

How Integrated are Theoretical and Applied Physics?

Henrique F. de Arruda,¹ Cesar H. Comin,² and Luciano da F. Costa^{3,*}

¹*Institute of Mathematics and Computer Science,
University of São Paulo, São Carlos, SP, Brazil.*

²*Department of Computer Science, Federal University of São Carlos - São Carlos, SP, Brazil*

³*São Carlos Institute of Physics, University of São Paulo, São Carlos, SP, Brazil*

How well integrated are more theoretically and application oriented works in Physics currently? This interesting question, which has several relevant implications, has been approached mostly in a more subjective way. Recent concepts and methods from network science are used in the current work in order to develop a more principled, quantitative and objective approach to gauging the integration and centrality of more theoretical/applied journals within the APS journals database, represented as a directed and undirected citation network. The results suggest a surprising level of integration between more theoretically and application oriented journals, which are also characterized by remarkably similar centralities in the network.

I. INTRODUCTION

Science has unfolded on the interface between theory and applications. For instance, in classical Greece, more considerable attention was given to theoretical constructs, such as Plato's *Theory of Forms* [1]. Contrariwise, Roman antiquity concentrated on applications and engineering [2]. Ever since, science progressed in both theory and applications which, as we believe, are equally important and complement one another. From time to time questions arise regarding the current level of integration between these two orientations [3, 4]. Such questions, which are perhaps implied by the own division of Physics into theoretical and applied branches, are relevant because they have several implications [3]. For instance, this could involve in important theoretical results not reaching the application and technological level, or vice-versa. Though much of the discussion about the integration between theoretical and applied science has been conducted in a relatively informal and subjective way, we now have the means for approaching it in a more principled, quantitative and, therefore, more scientific way.

A common approach for quantifying the integration between theory and application is to analyze citation patterns between scientific publications and patents [5, 6]. Such an analysis can focus on how science influences technology [4, 6] as well as on the benefits that technological advancements can bring to science [7].

The compilation and dissemination of statistics about publications, such as the APS database, as well as other initiatives [8], can be efficiently mapped into intuitive representations, such as graphs or networks [9, 10]. These representations are capable of making explicit the connectivity and relationship between different scientific areas. Powerful methods have also become available that can be used to analyze these representations efficiently. The coming of age of these research possibilities is now

directly reflected in the area of *Scientometry* [11–13].

The current work reports precisely an attempt at investigating the integration between applied and theoretical science, especially from the perspective of Physics. More specifically, we consider the APS database [8], containing several journals with diverse focuses. A single network is constructed from this database, where each node represent an article (all available articles from all journals in the database), while the links stand for citations between these articles. Then, two specific journals were selected as presenting dominant focus on theory (Phys. Rev. D) and applications (Phys. Rev. Applied), according to the description of their respective scopes. Because the latter journal is very recent, and therefore includes less publications, we defined a second experiment in which Phys. Rev. Applied was substituted by *Phys. Rev. B*, which is older and has many more articles. The integration or proximity between the different journals was quantified in terms of the multiplicity of shortest paths between their respective body of articles (Figure 1a), as well as by the betweenness centrality [14] (Figure 1b) of the considered groups of articles (Phys. Rev. D, Phys. Rev. Applied, and control groups). The basic hypothesis here is that the existence of multiple shortest paths of comparable lengths between two sets of articles indicates, to a good extent, integration and relationship between these two sets. These measurements are complemented by the betweenness of a journal, which is capable of expressing how many of the shortest paths between pairs of articles from the overall database go through that journal.

The results obtained from the aforementioned investigation revealed that journals of more theoretical and applied nature are about as tightly interrelated as any two randomly chosen sets of articles with same size. These results suggest that theoretical and applied physics are therefore closely related and integrated. Though these results are specific to the adopted database, network representation and measurements, they nevertheless provide quantitative, objective insights into the relevant question of integration between applied and theoretical physics. It is hoped that further studies can be developed considering larger databases and other areas, so as to achieve

*Electronic address: ldfcosta@gmail.com

broader and more conclusive results.

