have to absorb the costs of accounting and billing. The anticipated impact of such resource sharing on publishers would force them to increase p even more and to charge heavily for copying, with the result that this is not a viable way of coping with the problem.

Recent studies of scholarly communication[13] have concluded that the system is basically healthy, though beset with numerous, non-fatal problems. The finding that librarians, book publishers and journal editors are much more concerned about the state of scholarly publication than the scholars themselves is not surprising; they may discuss their concern with the state and outlook of their particular enterprises as concern about the state of scholarly publication. Quite possibly that concern is justified, because prices for scholarly publications have climbed at almost 40% per yr, and library budgets by only 6% at best. Unit sales have in many cases fallen sharply[10].

Despite the widespread belief that conventional paper-form books and journals will be rendered technologically obsolete, some publishers and librarians are convinced that demand for publishing print-outs from electronically stored textual data bases lies in the future. Both libraries and existing publishing firms may be displaced by “middlemen”, playing the kind of role for articles that Lockheed is currently playing for bibliographic references. The user, possibly using a regional, national or multinational bibliographic system, specifies an article or item in a scholarly book he wants to read, either by keying in the request himself through a private terminal in his office, through an input device attached to his home TV or through the services of an intermediary whom he instructs by telephone, mail or face to face. The electronic request is processed by a system such as SDC which copies the computerized file of the needed item with the “publisher’s” permission, into a private computer file at the user’s site. He can then make a hard copy on his local printer or use a soft display at his convenience, and at his cost. The intermediary may automatically subtract the cost of this service to that user
from his account and add the amount due to the “publisher” for permission. This system may work.

To meet the needs of professionals for overviews and perspectives, the idea of “compacts” is being seriously considered. The International Federation for Information Processing is planning to access a textual database through scientific and technical data communications networks in various parts of the world. This is a joint venture with North-Holland, with the test phase being executed on the packet switching network EURONET. This will make available database publishing directly to professionals rather than to intermediaries, such as libraries. The user may be able to choose from the following options:

- The compact/abstract file of papers to be published in full later.
- The calendar of forthcoming events and meetings in the field.
- New product and systems information.
- The multi-lingual dictionary and mnemonics file.
- Short reports and news items from associated institutes and groups.
- A regular column “from the editor”.
- List of addresses of organizations, companies and groups in the field.

5. CONCLUSIONS

The most fundamental consideration in scholarly communication is that the growth of scholarly knowledge is proportional to the growth of the scholarly population[14]. Insofar as the world population growth curve is probably logistic—unless we expand into space colonies and maintain our present geometric growth pattern indefinitely—the growth curve of the scholarly population is logistic. The latter may continue to accelerate for some time after the former has begun to decelerate due to most of the world’s people catching up to the rest in levels of education. Unless technological advances amplify the productivity of individual creative scholars—a distinct possibility—the growth curve of scholarly knowledge is also likely to be logistic.

It is neither a need nor a possibility for every member in the scholarly population to communicate with every one else. Unless his ability to assimilate the contributions of others like his ability to add to, is amplified by technology, there is a fixed number of others any given scholar can fruitfully communicate with. As a result, scholars group themselves into communicating classes of fixed size. The number of such classes then grows in proportion to the number of scholars; for small periods of time the number of different groupings and their growth could be much larger.

Several such groupings or communicating classes may be aggregated into larger communities. Each class has a need for nearly simultaneous communication and for an organized cumulative record. The former is now served by meetings of professional societies and those scholarly journals that stress their function as “newsletters”[15], and despite rising costs, scholars have not yet complained loudly. The latter are now served by the more archival journals and libraries, augmented with selective dissemination and on-line bibliographic search services. Again, despite the rising costs and curtailed services (especially in availability of books), scholars have not been seriously hindered. Indeed, as pointed out in Section 3, the economic squeeze on both publishers and libraries is likely to increase the percentage of available publications that are of high quality, which scholars will welcome. The traditional publisher-library interaction, however may well cause both these institutions to be supplanted by the kind of systems and brokers suggested at the end of the last section. Indeed, computer conferencing has the potential for connecting scholars “directly” with one another—in the same sense that the telephone system does—while meeting both their needs for nearly simultaneously communication and for an organized cumulative record. It does the former by providing both synchronous and store-and-forward capabilities. It does the latter by on-line storage and retrieval capabilities. In principle, all functions now performed by publishers and libraries with paper technology could also be performed by suitably extended computer conferencing systems.

The new intermediaries will be the managers of computer conferencing systems, conference organizers and leaders, and the new firms at the intersection of the computer and communications industries (e.g. Tymnet, Telenet, National Post Offices, etc.). The key intellectual
problem concerns the formation of groupings of scholars to engage in conferences for limited periods of time and for longer periods according to dynamics describable by systems of non-linear differential equations for the number of specialists in a field and the amount of knowledge in that specialty. Such dynamical systems have characteristic equilibrium points. Those which are stable correspond to a specialty cluster what we have here called a communicating class. This is likely to occur spontaneously. Random fluctuations in the tight coherence or uniformity of adherence to a paradigm within such groups corresponds to changes in the parameters of the system of differential equations describing the dynamical system. This may shift the solutions sufficiently far from their original equilibria that correspond to the formation of wholly novel communicating classes or specialty clusters. If all scholars were using a global computer conferencing infrastructure, then such new groupings would show up as participants in a new conference, after these have settled. A new potential "invisible college" will have been formed, and with it the electronic equivalent of their professional society specialty interest group, specialized journals and books. If this occurs, then it is because the potential inherent in these emerging technologies make it less costly and more convenient for scholars to communicate, and because the growth of knowledge has also generated the know-how and the excitement for organizing and implementing in practical ways such innovations.

Were this to happen, then know-how for increasing scholars' productivities and abilities to assimilate the contributions of more peers that might arise from research in artificial intelligence could be transformed into innovations more readily.

Discontinuous changes in the libraries-publishers dynamics are likely to occur if and when the steadily decreasing costs and widespread use of computer conferencing or something equivalent fall, respectively below and above the steadily rising costs and circulation of publications based on conventional technologies. As indicated in Section 3 above, there are a number of cushioning factors like economy of scale (on both sides), subsidies, etc. which will prevent drastic changes in the current system as new technologies augment the current one.

The movement of both the classical partners to fill in the needs for intermediate services will be accelerated, and the possibility for the entry of new intermediaries will grow.

Some aspects of libraries are likely to play roles similar to those of research museums, while others will become part of conferencing management and organization. Scholarly communication is not likely to change discontinuously, but may gradually become more efficient and effective in accelerating the growth of scholarly knowledge. Hopefully, such communications will increasingly include those who could use scholarly knowledge to cope with the complex problems of our world and not only bring such knowledge to bear, but also to shape its growth which would help us to shape our future and enable our species to self-actualize its human potential.

Acknowledgment—The authors wish to thank Melvin Voigt, Donald King, Harry Oelgeschlager and Bernard Fry for their helpful discussions and to the National Science Foundation for the partial support of M. Kochen under Grant IST-78-16629 in preparing this paper.

REFERENCES


