Krzysztof Apt
Centrum voor Wiskunde en Informatica
Postbus 94079, 1090 GB Amsterdam
K.R.Apt@cwi.nl

Opinion Serials crisis

Towards free access to scientific literature

Internet has profoundly changed the way scientific research is carried out. This in turn has altered our attitudes concerning access to scientific information. In this article we review a number of recent initiatives in various fields of science. We argue that the scientists in general, and mathematicians in particular, should actively support a transition towards free access to scientific literature.

In the software world a bizarre dichotomy established itself. On the one hand we have the Microsoft products, such as Windows, Word, and Excel that are not free and the source code of which is not available. On the other hand we also have the Linux operating system and other so-called GNU software that is free and the source code of which is open.

In the world of scientific publishing a similar dichotomy starts to arise with much more ominous side effects. On the one side commercial publishers, whose scientific journals are by no means free and to which the electronic access is beyond the pocket of many libraries even in the most prosperous countries of the world, consolidate their gains. On the other side learned societies and various groups of scientists launch an increasing number of initiatives that provide free electronic access to scientific knowledge.

This increasing discrepancy and the looming changes on the horizon start to attract a growing attention of the public opinion. On April 5, 2001 the Nature magazine started a debate on the impact of the electronic access on the dissemination of scientific literature (see [19]) inviting all parties involved in these matters to express their views in 1,000 word long articles.

There is a lot at stake, both from the economic and from the scientific point of view. The scientists in general, and mathematicians in particular, should realize the enormous advantages the free electronic access to scientific literature brings with itself. Not only will it make our research more efficient and of better quality, but it will also help scientists from less prosperous countries *and* from less affluent universities in prosperous countries to access the very same knowledge we use in our research. By passively cooperating with the current publishing system we delay the impending changes, burden the society that funds our research with excessive costs of library subscriptions to expensive journals, and deny equal opportunity in access to scientific research.

We argue here that in reality, in contrast to what the commercial publishers tell us, volunteer work and public funds suffice to properly organize scientific knowledge on the internet, that free scientific publishing (FSP) is increasingly taking place across several scientific disciplines, and that the resistance to the way the commercial publishers provide access to the scientific information on the web is growing. In what follows we review some of these developments. Because of the scope we shall focus on the subject of mathematics, only occasionally venturing into other domains of science.

Internet conquers scientific world

Already in 1998 a Yale University librarian estimated that there were between 5,000 and 10,000 electronic journals as opposed to 325 in 1995 (see [26]). Three years later practically all scientific journals of importance are also available electronically. However, the internet access to the overwhelming majority of these journals is by no means

free. If the costs of this access were negligible, the bill could be easily picked up by any university or research institute library and the problem would be solved. Small, interested companies could share the costs if necessary. Unfortunately, the reality is somewhat different: it is often the conglomerates of large and rich university libraries that have to share the costs to afford an access to the portals built over the journal web sites.

An important role is played here by Reed Elsevier. After its recent acquisition of Harcourt it will control about 20% of the market in scientific journals by owning a staggering total of 1700 titles, as opposed to all university presses that own about 700. Now, Elsevier journals are among the most expensive ones. According to the information available on [1], in 2000, twenty out of the 25 most expensive journals were owned by Elsevier. In contrast, the journals owned by the university presses and learned societies are usually the least expensive.

In Mathematics, according to detailed studies carried out in 1997 (and updated in 2000) by Rob Kirby from the Mathematics Department of the University of Berkeley (see [22]) the price ratio between comparable quality journals can be 10:1, with the most expensive journals owned by Springer.[1] It should be added here that Springer books, in contrast to those of Elsevier (and of Kluwer), are usually very reasonably priced, though in Mathematics mostly when on sale.

The entry '"most expensive" Elsevier' yields in Google ([2]) about 900 revealing web pages about journal costs, crisis in scholarly publishing, journal cancellations etcetera. In response to this storm of critique about the pricing, Elsevier started to offer subscription contracts directly to whole universities trying to circumvent the grumbling librarians. These packaged deals lower the prices of individual journals, though it's not clear by how much because of their usual non-disclosure proviso's. On the other hand, these deals oblige the libraries to maintain their collection for a couple of years and a fixed annual increase of prices is built into them. Various deals offered by learned societies don't have such offensive clauses.

