Reset reprinting of Sandia Laboratory Research Report SC-RR-71-0455, July 1971. See also the summarized version appearing in *Notices of the American Mathematical Society*, **18**:723-725, 1971. Reset 1995 with minor corrections by Jon Doyle (doyle@csc.ncsu.edu). This work was supported by the Atomic Energy Commission.

THE SUPER-CLASSICS OF MATHEMATICS

Joseph A. Schatz

July 1971

Contents

Introduction

This report consists of two parts. Part I is a list of super-classic mathematics papers. By definition, these are the papers which were cited 50 or more times in the material from which *A Mathematics Citation Index* [9] was compiled. The first section of Part II describes what this material is. The remainder of Part II consists of some comments about the collection of super-classics.

These papers were chosen because they are atypical (in the number of times they are cited). However, oddly enough, in many ways this collection resembles a random sample of citations. For this reason, an examination of the super-classics furnishes an example of the kind of information which can be obtained from citation counts.

I. The super-classics

1. Lars Ahlfors and Arne Beurling, Conformal invariants and function-theoretic null-sets, Acta Math. 83 (1950), 101-129. MR 12, #171.

2. Armand Borel, Sur la cohomologie des espaces fibrés principaux et des espaces homogènes de groupes de Lie compacts, Ann. of Math. (2) 57 (1953), 115-207. MR 14, #490.

3. E. Čech, On bicompact spaces, Ann. of Math. (2) 38 (1937), 823-844.

4. I. S. Cohen, On the structure and ideal theory of complete local rings, Trans. Amer. Math. Soc. 59 (1946) 54-106. MR 7, #509.

5. Jean Dieudonné et Laurent Schwartz, La dualité dans les espaces ($\mathcal{F}$) et ($\mathcal{LF}$), Ann. Inst. Fourier (Grenoble) 1 (1949), 61-101 (1950). MR 12, #417.

6. Samuel Eilenberg and Saunders Mac Lane, Cohomology theory in abstract groups. I, Ann. of Math. (2) 48 (1947), 51-78. MR 8, #367.

7. Lars Gårding, Dirichlet's problem for linear elliptic partial differential equations, Math. Scand. 1 (1953), 55-72. MR 16, #366.

8. I. M. Gel'fand, Normierte Ringe, Mat. Sb. 9 (51) (1941), 3-24. MR 3, #51.

9. K. Gödel, Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme. I, Monatsh. Math. Phys. 38 (1931), 173-198.

10. Alexandre Grothendieck, Sur quelques points d'algèbre homologique, Tôhoku Math. J. (2) 9 (1957), 119-221. MR 21 #1328.

11. Marshall Hall, Projective planes, Trans. Amer. Math. Soc. 54 (1943), 229-277. MR 5, #72.

12. P. Hall, A contribution to the theory of groups of prime-power order, Proc. London Math. Soc. (2) 36 (1933), 29-95.

13. Edwin Hewitt, Rings of real-valued continuous functions. I, Trans. Amer. Math. Soc. 64 (1948), 45-99. MR 10, #126.

14. Lars Hörmander, On the theory of general partial differential operators, Acta Math. 94 (1955), 161-248. MR 17, #853.

15. Kenkichi Iwasawa, On some types of topological groups, Ann. of Math. (2) 50 (1949), 507-558. MR 10, #679.

16. N. Jacobson, The radical and semi-simplicity for arbitrary rings, Amer. J. Math. 67 (1945), 300-320. MR 7, #2.

17. Shizuo Kakutani, Concrete representation of abstract (M)-spaces. (A characterization of the space of continuous functions.), Ann. of Math. (2) 42 (1941), 994-1024. MR 3, #205.

18. M. G. Kreĭn and M. A. Rutman, Linear operators leaving invariant a cone in a Banach space, Uspehi Mat. Nauk 3 (1948), no. 1 (23), 3-95; English transl., Amer. Math. Soc. Transl. no. 26 (1950); Reprint of translation, Amer. Math. Soc. Transl. (1) 10 (1962), 199-325. MR 10, #256; MR 12, #341.