This article starts by presenting the database, network representation and measurements, and then present and discuss the obtained results. Prospects for further investigations are also highlighted.

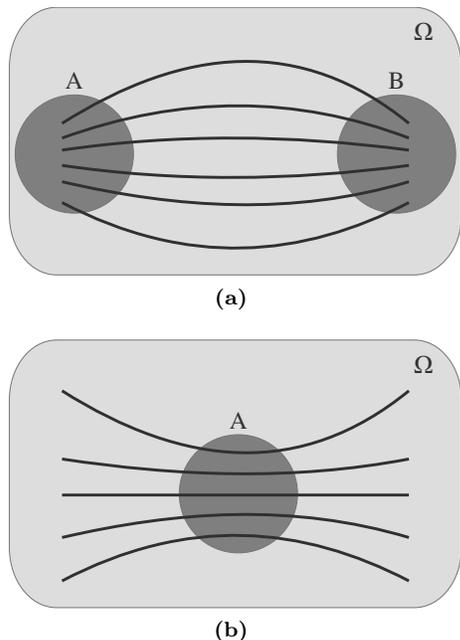


FIG. 1: The integration between two journals A and B can be quantified in terms of the multiplicity of shortest paths between them (a). The integration of a given journal A with the remainder journals can be further quantified by considering the betweenness centrality of its constituent nodes (b).

II. MATERIALS AND METHODS

In the APS database [8], each article is represented by its own Digital Object Identifier (DOI), and the citations are organized as a spreadsheet. For example, if a given article A cites B , the pair of DOIs for A and B are incorporated as a row into the table. Additionally, some meta-data are also included, such as the year of publication and the published journal. We mapped the citation table into a directed network, in which each article represents a node, and the edges correspond to the citations. The directions follow from the citing to the cited articles. Though this type of directed representation of the citations is inherently more related to the idea of scientific citation, we also consider *undirected* versions of these networks. The motivation for doing so is that when one article cites another, it is somehow implied that not only the citing article shares elements with the cited work, but that the reciprocal also takes place. Indeed, the directionality of citations is purely an implication of temporal causality, in the sense that only the later work can cite the earlier article.

TABLE I: The adopted groups of articles and some of their properties. The used dataset comprises articles from 1893 to 2016. Control (A) and Control (B) represent the Control group of articles when, respectively, Phys. Rev. Applied and Phys. Rev. B is employed.

Group	Number of articles	Years
Applied (Rev. App.)	703	2014–2016
Applied (Rev. B)	176,381	1970–2016
Theoretical (Rev. D)	80,427	1970–2016
Control (A)	509,325	1893–2016
Control (B)	333,647	1893–2016

Starting with this network, we divided the nodes (articles) into three groups: *applied physics*, *theoretical physics*, and *control*. The first two groups correspond to Phys. Rev. Applied (alternatively Phys. Rev. B) and Phys. Rev. D, respectively, and articles from the remaining journals were used to define control groups. The theoretical physics area was represented by the journal *Phys. Rev. D*, as this journal includes more academic-related areas, such as field theory, and quantum field theory. The applied physics articles were obtained from *Phys. Rev. Applied* because this journal typically publishes articles on applications, including engineering and technology. Because this journal is recent, and therefore includes a relatively small number of articles, we also considered a second experiment where this journal was substituted by *Phys. Rev. B*, which is arguably the next more application-oriented journal in the APS database. These two experiments are henceforth referred to as *Case Study #1* and *Case Study #2*. Observe that this choice is particular, and other possibilities could be taken into account, including alternative databases. Though the results reported in this article are specific to these choices, we believe them to be a reasonable indication of what could be found for other choices. An overview of the characteristics of the areas is shown in Table I. Note that the number of articles varies for each group, with the applied group being smaller as a consequence of its more recent introduction.

