Community Structure of the Physical Review Citation Network

P. Chen and S. Redner
Center for Polymer Studies, and Department of Physics, Boston University, Boston, MA, 02215

We investigate the community structure of physics subfields in the citation network of all Physical Review publications between 1893 and August 2007. We focus on well-cited publications (those receiving more than 100 citations), and apply modularity maximization to uncover major communities that correspond to clearly-identifiable subfields of physics. While most of the links between communities connect those with obvious intellectual overlap, there sometimes exist unexpected connections between disparate fields due to the development of a widely-applicable theoretical technique or by cross fertilization between theory and experiment. We also examine communities decade by decade and also uncover a small number of significant links between communities that are widely separated in time.

PACS numbers: 02.50.Ey, 05.40.-a, 05.50.+q, 89.65.-s

I. INTRODUCTION

In this work, we study the community structure of citations within the Physical Review (PR) family of journals from its inception in 1893 until August 2007. The journal consisted of just the Physical Review through 1969 and then split into 4 branches in 1970: Physical Review A (PRA): atomic, molecular, optical physics; Physical Review B (PRB): solid-state, condensed-matter physics; Physical Review C (PRC) nuclear physics; Physical Review D (PRD): particle physics. In 1990, there was yet another split of Physical Review E (PRE): statistical physics from PRA. The journal also includes a letters section, Physical Review Letters (PRL), that was introduced in 1958, a review section, Reviews of Modern Physics (RMP), that was introduced in 1929, and two recent special topics journals. For PR publications, we ask: can articles be naturally grouped into distinct subfields, with a high density of citations among papers within a given subfield and sparser citations across subfields? By the very nature of physics research and also as revealed by the data, this partitioning into subfields is self-evident. While anecdotal information exists about the identity and evolution of some of the more prominent subfields of physics [1], here we determine their quantitative properties, such as their size, time history, and citation impact. A related work recently studied the evolution of scientific fields through the PACS (Physics and Astronomy Classification Scheme) numbers of each article [2].

There are a variety of compelling reasons for studying the community structure in complex networks. In social networks, the partitioning of acquaintances into communities represents a basic fact about human interactions [3]. In metabolic networks, community structure may help identify basic reaction modules [4]. In the web, community structure reveals connections to web pages on related topics [5]. For the citation network, its underlying community structure may help us understand both the obvious and the more subtle interrelations between subfields, as well as the growth and the ebb of subfields.

To identify communities within networks, a variety of methods have been developed, such as the Kernighan-Lin algorithm [6], spectral partitioning [7, 8], and hierarchical clustering [3, 9]. While these methods sometimes work reasonably well, they can fail to identify communities when applied to networks that lie outside of the domain of their immediate application [10]. More recent work has led to the formulation of new and powerful methods to detect communities in complex networks, both with undirected [2, 11–15] and directed [15] links. A systematic review of these developments is given in [13]. One particularly useful approach exploits the concept of modularity [16]. Compared to the earlier community detection methods, the use of this metric to identify communities requires no extra knowledge beyond the network structure itself, involves no subjective judgments, and can be applied to any type of network.

In the next section, we outline the modularity maximization approach that we use to resolve community structure. We test the robustness of this method by using an alternative bottom-up network community partitioning algorithm [17]. We also check the significance of our results by applying community detection to a randomized version of the citation network. In Sec. III, we apply this algorithm to the Physical Review citation network to resolve its major communities and the connections between them. In Sec. IV, we study the structure of individual communities. In Sec. V, we partition PR publications into eight decadal sets by year of publication: 1927–1936, 1937–1946, . . ., 1997–2006 to study the time evolution of physics fields. Finally, we summarize in Sec. VI

II. COMMUNITY DETECTION BY MODULARITY MAXIMIZATION

For detecting communities within networks, we want to determine sets of vertices that are more strongly connected to each other but less connected to the rest of the network. For this purpose, we use the modularity Q [23], that is defined by

$$Q = \frac{1}{2L} \sum_{i,j} \left[A_{i,j} - \frac{k_i k_j}{2L} \right] \delta(i,j). \tag{1}$$

Here $A_{i,j}$ is the ij^{th} element of the adjacency matrix ($A_{i,j} = 1$ if a link exists between i and j and ($A_{i,j} = 0$ otherwise), L is the total number of network links, k_i is the degree of node i, and $\delta(i,j)$ equals 1 if i and j belong to the same group, and equals 0 otherwise. The modularity Q gives the difference between the number of links between groups in the actual network and the expected number of links between these same groups in an equivalent random network with the same link density. A modularity $Q = 0$ corresponds to a random network, in which two nodes are connected with probability that is proportional to their respective degrees. Empirical data indicates that a modularity value $Q \gtrsim 0.3$ is indicative of true community structure [24, 25], and the largest modularity that has been observed in real-world examples is 0.7 [16]. Thus we use modularity maximization as the criterion to divide a network into communities.

For large networks it is computationally impractical to maximize the modularity over all possible partitions of the network and one must resort to approximate methods [26]. Here we apply the eigenvector approach of Newman [16]. Suppose that the network contains N nodes and L links. We first focus on dividing the network into two communities, and then generalize to an arbitrary number of groups. Denote the two communities as 1 and 2, and let $s_i = 1$ if node i belongs to group 1 and $s_i = -1$ if i belongs to 2. Eq. (1) can be rewritten as

$$\begin{aligned} Q &= \frac{1}{2L} \sum_{i,j} \left[A_{i,j} - \frac{k_i k_j}{2L} \right] \frac{1}{2}(1 + s_i s_j) \\ &= \frac{1}{4L} \sum_{i,j} \left[A_{i,j} - \frac{k_i k_j}{2L} \right] s_i s_j \equiv \frac{1}{4L} \, \mathbf{s^T B s} \, . \end{aligned} \tag{2}$$

In going to the second line, we use $\sum_{i,j} A_{i,j} = 2L$, so that $\sum_{i,j} \left[A_{i,j} - \frac{k_i k_j}{2L} \right] = 0$. In the last line of Eq. (2), $\mathbf{s}$ is the vector whose elements are the s_i, and $\mathbf{B}$ is the symmetric modularity matrix with elements

$$B_{i,j} = A_{i,j} - \frac{k_i k_j}{2L} \; .$$

The sum of each row and column of $\mathbf{B}$ equals zero, so that $(1, 1, 1, \ldots, 1)$ is necessarily an eigenvector of this matrix with zero eigenvalue. Let $\mathbf{u}_i$ be the complete orthonormal set of eigenvectors of $\mathbf{B}$. We can then write $\mathbf{s} = \sum_i a_i \mathbf{u}_i$, with $a_i = \mathbf{u}_i^T \cdot \mathbf{s}$, so that the modularity becomes

$$Q = \sum_i a_i \mathbf{u}_i^T \; \mathbf{B} \; \sum_j a_j \mathbf{u}_j = \sum_i (\mathbf{u}_i^T \cdot \mathbf{s})^2 \beta_i \; . \tag{3}$$

Here β_i is the (ordered) eigenvalue of $\mathbf{B}$ that corresponds to the eigenvector $\mathbf{u}_i$,with $\beta_1 \geq \beta_2 \geq \ldots \geq \beta_n$. To maximize the modularity, Eq. (3) suggests that we choose $\mathbf{s}$ to be parallel to $\mathbf{u}_1$.

