University of Warsaw
Faculty of Mathematics, Informatics and Mechanics

Tomasz Wąs

Axiomatization of the Walk-Based Centrality Measures

PhD Dissertation

Supervisor
dr hab. Marcin Dziubiński
Institute of Informatics
University of Warsaw

Auxiliary Supervisor
dr Oskar Skibski
Institute of Informatics
University of Warsaw

January, 2022

Author’s declaration:
I hereby declare that this dissertation is my own work.

January 27, 2022

. . . .
Tomasz Wąs

Supervisor’s declaration:
The dissertation is ready to be reviewed.

January 27, 2022

. . . .
dr hab. Marcin Dziubiński

Abstract

Centrality measures are one of the fundamental tools of network science [43]. Their role is to assign to every node of a network a value that reflects the importance of this node. Centrality analysis finds numerous applications in the wide variety of fields: from social studies [32] and economics [7], through biology [42] and physics [80], to transportation [34] and computer science [49].

However, what it means for a node to be important heavily depends on the context of a particular application. This, along with the ever growing number of proposed measures, makes the choice of a centrality to use a difficult task. Since different centrality measures return very different results, this problem is of utmost importance. Hence, there is a need for research that will provide a better understanding of centrality measures and will help decide upon a centrality measure to use in a specific application.

A method that allows for achieving this goal is axiomatization. In this approach, we introduce a set of simple properties, called *axioms*, that characterize a given centrality measure. Then we formally prove that only this particular measure satisfies all of the axioms at the same time. In this way, we obtain an intuitive characterization of the centrality measure in question and its theoretical foundation. Moreover, by analyzing whether the axioms are desired, we can decide if this measure is well-suited for a particular application at hand.

In recent years, the axiomatic approach to centrality measures has been gaining popularity in the literature [10, 11, 12]. Axiomatic characterizations have been created for many centrality measures, such as closeness [77], beta measure [21], or attachment [76]. However, for many centrality measures such characterizations are still missing. Until recently, PageRank—one of the most popular centrality measures—was also lacking its axiomatization.

Against this background, the main contributions of this thesis are as follows: First, we introduce the first in the literature axiomatic characterization of PageRank. Next, we create a coherent axiomatization of three centrality measures: decay centrality, PageRank, and a novel measure—random walk decay centrality. Our analysis shows that while random walk decay centrality retains a majority of PageRank's properties, it may be more desirable than PageRank in some settings. Finally, we generalize our axiomatization of PageRank to create a consistent axiom system for four classic feedback centralities: eigenvector and Katz centralities, Seeley index, and PageRank.

Keywords: Network Science, Centrality Measures, Axiomatic Characterization, PageRank, Random Walks on Graphs

Streszczenie

Miary centralności są jednym z podstawowych narzędzi analizy sieci [43]. Nadają one każdemu wierzchołkowi wartość, określającą jak jest on istotny. Analiza centralności znajduje wiele zastosowań w szerokiej gamie dziedzin: od nauk społecznych [32] i ekonomii [7], przez biologię [42] i fizykę [80], po komunikację [34] i informatykę [49].

Jednak to, co rozumiemy przez istotność wierzchołka, zależy od kontekstu konkretnego zastosowania. W związku z tym oraz stale rosnącą liczbą zaproponowanych miar, wybór odpowiedniej z nich do danej sytuacji jest niezwykle trudny. Różne miary centralności potrafią dać znacząco różne wyniki, problem ten ma więc zasadnicze znaczenie. Niezbędne jest zatem opracowanie teorii, która pozwoli lepiej zrozumieć te miary i pomoże wskazać właściwą z nich do zastosowania w konkretnej sytuacji.

Metodą, którą możemy się w tym celu posłużyć, jest aksjomatyzacja. Polega ona na opracowaniu zbioru kilku prostych i intuicyjnych własności, *aksjomatów*, charakterystycznych dla danej miary centralności. Następnie formalnie udowadniamy, że tylko ta konkretna miara spełnia wszystkie te aksjomaty jednocześnie. W ten sposób otrzymujemy jej intuicyjną charakterystykę i teoretyczne podstawy do jej stosowania. Możemy bowiem stwierdzić czy powinniśmy zastosować miarę centralności w danej sytuacji, decydując czy charakteryzujące ją aksjomaty są pożądane.

Aksjomatyczne podejście do analizy miar centralności zyskuje popularność w literaturze ostatnich lat [10, 11, 12]. Charakteryzacje zostały opracowane dla takich centralności, jak centralność bliskości (ang. *closeness centrality*) [77], miara beta (ang. *beta measure*) [21], czy centralność łączenia (ang. *attachment centrality*) [76]. Jednak wiele znanych miar nadal nie ma swoich charakteryzacji. Do niedawna centralnością bez aksjomatyzacji był również PageRank – jedna z najpopularniejszych miar centralności.

Niniejsza rozprawa wpisuje się w ten nurt i oferuje następujący wkład: Po pierwsze, prezentujemy pierwszą w literaturze aksjomatyczną charakteryzację PageRanka. Po drugie, tworzymy spójne aksjomatyzacje trzech miar centralności: centralności zanikania (ang. *decay centrality*), PageRanka oraz nowej miary – centralności zanikania błądzenia losowego (ang. *random walk decay centrality*). Z naszej analizy wynika, że własności centralności zanikania błądzenia losowego w znacznej mierze pokrywają się z własnościami PageRanka, lecz jednocześnie może się ona lepiej sprawdzać w niektórych zastosowaniach. Po trzecie, uogólniamy aksjomatyzację PageRanka, aby stworzyć pierwszy spójny system aksjomatyczny dla czterech podstawowych centralności zwrotnych (ang. *feedback centralities*).

Słowa kluczowe: analiza sieci, miary centralności, charakterystyka aksjomatyczna, PageRank, błądzenie losowe po grafie

Acknowledgement

First of all, I would like to thank dr Oskar Skibski—my guide in a world of research and a role model of a scientist, a teacher, and a popularizer of science. It was a privilege to have a mentor that is at the same time a demanding supervisor and a friend on which you can always count.

Next, I would like to thank prof. Marcin Dziubiński for his many wise comments and help whenever I needed it. I want to thank all of the members of my research group, so also dr Dorota Celińska-Kopczyńska, dr Tomasz Michalak, and dr Piotr Skowron, for creating friendly and encouraging atmosphere for scientific work in these four years.

I would like to thank my family for believing in me and their constant support. Specifically, I want to thank my Mom for motivating me to reach ever higher, my Dad for many important life wisdoms, and my brothers for fostering my curiosity from my earliest years.

Finally, I would like to thank my dearest friends, Adam Oniszczuk, Aga Sitko, Antoni Zawodny, Dawid Nguyen, Jakub Trojgo, Jakub Wieczorek, Max Nguyen, Stanisław Łaniewski, and Wiola Osuch, to name a few, with whom I have learned so many lessons.

Contents

1 Introduction **1**
1.1 Related Work 3

2 Preliminaries **7**
2.1 Graphs 7
2.2 Centrality Measures 9
2.2.1 Classic Centrality Measures 9
2.2.2 Feedback Centralities 11
2.2.3 Walk Interpretations 12

3 Axiomatization of PageRank **19**
3.1 Axioms 19
3.2 Proof of Uniqueness 22
3.2.1 Part 1: PageRank Satisfies Axioms 22
3.2.2 Part 2: Axioms Imply PageRank 26
3.2.3 Independence of Axioms 37
3.3 Related Axiomatizations 44

4 Random Walk Decay Centrality **47**
4.1 Definitions 47
4.2 Axioms 49
4.3 Proofs of Uniqueness 51
4.3.1 Random Walk Decay Centrality 51
4.3.2 Personalized Decay Centrality 81
4.3.3 PageRank 85
4.4 Comparison of PageRank and RWD 87
4.4.1 Strategy-Proofness 87
4.4.2 Diversity 88
4.4.3 Real-World Example 89

5 An Axiom System for Feedback Centralities **91**
5.1 Additional Notation 91
5.2 Axioms 93

5.3 Proofs of Uniqueness 96
5.3.1 Seeley Index and Eigenvector Centrality 96
5.3.2 PageRank and Katz Centrality 111

Chapter 1

Introduction

In ever more interconnected world it is important, as never before, to understand the networks that surround us: transportation and communication networks, the network of webpages which we browse on the Internet, or the social networks that we are part of (both those in real life and those on virtual platforms), to name just a few examples. One of the key aspects in the analysis of these networks, is to understand how important are their particular elements.

Centrality measures were developed to answer these questions by formalizing our various intuitions on what it means to be important in a network. Each such measure is a function that assigns to each node in a network a real value that reflects its importance. Centrality analysis finds numerous applications in the wide variety of fields: from social studies [32] and economics [7], through biology [42] and physics [80], up to transportation [34] and computer science [49].

To date, more than three hundreds different centrality measures have been proposed in the literature [44]. However, three of them, *degree centrality*, *closeness centrality*, and *betweenness centrality*, are considered the most classic [30], and can be found in all standard tools for network analysis [24, 73]. Degree centrality simply counts the number of connections a node has in a network. Closeness centrality looks at the sum of distances to a node from all of the other nodes—the smaller this sum, the more central the node is. Finally, betweenness centrality aims to identify nodes that control the communication flow in a network by measuring how often a node lays on the shortest path between two other nodes.

A common feature of these three centrality measures is their reliance on the shortest paths in a network. This is a well-founded approach when the processes that occur in a network follow the shortest paths as well, which is the case in transportation or communication networks [17]. However, for many processes in real-life networks this assumption is clearly not adequate. Imagine surfers browsing through the Internet—they usually do not know how to reach another site in the optimal number of links [40]. Similarly, the news traveling through a social network move in a complex, seemingly random, way [55].

As a consequence, in this work, we mainly focus on walk-based centrality measures. This large class includes all centrality measures that assess the importance of nodes based not only on shortest paths, but on all paths, with possibly repeating nodes, called *walks*. In this way, they are able to capture the importance of a node in networks with complex flows and processes [17].

Feedback centralities form an especially appealing subclass of walk-based centrality measures. They stem from the assumption that a node is as important as important are its neighbours. Therefore, these measures are usually defined by recursive equations that bind centrality of a node with centralities of nodes it is connected to. One of the most famous feedback centrality is *PageRank*. Originally invented by Page et al. [62] to determine the importance of webpages for then newly created web search engine Google, it had a significant impact on the development of the Internet.

The success of PageRank inspired researchers from a variety of fields to use it in their own network problems. PageRank has been applied to indicate the most influential users in social media networks [86], to assess prestige of scientific journals in the citation network [15], to find the key proteins in metabolic networks [42], or even to determine the best tennis players in the history based on the network of their matches [67].

However, what it means to be important in a network depends heavily on the context of a particular application. Thus, it is not clear if it is PageRank that should be used in all of these settings, or more general, if a centrality measure that solves one problem should be applied to another. Since different measures lead to very different results, the problem of choosing the appropriate measure is of utmost importance. Therefore, there is a need for a research that will allow for better understanding of centrality measures and will help in choosing a right measure for a specific application.

For this purpose, axiomatic analysis of centrality measures has been developed [10, 13, 69]. In this approach, we consider a set of simple and intuitive properties, called *axioms*, that characterize a given centrality measure. Then, we formally prove that only this centrality measure can satisfy all of the proposed axioms at the same time. In this way, we obtain a unique characterisation of the measure and its theoretical foundation. Moreover, by analyzing whether the axioms are desired, we can decide if this measure is well-suited for a particular application at hand.

In this work, we follow axiomatic approach and create the first axiomatic characterization of PageRank in the literature. More in detail, we introduce six axioms that characterize it: *Node Deletion*, *Edge Deletion*, *Edge Multiplication*, *Edge Swap*, *Node Redirect*, and *Baseline*. Each of them is satisfied by many centrality measures, however, what is important, PageRank is the only centrality measure that satisfies all of them.

In this way, we obtain a clear and intuitive characterization of PageRank. Moreover, our axiomatization can help to decide whether PageRank should be used in a particular application. As an example, consider a network of tennis players analyzed by Radicchi [67]: for every tennis match in which player A won with player B, there is a directed edge from node B to node A. In such network PageRank indicates some node, v, as the most important one. Now, one of the axioms with which we characterized PageRank is Edge Multiplication. It states that creating additional copies of outgoing edges of a node does not affect the centrality of any node in the network. In the tennis player network example, it implies for instance that creating 9000 additional copies of outgoing edges of v does change the importance of any tennis player. Thus, v is still the most important. However, now v has 9000 times more lost matches—more than any other player. Therefore, a direct consequence of our characterization is the fact that PageRank should not be used in such applications.

An axiomatic characterization can also be a vital tool to compare the similarities and differences of several measures. Especially, if a coherent axiom system is proposed for multiple centrality measures. Following this approach, we create a consistent axiomatizations of three measures: decay centrality, PageRank, and a novel measure—random walk decay centrality. Specifically, we propose six axioms that uniquely characterize random walk decay centrality. Then, we show that if from this characterization we remove one axiom, called *Random Walk Property*, and instead add a new, similar axiom, *Shortest Paths Property*, then we obtain a unique characterization of standard decay centrality. Similarly, if in axiomatization of random walk decay centrality we exchange axiom *Lack of Self-Impact* for new axiom, *Edge Swap*, we get a unique characterization of PageRank. From our analysis we see that random walk decay centrality retains the majority of PageRank properties. However, because of their differences random walk decay centrality may be more desirable in certain settings. In particular, we show that it is less prone to manipulation and can positively account for diversity in a network.

In a similar manner, by extending our axiomatization of PageRank, we create a coherent axiom system for four classic feedback centralities: eigenvector centrality, Katz centrality, Seeley index, and PageRank. More in detail, we propose a set of seven axioms. Three general ones: *Locality*, *Edge Deletion*, and *Node Combination*, are satisfied by all four centralities. Two axioms: *Edge Multiplication* and *Edge Compensation*, consider manipulation of edges incident to a node. Finally, two axioms: *Baseline* and *Cycle*, specify centrality of a node in simple, borderline cases. We prove that each of the four centralities is uniquely characterized by a subset of five of our axioms: three general ones, one one-node-modification axiom, and one borderline axiom.

The rest of this work is structured as follows: In Section 1.1, we discuss the works related to axiomatization of centrality measures. The basic graph definitions and considered centrality measures are introduced in Chapter 2. Chapter 3 is devoted to PageRank and its axiomatic characterization. In Chapter 4, we introduce random walk decay centrality and provide its axiomatization that is consistent with axiomatization of standard decay centrality and PageRank. We conclude in Chapter 5, where we build upon our axiomatization of PageRank and create a coherent axiom system for four main feedback centralities.

1.1 Related Work

In this section, we present the existing literature devoted to characterization and classification of centrality measures. We begin with the works that, like us, follow an axiomatic approach.

Axiomatic approach to centrality measures

Axiomatic approach is one of the foundations on which mathematical theories are build. Many times it played a pivotal role in the development of mathematics. Euclid's five axioms of geometry [38], Peano's axioms of natural numbers [65], Zermelo–Fraenkel axiomatic set theory [45], or Kolmogorov's axiomatic theory of probability [51] are just a few notable examples. Today, it is often used in social choice theory [2], coalitional game theory [26], or strategic network analysis [85].

An axiomatic approach as a mean of studying centrality measures was first proposed by Sabidussi [69]. He introduced several axioms and argued that any reasonable centrality measure should satisfy all of them. Building upon this, he rejected some of the measures proposed in the contemporary literature, as not fit to be used as measures of centrality. Nieminen [58] followed a similar approach but considering directed, rather than undirected, graphs. More recently, a similar technique was employed by Landherr, Friedl, and Heidemann [54]. They proposed three properties that capture how adding a new edge to a graph should affect the centrality or relative ranking of certain nodes. Then, they checked which of these properties are satisfied by degree, closeness, betweenness, eigenvector, and Katz centralities. In a similar way, Boldi and Vigna [13] proposed three other general axioms and tested whether they are satisfied by ten of the popular centrality measures including degree, closeness, betweenness, and PageRank. Furthermore, Boldi, Luongo, and Vigna [11] extended the list of considered axioms by *Rank Monotonicity*: an axiom stating that adding an edge incident to a node cannot decrease its ranking. Finally, Riveros and Salas [68] considered five general axioms in the context of the class of centrality measures based on the subgraph counts.

Another approach, which is followed in this thesis, is to construct axiomatic characterizations of particular centrality measures. Many centrality measures have been axiomatized in such a way: beta measure was characterized by Brink [21], harmonic and decay centralities by Garg [31], and Skibski et al. [76] characterized attachment centrality. In all of these papers, also degree centrality has been characterized as a baseline to which other measures can be compared.

Probably due to their complexity, feedback centralities received especially significant attention when it comes to axiomatic characterizations. Eigenvector centrality has been characterized by Kitti [48]. Both Dequiedt and Zenou [25] and Wąs and Skibski [83] characterized eigenvector and Katz centralities in a joint axiomatic characterization. Seeley index, sometimes called simplified PageRank, has been axiomatized independently by Altman and Tennenholtz [1] and Palacios-Huerta and Volij [63] (to be precise, Palacios-Huerta and Volij axiomatized *invariance method*, which is a measure of importance of scientific journals that is equivalent to Seeley index). However, until recently, PageRank has not been yet characterized. In Section 3.3, we discuss in details the differences between axiomatizations of Seeley index by Altman and Tennenholtz [1] and Palacios-Huerta and Volij [63] and our axiomatization of PageRank.

A more challenging task is to create a coherent axiomatization for a whole class of centrality measures. In this way, Skibski and Sosnowska [77] characterized distance based centralities. They proposed a set of axioms that are satisfied by all centrality measures in this class together with axioms specific for particular measures, such as degree centrality, closeness centrality, or decay centrality. As a result, they obtained a unique characterizations for all of these measures and a framework that highlights their key similarities and differences. Until now, a similar construction for feedback centralities has not been developed.

A similar, yet fundamentally different, approach was taken by Bloch, Jackson, and Tebaldi [10]. They introduced a notion of *nodal statistics*—sequences of simple data regarding a node, e.g., numbers of nodes at a given distance. Then, they proposed several axioms that specify a relation between a centrality measure and an arbitrary nodal statistic. As they proved, the only centrality measures satisfying all of their axioms are the sums of discounted elements of a given nodal statistic. Thus, they showed that specifying a nodal statistic and a discounting scheme, characterizes different measures such as degree centrality, decay centrality, Katz centrality, or PageRank.

Other characterization approaches

Apart from axiomatic approach, other attempts to characterize and classify centrality measures were also made. One of the first influential papers devoted to that matter is that of Freeman [30]. He identified three premises for centrality, based on communication in a network, and presented a centrality measure to fit each premise. These were: direct communication activity (degree centrality), control over communication (betweenness centrality), and effectiveness of communication (closeness centrality).

A different and, in a sense, orthogonal approach was taken by Borgatti [17]. He considered structurally different types of flows that occur in networks. For instance, one can argue that a package delivery in a transportation network is characterized by intrinsically different flows than gossip propagation in a social network. Then, a centrality measure can be regarded as an assessment of a role that a node has in network communication given a certain type of flows. For example, if each node sends one message to each of its neighbours, degree centrality is the number of messages received by a node. In general, we can now classify centrality measures based on a type of flows for which they are suitable.

Borgatti and Everett [18] added to this picture the distinction inspired by Freeman: Are we more interested in the control over flows in a network, i.e., *medial* centralities, (e.g., betweenness centrality) or the flows' efficiency, i.e., *radial* centralities (e.g., closeness centrality)? Moreover, they distinguished between the centralities oriented towards the total volume of received flows and the ones oriented towards the length of these flows. In this way, they obtained multidimensional classification of centrality measures based on both flow types and specific aspects of flow propagation that we want to capture. However, it is worth noting that there can be multiple centrality measures in a single grid of this classification. Also, there are many centrality concepts that do not fit in this framework at all.

Another way of organizing the space of centrality measures was considered by Schoch and Brandes [70]. They expressed many of the known centrality measures in the unified framework based on the path algebras. Next, they showed that each centrality measure defined in such a way *preserves neighbourhood-inclusion pre-order* and argued that such property is a good indicator whether a node index can be regarded as a centrality measure.

A different approach is to distinguish certain classes of centrality measures based on their common characteristics. In this manner, Koschützki et al. [52] established several such classes, including feedback centralities or *vitality indices*—a class in which the centrality of a node is equal to the loss in the value of some graph function that would result from the removal of this node. Everett and Borgatti [28] further studied the class of vitality indices, specifically its subclass in which a graph function is the total centrality of all nodes in a graph for a certain centrality measure. In turn, Skibski, Michalak, and Rahwan [75] characterized the class of game theoretic centralities and its subclasses of separable, induced, and edge-induced game theoretic centralities. Moreover, Baeza-Yates, Boldi, and Castillo [5] established the class of centrality measures based on dumping factors, Brandes and Fleischer [19] focused on centrality measures based on current flows, and Vigna [79] studied the class of centrality measures based on spectral graph theory (which is roughly the same as the class of feedback centralities) and its complicated history.

We end this section by noting the existence of a plethora of papers that consider which centrality measures and on what conditions are fit to particular settings, e.g., psychological networks [20, 36], urban networks [41], protein interaction networks [3], or online social media networks [33, 39]. Another line of research, with similar, or even greater, number of papers, studies the similarities and the differences between centrality measures in general, based on empirical evaluations on various real-life and synthetic networks [6, 14, 47, 56, 60, 71, 74, 78].

Chapter 2

Preliminaries

In this section, we introduce our notation and the basic notions used throughout this work: graphs and centrality measures. To explain the intuition behind this concepts we will often refer to the World Wide Web network example, however the obtained results apply to an arbitrary setting.

2.1 Graphs

We consider directed multigraphs with possible self-loops. Each such multigraph (hereinafter called simply a graph) is a pair, (V,E), where V is a set of nodes and E is a multiset of edges. In the World Wide Web network, nodes would represent pages and the edges hyperlinks between them (note that there can be multiple hyperlinks from one page to another). To emphasize the fact that E is a multiset, we will use double brackets when we list its elements (with possible duplicates), e.g., $\{\!\{(u,u),(u,v),(u,v)\}\!\}$. Also, by $\sqcup$ (and $-$) we denote the sum (and the difference) of multisets. Finally, for any multiset E by $k \cdot E$ we will understand the sum of k copies of multiset E.

Each edge is an ordered pair of nodes in V. Edge (u,v) *starts* in node u, for which it is an *outgoing edge*, and *ends* in node v, for which it is an *incoming edge*. The multiset of all outgoing edges (or incoming edges) of node v is denoted by $\Gamma_v^+(G)$ (or $\Gamma_v^-(G)$), i.e.,

$$\Gamma_v^+(G) = \{\!\{(s,t) \in E : s = v\}\!\} \quad \text{and} \quad \Gamma_v^-(G) = \{\!\{(s,t) \in E : t = v\}\!\}.$$

Also, a set of all edges *incident* to v is a set of all outgoing and incoming edges of v, i.e., $\Gamma_v^\pm(G) = \{\!\{(s,t) \in E : s = v \vee t = v\}\!\}$. If for two nodes, u,v, it holds that $\Gamma_u^+(G) = \{\!\{(u,w) : (v,w) \in \Gamma_v^+(G)\}\!\}$, i.e, they share exactly the same outgoing edges, such nodes are called *out-twins*. The number of outgoing (incoming) edges of node v is its *out-degree* (*in-degree*) and it is denoted by $\deg_v^+(G) = |\Gamma_v^+(G)|$ ($\deg_v^-(G) = |\Gamma_v^-(G)|$). If v does not have any outgoing edges, i.e., $\deg_v^+(G) = 0$, then v is called a *sink*. If v does not have any incoming edges, it is a *source*. Finally, a node is *isolated* if it is a sink and a source at the same time.

The *multiplicity* of edge (u,v), denoted by $\mu_G(u,v)$, is the number of times edge (u,v) appears in E. The *adjacency matrix* of a graph is a $|V| \times |V|$ matrix, $A = (a_{u,v})_{u,v \in V}$, in which entries are equal to the number of edges between nodes, i.e., $a_{u,v} = \mu_G(v,u)$. A real value r is an *eigenvalue* of matrix A if there exists a non-zero vector $\mathbf{x} \in \mathbb{R}^V$ such that $A\mathbf{x} = r\mathbf{x}$; such vector $\mathbf{x}$ is called an *eigenvector*. The *principal eigenvalue*, denoted by λ, is the largest eigenvalue.

A *walk* is a sequence of nodes, $\omega = (\omega(0),\dots,\omega(k))$, such that each consecutive nodes are connected by an edge, i.e., $(\omega(i),\omega(i+1)) \in E$, for every $i \in \{0,\dots,k-1\}$.

The walk *starts* at node $\omega(0)$ and *ends* at $\omega(k)$ and the *length* of the walk is k. The set of all possible walks of length k on graph G is denoted by $\Omega_k(G)$. If the nodes in the walk are pairwise distinct, i.e., $\omega(i) \neq \omega(j)$, for every $i, j \in \{0, \ldots, k\}$ such that $i \neq j$, then the walk is a *path*. If $\omega(0) = \omega(k)$ and all other nodes in the walk are pairwise distinct, then such walk is a *cycle*.

If for two nodes, u and v, there exist a walk from u to v, i.e., the walk of positive length that starts at u and ends at v, then v is a *successor* of u and u is a *predecessor* of v. For $u \neq v$ such that u is a predecessor of v, the length of the shortest walk from u to v is called a *distance* from u to v and it is denoted by $dist(u,v)$. If u is not a predecessor of v we say that $dist(u,v) = \infty$ and we set $dist(u,u) = 0$ for every node u. If this distance is equal to 1, i.e., $(u,v) \in E$, then v is a *direct successor* of u and u is a *direct predecessor* of v. The set of all successors (or predecessors) of node v is denoted by $S_v(G)$ (or $P_v(G)$) and the set of all of its direct successors (or predecessors) is denoted by $S^1_v(G)$ (or $P^1_v(G)$).

A subset of nodes, $U \subseteq V$, forms a *connected component* if each successor and each predecessor of every node in U also belongs to U, i.e., $P_v(G), S_v(G) \subseteq U$, for every $v \in U$, and U is minimal (it is impossible to subtract one or more nodes from U, while preserving this property). A subset of nodes, $U \subseteq V$, forms a *strongly connected component* if for every two nodes in U one is a predecessor and a successor of the other, i.e., $u \in P_v(G) \cap S_v(G)$, for every $u, v \in U$, and U is maximal (it is impossible to add one or more nodes to U from $V \setminus U$, while preserving this property). We say that graph $G = (V,E)$ is *(strongly) connected* if V is a (strongly) connected component.

We also consider *node weights* that can be used to include additional information about nodes in a graph. For example, in a context of World Wide Web network, node weights can model personal preferences of a user [62], how well a page fits into a given topic [37], or the fact that a page is trusted [35]. If no such information is available, one can assume uniform weights for all nodes. Formally, we define a *weighted graph* as a pair (G,b) where $G = (V,E)$ is a graph and b is a node weights function $b : V \to \mathbb{R}_{\geq 0}$ that assigns non-negative weight to each node. Also, to denote small weighted graphs, we will use the following simplified notation:

$$(G,b) = \Big((\{v_1, \ldots, v_n\}, \{\!\{e_1, \ldots, e_m\}\!\}), [b_1, \ldots, b_n] \Big)$$

which means $G = (\{v_1, \ldots, v_n\}, \{\!\{e_1, \ldots, e_m\}\!\})$ and $b(v_i) = b_i$ for $i \in \{1, \ldots, n\}$. The set of all possible weighted graphs will be denoted by $\mathcal{G}$.

Let us introduce some additional shorthand notation. For a subset of nodes $U \subseteq V$ by b_U we will understand node weights b with the domain restricted to nodes in U and by b_{-U} restricted to the set of nodes $V \setminus U$. If U contains one element, i.e., $U = \{u\}$, we will skip parenthesis and simply write b_u and b_{-u}. Also, for a constant $x \in \mathbb{R}_{\geq 0}$, we define $x \cdot b$ as follows: $(x \cdot b)(v) = x \cdot b(v)$, for every $v \in V$. Furthermore, for every two node weights with possibly different domains, $b : V \to \mathbb{R}_{\geq 0}, b' : V' \to \mathbb{R}_{\geq 0}$, we define $b + b' : V \cup V' \to \mathbb{R}_{\geq 0}$ as $(b+b')(v) = b(v) + b'(v)$, if $v \in V \cap V'$, $(b + b')(v) = b(v)$, if $v \in V \setminus V'$, and $(b + b')(v) = b'(v)$, if $v \in V' \setminus V$, for every $v \in V \cup V'$. For example, $(b_{-v} + 2b_v)$ are node weights obtained from b by doubling weight of node v. Finally, for every graph G and node weights b, by $b(G)$ we denote the sum of all node weights in a graph.

Three particular types of node weights are used throughout this work: *uniform node weights*, $\mathbf{1}$, *zero node weights*, $\mathbf{0}$, and *unit node weights centered on v*, i.e., $\mathbb{1}_v$. Uniform node weights, $\mathbf{1}$, give each node weight equal to one, i.e., $\mathbf{1}(v) = 1$, for every $v \in V$. Similarly, in zero node weights, $\mathbf{0}$, each node has zero weight, i.e.,

$\mathbf{0}(v) = 0$, for every $v \in V$. Finally, for every $v \in V$, unit node weights centered on v, $\mathbb{1}_v$, are such node weights that $\mathbb{1}_v(v) = 1$ and $\mathbb{1}_v(u) = 0$, for every $u \in V \setminus \{v\}$.

For two graphs, $G = (V, E)$ and $G' = (V', E')$, and node weights b, b' graph (G', b') is *isomorphic* to (G, b) with *isomorphism* $f : V \to V'$, if f is a bijection and it holds that $E' = \{\!\{(f(u), f(v)) : (u, v) \in E\}\!\}$ and $b'(f(v)) = b(v)$, for every $v \in V$. Two graphs, $G = (V, E)$ and $G' = (V', E')$, are called *disjoint* if their sets of nodes are disjoint, i.e., $V \cap V' = \emptyset$. For such graphs their *sum* is defined as $G + G' = (V \cup V', E \sqcup E')$. For out-twins u and v the operation of *redirecting node u into node v* results in a graph with node u removed and its incoming edges along with its weight transferred to node v. For example, in the setting of World Wide Web network this can be obtained by setting an URL redirection on page u to page v, which results in automatic transfer of all internet traffic to page v. Formally,

$$R_{u \to v}(G, b) = \Big(\big(V \setminus \{u\},\ E - \Gamma^{\pm}_u(G) \sqcup \{\!\{(v, w) : (v, u) \in \Gamma^{-}(G) \wedge v \neq u\}\!\}\big),\ b_{-u} + b(u) \cdot \mathbb{1}_v\Big).$$

2.2 Centrality Measures

A *centrality measure* is a function, F, that for a given graph, $G = (V, E)$, node weights b, and node $v \in V$ returns a real non-negative value $F_v(G, b)$ that represents the importance of node v in weighted graph (G, b).

For directed graphs, most centralities can be defined in two ways: by focusing on predecessors or, more generally, paths ending at a given node, or by focusing on successors and paths starting at a given node. In this work we assume the former version. However, we note that via symmetry, all results can be directly translated to the latter one.

2.2.1 Classic Centrality Measures

Classic centrality measures are based on distances between nodes. Most of them were proposed for graphs without node weights. However, they can usually be easily adapted to this richer setting [53]. In such a case, we can talk about *personalized* [87] or *weighted* [61] centrality measures.

The simplest, yet one of the most popular centrality measures, is *degree centrality* [59]. It assesses a node simply by looking at the number of its incoming edges. Formally, degree centrality is defined as

$$D_v(G) = |\Gamma^{-}_v(G)| = \sum_{u \in P^1_v(G)} \mu_G(u, v).$$

In *personalized degree centrality* instead of taking simply a number of incoming edges, we sum the weights of their starts. Formally,

$$D_v(G, b) = \sum_{u \in P^1_v(G)} b(u) \cdot \mu_G(u, v).$$

Closeness centrality [8] aims to find nodes which are at the center of a graph. To this end, for each node it sums the distances from all other nodes in the graph. Now, nodes with a small total sum, i.e., nodes which are close to all other nodes, are considered most central. Formally, closeness centrality is defined for strongly connected graphs as

$$C_v(G) = \frac{1}{\sum_{u \in V \setminus \{v\}} dist(u, v)}.$$

In *personalized closeness centrality* distance from each node is multiplied by the weight of this node. In this way, the distances from nodes with large weights are more significant. Formally,

$$C_v(G,b) = \frac{1}{\sum_{u \in V \setminus \{v\}} b(u) \cdot dist(u,v)}.$$

Decay centrality [43] is a modification of closeness centrality that works for arbitrary graphs, not necessarily strongly connected. Here, instead of looking at the sum of distances, each node at distance k contributes a^k for some $a \in (0,1)$. Formally, for a decay factor $a \in (0,1)$, decay centrality is defined as

$$Y_v(G) = \sum_{u \in V \setminus \{v\}} a^{dist(u,v)}.$$

Here, we assume that if $dist(u,v) = \infty$, then $a^{dist(u,v)} = 0$. Decay centrality can also be considered an extension of degree centrality that counts not only direct predecessors, but also further predecessors with decreasing weights. *Personalized decay centrality* is defined as

$$Y_v(G,\beta) = \sum_{u \in V} b(u) \cdot a^{dist(u,v)}. \tag{2.1}$$

Personalized decay centrality introduces two modifications to the original definition. Firstly, the contribution of node u to the centrality of v (i.e., $a^{dist(u,v)}$) is now multiplied by the weight of u. Secondly, we now sum over all nodes ($\sum_{u \in V}$), rather than over all nodes other than v ($\sum_{u \in V \setminus \{v\}}$). To understand the rationale behind the latter modification, consider an extreme scenario in which only a single node, say v, has a positive weight. Here, if we sum over all nodes other than v, then any node with a connection to v would have a positive centrality, whereas v itself would have a centrality equal to zero, as all nodes not connected to v—a rather unintuitive outcome in most interpretations of node weights. Note that if all nodes have unit weights, then $Y_v(G,\mathbf{1}) = Y_v(G) + 1$.

Finally, *betweenness centrality* [29] is also based on the notion of shortest paths. However, its goal is to measure how often a specific node is an intermediary between other nodes. To this end, for every pair of nodes, s,t, it takes the fraction of shortest paths between them that goes through a node in question, and then sums it for all such pairs. Formally, betweenness centrality is defined as

$$B_v(G) = \sum_{s,t \in V \setminus \{v\} : \sigma_{st} \neq 0} \frac{\sigma_{st}(v)}{\sigma_{st}},$$

where σ_{st} is the number of shortest paths from s to t and $\sigma_{st}(v)$ is the number of such shortest paths that goes through v. The condition that $\sigma_{st} \neq 0$ is necessary in order to make the definition correct for graphs which are not strongly connected. *Personalized betweenness centrality* is defined as

$$B_v(G,\beta) = \sum_{s,t \in V : \sigma_{st} \neq 0} b(s)b(t) \cdot \frac{\sigma_{st}(v)}{\sigma_{st}}.$$

Firstly, the fractions of shortest paths between from s to t that pass through v are multiplied by the weights of nodes s and t. In this way, being an intermediary on the path between two nodes with large weights is more significant. Secondly, similarly to personalized decay centrality, we also include paths starting and ending at v, so that the weight of v can positively affect the centrality of v.

2.2.2 Feedback Centralities

Feedback centralities is the class of centrality measures on which we put the main focus in this work. These centrality measures assess the importance of a node by looking at its direct predecessors and their importance.

Arguably, the most classic feedback centrality is *Eigenvector centrality* [16]. Here, the importance of a node is proportional to the total importance of its direct predecessors. Formally, eigenvector centrality is defined through the recursive equation:

$$EV_v(G,b) = \frac{1}{\lambda} \sum_{u \in P_v^1(G)} \mu_G(u,v) \cdot EV_u(G,b). \tag{2.2}$$

Observe that the system of these recursive equations does not have a unique solution, as each valid solution multiplied by a scalar is still a valid solution. Hence, some additional normalization condition is usually added to make the centrality measure well defined. For example, it is often assumed that the sum of centralities of all nodes is equal to 1 or $|V|$. In this paper, we use a normalization that stems from the walk interpretation of eigenvector centrality, which is more consistent with other feedback centralities. It is discussed in detail in Section 2.2.3. Eigenvector centrality is usually defined only for strongly connected graphs, because otherwise the solution to the system of recursive equations is not unique even with the condition on the sum of all centralities. Our normalization condition allows us to relax this constraint by allowing also sums of disjoint strongly connected graphs with the same principal eigenvalue. We denote the class of all such graphs by $\mathcal{G}^{EV}$.

Another centrality based on a similar principle is *Katz centrality* [46]. Here, the importance of a node is mostly determined by the total importance of its direct predecessors. However, an additional small basic importance is added to every node. In this work, the node weights represent such a basic importance, but if weights are not provided, a fixed positive constant is used. Formally, for a decay factor $a \in \mathbb{R}_{\geq 0}$, Katz centrality is defined as a unique function that satisfies the following recursive equation:

$$K_v^a(G,b) = a \cdot \left(\sum_{u \in P_v^1(G)} \mu_G(u,v) \cdot K_u^a(G,b) \right) + b(v). \tag{2.3}$$

Adding a basic importance shifts the emphasis from the total importance of the direct predecessors back to their number. This is because each edge (u,v) contributes to the centrality of v an additional value $a \cdot b(u)$ which is independent of the position or the importance of a predecessor. That is why Katz centrality is sometimes seen as a middle-ground between degree and eigenvector centralities. For a fixed a, Katz centrality is uniquely defined for all graphs with $\lambda < 1/a$. We denote the class of such graphs by $\mathcal{G}^{K(a)}$.