19. J. Leray et J. Schauder, Topologie et équations fonctionnelles, Ann. Sci. École Norm. Sup. (3) 51 (1934), 45-78.

20. Edwin E. Moise, Affine structures in 3-manifolds. V. The triangulation theorem and Hauptvermutung, Ann. of Math. (2) 56 (1952), 96-114. MR 14, #72.

21. F. J. Murray and J. von Neumann, On rings of operators, Ann. of Math. (2) 37 (1936), 116-229.

22. Louis Nirenberg, Remarks on strongly elliptic partial differential equations, Comm. Pure Appl. Math. 8 (1955), 649-675. MR 17, #742.

23. D. Rees, On semi-groups, Proc. Cambridge Philos. Soc. 36 (1940), 387-400. MR 2, #127.

24. Jean-Pierre Serre, Homologie singulière des espaces fibrés. Applications, Ann. of Math. (2) 54 (1951), 425-505. MR 13, #574.

25. Jean-Pierre Serre, Groupes d'homotopie et classes de groupes abéliens, Ann. of Math. (2) 58 (1953), 258-294. MR 15, #548.

26. Jean-Pierre Serre, Faisceaux algébriques cohérents, Ann. of Math. (2) 61 (1955), 197-278. MR 16, #953.

27. A. H. Stone, Paracompactness and product spaces, Bull. Amer. Math. Soc. 54 (1948), 977-982. MR 10, #204.

28. M. H. Stone, The theory of representations for Boolean algebras, Trans. Amer. Math. Soc. 40 (1936), 37-111.

29. M. H. Stone, Applications of the theory of Boolean rings to general topology, Trans. Amer. Math. Soc. 41 (1937), 375-481.

30. René Thom, Espaces fibrés en sphères et carrés de Steenrod, Ann. Sci. École Norm. Sup. (3) 69 (1952), 109-182. MR 14, #1004.

31. René Thom, Quelques propriétés globales des variétés différentiables, Comment. Math. Helv. 28 (1954),17-86. MR 15, #890.

32. A. D. Wallace, The structure of topological semigroups, Bull. Amer. Math. Soc. 61 (1955), 95-112. MR 16, #796.

33. H. Weyl, Über gewöhnliche Differentialgleichungen mit Singularitäten und die zugehörigen Entwicklungen willkürlicher Funktionen, Math. Ann. 68 (1910), 220-269.

34. George W. Whitehead, A generalization of the Hopf invariant, Ann. of Math. (2) 51 (1950), 192-237. MR 12, #847.

35. J. H. C. Whitehead, Simplicial speces, nuclei and m-groups, Proc. London Math. Soc. (2) 45 (1939), 243-327.

36. J. H. C. Whitehead, Combinatorial homotopy. I, Bull. Amer. Math. Soc. 55 (1949), 213-245. MR 11, #48.

37. E. Witt, Treue Darstellung Liescher Ringe, J. Reine Angew. Math. 177 (1937), 152-160.

II. Discussion

1 SCIMP and *A Mathematics Citation Index*

SCIMP is an acronym for Sample Citation Index of Mathematics Papers. A citation index is a list of papers, with each item on the list followed by a list of the papers which have cited (referred to) the given item. SCIMP was compiled from the references in about 25,000 papers taken from 48 serial publications during the period 1950-1965. This constitutes about 5% of the total mathematics literature, about 10% of the mathematics literature published since *Mathematical Reviews* started, and about 20% of the periodical literature for the period covered.

A Mathematics Citation Index is the portion of SCIMP obtained by eliminating most doubtful and incomplete citations, and citations of books. It consists of more than 100,000 citations of more than 30,000 papers.

Section 6 contains some comments on how the super-classics were obtained from the SCIMP corpus. The choice of 50 as the number of citations to qualify as a super-classic is completely arbitrary. The number of papers with n citations varies smoothly with n. In this sense the super-classics are merely a statistical aberration.