For the citation network, we need to extend the above approach to directed networks, in which all links are directed by virtue of publications being able to cite in the past. For this purpose, we use a generalization, also developed by Newman [27], in which the modularity is now given by

$$Q = \frac{1}{2L} \sum_{i,j} \left[A_{i,j} - \frac{k_i^{\text{in}} k_j^{\text{out}}}{L} \right] (1 + s_i s_j) \equiv \frac{1}{2L} \sum_{i,j} s_i B_{ij}^{(d)} s_j \; , \tag{4}$$

where the elements of the directed modularity matrix $\mathbf{B}^{(d)}$ are:

$$B_{ij}^{(d)} = A_{ij} - \frac{k_i^{\text{in}} k_j^{\text{out}}}{L}$$

We thus maximize the modularity in Eq. (3) using the eigenvalues and eigenvectors of the directed modularity matrix.

Since the value of each s_i can only be 1 or -1 by definition, $\mathbf{s}$ cannot be parallel to $\mathbf{u}_1$ except for special conditions. A natural alternative is to choose the sign of s_i according to the sign of its corresponding element in $\mathbf{u}_1$. For example, if the elements $n_1, n_2, \ldots, n_j$ of $\mathbf{u}_1$ are negative while all rest are positive, we choose $s_{n_1} = s_{n_2} = \ldots s_{n_j} = -1$ and all other s_i equal to $+1$. This partitioning then also determines the sizes of the two groups. As long as the largest eigenvalue is positive, there is a meaningful division of the network into subgroups. However, when the largest eigenvalue is zero, which corresponds to the eigenvector $(1, 1, 1, \ldots, 1)$, the division is trivial — all nodes are partitioned into one group and none in the other. This point defines a natural stopping criterion for the algorithm.

The steps to detect communities in a directed network at some intermediate stage of division therefore are:

1. Calculate the modularity for a subgroup.
2. If the leading eigenvector is negative or zero, the subgroup is indivisible.
3. If the leading eigenvector is positive, calculate the modularity of the entire network, assuming that the division is applied.
4. If the global modularity increases, perform the division. If not, abandon the division, mark this subgroup as indivisible, and process another divisible subgroup.

Repeat steps 1–4 until Q reaches its maximum (or equivalently, all subgroups are as indivisible).

To check the robustness of the results, we also apply a bottom-up algorithm that was recently introduced by Blondel et al. [17]. Here each network node is initially assigned to a distinct community. Then, for each node i, the gain/loss of the modularity is calculated after assigning i to be in the community of one of its neighbors. Node i is then assigned to the community that maximizes the increase of the modularity. If there is no increase, then node i remains in its original community. A single pass through all network nodes defines the first stage of joinings. This joining step is iterated for each community in the network and continues, stage by stage, until a maximal modularity is achieved. This algorithm is computationally more efficient than the previous top-down approach, but has the disadvantage of requiring considerably more computer memory. Both algorithms successfully detect the same communities on simple artificial networks, such as a collection of complete graphs that are each weakly connected to each other. We also compared the results of the two algorithms when applied to the PR citation data. We find that approximately 87.4% of the nodes are assigned to same community by both algorithms. This result gives a sense of the resolution with which we can define communities.

III. COMMUNITIES IN THE PHYSICAL REVIEW CITATION NETWORK

Our PR citation network consists of 433,452 articles published between 1893 through August 2007 with at least one citation (the nodes) and 4,370,203 total citations among these publications (the links). To keep the scope manageable we restrict ourselves to well-cited PR publications, defined as those with more than 100 citations. This restriction reduces the network to $N = 2,920$ publications and $L = 11,749$ citations. All citations to publications outside this highly-cited set and citations from "external" PR publications to this highly-cited set are excluded. The mean degree for this subnetwork is $2L/N \approx 8.05$. This small value — smaller than the mean degree for the entire network — stems from most of the citations to highly-cited publications coming from articles outside this group. We also exclude RMP from the dataset because this journal is devoted to review articles and the citation pattern from such articles is quite different from other branches of PR. Generally, RMP papers contain many more citations and a much broader range of citations that all other PR publications. The effect of RMP articles is therefore to enhance citations between papers in disparate fields. For this reason, we exclude RMP when analyzing the community structure of citations.

TABLE I: Top-10 cited communities in the PR citation network. Here F denotes the fraction of the highly-cited network comprised by each community and N is the number publications in each community.

rank	F	N	subject
1	6.54%	191	Elementary particles
2	5.99%	175	Correlated electrons
3	5.86%	171	Quantum information
4	5.03%	147	Theories of high-Tc and type-II superconductors
5	4.97%	145	Quantum diffusion, quantum Hall effect, two-dimensional melting
6	4.59%	134	Statistical physics, Monte Carlo method, gauge theory, quarks
7	4.32%	126	High-Tc oxides
8	4.04%	118	Theories of superconductivity
9	2.95%	86	Strong-coupling theory of superconductivity
10	2.60%	76	Semiconductors and quantum dots

As modularity maximization is applied, the highly-cited (as defined above) PR citation network divides into subgroups and the process stops when the modularity no longer increases when additional division attempts are made.

The evolution of the initial citation network during these first few steps is illustrated in Fig. 1. After the network has been divided once, it is, in principle, arbitrary which subnetwork to divide next. We always divide the subnetwork that leads to the largest increase in the overall modularity.

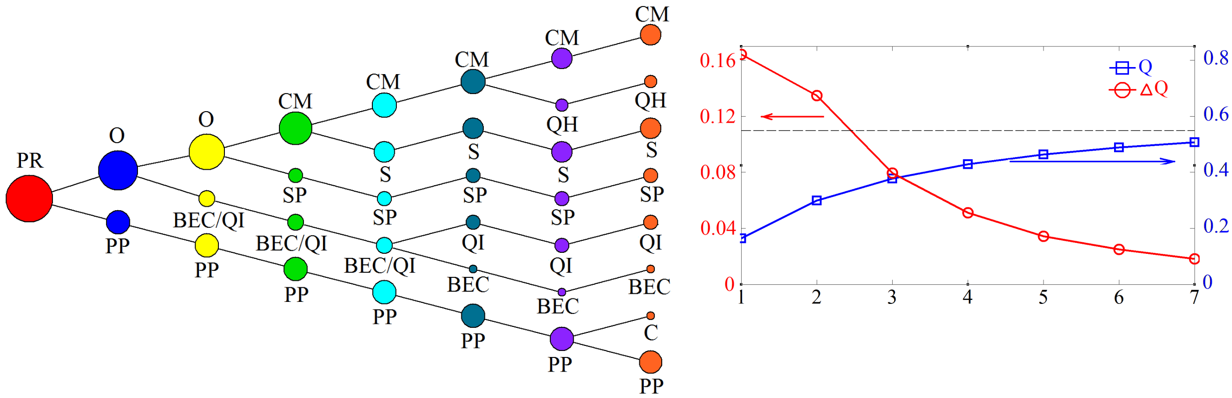

FIG. 1: (left) Evolution of the citation network after the first few division events. The labels on this branching tree have the following correspondences: (O) other, (CM) condensed-matter physics, (QH) quantum Hall effect, (S) superconductivity, (SP) statistical physics (QI) quantum information theory, (BEC) Bose-Einstein condensation, (C) cosmology, (PP) particle physics. (right) The modularity Q (squares) and the change of modularity ΔQ (circles) in the first 7 division steps, with different scales for ΔQ (left) and Q (right). The dotted line is the final modularity.

At the end of the modularity maximization procedure, there are 274 distinct communities, and the network modularity is $Q = 0.543$ ($Q = 0.514$ if RMP papers are included). For these 274 communities, the largest has 191 members (publications) and the smallest has only a single member. The 10 largest communities (listed in Table I) contain 1369 publications and comprise 46.9% of the highly-cited subnetwork. Figure 2 shows the 61 communities that consist of more than 5 publications; these are labeled numerically corresponding to the titles of the communities that are listed in Table VII in the appendix. We manually organized these communities so that publications in the eight major groupings that emerge after the first seven division steps of the PR network (Fig. 2) are spatially clustered.