In both eigenvector and Katz centralities, the whole importance of a node is "copied" to all of its direct successors. In turn, in *Seeley index* [72], which is also known as *Katz prestige* [43] or *simplified PageRank* [62], a node splits its importance equally among its successors. Hence, the importance of predecessors is divided by their out-degree. Formally, Seeley index is defined by the following recursive equation:

$$SI_v(G,b) = \sum_{u \in P_v^1(G)} \frac{\mu_G(u,v)}{\deg_u^+(G)} \cdot SI_u(G,b). \tag{2.4}$$

Similarly to eigenvector centrality, this system of equations does not have a unique solution. We will discuss our normalization condition in Section 2.2.3. Seeley index is also usually defined only for strongly connected graphs. In this work, we relax this assumption and consider sums of disjoint strongly connected graphs; we denote the class of all such graphs by $\mathcal{G}^{SI}$.

Finally, *PageRank* [62] was proposed as a modification of Seeley index, with the addition of a basic importance to each node. In this way, for a decay factor $a \in [0,1)$, PageRank is defined for all graphs as a unique solution to the system of the following recursive equations:

$$PR^a_v(G,b) = a \cdot \left(\sum_{u \in P^1_v(G)} \frac{\mu_G(u,v)}{\deg^+_u(G)} \cdot PR^a_u(G,b) \right) + b(v). \tag{2.5}$$

2.2.3 Walk Interpretations of Feedback Centrality Measures

In this section we show, how feedback centralities can be alternatively defined using walks on a graph.

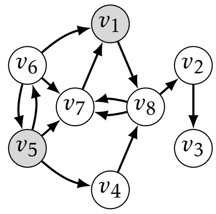
To illustrate the way PageRank works, Page et al. [62] proposed the *random surfer model* in which PageRank of a node is identified with the average time spent in this node in the infinite random walk on the entire network. Here, building on the work of Bianchini, Gori, and Scarselli [9] we consider slightly different walk interpretation of PageRank that we name *busy random surfer model*, which we then extend to other feedback centralities.

Imagine a surfer that browses the pages on the World Wide Web network. At step 0, she starts from the randomly chosen page with the probability of choosing each page proportional to its basic importance (node weight). Then, in each step, with probability a she chooses, uniformly at random, one hyperlink on the page she currently sees and follows it to the next page. At the same time, with probability $1-a$ (or if she arrives at a sink) the surfer gets bored and stop browsing all together. In such a walk, the PageRank of a page is equal to the expected number of times the surfer visits the page, multiplied by the total basic importance of all pages (sum of all node weights in a graph).

To formalize this model, for a graph $G = (V,E)$ and decay factor $a \in [0,1]$, let us define the probability that node $v \in V$ is visited at step $t \in \mathbb{N}$ as

$$p^a_{G,b}(v,t) = \sum_{\omega \in \Omega_t(G): \omega(t)=v} \frac{b(\omega(0))}{b(G)} \cdot \prod_{i=0}^{t-1} \frac{a \cdot \mu_G\big(\omega(i), \omega(i+1)\big)}{\deg^+_{\omega(i)}(G)}, \quad \text{if } b(G) > 0, \tag{2.6}$$

and $p^a_{G,b}(v,t) = 0$, otherwise. Let us analyze the right hand side of this equation. Firstly, $b(\omega(0))/b(G)$ is the probability that the walk starts at node $\omega(0)$. The fraction $\mu_G\big(\omega(i), \omega(i+1)\big)/\deg^+_{\omega(i)}(G)$ is the probability of choosing an edge going from node $\omega(i)$ to $\omega(i+1)$ out of all outgoing edges of node $\omega(i)$ assuming we pick them uniformly at random. If we multiply this fraction by a, the probability that the walk does not end at step i, we obtain the probability that the walk in node $\omega(i)$ at step i moves to a node $\omega(i+1)$ in the next step. Now, taking the product of these fractions over all steps from 0 to $t-1$ and multiplying it by the probability that the random walk starts at $\omega(0)$, i.e., $b(\omega(0))/b(G)$, gives us the probability that the random walk follows ω up to step t. Thus, summing over all possible walks of length t that end at v we indeed obtain the probability that node v is visited at step t.



(G, b)

Figure 2.1: An example graph. Grey nodes have weights equal to 1 and the weight of white nodes is 0.

Example 1. *Consider graph (G,b) from Fig. 2.1. Let us calculate the value of $p^a_{G,b}(v_1,3)$, i.e., the probability that node v_1 is visited at step 3. To this end, let us consider all walks of length 3 that end at v_1. There are five such walks: $\omega_1 = (v_1, v_8, v_7, v_1)$, $\omega_2 = (v_4, v_8, v_7, v_1)$, $\omega_3 = (v_5, v_6, v_7, v_1)$, $\omega_4 = (v_6, v_5, v_6, v_1)$, and $\omega_5 = (v_6, v_5, v_7, v_1)$. Observe that walks ω_2, ω_4, and ω_5 start at a node with zero weight, i.e., we have $b(\omega_2(0)) = b(\omega_4(0)) = b(\omega_5(0)) = 0$, hence their input in the sum on the right hand side of Eq. (2.6) will be also zero.*

The probability that the random walk starts at v_1 is equal to $b(v_1)/b(G) = 1/2$. Node v_1 has only one outgoing edge, to v_8, hence conditional on the random walk continuing, for which there is a probability a, the random walk goes to v_8 with probability 1. Two out of three outgoing edges of v_8 go to v_7, thus the probability that the random walk moves next to v_7 is $a \cdot 2/3$. Finally, v_7 again has only one outgoing edge, to v_1, thus the random walk moves from v_7 to v_1 with probability a. Therefore, the probability that the random walk starts with sequence ω_1 is equal to $1/2 \cdot a \cdot 2a/3 \cdot a = a^3/3$.

$$\textit{sequence } (\omega_1): \quad \circ \xrightarrow{1/2} v_1 \xrightarrow{a} v_8 \xrightarrow{2a/3} v_7 \xrightarrow{a} v_1.$$

Now, the probability that the random walk starts at v_5 is equal to $b(v_5)/b(G) = 1/2$. From there, it goes to v_6 with probability $a/3$ (node v_5 has three outgoing edges and one goes to v_6). Then it moves to v_7 with probability $a/2$. And again goes from v_7 to v_1 with probability a. Thus, we get that the probability that the random walk starts with sequence ω_3 is equal to $1/2 \cdot a/3 \cdot a/2 \cdot a = a^3/12$.

$$\textit{sequence } (\omega_3): \quad \circ \xrightarrow{1/2} v_5 \xrightarrow{a/3} v_6 \xrightarrow{a/2} v_7 \xrightarrow{a} v_1.$$

Summing up, we get that $p^a_{G,b}(v_1, 3) = a^3/3 + a^3/12 = 5a^3/12$.

Using this walk interpretation we can alternatively define PageRank of a node as the expected number of visits at the node multiplied by the total weight of all the nodes, i.e.,

$$PR^a_v(G,b) = b(G) \cdot \sum_{t=0}^{\infty} p^a_{G,b}(v,t). \tag{2.7}$$

In the following theorem, we prove that the PageRank defined by Eq. (2.7) indeed satisfies Eq. (2.5). Since PageRank is a unique solution to Eq. (2.5) it means that both definitions are equivalent.

Theorem 1. *For every decay factor $a \in [0,1)$, PageRank defined by Eq. (2.7) satisfies PageRank recursive equation (Eq. (2.5)).*

Proof. Let us denote $F_v(G,b) = b(G) \cdot \sum_{t=0}^{\infty} p^a_{G,b}(v,t)$ and prove that centrality F indeed satisfies Eq. (2.5), for every graph $G = (V,E)$, node weights b, and node $v \in V$.

If $b(G) = 0$, then $F_v(G,b) = 0$ and Eq. (2.5) is trivially satisfied. Thus, let us assume that $b(G) > 0$.

Observe that for each walk $\omega \in \Omega_t(G)$ that ends at v, i.e., $\omega(t) = v$, if $t > 0$, then $\omega(t-1)$ must be one of the direct predecessors of v, say u. The probability that the random walk does not end at step $t-1$ and moves from u to v is $a \cdot \mu_G\big(u,v\big)/\deg_u^+(G)$. Thus, for the value $p^a_{G,b}(v,t)$ we get

$$p^a_{G,b}(v,t) = \sum_{u \in P_v^1(G)} a \frac{\mu_G(u,v)}{\deg_u^+(G)} p^a_{G,b}(u,t-1).$$

Taking the sum over all $t \geq 1$ we get

$$\sum_{t=1}^{\infty} p^a_{G,b}(v,t) = \sum_{u \in P_v^1(G)} a \frac{\mu_G(u,v)}{\deg_u^+(G)} F_u(G,b)/b(G).$$

Observe that $p^a_{G,b}(v,0) = b(v)/b(G)$, thus adding both equations sidewise we obtain

$$\sum_{t=0}^{\infty} p^a_{G,b}(v,t) = b(v)/b(G) + \sum_{u \in P_v^1(G)} a \frac{\mu_G(u,v)}{\deg_u^+(G)} F_u(G,b)/b(G)$$

and by multiplying both sides by $b(G)$ we obtain the thesis. □

Now, the sum in Eq. (2.7) converges for every $a \in [0,1)$. However, for $a = 1$ the sum does not have to converge. In such a case, one can consider a partial sum up to step T, i.e., $b(G) \cdot \sum_{t=0}^{T} p^a_{G,b}(v,t)$, and observe that dividing it by T would only scale the sum for all nodes, not affecting the relations between them. Now, taking the limit of such "partial averages" instead of partial sums results in a value that converges for every node of every weighted graph. As a result, instead of looking at the expected number of visits at a node, we look at the average time spent at that node, which is the stationary distribution of the random walk. In this way, we obtain the walk-based definition of Seeley index, i.e.,

$$SI_v(G,b) = b(G) \cdot \lim_{T \to \infty} \frac{\sum_{t=0}^{T} p^1_{G,b}(v,t)}{T}. \tag{2.8}$$

In the following theorem, we prove that the definition given by Eq. (2.8) is coherent with Seeley index recursive equation. Moreover, we show that Eq. (2.8) implies that the sum of Seeley indices of all the nodes is equal to the sum of node weights in a graph. Hence, for strongly connected graphs the definition by Seeley index recursive equation and normalization condition $\sum_{v \in V} SI_v(G,b) = b(G)$ is equivalent to the walk-based definition of Seeley index.

Theorem 2. *Seeley index defined on $\mathcal{G}^{SI}$ by Eq. (2.8) satisfies Seeley index recursive equation (Eq. (2.4)) and for every graph $G = (V,E)$ and node weights b such that $(G,b) \in \mathcal{G}^{SI}$, it holds that $\sum_{v \in V} SI_v(G,b) = b(G)$.*

Proof. Let us denote $F_v(G,b) = b(G) \cdot \lim_{T \to \infty} \sum_{t=0}^{T} p^1_{G,b}(v,t)/T$ and prove that for every graph $G = (V,E)$, node weights b such that $(G,b) \in \mathcal{G}^{SI}$, and node $v \in V$, it satisfies Seeley index recursive equation (Eq. (2.4)) and that $\sum_{v \in V} F_v(G,b) = b(G)$. If $b(G) = 0$, then $F_v(G,b) = 0$ and Eq. (2.4) is trivially satisfied. Also, $\sum_{v \in V} F_v(G,b) = 0 = b(G)$. Thus, let us assume that $b(G) > 0$.

First, let us focus on the second part of the thesis, i.e., that the sum of centralities of all nodes is equal to the sum of weights of all nodes. To this end, observe that for every $t > 0$, every walk $\omega \in \Omega_t(G)$ that ends at v, i.e., $\omega(t) = v$, must have visited one of the direct predecessors of v, say u, at step $t-1$ and then move to v through edge (u,v). Thus, if we look at the value $p^1_{G,b}(v,t)$ we obtain that

$$p^1_{G,b}(v,t) = \sum_{u \in P^1_v(G)} \frac{\mu_G(u,v)}{\deg^+_u(G)} p^1_{G,b}(u,t-1). \tag{2.9}$$

Let us sum both sides of Eq. (2.9) for all nodes $v \in V$. As a result, on the right hand side, for every $u \in V$, the term $\mu_G(u,v)/\deg^+_u(G) \cdot p^1_{G,b}(u,t-1)$ appears exactly once, for every $v \in S^1_u(G)$. Hence, these terms sum to $p^1_{G,b}(u,t-1)$. In this way, we obtain

$$\sum_{v \in V} p^1_{G,b}(v,t) = \sum_{u \in V} p^1_{G,b}(u,t-1). \tag{2.10}$$

This means that for each step $t \in \mathbb{N}$ the sum of terms $p^1_{G,b}(v,t)$ for all nodes is constant. Since $p^1_{G,bs}(v,0) = b(v)/b(G)$ for every $v \in V$, this sum is always equal to one, i.e., $\sum_{v \in V} p^1_{G,b}(v,t) = 1$. Thus, when we sum it for all $t \in \{0,\dots,T\}$, multiply by $b(G)/T$, and take a limit in the infinity we still obtain that

$$\sum_{v \in V} F_v(G,b) = \sum_{v \in V} \lim_{T\to\infty} b(G) \cdot \sum_{t=0}^{T} \frac{p^1_{G,b}(v,t)}{T} = b(G) \cdot \lim_{T\to\infty} \sum_{t=0}^{T} \frac{\sum_{v \in V} p^1_{G,b}(v,t)}{T} = b(G).$$

Now, let us move to the first part of the thesis, i.e., that centrality F satisfies Seeley index recursive equation (Eq. (2.4)), for every graph $G = (V,E)$, node weights b, and node $v \in V$.

Let us sum both sides of Eq. (2.9) for all $t \in \{1,\dots,T\}$, to get

$$\sum_{t=1}^{T} p^1_{G,b}(v,t) = \sum_{u \in P^1_v(G)} \frac{\mu_G(u,v)}{\deg^+_u(G)} \sum_{t=0}^{T-1} p^1_{G,b}(u,t).$$

Now, if we add $p^1_{G,b}(v,0) = b(v)/b(G)$ to both sides of the equation and multiply each side by $b(G)/T$, we obtain

$$b(G) \cdot \sum_{t=0}^{T} \frac{p^1_{G,b}(v,t)}{T} = \frac{b(v)}{T} + \sum_{u \in P^1_v(G)} \frac{\mu_G(u,v)}{\deg^+_u(G)} b(G) \cdot \sum_{t=0}^{T-1} \frac{p^1_{G,b}(u,t)}{T}$$

When T approaches infinity, $b(v)/T$ approaches zero. Hence,

$$F_v(G,b) = \sum_{u \in P^1_v(G)} \frac{\mu_G(u,v)}{\deg^+_u(G)} b(G) \cdot \lim_{T\to\infty} \sum_{t=0}^{T-1} \frac{p^1_{G,b}(u,t)}{T} \tag{2.11}$$

Thus, it remains to show that $p^1_{G,b}(u,T)/T$ approaches zero as well, for every $u \in V$. To this end, observe that, by Eq. (2.10), the sum of $p^1_{G,b}(u,t)$ for all $u \in V$ is constant for all $t \in \mathbb{N}$. Thus, $p^1_{G,b}(u,t)$ is bounded by this sum. Hence, term $p^1_{G,b}(u,T)/T$ indeed approaches zero as T approaches infinity. Therefore, by Eq. (2.11), we get that $F_v(G) = \sum_{u \in P^1_v(G)} \mu_G(u,v)/\deg^+_u(G) \cdot F_u(G,b)$ which is exactly Seeley index recursive equation (Eq. (2.4)). □

Now, let us move to Katz and eigenvector centralities. Again let us start with the intuition behind their walk interpretations. To this end, let us modify the busy random surfer model and propose the *parallel busy random surfer model*.

Like before, imagine the surfer that starts browsing the World Wide Web network from a random page (with the probability of choosing each page proportional to its basic importance). Then, the surfer reads the whole current page and for each hyperlink on that page there is a probability a that she clicks on it and opens it in a new tab of her browser. Next, when she finishes reading the current page, she closes the tab with that page and moves to the first tab that is opened, where she repeats the process. If there are no opened tabs when she closes the current tab, she ends browsing all together. In such a process, Katz centrality of a page is the expected number of times the surfer reads this page, multiplied by the sum of the basic importance of all pages.

To formalize this model, for every graph $G = (V,E)$ and decay factor $a \in \mathbb{R}_{\geq 0}$, let us define the weighted sum of walks that visit node $v \in V$ at step $t \in \mathbb{N}$ as

$$w^a_{G,b}(v,t) = \sum_{\omega \in \Omega_t(G):\omega(t)=v} \frac{b(\omega(0))}{b(G)} \cdot \prod_{i=0}^{t-1} a \cdot \mu_G\Big(\omega(i), \omega(i+1)\Big), \quad \text{if } b(G) > 0, \tag{2.12}$$

and $w^a_{G,b}(v,t) = 0$, otherwise. The only difference between Eq. (2.12) and Eq. (2.6) is that here, we do not divide $\mu_G(\omega(i),\omega(i+1))$ by $\deg^+_{\omega(i)}$ since we do not choose one of the outgoing edges, but decide upon each of them independently. In other words, $w^a_{G,b}(v,t)$ is $a^t \cdot b(\omega(0))/b(G)$ multiplied by the number of walks of length t that ends at v if we differentiate between the walks that not only moves through different nodes, but also use different edges. Note that although the probabilistic interpretation is well-founded only for $a \in [0,1]$, Eq. (2.12) is valid for any $a \in \mathbb{R}$.

Example 2. *In a similar way to Example 1, let us consider graph (G,b) from Fig. 2.1 and calculate the value of $w^a_{G,b}(v_1,3)$, i.e., the weighted sum of walks that visit node v_1 at step 3. Again we have five walks of length 3 that end at v_1, i.e., $\omega_1 = (v_1,v_8,v_7,v_1)$, $\omega_2 = (v_4,v_8,v_7,v_1), \omega_3 = (v_5,v_6,v_7,v_1), \omega_4 = (v_6,v_5,v_6,v_1)$, and $\omega_5 = (v_6,v_5,v_7,v_1)$. As before, since ω_2, ω_4, and ω_5 start at a node with zero weight, they will have zero input in the sum on the right hand side of Eq.* (2.12).

We start at node v_1 with $b(v_1)/b(G) = 1/2$. Node v_1 has one outgoing edge to v_8. Then, there are two edges from v_8 to v_7. Finally, v_7 again has one outgoing edge to v_1. Therefore, the weighted sum of walks that follow sequence ω_1 is equal to $1/2 \cdot a \cdot 2a \cdot a = a^3$.

$$\textit{sequence } (\omega_1): \quad \circ \xrightarrow{\;1/2\;} v_1 \xrightarrow{\;a\;} v_8 \xrightarrow{\;2a\;} v_7 \xrightarrow{\;a\;} v_1.$$

In sequence ω_3 all consecutive nodes are connected by exactly one edge. Again, at v_5 we have $b(v_5)/b(G) = 1/2$, thus we get that the weighted sum of walks that follow sequence ω_3 is equal to $a^3/2$.

$$\textit{sequence } (\omega_3): \quad \circ \xrightarrow{\;1/2\;} v_5 \xrightarrow{\;a\;} v_6 \xrightarrow{\;a\;} v_7 \xrightarrow{\;a\;} v_1.$$

Summing up, we get that $w^a_{G,b}(v_1,3) = a^3 + a^3/2 = 3a^3/2$.

For every $a < 1/\lambda$, this walk interpretation allows us to alternatively define Katz centrality as the expected number of visits at a node multiplied by the total weights of all nodes, i.e.,

$$K^a_v(G,b) = b(G) \cdot \sum_{t=0}^{\infty} w^a_{G,b}(v,t). \tag{2.13}$$

In the following theorem we prove that Katz centrality defined in such a way indeed satisfies Katz centrality recursive equation (Eq. (2.3)) which implies that both definitions are equivalent.

Theorem 3. *For every decay factor $a \in \mathbb{R}_{>0}$, Katz centrality measure defined on $\mathcal{G}^{K(a)}$ by Eq.* (2.13) *satisfies Katz centrality recursive equation (Eq.* (2.3)*).*

Proof. We prove that centrality measure defined as $F_v(G,b) = b(G) \cdot \sum_{t=0}^{T} w_{G,b}^{a}(v,t)$ satisfies Eq. (2.3), for every graph $G = (V,E)$, node weights b such that $(G,b) \in \mathcal{G}^{K(a)}$, and node $v \in V$. If $b(G) = 0$, then $F_v(G,b) = 0$ and Eq. (2.3) is trivially satisfied. Thus, let us assume that $b(G) > 0$.

Observe that for every walk $\omega \in \Omega_t(G)$ such that $\omega(t) = v$ in order to arrive at v at step $t \geq 1$, it must visit a direct predecessor of v, say u, at step $t-1$ and then follow edge (u,v). Thus, for $w_{G,b}^{a}(v,t)$ we obtain that

$$w_{G,b}^{a}(v,t) = a \cdot \sum_{u \in P_v^1(G)} \mu_G(u,v) \cdot w_{G,b}^{a}(u,t-1).$$

Summing both sides for all $t \in \{1,2,\dots\}$ we get

$$\sum_{t=1}^{\infty} w_{G,b}^{a}(v,t) = a \cdot \sum_{u \in P_v^1(G)} \mu_G(u,v) \cdot \sum_{t=0}^{\infty} w_{G,b}^{a}(u,t)$$

Finally, let us add $w_{G,b}^{a}(v,0) = b(v)/b(G)$ to both sides of the equation and multiply both sides by $b(G)$ to obtain that $F_v(G,b) = a \cdot (\sum_{(u,v) \in \Gamma_v^-(G)} \mu_G(u,v) \cdot F_u(G)) + b(v)$ which is Katz centrality recursive equation (Eq. (2.3)). □

The sum in Eq. (2.13) converges if $a < 1/\lambda$. In the case of $a = 1/\lambda$, the sum does not converge. However, in a way similar to the way we defined Seeley index, instead of looking at the expected number of visits at a node, we can look at the average time spent at it. Thus, if we take the limit of "partial averages" of the weighted sum of walks that visit a node, we obtain the walk definition of eigenvector centrality, i.e.,

$$EV_v(G) = b(G) \cdot \lim_{T \to \infty} \sum_{t=0}^{T} \frac{w_{G,b}^{1/\lambda}(v,t)}{T}. \tag{2.14}$$

In the following theorem, we prove that eigenvector centrality defined in such a way indeed satisfies Eigenvector centrality recursive equation (Eq. (2.2)).

Theorem 4. *Eigenvector centrality defined on $\mathcal{G}^{EV}$ by Eq.* (2.14) *satisfies Eigenvector centrality recursive equation (Eq.* (2.2)*).*

Proof. We prove that the centrality $F_v(G,b) = b(G) \cdot \lim_{T \to \infty} \sum_{t=0}^{T} w_{G,b}^{1/\lambda}(v,t)/T$ satisfies eigenvector centrality recursive equation (Eq. (2.2)), for every graph $G = (V,E)$, node weights b such that $(G,b) \in \mathcal{G}^{EV}$, and node $v \in V$. If $b(G) = 0$, then $F_v(G,b) = 0$ and Eq. (2.2) is trivially satisfied. Thus, let us assume that $b(G) > 0$.

Then, observe that for every walk $\omega \in \Omega_t(G)$ such that $\omega(t) = v$, the walk must have visited a direct predecessor of v, say u, at step $t-1$ and then move through edge (u,v). Hence, for value $w_{G,b}^{1/\lambda}(t)$ we get

$$w_{G,b}^{1/\lambda}(v,t) = \frac{1}{\lambda} \sum_{u \in P_v^1(G)} \mu_G(u,v) \cdot w_{G,b}^{1/\lambda}(u,t-1).$$

If we sum both sides for all $t \in \{1,\ldots,T\}$ we obtain

$$\sum_{t=1}^{T} w_{G,b}^{1/\lambda}(v,t) = \frac{1}{\lambda} \sum_{u \in P_v^1(G)} \mu_G(u,v) \cdot \sum_{t=0}^{T-1} w_{G,b}^{1/\lambda}(u,t).$$

Adding $w_{G,b}^{1/\lambda}(v,0) = b(v)/b(G)$ and multiplying both sides by $b(G)/T$ yields

$$b(G) \cdot \sum_{t=0}^{T} \frac{w_{G,b}^{1/\lambda}(v,t)}{T} = \frac{b(v)}{T} + \sum_{u \in P_v^1(G)} \frac{\mu_G(u,v)}{\lambda} b(G) \cdot \sum_{t=0}^{T-1} \frac{w_{G,b}^{1/\lambda}(u,t)}{T}. \tag{2.15}$$

As T approaches infinity, $b(v)/T$ approaches zero. In order to show that value $w_{G,b}^{1/\lambda}(u,T)/T$ approaches zero as well, observe that for each step $t \in \mathbb{N}$, the vector of values $w_{G,b}^{1/\lambda}(u,t)$ for all nodes $u \in V$ is a vector of values $w_{G,b}^{1/\lambda}(u,0)$ for all $u \in V$ multiplied t times by adjacency matrix of G and divided t times by λ. The norm of such vector is bounded [57], thus its coordinates are also bounded. Hence, the value $w_{G,b}^{1/\lambda}(u,T)/T$ indeed approaches zero as T approaches infinity. Thus, taking limit in Eq. (2.15) we obtain $F_v(G,b) = 1/\lambda \cdot \sum_{u \in P_v^1(G)} \mu_G(u,v) \cdot F_u(G,b)$ which is exactly eigenvector centrality recursive equation (Eq. (2.2)). □

In the following chapters, if not stated otherwise, we use Eqs. (2.7)–(2.14) as the definition of all four feedback centralities.

Chapter 3

Axiomatization of PageRank

We begin with the axiomatic characterization of arguably the most important walk-based centrality measure in computer science—PageRank. More in detail, in this chapter, we propose six simple properties: *Node Deletion*, *Edge Deletion*, *Edge Multiplication*, *Edge Swap*, *Node Redirect*, and *Baseline*, and prove that PageRank is the only centrality measure that satisfies all of them.

We begin by introducing and explaining our axioms in Section 3.1. Then, in Section 3.2, we prove that they uniquely characterize PageRank and that they are independent. Finally, in Section 3.3, we compare our characterization with two axiomatizations of Seeley Index from the literature.

The content of this chapter is an extended version of the paper published in the proceedings of the IJCAI-18 conference [84].

3.1 Axioms

Let us present our main result of this chapter: the axiomatic characterization of PageRank. Specifically, we introduce six simple properties, or *axioms*, that PageRank satisfies. Some of these axioms are satisfied also by other known centrality measures. However, what is important, PageRank is the unique centrality measure that satisfies all of them, which is stated in Theorem 5. The axioms are:

- ***Node Deletion*** *(removing an isolated node does not affect centralities of other nodes in the graph): For every graph $G = (V,E)$, node weights b, and isolated node $u \in V$, it holds that*

$$F_v(G,b) = F_v((V \setminus \{u\},\ E),\ b_{-u}), \quad \textit{for every } v \in V \setminus \{u\}.$$

- ***Edge Deletion*** *(removing an edge does not affect centralities of nodes which are not successors of the start of this edge): For every graph $G = (V,E)$, node weights b, and edge $(u,w) \in E$, it holds that*

$$F_v(G,b) = F_v((V,\ E - \{\!\{(u,w)\}\!\}),\ b), \quad \textit{for every } v \in V \setminus S_u(G).$$

- ***Edge Multiplication*** *(creating additional copies of the outgoing edges of a node does not affect the centrality of any node in the graph): For every graph $G = (V,E)$, node weights b, node $u \in V$, and $k \in \mathbb{N}$, it holds that*

$$F_v(G,b) = F_v((V,\ E \sqcup k \cdot \Gamma_u^+(G)),\ b), \quad \textit{for every } v \in V.$$

- ***Edge Swap*** *(swapping ends of two outgoing edges of nodes with equal centralities and out-degrees does not affect the centrality of any node in the graph): For every graph $G = (V, E)$, node weights b, and edges $(u, u'), (w, w') \in E$ such that $F_u(G, b) = F_w(G, b)$ and $\deg^+_u(G) = \deg^+_w(G)$, it holds that*

$$F_v(G, b) = F_v((V,\ E - \{\!\!\{(u, u'), (w, w')\}\!\!\} \sqcup \{\!\!\{(u, w'), (w, u')\}\!\!\}),\ b), \quad \textit{for every } v \in V.$$

- ***Node Redirect*** *(redirecting a node into its out-twin sums up their centralities and does not affect the centrality of other nodes in the graph): For every graph $G = (V, E)$, node weights b, and out-twins $u, w \in V$, it holds that*

$$F_v(G, b) = F_v(R_{u\to w}(G, b)), \quad \textit{for every } v \in V \setminus \{u, w\},$$

and $F_u(G, b) + F_w(G, b) = F_w(R_{u\to w}(G, b))$.

- ***Baseline*** *(the centrality of an isolated node is equal to its weight): For every graph $G = (V, E)$, node weights b, and isolated node $v \in V$, it holds that $F_v(G, b) = b(v)$.*

The first five axioms are *invariance axioms*. Each invariance axiom is characterized by a graph operation, additional conditions on a graph, and a set of nodes. Given this, the axiom states that if the conditions are satisfied, then the graph operation does not affect the centrality of nodes in question. Each axiom is named after the graph operation it considers. Node Deletion and Edge Deletion concerns removing an isolated node or an edge. Edge Multiplication and Edge Swap focus on edge modifications: the former axiom considers replacing each outgoing edge of one node by multiple copies; the later one concerns swapping the ends of outgoing edges of two nodes with the same centrality and out-degree. Finally, Node Redirect considers redirecting a node into its out-twin.

The invariance axioms characterize PageRank up to a scalar multiplication, i.e., they are satisfied not only by PageRank, but also by PageRank multiplied by some constant. In order to uniquely characterize PageRank, the last, sixth axiom, called Baseline, specifies the centrality of an isolated node. The following theorem presents our main result of this chapter.

Theorem 5. *A centrality measure satisfies Node Deletion, Edge Deletion, Edge Multiplication, Edge Swap, Node Redirect, and Baseline if and only if it is PageRank.*

The proof of Theorem 5 is presented in Section 3.2. Before, we provide the interpretation of the axioms with respect to the hyperlink network.

The first two axioms—Node Deletion and Edge Deletion—identify elements (pages and links) which are irrelevant for the importance of a page in question. Node Deletion considers a page with no links or backlinks (e.g., a resource hidden on the server). The axiom states that such a page does not have any impact on the rest of the network and its removal does not affect the importance of all the remaining pages.

For the Edge Deletion, imagine that there is a page A from which it is not possible to reach page B through a sequence of links (the studies show that such pairs of pages are very common [22]). For example, imagine that page A has only links to its subpages that do not have external links themselves. The axiom states that the links on A do not have an impact on the importance of B. Hence, if we remove one of them, it will not affect the importance of B. In particular, if A cannot be reached from A, i.e., if it is not possible to enter page A again after leaving through one of its links, then links of A does not affect also its own importance. We note that Edge Deletion combined with Node Deletion implies that the importance of a page depends solely on the part of the WWW network from which this page can be reached.

Our next axiom, Edge Multiplication, concerns multiplying the whole content of the page several times. This operation naturally increases the number of backlinks for many pages. The axiom states that the importance of these pages, as well as all other pages in the network, do not change. This means that the absolute number of links on a page does not matter as long as the proportion of links to other pages remains the same. Looking from a different perspective, Edge Multiplication can be interpreted as robustness to manipulations by creating a large number of backlinks. Regardless of the number of links, the impact of a page is fixed to some extent. As there is no cost of creating a link on the World Wide Web, avoiding such a manipulation lays at the foundation of PageRank.

For the next axiom, Edge Swap, consider a case where there are two equally important pages with an equal number of links. The axiom states that the links from these pages have equal impact. It does not matter for the importance of any page from which of these two pages it has a backlink. Hence, ends of edges can be swapped without affecting the importance of pages they link to and any other pages in the network.

For Node Redirect, imagine that there are two copies of the same page, i.e., two pages with identical content and links. Their backlinks, however, can differ. Node Redirect states that URL redirecting, i.e., removing one of the copies and redirecting its incoming traffic to the other one, does not change the importance of other pages. Moreover, the total importance of both pages will also remain intact. At a high level, this axiom concerns a simple manipulation technique through creating several copies of the same page: the axiom states that merging them into one page does not change importance of any other page in the network.

Edge Multiplication and Node Redirect identify two manipulation techniques that do not affect PageRank of a page. However, we note that PageRank is not resilient to other types of manipulations. In particular, by modifying links a page may increase its PageRank (in Chapter 4, we discuss an alternative centrality measure that is resistant to such manipulation).

Our last axiom, Baseline, concerns a page without any links nor backlinks. Such a page does not profit from the network structure, as it is not connected to any other page. Hence, the axiom states that its centrality is equal to its basic importance.

Example 3. *As an illustration of our axioms consider graphs in Fig. 3.1. Assume that centrality measure F satisfies our five invariance axioms. Then, the centrality of node v_1 is not affected by the operations that transform graph G_e into G_a and equals the sum of centralities of v_1 and v_7 in graph G_f, i.e., $F_{v_1}(G_a, b) = F_{v_1}(G_f, b) + F_{v_7}(G_f, b)$. More in detail:*

- $F_{v_1}(G_a, \mathbf{1}) = F_{v_1}(G_b, \mathbf{1})$ *from* Node Deletion. *Node v_3 is isolated in graph G_b, hence its deletion does not affect the centralities of the remaining nodes.*
- $F_{v_1}(G_b, \mathbf{1}) = F_{v_1}(G_c, \mathbf{1})$ *from* Edge Deletion. *The only successor of node v_4 is node v_3, thus deleting edge (v_4, v_3) does not affect centralities of nodes other than v_3.*
- $F_{v_1}(G_c, \mathbf{1}) = F_{v_1}(G_d, \mathbf{1})$ *from* Edge Multiplication, *since G_d is obtained from G_c by doubling the edges of node v_6.*
- $F_{v_1}(G_d, \mathbf{1}) = F_{v_1}(G_e, \mathbf{1})$ *from* Edge Swap. *Nodes v_5 and v_6 have both 4 outgoing edges. If they have the same centrality (note that they have incoming edges only*

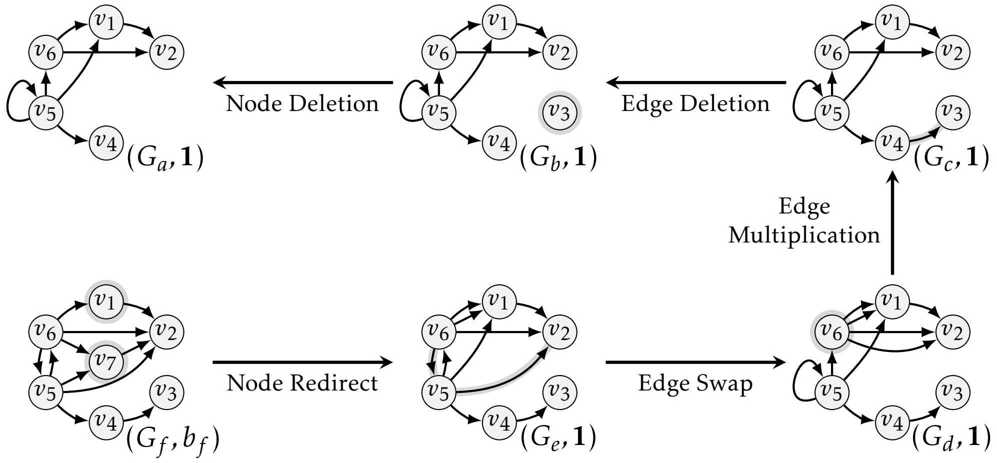


Figure 3.1: An illustration of invariance axioms. We set $b_f(v_i) = 1$ for $i \in \{1, \ldots, 6\}$ and $b_f(v_7) = 0$.

from themselves), then exchanging edges $(v_5, v_2), (v_6, v_5)$ for edges $(v_6, v_2), (v_5, v_5)$ does not affect the centrality of any node.

- $F_{v_1}(G_e, \mathbf{1}) = F_{v_1}(G_f, b_f) + F_{v_7}(G_f, b_f)$ *from* Node Redirect. *Nodes v_1 and v_7 both have only one edge to node v_2, hence they are out-twins. Therefore, redirecting v_7 into v_1 does not affect the centrality of other nodes and the centrality of node v_1 becomes the sum of the centralities of v_1 and v_7.*

3.2 Proof of Uniqueness

In this section, we present the proof of Theorem 5. It consists of two parts. In the first part (Section 3.2.1), we prove that for every decay factor $a \in [0, 1)$, PageRank satisfies all the six axioms. In the second part (Section 3.2.2), we prove that if a centrality measure satisfies all the six axioms, then it must be equal to PageRank for some decay factor $a \in [0, 1)$. Both parts combined imply Theorem 5. Additionally, in Section 3.2.3, we show that all the six axioms are independent and necessary in our characterization.