2 Subjects

Perhaps the most striking feature of the list of super-classics is the over-representation of algebraic topology. In this connection one can observe the vagaries of subject classification. The last paper which I would classify as algebraic topology was reviewed in volume 16 of *Mathematical Reviews*. It was not until volume 17 that *Mathematical Reviews* had a heading for algebraic topology. Thus none of these papers was classified as algebraic topology by *Mathematical Reviews*. At the other extreme, I have the uncomfortable feeling that some of the papers which I would call algebraic topology should now be more properly classified under one of the newer headings in *Mathematical Reviews*. It is to avoid revealing this embarrassing consequence of an aging education that I have refrained from saying how many papers in algebraic topology I think there are. One is tempted to generalize and say that a subject classification scheme is a scheme which insures that the most interesting papers will rapidly become misclassified. Let me hasten to add that no criticism is intented of the magnificent work done by *Mathematical Reviews* in classifying papers. Rather it is more likely that the super-classics in algebraic topology stimulated enough interest that the number of papers being reviewed warranted adding a new heading.

The next largest representation of papers is in functional analysis. This is closely followed by algebra. Is this a fair representation of the major interests of mathematics during the period 1950-1965? It certainly is not a fair representation of the bulk of material published during this period. Can one detect in this list a bias against "hard" analysis and geometry, for instance?

3 Journals

The list of journals represented is even more strongly skewed than the list of subjects. The *Annals of Mathematics* accounts for 11 of the 37 papers. This astonishing showing is not entirely due to the *Annals'* concentration on algebraic topology; the *Annals* is also well represented by super-classics in functional analysis and algebra. It is interesting to note that the *Annals'* popularity in citations is not a new fact. In order to explain this a slight detour is necessary.

Librarians are perennially faced with the problem of deciding which journals to purchase. In 1927 P. L. K. Gross and E. M. Gross [5] recommended that citation counts be used to aid in making this decision. To this end, they examined 3633 citations from the 1926 volume of the *Journal of the American Chemical Society*. A similar study for mathematics using 2165 citations from a number of journals was published by E. S. Allen in 1929 [2]. The largest study of this kind was made by C. H. Brown in 1954 [3]. Brown studied 8 fields of science with mathematics represented by 3348 citations of serial publications [3, page 161]. Similar counts are not yet available from SCIMP.

Brown found that from 1929 to 1954, the *Annals of Mathematics* had gone from tenth to the first most cited journal [3, page 75]. Brown listed the *Transactions of the American Mathematical Society* as the second most cited mathematical journal. We also find it second with 5 super-classics. From this point on our list differs from Brown's. The *Bulletin of the American Mathematical Society* has 3 representatives. There are 2 representatives from *Acta Mathematica* and *Annales Scientifiques de l'École Normale Supérieure*. Finally, there are 14 journals with 1 representative each.

The fact that Browm found the *Annals of Mathematics* and the *Transactions of the American Mathematical Society* the first and second most "used" journals with relatively little effort may suggest that examining large bodies of citations to determine what "use" is being made of the mathematics literature is a needless bit of caution, but I do not believe that this is the case. In section 6 I shall give one reason why information derived from citation counts should be viewed with skepticism. In addition to this type of objection there is always the question of whether the method of drawing conclusions is "valid." L. M. Raisig [8] presents the case that the techniques used by Brown to conclude which journals are most valuable are not the most appropriate ones. An ASLIB study [1] using a completely different method of deciding which mathematical journals are most used lists the *Annals* as tied for eighth place and the *Transactions* not in the top ten.

4 People

J.-P. Serre has three papers on the list. There are three people with two papers on the list: M. H. Stone, R. Thom, and J. H. C. Whitehead. I found the names missing from the list more surprising than the names on the list. It is an interesting exercise to consider who you would expect to be on the list, and to speculate as to why they are missing.

Six of the super-classics were papers by two authors. There were no super-classics with three or more authors. As the table below shows, this is a fair sample of the SCIMP citations.

	Multiple Authors (percent)	3 or more authors (percent)
Super-classics	16.2	0
Journal citations in SCIMP dated 1938 or earlier	8.72	.45
Journal citations in SCIMP dated 1939 or later	15.2	1.19
Source papers in SCIMP	14.5	1.21
Book citations in SCIMP	16.9	1.74

These figures show the increasing incidence of multiple authorship with time that one would expect. If one is willing to read significance into $14.5 < 15.2 < 16.9$, one could conclude that multiple-authored papers are somewhat more popular than average.