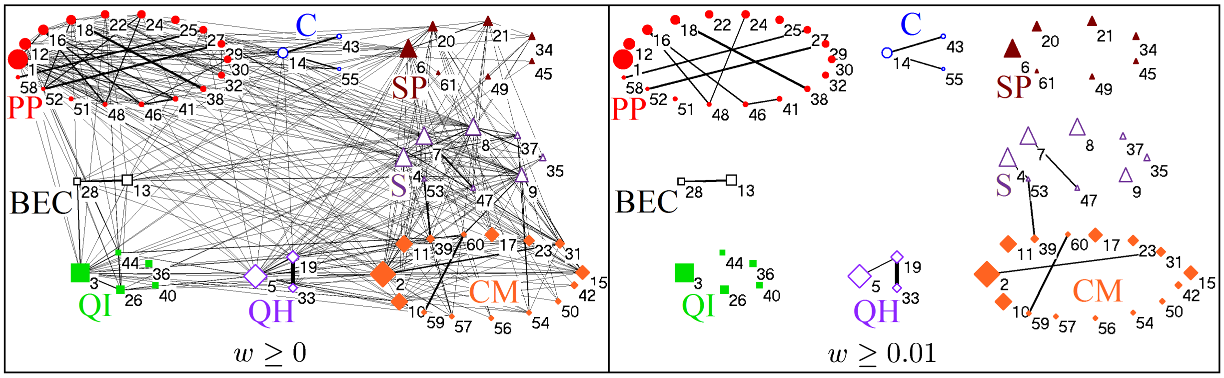

FIG. 2: The 61 highly-cited citation communities with more than 5 publications. Symbols refer to the major groupings: particle physics (dots), cosmology (circles), statistical physics (solid triangles), superconductivity (open triangles), condensed-matter physics (solid diamonds), quantum Hall effect (open diamonds), quantum information (closed squares), and Bose-Einstein condensation (open squares). The numerical labels correspond to the communities listed in table VII. The bond thickness indicates the intensity of citations between communities. The left panel shows all inter-community links, while the right panel shows only links with weight $w \geq 0.01$ (see text).

To indicate the relative importance of these communities, the size of each symbol in Fig. 2 is proportional to the number of publications within this community, while the width of each link is proportional to its relative weight. The weight w_{ij} of link ij is defined as $w_{ij} \equiv k_{ij}/(n_i n_j)$, where k_{ij} is the number of citations between communities i and j, and n_i is the number of nodes in the $i^{\rm th}$ community. Thus w_{ij} equals the fraction of all the potential $n_i(n_i-1)/2$ links between communities that actually exist. The average value of the link weight for the 61 largest communities is 0.206, while the strongest crosslink between communities — with weight $w_{ij} = 0.056$ — occurs between communities 19 (fractional quantum Hall effect, theory) and 33 (fractional quantum Hall effect, experiment). From the perspective of modularity maximization, it is almost immaterial whether these two communities are considered as separate or unified; joining them decreases the modularity by only 1.36×10^{-5}.

By the nature of the partitioning into communities, the links that remain between communities at the end of modularity maximization should be weak. In fact, only 17 out of 393 crosslinks have a weight that exceeds 0.01 (right side of Fig. 2). Moreover, 16 out of these remaining crosslinks join communities within the same major groupings that emerged in the initial few division steps. The only crosslink between different groupings joins communities 39 (charge density waves) and 53 (Hubbard model) and it consists of 3 citations. These connections arise from the mathematical similarity between one-dimensional charge-density wave systems and domain wall motion in the hall-filled Hubbard model. Specifically, the equation of motion used in “Incommensurate antiferromagnetism in the two-dimensional Hubbard model” (PRL 64, 1445 (1990)) [18] and “Continuum model for solitons in polyacetylene” (PRB 21, 2388 (1980)) [19] (both from Hubbard model community) is same as the one used for domain wall motion in “Solitons in polyacetylene” [20] (PRL 42, 1698 (1979)) and in “Particle spectrum in model field theories from semiclassical functional integral techniques” [21] (PRD 11, 3424 (1975)). These latter two publications are both in charge-density wave community.

We also quantify the cohesiveness of each community by the intensity of intra-community links. We therefore define the intra-community link weight $w_i \equiv 2l_i/[n_i(n_i-1)]$ as the fraction of potential links that actually exist within a community. Here l_i is the number of citations (links) inside community i. The largest value of the intra-community link weight is $w_i = 0.8$ for community 59 (atomistic metallic contacts). The large difference between the inter- and intra-community link weights suggest that the partitioning that results from modularity maximization is meaningful.

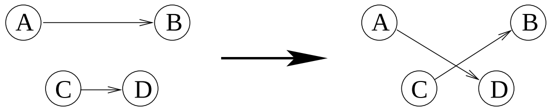

FIG. 3: A single random rewiring step. In each panel, time is increasing horizontally to the right.

To test the significance of the communities found by modularity maximization, we also applied this algorithm to a randomized version of the citation network. We construct such randomized networks by randomly rewiring links so as to preserve the degree of each node (the number of citations to each publication) and the time ordering of the links. As illustrated in Fig. 3, a single rewiring step consists of first selecting two links at random, AB and CS, and exchanging the targets of these two links so that the links become AD and CB. The rewiring step is performed only if citing paper A is earlier than cited paper D and C is earlier than B, so as to preserve the time-ordering of the links. This rewiring step is applied $10 \times L$ times, where L is the total number of network links. Thus each link is rewired ten times, on average. In this way, we preserve the in-degree and out-degree of each network node, as well as the time ordering of every link, but mix the global connectivity pattern. Thus community structure will be significantly reduced by this repeated exchange of links.

Applying modularity maximization to these randomized networks, we find that the modularity ranges from 0.18 to 0.25 for 20 different networks that were randomized by rewiring. These values are significantly smaller than the modularity of $Q = 0.543$ for the actual PR network. Moreover, after the final division of the randomized networks into communities, the number of inter-community crosslinks is a factor 8.7 larger than that in the real PR citation network. Thus communities are much more cleanly defined in the PR citation network than in the randomized networks. This test strongly suggests that the community structure we find in the PR citation network is a real feature that arises from the correlations between citations.

IV. STRUCTURE OF INDIVIDUAL COMMUNITIES

The individual communities within the PR citation network have a wide range of structures, ranging from tightly knit to barely classifiable as a single entity. To illustrate this diversity, we again focus on the 61 most prominent PR citation communities that contain 5 or more publications (Table VII). Apply modularity maximization to each community separately, we find modularity values that range from 0.16 to 0.50 for the communities that contain more than 25 publications. (A modularity value of zero does occur for 13 of the smallest communities; we ignore them because of their small size in the following discussion.) As mentioned in section II a modularity greater than 0.3 was empirically found indicate community structure within a network [24, 25]. Among the 61 communities listed in table VII, 23 of them have a maximum modularity larger than 0.3.

A. Most Cohesive Communities

Let us focus on the extreme cases. The most tightly-connected communities are high-temperature superconductivity (high-Tc), with a modularity value of 0.194 and Bose-Einstein condensation (BEC) with a modularity value 0.217. These two communities are illustrated in Fig. 4, with the titles of the top-5 cited articles in each of them given in table II. The communities are visualized by the Kamada-Kawai algorithm [22] in which nodes are treated as identically-charged particles and the edges are identical springs and the algorithm arranges nodes to minimize the energy of the system. As is visually apparent in Fig. 4, these two communities are strongly interconnected and they do not contain any visually discernible substructure.

TABLE II: The top-5 cited papers in the high-Tc and BEC communities.