3.2.1 Part 1: PageRank Satisfies Axioms

In this part, we show that PageRank, for every decay factor $a \in [0, 1)$, indeed satisfies Node Deletion, Edge Deletion, Edge Multiplication, Edge Swap, Node Redirect, and Baseline.

All axioms, except for Baseline, are invariance axioms, so we need to show that a specific graph operation does not affect PageRank of the node in question. To this end, we will either look at the walk interpretation of PageRank or PageRank recursive equation (Eq. (2.5)). In the former case, based on Theorem 1, it is enough to show that a specific operation does not affect the expected number of visits in the busy random surfer model (or, if the sum of node weights decreases, then the expected number of visits increases accordingly); we will do this for four axioms: Node Deletion, Edge Deletion, Edge Multiplication, and Node Redirect. In the later case, we show that the system of recursive equations after the graph modification has the same solution; we will do this for Edge Swap. Finally, Baseline easily follows from PageRank recursive equation.

Fix an arbitrary graph $G = (V,E)$ and node weights b. We consider each axiom separately.

Node Deletion

Let u be an isolated node and v be an arbitrary node other than u. Consider a graph obtained from (G,b) by removing node u: $(G',b') = ((V \setminus \{u\},E),b_{-u})$. We have to prove that $PR^a_v(G,b) = PR^a_v(G',b')$. To this end, based on Theorem 1, it is enough to prove that the probability of visit at node v at time $t \geq 0$ multiplied by the sum of node weights is the same for both graphs, i.e.,

$$p^a_{G,b}(v,t) \cdot b(G) = p^a_{G',b'}(v,t) \cdot b'(G'). \tag{3.1}$$

Observe that if $b(G) = 0$, then $b'(G') = 0$ as well. Thus, $p^a_{G,b}(v,t) \cdot b(G) = 0 = p^a_{G',b'}(v,t) \cdot b'(G')$, for every $t \in \mathbb{N}$. Hence, let us assume that $b(G) > 0$. Then, each walk $\omega \in \Omega_t(G)$ that ends at v, i.e., $\omega(t) = v$, belongs also to $\Omega_t(G')$ and vice versa. Let us take an arbitrary such walk, ω. Since u is isolated, we know that it is not possible for ω to visit u at any step. Moreover, for every pair of nodes $s,t \in V \setminus \{u\}$, we have that both $\mu_G(s,t) = \mu_{G'}(s,t)$ and $\deg^+_s(G) = \deg^+_s(G')$. Thus,

$$\prod_{i=0}^{t-1} \frac{a \cdot \mu_G\big(\omega(i),\omega(i+1)\big)}{\deg^+_{\omega(i)}(G)} = \prod_{i=0}^{t-1} \frac{a \cdot \mu_{G'}\big(\omega(i),\omega(i+1)\big)}{\deg^+_{\omega(i)}(G')}.$$

Multiplying both sides by $b(\omega(0))$ we obtain that

$$b(G) \cdot \frac{b(\omega(0))}{b(G)} \cdot \prod_{i=0}^{t-1} \frac{a \cdot \mu_G\big(\omega(i),\omega(i+1)\big)}{\deg^+_{\omega(i)}(G)} = b(G') \cdot \frac{b(\omega(0))}{b(G')} \cdot \prod_{i=0}^{t-1} \frac{a \cdot \mu_{G'}\big(\omega(i),\omega(i+1)\big)}{\deg^+_{\omega(i)}(G')}.$$

Now, according to Eq. (2.6) by summing both sides for all walks in $\Omega_t(G)$ such that $\omega(t) = v$, we get the Eq. (3.1).

Edge Deletion

Fix edge $(u,u') \in E$ and let v be an arbitrary node which is not a successor of u, i.e., $v \notin S_u(G)$. Consider a graph obtained from (G,b) by removing edge (u,u'), i.e., $(G',b) = ((V,E - \{\!\{(u,u')\}\!\}),b)$. We have to prove that $PR^a_v(G,b) = PR^a_v(G',b)$. Note that the sum of weights is the same in both graphs. Hence, based on Theorem 1, to prove that $PR^a_v(G,b) = PR^a_v(G',b)$ it is enough to show that the probability of visit at node v at step $t \geq 0$ is also the same in both graphs, i.e.,

$$p^a_{G,b}(v,t) = p^a_{G',b}(v,t). \tag{3.2}$$

Consider an arbitrary walk $\omega \in \Omega_t(G)$ such that $\omega(t) = v$. Since $v \notin S_u(G)$, there is no path from u to v. Hence, we know that if a walk visits v at step t, then it could not have visited u before. Thus, each such walk belongs also to $\Omega_t(G')$ and each walk in $\Omega_t(G')$ that ends at v belongs to $\Omega_t(G)$. Moreover, for all $s,t \in V \setminus \{u\}$ we have that both $\mu_G(s,t) = \mu_{G'}(s,t)$ and $\deg^+_s(G) = \deg^+_s(G')$. Hence,

$$\prod_{i=0}^{t-1} \frac{a \cdot \mu_G\big(\omega(i),\omega(i+1)\big)}{\deg^+_{\omega(i)}(G)} = \prod_{i=0}^{t-1} \frac{a \cdot \mu_{G'}\big(\omega(i),\omega(i+1)\big)}{\deg^+_{\omega(i)}(G')}.$$

Thus, from Eq. (2.6) we get that Eq. (3.2) holds.

Edge Multiplication

Let $u, v \in V$ be two arbitrary nodes and $k \in \mathbb{N}$ a natural number. Also, let us denote the graph obtained from (G, b) by adding k copies of outgoing edges of u, i.e., $(G', b) = ((V, E \sqcup k \cdot \Gamma_u^+(G)), b)$. Observe that for every $t \in \mathbb{N}$, the sets of walks of length t in both graphs are identical, i.e., $\Omega_t(G) = \Omega_t(G')$. Moreover, for every nodes $s, t \in V$, we have that

$$\frac{\mu_G(s,t)}{\deg_s^+(G)} = \frac{\mu_{G'}(s,t)}{\deg_s^+(G')}$$

(if $s = u$, then both the numerator and the denominator are multiplied by k). Thus, from Eq. (2.6) we get that $p^a_{G,b}(v,t) = p^a_{G',b}(v,t)$, for every $t \in \mathbb{N}$. Hence, from Theorem 1 we get that $PR^a_v(G,b) = PR^a_v(G',b)$.

Edge Swap

Consider two edges $(u,u'), (w,w') \in E$ such that $PR^a_u(G,b) = PR^a_w(G,b)$ and also $\deg_u^+(G) = \deg_w^+(G)$ and the graph obtained from (G,b) by swapping the ends of these edges, i.e., $(G',b) = ((V,\ E - \{\!\{(u,u'),(w,w')\}\!\} \sqcup \{\!\{(u,w'),(w,u')\}\!\}),\ b)$. We have to prove that $PR^a(G,b) = PR^a(G',b)$. To this end, for every $v \in V$, let us define $x_v = PR^a_v(G,b)$ and prove that $(x_v)_{v \in V}$ satisfies the system of PageRank recursive equations (Eq. (2.5)) for graph (G',b), i.e., that

$$x_v = a \cdot \left(\sum_{s \in P_v^1(G')} \frac{\mu_{G'}(s,v)}{\deg_s^+(G)} \cdot x_s \right) + b(v), \tag{3.3}$$

holds for every $v \in V$. Since this system of equations has exactly one solution which is $PR^a(G',b)$, this will prove that $PR^a_v(G',b) = x_v = PR^a_v(G,b)$, for every $v \in V$.

Take an arbitrary node $v \in V$. From PageRank recursive equation (Eq. (2.5)) for graph (G,b) we have

$$x_v = a \cdot \left(\sum_{s \in P_v^1(G)} \frac{\mu_G(s,v)}{\deg_s^+(G)} \cdot x_s \right) + b(v). \tag{3.4}$$

From the definition of graph (G',b) we know that the out-degree of every node is the same in (G,b) as in (G',b). Moreover, if $v \notin \{u', w'\}$, then it has the same set of incoming edges in both graphs. Hence, replacing in Eq. (3.4) $\deg_s^+(G)$, $\mu_G(s,v)$, and $P_v^1(G)$ with $\deg_s^+(G')$, $\mu_{G'}(s,v)$, and $P_v^1(G')$, respectively, proves Eq. (3.3).

Assume $v = u'$ (for $v = w'$ the proof is analogous). Consider the sum on the right-hand side of the Eq. (3.4) for node $v = u'$. In graph (G',b) node u' has an edge (u,u') replaced with (w,u'). Thus,

$$\mu_{G'}(u,u') + \mu_{G'}(w,u') = (\mu_G(u,u') - 1) + (\mu_G(w,u') + 1) = \mu_G(u,u') + \mu_G(w,u').$$

Moreover, from the initial assumption we know that $x_u = x_w$ and $\deg_u^+(G) = \deg_w^+(G)$. Hence, we get that

$$\sum_{s \in P_{u'}^1(G) \setminus \{u,w\}} \frac{\mu_G(s,u')}{\deg_s^+(G)} \cdot x_s + \frac{\mu_G(u,u') + \mu_G(w,u')}{\deg_u^+(G)} \cdot x_u =$$

$$\sum_{s \in P_{u'}^1(G') \setminus \{u,w\}} \frac{\mu_{G'}(s,u')}{\deg_s^+(G')} \cdot x_s + \frac{\mu_{G'}(u,u') + \mu_{G'}(w,u')}{\deg_u^+(G')} \cdot x_u,$$

where we also used the fact that the rest of incoming edges of u' and the out-degree of all nodes are the same in (G,b) as in (G',b). This combined with Eq. (3.4) proves Eq. (3.3).

Node Redirect

Let u,w be out-twins. Consider a graph obtained from (G,b) by redirecting u into w, i.e., $(G',b') = R_{u\to w}(G,b)$. We have to prove that $PR^a_v(G',b') = PR^a_v(G,b)$, for every $v \in V\setminus\{u,w\}$, and that $PR^a_w(G',b') = PR^a_u(G,b)+PR^a_w(G,b)$. Note that the sum of weights is the same in both graphs. Hence, based on Theorem 1, it is enough to show that the probability of visit at node v at step $t \in \mathbb{N}$ is the same in both graphs for every $v \in V \setminus \{u,w\}$ and that this probability for node w in graph (G',b') is the sum of probabilities of visits for u and w in graph (G,b), i.e.,

$$p^a_{G',b'}(v,t) = \begin{cases} p^a_{G,b}(v,t), & \text{for every } v \in V \setminus \{u,w\}, \\ p^a_{G,b}(u,t) + p^a_{G,b}(w,t), & \text{for } v = w. \end{cases} \tag{3.5}$$

If $b(G) = 0$, then from Eq. (2.6) we get that $p^a_{G',b'}(w,t) = 0 = p^a_{G,b}(u,t) + p^a_{G,b}(w,t)$ and $p^a_{G',b'}(v,t) = 0 = p^a_{G,b}(v,t)$, for every $v \in V \setminus \{u,w\}$. Hence, let us assume that $b(G) > 0$.

Then, we will prove Eq. (3.5) by induction on t. From Eq. (2.6) for $t = 0$ we have that $p^a_{G',b'}(v,0) = b'(v)/b'(G) = b(v)/b(G) = p^a_{G,b}(v,0)$, for every $v \in V \setminus \{u,w\}$, and $p^a_{G',b'}(w,0) = b'(w)/b'(G) = (b(u)+b(w))/b(G) = p^a_{G,b}(u,0) + p^a_{G,b}(w,0)$.

Now, let us assume that Eq. (3.5) holds for some $t \geq 0$ and fix $v \in V$. Observe that for every walk $\omega \in \Omega_{t+1}(G')$ that ends at v, i.e., $\omega(t+1) = v$, the walk must have visited a direct predecessor of v, say s, at step $t-1$ and then move through edge (s,v). From this we get the equation $p^a_{G',b'}(v,t+1) = a\sum_{s\in P^1_v(G')} \frac{\mu_{G'}(s,v)}{\deg^+_s(G')} \cdot p^a_{G,b}(s,t)$, which we write as

$$p^a_{G',b'}(v,t+1) = a\frac{\mu_{G'}(w,v)}{\deg^+_w(G')} \cdot p^a_{G',b'}(w,t) + a \sum_{s\in P^1_v(G')\setminus\{w\}} \frac{\mu_{G'}(s,v)}{\deg^+_s(G')} \cdot p^a_{G',b'}(s,t). \tag{3.6}$$

Analogously, for graph (G,b) we get

$$p^a_{G,b}(v,t+1) = a\frac{\mu_G(u,v)}{\deg^+_u(G)} p^a_{G,b}(u,t) + a\frac{\mu_G(w,v)}{\deg^+_w(G)} p^a_{G,b}(w,t) + a \sum_{s\in P^1_v(G)\setminus\{u,w\}} \frac{\mu_G(s,v)}{\deg^+_s(G)} p^a_{G,b}(s,t).$$

Since u and w are out-twins we have $\mu_G(u,v) = \mu_G(w,v)$ and $\deg^+_u(G) = \deg^+_w(G)$. Moreover, from the inductive assumption we get $p^a_{G,b}(u,t) + p^a_{G,b}(w,t) = p^a_{G',b'}(w,t)$ as well as $p^a_{G,b}(s,t) = p^a_{G',b'}(s,t)$, for every $s \in V \setminus \{u,w\}$. Also, the redirection does not affect the out-degree of any node. Thus, equivalently

$$p^a_{G,b}(v,t+1) = a\frac{\mu_G(w,v)}{\deg^+_w(G')} \cdot p^a_{G',b'}(w,t) + a \sum_{s\in P^1_v(G)\setminus\{u,w\}} \frac{\mu_G(s,v)}{\deg^+_s(G')} p^a_{G',b'}(s,t). \tag{3.7}$$

If $v \in V \setminus \{u,w\}$, then for every node $s \in V \setminus \{u\}$, the edges going from s to v are not affected by the redirection, i.e., $\mu_G(s,v) = \mu_{G'}(s,v)$ and $P^1_v(G) \setminus \{u\} = P^1_v(G')$. Thus, from Eq. (3.6) and Eq. (3.7) we have that $p^a_{G',b'}(v,t+1) = p^a_{G,b}(v,t+1)$.

It remains to prove that $p^a_{G',b'}(w,t+1) = p^a_{G,b}(u,t+1) + p^a_{G,b}(w,t+1)$. To this end, let us add Eq. (3.7) for $v = u$ and $v = w$ sidewise. We get

$$p^a_{G,b}(u,t+1) + p^a_{G,b}(w,t+1) = a \sum_{s\in P^1_u(G)\cup P^1_w(G)\setminus\{u\}} \frac{\mu_G(s,u)+\mu_G(s,w)}{\deg^+_s(G')} p^a_{G',b'}(s,t). \tag{3.8}$$

Recall that the redirection transfers all incoming edges of u into w. Thus, for every $s \in V \setminus \{u\}$, we have that $\mu_{G'}(s,w) = \mu_G(s,u) + \mu_G(s,w)$ and $P^1_w(G') = P^1_u(G) \cup P^1_w(G) \setminus \{u\}$. Therefore, combining Eq. (3.6) and Eq. (3.8) we get that $p^a_{G',b'}(v,t+1) = p^a_{G,b}(u,t+1) + p^a_{G,b}(w,t+1)$. Hence, the induction hypothesis holds.

Baseline

Assume v is an isolated node in (G,b). Since node v has no incoming edges, from PageRank recursive equation (Eq. (2.5)) we get that $PR^a_v(G,b) = b(v)$. This proves that PageRank satisfies Baseline.

3.2.2 Part 2: Axioms Imply PageRank

In the main part of the proof, we show that if a centrality measure, F, satisfies first five axioms, i.e., Node Deletion, Edge Deletion, Edge Multiplication, Edge Swap, and Node Redirect, then it is equal to PageRank for some decay factor up to a scalar multiplication. Formally, we show that there exist two constants $c_F \in \mathbb{R}_{\geq 0}$ and $a_F \in [0,1)$, such that for every graph $G = (V,E)$, node weights b, and node $v \in V$, it holds that $F_v(G,b) = c_F \cdot PR^{a_F}_v(G,b)$. If F satisfies Baseline as well, then $c_F = 1$ and $F_v(G,b) = PR^{a_F}_v(G,b)$.

We start by introducing a class of *k-arrow graphs* that will be used in the proof. Each graph in this class consists of $k+1$ nodes—one source and k sinks—and k edges connecting the source to all the sinks. Moreover, only the source has a positive weight, i.e., the weight of each sink is zero.

Definition 1. *Graph $G = (V,E)$ with weights b is a k-arrow graph if $V = \{u, v_1, \ldots, v_k\}$, $E = \{\!\{(u,v_1),\ldots,(u,v_k)\}\!\}$ and $b(v_i) = 0$, for every $i \in \{1,\ldots,k\}$.*

In particular, every 1-arrow graph is of the form $((\{u,v\}, \{\!\{u,v\}\!\}), [x,0])$, for some nodes u,v and $x \in \mathbb{R}_{\geq 0}$.

Our proof has the following structure:

- First, we show that the centrality measure F satisfies two basic properties: *Locality*—the centrality of a node depends only on the part of the graph connected to it (Lemma 6) and *Source Node*—the centrality of a source is equal to its weight multiplied by some non-negative constant (Lemma 7). Moreover, this constant is the same for every source in every graph: it will be our constant c_F.

- Then, we show that the centrality of a sink in a 1-arrow graph is equal to the centrality of a source multiplied by some non-negative constant (Lemma 8). Moreover, this constant is the same for every 1-arrow graph and lies in the interval $[0,1)$ (Lemma 9): it will be our constant a_F.

- Having defined a_F and c_F, we turn our attention to proving that in every graph the centrality of any node is equal to PageRank with decay factor a_F multiplied by c_F. We do it by considering increasingly complex graphs. Specifically, we start with 1-arrow graphs (Lemma 10) and then k-arrow graphs (Lemma 11). Furthermore, we consider arbitrary graphs with no cycles (Lemma 12) and, ultimately, all possible graphs (Lemma 13).

Finally, if F additionally satisfies Baseline, then $c_F = 1$ and $F(G,b) = PR^{a_F}(G,b)$, for every graph (G,b) (Lemma 14).

First, let us focus on a basic property of *Locality*, which is implied by our first two axioms: Node Deletion and Edge Deletion. Locality states that if a graph consists of several disjoint parts (also called *connected components*), then the centrality of a node can be calculated by looking only at the part it is in [76].

Lemma 6. *(Locality) If a centrality measure, F, satisfies Node Deletion and Edge Deletion, then for every two disjoint graphs, $G = (V,E)$ and $G' = (V',E')$, node weights, b and b', and node $v \in V$, it holds that*

$$F_v((G + G', b + b')) = F_v(G, b).$$

Proof. Fix $v \in V$ and consider an arbitrary edge $(u,w) \in E'$. Since graphs G and G' are disjoint, there is no path from u to v. In particular, u is not a successor of v. Hence, by Edge Deletion, if we remove edge (u,w), the centrality of node v will remain unchanged. Using this argument for all the edges from E' we get that removing them does not affect the centrality of node v, i.e.,

$$F_v((G + G', b + b')) = F_v((V \cup V', E \sqcup E'), b + b') = F_v((V \cup V', E), b + b'). \tag{3.9}$$

Now, in graph $(V \cup V', E)$ all nodes from V' are isolated. Hence, by Node Deletion, removing these nodes does not affect the centrality of node v as well, i.e.,

$$F_v((V \cup V', E), b + b') = F_v((V,E), b) = F_v(G, b). \tag{3.10}$$

Combining Eq. (3.9) and Eq. (3.10) yields the thesis. □

In the second lemma, we prove that if a centrality measure satisfies Node Deletion, Edge Deletion, and Node Redirect, then it also satisfies the property of *Source Node*: the centrality of a source, i.e., a node without incoming edges, is proportional to its weight. This property is similar to Baseline, but there are two differences: First, Baseline applies only to isolated nodes and Source Node applies to all sources. Second, Baseline implies that the centrality of a node is equal, not merely proportional, to its weight.

Lemma 7. *(Source Node) If a centrality measure, F, satisfies Node Deletion, Edge Deletion, and Node Redirect, then there exists a constant $c_F \in \mathbb{R}_{\geq 0}$ such that for every graph $G = (V,E)$, weights b, and every source $v \in V$, it holds that*

$$F_v(G, b) = c_F \cdot b(v).$$

Specifically, $c_F = F_w((\{w\},\emptyset),[1])$ for an arbitrary node w.

Proof. We begin by considering graphs with one node and zero edges, i.e., graphs of the following form: $((\{v\},\emptyset),[x])$ for some node v and $x \in \mathbb{R}_{\geq 0}$. We will later show the relation between such graphs and sources in arbitrary graphs.

Consider two graphs $((\{u\},\emptyset),[x])$ and $((\{v\},\emptyset),[y])$ for arbitrary $u \neq v$ and $x, y \in \mathbb{R}_{\geq 0}$. Let (G,b) be their sum, i.e., $(G,b) = ((\{u,v\},\emptyset),[x,y])$. Since both u and v are isolated in (G,b), from Node Deletion we know that their centralities are the same as in the original graphs. In particular,

$$F_u(G,b) + F_v(G,b) = F_u((\{u\},\emptyset),[x]) + F_v((\{v\},\emptyset),[y]). \tag{3.11}$$

Nodes u and v are out-twins in (G,b) (both have the same empty set of outgoing edges). Thus, by Node Redirect, redirecting node v into u increases the centrality of u by the centrality of v. Such a redirecting results in graph $((\{u\},\emptyset),[x+y])$, so we get

$$F_u((\{u\},\emptyset),[x+y]) = F_u(R_{v\to u}(G,b)) = F_u(G,b) + F_v(G,b). \tag{3.12}$$

Combining Eq. (3.11) and Eq. (3.12) we get

$$F_u((\{u\},\emptyset),[x+y]) = F_u((\{u\},\emptyset),[x]) + F_v((\{v\},\emptyset),[y]). \tag{3.13}$$

We make the following observations:

(a) $F_v((\{v\},\emptyset),[0]) = 0$, for every v (from Eq. (3.13) with $y = 0$);

(b) $F_v((\{v\},\emptyset),[y]) = F_u((\{u\},\emptyset),[y])$, for every $u \neq v$ and $y \in \mathbb{R}_{\geq 0}$ (from Eq. (3.13) with $x = 0$ and (a));

(c) $F_v((\{v\},\emptyset),[x+y]) = F_v((\{v\},\emptyset),[x]) + F_v((\{v\},\emptyset),[y])$, for every v and $x,y \in \mathbb{R}_{\geq 0}$ (from Eq. (3.13) and (b)).

Note that (b) implies that the centrality of v in the weighted graph $((\{v\},\emptyset),[x])$ depends solely on weight x. In other words, there exists a function $f : \mathbb{R}_{\geq 0} \to \mathbb{R}$ such that $F_v((\{v\},\emptyset),[x]) = f(x)$. Since centralities are non-negative, we know that f is also non-negative, i.e., $f(x) \geq 0$, for every $x \in \mathbb{R}_{\geq 0}$. On the other hand, from (c) we know that f is additive, i.e., $f(x+y) = f(x) + f(y)$, for every $x,y \in \mathbb{R}_{\geq 0}$. Non-negativity and additivity combined imply that f is linear [23], i.e., $f(x) = c_F \cdot x$, for some $c_F \in \mathbb{R}_{\geq 0}$ and for every $x \in \mathbb{R}_{\geq 0}$. In effect, we know that there exists $c_F \in \mathbb{R}_{\geq 0}$ such that for every node v we have

$$F_v((\{v\},\emptyset),[x]) = c_F \cdot x. \tag{3.14}$$

Now, let $G = (V,E)$ be an arbitrary graph with node weights b and v be a source in G. Since v has no incoming edges, we know that it is not its own successor. Hence, by Edge Deletion, removing its outgoing edges does not affect its centrality, i.e., $F_v(G,b) = F_v((V,E-\Gamma_v^+(G)),b)$. In the resulting graph, v is isolated. Thus, from Lemma 6 (Locality) we have that $F_v((V,E-\Gamma_v^+(G)),b) = F_v((\{v\},\emptyset),[b(v)])$. This combined with Eq. (3.14) yields the thesis.

Finally, by taking Eq. (3.14) for $v = w$ and $x = 1$ we get that $c_F = F_w((\{w\},\emptyset),[1])$, which concludes the proof. □

Consider an arbitrary graph (G,b) in which node v is a source. Since the set of incoming edges of v is empty, i.e., $\Gamma_v^-(G) = \emptyset$, from PageRank recursive equation (Eq. (2.5)) we know that $PR_v^{a_F}(G,b) = b(v)$, for every decay factor $a_F \in [0,1)$. This implies that—regardless of the decay factor a_F—centrality of v is equal to PageRank of v multiplied by c_F, i.e., $F_v(G,b) = c_F \cdot PR_v^{a_F}(G,b)$.

Now, let us focus on 1-arrow graphs. Recall that in a 1-arrow graph there is one sink and one source connected by an edge and the sink has zero weight. In the next lemma, we prove that the centrality of the sink is proportional to the weight of the source.

Lemma 8. *If a centrality measure, F, satisfies Node Deletion, Edge Deletion, Edge Multiplication, Edge Swap, and Node Redirect, then there exists a constant $d_F \in \mathbb{R}_{\geq 0}$ such that for every 1-arrow graph $(G,b) = ((\{u,v\},\{\!\{(u,v)\}\!\}),[x,0])$, it holds that*

$$F_v(G,b) = d_F \cdot x.$$

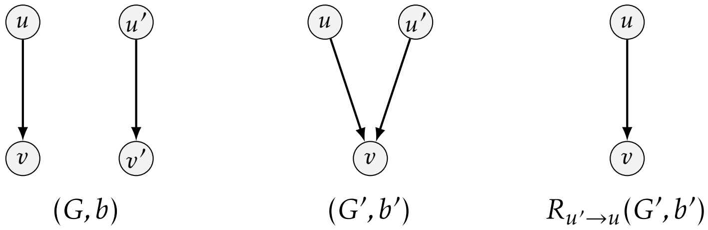


Figure 3.2: Graphs illustrating the proof of Lemma 8.

Proof. Our proof has a similar structure as the proof of Lemma 7, but instead of graphs with a single node we consider 1-arrow graphs.

First, let us consider two arbitrary 1-arrow graphs $((\{u,v\},\{\!\!\{(u,v)\}\!\!\}),[x,0])$ and $((\{u',v'\},\{\!\!\{(u',v')\}\!\!\}),[y,0])$ such that node u,v,u',v' are distinct. Let G be their sum, i.e., $(G,b) = ((\{u,v,u',v'\},\{\!\!\{(u,v),(u',v')\}\!\!\}),[x,0,y,0])$ (see Fig. 3.2 for an illustration). From Lemma 6 (Locality) we know that centralities of all nodes in (G,b) are the same as in the original graphs. In particular,

$$F_v(G,b) + F_{v'}(G,b) = F_v((\{u,v\},\{\!\!\{(u,v)\}\!\!\}),[x,0]) + F_{v'}((\{u',v'\},\{\!\!\{(u',v')\}\!\!\}),[y,0]). \quad (3.15)$$

Moreover, nodes v and v' are out-twins in (G,b) (both have the same empty set of outgoing edges). Hence, by Node Redirect, redirecting v' into v increases the centrality of v by the centrality of v'. Formally, let us denote the resulting graph by $(G',b') = ((\{u,v,u'\},\{\!\!\{(u,v),(u',v)\}\!\!\}),[x,0,y])$. Then,

$$F_v(G',b') = F_v(R_{v'\to v}(G,b)) = F_v(G,b) + F_{v'}(G,b). \quad (3.16)$$

Furthermore, nodes u and u' are out-twins in (G',b') (both have only one edge to node v). Hence, again by Node Redirect, redirecting u' into u does not affect the centrality of node v. Such a redirecting results in graph $((\{u,v\},\{\!\!\{(u,v)\}\!\!\}),[x+y,0])$ which is a 1-arrow graph (see Fig. 3.2), and we get

$$F_v((\{u,v\},\{\!\!\{(u,v)\}\!\!\}),[x+y,0]) = F_v(R_{u'\to u}(G',b')) = F_v(G',b'). \quad (3.17)$$

Combining Eqs. (3.15)–(3.17) we have

$$F_v((\{u,v\},\{\!\!\{(u,v)\}\!\!\}),[x+y,0]) = F_v((\{u,v\},\{\!\!\{(u,v)\}\!\!\}),[x,0]) + F_{v'}((\{u',v'\},\{\!\!\{(u',v')\}\!\!\}),[y,0]). \quad (3.18)$$

From Eq. (3.18) we make the following observations:

(a) $F_{v'}((\{u',v'\},\{\!\!\{(u',v')\}\!\!\}),[0,0]) = 0$ (from Eq. (3.18) with $y = 0$);

(b) $F_v((\{u,v\},\{\!\!\{(u,v)\}\!\!\}),[0,0]) = 0$ (from (a) and the fact that u',v' were chosen arbitrarily);

(c) $F_v((\{u,v\},\{\!\!\{(u,v)\}\!\!\}),[y,0]) = F_{v'}((\{u',v'\},\{\!\!\{(u',v')\}\!\!\}),[y,0])$, for every $y \in \mathbb{R}_{\geq 0}$ (from Eq. (3.18) with $x = 0$ and (b));

(d) $F_v((\{u,v\},\{\!\!\{(u,v)\}\!\!\}),[x+y,0]) = F_v((\{u,v\},\{\!\!\{(u,v)\}\!\!\}),[x,0]) + F_v((\{u,v\},\{\!\!\{(u,v)\}\!\!\}),[y,0])$, for every $x,y \in \mathbb{R}_{\geq 0}$ (from Eq. (3.18) and (c)).

From the fact that nodes were chosen arbitrarily, we know that (c) holds for every four pairwise distinct nodes u,v,u',v'. Assume they are not distinct, i.e., $u \neq v$ and

$u' \neq v'$, but $\{u,v\} \cap \{u',v'\} \neq \emptyset$. Then, let us take two new nodes u'',v'', different than u,v,u',v'. Using (c) two times we get that

$$\begin{aligned} F_v((\{u,v\},\{\!\{(u,v)\}\!\}),[y,0]) &= F_{v''}((\{u'',v''\},\{\!\{(u'',v'')\}\!\}),[y,0]) \\ &= F_{v'}((\{u',v'\},\{\!\{(u',v')\}\!\}),[y,0]), \end{aligned}$$

for every $y \in \mathbb{R}_{\geq 0}$ which implies (c) also for not pairwise distinct nodes. Furthermore, this means that the centrality of a sink in a 1-arrow graph depends only on the weight of the source, i.e., there exists a function $f : \mathbb{R}_{\geq 0} \to \mathbb{R}$ such that $F_v(\{u,v\},\{\!\{u,v\}\!\},[x,0]) = f(x)$. Since centralities are non-negative, we know that f is also non-negative, i.e., $f(x) \geq 0$, for every $x \in \mathbb{R}_{\geq 0}$. On the other hand, from (d) we know that f is additive, i.e., $f(x+y) = f(x) + f(y)$, for every $x,y \in \mathbb{R}_{\geq 0}$. Now, non-negativity combined with additivity implies that f is linear [23], i.e., $f(x) = d_F \cdot x$. This concludes the proof. □

In the next lemma, building upon Lemma 8, we show that there exists a constant $a_F \in [0,1)$ such that the centrality of a sink in 1-arrow graph equals $a_F \cdot c_F \cdot x$ where x is the weight of the source (see Lemma 7 for the definition of c_F).

Lemma 9. *If a centrality measure, F, satisfies Node Deletion, Edge Deletion, Edge Multiplication, Edge Swap, and Node Redirect, then there exists a constant* $a_F \in [0,1)$ *such that for every 1-arrow graph* $(G,b) = ((\{u,v\},\{\!\{(u,v)\}\!\}),[x,0])$ *it holds that* $F_v(G,b) = a_F \cdot c_F \cdot x$. *Specifically,* $a_F = F_w((\{w',w\},\{\!\{(w',w)\}\!\}),[1,0])/c_F$ *for arbitrary nodes* w,w' *if* $c_F > 0$ *and* $a_F = 0$, *otherwise.*

Proof. So far, we have proved that there exist constants $c_F, d_F \in \mathbb{R}_{\geq 0}$ such that for every 1-arrow graph the centrality of the source equals $c_F \cdot x$ (Lemma 7) and the centrality of the sink equals $d_F \cdot x$ (Lemma 8), where x is the weight of the source. Hence, to show that the centrality of the sink equals $a_F \cdot c_F \cdot x$ for some $a_F \in [0,1)$ it is enough to prove that if $c_F = 0$, then $d_F = 0$, and if $c_F \neq 0$, then $d_F < c_F$.

Assume $c_F = 0$. Consider graph $(G,b) = ((\{u,v,u',v'\},\{\!\{(u,v),(u',v')\}\!\}),[1,0,0,0])$ which is a sum of two 1-arrow graphs. Since $c_F = 0$, from Lemma 7 (Source Node) we know that $F_u(G,b) = 0 = F_{u'}(G,b)$ and from Lemma 8 and Lemma 6 (Locality), $F_v(G,b) = d_F$. Nodes u and u' each have one outgoing edge. Thus, from Edge Swap we know that swapping the ends of these edges does not affect the centralities in the graph. Formally, for graph $(G',b') = ((\{u,v,u',v'\},\{\!\{(u,v')(v,u')\}\!\}),[1,0,0,0])$ we know that $F_v(G',b') = F_v(G,b)$. However, from Lemma 8 and Lemma 6 (Locality) we get that $F_v(G',b') = 0$. Hence, $d_F = 0$.

Assume now that $c_F \neq 0$. To show that $d_F < c_F$, we will show that in one particular 1-arrow graph the centrality of the sink is strictly smaller than the centrality of the source which building upon the above general results will imply the thesis. Let u,v be two arbitrary distinct nodes and let y be the centrality of v in graph $((\{v\},\{\!\{(v,v)\}\!\}),[1])$ divided by c_F, i.e., $y = F_v((\{v\},\{\!\{(v,v)\}\!\}),[1])/c_F$. We will show that in 1-arrow graph $((\{u,v\},\{\!\{(u,v)\}\!\}),[y,0])$ the centrality of the sink is smaller than the centrality of the source, i.e.,

$$F_u((\{u,v\},\{\!\{(u,v)\}\!\}),[y,0]) > F_v((\{u,v\},\{\!\{(u,v)\}\!\}),[y,0]). \tag{3.19}$$

To this end, we begin by proving that the centralities of the source and the sink in graph $(G,b) = ((\{u,v\},\{\!\{(u,v)\}\!\}),[y,1])$ are equal, i.e.,

$$F_u((\{u,v\},\{\!\{(u,v)\}\!\}),[y,1]) = F_v((\{u,v\},\{\!\{(u,v)\}\!\}),[y,1]). \tag{3.20}$$

To prove Eq. (3.20), first let us consider graph

$$(G',b') = ((\{u,v,w\},\{\!\{(u,w),(v,v)\}\!\}),[y,1,0]),$$

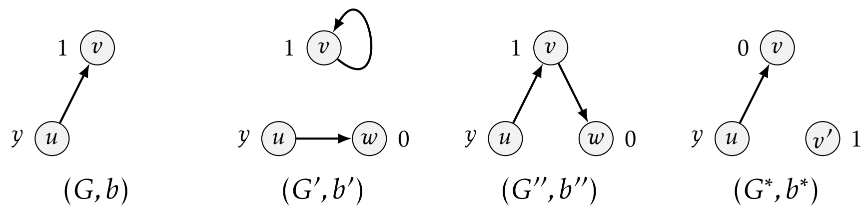


Figure 3.3: Graphs illustrating the proof of Lemma 9. The weights of nodes are shown.

which is the sum of two graphs: 1-arrow graph $((\{u,w\}, \{\!\{(u,w)\}\!\}), [y,0])$ and already mentioned one-node graph $((\{v\}, \{\!\{(v,v)\}\!\}), [1])$ (see Fig. 3.3 for an illustration). From Lemma 6 (Locality) and the definition of y we know that the centrality of node v equals $c_F \cdot y$. On the other hand, u is a source in (G', b'). Thus, from Lemma 7 (Source Node) we know that its centrality also equals $c_F \cdot y$. Hence, centralities of both nodes are equal, i.e.,

$$F_u(G', b') = F_v(G', b').$$

Since u and v have the same centralities in (G', b') and both have one outgoing edge, from Edge Swap we know that exchanging the ends of these edges does not affect the centralities in the graph. Let (G'', b'') be the resulting graph, i.e., $(G'', b'') = ((\{u,v,w\}, \{\!\{(u,v),(v,w)\}\!\}), [y,1,0])$ (see Fig. 3.3). We get that

$$F_u(G'', b'') = F_v(G'', b'').$$

Furthermore, since v is not its own successor and is not a successor of u, i.e., $v \notin S_u(G'', b'')$, $v \notin S_v(G'', b'')$, from Edge Deletion we can remove edge (v,w) from (G'', b'') without affecting centralities of nodes u, v. Moreover, after deleting edge (v,w) node w is isolated. Thus, from Node Deletion we know that w can also be deleted without affecting these centralities. Observe that in this way we obtain graph $(G,b) = ((\{u,v\}, \{\!\{(u,v)\}\!\}), [y,1])$ which proves Eq. (3.20).