5 The aura of the papers

Of the 37 super-classics, 10 (or 27.0%) were published before *Mathematical Reviews* started. It is this number which we discuss in this section. The question which we are hinting at, but which we shall not discuss, is: How much of the published literature need be saved?

A number of ways of attacking this question have been considered. One way is to estimate the "half-life" of the literature in a scientific field (see Burton and Kebler [4]); a half-life of x years means roughly that half of the citations given are of papers less than x years old. Another way is to estimate the size of the research front (see Price [7]); this is supposed to measure the greater likelihood of a recent paper being cited than of an older paper being cited. The counts of the SCIMP data from which these numbers can be computed are not completed, but the direction they are taking is clear. It is to this question that we turn.

Until the end of 1965 *Mathematical Reviews* had listed about 171,000 items. May [6] has estimated the number of items published before *Mathematical Reviews* started as 235,000. However, only 24.7% of the journal-like citations in the SCIMP corpus were to publications dated prior to the start of *Mathematical Reviews*. Thus 24.7% of the journal-like citations were to the earlier 58% of the literature. For the super-classics, the corresponding figures are: 664 citations to papers published before *Mathematical Reviews* started, and 1885 citations to papers published afterwards. Thus 26.1% of the citations of the super-classics were citations of papers published before *Mathematical Reviews* started. This close agreement with the SCAMP corpus seems at least mildly surprising.

These figures suggest that although mathematics has a substantial research front, it also has a healthy regard for at least some of the older periodical literature. This confirms the findings of Brown [3, page 161].

6 The processing

The 37 super-classics were cited 2549 times for an average of 68.9 citations per paper. However the descriptions of these 37 papers were given in 318 ways; this is an average of 8.60 descriptions per paper. Twelve of the papers had 10 or more different descriptions. If these citations had been

considered to be citations of 318 papers, we would have concluded that these papers were cited an average of 8.01 times.

Citation counts furnish information on the nature of scientific literature which can be obtained in no other way. The list of super-classics in Part I is an example. It would be even more useful to know, say, what percentage of the literature 20 or more years old will ever be used again. Citation counts give us a method of attacking questions like this. However, the numbers in the preceding paragraph show that, unless care is exercised, the information obtained can be considerably distorted. In particular, failure to "clump" the various descriptions of the same paper results in overestimates for the number of papers cited and underestimates for the average number of citations per paper. I feel that this is one of the major reasons why the small amount of information available on the use made of scientific literature should be viewed with considerable skepticism. Most of the reasons why a given paper is described in many ways are not hard to come by: different ways of referring to a given Journal, variations in an author's initials or name, difficulty of deciding which year a given volume of a Journal was issued, missing information, and incorrect information. The question is not whether these things happen, but how frequently.

The citations of the super-classics should be relatively free from these defects since often cited papers should tend to be cited in the same way. Nonetheless, if we had used only the most popular description of each of these papers, we would have lost 543 citations or 21.3% of the total: moreover, our list of super-classics would have been reduced from 37 to 21, a reduction of 43.2%; of the remaining 21 super-classics, only 3 would have pre-dated *Mathematical Reviews* severely distorting the age distribution of the super-classics.

A sizeable amount of information is available about these "normal" difficulties in "clumping' the citations in SCIMP. Because of the difficulties of discussing this data, it seems inappropriate to do so at this time. Let it suffice to say that it appears that about 15% of the citations contain incorrect or incomplete information.

In addition to the "normal" impediments to "clumping" there is a continuing supply of "unusual" difficulties. I shall close this discussion by presenting 3 examples which illustrate some of the problems involved in clumping.

Super-classic #5 by Dieudonné and Schwartz was cited 72 times in 21 different ways. The most common description was used 23 times. The journal was abbreviated in 6 different ways. The date was given in 1949 in 47 citations, as 1950 in 24 citations, and was not given in one. The volume was not given in 11 citations. No page description was given in 9 citations; page 16 was given in 1 citation.