# cites	title of high-T_c paper	authors	reference
844	Effective Hamiltonian for the superconducting ...	Zhang and Rice	PRB 37, 3759 (1988)
643	Superconductivity at 93K in a new mixed-phase ...	Wu et al.	PRL 58, 908 (1987)
635	Density matrix formulation for quantum ...	White	PRL 69, 2863 (1992)
634	Effects of double exchange in magnetic crystals	de Gennes	PR 118, 141 (1960)
558	Theory of high-Tc superconductivity in oxides	Emery	PRL 58, 2794 (1987)
# cites	title of BEC paper		
1119	Bose-Einstein condensation in a gas of sodium atoms	Davis et al.	PRL 75, 3969 (1995)
839	Evidence of Bose-Einstein condensation in an ...	Bradley et al.	PRL 75, 1687 (1995)
512	Cold bosonic atoms in optical lattices	Jaksch et al.	PRL 81, 3108 (1998)
388	Vortex formation in a stirred Bose-Einstein condensate	Madison et al.	PRL 84, 806 (2000)
353	Bose-Einstein condensation of Lithium: ...	Bradley et al.	PRL 78, 985 (1997)

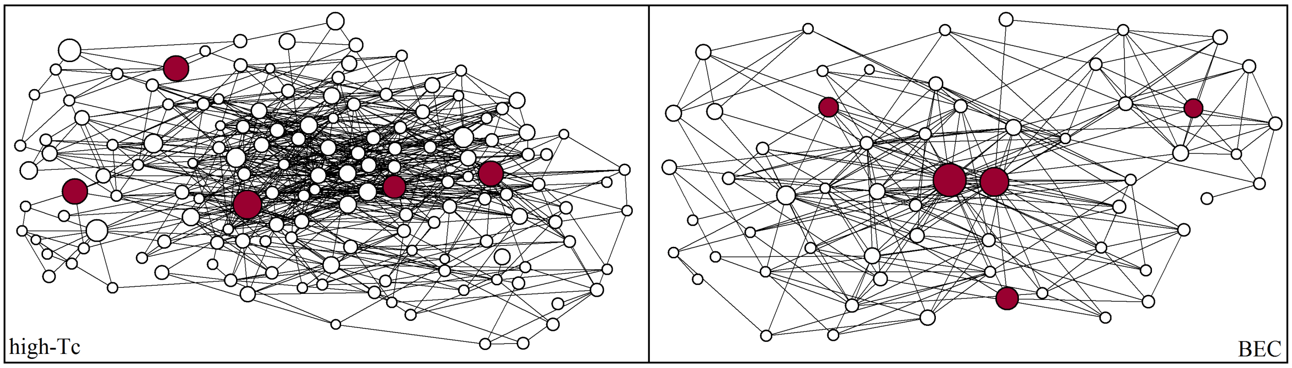

FIG. 4: Visualization of communities 4 and 13 in Fig. 2: High-Tc and Bose-Einstein condensation, respectively. The top-5 cited papers in each community are denoted by dots. The size of each symbol is proportional to its number of citations.

B. Least Cohesive Community

On the other hand, the communities with the highest modularity values are quite different in character. If these communities were treated as isolated entities in the absence of the rest of the network, modularity maximization would divide them into still smaller communities. The resulting large value of the modularity that is found after these communities are divided suggests that the substructure within each community is significant. The least-cohesive such example is community 6 (statistical physics, Monte Carlo methods, gauge theory, and quarks), with modularity 0.498. This community is shown in Fig. 5 and table III lists the top-5 cited papers for this community.

It may seem surprising, at first sight, that papers such as "Thermal fluctuations, quenched disorder, phase transitions, and transport in type-II superconductors" and "Dynamic scaling of growing interfaces" are in the same community as "Confinement of quarks". The origin of the connection between these publications is illustrated in Fig. 5, where the community appears to have three distinct modules. Here we manually arrange the nodes so that the connections between these modules are highlighted. Statistical physics papers (SP) appear in the large ellipse are on the left, Monte Carlo and gauge theory papers (MC) are in the middle, and quark-related papers (Q) are on the right. There are 22 links between the SP and MC modules; for visual clarity, only the 6 most significant links (in which one or both of the link ends attach to a node with more than 500 citations) are shown. Also shown are all of the 6 links between the MC and Q modules.

TABLE III: Statistical physics, Monte-Carlo method, gauge theory, and quarks

# cites	title	authors	reference
1009	Dynamic scaling of growing interfaces	Kardar et al.	PRL 56, 889 (1986)
852	Absence of ferromagnetism or antiferromagnetism ...	Mermin and Wagner	PRL 17, 1133 (1966)
783	Thermal fluctuations, quenched disorder, phase ...	Fisher et al.	PRB 43, 130 (1991)
660	Crystal statistics.I. a two-dimensional model with ...	Onsager	PR 65, 117 (1944)
655	Confinement of quarks	Wilson	PRD 10, 2445 (1974)

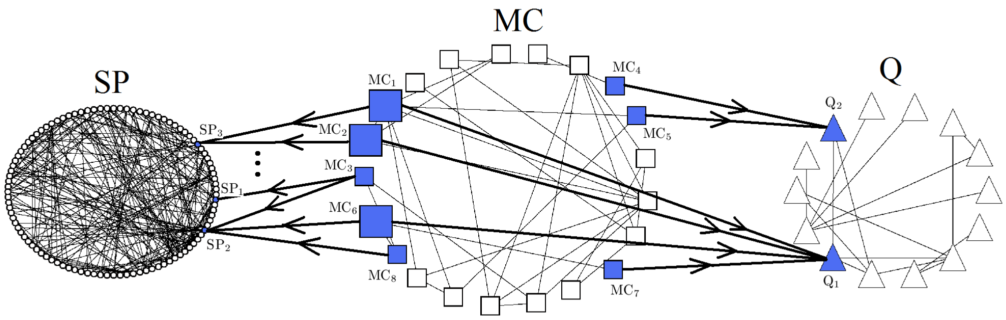

FIG. 5: Community 6 in Fig. 2: statistical physics, Monte Carlo methods, gauge theory, and quarks, that is divided into the modules of statistical physics (SP), Monte Carlo methods and gauge theory (MC), and quarks (Q). Filled symbols denote the labeled nodes in table IV. The larger nodes directly join the three modules through their citations.

To highlight the connections between these modules, we focus on the subset of citations, marked by thick lines in Fig. 5, that join publications across all three modules. We label the nodes at the ends of these links by SP$_1$–SP$_3$, MC$_1$–MC$_8$, Q$_1$ and Q$_2$, respectively, for statistical physics, Monte Carlo, and quark publications. The subscripts order these nodes by their number of citations. Among these highly-cited nodes, MC$_1$, MC$_2$, and MC$_6$ are exceptional because they directly link modules SP and Q. In MC$_1$: "New Monte Carlo technique for studying phase transitions" (PRL 61, 2635 (1988)) and MC$_2$: "Optimized Monte Carlo data analysis" (PRL 63, 1195 (1989)), Ferrenberg and Swendsen introduced an efficient Monte Carlo technique to extract information at many temperatures from data that is generated at a single temperature. This technique is especially useful in the Ising model and in lattice gauge theory, where significant computational resources are needed. The former is a model of ferromagnetism and is widely used in statistical physics. The latter is an important tool for studying the confinement of color-charged particles (such as quarks). Publication MC$_6$ "Order and disorder in gauge systems and magnets" (PRD 17, 2637 (1978)) showed that the renormalization-group equations of four-dimensional gauge theories (used in Q$_1$) have the same structure as those of two-dimensional spin systems (discussed in SP$_2$). Because of the shared techniques (large-scale Monte Carlo simulations) and similar mathematical structure (the analogies between gauge theories and spin models), the SP and Q modules, which might seem to belong to disparate physics fields, are actually connected by module MC.