Now, let us go back to proving Eq. (3.19). To this end, observe that graph $((\{u,v\}, \{\!\{(u,v)\}\!\}), [y,0])$ is obtained from (G,b) by changing the weight of v to zero. Since u is a source also in this graph and its weight did not change, from Lemma 7 (Source Node) we know that its centrality is the same as in (G,b), i.e.,

$$F_u((\{u,v\}, \{\!\{(u,v)\}\!\}), [y,0]) = F_u((\{u,v\}, \{\!\{(u,v)\}\!\}), [y,1]). \tag{3.21}$$

Now, let us turn our attention to the centrality of node v. Let us consider graph $(G^*, b^*) = ((\{u,v,v'\}, \{\!\{(u,v)\}\!\}, [y,0,1])$ (see Fig. 3.3). Since v' is a source in this graph, we know that $F_{v'}(G^*, b^*) = c_F$. Moreover, both v and v' do not have any outgoing edges in (G^*, b^*). Thus, by Node Redirect, redirecting v' into v increases the centrality of v by the centrality of v', i.e., by c_F. Such redirecting results in graph $((\{u,v\}, \{\!\{(u,v)\}\!\}), [y,1])$. Hence, we get

$$F_v((\{u,v\}, \{\!\{(u,v)\}\!\}), [y,0]) = F_v((\{u,v\}, \{\!\{(u,v)\}\!\}), [y,1]) - c_F. \tag{3.22}$$

Since we assumed that $c_F > 0$, Eq. (3.22) combined with Eq. (3.20) and Eq. (3.21) implies Eq. (3.19).

Our proof implies that $F_w((\{w',w), \{\!\{(w',w)\}\!\}), [1,0]) = a_F \cdot c_F$ for arbitrary nodes w, w'. Hence, if $c_F > 0$, we get that $a_F = F_w((\{w',w), \{\!\{(w',w)\}\!\}), [1,0])/c_F$, and if $c_F = 0$, a_F can be defined arbitrarily: we will assume $a_F = 0$. □

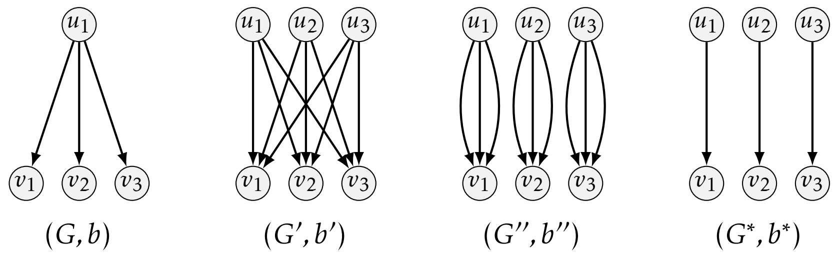


Figure 3.4: Example graphs illustrating the proof of Lemma 11 for $k = 3$.

A direct corollary from Lemma 7 (Source Node) and Lemma 9 is the fact that in 1-arrow graphs centralities are equal to PageRank with decay factor a_F multiplied by c_F (see these lemmas for the definitions of a_F and c_F).

Lemma 10. *If a centrality measure, F, satisfies Node Deletion, Edge Deletion, Edge Multiplication, Edge Swap, and Node Redirect, then for every 1-arrow graph* $(G,b) = ((\{u,v\},\{\!\!\{(u,v)\}\!\!\}),[x,0])$, *it holds* $F(G,b) = c_F \cdot PR^{a_F}(G,b)$. *Specifically,* $F_u(G,b) = c_F \cdot x$ *and* $F_v(G,b) = a_F \cdot c_F \cdot x$.

Proof. From Lemma 7 (Source Node) we know that $F_u(G,b) = c_F \cdot x$ and from Lemma 9 we know that $F_v(G,b) = a_F \cdot c_F \cdot x$. On the other hand, from PageRank recursive equation (Eq. (2.5)) we have that $PR^{a_F}_u(G,b) = x$ (node u has weight x and no incoming edges) and $PR^{a_F}_v(G,b) = a_F \cdot PR^{a_F}_u(G,b) = a_F \cdot c_F \cdot x$ (node v has a zero weight and only one incoming edge from node u with $\deg^+_u(G) = 1$). This concludes the proof. □

In the next lemma, we extend the result from Lemma 10 concerning 1-arrow graphs to k-arrow graphs. More in detail, we show that the centrality of every sink in k-arrow graph equals $a_F \cdot c_F \cdot x/k$, where x is the weight of the source. Hence, the centrality $a_F \cdot c_F \cdot x$ of the sink in a 1-arrow graph is split equally among all k sinks.

Lemma 11. *If centrality measure F satisfies Node Deletion, Edge Deletion, Edge Multiplication, Edge Swap, and Node Redirect axioms, then for every k-arrow graph* $(G,b) = ((\{u,v_1,\dots,v_k\},\{\!\!\{(u,v_1),\dots,(u,v_k)\}\!\!\}),[x,0,\dots,0])$, *it holds that* $F(G,b) = c_F \cdot PR^{a_F}(G,b)$. *Specifically,* $F_u(G,b) = c_F \cdot x$ *and* $F_{v_i}(G,b) = a_F \cdot c_F \cdot x/k$, *for every* $i \in \{1,\dots,k\}$.

Proof. First, let us consider an arbitrary k-arrow graph and denote it by $(G,b) = ((\{u_1,v_1,\dots,v_k\},\{\!\!\{(u_1,v_1),\dots,(u_1,v_k)\}\!\!\}),[x,0,\dots,0])$ (note that for the notational convenience we will denote the source node by u_1, not u). See Fig. 3.4 for an illustration. We need to prove that $F_{v_i}(G,b) = a_F \cdot c_F \cdot x/k$, for every $i \in \{1,\dots,k\}$. To this end, through a series of invariance operations, we will show that splitting the source of a k-arrow graph into k separate sources, each with 1 edge and $1/k$ of the original weight does not affect the centralities of sinks. In so doing, we obtain k separate 1-arrow graphs and Lemma 10 will imply the thesis.

First, take $k-1$ distinct nodes, $u_2,\dots,u_k$, that do not appear in (G,b) and consider a graph obtained from (G,b) by splitting u_1 into k nodes, $u_1,\dots,u_k$, each with the same edges as u_1 in (G,b) and $1/k$ of the original weight (see Fig. 3.4), i.e., let

$$(G',b') = \left(\left(\{u_1,\dots,u_k,v_1,\dots,v_k\}, \bigsqcup_{i=1}^{k} \{\!\!\{(u_i,v_1),\dots,(u_i,v_k)\}\!\!\}\right), [x/k,\dots,x/k,0,\dots,0]\right).$$

More formally, we have $b'(u_i) = x/k$ and $b'(v_i) = 0$, for every $i \in \{1,\dots,k\}$. Graph (G',b') contains k identical sources $(u_1,\dots,u_k)$ and k identical sinks $(v_1,\dots,v_k)$. All sources are out-twins, so, by Node Redirect, redirecting one of the nodes $u_2, \dots, u_k$ into u_1 does not affect the centralities of the sinks. This operation does not change edges of u_1 and the remaining sources, hence they are still out-twins. By performing all such redirections one by one, i.e., redirecting u_2 into u_1, u_3 into u_1, and so on, we will eventually obtain the original graph (G, b). Hence, we have $F_{v_i}(G', b') = F_{v_i}(G, b)$, for every $i \in \{1, \dots, k\}$.

Next, fix arbitrary $i, j \in \{1, \dots, k\}$ and consider replacing in (G', b') edges (u_i, v_j) and (u_j, v_i) with (u_i, v_i) and (u_j, v_j), i.e., an edge swap. Both sources u_i and u_j have exactly k outgoing edges and, by Lemma 7 (Source Node), they have the same centrality equal to x/k. Hence, by Edge Swap, this operation does not affect centralities in the graph. Moreover, it does not affect the number of outgoing edges of any node. Hence, by sequentially replacing edges $(u_i, v_j), (u_j, v_i)$ with $(u_i, v_i), (u_j, v_j)$ for all (unordered) pairs $i, j \in \{1, \dots, k\}$ we obtain graph

$$(G'', b'') = ((\{u_1, \dots, u_k, v_1, \dots, v_k\}, k \cdot \{\!\{(u_1, v_1), \dots, (u_k, v_k)\}\!\}), [x/k, \dots, x/k, 0, \dots, 0]),$$

(see Fig. 3.4) and we know that centralities of sinks did not change, i.e., we get that $F_{v_i}(G'', b'') = F_{v_i}(G', b')$, for every $i \in \{1, \dots, k\}$.

In graph (G'', b'') each source u_i has k edges, all to the same node v_i. From Edge Multiplication we know that replacing these k edges with only one edge does not affect centralities in the graph. Hence, for a graph

$$(G^*, b^*) = ((\{u_1, \dots, u_k, v_1, \dots, v_k\}, \{\!\{(u_1, v_1), \dots, (u_k, v_k)\}\!\}), [x/k, \dots, x/k, 0, \dots, 0]),$$

we get $F_{v_i}(G^*, b^*) = F_{v_i}(G'', b'')$, for every $i \in \{1, \dots, k\}$.

Finally, observe that graph G^* is a sum of k separate 1-arrow graphs. Hence, from Lemma 6 (Locality) and Lemma 10 we get that

$$F_{v_i}(G^*, b^*) = F_{v_i}((\{u_i, v_i\}, \{\!\{(u_i, v_i)\}\!\}), [x/k, 0]) = a_F \cdot c_F \cdot x/k,$$

for every $i \in \{1, \dots, k\}$. As a result, we showed that the centrality of every sink v_i in (G^*, b^*) is the same as in (G'', b''), (G', b') and eventually in (G, b), so we have $F_{v_i}(G, b) = a_F \cdot c_F \cdot x/k$.

Now, equality $F_{u_1}(G, b) = c_F \cdot x$ comes directly from Lemma 7 (Source Node). Moreover, from PageRank recursive equation (Eq. (2.5)) we have $PR_{u_1}(G, b) = x$ (u_1 is a node with no incoming edges and weight x) and

$$PR_{v_i}(G, b) = a_F \cdot PR_{u_1}(G, b)/k = a_F \cdot x/k, \quad \text{for every } i \in \{1, \dots, k\}$$

(v_i has a zero weight and one incoming edge from node u which has $\deg^+_{u_1}(G) = k$). This shows that $F(G, b) = c_F \cdot PR(G, b)$ and concludes the proof. □

Now, let us turn our attention to more complex graphs. In the following lemma we consider an arbitrary graph with no cycles and prove that the centralities are equal to PageRank with decay factor a_F multiplied by c_F (recall that these constants where defined in Lemmas 7 and 9).

Lemma 12. *If a centrality measure, F, satisfies Node Deletion, Edge Deletion, Edge Multiplication, Edge Swap, and Node Redirect, then for every graph with no cycles (G, b), it holds that $F(G, b) = c_F \cdot PR^{a_F}(G, b)$.*

Proof. We will use induction on the number of predecessors of a node in a graph. If $P_v(G) = \emptyset$, then node v is a source in G. Hence, from Lemma 7 (Source Node) we have $F_v(G, b) = c_F \cdot b(v) = c_F \cdot PR^{a_F}_v(G, b)$.

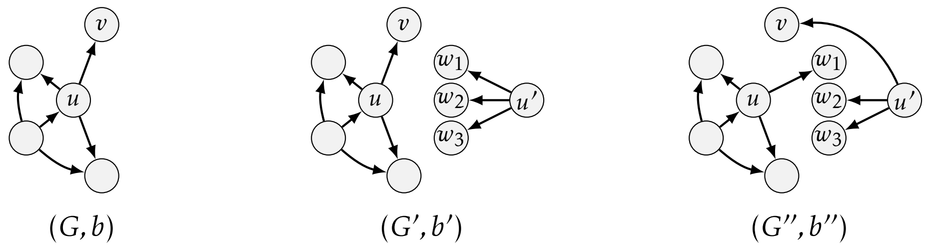


Figure 3.5: Example graphs illustrating the first part of the proof of Lemma 12 for $k = 3$.

Take a graph, (G, b), and a node, v, with non-empty set of predecessors, i.e., $P_v(G) \neq \emptyset$. Let u be an arbitrary predecessor of v. Clearly, $P_u(G) \subseteq P_v(G)$, but since graph has no cycles also $u \notin P_u(G)$. This implies that u has less predecessors than v, so from the inductive assumption, we get that $F_u(G, b) = c_F \cdot PR_u^{a_F}(G, b)$, for every $u \in P_v(G)$. Thus, to show that $F_v(G, b) = c_F \cdot PR_v^{a_F}(G, b)$, based on PageRank recursive equation (2.5) it is enough to prove that

$$F_v(G, b) = a_F \cdot \left(\sum_{(u,v) \in \Gamma_v^-(G)} \frac{F_u(G, b)}{\deg_u^+(G)} \right) + c_F \cdot b(v). \tag{3.23}$$

Note that from Edge Deletion and the fact that v is not a predecessor of itself nor its direct predecessors, we know that the outgoing edges of node v do not affect neither $F_v(G, b)$, $\Gamma_v^-(G)$, $F_u(G, b)$, nor $\deg_u^+(G)$, for every $(u, v) \in \Gamma_v^-(G)$. Hence, they do not affect Eq. (3.23). Consequently, in what follows, we will assume that v has no outgoing edges, i.e., it is a sink.

First, let us assume that v has a zero weight, i.e., $b(v) = 0$, and only one incoming edge, (u, v) (see Fig. 3.5 for an illustration). Let us denote the number of outgoing edges of u by k and its PageRank by x, i.e., $k = \deg_u^+(G)$ and $x = PR_u^{a_F}(G, b)$. From the inductive assumption we know that $F_u(G, b) = c_F \cdot x$. Hence, to prove Eq. (3.23), we need to show that $F_v(G, b) = a_F \cdot c_F \cdot x/k$. This, combined with Lemma 11 is equivalent to proving that the centrality of v is equal to the centrality of a sink in a k-arrow graph in which the source has weight x. To prove this, consider adding such a graph to (G, b), i.e., let

$$(G', b') = ((V', E'), b') = (G + (\{u', w_1, \dots, w_k\}, \{\!\{(u', w_1), \dots, (u', w_k)\}\!\}), b + x \cdot \mathbb{1}_{u'}).$$

From Lemma 6 (Locality), we know that the centrality of v did not change, i.e., $F_v(G, b) = F_v(G', b')$. Let us turn our attention to node u and the source of a k-arrow graph, u'. For u, from Lemma 6 (Locality) we know that $F_u(G', b') = F_u(G, b) = c_F \cdot x$. For u', from Lemma 7 (Source Node) we have that $F_{u'}(G', b') = c_F \cdot x$. Hence, u and u' have equal centralities and equal numbers of outgoing edges. In effect, from Edge Swap we know that we can replace edges (u, v) and (u', w_1) with edges (u, w_1) and (u', v) and centralities in the graph will not change. Such a swap results in a graph $(G'', b'') = ((V', E' - \{\!\{(u, v), (u', w_1)\}\!\} \sqcup \{\!\{(u, w_1), (u', v)\}\!\}), b')$ in which v is a sink in a part of the graph which is a k-arrow graph (see Fig. 3.5). In this k-arrow graph the source has weight x, hence node v has the centrality $a_F \cdot c_F \cdot x/k$. Formally, we proved that

$$F_v(G, b) = F_v(G', b') = F_v(G'', b'') = a_F \cdot c_F \cdot x/k = a_F \cdot \frac{F_u(G, b)}{\deg_u^+(G)}, \tag{3.24}$$

where the consecutive equalities come from Lemma 6 (Locality), Edge Swap, Lemma 11 combined with Locality, and the definition of constants x and k.

Figure 3.6: Example graphs illustrating the second part of the proof of Lemma 12 for $m = 3$.

Now, assume that v has m ($m \geq 1$) incoming edges and possibly non-zero weight. In such a case let us split node v into $m+1$ separate nodes, one with the original weight of v and no incoming edges and m nodes, each with zero weight and one incoming edge (see Fig. 3.6 for an illustration). Formally, assume $\Gamma_v^-(G) = \{\!\{(u_1, v), \ldots, (u_m, v)\}\!\}$ (note that u_i may not be pairwise different) and consider adding nodes $w_1, \ldots, w_m$ to the graph and replacing edges $\Gamma_v^-(G)$ with $(u_1, w_1), \ldots, (u_m, w_m)$. Let us denote the resulting graph by

$$(G', b') = ((V \cup \{w_1, \ldots, w_m\}, E - \Gamma_v^-(G) \sqcup \{\!\{(u_1, w_1), \ldots, (u_m, w_m)\}\!\}), b + \mathbf{0}).$$

Clearly, we have $(G, b) = R_{w_m \to v}\left(\ldots\left(R_{w_1 \to v}(G', b')\right)\right)$. Moreover, node v and nodes $w_1, \ldots, w_m$ are out-twins in (G', b') (they are all sinks). Hence, from Node Redirect we get that the centrality of v is the sum of centralities of v and $w_1, \ldots, w_m$ in (G', b'), i.e.,

$$F_v(G, b) = F_v(G', b') + F_{w_1}(G', b') + \cdots + F_{w_m}(G', b'). \tag{3.25}$$

Furthermore, from Node Redirect we know that centralities of nodes other than v did not change, i.e., $F_u(G, b) = F_u(G', b')$. Since the out-degrees of these nodes did not change either, from our analysis of nodes with a single edge and Eq. (3.24), in particular, we get that

$$F_{w_i}(G', b') = a_F \cdot \frac{F_{u_i}(G', b')}{\deg^+_{u_i}(G')} = a_F \cdot \frac{F_{u_i}(G, b)}{\deg^+_{u_i}(G)}. \tag{3.26}$$

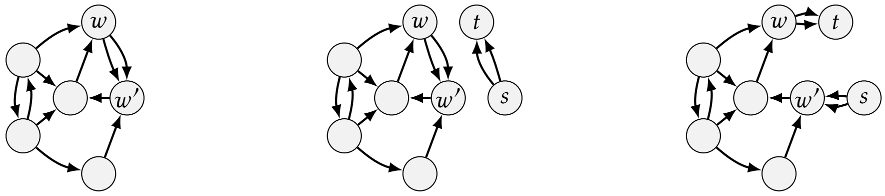
Finally, Lemma 7 (Source Node) implies that $F_v(G', b') = c_F \cdot b(v)$ (recall that v has no incoming edges in (G', b')). This combined with Eq. (3.25) and Eq. (3.26) concludes the proof. □

We are now ready to prove that in every graph centrality measure satisfying the axioms is equal to PageRank with the decay factor a_F multiplied by c_F.

Lemma 13. *If a centrality measure, F, satisfies Node Deletion, Edge Deletion, Edge Multiplication, Edge Swap, and Node Redirect, then for every graph (G, b), it holds that $F(G, b) = c_F \cdot PR^{a_F}(G, b)$.*

Proof. Take an arbitrary graph $G = (V, E)$ and weights b. We will prove the thesis by induction on the number of cycles in G. If there are no cycles in graph G, then the thesis follows from Lemma 12.

Assume otherwise. Fix node w that belongs to at least one cycle and let x_w be its PageRank in (G, b), i.e., $x_w = PR^{a_F}_w(G, b)$. Consider graph (G', b') obtained from (G, b) by adding two-node graph consisted of node s with weight x_w , node t with



(G, b) (G', b') (G'', b')

Figure 3.7: Example graphs illustrating the proof of Lemma 13.

weight 0, and edges from s to t in the number equal to $\deg^+_w(G)$ (see Fig. 3.7 for an illustration). Formally,

$$(G', b') = (G + (\{s, t\}, \deg^+_w(G) \cdot \{\!\{(s, t)\}\!\}), b + x_w \cdot \mathbb{1}_s).$$

We know that PageRank satisfies our axioms (Section 3.2.1), hence it also satisfies Lemma 6 (Locality). Thus,

$$PR^{a_F}_w(G', b') = x_w \quad \text{and} \quad PR^{a_F}_s(G', b') = x_w, \tag{3.27}$$

where the first equation holds from Locality and the second one from PageRank recursive equation (Eq. (2.5)). Moreover, nodes w and s have the same number of outgoing edges in graph (G', b') equal to $\deg^+_w(G)$. Therefore, if we swap the ends of all of their outgoing edges, then, by Edge Swap, this operation will not affect PageRank of any node. Let us define graph (G'', b') as a result of such an operation (see Fig. 3.7 for an illustration). Formally, let $G'' = (V \cup \{s, t\}, E - \Gamma^+_u(G) \sqcup \{\!\{(w, t), (s, w') : (w, w') \in \Gamma^+_u(G)\}\!\})$. From the fact that PageRank satisfies Edge Swap we get that

$$PR^{a_F}_v(G'', b') = PR^{a_F}_v(G', b'), \quad \text{for every } v \in V. \tag{3.28}$$

Observe that in graph (G'', b') all of the outgoing edges of w go to node t. Hence, graph (G'', b') has less cycles than graph (G, b) (every cycle in the former graph is also a cycle in the later one, but the former graph does not contain cycles with w). Hence, from the inductive assumption we know that

$$F_v(G'', b') = c_F \cdot PR^{a_F}_v(G'', b'), \quad \text{for every } v \in V. \tag{3.29}$$

Thus, combining Eqs. (3.27)–(3.29) we obtain that

$$F_w(G'', b') = c_F \cdot PR^{a_F}_w(G'', b') = c_F \cdot x_w = c_F \cdot PR^{a_F}_s(G'', b') = F_s(G'', b').$$

Thus, nodes w and s have equal centralities and equal number of outgoing edges in graph (G'', b'). Therefore, again from Edge Swap, this time for centrality F, we get that

$$F_v(G', b') = F_v(G'', b'), \quad \text{for every } v \in V.$$

Combining this with Eq. (3.28) and Eq. (3.29) yields $F_v(G', b') = c_F \cdot PR_v(G', b')$. Hence, the thesis follows from Lemma 6 (Locality). □

So far, we have proved that if F satisfies Node Deletion, Edge Deletion, Edge Multiplication, Edge Swap, and Node Redirect, then $F(G, b) = c_F \cdot PR^{a_F}(G, b)$, for every graph (G, b). In the last lemma of this section, we show that if F also satisfies Baseline, then $c_F = 1$; hence, $F(G, b) = PR^{a_F}(G, b)$.

Lemma 14. *If centrality measure F satisfies Node Deletion, Edge Deletion, Edge Multiplication, Edge Swap, Node Redirect, and Baseline, then for every graph (G, b), it holds that $F(G, b) = PR^{a_F}(G, b)$.*

Proof. If centrality measure F satisfies Node Deletion, Edge Deletion, Edge Multiplication, Edge Swap, and Node Redirect, then from Lemma 13 we know that $F_v(G, b) = c_F \cdot PR_v^{a_F}(G, b)$ for c_F and a_F defined as follows: $c_F = F_w((\{w\}, \emptyset), [1])$ and $a_F = F_w((\{w', w\}, \{\!\{(w', w)\}\!\}), [1, 0])/c_F$ if $c_F > 0$ and $a_F = 0$, otherwise, for arbitrary nodes w, w'. Now, from Baseline we have that $F_w((\{w\}, \emptyset), [1]) = 1$, which implies $c_F = 1$ and concludes the proof. □

3.2.3 Independence of Axioms

In this section, we show that all six axioms used in our characterization of PageRank are necessary, i.e., if we remove any one of them, then the remaining axioms will be satisfied also by some centrality measure other than PageRank. In other words, we will show that axioms are independent and no axiom is implied by the others. To this end, in the following theorem for each axiom we show that there exists a centrality, other than PageRank, that satisfies all other axioms.

Theorem 15. *(Independence of Axioms) From six axioms: Node Deletion, Edge Deletion, Edge Multiplication, Edge Swap, Node Redirect, and Baseline, none is implied by a combination of five others. Specifically:*

- *A centrality measure F defined for every graph (G, b) and node v as*

$$F_v(G, b) = PR_v^{a(G,b)}(G, b), \qquad \text{where } a(G, b) = 1/(2 + b(G))$$

  satisfies Edge Deletion, Edge Multiplication, Edge Swap, Node Redirect, and Baseline, but does not satisfy Node Deletion.

- *A centrality measure F defined for every graph (G, b) and node v and an arbitrary $a \in (0, 1)$ as*

$$F_v^a(G, b) = \begin{cases} 2 \cdot PR_v^a(G, b) - b(v), & \text{if } v \text{ is a sink,} \\ PR_v^a(G, b), & \text{otherwise} \end{cases}$$

  satisfies Node Deletion, Edge Multiplication, Edge Swap, Node Redirect, and Baseline, but does not satisfy Edge Deletion.

- *A centrality measure F defined for every graph (G, b) and node v and an arbitrary $a \in (0, 1)$ as*

$$F_v^a(G, b) = a \cdot \left(\sum_{u \in P_v^1(G)} \frac{\mu_G(u, v)}{\deg_u^+(G) + 1} \cdot F_u^a(G, b) \right) + b(v)$$

  satisfies Node Deletion, Edge Deletion, Edge Swap, Node Redirect, and Baseline, but does not satisfy Edge Multiplication.

- *A centrality measure F defined for every graph (G, b) and node v as*

$$F_v(G, b) = \sum_{u \in P_v^1(G)} \frac{\mu_G(u, v)}{\deg_u^+(G)} \cdot b(u) + b(v)$$

  satisfies Node Deletion, Edge Deletion, Edge Multiplication, Node Redirect, and Baseline, but does not satisfy Edge Swap.

- *A centrality measure F defined for every graph (G, b) and node v as*

$$F_v(G,b) = \sum_{u \in P_v^1(G)} \frac{\mu_G(u,v)}{\deg_u^+(G)} + b(v)$$

satisfies Node Deletion, Edge Deletion, Edge Multiplication, Edge Swap, and Baseline, but does not satisfy Node Redirect.

- *A centrality measure F defined for every graph (G, b) and node v and an arbitrary $a \in (0,1)$ as*

$$F_v^a(G,b) = 2 \cdot PR_v^a(G,b)$$

satisfies Node Deletion, Edge Deletion, Edge Multiplication, Edge Swap, and Node Redirect, but does not satisfy Baseline.

We divide the proof of Theorem 15 into six lemmas, Lemmas 16–21, each characterizing a centrality different than PageRank that satisfies five out of six of our axioms.

Lemma 16. *A centrality measure F defined for every graph (G, b) and node v as*

$$F_v(G,b) = PR_v^{a(G,b)}(G,b), \qquad \text{where } a(G,b) = 1/(2 + b(G))$$

satisfies Edge Deletion, Edge Multiplication, Edge Swap, Node Redirect, and Baseline, but does not satisfy Node Deletion.

Proof. Observe that for any two graphs (G, b), (G', b'), and a node v, if $b(G) = b'(G')$ and $PR_v^a(G,b) = PR_v^a(G',b')$, for every $a \in [0,1)$, then $a(G,b) = a(G',b')$ and

$$F_v(G,b) = PR^{a(G,b)}(G,b) = PR_v^{a(G',b')}(G',b') = F_v(G',b'). \tag{3.30}$$

We consider each axiom separately. Fix an arbitrary graph $G = (V, E)$ and node weights b.

- For Edge Deletion, consider an arbitrary edge $(u, w) \in E$ and take graph $(G', b') = ((V, E - \{\!\{(u,w)\}\!\}), b)$. Clearly, $b(G) = b'(G')$. Since PageRank satisfies Edge Deletion, for every $v \in V \setminus S_u(G)$, we have $PR_v^a(G,b) = PR_v^a(G',b')$ for $a \in [0,1)$ and Eq. (3.30) implies $F_v(G,b) = F_v(G',b')$.

- For Edge Multiplication, consider node $u \in V$, constant $k \in \mathbb{N}$, and graph $(G', b') = ((V, E \sqcup k \cdot \Gamma_u^+(G)), b)$. Clearly, $b(G) = b'(G')$. Since PageRank satisfies Edge Multiplication, Eq. (3.30) implies $F_v(G,b) = F_v(G',b')$, for every $v \in V$.

- For Edge Swap, consider $(u, u'), (w, w') \in E$ such that $\deg_u^+(G) = \deg_w^+(G)$ and $F_u(G,b) = F_w(G,b)$. The latter implies that $PR_u^{a(G,b)}(G,b) = PR_w^{a(G,b)}(G,b)$ as well. Let $(G', b') = ((V,\ E - \{\!\{(u,u'),(w,w')\}\!\} \sqcup \{\!\{(u,w'),(w,u')\}\!\}),\ b)$. Clearly, $b(G) = b'(G')$. Since PageRank satisfies Edge Swap, Eq. (3.30) implies that $F_v(G,b) = F_v(G',b')$, for every $v \in V$.

- For Node Redirect, assume $u, w \in V$ are out-twins. Let $(G', b') = R_{u \to w}(G, b)$. Clearly, $b(G) = b'(G')$. Since PageRank satisfies Node Redirect, for every node $v \in V \setminus \{u, w\}$, we have $PR_v^a(G,b) = PR_v^a(G',b')$, for every $a \in [0,1)$, and Eq. (3.30) implies $F_v(G,b) = F_v(G',b')$. Analogously, we get that

$$F_u(G,b) + F_w(G,b) = PR_u^{a(G,b)}(G,b) + PR_w^{a(G,b)}(G,b) = PR_w^{a(G',b')}(G',b') = F_w(G',b').$$

- For Baseline, assume v is isolated. Since PageRank satisfies Baseline we get that $F_v(G,b) = b(v)$.

Finally, consider Node Deletion. Let $(G,b) = ((\{u,v,w\}, \{\!\{(u,v)\}\!\}), [1,0,1])$. We have $a(G,b) = 1/4$, so $F_v(G,b) = PR_v^{1/4}(G,b) = 1/4$. Note that w is isolated in (G,b). Now, if we delete node w we will obtain $(G',b') = ((\{u,v\}, \{\!\{(u,v)\}\!\}), [1,0])$. Here, we have $a(G',b') = 1/3$, so $F_v(G,b) = PR_v^{1/3}(G,b) = 1/3$. Thus, Node Deletion is not satisfied. □

Lemma 17. *A centrality measure F defined for every graph (G,b) and node v and an arbitrary $a \in (0,1)$ as*

$$F_v^a(G,b) = \begin{cases} 2 \cdot PR_v^a(G,b) - b(v), & \text{if } v \text{ is a sink,} \\ PR_v^a(G,b), & \text{otherwise,} \end{cases}$$

satisfies Node Deletion, Edge Multiplication, Edge Swap, Node Redirect, and Baseline, but does not satisfy Edge Deletion.

Proof. Fix $a \in (0,1)$. Observe that for any two graphs $(G,b),(G',b')$, and node v, if it holds that node v is a sink in (G,b) if and only if v is also a sink in (G',b') and $PR_v^a(G,b) = PR_v^a(G',b')$, then

$$\begin{cases} F_v^a(G,b) = 2 \cdot PR_v^a(G,b) - b(v) = 2 \cdot PR_v^a(G',b') - b'(v) = F_v^a(G',b'), & \text{if } v \text{ is a sink,} \\ F_v^a(G,b) = PR_v^a(G,b) = PR_v^a(G',b') = F_v^a(G',b'), & \text{otherwise.} \end{cases} \tag{3.31}$$

We consider each axiom separately. Fix an arbitrary graph $G = (V,E)$ and node weights b.

- For Node Deletion, assume u is an isolated node. Let $(G',b') = ((V \setminus \{u\}, E), b)$. Fix $v \in V \setminus \{u\}$. Note that v is a sink in (G,b) if and only if it is a sink in (G',b'). Since PageRank satisfies Node Deletion, from Eq. (3.31) we have that $F_v^a(G,b) = F_v^a(G',b')$.
- For Edge Multiplication, consider node $u \in V$, constant $k \in \mathbb{N}$, and graph $(G',b') = ((V, E \sqcup k \cdot \Gamma_u^+(G)), b)$. Fix $v \in V$. Note that v is a sink in (G,b) if and only if it is a sink in (G',b'). Since PageRank satisfies Edge Multiplication, Eq. (3.31) implies $F_v^a(G,b) = F_v^a(G',b')$.
- For Edge Swap, consider edges $(u,u'),(w,w') \in E$ such that $F_u(G,b) = F_w(G,b)$ and $\deg_u^+(G) = \deg_w^+(G)$. Since u and w are not sinks, by the definition of F, this means that also $PR_u^a(G,b) = PR_w^a(G,b)$. Next, let us consider graph $(G',b') = ((V, E - \{\!\{(u,u'),(w,w')\}\!\} \sqcup \{\!\{(u,w'),(w,u')\}\!\}), b)$. Fix $v \in V$. Note that v is a sink in (G,b) if and only if it is a sink in (G',b'). Since PageRank satisfies Edge Swap, Eq. (3.31) implies $F_v^a(G,b) = F_v^a(G',b')$.
- For Node Redirect, assume $u,w \in V$ are out-twins. Let $(G',b') = R_{u \to w}(G,b)$. Fix $v \in V \setminus \{u,w\}$. Note that v is a sink in (G,b) if and only if it is a sink in (G',b'). Since PageRank satisfies Node Redirect, Eq. (3.31) implies that $F_v(G,b) = F_v(G',b')$. Analogously, we get that

  $$F_u^a(G,b) + F_w^a(G,b) = 2 \cdot PR_u^a(G,b) - b(u) + \cdot PR_w^a(G,b) - b(w) = \\ 2 \cdot PR_w^a(G',b') - b'(w) = F_w^a(G',b')$$

  if u and w are sinks and if not, then

  $$F_u^a(G,b) + F_w^a(G,b) = PR_u^a(G,b) + PR_w^a(G,b) = PR_w^a(G',b') = F_w^a(G',b').$$

- For Baseline, assume v is isolated. Since v is as sink and PageRank satisfies Baseline, we get that $F^a_v(G,b) = 2 \cdot PR^a_v(G,b) - b(v) = 2 \cdot b(v) - b(v) = b(v)$.

Finally, consider Edge Deletion. Let $(G,b) = ((\{u,v,w\}, \{\!\{(u,v),(v,w)\}\!\}), [1,0,0])$. In (G,b) node v is not a sink, hence $F^a_v(G,b) = PR^a_v(G,b) = a$. Note that v is not its own successor in G. Now, if we delete edge (v,w), then we obtain a graph $(G',b') = ((\{u,v,w\}, \{\!\{(u,v)\}\!\}), [1,0,0])$. Since v is a sink, $F_v(G',b') = 2 \cdot PR^a_v(G,b) = 2 \cdot a$. Thus, Edge Deletion is not satisfied. □

Lemma 18. *A centrality measure F defined for every graph (G,b) and node v and an arbitrary $a \in (0,1)$ as*

$$F^a_v(G,b) = a \cdot \left(\sum_{u \in P^1_v(G)} \frac{F^a_u(G,b)}{\deg^+_u(G)+1} \right) + b(v) \tag{3.32}$$

satisfies Node Deletion, Edge Deletion, Edge Swap, Node Redirect, and Baseline, but does not satisfy Edge Multiplication.

Proof. Fix $a \in (0,1)$. Let us start by showing that the centrality F^a is well defined, i.e., that the Eq. (3.32) has a unique solution. To this end, consider a graph operation f that for every graph $G = (V,E)$ with node weights b adds a new node t with a zero weight and adds one edge from every node in V to t, i.e.,

$$f(G,b) = ((V \cup \{t\}, E \sqcup \{\!\{(v,t) : v \in V\}\!\}), b + \mathbf{0}).$$

Observe that for every weighted graph (G,b) and node v, the PageRank recursive equation (Eq. (2.5)) for graph $f(G,b)$ and node v is the same as Eq. (3.32). Hence, we get that centrality F^a in graph (G,b) is equal to PageRank in graph $f(G,b)$, i.e., $F^a(G,b) = PR^a(f(G,b))$.

Consider a second graph operation h. Observe that if for some graph (G,b) and node v we have $f(h(G,b)) = h(f(G,b))$ and $PR^a_v(f(G,b)) = PR^a_v(h(f(G,b)))$, then

$$F^a_v(G,b) = PR^a_v(f(G,b)) = PR^a_v(h(f(G,b))) = PR^a_v(f(h(G,b))) = F^a_v(h(G,b)). \tag{3.33}$$

Now, we are ready for the axiomatic analysis of centrality F^a. We will consider each axiom separately. Fix an arbitrary graph $G = (V,E)$ and node weights b.

- For Node Deletion, assume u is an isolated node. Let us define a graph operation h that for an arbitrary graph $(G',b') = ((V',E'),b')$ deletes node u and all its edges, i.e., let $h(G',b') = ((V' \setminus \{u\}, E' - \Gamma^+_u(G') - \Gamma^-_u(G')), b'')$ where $b''(v') = b'(v')$, for every $v' \in V'$. Fix $v \in V \setminus \{u\}$.
  - Clearly, operations f and h are commutative, so $f(h(G,b)) = h(f(G,b))$.
  - Consider PageRank of v in graph $f(G,b)$. Note that in $f(G,b)$ node u is no longer isolated, but has one outgoing edge to node t. Nevertheless, since PageRank satisfies Edge Deletion, we know that removing this edge does not affect PageRank of v. If edge (u,t) is removed, then u is isolated, so since PageRank satisfies Node Deletion, this node can also be removed without affecting PageRank of v. Hence, we get that $PR^a_v(f(G,b)) = PR^a_v(h(f(G,b)))$.