In order to gauge the distance between two different areas of study, we employed the distribution of minimal paths, which were computed using the *Dijkstra* algorithm [15]. Sets of nodes with the same size were randomly selected for the three groups, and the shortest path distances were calculated. The number of selected articles was chosen as 705, which is approximately the number of items in the smallest group. Note that we did not include in our statistics the pairs of nodes that do not have paths connecting them. We used the measurement $1/(\text{Path length})$, henceforth called *cohesion*, instead of the shortest paths so as to have a measurement that is directly related to the relationship between articles (the shortest path length would be *inversely* related, being less intuitive).

As an additional measurement, we considered the betweenness centrality measurement [14], which quantifies

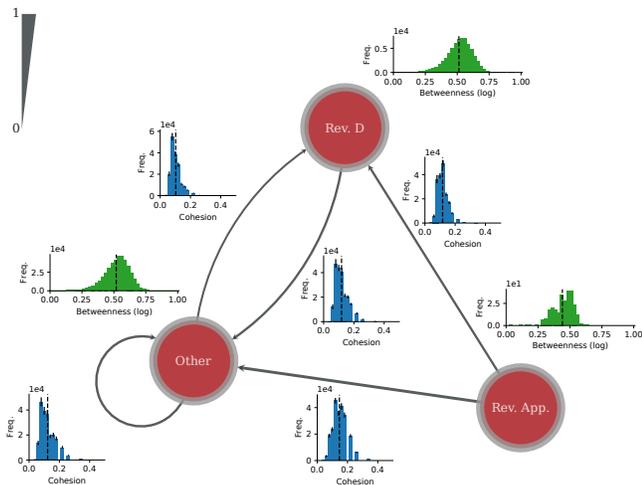


FIG. 2: Cohesion (blue), defined as the reciprocal of the shortest path length, and betweenness centrality (green) relationships between Phys. Rev. D, Phys. Rev. Applied and Other APS journals for the directed network. The legend at the upper right corner indicates the width range of the links. Observe that all links have similar width, suggesting symmetric relationships between all groups.

the centrality of network nodes according to the number of shortest paths to which they belong. This measurement is computed as follows

$$B_k = \sum_{ij} \frac{\sigma_{ij}^k}{\sigma_{ij}}, \quad (1)$$

where σ_{ij}^k is the number of shortest paths that connect the nodes i and j and crosses k , and σ_{ij} is the total number of shortest paths connecting i and j .

III. RESULTS AND DISCUSSION

The obtained results for directed and undirected versions of the article networks are reported and discussed respectively in the following two sections.

Case Study #1 - Phys. Rev. Applied

Figure 2 shows the results obtained for the directed version of the considered networks. For the sake of better integration and interpretation of the results, we combined the cohesion histograms with the digraph representing the relationships between the three considered cases, which include: Phys. Rev. D, Phys. Rev. Applied, and the other journals. The width of each link reflects the average of the respective histogram. The average values for each histogram are also identified. The histograms of betweenness centralities obtained for each of the three groups are shown in green.

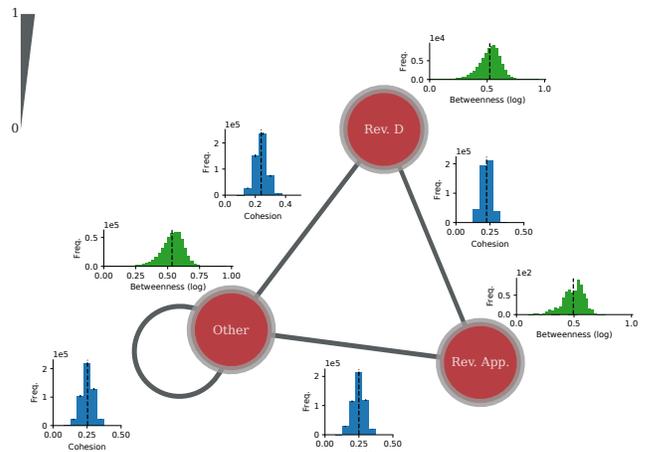


FIG. 3: Cohesion (blue) and betweenness centrality (green) relationships between Phys. Rev. D, Phys. Rev. Applied and Other APS journals for the undirected network. The legend at the upper right corner indicates the width range of the links.