TABLE IV: Bridge nodes for statistical physics, Monte Carlo method, gauge theory, and quarks community

label	# cites	title	authors	reference
SP$_1$	852	Absence of ferromagnetism or antiferromagnetism ...	Mermin and Wagner	PRL 17, 1133 (1966)
SP$_2$	602	Renormalization, vortices, and ...	Jose and Kadanoff	PRB 16, 1217 (1977)
SP$_3$	172	Bounded and inhomogeneous Ising models: ...	Ferdinand and Fisher	PR 185, 832 (1969)
MC$_1$	404	New Monte Carlo technique for studying phase ...	Ferrenberg and Swendsen	PRL 61, 2635 (1988)
MC$_2$	301	Optimized Monte Carlo data analysis	Ferrenberg and Swendsen	PRL 63, 1195 (1989)
MC$_3$	135	Monte Carlo study of the planar spin model	Tobochnik and Chester	PRB 20, 3761 (1979)
MC$_4$	129	High-temperature Yang-Mills theories and ...	Appelquist and Pisarski	PRD 23, 2305 (1981)
MC$_5$	125	Critical properties from Monte Carlo coarse ...	Binder	PRL 47, 693 (1981)
MC$_6$	120	Order and disorder in gauge systems and magnets	Fradkin and Susskind	PRD 17, 2637 (1978)
MC$_7$	104	Impossibility of spontaneously breaking local ...	Elitzur	PRD 12, 3978 (1975)
MC$_8$	103	Topological excitations and Monte Carlo ...	DeGrand and Toussaint	PRD 22, 2478 (1980)
Q$_1$	655	Confinement of quarks	Cardona et al.	PR 154, 696 (1967)
Q$_2$	108	Lattice models of quark confinement at high ...	Susskind	PRD 20, 2610 (1979)

The example of the statistical physics, Monte-Carlo method, gauge theory, and quarks community suggests that the reducibility of a community by modularity maximization depends on the nature of its embedding in the rest of the network. When this particular community is isolated it is reducible, but it is irreducible when considered part of the entire citation network. It is worthwhile to understand this feature in an idealized example. Thus consider two complete graphs of 100 nodes each in which the fraction of all possible links between these two graphs is p. When this twinned complete graph is isolated, a value of $p < 0.9$ is sufficient to induce modularity maximization to divide this network into its two constituent complete graphs. However, if the double-graph system is embedded in a larger network, the value of p needed for division to occur quickly decreases with the size of the larger network.

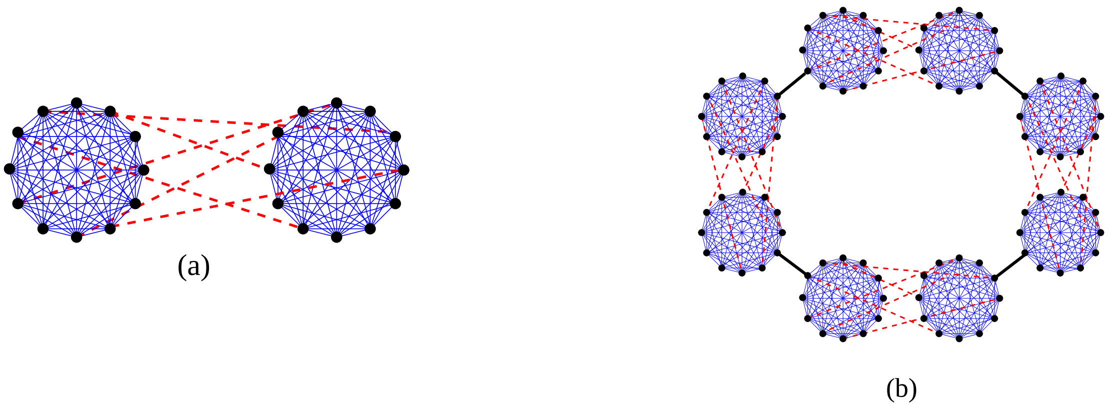

FIG. 6: (a) A twinned complete graph that consists of two complete graphs, with a fraction p of the possible links (dashed) between them also present. (b) Four twinned complete graphs interconnected by a single link (heavy solid lines).

For example, consider four identical copies of the twinned complete graph, in which each such graph is connected to another by a single link (Fig. 6). In this case, modularity maximization for the entire network immediately breaks all the single links, as one would naturally expect. However, the individual complete graph twins remain joined unless the value of p is less than 0.2. Thus there is a wide range of p values for which an isolated twinned complete graph will be split by modularity maximization, but will remain intact if this same twinned complete graph is part of a larger network.

V. TIME EVOLUTION

The community structure presented in the previous sections represents an integrated view of the PR citation network over the 114-year period 1893–2006. Physics is an evolving discipline, however, and it is revealing to study how the community structure of the citation network evolves over time. To this end we partitioned all PR publications into eight decadal sets by year of publication: 1927–1936, 1937–1946, . . ., 1997–2006. Our choice of decades was dictated by 2006 representing the last year for which complete citation data was available. We also restrict our analysis to roughly the top-3000 cited papers in each decade. More precisely, we adjust the threshold for number of citations to make the total papers in each decade to be as close to 3000 as possible (the actual number in each decade ranged from 2751 to 3046). Because the Physical Review has been growing roughly exponentially with time since 1893 (with somewhat slower growth after World War II than before [28]), each decadal snapshot represents a different fraction of the total number of publications in this period.

Table V lists basic information about the publications in each of the eight decades. The fraction of publications that we analyze (column P) becomes quite small in the last four decades so that our results from later decades are biased toward highly-cited papers. We did not analyze the period before 1927 because the number of papers is too small to resolve them into meaningful communities. We categorize these communities as belonging to PRA/E (atomic, molecular, optical, statistical), PRB (condensed-matter), PRC (nuclear), or PRD (particle). We treat Physical Review E (statistical physics) and Physical Review A (atomic, molecular, optical physics) together, since PRE was split off from the latter only in 1990. To assign communities to categories, we adopt the following procedure: after 1970, papers that appeared in PRA/E, PRB, PRC, and PRD can be unambiguously be assigned their subject category. For papers in the letters section (PRL) or for papers published before 1970, we count the fraction of their citations that come from PRA/E, PRB, PRC, and PRD. We then assign a paper the category of the plurality of its citations. Finally, a community can be categorized by the plurality of all the citations to its constituent publications.

There are some ambiguities with this procedure that should be noted. First, we can only use citing papers after 1970 (that appear in the four sections of Physical Review) to categorize papers that were published before 1970. Thus citations from papers that were published before 1970 are not used. Second, the categorization of some communities is not definitive. For example, the category fractions for community 401 (Pion & nucleon interactions) is 24%, 11%, 34%, and 32%, respectively, for PRA/E, PRB, PRC, and PRD. Since PRC has the largest share, the community is categorized as being part of PRC. Fortunately, for the 109 decadal communities in Fig. 7, 101 of them have more than a 50% share in one category.

TABLE V: The eight decadal publication sets 1927–36 to 1997–2006. Here T is the citation threshold for inclusion in the dataset; N is number of PR papers in each decade, M is number of decadal papers analyzed, $P = M/N$, and Q is the maximum modularity in each decade. The last 4 columns are the fraction of papers that belong to the categories PRA/E, PRB, PRC, and PRD (see text).