  As a result, by Eq. (3.33), $F^a_v(G,b) = F^a_v(h(G,b))$.

- For Edge Deletion, consider an arbitrary edge $(u,w) \in E$. Let us define a graph operation h that for an arbitrary graph $(G',b') = ((V',E'),b')$ deletes edge (u,w) from the graph, i.e., let $h(G',b') = ((V',E' - \{\!\{(u,w)\}\!\}),b')$. Also, fix $v \notin S_u(G)$.
  - Clearly, operations f and h are commutative, so $f(h(G,b)) = h(f(G,b))$.
  - Note that $v \notin S_u(G)$ implies $v \notin S_u(f(G))$. Hence, since PageRank satisfies Edge Deletion, we get that $PR^a_v(f(G,b)) = PR^a_v(h(f(G,b)))$.

  As a result, by Eq. (3.33), $F^a_v(G,b) = F^a_v(h(G,b))$.

- For Edge Swap, let us assume that $(u,u'),(w,w') \in E$ are two edges such that $F^a_u(G,b) = F^a_w(G,b)$ and $\deg^+_u(G) = \deg^+_w(G)$. Let us define a graph operation h such that for an arbitrary graph $(G',b') = ((V',E'),b')$ operation h deletes edges $(u,u'),(w,w')$ and adds edges $(u,w'),(w,u')$. Formally, we have that $h(G',b') = ((V',E' - \{\!\{(u,u'),(w,w')\}\!\} \sqcup \{\!\{(u,w'),(w,u')\}\!\}),b')$. Fix $v \in V$.
  - Clearly, operations f and h are commutative, so $f(h(G,b)) = h(f(G,b))$.
  - $PR^a_u(f(G,b)) = PR^a_w(f(G,b))$ from the definition of F^a and the fact that $F^a_u(G,b) = F^a_w(G,b)$. Also, $\deg^+_u(f(G)) = \deg^+_w(f(G)) = \deg^+_u(G) + 1$. PageRank satisfies Edge Swap, hence $PR^a_v(f(G,b)) = PR^a_v(h(f(G,b)))$.

  As a result, by Eq. (3.33), $F^a_v(G,b) = F^a_v(h(G,b))$.

- For Node Redirect, assume $u,w \in V$ are out-twins. Let us define an operation h as redirecting u into w, i.e., $h = R_{u\to w}$. Take an arbitrary $v \in V \setminus \{u,w\}$.
  - Clearly, operations f and h are commutative, so $f(h(G,b)) = h(f(G,b))$.
  - Note that the fact that u,w are out-twins in (G,b) implies that they are also out-twins in $f(G,b)$. Hence, since PageRank satisfies Node Redirect, we have $PR^a_v(f(G,b)) = PR^a_v(h(f(G,b)))$.

  As a result, by Eq. (3.33), $F^a_v(G,b) = F^a_v(h(G,b))$. Analogously, we get that

$$F^a_u(G,b) + F_w(G,b) = PR^a_u(f(G,b)) + PR^a_w(f(G,b)) = \\ PR^a_w(h(f(G,b))) = PR^a_w(f(h(G,b))) = F^a_w(h(G,b)).$$

- For Baseline, assume that node v is isolated. From PageRank recursive equation (Eq. (2.5)) we get that $F^a_v(G,b) = b(v)$.

Finally, consider Edge Multiplication. Let $(G,b) = ((\{u,v\},\{\!\{(u,v)\}\!\}),[1,0])$. We have that $F^a_u(G,b) = 1$ and $F^a_v(G,b) = a/2$. Now, if we add an additional copy of outgoing edges of node u we obtain $(G',b') = ((\{u,v\},\{\!\{(u,v),(u,v)\}\!\}),[1,0])$. Here, we have that $F^a_u(G,b) = 1$ and $F^a_v(G,b) = 2/3 \cdot a$. Thus, Edge Multiplication is not satisfied. □

Lemma 19. *A centrality measure F defined for every graph (G,b) and node v as*

$$F_v(G,b) = \sum_{u \in P^1_v(G)} \frac{\mu_G(u,v)}{\deg^+_u(G)} \cdot b(u) + b(v) \tag{3.34}$$

satisfies Node Deletion, Edge Deletion, Edge Multiplication, Node Redirect, and Baseline, but does not satisfy Edge Swap.

Proof. We will consider each axiom separately. Fix an arbitrary graph $G = (V, E)$ and node weights b.

- For Node Deletion, assume u is an isolated node. Since u is isolated, it is not a direct predecessor of any node and we get that $F_v(G, b) = F_v((V \setminus \{u\}, E), b)$, for every $v \in V \setminus \{u\}$.

- For Edge Deletion, consider an arbitrary edge, $(u, w) \in E$. For every $v \notin S_u(G)$, node u is not a direct predecessor and we get $F_v(G, b) = F_v((V, E - \{\!\{(u, w)\}\!\}), b)$.

- For Edge Multiplication, consider node $u \in V$ and constant $k \in \mathbb{N}$. Also, let $(G', b') = ((V, E \sqcup k \cdot \Gamma_u^+(G), b)$. Note that $\mu_G(u, v)/\deg_u^+(G) = \mu_{G'}(u, v)/\deg_u^+(G')$, for every $u, v \in V$. Also, for every $v \in V$, the weight and the set of direct predecessors in both graphs is the same, i.e.,$b(v) = b'(v)$ and $P_v^1(G) = P_v^1(G')$. Hence, from Eq. (3.34) we get that $F_v(G, b) = F_v(G', b')$, for every $v \in V$.

- For Node Redirect, assume $u, w \in V$ are out-twins. Let $(G', b') = R_{u \to w}(G, b)$. Note that $\deg_u^+(G) = \deg_w^+(G)$ and $\mu_G(u, v) = \mu_G(w, v)$, for every $v \in V$. Moreover, the out-degrees of all nodes in (G, b) are the same as in (G', b'). Hence, for every $v \in V \setminus \{u, w\}$, we have

$$F_v(G, b) = \sum_{s \in P_v^1(G) \setminus \{u, w\}} \frac{\mu_G(s, v)}{\deg_s^+(G)} \cdot b(s) + \frac{\mu_G(w, v)}{\deg_w^+(G)} \cdot (b(u) + b(w)) + b(v) = F_v(G', b').$$

  For node w, we have that $\mu_G(u, u) + \mu_G(u, w) = \mu_G(w, u) + \mu_G(w, w) = \mu_{G'}(w, w)$. So we get that

$$F_u(G, b) + F_w(G, b) =$$
$$\sum_{s \in V \setminus \{u, w\}} \frac{\mu_G(s, u) + \mu_G(s, w)}{\deg_s^+(G)} \cdot b(s) + \frac{\mu_G(u, u) + \mu_G(u, w)}{\deg_w^+(G)} \cdot (b(u) + b(w)) = F_w(G', b').$$

- For Baseline, assume v is isolated. Since v has no direct predecessors, we have $F_v(G, b) = b(v)$.

Finally, consider Edge Swap. Let $(G, b) = ((\{u, v, w\}, \{\!\{(u, v), (v, w)\}\!\}), [1, 0, 0])$. We have that $F_u(G, b) = 1$, $F_v(G, b) = 1$, and $F_w(G, b) = 0$. Observe that nodes u and v both have exactly one outgoing edge and equal centralities. Now, observe that if we replace edges (u, v) and (v, w) with edges (u, w) and (v, v), then we obtain graph $(G', b') = ((\{u, v, w\}, \{\!\{(u, w), (v, v)\}\!\}), [1, 0, 0])$. Here, we have that $F_w(G', b') = 1$. Thus, Edge Swap is not satisfied. □

Lemma 20. *A centrality measure F defined for every graph (G, b) and node v as*

$$F_v(G, b) = \sum_{u \in P_v^1(G)} \frac{\mu_G(u, v)}{\deg_u^+(G)} + b(v) \tag{3.35}$$

satisfies Node Deletion, Edge Deletion, Edge Multiplication, Edge Swap, and Baseline, but does not satisfy Node Redirect.

Proof. We will consider each axiom separately. Fix an arbitrary graph $G = (V, E)$ and node weights b.

- For Node Deletion, assume u is an isolated node. Since u is isolated, it is not a direct predecessor of any node and we get that $F_v(G,b) = F_v((V \setminus \{u\}, E), b)$, for every $v \in V \setminus \{u\}$.

- For Edge Deletion, consider an arbitrary edge, $(u,w) \in E$. For every $v \notin S_u(G)$, node u is not a direct predecessor, thus $F_v(G,b) = F_v((V, E - \{\!\{(u,w)\}\!\}), b)$.

- For Edge Multiplication, consider node $u \in V$, constant $k \in \mathbb{N}$, and graph $(G',b') = ((V, E \sqcup k \cdot \Gamma_u^+(G), b)$. Observe that for every pair of nodes $u,v \in V$, we have $\mu_{G'}(u,v)/\deg_u^+(G') = \mu_G(u,v)/\deg_u^+(G)$. Also, for every $v \in V$, the weight and the set of direct predecessors in both graphs is the same, i.e.,$b(v) = b'(v)$ and $P_v^1(G) = P_v^1(G')$. Hence, from Eq. (3.35) we get that $F_v(G,b) = F_v(G',b')$, for every $v \in V$.

- For Edge Swap, take $(u,u'),(w,w') \in E$ such that $F_u(G,b) = F_w(G,b)$ and also $\deg_u^+(G) = \deg_w^+(G)$. Let $(G',b') = ((V, E - \{\!\{(u,u'),(w,w')\}\!\} \sqcup \{\!\{(u,w'),(w,u')\}\!\}), b)$. Note that the out-degree and the weight of every node in (G,b) is the same as in (G',b'), i.e.,$\deg_v^+(G) = \deg_v^+(G')$ and $b(v) = b'(v)$, for every $v \in V$. Hence, for every node $v \in V \setminus \{u',w'\}$, we get that $F_v(G,b) = F_v(G',b')$. Observe that $\deg_w^+(G') = \deg_u^+(G')$. Also, $\mu_G(w,u') + \mu_G(u,u') = \mu_{G'}(w,u') + \mu_{G'}(u,u')$. Hence,

$$F_{u'}(G,b) = \sum_{s \in P_{u'}^1(G') \setminus \{u,w\}} \frac{\mu_{G'}(s,u')}{\deg_s^+(G')} + \frac{\mu_{G'}(w,u') + \mu_{G'}(u,u')}{\deg_w^+(G')} + b'(v) = F_{u'}(G',b').$$

  For w' we get analogously $F_{w'}(G,b) = F_{w'}(G',b')$.

- For Baseline, assume v is isolated. Since v has no direct predecessors, we have $F_v(G,b) = b(v)$.

Finally, consider Node Redirect. Let $(G,b) = ((\{u,v,w\}, \{\!\{(u,v),(w,v)\}\!\}), [0,0,0])$. We have $F_v(G,b) = 2$. Note that u and w are out-twins. Now, if we redirect node u into w we will get graph $(G',b') = ((\{v,w\}, \{\!\{(w,v)\}\!\}), [0,0])$. Here, we have that $F_v(G',b') = 1$. Thus, Node Redirect is not satisfied. □

Lemma 21. *A centrality measure F defined for every graph (G,b) and node v and an arbitrary $a \in (0,1)$ as*

$$F_v^a(G,b) = 2 \cdot PR_v^a(G,b)$$

satisfies Node Deletion, Edge Deletion, Edge Multiplication, Edge Swap, and Node Redirect, but does not satisfy Baseline.

Proof. Fix $a \in (0,1)$. Observe that for any two graphs (G,b), (G',b') and a node v, if $PR_v^a(G,b) = PR_v^a(G',b')$, then

$$F_v^a(G,b) = 2 \cdot PR^a(G,b) = 2 \cdot PR_v^a(G',b') = F_v^a(G',b'). \tag{3.36}$$

We will consider each axiom separately. Fix an arbitrary graph $G = (V,E)$ and node weights b.

- For Node Deletion, assume $u \in V$ is an isolated node. Let us consider graph $(G',b') = ((V \setminus \{u\}, E), b)$. Since PageRank satisfies Node Deletion, Eq. (3.36) implies $F_v^a(G,b) = F_v^a(G',b')$, for every $v \in V \setminus \{u\}$.

- For Edge Deletion, consider edge, $(u,w) \in E$. Let $(G',b') = ((V, E - \{\!\{(u,w)\}\!\}), b)$. Since PageRank satisfies Edge Deletion, Eq. (3.36) implies $F_v^a(G,b) = F_v^a(G',b')$, for every $v \in V \setminus S_u(G)$.

- For Edge Multiplication, consider node $u \in V$, constant $k \in \mathbb{N}$ and graph $(G', b') = ((V, E \sqcup k \cdot \Gamma_u^+(G)), b)$. Since PageRank satisfies Edge Multiplication, Eq. (3.36) implies $F_v^a(G, b) = F_v^a(G', b')$, for every $v \in V$.

- For Edge Swap, take $(u, u'), (w, w') \in E$ such as $F_u^a(G, b) = F_w^a(G, b)$ and also $\deg_u^+(G) = \deg_w^+(G)$. From the definition of F^a we have $PR_u^a(G, b) = PR_w^a(G, b)$ as well. Let $(G', b') = ((V,\ E - \{\!\{(u, u'), (w, w')\}\!\} \sqcup \{\!\{(u, w'), (w, u')\}\!\}),\ b)$. Since PageRank satisfies Edge Swap, Eq. (3.36) implies $F_v^a(G, b) = F_v^a(G', b')$, for every $v \in V$.

- For Node Redirect, assume $u, w \in V$ are out-twins. Let $(G', b') = R_{u \to w}(G, b)$. Since PageRank satisfies Node Redirect, by Eq. (3.36), $F_v^a(G, b) = F_v^a(G', b')$, for every $v \in V \setminus \{u, w\}$. Analogously, we get that

$$F_u^a(G, b) + F_w^a(G, b) = 2 \cdot (PR_u^a(G, b) + PR_w^a(G, b)) = \\ 2 \cdot PR_w^a(R_{u \to w}(G, b)) = F_w^a(R_{u \to w}(G, b)).$$

Finally, consider Baseline. Observe that $F_v^a((\{v\}, \emptyset), [1]) = 2$. Hence, the axiom is not satisfied. □

3.3 Related Axiomatizations

In their seminal paper, Page et al. [62] introduced PageRank as an extension of a centrality measure that they called *simplified PageRank*, which is equivalent to Seeley index. Indeed, both centrality measures are closely connected, which manifests itself, inter alia, through their walk interpretations (see Section 2.2.3 for details). Therefore, in this section we present a detailed comparison of our axiomatic characterization of PageRank to existing axiomatizations of Seeley index. There are two such axiomatizations in the literature: one by Palacios-Huerta and Volij [63] and one by Altman and Tennenholtz [1].

Axiomatizing The Invariant Method

Palacios-Huerta and Volij [63] considered the problem of measuring the importance of scientific journals based on the journal citation network. In a citation network, nodes represent journals and an edge (A, B) represents a reference of journal A to journal B. The authors presented an axiomatization of the *invariant method* [66] which is equal to Seeley index of a journal in a citation network divided by the number of articles it has published. To this end, they proposed the following four axioms:

Invariance with Respect to Reference Intensity: *Multiplying the references in every journal by arbitrary constants, specific for each journal, does not affect the importance of any journal.*

This axiom, satisfied by both the invariant method and Seeley index, is also satisfied by PageRank. It is equivalent to Edge Multiplication with the only difference that it is formulated as an operation on all nodes at once.

Weak Homogeneity: *Imagine that there are only two journals, A and B, and both have the same number of articles and references (some to themselves, some to the other journal). If journal A has x times more references from B than B from A, then A is x times more important.*

This axiom is satisfied by both the invariant method and Seeley index. However, it is not satisfied by PageRank: PageRank is not proportional to the impact of predecessors as it takes into consideration also the weight of a node.

Weak Consistency: *Assume that every journal has the same number of articles and references. Consider deleting one of the journals, A, and for every journal B citing A, redirecting all references of B to A to journals that were originally cited by A, preserving the proportions in the numbers of citations (e.g., if B had six references to A and A had two references to C and one to D, then four references are added from B to C and two from B to D). Now, the importance of any journal in the resulting network is the same as in the original network.*

This axiom, satisfied by both the invariant method and Seeley index, is based on the fact that in both of these measures intermediaries transfer further the whole importance they got from their predecessors. In particular, if a node is added in the middle of an edge, then the importance of all nodes remain the same. PageRank does not satisfy this axiom, since the decay factor decreases the importance transferred by intermediaries.

Invariance to Splitting of Journals: *Consider an operation of splitting a journal into k identical copies in a way that each copy has exactly* $1/k$ *of the original edges to each cited journal. Now, if every journal is split into an arbitrary number of copies, then every copy will have the same importance as the original journal in the original network.*

This axiom is satisfied by the invariant method, but it is not satisfied by Seeley index: according to Seeley index, the total importance of all copies equals the original importance of a journal. The same is true also for PageRank. Node Redirect is based on the same idea, but instead of splitting all nodes into several copies, it considers merging two copies with possible different incoming edges into one node.

Axiomatizing The Ranking

Altman and Tennenholtz [1] proposed an axiomatization of the ranking of nodes that result from Seeley index. As a result, axioms are of the different nature, as they concern the relation between centralities of different nodes, but not the specific values. The authors proposed five such axioms:

Isomorphism: *In two isomorphic graphs the ranking is the same with respect to the isomorphism.*

This axiom, proposed originally by Sabidussi [69], is satisfied by all reasonable centrality measures, including all measures introduced in this paper.

Self Edge: *Adding a self-loop to a node can only increase its position in the ranking and does not affect the ranking of other nodes.*

This axiom is not satisfied by PageRank, because adding a self-loop to a node can significantly decrease the centrality of its direct successors and change their ranking with respect to other nodes.

Vote by Committee: *Splitting a node into* $k+1$ *parts in a way that one node has the original incoming edges and k outgoing edges to other parts and k nodes have one* *incoming edge each and the original outgoing edges of the node does not affect the ranking of nodes in the graph.*

This axiom is based on a similar principle as Weak Consistency from the axiomatization of the invariant method. Similarly, it is not satisfied by PageRank since the decay factor decreases the importance transferred by the intermediaries. Hence, the direct predecessors of the split node may end up with a lower ranking.

Collapsing: *Redirecting a node into its out-twin does not affect the ranking of other nodes.*

This axiom, very similar to Node Redirect, is satisfied also by PageRank.

Proxy: *Assume that there is a node v with k incoming and k outgoing edges such that all direct predecessors are different and have equal centralities. Removing node v and adding k edges from direct predecessors to direct successors, one edge to each node, does not affect the ranking of any node in a graph.*

This axiom, similar to Vote by Committee, is not satisfied by PageRank because of the same reason: the decay factor decreases the importance transferred by the intermediaries.

As we can see, in both axiomatization there is an axiom that concerns splitting or merging out-twin nodes, as in Node Redirect. Also, in the axiomatization of the invariant method, there is an axiom equivalent to Edge Multiplication. However, most axioms in both axiomatizations are not satisfied by PageRank. That is why the axiomatic characterizations described in this section cannot be easily extended for the characterization of PageRank.

Chapter 4

Random Walk Decay Centrality

In the previous chapter, we focused on the axiomatic characterization of PageRank. As we have argued in the Introduction, various properties of PageRank make it suitable for particular applications, but at the same time, the same properties may be undesirable in other settings. In particular, PageRank of a node can be increased by changing the outgoing edges of the node. In the social media network, where users can decide who they want to follow, this allows for strategic manipulation of one's connections in order to increase their importance score.

In this chapter, we introduce an alternative for PageRank, named *random walk decay centrality* (RWD), which is robust to changes in the outgoing edges of the assessed node. We prove that it can be uniquely characterized by six axioms: *Random Walk Property*, *Locality*, *Sink Merging*, *Lack of Self-Impact*, *Directed Leaf Proportionality*, and *One-Node Graph*. Moreover, we show that exchanging one of the axioms, Random Walk Property, for a new axiom, *Shortest Paths Property*, results in a unique characterization of standard personalized decay centrality from Section 2.2.1. Similarly, if in the axiomatization of RWD we exchange axiom Lack of Self-Impact for Edge Swap from Section 3.1, then we obtain another axiomatic characterization of PageRank. Based on the analysis of our axioms, we argue that RWD has properties, violated by PageRank, that can be desirable in various setting.

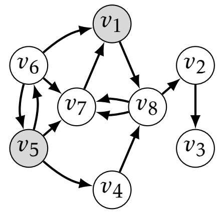
This chapter is organized as follows. First, we introduce an additional notation as well as the definition of random walk decay centrality in Section 4.1. Then, in Section 4.2, we introduce axiomatic characterizations of RWD, personalized decay centrality, and PageRank. Section 4.3 is devoted to the proofs that our axiomatizations indeed uniquely characterize respective measures. Finally, in Section 4.4, we take a closer look at the differences in axioms satisfied by PageRank and RWD and discuss the implications.

The content of this chapter is an extended version of the paper published in the proceedings of the AAAI-19 conference [81].

4.1 Definitions

In this section we introduce additional notation and the definition of random walk decay centrality.

As the name suggest, random walk decay centrality is primarily based on walks. Recall that by $p^a_{G,b}(v,t)$ we denote that probability that node v is visited at step t of the random walk on graph (G,b), assuming that in each step of the walk there is probability $a \in [0,1]$ that it does not stop (Eq. (2.6)). Now, we are interested not only in the fact that node v is visited at step t, but also we want to know how many



(G, b)

Figure 4.1: An example graph. Grey nodes have weights equal to 1 and the weight of white nodes is 0.

times v was visited before. Thus, let us denote the probability that the random walk visits node v for the k-th time at step t by $p^a_{G,b}(v,t,k)$. Formally, let

$$p^a_{G,b}(v,t,k) = \sum_{\substack{\omega \in \Omega_t(G):\omega(t)=v,\\ |\{s \le t:\omega(s)=v\}|=k}} \frac{b(\omega(0))}{b(G)} \cdot \prod_{i=0}^{t-1} \frac{a \cdot \mu_G\big(\omega(i), \omega(i+1)\big)}{\deg^+_{\omega(i)}(G)}, \quad \text{if } b(G) > 0, \tag{4.1}$$

and $p^a_{G,b}(v,t,k) = 0$, otherwise. Note that the only difference between Eq. (4.1) and Eq. (2.6) is that we restrict the sum to the walks that visited v exactly k times, i.e, we take $\omega \in \Omega_t(G)$ such that $|\{s \le t : \omega(s) = v\}| = k$. Clearly, it holds that $\sum_{k\ge 1} p^a_{G,b}(v,t,k) = p^a_{G,b}(v,t)$.

Example 4. *Consider once again graph (G,b) from Fig. 2.1 (for readers convenience, presented also in Fig. 4.1). Let us calculate the value of $p^a_{G,b}(v_1,3,1)$, i.e., the probability that node v_1 is visited for the first time at step 3. To this end, let us consider all walks of length 3 that end at v_1. There are five such walks: $\omega_1 = (v_1,v_8,v_7,v_1)$, $\omega_2 = (v_4,v_8,v_7,v_1)$, $\omega_3 = (v_5,v_6,v_7,v_1)$, $\omega_4 = (v_6,v_5,v_6,v_1)$, and $\omega_5 = (v_6,v_5,v_7,v_1)$. In walks ω_2–ω_5, node v_1 is indeed visited for the first time at step 3. However, in ω_1, at step 3, it is visited for the second time (first time being at step 0), thus we do not count this walk to calculate $p^a_{G,b}(v_1,3,1)$. Also, walks ω_2, ω_4, and ω_5 start at a node with zero weight, i.e., we have that $b(\omega_2(0)) = b(\omega_4(0)) = b(\omega_5(0)) = 0$, hence their input is also equal to zero. Therefore, the probability that the random walk visits node v_1 for the first time at step 3 is equal to the probability that the random walk starts with sequence ω_3, which (as established in Example 1) is equal to $1/2 \cdot a/3 \cdot a/2 \cdot a = a^3/12$. Thus, $p^a_{G,b}(v_1,3,1) = a^3/12$.*

As stated in Section 2.2.3, PageRank is defined as the expected number of visits at a node, regardless of the number of times it has been visited, i.e.,

$$PR^a_v(G,b) = b(G) \cdot \sum_{t\ge 0} p^a_{G,b}(v,t) = b(G) \cdot \sum_{t\ge 0} \sum_{k\ge 1} p^a_{G,b}(v,t,k). \tag{4.2}$$

In contrast, *random walk decay centrality* is defined as the expected number of the first time visits at a node or, equivalently, as the probability that a node will be visited at all, i.e.,

$$RWD^a_v(G,b) = b(G) \cdot \sum_{t\ge 0} p^a_{G,b}(v,t,1). \tag{4.3}$$

From random walk definitions of PageRank and RWD we obtain the following relation between these measures: for every node, $v \in V$, its RWD is equal to PageRank of v in the graph obtained from the original graph by removing all outgoing edges of v.

Proposition 22. *For every decay factor $a \in [0,1)$, graph $G = (V,E)$, node weights b, and node $v \in V$, it holds that*

$$RWD_v^a(G,b) = PR_v^a((V, E - \Gamma_v^+(G)), b).$$

Proof. Denote $G' = (V, E - \Gamma_v^+(G))$. Observe that walks that end at v and do not visit v before are unaffected by removal of outgoing edges of v. Hence, from Eq. (4.1) we have $p^a_{G',b}(v,t,1) = p^a_{G,b}(v,t,1)$, for every $t \in \mathbb{N}$. On the other hand, v in graph G' is a sink, thus it cannot be visited by a walk more than once. Hence, $p^a_{G',b}(v,t,k) = 0$, for every $t \in \mathbb{N}$ and $k > 1$. Summing over all $k \geq 1$ we get

$$\sum_{k \geq 1} p^a_{G',b}(v,t,k) = p^a_{G,b}(v,t,1).$$

Hence, the thesis follows from Eq. (4.2) and Eq. (4.3). □

4.2 Axioms

In this section, we introduce our axioms and use them to characterize random walk decay centrality, personalized decay centrality, and PageRank.

First, let us introduce the following six axioms.

- ***Random Walk Property*** *(centrality of a node depends only on its visit probabilities and the sum of node weights in the graph): For every two graphs $G = (V,E)$, $G' = (V',E')$, node weights b, b', and node $v \in V \cap V$ such that $b(G) = b'(G')$ and $p^1_{G,b}(v,t,k) = p^1_{G',b'}(v,t,k)$, for every $t,k \in \mathbb{N}$, it holds that*

  $$F_v(G,b) = F_v(G',b').$$

- ***Locality*** *(centrality of a node depends only on the connected component of this node): For every two disjoint graphs $G = (V,E)$, $G' = (V',E')$ and node weights b, b', it holds that*

  $$F_v(G,b) = F_v(G+G', b+b'), \quad \textit{for every } v \in V.$$

- ***Sink Merging*** *(redirecting a sink into another sink without a common predecessor sums up their centralities and does not affect the centrality of other nodes): For every graph $G = (V,E)$, node weights b, and sinks $u,w \in V$ such that $P_u(G) \cap P_w(G) = \emptyset$, it holds that*

  $$F_v(G,b) = F_v(R_{u\to w}(G,b)), \quad \textit{for every } v \in V \setminus \{u,w\},$$

  and $F_u(G,b) + F_w(G,b) = F_w(R_{u\to w}(G,b))$.

- ***Lack of Self-Impact*** *(centrality of a node does not depend on the outgoing edges of this node): For every graph $G = (V,E)$, node weights b, and edge $(v,u) \in E$, it holds that*

  $$F_v(G,b) = F_v((V, E - \{\!\{(v,u)\}\!\}), b).$$

- ***Directed Leaf Proportionality*** *(adding an edge from a sink to an isolated node increases the centrality of this node by the centrality of the sink times a constant): There exists a constant, $a \in \mathbb{R}_{\geq 0}$, such that for every graph $G = (V,E)$, node weights b, sink $u \in V$, and isolated node $v \in V$, it holds that*

  $$F_v((V, E \sqcup \{\!\{(u,v)\}\!\}), b) - F_v(G,b) = a \cdot F_u(G,b).$$

- ***One-Node Graph*** *(a node with unit weight in a graph without any edges nor other nodes has centrality equal to one): For every node v, it holds that*

$$F_v((\{v\},\emptyset),[1]) = 1.$$

The first axiom, Random Walk Property, allows us to restrict our attention to only those centrality measures that are based solely on the visit probabilities and the sum of node weights in the graph.

The next three axioms are invariance axioms. The first one of them, Locality, initially proposed by Skibski et al. [76], can be seen as a stronger version of Node Deletion from Section 3.1. Node Deletion states that removing an isolated node does not affect the centrality of other nodes. In turn, Locality allows for removing a whole part of a graph without affecting the centrality of other nodes as long as there is no connection, even disregarding the direction of edges, between these nodes and the removed part. Note that Node Deletion and Edge Deletion imply Locality (Lemma 6). However, Locality implies only Node Deletion and not Edge Deletion. Sink Merging, i.e., our next axiom, is a weaker version of Node Redirect from Section 3.1. Node Redirect states that redirection of arbitrary out-twins results in summing their centralities and does not affect the other nodes. In Sink Merging, both nodes have to be sinks and they cannot share any predecessors as well. The last invariance axiom is Lack of Self-Impact. It states that the centrality of a node does not depend on its outgoing edges. In the setting in which nodes can decide upon their outgoing edges this axiom translates to unmanipulability (or strategy-proofness) property of centrality measures (see Section 4.4 for details).

The fifth axiom, Directed Leaf Proportionality, is inspired by Leaf Proportionality axiom proposed by Skibski and Sosnowska [77] for undirected and unweighted graphs. The axiom binds the centrality of a leaf, i.e., a sink with one incoming edge, with the centrality of its direct predecessor, when this direct predecessor has only one outgoing edge.

The first five axioms uniquely characterize RWD up to a scalar multiplication (see Lemma 38). Our last axiom, One-Node Graph, plays a similar role as Baseline axiom in characterization of PageRank from Chapter 3, i.e, it specifies the centrality in a simple borderline case: a graph with one node, unit weight, and no edges. As a result, RWD is uniquely characterized, a proof of which is presented in Section 4.3.1.

Theorem 23. *A centrality measure satisfies Random Walk Property, Locality, Sink Merging, Lack of Self-Impact, Directed Leaf Proportionality, and One-Node Graph if and only if it is random walk decay centrality.*

Personalized decay centrality, as defined in Section 2.2.1, satisfies five out of six of our axioms, i.e., Locality, Sink Merging, Directed Leaf Property, Lack of Self-Impact, and One-Node Graph. However, unsurprisingly, it does not satisfy Random Walk Property since it is based on distances in a graph rather than visit probabilities. In order to obtain a unique characterization, let us introduce *Shortest Paths Property* axiom that is analogous to Random Walk Property, but takes distances into account instead of probabilities of visits. Our axiom is a direct translation of the definition of the class of *distance based centralities* by Skibski and Sosnowska [77] to weighted and directed graphs.

> ***Shortest Paths Property*** *(centrality of a node depends only on the distances to it from other nodes and the sum of node weights in the graph): For every two graphs, $G = (V,E)$, $G' = (V',E')$, node weights b, b', and node*

$v \in V \cap V'$ such that $b(G) = b'(G')$ and $|\{u \in V : dist_G(u,v) = k \wedge b(u) = x\}| = |\{u \in V' : dist_{G'}(u,v) = k \wedge b'(u) = x\}|$, for every $k \in \mathbb{N}$ and $x \in \mathbb{R}$, it holds that

$$F_v(G,b) = F_v(G',b').$$

The following theorem states that if in the axiomatic characterization of RWD we exchange Random Walk Property for Shortest Path Property, then we obtain the unique characterization of personalized decay centrality. Observe that Shortest Paths Property implies Lack of Self-Impact: the distance from other nodes to a node does not depend on its outgoing edges. Hence, the latter axiom is omitted from the theorem.

Theorem 24. *A centrality measure satisfies Shortest Paths Property, Locality, Sink Merging, Directed Leaf Proportionality, and One-Node Graph if and only if it is personalized decay centrality.*

The proof of the theorem is presented in Section 4.3.2.

Finally, let us consider PageRank. Observe that it also satisfies five out of six axioms from the characterization of RWD, namely: Random Walk Property, Locality, Sink Merging, Directed Leaf Proportionality, and One-Node Graph. However, observe that adding a loop to a sink increases its PageRank. Therefore, PageRank violates Lack of Self-Impact.

In order to extend these five axioms to a unique characterization of PageRank we incorporate Edge Swap axiom from Section 3.1.

Theorem 25. *A centrality measure satisfies Random Walk Property, Locality, Sink Merging, Edge Swap, Directed Leaf Proportionality, and One-Node Graph if and only if it is PageRank.*

The proof of the theorem is given in Section 4.3.3. The consequences of the differences in axioms characterizing RWD and PageRank are discussed in detail in Section 4.4.

4.3 Proofs of Uniqueness

In this section, we present the proofs of Theorems 23–25. We start with the unique characterization of RWD (Theorem 23), then we move to personalized decay centrality (Theorem 24), and finally we address the unique characterization of PageRank (Theorem 25).

4.3.1 Random Walk Decay Centrality (Theorem 23)

Let us begin by proving that RWD satisfies all six of our axioms.

Lemma 26. *For every decay factor $a \in [0,1)$, random walk decay centrality satisfies Random Walk Property, Locality, Sink Merging, Lack of Self-Impact, Directed Leaf Proportionality, and One-Node Graph.*

Proof. Let us consider an arbitrary graph $G = (V,E)$ and node weights b, and consider the axioms one by one.

- For Random Walk Property, consider graph $G' = (V',E')$, node weights b', and node $v \in V \cap V'$ such that $b(G) = b'(G')$ and $p^1_{G,b}(v,t,k) = p^1_{G',b'}(v,t,k)$, for every $t,k \in \mathbb{N}$. Observe that from Eq. (4.1) we get that

$$p^a_{G,b}(v,t,k) = a^t \cdot p^1_{G,b}(v,t,k) = a^t \cdot p^1_{G',b'}(v,t,k) = p^a_{G',b'}(v,t,k),$$

for every $t, k \in \mathbb{N}$. Hence, the axiom follows from Eq. (4.3).

- For Locality, take graph $G' = (V', E')$ such that $V \cap V' = \emptyset$, weights b', and node $v \in V$. Observe that in graph $(G + G', b + b')$ any walk that starts at one of the nodes in V' cannot visit nodes in V and vice versa. Thus, for every $t \in \mathbb{N}$, we have that $\{\omega \in \Omega_t(G + G') : \omega(t) = v \wedge |\{s \leq t : \omega(t) = v\}| = 1\} = \{\omega \in \Omega_t(G) : \omega(t) = v \wedge |\{s \leq t : \omega(t) = v\}| = 1\}$. Since proportions of numbers of the outgoing edges and the out-degrees of nodes in V are the same in both (G, b) and $(G+G', b+b')$, by Eq. (4.1), this implies that $p^a_{G+G',b+b'}(v, t, 1)/(b(G)+b'(G')) = p^a_{G,b}(v, t, 1)/b(G)$. Hence, Locality follows from Eq. (4.3).

- For Sink Merging, consider two sinks $u, w \in V$ such that $P_u(G) \cap P_w(G) = \emptyset$. Fix $v \in V \setminus \{u, w\}$. Since u and w are sinks, it holds that every walk of length t that visits v for the first time at step t, it does not visit u nor w before that. Hence, the same walk is also present in graph $R_{u\to w}(G, b)$. Furthermore, since the probability of moving between any two nodes in $V \setminus \{u, w\}$ is also unaffected, from Eq. (4.1) we have $p^a_{R_{u\to w}(G,b)}(v, t, 1) = p^a_{G,b}(v, t, 1)$. Thus, from Eq. (4.3) we get that $RWD^a_v(R_{u\to w}(G, b)) = RWD^a_v(G, b)$.

  It remains to show that $RWD^a_w(R_{u\to w}(G, b)) = RWD^a_u(G, b) + RWD^a_w(G, b)$. To this end, observe that from Proposition 22 we get $RWD^a_u(G, b) = PR^a_u(G, b)$, $RWD^a_w(G, b) = PR^a_w(G, b)$, and also $RWD^a_w(R_{u\to w}(G, b)) = PR^a_w(R_{u\to w}(G, b))$. Thus, the axiom follows from the fact that PageRank satisfies Sink Merging (Lemma 44).

- For Lack of Self-Impact, consider edge $(v, u) \in E$ and let $G' = (V, E - \{\!\{(v, u)\}\!\})$. Observe that for every $t \in \mathbb{N}$, walks of length t that visit v for the first time at step t are unaffected by removal of edge (v, u). Hence, from Eq. (4.1) we have $p^a_{G',b}(v, t, 1) = p^a_{G,b}(v, t, 1)$, for every $t \in \mathbb{N}$. Thus, the axiom follows from Eq. (4.3).

- For Directed Leaf Proportionality, consider sink $u \in V$ and an isolated node $v \in V$. Also, denote $G' = (V, E \sqcup \{\!\{(u, v)\}\!\})$. Now, from Proposition 22 we get $RWD^a_v(G, b) = PR^a_v(G, b)$, $RWD^a_u(G, b) = PR^a_u(G, b)$, and $RWD^a_v(G', b) = PR^a_v(G', b)$. Thus, since PageRank satisfies Directed Leaf Proportionality (Lemma 44), we get that RWD satisfies it as well.