At the same time a growing number of recent initiatives aim at making the internet access to scientific literature free.

Los Alamos National Laboratory e-Print archives (LANL)

One of the first initiatives towards free scientific publishing on the internet came from the physicist Paul Ginsparg. In 1991 he founded the Los Alamos National Laboratory e-Print archive, [3] (LANL). It is a fully automated electronic archive and distribution server for research papers. Ten years later it includes sections on physics, mathematics, nonlinear sciences, and computer science (into which an earlier Computation and Language E-Print Archive was absorbed). In these archives one can freely search, browse, and post articles, or even electronic proceedings of a conference. Since LANL is freely accessible to everyone, it is a modern age global home page with a simple alert system thanks to which one can monitor postings in any of its areas. The archives contain about 150,000 articles in physics and 15,000 in mathematics with some two million visits per week. The impact of the LANL archives becomes apparent once one reviews the initiatives of the mathematicians in electronic publishing.

Mathematics journals, new style

In the past few years mathematicians came up with a number of important FSP journals. In all of them, papers appear a few days after acceptance.

The earliest electronic mathematics journals (EMJ) started in 1994 through the initiatives reported on the EMJ mailing list (see [4]). Eventually these journals entered The Electronic Library of Mathematics (ELibM) of the European Mathematical Society.

ELibM contains online journals, conference proceedings, and monographs in the field of mathematics and also collections of papers of William Rowan Hamilton and of Georg Friedrich Bernhard Riemann. All material is in electronic form and access is free. The website of ELibM, [5], provides an entry point to 47 journals (with 8 more entering currently) freely available electronically. About half of them also have a print edition. These electronic services are available worldwide from more than 40 mirror sites of ELibM that replicate the whole server contents.

In what follows, we discuss some of the FSP journals, focusing on their economic underpinnings and on the way they are organized.

In 1996 Ulf Rehmann from the University of Bieleveld established a high quality FSP mathematical journal Documenta Mathematica, [6]. In [25] Louis, Schneider and Rehmann provide an account of the costs of running the Documenta Mathematica based on their four years of experience with the journal. In their detailed analysis they consider

- the depreciation and maintenance costs for a PC,
- network costs for running the server,
- technical maintenance costs,

and conclude 'including hidden costs' with the revealing amount of 200 Euro per year. The refereeing and editing is done, as for other scientific journals, for free, and the production work like manuscript adaption etc. is mostly automated.

The creation of the LANL archives eventually inspired some mathematicians to create so-called overlay (and thus FSP) journals. The articles in such journals consist just of pointers to the LANL archives. In all three examples that we will mention the journals also exist in a printed form, in one bound volume per year available at a nominal price. The first overlay journal is the Advances in Theoretical and Mathematical Physics, [7]. It was founded in 1997 by the 1982 Fields medal winner S.-T. Yau.

Then in 1997 a group of prominent topologists founded another overlay journal *Geometry and Topology*, (*G&T*) [8]. One of the founders, Colin Rourke, offered this account of his experience of running the journal (see Birman [17]): "There were no secretarial or setting up costs. Computer costs for running a journal the size of *G&T* are negligible, given the fact that universities are already networked and provide good computing facilities for their staff. I estimate that the size of the Warwick Maths computing system is about four orders of magnitude greater than that needed to run *G&T*."

Encouraged by the success of *G&T* Joan Birman and her colleagues founded last year another overlay journal, *Algebraic & Geometric Topology* (see [9]) that is expected to compete for strong papers with expensive, commercial journals in the same area. As of May 8, 2001 also Annals of Mathematics became an overlay journal.

Just in case somebody might wonder whether the production of an FSP journal scales up. Rehmann also describes in [27] how he produced within a couple of months as an Extra Volume of Documenta Mathematica the 2400 pages (3 hardcover volumes) of the Proceedings of the 1998 International Mathematical Congress in Berlin (attended by some 3000 participants). He wrote to me: "I was offered 50.000 DM

1 In 1999 the price of *Annals of Mathematics* was about $.10 per page while that of *Inventiones Mathematicae* just under under $1.00 per page. In the meantime *Annals* became freely available electronically, see below.

for the production by the congress organizers — it turned out that I just needed 600 DM — just to go and visit the printer's facilities in Bavaria to make sure that he can do the job. Everything else was done by TEX, Unix, Perl and their friends."