Decades	T	N	M	P	Q	PRA/E	PRB	PRC	PRD
1927-1936	2	3908	2751	70%	0.50	84.1%	15.9%	0%	0%
1937-1946	1	3530	3007	85%	0.51	39.9%	0%	51.4%	8.7%
1947-1956	14	12692	2994	24%	0.40	9.6%	27.3%	43.0%	20.1%
1957-1966	25	20642	3046	15%	0.52	10.3%	48.7%	11.9%	29.2%
1967-1976	34	43628	2982	7%	0.56	12.1%	57.1%	0%	30.8%
1977-1986	44	54475	2950	5%	0.59	6.7%	68.2%	0%	25.1%
1987-1996	59	103774	2997	3%	0.61	15.7%	79.6%	0%	4.6%
1997-2006	42	151693	2954	2%	0.64	41.5%	31.6%	3.7%	23.2%

Figure 7 illustrates these time-resolved communities decade by decade since 1927, showing both connections within each decade and connections between communities in different decades. Along the vertical axis, the communities are arranged according to the categories PRA/E, PRB, PRC, and PRD. There are 109 decadal communities with more than 50 publications and their subjects are listed in tables VIII to IX (see appendix). The node size again is proportional to the number of publications in each community. For visual clarity, we do not show links with weights $w_{ij} \leq 0.001$. The strongest connections occur between communities within the same decade and between communities in consecutive decades. These short-range temporal connections are natural to expect because of the average citation lifetime is only 6 years [28] and because intellectually-close publications may be in neighboring decadal datasets. Because of the arbitrariness of the partitioning into decades, a single subfield may appear as multiple communities in adjacent decades. Consequently, links between communities in adjacent decades should be viewed as equivalent to links within a given decade. An indication that two communities in adjacent decades are really a single community is that their topics are similar and there is an above-average number of interconnecting links between them. Thus, for example, the communities “Conductance fluctuations/scaling” (703 in Fig. 7) from 1977-1986 and community “Conductance fluctuations” (labeled 814) from 1987-1996 are likely part of the same community. The intensity of citations (the number divided by the product of the two community sizes) between them is 10.2 times larger than the average intensity of citations between any two communities.

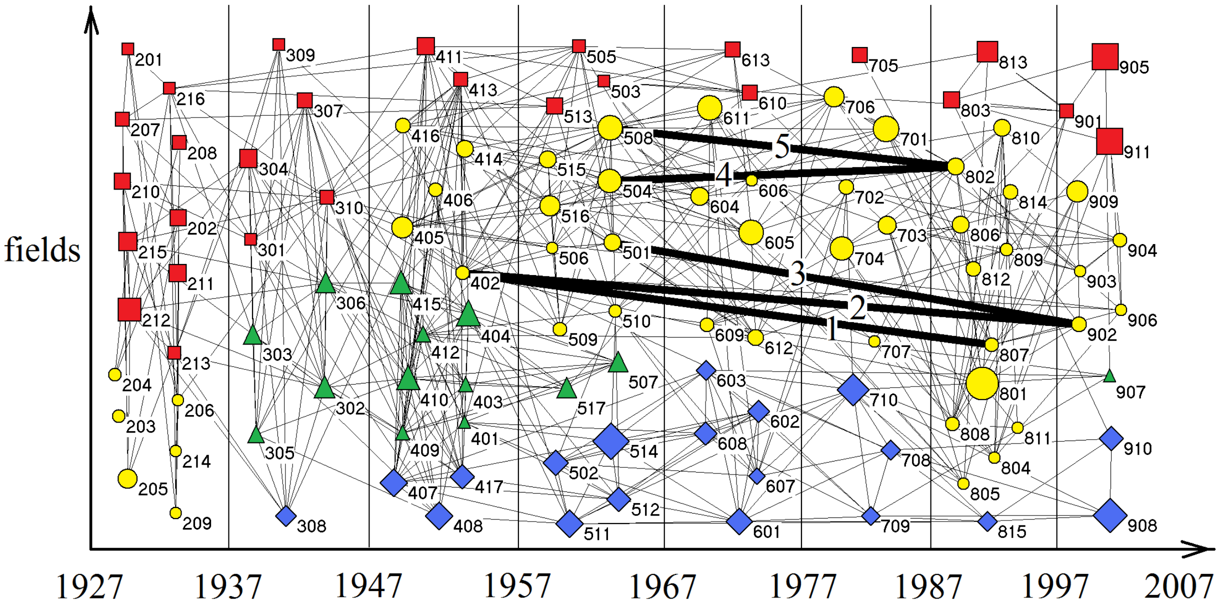

FIG. 7: Citation communities decade by decade. Communities are arranged from top to bottom categories by PRA/E; (squares), PRB (circles), PRC (triangles), or PRD (diamonds). The thick lines indicate the top-5 long-range temporal connections between communities.

There are a variety of interesting anomalies in the decadal data. During the decades 1927–36 and 1997–2006 atomic physics (PRA) was the largest field. From 1927–36, the activity in this field was likely due to the revolution in quantum mechanics. Between 1997–2006, the upsurge in atomic physics is due to the developments in Bose-Einstein condensation and in quantum information theory. Condensed-matter physics (PRB) is relatively stable throughout all eight decades, but with two noteworthy features. First, in 1987–96, condensed-matter physics represented 79.6% of all highly-cited papers in this decade (defined as more than 59 citations, see table V). Moreover 40.1% of all papers were related to high-temperature superconductivity. Second, during 1937–46, none of the communities shown in Fig. 7 belong to condensed-matter physics, while 51.4% of all papers in this decade were in nuclear physics.

The fraction of highly-cited nuclear physics publications reached a maximum during 1937–46, remained significant for the following decade, but then almost completely disappeared from the most highly-cited papers after 1967 (only 1 community in 1997–2006 can be categorized as belonging to PRC). Last, particle physics (PRD) has had a fraction of between 20% − −30% of the highest-cited papers, except for 1987–1996 (4.6%). The small share for this field between 1987–96 was caused by the upsurge in high-temperature superconductivity, Bose-Einstein condensation, and quantum information. Amazingly, these three topics comprised 63.1% of all highly-cited papers in this decade.

Another interesting feature that is visible in Fig. 7 is the existence of a very small number of long-range links in time. In fact, there are only five links that span more than two decades in time and have a link weight greater then 0.004 (table VI). These long-range links are indicative of significant scientific developments. For example, links 1, 2, and 3 in Fig. 7 can be traced to the phenomenon of colossal magnetoresistance (CMR). Magnetoresistance is the change in electrical resistance of a material when a magnetic field is applied. In 1951, Zener explained the magnetoresistance of manganese oxides by the double exchange mechanism [29]. It was only in the mid-1980s, that advances in nanotechnology allowed experimentalists to create the types materials proposed in earlier theoretical studies. In fact, these materials exhibited "giant" magnetoresistance (GMR), a phenomenon that led to the development of modern hard-disk industry and to the Nobel prize in physics to Fert and Grunberg in 1988 for their discoveries in this field. Subsequent high-precision fabrication techniques led to the even more extreme phenomenon of "colossal" magnetoresistance (CMR). The three long-range links mentioned above owe their existence to the delayed experimental renaissance of theoretical ideas.

Links 4 and 5 can be attributed to two famous condensed-matter physics papers — "Self-consistent equations including exchange and correlation Effects" by Kohn and Sham (KS) and "Inhomogeneous electron gas" by Hohenberg and Kohn (HK). Link 4 consists of 33 citations between the communities "Correlated electron systems" and "Pseudopotentials", and 26 out of them are solely based on KS. This publication is, in fact, the most-cited in all Physical Review, with 4849 *internal* citations (citations from other PR publications to KS) as of the fall of 2009. Similarly link 5 consists of 25 citations between communities "Many-body systems" and "Pseudopotentials", of which 21 of them are due to HK. This publication is the second most-cited PR paper. These two papers continue to be heavily cited more than 40 years after publication because of the wide usage of the approximation methods for dealing inhomogeneous and interacting electron systems.