Though, in principle, all directed interconnections could be expected to appear in this digraph, some of these connections that were too small (e.g. between *other* and *applied*) have been discarded. Generally, there is little difference between the width of the links in the obtained digraph, which indicates a good overall uniformity of connections between the three considered groups. This result suggests a good integration between the more theoretical and applied articles in Physics.

Regarding the betweenness centrality of the nodes, which is also shown in Figure 2, similar results were again obtained, with the *Applied* group having a density less symmetric than the other cases, which is a possible consequence of the smaller size of this group. In addition, the average betweenness centrality observed for this group is slightly smaller than those of the other two groups.

The results for the undirected versions of the considered networks are presented in Figure 3, together with the respective cohesion histograms. The interconnections between the three groups resulted again very similar.

Case Study #2 - Phys. Rev. B

Figure 4 shows the results for the directed version of the APS network. For this case study, the three considered groups are: Phys. Rev. B, Phys. Rev. D, and the remaining journals. As before, the width of each link reflects the average of the respective histogram. Observe that, respectively to the previous results, more links appear in this figure as a consequence of the larger number of articles in Phys. Rev. B. The cohesion histograms are similar for all cases, indicating that citations between journals have no preferred direction. The betweenness

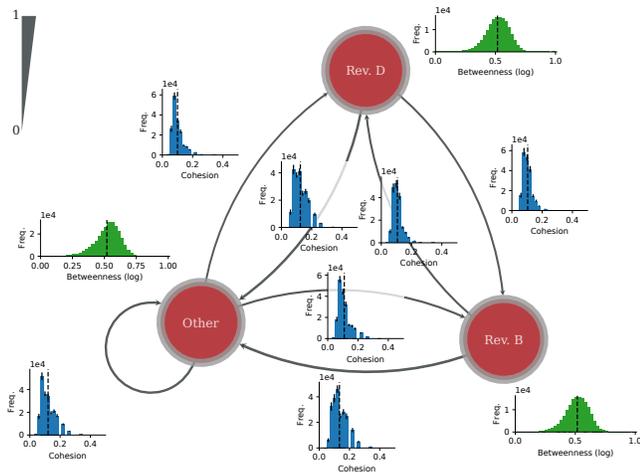


FIG. 4: Cohesion (blue) and betweenness centrality (green) relationships between Phys. Rev. D, Phys. Rev. B and Other APS journals for the directed network. The legend at the upper right corner indicates the width range of the links.

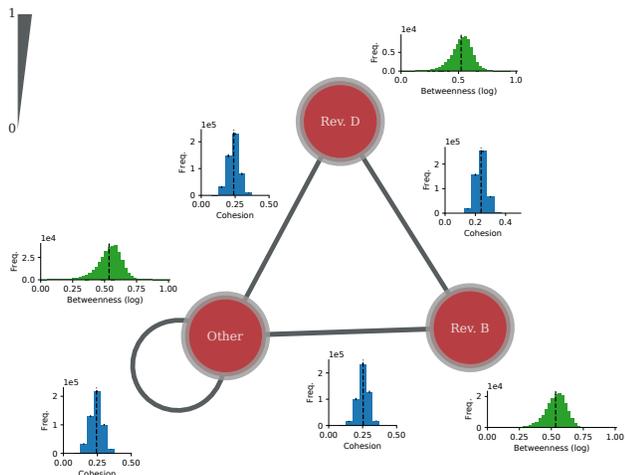


FIG. 5: Cohesion (blue) and betweenness centrality (green) relationships between Phys. Rev. D, Phys. Rev. Applied and Other APS journals for the directed network. The legend at the upper right corner indicates the width range of the links.

histograms for the groups are also similar. These results are in agreement with those obtained for case study #1, therefore conforming a good integration between theoretical and applied articles.