The winds of change

No matter how many free or almost free scientific journals will be created in the near future, the current situation in scientific publishing will not drastically change as long as scientific journals with exorbitant subscription prices continue to exist. A number of initiatives aim at changing this situation.

Editorial boards

To start with some editorial boards decided to move from a commercial to a non-commercial publisher. Let us quote here two examples, both involving Dutch publishers.

Michael Rosenzweig of the University of Arizona started in 1986 a journal in the field of evolutionary ecology. The initial subscription price was $100 for the libraries. Twelve years later, during which the journals changed the hands twice and became a property of Kluwer, the subscription price for libraries became $800. After failed negotiations, the whole editorial board of the journal left Kluwer and started a new journal *Evolutionary Ecology Research* published privately with the accepted papers freely available on the web. The matter was widely reported in the press, including The New York Times (see [29]). The current subscription price for the libraries? Back to the past: $125.

Another example concerns the Journal of Logic Programming (JLP) of Elsevier. In the period from 1992 till 1998 its price per page increased by 102% and its subscription price for the libraries reached $936. This led the editors to negotiate a price decrease with the publisher. No compromise solution was found and 16 months later, in November 1999, the entire Editorial Board of 50 computer scientists left Elsevier. They founded a new journal, Theory and Practice of Logic Programming (TPLP), published by the Cambridge University Press as of January 2001 at the 60% lower price per page. An account of how this happened can be found in Birman [17].

This decision was awarded during the 91st annual conference of the Special Libraries Association held in June 2000 in Philadelphia. Maurice Bruynooghe, the past Editor-in-Chief of the JLP was invited to attend the conference and to receive the most prestigious award of the Physics-Astronomy-Mathematics Division for his (and in extension the whole Editorial Board of the JLP and the Association for Logic Programming, [10]) efforts to bring the price of the JLP down and to move to another publisher when the negotiations, failed (see [28]). This action of the editors was cited in New York Times (see [23]). TPLP is not an FSP journal but the Cambridge University Press allows the authors to post the accepted versions of their papers in Computing Research Repository (CoRR) that forms a part of the LANL archives.

Libraries

Then libraries, alarmed by the dramatic increase of the subscription prices to the scientific journals, founded in 1998 the Scholarly Publishing and Academic Resources Coalition, SPARC, [11]. SPARC encourages competition in the scholarly communications market by supporting

illustratie: Ryu Tajiri

creation of new, inexpensive scientific journals. By now more than 200 libraries joined SPARC. Recently SPARC produced a most informative Declaring Independence Handbook [18]. One can read in it that in mathematics and computer science the prices of scientific journals went up by 35% in the period 1996–2000. The Handbook is a step-by-step guide for editors of scientific journals that explains what options they have if they find their journal is too expensive. It also reports several new initiatives and cost-effective projects in scientific publishing that for the lack of space have to be omitted here.

Grassroot movements

Further, whole grassroot movements started that aim at radically different ways of sharing scientific information. One of them, the ELectronic Society for Social Scientists, [12], originated by the economists, aims at:

"the provision of electronic publications of high quality, wide diffusion and low cost for the direct benefit of the academic community."

This recent initiative is supported among others by the Scottish Confederation of University and Research Libraries, the Conference of Heads of University Departments of Economics which includes all economics departments in the United Kingdom and more than ninety scientists from USA, Canada, Germany, Belgium, Israel, United Kingdom and Switzerland. The organizers report over 1000 positive responses to this initiative.

Perhaps the most spectacular action was launched by the scientists in medicine and life sciences. The U.S. National Institutes of Health launched in 1999 a central repository, PubMed Central, whose mission is to provide a comprehensive electronic archive of the peer-reviewed literature in the biological sciences. Once it takes shape it will have a dramatic effect on the way research is done in these sciences, since it will make possible to search all papers present in the archives for relevant contents. The initiative has been massively endorsed by the scientists through the Public Library of Science project, [13]. We can read there:

"We support the establishment of an online public library that would provide the full contents of the published record of research and scholarly discourse in medicine and the life sciences research in a freely accessible, fully searchable, interlinked form."