Super-classic #18 by Kreĭn and Rutman has an "unusual" problem. However, it does not seem fair to consider this as two (or three) papers rather than one. There is also the additional difficulty of trying to describe the American Mathematical Society translation. Neither of the authors suffers from transliteration problems. They wind up with 60 citations using 24 descriptions; the most common is 13 citations of the Russian original.

Super-classic #22 by Nirenberg is another "unusual" problem. Although the pagination is as given above, at the top of the first page (649) it says that the pagination is 648-674. Of the 64 citations of this paper, 37 believed the paper started on page 648; 25 citations (and *Mathematical Reviews*) gave the correct page; 1 citation gave page 643; and 1 citation gave no pagination. There was also the usual sprinkling of differences (1 wrong volume, 7 citations with no initial).

The blunder in super-classic #22 was easily identified. Another story is in order to correct the impression that all blunders can be identified so easily. A paper by Fessenkoff was listed by

Mathematical Reviews on page 117 of volume 4 with the identification "19, 28-49 (1942)." An erratum on page 340 of volume 4 says that the "19" should be "20." Another erratum on page 328 of volume 5 says that "according to the final volume index the correct quotation is: vol. 19, No. 4, pp. 28-29 [sic] (1942). The pagination of this particular issue starts with p. 1." However, there is another erratum (page 855 of volume 10) which says that in the erratum in volume 5 read "20" for "19." Finally, the 20-volume Author Index gives the citation as "19, No. 4, 28-49 (1942)" and refers to the errata in volumes 4 and 5 but not the erratum in volume 10. The point of this story is not that *Mathematical Reviews* is careless, but that reference identification causes problems for even the best of us.

7 Further comments

Twenty-seven of the super-classics were reviewed by *Mathematical Reviews*. Most of these reviews give the reader no indication of the "exalted" nature of these papers. It is interesting to speculate whether this is more a result of the unpredictability of the future, or of the intentional blandness of the reviews in *Mathematical Reviews*. If the latter is the case, perhaps this blandness has been overdone. This comment is offered with considerable trepidation, since I feel that one should exercise extreme caution in changing a product as excellent as *Mathematical Reviews*.

It will not have failed to occur to the astute reader that if citation counts can be used to pick out the authors of the super-classics, then they can also be used to "rate" lesser mortals. I believe that citation counts are useful for determining characteristics of the literature of mathematics. However, I am unconvinced that any single "objective" technique is useful for rating individuals.

As has been stated, this list of super-classics was compiled from the SCIMP corpus. Because of this it will reflect any biases which are present in the journals used. It is possible, that a different selection of journals would result in a different list.

References

[1] ASLIB Research Department. Further minor by-products of research. In *ASLIB Proceedings 16*, pages 116–118, 1964.

[2] E. S. Allen. Periodicals for mathematicians. *Science*, 70:592–594, 1929.

[3] C. H. Brown. *Scientific Serials*, volume 16 of *ACRL Monographs*. Association of College and Reference Libraries, Chicago, 1956.

[4] R. E. Burton and R. W. Kebler. The "half-life" of some scientific and technical literatures. *American Documentation*, 11:18–22, 1960.

[5] P. L. K. Gross and E. M. Gross. College libraries and chemical education. *Science*, 66:385–389, 1927.

[6] Kenneth O. May. Quantitative growth of the mathematical literature. *Science*, 154:1672–1673, 1966. **MR** 40, #7070.

[7] Derek J. de Solla Price. Citation measures of hard science, soft science, technology, and non-science. In C. E. Nelson and D. K. Pollock, editors, *Communication among Scientists and Engineers*, pages 3–22. D. C. Heath and Company, Lexington, Massachusetts, 1970.

[8] L. M. Raisig. Mathematical evaluation of the scientific serial. *Science*, 131:1417–1419, 1960.

[9] Joseph A. Schatz. A mathematics citation index. Research Report SC-RR-70-910, Sandia Laboratories, 1970.