- Finally, for One-Node Graph assume that $(G, b) = ((\{v\}, \emptyset), [1])$. Observe that in such a case, from Eq. (4.1) we have $p^a_{G,b}(v, 0, 1) = 1$. Also, since there are no walks of positive length in G, we have that $p^a_{G,b}(v, t, 1) = 0$, for every $t > 0$. Hence, the axioms follows from Eq. (4.3).

□

Next, we move to the main part of the proof in which we show that if centrality measure F satisfies Random Walk Property, Locality, Sink Merging, Lack of Self-Impact, and Directed Leaf Proportionality, then it is equal to RWD up to a scalar multiplication. Observe that in contrast to the proof of unique characterization of PageRank in Section 3.2, here the decay factor of a centrality measure is directly given by constant a in Directed Leaf Proportionality. Thus, formally, we prove that there exists a constant, $c_F \in \mathbb{R}_{\geq 0}$, such that for every graph $G = (V, E)$, node weights b, and node $v \in V$, it holds that $F_v(G, b) = c_F \cdot RWD^a_v(G, b)$. Then, by adding One-Node Graph axiom, we obtain that $c_F = 1$ and $F_v(G, b) = RWD^a_v(G, b)$.

First, we prove that for every graph in which all nodes have zero weight it holds that each node has zero centrality.

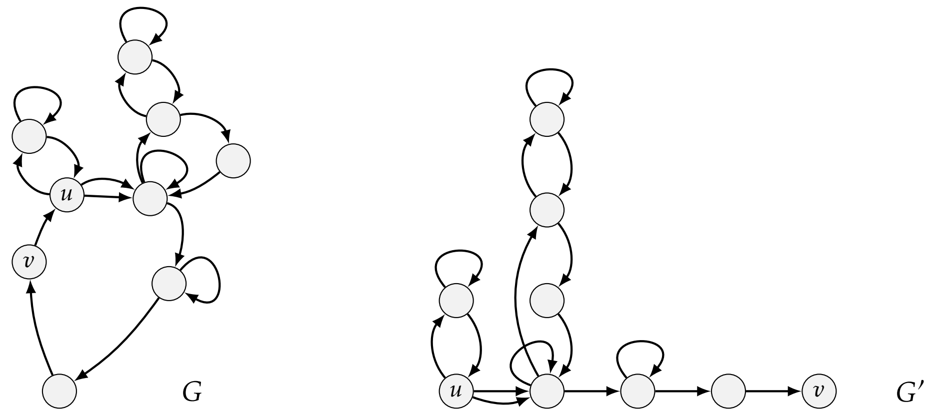


Figure 4.2: An example cactus, G, and a broken cactus, G'. Observe that by adding edge (v,u) to G' we obtain graph G.

Lemma 27. *If a centrality measure, F, satisfies Random Walk Property, Locality, and Sink Merging, then for every graph $G=(V,E)$, node weights b such that $b(G)=0$, node $v \in V$, and $a \in [0,1)$, it holds that*

$$F_v(G,b) = 0 = RWD^a_v(G,b).$$

Proof. For the second equation, observe that from Eq. (4.3) we get $RWD^a_v(G,b)=0$, for every graph $G=(V,E)$, node weights b such that $b(G)=0$, node $v \in V$, and $a \in [0,1)$.

For the first equation, let us observe that for every two graphs $G=(V,E)$ and $G'=(V',E')$, node weights b,b' such that $b(G)=b'(G')=0$, and nodes $v \in V$ and $v' \in V'$, by Eq. (4.1), it holds that $p^1_{G,b}(v,t,k) = 0 = p^1_{G',b'}(v',t,k)$, for every $t,k \in \mathbb{N}$. Hence, by Random Walk Property, $F_v(G,b) = F_{v'}(G',b')$. Thus, it suffices to show that there exists a node with a zero centrality in a graph in which all nodes have zero weights.

To this end, let us consider graph $(G,b) = ((\{u,v\},\emptyset),[0,0])$. Also, let us denote the graph obtained by redirecting u into v by $(G',b') = R_{u\to v}(G,b) = ((\{v\},\emptyset),[0])$. Since we have that both u and v are sinks in graph G, from Sink Merging we obtain that $F_v(G',b') = F_v(G,b) + F_u(G,b)$. On the other hand, observe that from Locality we have that $F_v(G',b') = F_v(G,b)$. Hence, $F_u(G,b)=0$. □

Therefore, in the remainder of the proof we will focus on graphs with at least one node with a positive weight.

Let us introduce three additional concepts that we will use in the proof: cactus graphs, broken cactus graphs, and visit probability generating functions. We begin with the definition of (directed) *cactus* graphs [64].

Definition 2. *Graph $G=(V,E)$ is a* cactus *if for every two distinct nodes $u,v \in V$ there exists exactly one path from u to v.*

Intuitively, cacti resemble undirected trees: If we look at all significantly different cycles (by significantly different we mean that one cannot be obtain by the other by a different choice of the first node) in a cactus, then every two such cycles can share at most one common node. Moreover, if we consider an undirected graph in which nodes represent cycles and edges the fact that corresponding cycles share a common node, then such graph is a tree. See Fig. 4.2 for an illustration.

Next, let us define related class of *broken cactus* graphs.

Definition 3. *Graph $G = (V, E)$ is a* broken cactus *with* start u *and* end v *if* $u, v \in V$, v *is a sink, and graph* $(V, E \sqcup \{\!\{(v, u)\}\!\})$ *is a cactus.*

Consider an arbitrary broken cactus $G = (V, E)$ with start u and end v. Since $(V, E \sqcup \{\!\{(v, u)\}\!\})$ is a cactus, there is exactly one path from w to v, for every $w \in V \setminus \{v\}$. It does not use edge (v, u), thus in graph G there is also exactly one path from w to v. In particular there is a path from the start, u, to the end, v. We call it the *main path* of broken cactus G.

Observe that for a given broken cactus G, its end and start, and hence also its main path, is uniquely determined. Specifically, the end of a broken cactus is the only sink in a graph. In turn, the start is a node with maximal distance to the start such that it has an outgoing edge to a node that is not its predecessor (in general, a node is on the main path if and only if it is the end or it has an outgoing edge to a node that is not its predecessor).

Alternatively, we can think of a broken cactus as of its main path with possible cacti attached to every node of the main path excluding its end. See Fig 4.2 for an illustration.

Apart from cacti and broken cacti, another important concept that is used in the proof is the *visit probability generating function*.

Definition 4. *For every graph $G = (V, E)$, node weights b, and node $v \in V$, the* visit probability generating function *is defined as a formal power series of the form*

$$P^v_{G,b}(x) = \sum_{t=0}^{\infty} x^t \cdot p^1_{G,b}(v, t). \tag{4.4}$$

Visit probability generating functions allow us to describe in an efficient yet formal way how often and at which steps a given node is visited by the random walk. Moreover, we will be able to express how visit probabilities of nodes are affected by certain graph operations.

With these concepts in hand, we can move to the proof itself. In short, we show that for every graph G and its node v that is a sink, there exists a collection of broken cacti such that the average of the visit probability generating functions of their ends is equal to the visit probability generating function of v (Proposition 33). Next, we prove that if centrality measure F satisfies our axioms, then it is equal to RWD (up to a scalar multiplication) for every end of a broken cactus (Lemma 36). Combining both facts with Random Walk Property, we obtain that F is equal to RWD for every sink in every graph. Thus, we obtain the thesis from Lack of Self-Impact.

More in detail, the proof is structured as follows:

- First, we introduce the notion of *limiting nodes*, i.e., nodes for which the visit probability decreases to zero over time, and prove that all of the predecessors of such nodes are also limiting (Propositions 28).

- Next, we prove several useful properties of visit probability generating functions (Proposition 29).

- Then, we focus on visit probability generating functions of the ends of broken cacti and prove several results for them:

  - We show that appending two broken cacti results in multiplying generating functions of their ends (Proposition 30).

  - Next, we prove that we can obtain an average of generating functions of the ends of broken cacti, if we add them together and redirect all of their ends into one (Proposition 31).
  - Then, we show how we can reduce the level of complexity of a broken cactus, by relating the generating function of its end to the generating function of the ends of broken cacti with fewer cycles (Proposition 32).
  - Finally, building upon these result, we show that for every graph and its node that is limiting, its generating function is an average of the generating functions of the ends of some broken cacti (Proposition 33).

- After establishing all needed properties of visit probability generating function, we consider an arbitrary centrality measure F that satisfies our axioms. In a series of lemmas we show that it is equal to RWD:
  - First, we consider properties that can be regarded as stronger version of Random Walk Property for sinks (Lemmas 34 and 35).
  - Next, we show that for every broken cactus the centrality of its end is equal to RWD up to a scalar multiplication (Lemma 36).
  - Finally, combining Lemma 36 and Proposition 33 we prove that for every graph the centrality is equal to RWD (up to a scalar multiplication) for an arbitrary sink (Lemma 37) and node (Lemma 38). Thus, from One-Node Graph we get the thesis (Lemma 39).

Visit Probability Generating Functions and Broken Cacti

Let us begin with the definition of *limiting* nodes.

Definition 5. *For every graph $G = (V,E)$ and node weights b, we say that node $v \in V$ is* limiting *if*

$$\sum_{t=0}^{\infty} p^{1}_{G,b}(v,t) < \infty.$$

The concept of limiting nodes corresponds to the concept of *transient states* in Markov Chains [27]. Observe that alternatively we can say that node v in graph (G,b) is limiting if there exists $r \in \mathbb{R}_{\geq 0}$ such that $P^{v}_{G,b}(1) = r$. One important property of limiting nodes is that all of their predecessors are also limiting.

Proposition 28. *For every graph $G = (V,E)$, node weights b such that $b(G) > 0$, and node $v \in V$, if v is limiting, then also u is limiting, for every $u \in P_v(G)$.*

Proof. Let us assume that the thesis is false, i.e., there exists graph $G = (V,E)$, node weights b such that $b(G) > 0$, and nodes v,u such that v is limiting and $u \in P_v(G)$, but u is not limiting. Since u is a predecessor of v, there exists a walk ω of length s that starts at u and ends at v and does not visit u in between. Hence, there exists some positive probability, $p^* > 0$, that when the random walk visits node u at step t, then the random walk follows walk ω and visit node v at step $t+s$. Thus, for every $t \in \mathbb{N}$, we have $p^{1}_{G,b}(v,t+s) \geq p^* \cdot p^{1}_{G,b}(u,t)$. Summing for all $t \geq 0$, we get

$$\sum_{t=s}^{\infty} p^{1}_{G,b}(v,t) \geq p^* \cdot \sum_{t=0}^{\infty} p^{1}_{G,b}(u,t).$$

However, since u is not limiting, the right hand side of this inequality is infinite. We have that $\sum_{t=0}^{\infty} p^{1}_{G,b}(v,t) \geq \sum_{t=s}^{\infty} p^{1}_{G,b}(v,t)$, hence we arrive at a contradiction. □

Now, let us move to visit probability generating functions and their properties. The most basic ones are listed in the following proposition.

Proposition 29. *For every graph $G = (V, E)$, node weights b such that $b(G) > 0$, and node $v \in V$, it holds that:*

(a) *(Isomorphism) $P^v_{G,b}(x) = P^{v'}_{G',b'}(x)$, for every graph $G' = (V', E')$ and weights b' such that graph (G', b') is isomorphic to (G, b) with isomorphism $f : V \to V'$ and $f(v) = v'$,*

(b) *(Recursive Equation) $P^v_{G,b}(x) = \frac{b(v)}{b(G)} + \sum_{u \in P^1_v(G)} x \cdot \frac{\mu_G(u,v)}{\deg^+_u(G)} P^u_{G,b}(x)$,*

(c) *(Sink Redirection) $P^v_{R_{u\to v}(G,b)}(x) = P^u_{G,b}(x) + P^v_{G,b}(x)$, for every $u \in V$, if u and v are sinks,*

(d) *(Weight Multiplication) $P^v_{G,r\cdot b}(x) = P^v_{G,b}(x)$, for every $r \in \mathbb{R}_{>0}$,*

(e) *(Graph Addition) $P^v_{G+G',b+b'}(x) = \frac{b(G)}{b(G)+b'(G')} P^v_{G,b}(x)$, for every graph $G' = (V', E')$ such that $V \cap V' = \emptyset$ and node weights b'.*

(f) *(Edge Multiplication) $P^v_{G',b}(x) = P^v_{G,b}(x)$, for every $u \in V$ and $k \in \mathbb{N}$ and graph $G' = (V, E \sqcup k \cdot \Gamma^+_u(G))$,*

(g) *(Cycle Decomposition) $P^v_{G,b}(x) = P^v_{G^\dagger,b}(x)/(1 - P^v_{G^*,\mathbb{1}_{v'}}(x))$ where $G^\dagger = (V, E - \Gamma^+_v(G))$ and $G^* = (V \cup \{v'\}, E - \Gamma^+_v(G) \sqcup \{\!\{(v', u) : (v, u) \in \Gamma^+_v(G)\}\!\})$ in which $v' \notin V$, and*

(h) *(Sink Bound) $P^v_{G,b}(1) \le 1$ if v is a sink and $P^v_{G,b}(1) = 1$ if v is the end of a broken cactus.*

Proof. Let us take an arbitrary graph $G = (V, E)$, node weights b, and node $v \in V$.

(a) *(Isomorphism)* Let (G', b') be a graph isomorphic to (G, b) with isomorphism f. Let us consider an arbitrary walk $\omega \in \Omega_t(G)$ that ends at v, i.e., $\omega(t) = v$. Then, by ω_f let us denote the walk of length t on graph (G', b') such that $\omega_f(i) = f(\omega(i))$, for every $i \in \{1, \dots, t\}$. Observe that in such a way we obtain a one-to-one correspondence between walks in $\Omega_t(G)$ and walks in $\Omega_t(G')$. Moreover, since the two graphs are isomorphic,

$$\frac{b(\omega(0))}{b(G)} \cdot \prod_{i=0}^{t-1} \frac{\mu_G(\omega(i), \omega(i+1))}{\deg^+_{\omega(i)}(G)} = \frac{b'(\omega_f(0))}{b'(G')} \cdot \prod_{i=0}^{t-1} \frac{\mu_{G'}(\omega_f(i), \omega_f(i+1))}{\deg^+_{\omega_f(i)}(G')}.$$

Thus, from Eq. (2.6) we have that $p^1_{G,b}(v,t) = p^1_{G',b'}(v',t)$, for every $t \in \mathbb{N}$, and the thesis follows from Eq. (4.4).

(b) *(Recursive Equation)* For $t > 0$ consider a walk $\omega \in \Omega_t(G)$ that ends at v, i.e., $\omega(t) = v$. Observe that node $\omega(t-1)$ is one of the direct predecessors of v, let say u. Then, in the last step, it follows one of the outgoing edges of u that goes to v, for which the probability is $\mu_G(u,v)/\deg^+_u(G)$. Thus, for the value $p^1_{G,b}(v,t)$ we get

$$p^1_{G,b}(v,t) = \sum_{u \in P^1_v(G)} \frac{\mu_G(u,v)}{\deg^+_u(G)} p^1_{G,b}(u, t-1).$$

Multiplying by x^t and taking the sum over all $t \geq 1$ we obtain

$$\sum_{t=1}^{\infty} x^t \cdot p^1_{G,b}(v,t) = \sum_{u \in P^1_v(G)} x \cdot \frac{\mu_G(u,v)}{\deg^+_u(G)} \sum_{t=1}^{\infty} x^{t-1} \cdot p^1_{G,b}(u,t-1).$$

Since $p^1_{G,b}(v,0) = b(v)/b(G)$, by adding this to both sides of the equation, we obtain the thesis from Eq. (4.4).

(c) *(Sink Redirection)* Let us assume that nodes $u,v \in V$ are sinks and let us denote graph $(G',b') = R_{u \to v}(G,b)$. Consider node $w \in V \setminus \{u,v\}$ and take an arbitrary walk $\omega \in \Omega_t(G)$ that ends at w. Since u and v are sinks, the walk that visits w at step t cannot visit u or v earlier. Hence, the probability associated with this walk is not affected by the redirection, i.e.,

$$\frac{b(\omega(0))}{b(G)} \cdot \prod_{i=0}^{t-1} \frac{\mu_G(\omega(i),\omega(i+1))}{\deg^+_{\omega(i)}(G)} = \frac{b'(\omega(0))}{b'(G')} \cdot \prod_{i=0}^{t-1} \frac{\mu_{G'}(\omega(i),\omega(i+1))}{\deg^+_{\omega(i)}(G')}.$$

Thus, from Eq. (2.6) we have that $p^1_{G,b}(w,t) = p^1_{G',b'}(w,t)$, for every $t \in \mathbb{N}$, which, by Eq. (4.4), means that $P^w_{G,b}(x) = P^w_{G',b'}(x)$, for every $w \in V \setminus \{u,v\}$.

Now, let us move to nodes u and v. From (b) we can sum the recursive equations for $P^u_{G,b}(x)$ and $P^v_{G,b}(x)$ to obtain

$$P^u_{G,b}(x) + P^v_{G,b}(x) = \frac{b(u)+b(v)}{b(G)} + \sum_{w \in P^1_u(G) \cup P^1_v(G)} \frac{\mu_G(w,u)+\mu_G(w,v)}{\deg^+_w(G)} \cdot P^w_{G,b}(x).$$

We know already that $P^w_{G,b}(x) = P^w_{G',b'}(x)$, for every node $w \in V \setminus \{u,v\}$. Also, since u and v are sinks it holds that $P^1_u(G) \cup P^1_v(G) \subseteq V \setminus \{u,v\}$. As for node weights, observe that redirection sums the weights of u in v and does not affect other weights. Thus, we have $b'(v) = b(u)+b(v)$ and $b'(G) = b(G)$. Finally, the out-degrees of all the nodes does not change and the incoming edges of v in G' are the sum of the incoming edges of u and v, i.e., $P^1_u(G') = P^1_u(G) \cup P^1_v(G)$ and $\mu_{G'}(w,u) = \mu_G(w,u) + \mu_G(w,v)$, for every $w \in P^1_u(G')$. Therefore, we get

$$P^u_{G,b}(x) + P^v_{G,b}(x) = \frac{b'(v)}{b'(G')} + \sum_{w \in P^1_u(G')} \frac{\mu_{G'}(w,u)}{\deg^+_w(G')} \cdot P^w_{G',b'}(x).$$

Hence, the thesis follows from (b) for node v in graph (G',b').

(d) *(Weight Multiplication)* Fix $r \in \mathbb{R}_{\geq 0}$. Observe that for every $t \geq 0$ and every walk $\omega \in \Omega_t(G)$ that ends at v, it holds that

$$\frac{b(\omega(0))}{b(G)} \cdot \prod_{i=0}^{t-1} \frac{\mu_G(\omega(i),\omega(i+1))}{\deg^+_{\omega(i)}(G)} = \frac{r \cdot b(\omega(0))}{r \cdot b(G)} \cdot \prod_{i=0}^{t-1} \frac{\mu_G(\omega(i),\omega(i+1))}{\deg^+_{\omega(i)}(G)}.$$

Hence, the thesis follows from Eq. (2.6) and Eq. (4.4).

(e) *(Graph Addition)* Let $G' = (V',E')$ be an arbitrary graph such that $V \cap V' = \emptyset$ and b' be arbitrary node weights. Let us denote $b'' = b + b'$. Consider an arbitrary walk $\omega \in \Omega_t(G+G')$ that ends at v. Since $b''(u) = b(u)$, for every $u \in V$, and edges of graph G are unaffected by adding graph G', we get

$$\frac{b''(\omega(0))}{b''(G+G')} \cdot \prod_{i=0}^{t-1} \frac{\mu_{G+G'}(\omega(i),\omega(i+1))}{\deg^+_{\omega(i)}(G+G')} = \frac{b(G)}{b(G)} \cdot \frac{b(\omega(0))}{b''(G+G')} \cdot \prod_{i=0}^{t-1} \frac{\mu_G(\omega(i),\omega(i+1))}{\deg^+_{\omega(i)}(G)}.$$

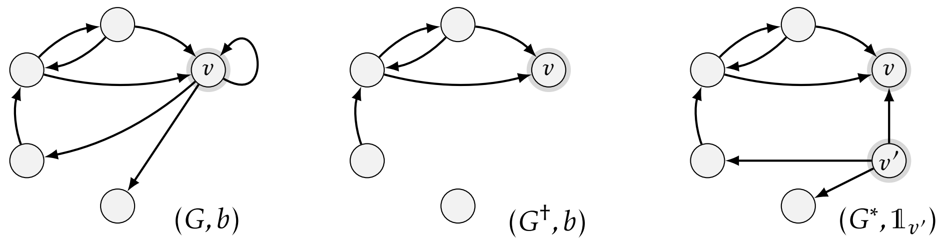


Figure 4.3: Example graphs illustrating Cycle Decomposition (Proposition 29g).

Observe that $\frac{b(\omega(0))}{b(G)} \cdot \prod_{i=0}^{t-1} \frac{\mu_G(\omega(i),\omega(i+1))}{\deg^+_{\omega(i)}(G)}$ is equal to the probability of walk ω in graph G. Also, observe that every walk on $G+G'$ that ends at v is a walk on G that ends at v and vice versa. Thus, summing the above equation for all such walks, from Eq. (2.6) we get that $p^1_{G+G',b+b'}(v,t) = \frac{b(G)}{b(G)+b'(G')} p^1_{G,b}(v,t)$, for every $t \in \mathbb{N}$. Thus, the thesis follows from Eq. (4.4).

(f) *(Edge Multiplication)* Fix $u \in V$ and $k \in \mathbb{N}$ and let $G' = (V, E \sqcup k \cdot \Gamma^+_u(G))$. Consider an arbitrary walk $\omega \in \Omega_t(G+G')$ that ends at v. Observe that

$$\frac{b(\omega(0))}{b(G)} \cdot \prod_{i=0}^{t-1} \frac{\mu_G(\omega(i),\omega(i+1))}{\deg^+_{\omega(i)}(G)} = \frac{b'(\omega(0))}{b'(G')} \cdot \prod_{i=0}^{t-1} \frac{\mu_{G'}(\omega(i),\omega(i+1))}{\deg^+_{\omega(i)}(G')},$$

because each time $\omega(i) = u$, both the numerator and the denominator are multiplied by $k+1$. Thus, the thesis follows from Eq. (2.6) and Eq. (4.4).

(g) *(Cycle Decomposition)* Let $G^\dagger = (V, E - \Gamma^+_v(G))$ be a graph with outgoing edges of v removed and $G^* = (V \cup \{v'\}, E - \Gamma^+_v(G) \sqcup \{\!\{(v',u) : (v,u) \in \Gamma^+_v(G)\}\!\})$ be a graph where node v is "uncycled", i.e., we change the start of each outgoing edge of v to an additional node v'. In this way, node v does not belong to any cycle in G^*. See Fig. 4.3 for an illustration.

Consider an arbitrary walk $\omega \in \Omega_t(G)$ that ends at v and visits it k times altogether. Observe that in such a walk there exists a step, $s \le t$, in which v is visited for the first time. Then, it holds that $\mathbb{1}_v(\omega(s))/\mathbb{1}_v(G) = 1$, thus

$$\frac{b(\omega(0))}{b(G)} \cdot \prod_{i=0}^{t-1} \frac{\mu_G(\omega(i),\omega(i+1))}{\deg^+_{\omega(i)}(G)} = \left(\frac{b(\omega(0))}{b(G)} \cdot \prod_{i=0}^{s-1} \frac{\mu_G(\omega(i),\omega(i+1))}{\deg^+_{\omega(i)}(G)}\right) \cdot \left(\frac{\mathbb{1}_v(\omega(s))}{\mathbb{1}_v(G)} \cdot \prod_{i=s}^{t-1} \frac{\mu_G(\omega(i),\omega(i+1))}{\deg^+_{\omega(i)}(G)}\right).$$

Observe that there is actually a one-to-one correspondence between walks in $\Omega_t(G)$ that ends at v and visits it k times altogether and pairs: a walk that visits v for the first time in some step $s \in \{0,\dots,t\}$ and a walk that starts at v and visits it for the k-th time at step $t-s$. Thus, summing the above equation for all such walks, from Eq. (4.1) we get that

$$p^1_{G,b}(v,t,k) = \sum_{s=0}^{t} p^1_{G,b}(v,s,1) \cdot p^1_{G,\mathbb{1}_v}(v,t-s,k).$$

Multiplying by x^t and summing for all $t \in \mathbb{N}$ we get that

$$\begin{aligned}\sum_{t=0}^{\infty} x^t \cdot p^1_{G,b}(v,t,k) &= \sum_{t=0}^{\infty}\sum_{s=0}^{t} x^s \cdot p^1_{G,b}(v,s,1)\cdot x^{t-s}\cdot p^1_{G,\mathbb{1}_v}(v,t-s,k)\\ &= \sum_{s=0}^{\infty} x^s \cdot p^1_{G,b}(v,s,1)\cdot \sum_{t=s}^{\infty} x^{t-s}\cdot p^1_{G,\mathbb{1}_v}(v,t-s,k)\\ &= \left(\sum_{t=0}^{\infty} x^t \cdot p^1_{G,b}(v,t,1)\right)\cdot\left(\sum_{t=0}^{\infty} x^t \cdot p^1_{G,\mathbb{1}_v}(v,t,k)\right). \qquad (4.5)\end{aligned}$$

Similarly, for every walk $\omega \in \Omega_t(G)$ that starts and ends at v and visits it exactly k times, there exists a step, $s \le t$, in which ω returns to v for the first time, i.e., visits it for the second time (since the first time is at start). Hence, there is a one-to-one correspondence between such walks and pairs: a walk that starts at v and returns to v for the first time in some step $s \in \{0,\dots,t\}$ and a walk that starts at v and visits it for the $k-1$-st time at step $t-s$. Thus, we get

$$p^1_{G,\mathbb{1}_v}(v,t,k) = \sum_{s=0}^{t} p^1_{G,\mathbb{1}_v}(v,s,2)\cdot p^1_{G,\mathbb{1}_v}(v,t-s,k-1).$$

Multiplying by x^t and summing for all $t \in \mathbb{N}$ we get that

$$\begin{aligned}\sum_{t=0}^{\infty} x^t \cdot p^1_{G,\mathbb{1}_v}(v,t,k) &= \sum_{s=0}^{\infty} x^s \cdot p^1_{G,\mathbb{1}_v}(v,s,2)\cdot \sum_{t=s}^{\infty} x^{t-s}\cdot p^1_{G,\mathbb{1}_v}(v,t-s,k-1)\\ &= \left(\sum_{t=0}^{\infty} x^t \cdot p^1_{G,\mathbb{1}_v}(v,t,2)\right)\cdot\left(\sum_{t=0}^{\infty} x^t \cdot p^1_{G,\mathbb{1}_v}(v,t,k-1)\right).\end{aligned}$$

Applying this reasoning $k-1$ number of times yields

$$\sum_{t=0}^{\infty} x^t \cdot p^1_{G,\mathbb{1}_v}(v,t,k) = \left(\sum_{t=0}^{\infty} x^t \cdot p^1_{G,\mathbb{1}_v}(v,t,2)\right)^{k-1}. \qquad (4.6)$$

Now, let us combine Eq. (4.5) and Eq. (4.6) to obtain

$$\sum_{t=0}^{\infty} x^t \cdot p^1_{G,b}(v,t,k) = \sum_{t=0}^{\infty} x^t \cdot p^1_{G,b}(v,t,1)\cdot\left(\sum_{t=0}^{\infty} x^t \cdot p^1_{G,\mathbb{1}_v}(v,t,2)\right)^{k-1}.$$

By summing for all $k \ge 1$, we get that

$$\sum_{t=0}^{\infty} x^t \cdot p^1_{G,b}(v,t) = \sum_{t=0}^{\infty} x^t \cdot p^1_{G,b}(v,t,1)\cdot\sum_{k=0}^{\infty}\left(\sum_{t=0}^{\infty} x^t \cdot p^1_{G,\mathbb{1}_v}(v,t,2)\right)^{k}. \qquad (4.7)$$

Thus, by geometric series transformation, we obtain

$$P^v_{G,b}(x) = \frac{\sum_{t=0}^{\infty} x^t \cdot p^1_{G,b}(v,t,1)}{1-\sum_{t=0}^{\infty} x^t \cdot p^1_{G,\mathbb{1}_v}(v,t,2)}.$$

Therefore, it remains to prove that (I) $P^v_{G^\dagger,b}(x) = \sum_{t=0}^{\infty} x^t \cdot p^1_{G,b}(v,t,1)$ and that (II) $P^v_{G^*,\mathbb{1}_{v'}}(x) = \sum_{t=0}^{\infty} x^t \cdot p^1_{G,\mathbb{1}_v}(v,t,2)$.

(I) Let us start with the former equality. Let us consider an arbitrary walk $\omega \in \Omega_t(G)$ that visits v for the first time at step t. Since ω does not visit v before step t, outgoing edges of this node does not affect the probability associated with walk ω, i.e.,

$$\frac{b(\omega(0))}{b(G)} \cdot \prod_{i=0}^{t-1} \frac{\mu_G(\omega(i), \omega(i+1))}{\deg^+_{\omega(i)}(G)} = \frac{b(\omega(0))}{b(G^\dagger)} \cdot \prod_{i=0}^{t-1} \frac{\mu_{G^\dagger}(\omega(i), \omega(i+1))}{\deg^+_{\omega(i)}(G^\dagger)}.$$

Hence, from Eq. (4.1) we have $p^1_{G,b}(v,t,1) = p^1_{G^\dagger,b}(v,t,1)$, for every $t \in \mathbb{N}$. In graph $G^\dagger$ node v is a sink, thus it cannot be visited more than once. Thus, $p^1_{G^\dagger,b}(v,t) = p^1_{G^\dagger,b}(v,t,1)$ and, by Eq. (4.4), $P^v_{G^\dagger,b}(x) = \sum_{t=0}^{\infty} x^t \cdot p^1_{G,b}(v,t,1)$.

(II) Now, let us show that $P^v_{G^*,\mathbb{1}_{v'}}(x) = \sum_{t=0}^{\infty} x^t \cdot p^1_{G,\mathbb{1}_v}(v,t,2)$. To this end, consider an arbitrary walk $\omega \in \Omega(G)_t$ that starts at v, ends at v, and does not visit v in between. By ω' let us denote the walk on G^* that starts at v' and then follows the same nodes as walk ω, i.e., $\omega'(i) = \omega(i)$, for every $i \in \{1,\dots,t\}$. Since outgoing edges of v' in G^* are the same as outgoing edges of v in G and the outgoing edges of other nodes are not changed, we get that

$$\frac{\mathbb{1}_v(\omega(0))}{\mathbb{1}_v(G)} \cdot \prod_{i=0}^{t-1} \frac{\mu_G(\omega(i), \omega(i+1))}{\deg^+_{\omega(i)}(G)} = \frac{\mathbb{1}_{v'}(\omega'(0))}{\mathbb{1}_{v'}(G^*)} \cdot \prod_{i=0}^{t-1} \frac{\mu_{G^*}(\omega'(i), \omega'(i+1))}{\deg^+_{\omega'(i)}(G^*)}.$$

Moreover, observe that in this way we obtain a one-to-one correspondence between walks in $\Omega(G)_t$ that start at v, end at v, and do not visit v in between and walks in $\Omega(G^*)_t$ that start at v', end at v, and do not visit v in between. Thus,

$$p^1_{G^*,\mathbb{1}_{v'}}(v,t,1) = p^1_{G,\mathbb{1}_v}(v,t,2), \quad \text{for every } t \in \mathbb{N}. \tag{4.8}$$

Observe that since v is a sink in G^* we have that $p^1_{G^*,\mathbb{1}_{v'}}(v,t) = p^1_{G^*,\mathbb{1}_{v'}}(v,t,1)$. Therefore, from Eq. (4.4) we obtain that $P^v_{G^*,\mathbb{1}_{v'}}(x) = \sum_{t=0}^{\infty} x^t \cdot p^1_{G,\mathbb{1}_v}(v,t,2)$.

(h) *(Sink Bound)* Assume that v is a sink. Then, observe that v cannot be visited more than once. Hence, $p^1_{G,b}(v,t) = p^1_{G,b}(v,t,1)$, for every $t \in \mathbb{N}$. Thus, we get $P^v_{G,b}(1) = \sum_{t=0}^{\infty} p^1_{G,b}(v,t,1)$. Observe that $\sum_{t=0}^{\infty} p^1_{G,b}(v,t,1)$ is the probability that node v is visited by the random walk at least once. Hence, it cannot be greater than one.

Now, assume that G is a broken cactus and v is its end. Since the end is a sink, $P^v_{G,b}(1) \le 1$ from the first part of the proof. Hence, it is limiting. In a broken cactus, all nodes are predecessors of its end. Thus, by Proposition 28, all the nodes in G are limiting, i.e., $\sum_{t=0}^{\infty} p^1_{G,b}(u,t) < \infty$, for every $u \in V$. Observe that this implies that also $\lim_{t\to\infty} p^1_{G,b}(u,t) = 0$, for every $u \in V$. Thus, summing for all nodes in a graph, we get $\lim_{t\to\infty} \sum_{u \in V} p^1_{G,b}(u,t) = 0$. Observe that $\sum_{u\in V} p^1_{G,b}(u,t)$ is the probability that the random walk does not stop before step t. If it goes to zero, then with probability 1 the random walk ends at some step. On the other hand, $a = 1$, hence the random walk can stop only if it arrives at a sink. The only sink in a broken cactus is its end, thus v is visited with probability 1, i.e., $\sum_{t=0}^{\infty} p^1_{G,b}(v,t,1) = 1$. Hence, $P^v_{G,b}(1) = 1$. □

In the next four propositions (Propositions 30–33), we consider visit probability generating functions of the ends of broken cacti. Let us begin however, by defining an operation of *appending* broken cacti. Intuitively, appending broken cactus G' to broken cactus G is adding them together and "merging" the end of G with the start of G' (see Fig. 4.4 for an illustration).

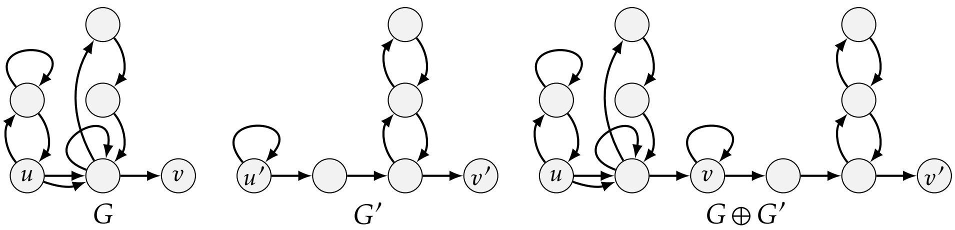


Figure 4.4: Example broken cacti, G, G', and the broken cactus obtained by appending G' to G, i.e., graph $G \oplus G'$.

Definition 6. *For two disjoint broken cacti $G = (V, E)$ and $G' = (V', E')$ with starts u, u' and ends v, v', respectively, appending G' to G results in a graph*

$$G \oplus G' = (V \cup V' \setminus \{u'\}, E \sqcup E' - \Gamma^{\pm}_{u'}(G) \sqcup E^v),$$

where $E^v = \{\!\!\{(v,w) : (u',w) \in \Gamma^+_{u'}(G) \wedge w \neq u'\}\!\!\} \sqcup \{\!\!\{(w,v) : (w,u') \in \Gamma^-_{u'}(G) \wedge w \neq u'\}\!\!\} \sqcup \mu_{G'}(u',u') \cdot \{\!\!\{(v,v)\}\!\!\}$.

In the following proposition we show that appending broken cacti results in multiplying generating functions of their ends.