In the open letter, signed by September 5th this year by 27,077 scientists from 170 countries (with 419 signatures by the Dutch scientists) the signatories from medicine and the life sciences pledge that "beginning in September, 2001, we will publish in, edit or review for, and personally subscribe to, only those scholarly and scientific journals that have agreed to grant unrestricted free distribution rights to any and all original research reports that they have published." This initiative was mentioned in numerous newspapers, including The Economist, The Guardian, Le Monde, de Volkskrant and NRC Handelsblad, and will likely have a dramatic impact on publishing in all fields of science.

Scientific progress through FSP

It is time that more scientists start appreciating what is at stake. The issue is who is to govern the access to the scientific literature: the scientists and public organizations or the commercial companies that treat our work as just any other commercial product. Scientific publishing can be a hugely lucrative business. In fact, the profit margin of Elsevier-Reed in the scientific sector is around 35% as opposed to an overage of 20% in all of its publishing interests (see [21]).

By submitting scientific papers to expensive journals we contribute to a system of knowledge distribution in which the universities and research laboratories are forced to buy back at a high cost products of their own creation and in which researchers from less prosperous universities are penalized.

In contrast, if we achieve free access to all scientific knowledge we shall serve the society that funds our research in the optimal way, minimizing administration and dependence on third parties, while maximizing access, search possibilities, availability, and knowledge dissemination, and promoting in this way scientific cooperation and equal opportunity. Any computer scientist can testify what a difference to research Citeseer ([14]) makes. It allows one to search for free through the computer science articles freely available on the web (over 5 million citations are compiled) for all those that refer *somewhere in their text* (and not only in title or abstract) for some keywords of importance. The costly ScienceDirect digital library of Elsevier offers far fewer search possibilities, with downloading limited to the articles published in the Elsevier journals to which one's library could afford to buy an access.

Free access to scientific knowledge is entirely in the spirit of the long standing tradition of the universities and research laboratories (with some exceptions within industry) to provide open access to the research results. Research in science has always involved a great deal of volunteer work. There is no reason to provide this voluntary work in the form of refereeing and editing scientific articles to commercial companies that aim at the maximization of their profits and the satisfaction of their shareholders instead of at the advancement of knowledge and a realization of a free access to it.

As a scientist who comes from Poland and who got his background in logic in sixties at the University of Wroc^{a}aw in Poland by studying an illegal photocopy of J.R. Shoenfield's book 'Mathematical Logic', I can testify that access to scientific knowledge is one of the many dramatic divides between poor and rich countries. Internet has created unprecedented opportunities to break some of these barriers. The less prosperous countries have increasing problems in funding their research, and that leads to their further isolation. Why should the mathematicians and physicists in the Czech Republic, Iran, Bulgaria, Cuba, Colombia, India, Russia, and Zambia have a more restrictive Internet access to mathematical papers than the Dutch ones? Now, from a recent article [20] in New York Times we can learn how the free access to the LANL archives has dramatically influenced the way research is conducted in scientific institutions in exactly these countries.

The economic analyses quoted here show that the financial aspects of creating high quality FSP journals are by now negligible and within the reach of any well-organized group of scientists. Subsidies and continuous funding will be needed to guarantee further progress through creation of better solutions for storing, filtering, using and distributing scientific knowledge. Nothing seems to suggest that public funds for these purposes will be inadequate.

Conclusions

The fact that nowadays most articles in computer science can and are fetched directly from the authors' personal pages using Google and Citeseer shows that the computer scientists have organized themselves outside of the electronic walls erected by the commercial publishers. In turn, the physicists embraced the LANL archives, the researchers in medicine and the life sciences expressed their support for the PubMed Central by endorsing in huge numbers the goals of the Public Library of Science project, and the economists launched the

ELectronic Society for Social Scientists. The trend is clear.

I argued in [16] that free scientific publishing can be achieved only if we change our attitudes towards scientific publishing. If we stop submitting papers to expensive journals, decline to be their editors and refuse to referee papers submitted to these journals, we plant the seeds of change. And if such 'extreme' measures seem to be radical, just demand to be paid adequately for all your refereeing work and use it to sponsor your students. And do voluntary work as of now for non-profit organizations only.