TABLE VI: The top-5 long-range links.

link	community name	community name
1	NMR (1947-1956)	Giant magnetoresistance (1987-1996)
2	NMR (1947-1956)	Manganite (1997-2006)
3	Quantum magnetism (1957-1966)	"Manganite" (1997-2006)
4	Correlated electron systems (1957-1966)	Pseudopotentials (1987-1996)
5	Many-body systems (1957-1966)	Pseudopotentials (1987-1996)

VI. SUMMARY

The recent availability of large citation datasets has made it feasible to study properties of citation networks that were inaccessible even one decade ago. One of the goals of this work is to understand the structure of the citation network of the most prominent US archival physics journal — the Physical Review (PR) family of journals. Our goal was to determine basic properties of the subfields of physics — their importance, their interconnections, and their evolution — by focusing on citations rather than on the content of the publications themselves. To identify subfields within the PR citation network, we used an algorithm that repeatedly attempts to partition the network into well-defined communities by maximizing a measure known as the modularity. Using this approach, the PR citation network was found to possess an underlying community structure that corresponds to clearly identifiable subfields.

We also we studied the structure of the individual communities that were identified by modularity maximization. By treating each community as an independent network and again applying modularity maximization, we found that these individual communities were diverse in their structures, ranging from tightly focused to barely being identifiable as a community. The most weakly defined communities consist of a small number of well-defined modules that are typically linked by publications that emphasize techniques (as in the example of the Statistical physics, Monte Carlo method, gauge theory, and quarks community discussed in Sec. IV. These weakly defined communities — when isolated from the rest of the network — will be divided further by modularity maximization. However, modularity maximization leaves these communities intact when they are considered as part of the entire network, so that some of the communities do not appear coherently organized around a single theme.

We also studied the time evolution of the citation network and found five exceptionally long-lived links between subfields that are separated by more than two decades in time (compared to the average PR citation age of approximately 6 years). These long-range links arise from one of two mechanisms: either (i) a long delay between theoretical insights and the development of experimental methods to implement these ideas (links 1, 2, and 3 in Fig.7), or (ii) the introduction of a widely-used new method (links 4 and 5). We also uncovered a number of anomalies in the evolution of the four major categories of Physical Review (PRA/E, PRB, PRC, and PRD) throughout the decades 1927–36 until 1997–2006 that can be traced to major historical events. The more prominent such examples include: (i) the flowering of nuclear physics in the period just before and just after WWII, where the share of PRC (nuclear physics) publications was maximal, (ii) the discovery of high-temperature superconductivity, where the share of highly-cited PRB (condensed-matter physics) publication rose dramatically in the 1980s, and (iii) the burgeoning of quantum information theory over the past decade, leading to a sharp increase in publications in PRA (atomic, molecular, optical physics).

Acknowledgments

We are grateful to Mark Doyle and Paul Dlug from the APS for providing the Physical Review citation data. We are also grateful for financial support from NSF grants DMR0535503 and DMR0906504 .

[1] A. Z. Capri, Quips, quotes, and quanta: an anecdotal history of physics, (World Scientific, Singapore; Hackensack, NJ 2007).
[2] M. Herrera, D. C. Roberts, N. Gulbahce, arXiv:0904.1234v1 (2009).
[3] S. Wasserman and K. Faust, Social Network Analysis, (Cambridge University Press, Cambridge, 1994).
[4] M. Girvan and M. E. J. Newman, Proc. Natl. Acad. Sci. USA **99**, 7821 (2002); Proc. Natl. Acad. Sci. USA **103**, 8577 (2006).
[5] G. W. Flake, S. R. Lawrence, C. L. Giles, and F. M. Coetzee, IEEE Computer **35**, 66 (2002).
[6] B. W. Kernighan and S. Lin, Bell Syst. Tech. J. **49**, 291 (1970).
[7] U. Elsner, Graph partitioning - a survey. Technical Report 97, Technische Universität Chemnitz (1997).
[8] Fjallstrom, Electronic Articles in Computer and Information Science **3**, 10 (1998).
[9] H. C. White, S. A. Boorman, and R. L. Breiger, Am. J. Sociol. **81**, 730 (1976).
[10] M. E. J. Newman, Eur. Phys. J. B **38**, 321 (2004).
[11] M. Rosvall, C. T. Bergstrom, arXiv:0812.1242v1 (2008).
[12] Y. Kim, S. W. Son, H. Jeong, arXiv:0902.3728v1 (2009).
[13] M. A. Porter, J. P. Onnela, and P. J. Mucha, arXiv:0902.3788v1 (2009).
[14] A. Lancichinetti and S. Fortunato, arXiv:0904.3940v2 (2009).
[15] E. A. Leicht, M. E. J. Newman, Phys. Rev. Lett. **100**, 118703 (2008).
[16] M. E. J. Newman, Phys. Rev. E **69**, 066133 (2005).
[17] V. D. Blondel, J.L. Guillaume, R. Lambiotte and E. Lefebvre, J. Stat. Mech. P10008, (2008).
[18] H. J. Schulz, Phys. Rev. Lett. **64**, 1445 (1990).
[19] H. Takayama, Y. R. Lin-liu and K. Maki, Phys. Rev. B **21**, 2388 (1980).
[20] W. P. Su, J. R. Schrieffer and A. J. Heeger, Phys. Rev. Lett. **42**, 1698 (1979).
[21] R. Dashen, B. Hasslacher, and A. Neveu, Phys. Rev. D **11**, 3424 (1975).
[22] T. Kamada and S. Kawai, Information Processing Letters **31**, 7 (1988).
[23] M. E. J. Newman and M. Girvan, Phys. Rev. E **69**, 026113 (2004).
[24] L. Danon, J. Duch, A. Diaz-Guilera, and A. Arenas, J. Stat. Mech. P09008, (2005).
[25] R. Guimera and L. A. N. Amaral, Nature **433**, 895 (2005).
[26] J. Duch and A. Arenas, Phys. Rev. E **72**, 027104 (2005).
[27] E. A. Leicht and M. E. J. Newman, Phys. Rev. Lett. **100**, 118703 (2008).
[28] S. Redner, Physics Today **58**, 49 (2005).
[29] C. Zener, Phys. Rev. **82**, 403 (1951).

Appendix

TABLE VII: Top 61 communities in the PR citation network. Here R denotes the citation rank, F and N are fraction and the number of all highly-cited publications in each community, and Q is the local modularity for each community. Communities with modularity values greater than 0.4 are generally multi-themed, as described in the text, while modularity values of zero occur only for the smallest communities.