Proposition 30. *For every two disjoint broken cacti $G = (V, E)$ and $G' = (V', E')$ with starts u, u' and ends v, v', respectively, it holds that*

$$P^{v'}_{G \oplus G', \mathbb{1}_u}(x) = P^{v}_{G, \mathbb{1}_u}(x) \cdot P^{v'}_{G', \mathbb{1}_{u'}}(x).$$

Proof. Consider an arbitrary walk, $\omega \in \Omega_t(G \oplus G')$, that starts at u, i.e., $\omega(0) = u$, and ends at v', i.e., $\omega(t) = v'$. Since node v lies on the main path between node u and v' in graph $G \oplus G'$, there exists a step $s \in \{0, \ldots, t\}$ such that at this step walk ω arrives at node v for the first time, i.e., $\omega(s) = v$ and $\omega(s') \neq v$, for every $s' < s$. In $G \oplus G'$, outgoing edges of nodes from $V \setminus \{v\}$ are the same as in G, outgoing edges of nodes from V' are the same as in G', and outgoing edges of v are the same as outgoing edges of u' in G'. Also, $\mathbb{1}_u(\omega(0))/\mathbb{1}_u(G \oplus G') = \mathbb{1}_u(\omega(0))/\mathbb{1}_u(G) = \mathbb{1}_{u'}(\omega(s))/\mathbb{1}_{u'}(G') = 1$. Thus, we have

$$\frac{\mathbb{1}_u(\omega(0))}{\mathbb{1}_u(G \oplus G')} \cdot \prod_{i=0}^{t-1} \frac{\mu_{G \oplus G'}(\omega(i), \omega(i+1))}{\deg^+_{\omega(i)}(G \oplus G')} =$$
$$\frac{\mathbb{1}_u(\omega(0))}{\mathbb{1}_u(G)} \cdot \prod_{i=0}^{s-1} \frac{\mu_G(\omega(i), \omega(i+1))}{\deg^+_{\omega(i)}(G)} \cdot \frac{\mathbb{1}_{u'}(\omega(s))}{\mathbb{1}_{u'}(G')} \cdot \prod_{i=s}^{t-1} \frac{\mu_G(\omega(i), \omega(i+1))}{\deg^+_{\omega(i)}(G)}.$$

Moreover, observe that there exists a one-to-one correspondence between walks of length t from u to v' in $G \oplus G'$ and the pairs: walk of length s from u to v in G and walk of length $t - s$ from u' to v' in G'. Hence, by Eq. (2.6), summing the above equation for all walks yields

$$p^1_{G \oplus G', \mathbb{1}_u}(v', t) = \sum_{s=0}^{t} p^1_{G, \mathbb{1}_u}(v, s) \cdot p^1_{G', \mathbb{1}_{u'}}(v', t - s).$$

Multiplying by x^t we get $x^t \cdot p^1_{G \oplus G', \mathbb{1}_u}(v', t) = \sum_{s=0}^{t} x^s \cdot p^1_{G, \mathbb{1}_u}(v, s) \cdot x^{s-t} \cdot p^1_{G', \mathbb{1}_{u'}}(v', t - s)$. Thus, by Eq. (4.4), summing for all t yields the thesis. □

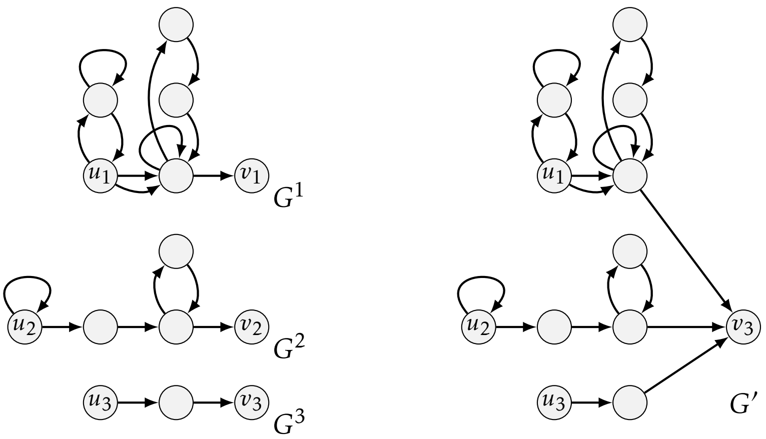


Figure 4.5: An illustration to Proposition 31 for $n = 3$ and example broken cacti: G^1, G^2 and G^3.

We have just considered adding two broken cacti together and "merging" start of one of them with the end of the other. In the next proposition, we consider adding broken cacti and "merging" their ends (see Fig. 4.5).

Proposition 31. *For every collection of pairwise disjoint broken cacti $G^1,\dots,G^n$ with starts $u_1,\dots,u_n$ and ends $v_1,\dots,v_n$, respectively, and constants $\alpha_1,\dots,\alpha_n \geq 0$ it holds that*

$$\sum_{i=1}^{n} \alpha_i \cdot P^{v_i}_{G^i,\mathbb{1}_{u_i}}(x) \Big/ \sum_{i=1}^{n} \alpha_i = P^{v_n}_{G',b'}(x)$$

where $(G',b') = R_{v_1 \to v_n}(\dots(R_{v_{n-1}\to v_n}(G^1 + \dots + G^n, \alpha_1 \cdot \mathbb{1}_{u_1} + \dots + \alpha_n \cdot \mathbb{1}_{u_n}))\dots)$.

Proof. Consider adding together graphs $(G^1, \alpha_1 \cdot \mathbb{1}_{u_1}),\dots,(G^n, \alpha_n \cdot \mathbb{1}_{u_n})$. Formally, let $G'' = G^1 + \dots + G^n$ and $b'' = \alpha_1 \cdot \mathbb{1}_{u_1} + \dots + \alpha_n \cdot \mathbb{1}_{u_n}$. Observe that from Proposition 29e (Graph Addition) we get that $P^{v_i}_{G'',b''}(x) = \alpha_i P^{v_i}_{G^i,\mathbb{1}_{u_i}}(x)/\sum_{j=1}^{n} \alpha_j$. Then, the thesis follows from Proposition 29c (Sink Redirection). □

In the next proposition, we show how we can reduce the level of complexity of a broken cactus by expressing the generating function of its end in terms of the generating functions of the ends of broken cacti with fewer cycles. More in detail, we consider broken cactus, G, such that its start, u, has $n+1$ direct successors: nodes $w_1,\dots,w_n$ and the end of G, node v (this means that the main path of G has length 1). Then, we obtain a set of simpler broken cacti in five steps:

1. We remove node v (graph $G^\ddagger$).
2. We follow Proposition 29g (Cycle Decomposition) and "uncycle" node u, i.e., we change the start of each outgoing edge of u to an additional node, u' (graph $G^\star$).
3. We split node u' into n nodes, $w'_1,\dots,w'_n$, such that each has outgoing edges to only one direct successor, $w_1,\dots,w_n$, respectively (graph G').
4. We split node u into n nodes, $u_1,\dots,u_n$, in such a way that each of them is now successor of only one, respective node from $w'_1,\dots,w'_n$ (graph G'').
5. As a result, we obtain graph with n disjoint connected components—we treat each of them as a separate graph (graphs $G^1,\dots,G^n$). As we prove, they are also broken cacti.

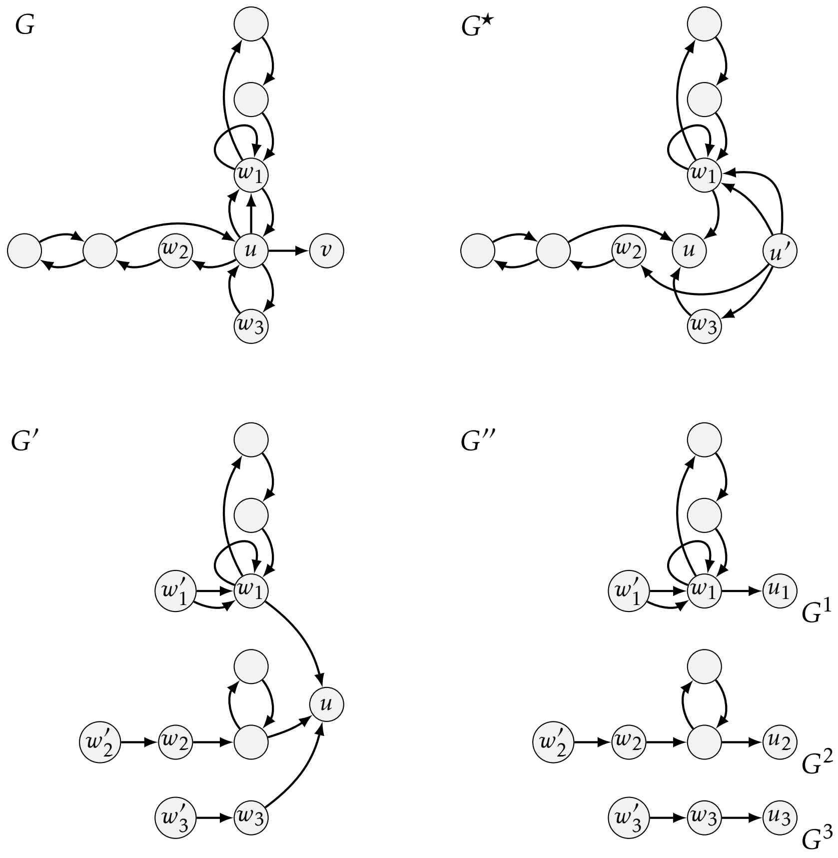


Figure 4.6: An illustration to Proposition 32 for an example broken cactus, G.

See Fig. 4.6 for an illustration. We prove that the generating function of the end of the original broken cactus G, can be expressed in terms of generating function of the ends of the obtained graphs.

Proposition 32. *For every broken cactus $G = (V, E)$ with start u and end v such that $S_u^1(G) = \{v, w_1, \dots, w_n\}$, nodes $w'_1, \dots, w'_n, u_1, \dots, u_n \notin V$, and graphs*

$$G^\ddagger = (V^\ddagger, E^\ddagger) = (V \setminus \{v\}, E - \Gamma_v^-(G))$$

$$G^\star = (V^\star, E^\star) = (V^\ddagger \cup \{u'\}, E^\ddagger - \Gamma_u^+(G^\ddagger) \sqcup \{\!\{(u', t) : (u, t) \in \Gamma_u^+(G^\ddagger)\}\!\}),$$

$$G' = (V', E') = (V^\star \setminus \{u'\} \cup \{w'_1, \dots, w'_n\}, E^\star - \Gamma_{u'}^+(G^\star) \sqcup \bigsqcup_{i=1}^{n} \mu_G(u, w_i) \cdot \{\!\{(w'_i, w_i)\}\!\}),$$

$$G'' = (V'', E'') = (V' \setminus \{u\} \cup \{u_1, \dots, u_n\}, E' - \Gamma_u^-(G') \sqcup \bigsqcup_{i=1}^{n} \{\!\{(s, u_i) : (s, u) \in \Gamma_u^-(G'), s \in S_{w'_i}(G') \cup \{w'_i\}\}\!\}),$$

$$G^i = (V^i, E^i) = (\{w'_i\} \cup S_{w'_i}(G''), \{\!\{(s, t) : (s, t) \in E'', t \in S_{w'_i}(G'')\}\!\}) \quad \textit{for every } i \in \{1, \dots, n\},$$

it holds that $G^1, \dots, G^n$ are disjoint broken cacti with less edges than G and

$$P^v_{G, \mathbb{1}_u}(x) = \frac{\mu_G(u, v)}{\deg^+_u(G)} \cdot x \Bigg/ \left(1 - \sum_{i=1}^{n} \frac{\mu_G(u, w_i)}{\deg^+_u(G)} P^{u_i}_{G^i, \mathbb{1}_{w'_i}}(x)\right). \tag{4.9}$$

Proof. Let us consider an arbitrary broken cactus G with start u and end v such that $S_u^1(G) = \{v, w_1, \dots, w_n\}$. First, we prove that graphs $G^1, \dots, G^n$ are disjoint, i.e., for every $i, j \in \{1, \dots, n\}$ such that $i \neq j$, it holds that $V^i \cap V^j = \emptyset$. To this end, we first show that any pair of distinct nodes from $w'_1,\dots,w'_n$ have exactly one shared successor in graph G' and it is node u, i.e.,

$$S_{w'_i}(G') \cap S_{w'_j}(G') = \{u\}, \quad \text{for every } i,j \in \{1,\dots,n\} \text{ such that } i \neq j. \tag{4.10}$$

Assume otherwise, i.e., there exist $i,j \in \{1,\dots,n\}$ such that $i \neq j$ and node $s \neq u$ such that $s \in S_{w'_i}(G') \cap S_{w'_j}(G')$. This means that in G' there are paths $(\pi(0),\dots,\pi(l))$ and $(\rho(0),\dots,\rho(k))$ such that $\pi(0) = w'_i$, $\rho(0) = w'_j$, and $\pi(l) = \rho(k) = s$. Without loss of generality let us assume that $k \geq l$. Since the only direct successors of w'_i and w'_j are w_i and w_j, respectively, we know that $\pi(1) = w_i$ and $\rho(1) = w_j$. Thus, in $G^\star$ there are paths $(u',\pi(1),\dots,\pi(l))$ and $(u',\rho(1),\dots,\rho(k))$ and they are disjoint. Hence, in $G^\ddagger$ we have disjoint paths $(u,\pi(1),\dots,\pi(l))$ and $(u,\rho(1),\dots,\rho(k))$ and they are also disjoint paths in G. Thus, they are also disjoint in graph $(V, E \sqcup \{\!\{(v,u)\}\!\})$. However, since G is a broken cactus, that graph is a cactus and we arrive at a contradiction, which means that Claim (4.10) holds. Observe that for every $i \in \{1,\dots,n\}$ it holds that $S_{w'_i}(G'') = S_{w'_i}(G') \setminus \{u\} \cup \{u_i\}$. Thus, from Claim (4.10) we get that sets of successors $S_{w'_1}(G''),\dots,S_{w'_n}(G'')$ are pairwise disjoint. Moreover, nodes $w'_1,\dots,w'_n$ are sources, thus $w'_i \notin S_{w'_j}(G')$, for every $i,j \in \{1,\dots,n\}$. Hence, we get that $V^1,\dots,V^n$ are pairwise disjoint as well.

Now, let us fix $i \in \{1,\dots,n\}$ and prove that G^i is a broken cactus. Assume otherwise. Then, in graph $(V^i, E^i \sqcup \{\!\{(u_i,w'_i)\}\!\})$, there are two distinct paths $(\pi(0),\dots,\pi(l))$ and $(\rho(0),\dots,\rho(k))$ such that $\pi(0) = \rho(0)$ and $\pi(l) = \rho(k)$. In what follows, we will show that this implies that in cactus $(V, E \sqcup \{\!\{(v,u)\}\!\})$ there are two distinct paths between the same pair of nodes as well, which is a contradiction. To this end, first observe that $(\pi(0),\dots,\pi(l))$ and $(\rho(0),\dots,\rho(k))$ are also paths in graph $(V'', E'' \sqcup \{\!\{(u_i,w'_i)\}\!\})$. Now, for every $i \in \{0,\dots,k\}$, let us define $\pi'(i) = \pi(i)$, if $\pi(i) \neq u_i$, and $\pi'(i) = u$, otherwise, and in the same way $\rho'(i) = \rho(i)$, if $\rho(i) \neq u_i$, and $\rho'(i) = u$, otherwise. Then, $(\pi'(0),\dots,\pi'(l))$ and $(\rho'(0),\dots,\rho'(k))$ are two distinct paths in graph $(V', E' \sqcup \{\!\{(u,w'_i)\}\!\})$. Next, analogously, for every $i \in \{0,\dots,k\}$, let us denote $\pi^\star(i) = \pi'(i)$, if $\pi'(i) \neq w'_i$, and $\pi^\star(i) = u'$, otherwise, and $\rho^\star(i) = \rho'(i)$, if $\rho'(i) \neq w'_i$, and $\rho^\star(i) = u'$, otherwise. Then, sequences $(\pi^\star(0),\dots,\pi^\star(l))$ and $(\rho^\star(0),\dots,\rho^\star(k))$ are two distinct paths in graph $(V^\star, E^\star \sqcup \{\!\{(u,u')\}\!\})$. Finally, for every $i \in \{0,\dots,k\}$, define $\pi^\circ(i) = \pi^\star(i)$ if $\pi^\star(i) \neq u'$, and $\pi^\circ(i) = u$, otherwise, and $\rho^\circ(i) = \rho^\star(i)$, if $\rho^\star(i) \neq u'$, and $\rho^\circ(i) = u$, otherwise. Then, sequences $(\pi^\circ(0),\dots,\pi^\circ(l))$ and $(\rho^\circ(0),\dots,\rho^\circ(k))$ are two distinct paths in $G^\ddagger$. Hence, they are also two distinct paths in $(V, E \sqcup \{\!\{(v,u)\}\!\})$. Observe that still $\pi^\circ(0) = \rho^\circ(0)$ and $\pi^\circ(l) = \rho^\circ(k)$. Thus, we arrive at a contradiction.

Next, let us show that for every $i \in \{1,\dots,n\}$, graph G^i has less edges than G, i.e., $|E^i| < |E|$. To this end, observe that in the first step of the construction we remove all incoming edges of node v, thus $|E^\ddagger| < |E|$. Then, in each next step, the total number of edges stays the same. Hence, $\sum_{i=1}^n |E^i| = |E^\ddagger|$. Thus, indeed, $|E^i| < |E|$.

Finally, in the remainder of the proof, we will prove Eq. (4.9). From Proposition 29b (Recursive Equation) we have that $P^v_{G,\mathbb{1}_u}(x) = x \cdot \mu_G(u,v)/\deg^+_u(G) \cdot P^u_{G,\mathbb{1}_u}(x)$. Hence, it suffices to show that

$$P^u_{G,\mathbb{1}_u}(x) = \left(1 - \sum_{i=1}^n \frac{\mu_G(u,w_i)}{\deg^+_u(G)} \cdot P^{u_i}_{G^i,\mathbb{1}_{w'_i}}(x)\right)^{-1}.$$

To this end, we follow Proposition 29g (Cycle Decomposition) and define graph $G^\dagger$, which is graph G with outgoing edges of u removed, and graph G^*, which is obtained from G by "uncycling" node u (the difference between G^* and $G^\star$ is that we do not remove node v in the former). Formally, let $G^\dagger = (V, E - \Gamma_u^+(G))$ and let $G^* = (V \cup \{u'\}, E - \Gamma_u^+(G) \sqcup \{\!\{(u', t) : (u, t) \in \Gamma_u^+(G)\}\!\})$. Then, from Proposition 29g (Cycle Decomposition) we have $P^u_{G,\mathbb{1}_u}(x) = P^u_{G^\dagger,\mathbb{1}_u}(x)/(1 - P^u_{G^*,\mathbb{1}_{u'}}(x))$. Since we consider unit node weights centered on u, we get that $p^1_{G^\dagger,\mathbb{1}_u}(u, 0) = 1$. Moreover, in graph $G^\dagger$ node u is a sink. Thus, by Proposition 29h (Sink Bound), $P^u_{G^\dagger,\mathbb{1}_u}(x) = 1$. Hence, in order to prove Eq. (4.9), it remains to prove that

$$\sum_{i=1}^{n} \frac{\mu_G(u, w_i)}{\deg_u^+(G)} \cdot P^u_{G^i,\mathbb{1}_{w_i'}}(x) = P^u_{G^*,\mathbb{1}_{u'}}(x).$$

To this end, we will prove two equations that combined yield the above equation. These will be the following: (I) $\sum_{i=1}^n \frac{\mu_G(u,w_i)}{\deg_u^+(G)} \cdot P^u_{G^i,\mathbb{1}_{w_i'}}(x) = P^u_{G',b}(x) \cdot \sum_{i=1}^n \frac{\mu_G(u,w_i)}{\deg_u^+(G)}$ and (II) $P^u_{G^*,\mathbb{1}_{u'}}(x) = P^u_{G',b}(x) \cdot \sum_{i=1}^n \frac{\mu_G(u,w_i)}{\deg_u^+(G)}$, where $b = \sum_{i=1}^n \frac{\mu_G(u,w_i)}{\deg_u^+(G)} \mathbb{1}_{w_i'}$.

(I) First, let us follow Proposition 31 and consider graph G''' that is obtained from broken cacti $G^1, \dots, G^n$ by adding them together and redirecting their ends to u_n. Also, let us take $\alpha_i = \mu_G(u, w_i)/\deg_u^+(G)$ for every $i \in \{1, \dots, n\}$. Formally, let

$$(G''', b''') = R_{u_1 \to u_n}(\dots(R_{u_{n-1}} \to R_{u_n}(G^1 + \dots + G^n, \alpha_1 \cdot \mathbb{1}_{w_1'} + \dots + \alpha_n \cdot \mathbb{1}_{w_n'}))\dots).$$

Then, observe that the only difference between graph G''' and graph G' is the fact that node u_n in G''' is labeled as u in G'. Observe also that $b''' = \sum_{i=0}^n \frac{\mu_G(u,w_i)}{\deg_u^+(G)} \mathbb{1}_{w_i'} = b$. Thus, by Proposition 31 and Proposition 29a (Isomorphism), indeed

$$\sum_{i=1}^{n} \frac{\mu_G(u, w_i)}{\deg_u^+(G)} \cdot P^u_{G^i,\mathbb{1}_{w_i'}}(x) = P^u_{G',b}(x) \cdot \sum_{i=1}^{n} \frac{\mu_G(u, w_i)}{\deg_u^+(G)},$$

(II) Next, let us prove that $P^u_{G^*,\mathbb{1}_{u'}}(x) = P^u_{G',b}(x) \cdot \sum_{i=1}^n \mu_G(u, w_i)/\deg_u^+(G)$. To this end, let us consider an arbitrary walk of length t in graph G^* that starts at u' and ends at u, i.e., $\omega \in \Omega_t(G^*)$ such that $\omega(0) = u'$ and $\omega(t) = u$. Since v is a sink in G^*, walk ω cannot visit v, because it would not be possible to end at u later on. Thus, as direct successors of node u' in G^* are $\{v, w_1, \dots, w_n\}$, it holds that there exists $k \in \{1, \dots, n\}$ such that $\omega(1) = w_k$. Now, let us define sequence ω' such that $\omega'(0) = w_k'$ and $\omega'(i) = \omega(i)$, for every $i \in \{1, \dots, t\}$. Then, observe that ω' is a walk on G'. In what follows, we will show that probability that the random walk follows ω in G^* is equal to $\sum_{i=1}^n \frac{\mu_G(u,w_i)}{\deg_u^+(G)}$ times the probability that the random walk follows ω' in G'. Observe that since outgoing edges of nodes other than u' do not change between graph G^* and G', for steps from 1 to t we get that

$$\prod_{i=1}^{t-1} \frac{\mu_{G^*}(\omega(i), \omega(i+1))}{\deg^+_{\omega(i)}(G^*)} = \prod_{i=1}^{t-1} \frac{\mu_{G'}(\omega'(i), \omega'(i+1))}{\deg^+_{\omega'(i)}(G')}. \tag{4.11}$$

On the other hand, we have that

$$\frac{\mu_G(u, w_k)/\deg_u^+(G)}{\sum_{i=1}^n \mu_G(u, w_i)/\deg_u^+(G)} = \frac{b(\omega'(0))}{b(G')}.$$

Also, observe that $\mathbb{1}_{u'}(\omega(0))/\mathbb{1}_{u'}(G^*) = 1$ and since w_k is the only direct successor of w_k' in G', $\mu_{G'}(\omega'(0))/\deg^+_{\omega'(0)}(G') = 1$ as well. Thus,

$$\frac{\mathbb{1}_{u'}(\omega(0))}{\mathbb{1}_{u'}(G^*)} \cdot \frac{\mu_G(u, w_k)}{\deg_u^+(G)} = \frac{\mu_{G'}(\omega'(0))}{\deg^+_{\omega'(0)}(G')} \cdot \frac{b(\omega'(0))}{b(G')} \cdot \sum_{i=1}^{n} \frac{\mu_G(u, w_i)}{\deg_u^+(G)}.$$

Finally, observe that $\mu_G(u,w_k) = \mu_{G^*}(u',w_k)$ and $\deg^+_u(G) = \deg^+_{u'}(G)$. Hence, combining the above equation with Eq. (4.11) we get that

$$\frac{\mathbb{1}_{u'}(\omega(0))}{\mathbb{1}_{u'}(G^*)}\cdot\prod_{i=0}^{t-1}\frac{\mu_{G^*}(\omega(i),\omega(i+1))}{\deg^+_{\omega(i)}(G^*)} = \sum_{i=1}^{n}\frac{\mu_G(u,w_i)}{\deg^+_u(G)}\cdot\left(\frac{b(\omega'(0))}{b(G')}\cdot\prod_{i=0}^{t-1}\frac{\mu_{G'}(\omega'(i),\omega'(i+1))}{\deg^+_{\omega'(i)}(G')}\right).$$

Observe that the relation between ω and ω' can be extended to a one-to-one correspondence between walks on G^* that starts at u' and ends at u and walks on G' that starts at a node from $w'_1,\dots,w'_n$ and ends at u. Therefore, from Eq. (4.1) and Eq. (4.4) we get that

$$P^u_{G^*,\mathbb{1}_{u'}}(x) = P^u_{G',b}(x)\cdot\sum_{i=1}^{n}\frac{\mu_G(u,w_i)}{\deg^+_u(G)},$$

which concludes the proof. □

In the final proposition of this part of the proof, we show that for every graph and its limiting node, its visit probability generating function is an average of the visit probability generating functions of the ends of some broken cacti.

Proposition 33. *For every graph $G=(V,E)$, node weights b such that $b(G)>0$, and limiting node $v\in V$, there exist constants $\alpha_1,\alpha_2,\dots,\alpha_n \geq 0$ and a collection of broken cacti, $G^1,\dots,G^n$, with starts $u_1,\dots,u_n$ and ends $v_1,\dots,v_n$, respectively, such that*

$$P^v_{G,b}(x) = \sum_{i=1}^{n}\alpha_i\cdot P^{v_i}_{G^i,\mathbb{1}_{u_i}}(x).$$

Proof. Let us prove the thesis by induction on the number of incoming edges of v and its predecessors, i.e., $\sum_{u\in P_v(G)\cup\{v\}}|\Gamma^-_u(G)|$. To this end, we will strengthen the induction hypothesis with three additional implications:

(a) If all nodes in G have rational weights, i.e., $b(u)\in\mathbb{Q}$, for every node $u\in V$, then there exist such a collection of broken cacti $G^1,\dots,G^n$ and constants $\alpha_1,\dots,\alpha_n$ that satisfy the thesis and $\alpha_1,\dots,\alpha_n$ are rational as well.

(b) If, in graph G, every node with positive weight, i.e., $u\in V$ such that $b(u)>0$, is a source, then there exists such a collection of broken cacti, $G^1,\dots,G^n$, with starts $u_1,\dots,u_n$ and ends $v_1,\dots,v_n$, respectively, that satisfies the thesis and nodes $u_1,\dots,u_n$ in graphs $G^1,\dots,G^n$ are sources as well.

(c) If $b(v)=0$, then there exists such a collection of broken cacti, $G^1,\dots,G^n$, with starts $u_1,\dots,u_n$ and ends $v_1,\dots,v_n$, respectively, that satisfies the thesis and for each $i\in\{1,\dots,n\}$ we have that $u_i\neq v_i$, i.e., the main path of each broken cacti has length of at least 1.

Let us begin with the induction basis. If there are no edges incoming to v and it predecessors, it means that v is a source. If $b(v)=0$, then let us take broken cactus $G_1=(\{s,t\},\{\!\{(s,t)\}\!\})$ for distinct $s,t\notin V$ and constant $\alpha_1=0$. Observe that indeed $P^v_{G,b}(x)=0=\alpha_1\cdot P^1_{G^1,\mathbb{1}_s}(x)$. Moreover, α_1 is rational (thus additional implication (a) holds), s is a source (thus (b) holds), and $s\neq t$ (thus (c) holds). Hence, let us assume that $b(v)>0$. Then, observe that $p^1_{G,b}(v,0)=b(v)/b(G)$ and $p^1_{G,b}(v,t)=0$, for every $t>0$. Hence, taking constant $\alpha_1=b(v)/b(G)$, and one one-node broken cactus $G^1=(\{v_1\},\emptyset)$ we get that indeed

$$P^v_{G,b}(x)=\alpha_1\cdot P^1_{G^1,\mathbb{1}_{v_1}}(x).$$

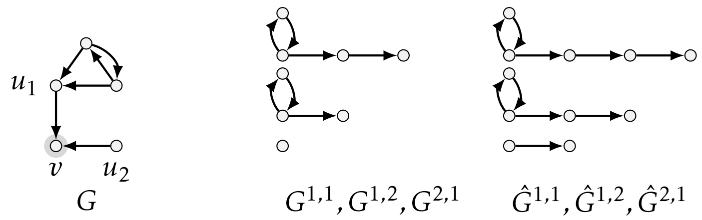


Figure 4.7: An illustration to the case (I) of the proof of Proposition 33 for an example graph, G.

Moreover, if all weights of nodes are rational, then constant $\alpha_1 = b(v)/b(G)$ is rational as well, hence additional implication (a) is satisfied. Also, in graph G^1 node v_1 is a source, thus additional implication (b) is satisfied as well. Finally, since $b(v) > 0$, additional implication (c) is also satisfied.

Therefore, let us focus on a case in which $\sum_{u \in P_v(G) \cup \{v\}} |\Gamma_u^-(G)| > 1$, i.e., there is at least one incoming edge of v. First, observe that if the probability of visiting v is zero, i.e., $\sum_{t=0}^{\infty} p_{G,b}^1(v,t) = 0$, then we can take broken cactus $G_1 = (\{s,t\}, \{\!\{(s,t)\}\!\})$ for distinct $s,t \notin V$ and constant $\alpha_1 = 0$. Again, $P_{G,b}^v(x) = 0 = \alpha_1 \cdot P_{G^1,\mathbb{1}_s}^1(x)$ and α_1 is rational (thus (a) holds), s is a source (thus (b) holds), and $s \neq t$ (thus (c) holds). Thus, in the remainder of the proof, let us assume that there is non-zero probability of visiting node v in the random walk. Then, we will consider two cases: the first one (I) in which node v is a sink, and the second one (II) in which v has outgoing edges.

(I) Let us begin with the case in which v is a sink. Then, let us denote the direct predecessors of v by $\{u_1, \dots, u_m\} = P_v^1(G)$ and fix arbitrary $i \in \{1, \dots, m\}$. Observe that u_i has less incoming edges to it and its predecessors then v, i.e., $\sum_{w \in P_v(G) \cup \{v\}} |\Gamma_w^-(G)| > \sum_{w \in P_{u_i}(G) \cup \{u_i\}} |\Gamma_w^-(G)|$. This is because each predecessor of u_i is a predecessor of v, but on the right hand side of the inequality we do not count incoming edges of v. Also, since v is limiting, from Proposition 28 we know that node u_i is limiting as well. Hence, from the inductive assumption we obtain that there exist constants $\alpha_{i,1}, \dots, \alpha_{i,n_i} \geq 0$ and such a collection of broken cacti $G^{i,1} = (V^{i,1}, E^{i,1}), \dots, G^{i,n_i} = (V^{i,n_i}, E^{i,n_i})$ with starts $u_{i,1}, u_{i,2}, \dots, u_{i,n_i}$ and ends $v_{i,1}, v_{i,2}, \dots, v_{i,n_i}$, respectively, that

$$P_{G,b}^{u_i}(x) = \sum_{j=1}^{n_i} \alpha_{i,j} \cdot P_{G^{i,j},\mathbb{1}_{u_{i,j}}}^{v_{i,j}}(x). \tag{4.12}$$

Now, for every $i \in \{1, \dots, m\}$ and every $j \in \{1, \dots, n_i\}$ let us modify graph $G^{i,j}$ by adding a new node $\hat{v}_{i,j}$ and edge $(v_{i,j}, \hat{v}_{i,j})$. In other words, we append a simple broken cactus $(\{s, \hat{v}_{i,j}\}, \{\!\{(s, \hat{v}_{i,j})\}\!\})$, for a new node s, to each broken cactus $G^{i,j}$. See Fig. 4.7 for an illustration. Formally, let $\hat{G}^{i,j} = G^{i,j} \oplus (\{s, \hat{v}_{i,j}\}, \{\!\{(s, \hat{v}_{i,j})\}\!\})$. In what follows, we will prove that graphs $\hat{G}^{i,j}$ with constants $(\mu_G(u_i, v)/\deg_{u_i}^+(G))\alpha_{i,j}$, for every $i \in \{1, \dots, m\}$ and every $j \in \{1, \dots, n_i\}$, and, if $b(v) > 0$, also graph $G^0 = (\{v_0\}, \emptyset)$ with constant $b(v)/b(G)$, constitute a collection such that

$$P_{G,b}^v(x) = \frac{b(v)}{b(G)} P_{G^0,\mathbb{1}_{v_0}}^{v_0}(x) + \sum_{i=1}^{m} \sum_{j=1}^{n_i} \frac{\mu_G(u_i, v)}{\deg_{u_i}^+(G)} \alpha_{i,j} \cdot P_{\hat{G}^{i,j},\mathbb{1}_{u_{i,j}}}^{\hat{v}_{i,j}}(x). \tag{4.13}$$

To this end, observe that for every $i \in \{1, \dots, m\}$ and $j \in \{1, \dots, n_j\}$, from Proposition 30 we know that generating function of the end of broken cactus $\hat{G}^{i,j}$ is a

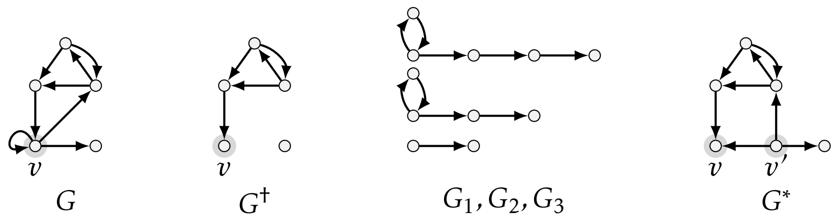


Figure 4.8: An illustration to the first part of case (II) of the proof of Proposition 33 for an example graph, G.

product of generating functions of the ends in broken cactus $G^{i,j}$ and broken cactus $(\{s, \hat{v}_{i,j}\}, \{\!\{(s, \hat{v}_{i,j})\}\!\})$. Generating function of the latter is simply x, therefore we obtain that $P^{\hat{v}_{i,j}}_{\hat{G}^{i,j},\mathbb{1}_{u_{i,j}}}(x) = x \cdot P^{v_{i,j}}_{G^{i,j},\mathbb{1}_{u_{i,j}}}(x)$. Inserting it into Eq. (4.13) we can transform it into

$$P^{v}_{G,b}(x) = \frac{b(v)}{b(G)} P^{v_0}_{G^0,\mathbb{1}_{v_0}}(x) + \sum_{i=1}^{m} x \cdot \frac{\mu_G(u_i, v)}{\deg^+_{u_i}(G)} \sum_{j=1}^{n_i} \alpha_{i,j} \cdot P^{v_{i,j}}_{G^{i,j},\mathbb{1}_{u_{i,j}}}(x).$$

Furthermore, observe that $P^{v_0}_{G^0,\mathbb{1}_{v_0}}(x) = 1$. From this and Eq. (4.12) we get

$$P^{v}_{G,b}(x) = \frac{b(v)}{b(G)} + \sum_{i=1}^{m} x \cdot \frac{\mu_G(u_i, v)}{\deg^+_{u_i}(G)} P^{u_i}_{G,b}(x).$$

Hence, Eq. (4.13) follows from Proposition 29b (Recursive Equation).

It remains to prove additional implications. For (a), observe that if all node weights are rational, then from the inductive assumption we know that constants $\alpha_{i,j}$, for every $i \in \{1, \dots, m\}$ and $j \in \{1, \dots, n_i\}$, are rational as well. Hence, constants $\mu_G(u_i, v)/\deg^+_{u_i}(G) \cdot \alpha_{i,j}$ are rational too. Also, $b(v)/b(G)$ is rational, thus implication (a) follows. For (b), observe that if all nodes with positive weights in (G, b) are sources, then from the inductive assumption the start of every broken cactus $G^{i,j}$ is a source. As a result, the start of every broken cactus $\hat{G}^{i,j}$ is a source as well. Since G_0 also starts with a source, implication (b) follows. Finally, for (c), observe that if $b(v) > 0$, then we do not include graph G^0 in the collection. Moreover, to all other graphs we have appended a broken cactus with the main path of length 1. Thus, the main path of each broken cactus in the collection has length at least 1.

(II) Now, let us assume that v is not a sink. Then, from Proposition 29g (Cycle Decomposition) we know that the generating function of v is equal to the fraction

$$P^{v}_{G,b}(x) = P^{v}_{G^\dagger,b}(x)/(1 - P^{v}_{G^*,\mathbb{1}_{v'}}(x)) \tag{4.14}$$

where $G^\dagger = (V, E - \Gamma^+_v(G))$ and $G^* = (V \cup \{v'\}, E - \Gamma^+_v(G) \sqcup \{\!\{(v', u) : (v, u) \in \Gamma^+_v(G)\}\!\})$ (see Fig. 4.8). Building upon this, we will construct two collections of broken cacti with constants: one in which the averaged generating function of the ends is equal to $P^{v}_{G^\dagger,b}(x)$, and one in which it is equal to $P^{v}_{G^\dagger,b}(x) \cdot P^{v}_{G^*,\mathbb{1}_{v'}}(x)/(1 - P^{v}_{G^*,\mathbb{1}_{v'}}(x))$. From Eq. (4.14) we get that added together they will yield averaged generating function equal to $P^{v}_{G,b}(x)$.

To this end, observe that in graph $G^\dagger$ node v is a sink and the number of edges incoming to v and its predecessors in $G^\dagger$ is at most the same as in graph G, i.e., $\sum_{w \in P_v(G^\dagger) \cup \{v\}} |\Gamma^-_w(G^\dagger)| \le \sum_{w \in P_v(G) \cup \{v\}} |\Gamma^-_w(G)|$. Hence, from case (I) of the proof we know that there exists a collection of broken cacti $G_1, \dots, G_m$ with starts $s_1, \dots, s_m$

and ends $w_1, \dots, w_m$ and constants $\beta_1, \dots, \beta_m$ such that the averaged generating function of their ends is equal to $P^{v}_{G^{\dagger},b}(x)$, i.e.,

$$P^{v}_{G^{\dagger},b}(x) = \sum_{j=1}^{m} \beta_j P^{w_j}_{G_j,\mathbb{1}_{s_j}}(x). \tag{4.15}$$

We will come back to this collection at the end of the proof. Now, let us turn our attention to graph G^*.