If no inexpensive, or preferably FSP, journal exists to which our paper could be submitted, perhaps a time has come to consult the Declaring Independence Handbook [18] and to ponder starting a new FSP journal. The Handbook mentions among other the project Euclid, [15], a joint initiative of the Cornell University Library and Duke University Press that focuses on non-profit journals in mathematics and statistic. Of course, not everybody would like to get involved in time consuming editorial work for a new journal. But ELibM could help here by posting information about the areas in which authors found that new FSP journals should be created. This might attract potential founding editors.

We all recognize the obvious advantages with which the Internet provided us but are reluctant to realize that we ourselves are still picking up the bill for our shortsighted actions. In Lawrence [24], by analyzing no less than 119,924 conference articles in computer science and related disciplines, the author found a clear correlation between the number of times an article is cited, and the probability that the article is freely available online. He concludes: "To maximize impact, minimize redundancy, and speed scientific progress, authors and publishers should aim to make research easy to access."

Here are, in contrast, the benefits offered by Elsevier to the authors publishing in their journals: 50 offprints per contribution free of charge, and 30% discount on all Elsevier Science books (that are typically three to five times more expensive than those published by the University Presses).

Acknowledgements

Information and insights provided by Ay Ong, Joan Birman, Maurice Bruynooghe, Ulf Rehmann and Laurent Siebenmann turned out to be most useful. Mark Peletier's helpful comments are appreciated.

References

1 http://db.arl.org/journals/FMPro

2 http://www.google.com

3 http://arXiv.org/

4 http://math.albany.edu:8800/hm/emj/

5 http://www-emis.cwi.nl/ELibM.html

6 http://www.mathematik.uni-bielefeld.de/documenta/

7 http://pascal.intlpress.com/journals/ATMP/

8 http://www.maths.warwick.ac.uk/gt/

9 http://www.maths.warwick.ac.uk/agt/

10 http://www.cwi.nl/projects/alp

11 http://www.arl.org/sparc

12 http://www.elsss.org.uk/

13 http://www.publiclibraryofscience.org

14 http://citeseer.nj.nec.com/cs

15 http://projecteuclid.org/

16 K.R. Apt, One More Revolution to Make: Free Scientific Publishing, *Communications of ACM*, Volume 44, Number 5, pages 25–28, 2001.
http://www.acm.org/pubs/citations/journals/cacm/2001-44-5/
http://xxx.lanl.gov/abs/cs.GL/0106022/

17 J. Birman, Scientific Publishing: A Mathematician's Viewpoint, *Notices of AMS*, Volume 47, Number 7, pages 770-774, August 2000.
http://www.ams.org/notices/200007/forum-birman.ps

18 Declaring Independence, A Guide To Creating Community-Controlled Science Journals, SPARC, January 2001. http://www.arl.org/sparc/DI/

19 'Future e-access to the primary literature', Nature, April 5, 2001.
http://www.nature.com/nature/debates/e-access/index.html

20 J. Glanz, 'Web Archive Opens a New Realm of Research', The New York Times, Science Section, May 1, 2001.
http://www.nytimes.com/2001/05/01/science/01ARCH.html

21 'Journal Wars', The Economist, May 10, 2001.

22 R. Kirby, http://www.math.berkeley.edu/ kirby/journals.html

23 D.D. Kirkpatrick, 'Librarians Unite Against Cost of Journals',
The New York Times, December 25, 2000.
http://www.nytimes.com/2000/12/25/business/25JOUR.html

24 S. Lawrence, 'Online or Invisible?', Nature, Volume 411, Number 6837, 2001, p. 521.
http://www.pnas.org/cgi/content/full/041601398v1

25 A.K. Louis, P. Schneider and U. Rehmann, DOCUMENTA MATHEMATICA, August 31, 1999.
http://www.mathematik.uni-bielefeld.de/~rehmann/bericht-eng.html

26 'Publishing, perishing, and peer review', The Economist, January 24, 1998.

27 U. Rehmann, 'Efficient Production of Mathematical Literature', A report given on the Workshop 'The Future of Mathematical Communication', Berkeley, December 3, 1999.
http://www.mathematik.uni-bielefeld.de/~rehmann/EP

28 SPARC e-news, May-June 2000, http://www.arl.org/sparc/

29 C. Kaesuk Yoon, The New York Times, 'Soaring Prices Spur a Revolt in Scientific Publishing', The New York Times, December 8, 1998.
http://www.nytimes.com/library/national/science/120898sci-journals.html