R	F	N	Q			R	F	N		
1	6.54%	191	.371	Elementary particles		32	0.75%	22	.312	Cross sections
2	5.99%	175	.240	Correlated electrons		33	0.72%	21	0	Fractional quantum Hall effect: experiment
3	5.86%	171	.398	Quantum information		34	0.68%	20	.318	Self organized criticality
4	5.03%	147	.194	Theories of high-Tc and type-II superconductors		35	0.68%	20	.279	Magnetism
5	4.97%	145	.436	Quantum diffusion, quantum Hall effect and two-dimensional melting		36	0.65%	19	.223	High-harmonic generation
6	4.59%	134	.498	Statistical physics, Monte Carlo method, gauge theory, and quarks		37	0.65%	19	0	Fermi surface
7	4.32%	126	.424	High-Tc oxides		38	0.62%	18	0	Quarks and infinite momentum
8	4.04%	118	.386	Theories of superconductivity		39	0.55%	16	.234	Charge density waves
9	2.95%	86	.441	Strong-coupling theory of superconductivity		40	0.55%	16	.191	Self-induced transparency
10	2.60%	76	.487	Semiconductors and quantum dots		41	0.48%	14	.308	Weak interactions
11	2.50%	73	.426	Material Science		42	0.45%	13	.165	Scanning tunneling microscope
12	2.33%	68	.366	Quantum field theory		43	0.45%	13	.168	Astrophysics
13	2.05%	60	.217	B.E. condensation		44	0.45%	13	0	Quantum optics
14	1.99%	58	.251	Cosmology		45	0.41%	12	.223	Stochastic resonance
15	1.82%	53	.481	Metals and alloys		46	0.38%	11	.286	Top quark
16	1.75%	51	.314	Symmetry breaking and gauge theories		47	0.34%	10	.238	Pinning in superconductors
17	1.71%	50	.320	Giant magnetoresistance		48	0.34%	10	.078	e^+e^- annihilation
18	1.51%	44	.322	High energy collisions		49	0.34%	10	.156	Synchronized chaos
19	1.40%	41	.206	Fractional quantum Hall effect: theory		50	0.27%	8	0	Surface transitions
20	1.34%	39	.381	Equilibrium statistical mechanics		51	0.27%	8	0	Atomic collision interactions
21	1.23%	36	.161	Diffusion limited aggregation		52	0.24%	7	0	Elastic scattering
22	1.06%	31	.447	Quantum mechanics		53	0.24%	7	0	Hubbard model
23	1.03%	30	.393	Crystal structure		54	0.24%	7	.223	Optical properties of metals
24	0.99%	29	.361	Nucleon-nucleon interactions		55	0.24%	7	.198	Reheating after inflation
25	0.99%	29	.305	Neutrino oscillations		56	0.24%	7	0	Magnetic anisotropy
26	0.99%	29	.326	Quantum teleportation		57	0.24%	7	0	Transport in disordered systems
27	0.92%	27	.428	Collective electronic properties/ nuclear fission		58	0.24%	7	.210	Nuclei
28	0.89%	26	.278	Quantum entanglement		59	0.21%	6	0	Atomistic-sized metallic contacts
29	0.82%	24	.282	Nuclear collision		60	0.21%	6	0	Carbon nanotubes
30	0.79%	23	.152	Elementary particle theories		61	0.21%	6	0	Networks
31	0.75%	22	.242	Metal surface structure						

TABLE VIII: Highly-cited publications by decades. The numerical label indicates the individual communities in Fig. 7 and N is number of publications in each community. Each is also categorized by its Phys. Rev. classification and by its topic.

1927–36					1937–46			
label	N	class	topic		label	N	class	topic
201	55	PRA/E	Ionization/electron scattering		301	52	PRA/E	Radioactivity
202	104	PRA/E	Electron spectroscopy		302	151	PRC	Nuclear moments
203	64	PRB	X-Ray diffraction		303	120	PRC	Radioactivity/nuclear reactions
204	64	PRA/E	Infrared spectrum		304	121	PRA/E	Nuclear physics/Cosmic rays
205	143	PRA/E	Atomic spectra		305	90	PRC	Gamma/beta rays
206	53	PRB	Photoelectric properties of metals		306	133	PRC	Neutron scattering
207	71	PRA/E	Ionization		307	85	PRA/E	Deuteron binding
208	79	PRA/E	Molecular spectra		308	84	PRD	Cosmic rays
209	57	PRB	Vacuum arc cathode		309	54	PRA/E	Mesons
210	106	PRA/E	Thermionic emission		310	71	PRA/E	Nuclear magnetic moments
211	117	PRA/E	X-ray spectra					
212	198	PRA/E	Cosmic ray/N & P interactions					
213	65	PRA/E	Cosmic radiation					
214	51	PRB	Photoelectric effect					
215	133	PRA/E	Slow neutrons					
216	53	PRA/E	Spectrography					

1947–56					1957–66			
label	N	class	topic		label	N	class	topic
401	53	PRC	Pion & nucleon interactions		501	121	PRB	quantum magnetism
402	77	PRB	Nuclear magnetic moments		502	117	PRD	Axial vector/weak interactions
403	68	PRC	Nuclear levels/scattering		503	52	PRA/E	Optical beams
404	201	PRC	Nuclear energy levels		504	208	PRB	Correlated electron systems
405	177	PRB	Semiconductors		505	69	PRA/E	Atomic structure
406	73	PRB	Mesons		506	58	PRB	Spin relaxation
407	140	PRD	Quantum electrodynamics		507	126	PRC	Scattering
408	138	PRD	Beta-ray spectra		508	227	PRB	Many-body systems
409	72	PRC	Nuclear reactions and levels		509	81	PRB	Bismuth
410	184	PRC	Properties of nuclei		510	62	PRB	High energy scattering
411	112	PRA/E	Precision QED tests		511	140	PRD	Parity nonconservation
412	83	PRC	High energy scattering		512	106	PRD	Pion interactions
413	72	PRA/E	Magnetic moments		513	96	PRA/E	Atomic collisions
414	110	PRB	Application of magnetism		514	254	PRD	Symmetries of elementary particles
415	167	PRC	QED/Nuclear reactions/scattering		515	111	PRB	Energy bands
416	89	PRB	Liquid helium		516	161	PRB	Superconductivity
417	109	PRD	Cosmic rays/nuclear scattering		517	126	PRC	Interaction of radiation with matter

1967–76					1977–86			
label	N	class	topic		label	N	class	topic
601	135	PRD	Weak interactions		701	256	PRB	Metal compounds
602	90	PRD	High-energy collisions		702	85	PRB	Soliton in polyacetylene
603	77	PRD	Strong interaction		703	132	PRB	Conductance fluctuations/scaling
604	132	PRB	Dielectric properties of complex materials		704	136	PRB	$2d$ electron gas
605	234	PRB	Metals		705	89	PRA/E	Diffusion limited aggregation
606	51	PRB	Two-dimensional systems		706	163	PRB	Correlated electron systems
607	52	PRD	Quarks		707	53	PRB	Semiconductor superlattices
608	98	PRD	Scattering theory		708	70	PRD	Cosmology
609	80	PRB	TTF-TCNQ		709	67	PRD	Particle physics
610	89	PRA/E	Self-induced transparency		710	194	PRD	Relativistic collisions/cosmology
611	237	PRB	Scaling theory					
612	105	PRB	Alloys					
613	89	PRA/E	Atomic structure					

TABLE IX: Continuation of table VIII.

1987–96					1997–2006			
label	N	class	topic		label	N	class	topic
801	407	PRB	High-Tc oxides		901	75	PRA/E	B.E.Condensation in optical lattices
802	112	PRB	Pseudopotentials		902	83	PRB	Manganites
803	96	PRA/E	B.E.Condensation in atomic gas		903	57	PRB	Electron transport in exotic systems
804	56	PRB	Fractional quantum Hall effect		904	70	PRB	Quantum magnetism/quantum dots
805	51	PRB	Photonics		905	255	PRA/E	Vortex formation in BEC
806	118	PRB	RG and Monte Carlo simulation		906	52	PRB	Spin-Hall effect
807	73	PRB	Giant magnetoresistance		907	52	PRC	High-energy nuclear collision
808	75	PRB	Flux lattice melting and High-Tc		908	214	PRD	Elemental particle theory
809	68	PRB	Quantum dots		909	184	PRB	High-Tc
810	119	PRB	SOC/pinning in superconductors		910	113	PRD	Cosmology
811	58	PRB	Vortex-glass superconductivity		911	256	PRA/E	Quantum information
812	84	PRB	Semiconductor epitaxial growth					
813	163	PRA/E	Quantum computation/teleportation					
814	89	PRB	Conductance fluctuations					
815	76	PRD	Particle physics					