The results for the undirected version of the APS network is shown in Figure 5. Again, the cohesion and betweenness histograms are similar in all considered groups.

IV. CONCLUSION

Since the beginnings of science, eventual inclinations toward theoretical or applied approaches have been subject of considerable debate. Such discussions are, in principle, justified by the desire to understand how science changes along time, and also because such eventual biases may have effects in the relationship between science and technology. Though much of these discussions have taken place in a more subjective way, the advent of network science paved the way to more principled, quantitative, and therefore objective approaches to the aforementioned problem. The current work reports a related attempt. More specifically, we applied the network science concepts of shortest path and betweenness centrality to quantify interrelationships between journals in networks derived from the interesting APS journal database. We considered pairwise relations involving potentially more theoretical (Phys. Rev. D) and applied (Phys. Rev. Applied and Phys. Rev. B) APS journals.

The obtained results indicates a substantial symmetry of relationship and centrality of all considered journals, suggesting that both the more theoretically and applied inclined journals tend to have similar roles and be similarly integrated within the body of APS publications. So, there seems, at least along the years covered in the APS database, a surprising integration and relationship between more theoretical and applied studies and results in the area of Physics. We expect that this finding extends potentially to other publications and areas. Further investigation is however required considering other databases, other types of relationships, other types of measurements, among other possibilities. It would also be interesting to verify how these results extend to other areas and even science as a whole. The proposed methodology is also not limited to theoretical/applied orientations of a given area, as it can also be applied to quantify the interrelationship and integration between distinct areas such as Physics and Biology, etc.

Acknowledgments

Henrique F. de Arruda acknowledges CAPES for sponsorship. Cesar H. Comin thanks FAPESP (grant no. 15/18942-8) for financial support. Luciano da F. Costa thanks CNPq (grant no. 307333/2013-2) and NAP-PRP-USP for sponsorship. This work has been supported also by FAPESP grants 11/50761-2 and 2015/22308-2.

[1] G. Fine, *On ideas: Aristotle's criticism of Plato's theory of forms* (Clarendon Press, 1993).

[2] D. Hill, *A history of engineering in classical and medieval times* (Routledge, 2013).

- [3] H. Brooks, *Research policy* **23**, 477 (1994).
- [4] M. Meyer, K. Debackere, and W. Glänzel, *Scientometrics* **85**, 527 (2010).
- [5] R. Tussen, R. Buter, and T. Van Leeuwen, *Scientometrics* **47**, 389 (2000).
- [6] N. Shibata, Y. Kajikawa, and I. Sakata, *Technological Forecasting and Social Change* **77**, 1147 (2010).
- [7] D. C. Gazis, *Research Policy* **8**, 244 (1979).
- [8] *Aps data sets for research* (2017), <https://journals.aps.org/datasets> (accessed December 14, 2017).
- [9] L. d. F. Costa, F. A. Rodrigues, G. Travieso, and P. R. Villas Boas, *Advances in physics* **56**, 167 (2007).
- [10] L. d. F. Costa, O. N. Oliveira Jr, G. Travieso, F. A. Rodrigues, P. R. Villas Boas, L. Antiqueira, M. P. Viana, and L. E. Correa Rocha, *Advances in Physics* **60**, 329 (2011).
- [11] A. Bonilla-Calero, *Library Review* **57**, 700 (2008).
- [12] G. LaRowe, S. Ambre, J. Burgoon, W. Ke, and K. Börner, *Scientometrics* **79**, 219 (2008).
- [13] D. R. Amancio, M. d. G. V. Nunes, O. Oliveira, and L. d. F. Costa, *Scientometrics* **91**, 827 (2012).
- [14] L. C. Freeman, *Sociometry* pp. 35–41 (1977).
- [15] E. W. Dijkstra, *Numerische mathematik* **1**, 269 (1959).