Again, v in graph G^* is a sink and there are at most as many edges incoming to v and its predecessors as in G, i.e., $\sum_{w \in P_v(G^*) \cup \{v\}} |\Gamma^-_w(G^*)| \le \sum_{w \in P_v(G) \cup \{v\}} |\Gamma^-_w(G)|$. Thus, from case (I) of a proof, there exist constants $\alpha_1, \dots, \alpha_n \ge 0$ and a collection of broken cacti $G^1 = (V^1, E^1), \dots, G^n = (V^n, E^n)$ with starts $u_1, \dots, u_n$ and ends $v_1, \dots, v_n$, respectively, such that

$$P^{v}_{G^*,\mathbb{1}_{v'}}(x) = \sum_{i=1}^{n} \alpha_i \cdot P^{v_i}_{G^i,\mathbb{1}_{u_i}}(x). \tag{4.16}$$

Observe that since graphs $G^1, \dots, G^n$ are broken cacti, from Proposition 29h (Sink Bound) we get that $P^{v_i}_{G^i,\mathbb{1}_{u_i}}(1) = 1$. Thus, by Eq. (4.16), $P^{v}_{G^*,\mathbb{1}_{v'}}(1) = \alpha_1 + \dots + \alpha_n$. Now, let us show that the fact that v is limiting implies that $P^{v}_{G^*,\mathbb{1}_{v'}}(1) < 1$. Assume otherwise. Then, since v is a sink in G^*, from Proposition 29h (Sink Bound) we have that $P^{v}_{G^*,\mathbb{1}_{v'}}(1) = 1$. Thus, $\sum_{t=0}^{\infty} p^1_{G^*,\mathbb{1}_{v'}}(v,t) = 1$. Since from the fact that v is a sink we have $p^1_{G^*,\mathbb{1}_{v'}}(v,t) = p^1_{G^*,\mathbb{1}_{v'}}(v,t,1)$, by Eq. (4.8), this means that also $\sum_{t=0}^{\infty} p^1_{G,\mathbb{1}_v}(v,t,2) = 1$. However, by Eq. (4.7) for $x = 1$, this implies that

$$\sum_{t=0}^{\infty} p^1_{G,\mathbb{1}_v}(v,t) = \sum_{t=0}^{\infty} p^1_{G,b}(v,t,1) \cdot \sum_{k=0}^{\infty} 1.$$

Thus, either $p^1_{G,b}(v,t,1) = 0$, for every $t \in \mathbb{N}$, or v is not limiting. Since we assumed that there is non-zero probability of visiting v, we arrive at a contradiction.

Thus, we know that $\alpha_1 + \dots + \alpha_n < 1$ and we can denote $\alpha_0 = 1 - (\alpha_1 + \dots + \alpha_n)$. Also, by additional implication (a), since weight $\mathbb{1}_{v'}$ is rational for every node, constants $\alpha_1, \dots, \alpha_n$ are rational as well. Hence, α_0 is also rational. Thus, there exists constant $q \in \mathbb{N}$ such that $q \cdot \alpha_0, \dots, q \cdot \alpha_n \in \mathbb{N}$. Next, observe that from additional implication (b) and the fact that v' in graph G^* is a source we know that nodes $u_1, \dots, u_n$ in graphs $G^1, \dots, G^n$ are sources themselves. Also, from additional implication (c) we know that in broken cacti $G^1, \dots, G^n$ the starts are distinct from the ends. Observe that if the start of a broken cactus is a source and it is not the end, then it has exactly one direct successor—the next node on the main path. For every $i \in \{1, \dots, n\}$ let us denote this direct successor of u_i by u'_i. Hence, by Proposition 29f (Edge Multiplication), without loss of generality, we can also assume that there is exactly $q \cdot \alpha_i$ edges from u_i to u'_i in graph G^i. Moreover, by Proposition 29a (Isomorphism), without loss of generality, we can assume that graphs $G^1, \dots, G^n$ are pairwise disjoint. In what follows, using Proposition 32, we will construct broken cactus G° such that the generating function of its end can be expressed in terms of the generating functions of the ends of graphs $G^1, \dots, G^n$.

To this end, we follow the same five steps of graph operations as in Proposition 32, but in the reversed order. Specifically,

1. we add graphs $G^1, \dots, G^n$ together (graph G''),
2. we redirect ends $v_1, \dots, v_n$ into one node, v^* (graph G'),

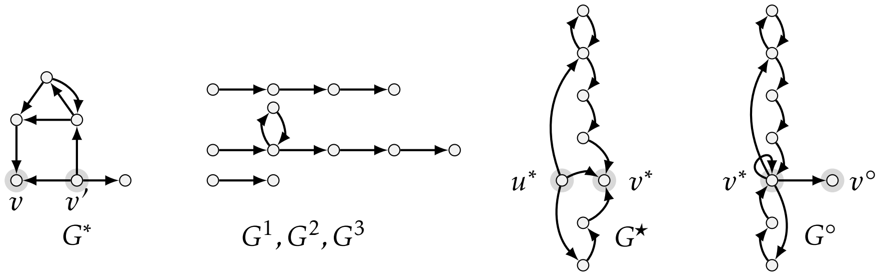


Figure 4.9: An illustration to the second part of case (II) of the proof of Proposition 33 for graph G from Fig. 4.8.

3. we combine starts $u_1,\dots,u_n$ into one node, u^* (graph $G^\star$),
4. we merge node u^* with node v^* (graph $G^\ddagger$), and
5. we add node v° along with $q\alpha_0$ edges from v^* to v° (graph G°).

See Fig. 4.9 for an illustration.

Formally, let us take $(G'', b'') = (G^1 + \dots + G^n, \alpha_1 \cdot \mathbb{1}_{u_1} + \dots + \alpha_n \cdot \mathbb{1}_{u_n})$ and also $(G', b') = R_{v_1 \to v^*}(\dots(R_{v_n \to v^*}(G'' + G^v, b'' + b^v))\dots)$, where $(G^v, b^v) = ((\{v\}, \emptyset), [0])$. Let us denote $G' = (V', E')$. Then, we remove nodes $u_1,\dots,u_n$ from graph G' and instead add node u^* and $q \cdot \alpha_i$ edges from node u^* to node u_i', for every $i \in \{1,\dots,n\}$. Formally, let $G^\star = (V^\star, E^\star)$, where $V^\star = V' \setminus \{u_1,\dots,u_n\} \cup \{u^*\}$ and

$$E^\star = E' - \bigsqcup_{i=1}^{n} \Gamma^+_{u_i}(G') \sqcup \bigsqcup_{i=1}^{n} q\alpha_i \cdot \{\!\{(u^*, u_i')\}\!\}.$$

Next, we merge node u^* with node v^*. Formally, let

$$G^\ddagger = (V^\ddagger, E^\ddagger) = (V^\star \setminus \{u^*\}, E^\star - \Gamma^+_{u^*}(G^\star) \sqcup \{\!\{(v^*, t) : (u^*, t) \in \Gamma^+_{u^*}(G^\star)\}\!\}).$$

Finally, let us add node v° and $q\alpha_0$ edges from v^* to v°, i.e., let

$$G^\circ = (V^\circ, E^\circ) = (V^\ddagger \cup \{v^\circ\}, E^\ddagger \sqcup q\alpha_0 \cdot \{\!\{(v^*, v^\circ)\}\!\}).$$

Observe that in this way graph G° is indeed a broken cactus graph, with start v^* and end v°. Moreover, observe that broken cacti $G^1,\dots,G^n$ are the result of the graph operations in the construction from Proposition 32 for graph G° (the operation in each step of the construction here is the opposite to the operation taken in the corresponding step of the construction from Proposition 32 and they are in the reversed order). Thus, from Proposition 32 we get that

$$P^{v^\circ}_{G^\circ, \mathbb{1}_{v^*}}(x) = \frac{\frac{\mu_{G^\circ}(v^*, v^\circ)}{\deg^+_{v^*}(G^\circ)} x}{1 - \sum_{i=1}^{n} \frac{\mu_{G^\circ}(v^*, u_i')}{\deg^+_{v^*}(G^\circ)} \cdot P^{v_i}_{G^i, \mathbb{1}_{u_i}}(x)} = \frac{\alpha_0 x}{1 - \sum_{i=1}^{n} \alpha_i \cdot P^{v_i}_{G^i, \mathbb{1}_{u_i}}(x)}.$$

Combining this with Eq. (4.16) we get that $P^{v^\circ}_{G^\circ, \mathbb{1}_{u^\star}}(x) = \alpha_0 x / (1 - P^{v}_{G^*, \mathbb{1}_{v'}}(x))$. Finally, let us take graph $G^\bullet$ isomorphic to G° and disjoint with graphs $G^1,\dots,G^n$ and $G_1,\dots,G_m$ with node $u^\bullet$ as its start and node $v^\bullet$ as its end. Then, from Proposition 29a (Isomorphism) we get

$$P^{v^\bullet}_{G^\bullet, \mathbb{1}_{u^\bullet}}(x) = \frac{\alpha_0 x}{1 - P^{v}_{G^*, \mathbb{1}_{v'}}(x)}. \tag{4.17}$$

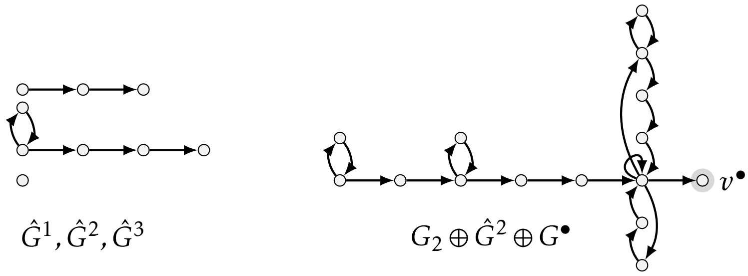


Figure 4.10: An illustration to the third part of case (II) of the proof of Proposition 33 for graph G from Fig. 4.8.

So far, we have obtained a broken cactus with generating function of its end $\alpha_0 x/(1-P^{v}_{G^*,\mathbb{1}_{v'}}(x))$ (Eq. (4.17)), a collection of broken cacti with averaged generating function of their ends $P^{v}_{G^*,\mathbb{1}_{v'}}(x)$ (Eq. (4.16)), and another collection with averaged generating function of their ends $P^{v}_{G^\dagger,b}(x)$ (Eq. (4.15)). Now, we will combine them together to obtain a collection of broken cacti with averaged generating function $P^{v}_{G^\dagger,b}(x) \cdot P^{v}_{G^*,\mathbb{1}_{v'}}(x)/(1-P^{v}_{G^*,\mathbb{1}_{v'}}(x))$.

To this end, recall that for every $i \in \{1,\ldots,n\}$, node u_i in graph G^i is a source with $q\alpha_i$ outgoing edges to node u_i'. Now, let us construct graph $\hat{G}^i$ which is G^i with node u_i removed (see Fig. 4.10). Formally, $\hat{G}^i = (V^i \setminus \{u_i\}, E^i - \Gamma^{+}_{u_i}(G^i))$ for every $i \in \{1,\ldots,n\}$. Observe that graph G^i can be seen as a broken cactus that is broken cactus $\hat{G}^i$ appended to a simple broken cactus $(\{u_i,t\}, q\alpha_i \cdot \{\!\{(u_i,t)\}\!\})$, for a new node t. Since the generating function of the end of the latter is simply x, from Proposition 30 we get that

$$P^{v_i}_{\hat{G}^i,\mathbb{1}_{u_i'}}(x) = \frac{1}{x} \cdot P^{v_i}_{G^i,\mathbb{1}_{u_i}}(x). \tag{4.18}$$

Thus, when we append broken cactus $G^\bullet$ to broken cactus $\hat{G}^i$, then from Proposition 30, Eq. (4.17), and Eq. (4.18) we get

$$P^{v^\bullet}_{\hat{G}^i \oplus G^\bullet,\mathbb{1}_{u_i'}}(x) = P^{v_i}_{G^i,\mathbb{1}_{u_i}}(x) \cdot \frac{\alpha_0}{1 - P^{v}_{G^*,\mathbb{1}_{v'}}(x)}.$$

Therefore, when we take a collection of broken cacti $\hat{G}^1 \oplus G^\bullet, \ldots, \hat{G}^n \oplus G^\bullet$ with constants $\alpha_1/\alpha_0, \ldots, \alpha_n/\alpha_0$, then from Eq. (4.16) we get that

$$\sum_{i=1}^{n} \frac{\alpha_i}{\alpha_0} P^{v^\bullet}_{\hat{G}^i \oplus G^\bullet,\mathbb{1}_{u_i'}}(x) = \frac{P^{v}_{G^*,\mathbb{1}_{v'}}(x)}{1 - P^{v}_{G^*,\mathbb{1}_{v'}}(x)}. \tag{4.19}$$

It remains, to multiply obtained generating function by $P^{v}_{G^\dagger,b}(x)$. To this end, we take a collection of broken cacti $G_1, \ldots, G_m$ and to each one of them append each graph from $\hat{G}^1 \oplus G^\bullet, \ldots, \hat{G}^n \oplus G^\bullet$ and for graph $G_j \oplus \hat{G}^i \oplus G^\bullet$ we take constant $\beta_j \cdot \alpha_i/\alpha_0$ (by Proposition 29a (Isomorphism), without loss of generality, we can assume that graphs $G_1, \ldots, G_m$ are pairwise disjoint and also disjoint from graphs $\hat{G}^1 \oplus G^\bullet, \ldots, \hat{G}^n \oplus G^\bullet$). As a result, we obtain that

$$\begin{aligned}
\sum_{j=1}^{m}\sum_{i=1}^{n} \frac{\beta_j \cdot \alpha_i}{\alpha_0} P^{v^\bullet}_{G_j \oplus \hat{G}^i \oplus G^\bullet,\mathbb{1}_{s_j}}(x) &= \sum_{j=1}^{m} \beta_j \cdot P^{w_j}_{G_j,\mathbb{1}_{s_j}}(x) \cdot \sum_{i=1}^{n} \frac{\alpha_i}{\alpha_0} \cdot P^{w_j}_{\hat{G}^i \oplus G^\bullet,\mathbb{1}_{u_i'}}(x) \\
&= \sum_{j=1}^{m} \beta_j \cdot P^{v^\bullet}_{G_j,\mathbb{1}_{s_j}}(x) \cdot \frac{P^{v}_{G^*,\mathbb{1}_{v'}}(x)}{1 - P^{v}_{G^*,\mathbb{1}_{v'}}(x)} \\
&= P^{v}_{G^\dagger,b}(x) \cdot P^{v}_{G^*,\mathbb{1}_{v'}}(x)/(1 - P^{v}_{G^*,\mathbb{1}_{v'}}(x)).
\end{aligned}$$

where the consecutive equalities come from Proposition 30, Eq. (4.19), and Eq. (4.15).

Finally, to this collection of graphs we add a collection of broken cacti $G_1, \dots, G_m$ with constants $\beta_1, \dots, \beta_m$. Then, from Eq. (4.14) we obtain that

$$\sum_{j=1}^{m} \beta_j P^{w_j}_{G_j, \mathbb{1}_{s_j}}(x) + \sum_{j=1}^{m} \sum_{i=1}^{n} \frac{\beta_j \cdot \alpha_i}{\alpha_0} P^{v^\bullet}_{G_j \oplus \hat{G}^i \oplus G^\bullet, \mathbb{1}_{s_j}}(x) = P^v_{G,b}(x).$$

It remains to prove the additional implications. For (a), observe that if all node weights are rational in (G,b), then it holds also in graph $(G^\dagger, b)$. Hence, from the proof of case (I) we get that constants $\beta_1, \dots, \beta_m$ are rational. Thus, also constants $\beta_j \cdot \alpha_i / \alpha_0$ are rational, for every $i \in \{1, \dots, n\}$ and $j \in \{1, \dots, m\}$. For (b), observe that if every node with positive weight is a source, then it holds also in graph $(G^\dagger, b)$. Hence, from the proof of case (I) we get that starts of broken cacti $G_1, \dots, G_m$ are sources as well. Since $\sum_{u \in P_v(G) \cup \{v\}} |\Gamma^-_u(G)| > 1$, we know that v has at least one incoming edge. Thus, its weight must be zero, which means that $p^1_{G^\dagger,b}(v,0) = 0$. Therefore, in each graph $G_1, \dots, G_m$ the start must be a distinct node from the end (otherwise in this graph the probability of visiting the end at step 0 would be 1). Thus, in graph $G_j \oplus \hat{G}^i \oplus G^\bullet$, for every $i \in \{1, \dots, n\}$ and $j \in \{1, \dots, m\}$, the start is also a source. Finally, for (c), observe that if $b(v) = 0$, then from the inductive assumption, we get that broken cactus G_j has main path of length at least 1 for every $j \in \{1, \dots, m\}$. Since in our collection each graph is a broken cactus from $G_1, \dots, G_m$ or a broken cactus obtained from appending a broken cactus to a graph from $G_1, \dots, G_m$, all of them have main paths of length at least 1. This concludes the proof. □

Centrality Measure

Now, let us move to the main part of the proof, in which we consider an arbitrary centrality measure F that satisfies our axioms.

In the first two lemmas, we show that if we restrict ourselves to sinks, we can strengthen Random Walk Property axiom. The original axiom states that for every two graphs with equal sum of node weights, if there is the same node in both graphs and it has the same visit probabilities, then it has the same centrality as well. First, we relax the assumption that the considered nodes have to be the same node. Specifically, we say that for every two graphs with equal sum of node weights, if there are two sinks in them that have the same visit probabilities, then their centrality is the same as well.

Lemma 34. *If a centrality measure, F, satisfies Random Walk Property, Locality, and Sink Merging, then for every two graphs $G = (V,E)$, $G' = (V',E')$, node weights b, b' such that $b(G) = b'(G') > 0$, and sinks $v \in V$, $v' \in V'$ such that $P^v_{G,b}(x) = P^{v'}_{G',b'}(x)$, it holds that*

$$F_v(G,b) = F_{v'}(G',b').$$

Proof. Let us consider an isolated node, $u \notin V \cup V'$, with zero weight, i.e., let $(G^u, b^u) = ((\{u\}, \emptyset), [0])$. Next, we add it to both (G,b) and (G',b') and redirect v into u in (G,b) and v' into u in (G',b'). Formally, $(\hat{G}, \hat{b}) = R_{v \to u}(G + G^u, b + b^u)$ and $(\hat{G}', \hat{b}') = R_{v' \to u}(G' + G^u, b' + b^u)$. From Locality and Lemma 27 we have that in both $(G + G^u, b + b^u)$ and $(G' + G^u, b' + b^u)$ node u has zero centrality. Therefore, from Locality and Sink Merging we obtain that $F_u(\hat{G}, \hat{b}) = F_v(G,b)$ and $F_u(\hat{G}', \hat{b}') = F_{v'}(G',b')$. Thus, it suffices if we prove that $F_u(\hat{G}, \hat{b}) = F_u(\hat{G}', \hat{b}')$

To this end, observe that graphs $(\hat{G}, \hat{b})$ and $(\hat{G}', \hat{b}')$ are isomorphic to graphs (G,b) and (G',b'), respectively. As a result, from Proposition 29a (Isomorphism) we obtain that $P^u_{\hat{G},\hat{b}}(x) = P^v_{G,b}(x) = P^{v'}_{G',b'}(x) = P^u_{\hat{G}',\hat{b}'}(x)$. Since u is a sink in both graphs, we know that it can be visited at most once by the random walk. Thus, $p_{\hat{G},\hat{b}}(u,t) = p_{\hat{G},\hat{b}}(u,t,1)$ and $p_{\hat{G}',\hat{b}'}(u,t) = p_{\hat{G}',\hat{b}'}(u,t,1)$, for every $t \in \mathbb{N}$. Hence, for every $t,k \in \mathbb{N}$, it holds that $p_{\hat{G},\hat{b}}(u,t,k) = p_{\hat{G}',\hat{b}'}(u,t,k)$. Since we also have that $\hat{b}(\hat{G}) = b(G) = b'(G') = \hat{b}'(\hat{G}')$, from Random Walk Property we get that $F_u(\hat{G},\hat{b}) = F_u(\hat{G}',\hat{b}')$. This concludes the proof. □

Next, we strengthen Random Walk Property even further. More in detail, we relax two assumptions: First, we remove the condition that the sum of node weights in both graphs has to be equal. Second, we say that the visit probabilities of both nodes do not have to be equal, it suffices if their are proportional, i.e., there exists a constant, $\alpha \in \mathbb{R}_{\geq 0}$, such that multiplying visit probabilities of one node by α results in visit probabilities of the other node. As a result, the centralities of both nodes do not have to be equal, but their ratio depends on α and the ratio of the sum of node weights in both graphs.

Lemma 35. *If a centrality measure, F, satisfies Random Walk Property, Locality, and Sink Merging, then for every two graphs $G = (V,E)$, $G' = (V',E')$, node weights b,b' such that $b(G), b'(G') > 0$, sinks $v \in V$, $v' \in V'$ and constant $\alpha \in (0,1]$ such that $P^v_{G,b}(x) = \alpha \cdot P^{v'}_{G',b'}(x)$, it holds that*

$$\frac{F_v(G,b)}{b(G)} = \alpha \frac{F_{v'}(G',b')}{b'(G')}.$$

Proof. First, we will show that for every graph $G = (V,E)$, node weights b, node $v \in V$, and constant $\alpha \in \mathbb{R}_{>0}$, it holds that

$$F_v(G, \alpha \cdot b) = \alpha \cdot F_v(G,b).$$

To this end, we will consider function $f : \mathbb{R}_{\geq 0} \to \mathbb{R}$ defined as $f(\alpha) = F_v(G, \alpha \cdot b)$. Observe that it is non-negative as a centrality is always non-negative. In what follows, we will prove that it is also additive. Both properties will imply that it is linear [23].

Let us take arbitrary $\alpha, \alpha' \in \mathbb{R}_{\geq 0}$ and prove that $f(\alpha + \alpha') = f(\alpha) + f(\alpha')$. Consider graph $G' = (V',E')$ and node weights b' such that $V' \cap V = \emptyset$ and (G',b') is isomorphic to (G,b) with isomorphism $g : V \to V'$ in which $g(v) = v'$. Observe that from Proposition 29a (Isomorphism) we have $P^v_{G,\alpha'\cdot b}(x) = P^{v'}_{G',\alpha'\cdot b'}(x)$. Since also $\alpha' \cdot b(G) = \alpha' \cdot b'(G')$, from Lemma 34 we get that $F_v(G, \alpha' \cdot b) = F_{v'}(G', \alpha' \cdot b')$.

Next, let us add together graphs $(G, \alpha \cdot b)$ and $(G', \alpha' \cdot b')$ and then redirect v' into v. Formally, let $(G'', b'') = R_{v' \to v}(G + G', \alpha \cdot b + \alpha' \cdot b')$. From Locality and Sink Merging we get that

$$F_v(G'',b'') = F_v(G, \alpha \cdot b) + F_{v'}(G', \alpha' \cdot b') = F_v(G, \alpha \cdot b) + F_v(G, \alpha' \cdot b). \tag{4.20}$$

On the other hand, from Proposition 29e (Graph Addition) we get that

$$P^v_{G+G',\alpha\cdot b+\alpha'\cdot b'}(x) = \frac{\alpha \cdot b(G) \cdot P^v_{G,\alpha\cdot b}(x)}{\alpha \cdot b(G) + \alpha' \cdot b'(G')} \text{ and } P^{v'}_{G+G',\alpha\cdot b+\alpha'\cdot b'}(x) = \frac{\alpha' \cdot b'(G) \cdot P^{v'}_{G',\alpha'\cdot b'}(x)}{\alpha \cdot b(G) + \alpha' \cdot b'(G')}.$$

Thus, when we redirect v' into v, then, by Proposition 29c (Sink Redirection),

$$P^v_{G'',b''}(x) = \frac{\alpha \cdot b(G) \cdot P^v_{G,\alpha\cdot b}(x) + \alpha' \cdot b'(G) \cdot P^{v'}_{G',\alpha'\cdot b'}(x)}{\alpha \cdot b(G) + \alpha' \cdot b'(G')}. \tag{4.21}$$

Moreover, from Proposition 29a (Isomorphism) and Proposition 29d (Weight Multiplication) we get $P^{v'}_{G',\alpha'\cdot b'}(x) = P^{v}_{G,\alpha\cdot b}(x)$. Hence, from Eq. (4.21) we obtain that $P^{v}_{G'',b''}(x) = P^{v}_{G,\alpha\cdot b}(x)$. Furthermore, observe that, by Proposition 29d (Weight Multiplication), this is equivalent to $P^{v}_{G'',b''}(x) = P^{v}_{G,(\alpha+\alpha')\cdot b}(x)$. Since we also have that $b''(G'') = \alpha \cdot b(G) + \alpha' \cdot b'(G') = (\alpha+\alpha') \cdot b(G)$, from Lemma 34 we obtain that $F_v(G'',b'') = F_v(G,(\alpha+\alpha')\cdot b)$. Now, combined with Eq. (4.20) this implies that $F_v(G,(\alpha+\alpha')\cdot b) = F_v(G,\alpha\cdot b) + F_v(G,\alpha'\cdot b)$. As a result, function f is additive. Since it is also non-negative, we get that it is linear. Hence, indeed, $F_v(G,\alpha\cdot b) = \alpha\cdot F_v(G,b)$ for every $\alpha \in \mathbb{R}_{\geq 0}$.

Now, let us move to the main part of the proof. Let us consider arbitrary graphs $G = (V,E)$, $G' = (V',E')$, node weights b, b', sinks $v \in V$, $v' \in V'$, and constant $\alpha \in (0,1]$ such that $P^{v}_{G,b}(x) = \alpha \cdot P^{v'}_{G',b'}(x)$. We will prove that $\frac{F_v(G,b)}{b(G)} = \alpha\frac{F_{v'}(G',b')}{b'(G')}$. To this end, let us first add a new node, $u \notin V'$, to graph G' and let us assign it such a weight that the total weight of resulting graph would be $b'(G')/\alpha$, i.e., a weight equal to $b'(G)/\alpha - b'(G)$. Formally, let $(G^u, b^u) = ((\{u\},\emptyset),[b'(G)/\alpha - b'(G)])$ and $(G'',b'') = (G'+G^u, b'+b^u)$. From Locality we have that

$$F_{v'}(G'',b'') = F_{v'}(G',b'). \tag{4.22}$$

On the other hand, from Proposition 29e (Graph Addition) we get that

$$P^{v}_{G'',b''}(x) = \frac{b'(G')}{b'(G') + \frac{b'(G')}{\alpha} - b'(G')} P^{v}_{G',b'}(x) = \alpha P^{v}_{G',b'}(x) = P^{v}_{G,b}(x).$$

Hence, if it holds that $b(G) = b''(G'')$, which means that $b(G) = b'(G')/\alpha$, then from Lemma 34 we get $F_v(G,b) = F_v(G'',b'')$. Thus, from Eq. (4.22) we obtain that $F_v(G,b) = F_v(G',b')$ and the thesis follows by sidewise division by $b(G) = b'(G')/\alpha$.

If, however, $b(G) \neq b''(G'')$, then let us accordingly scale the node weights of graph (G,b), i.e., let $b^* = \frac{b''(G'')}{b(G)} \cdot b$. From the first part of the proof, we have

$$F_v(G,b^*) = \frac{b''(G'')}{b(G)} \cdot F_v(G,b). \tag{4.23}$$

Now, by Proposition 29d (Weight Multiplication), $P^{v}_{G,b^*}(x) = P^{v}_{G,b}(x) = P^{v}_{G'',b''}(x)$. Moreover, observe that $b^*(G) = b''(G'')/b(G) \cdot b(G) = b''(G'')$. Thus, from Lemma 34 we obtain that $F_v(G,b^*) = F_v(G'',b'')$. In turn, by Eq. (4.23), this implies that $F_v(G,b)/b(G) = F_v(G'',b'')/b''(G'')$. Finally, observe that from Eq. (4.22) and the fact that $b''(G'') = b'(G')/\alpha$ we obtain that $F_v(G,b)/b(G) = \alpha F_v(G',b')/b'(G')$, which concludes the proof. □

In the central lemma of this part of the proof, we show that the centrality of the end of every broken cactus is equal to its RWD (up to a scalar multiplication).

Lemma 36. *If a centrality measure, F, satisfies Random Walk Property, Locality, Sink Merging, and Directed Leaf Proportionality, then there exists a constant, $c_F \in \mathbb{R}_{\geq 0}$, such that for every broken cactus $G = (V,E)$ with the start u and the end v, it holds that*

$$F_v(G,\mathbb{1}_u) = c_F \cdot RWD^a_v(G,\mathbb{1}_u).$$

Proof. By c_F let us denote the centrality of a node with unit weight in a graph without any other nodes nor edges. Formally, let $c_F = F_v((\{v\},\emptyset),[1])$.

We will follow the proof by induction. To this end, we strengthen the induction hypothesis by considering not only the centrality of the end of the broken cactus,

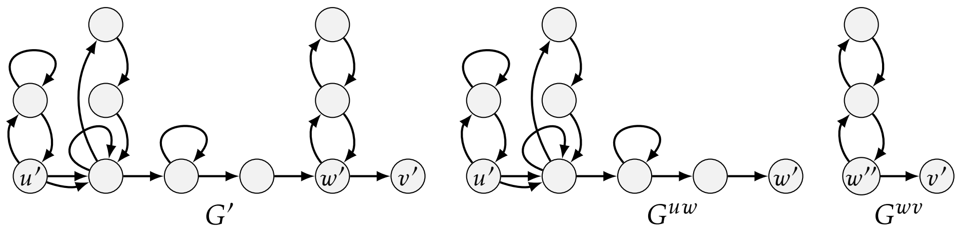


Figure 4.11: An illustration to the first part of the proof of Lemma 36 for an example broken cactus, G'.

but also the effect that broken cacti appending has on the centrality of the end. Specifically, we will show that for every two disjoint broken cacti $G = (V,E)$ and $G' = (V',E')$ with starts u,u' and ends v,v', respectively, it holds that

$$F_{v'}(G \oplus G', \mathbb{1}_u) = F_v(G, \mathbb{1}_u) \cdot RWD^a_{v'}(G', \mathbb{1}_{u'}) \text{ and } F_{v'}(G', \mathbb{1}_{u'}) = c_F \cdot RWD^a_{v'}(G', \mathbb{1}_{u'}).$$

Since G' is an arbitrary broken cactus, this will imply the original thesis.

The induction is on the number of edges in graph G'. If $|E'| = 0$, then broken cactus G' consists of a single isolated node, i.e., for an arbitrary node, v', we have $G' = ((\{v'\}, \emptyset))$. From Lemma 34 we have that $F_{v'}(G', \mathbb{1}_{v'}) = c_F$. Also, observe that $RWD^a_{v'}(G', \mathbb{1}_{v'}) = 1$, therefore $F_{v'}(G', \mathbb{1}_{v'}) = c_F \cdot RWD^a_{v'}(G', \mathbb{1}_{v'})$. Moreover, observe that the graph obtained by appending G' to G, i.e., $G \oplus G'$, is just graph G. Thus, $F_v(G \oplus G', \mathbb{1}_u) = F_v(G, \mathbb{1}_u)$, which concludes the proof of the induction basis.

Thus, let us assume that $|E'| > 0$. Since the end of a broken cactus is a sink, this means that $u' \neq v'$. First, let us assume that u' is not a direct predecessor of v', i.e., $(u', v') \notin E'$. Then, there exists another node on the main path of graph G' that is a direct predecessor of v'. Let us denote it by w'. Observe that this means that G' can itself be viewed as the result of the appending of two broken cacti: the first one, $G^{uw} = (V^{uw}, E^{uw})$, that starts at u' and ends at w', and the second one, $G^{wv} = (V^{wv}, E^{wv})$, that starts at w' and ends at v'. See Fig. 4.11 for an illustration. Formally, take $w'' \notin V \cup V'$ and let $V^{wv} = S_{w'}(G') \setminus \{w'\} \cup \{w''\}$ and $V^{uw} = V' \setminus V^{wv}$. Also, let $E^{uw} = \bigsqcup_{s \in V^{uw} \setminus \{w'\}} \Gamma^+_s(G')$ and $E^{wv} = E' - E^{uw} - \Gamma^{\pm}_{w'} \sqcup E^+_{w''} \sqcup E^-_{w''} \sqcup E^{\circ}_{w''}$, where $E^+_{w''} = \{\!\{(w'', t) : (w', t) \in \Gamma^+_{w'}(G'), t \neq w'\}\!\}$, $E^-_{w''} = \{\!\{(s, w'') : (s, w') \in \Gamma^-_{w'}(G'), s \in V^{wv}\}\!\}$ and $E^{\circ}_{w''} = \mu_{G'}(w', w') \cdot \{\!\{(w'', w'')\}\!\}$. Observe that indeed $G^{uw} \oplus G^{wv} = G'$. Moreover, both G^{uw} and G^{wv} have less edges than G', thus by the inductive assumption, $F_{w'}(G^{uw}, \mathbb{1}_{u'}) = c_F \cdot RWD^a_{w'}(G^{uw}, \mathbb{1}_{u'})$ and

$$\begin{aligned} F_{v'}(G', \mathbb{1}_{u'}) &= F_{w'}(G^{uw}, \mathbb{1}_{u'}) \cdot RWD^a_{v'}(G^{wv}, \mathbb{1}_{w''}) \\ &= c_F \cdot RWD^a_{w'}(G^{uw}, \mathbb{1}_{u'}) \cdot RWD^a_{v'}(G^{wv}, \mathbb{1}_{w''}). \end{aligned}$$

Since RWD also satisfies our axioms (Lemma 26) and $c_{RWD} = 1$, from this equation for $F = RWD^a$ we get

$$RWD^a_{v'}(G', \mathbb{1}_{u'}) = RWD^a_{w'}(G^{uw}, \mathbb{1}_{u'}) \cdot RWD^a_{v'}(G^{wv}, \mathbb{1}_{w''}). \tag{4.24}$$

Therefore, we get also that $F_{v'}(G', \mathbb{1}_{u'}) = c_F \cdot RWD^a_{v'}(G', \mathbb{1}_{u'})$. Now, consider appending G' to G. Again from the inductive assumption, used two times, we have

$$\begin{aligned} F_{v'}(G \oplus G', \mathbb{1}_u) &= F_{w'}(G \oplus G^{uw}, \mathbb{1}_u) \cdot RWD^a_{v'}(G^{wv}, \mathbb{1}_{w''}) \\ &= F_v(G, \mathbb{1}_u) \cdot RWD^a_{w'}(G^{uw}, \mathbb{1}_{u'}) \cdot RWD^a_{v'}(G^{wv}, \mathbb{1}_{w''}). \end{aligned}$$

Thus, we get the thesis from Eq. (4.24).

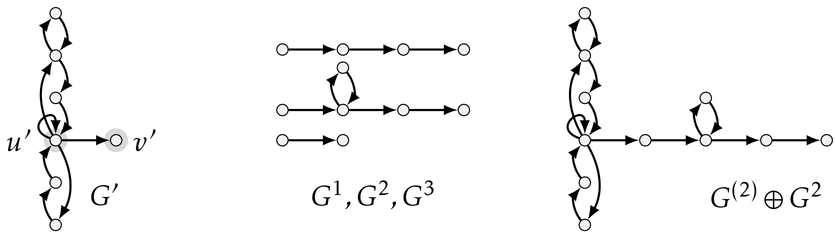


Figure 4.12: An illustration to the first part of case (II) of the proof of Lemma 36 for an example broken cactus, G'.

Hence, let us assume that $(u',v') \in E'$, i.e., the main path of G' is of length 1. Now, we consider two cases: the first one (I) in which there is no cycle in G', and the second one (II) in which there is at least one cycle in G'. Let us begin with the former case.

(I) If there are no cycles in broken cactus G' and its end is a direct successor of its start, it means that edges in G' can go only from node u' to v'. Hence, it is of the form $G' = (\{u',v'\}, k \cdot \{\!\{(u',v')\}\!\})$ for some $k \geq 1$. If $k > 1$, then let us reduce the number of edges (u',v') to one, i.e, let $G'' = (\{u',v'\}, \{\!\{(u',v')\}\!\})$. From Proposition 29f (Edge Multiplication) we get that $P^{v'}_{G',\mathbb{1}_{u'}}(x) = P^{v'}_{G'',\mathbb{1}_{u'}}(x)$. Thus, from Lemma 34 we have $F_{v'}(G',\mathbb{1}_{u'}) = F_{v'}(G'',\mathbb{1}_{u'})$. In the same way, $F_v(G \oplus G',\mathbb{1}_u) = F_v(G \oplus G'',\mathbb{1}_u)$. Since G'' has less edges than G', we get the thesis from the inductive assumption.

Thus, let us assume that $k = 1$, i.e., $G' = (\{u',v'\}, \{\!\{(u',v')\}\!\})$. In such a case, let us remove all edges from the graph, i.e., let $G'' = (\{u',v'\}, \emptyset)$. Since now node v' is isolated, from Locality and Lemma 27 we have $F_{v'}(G'',\mathbb{1}_{u'}) = 0$. Moreover, from Locality and the induction basis, we have $F_{u'}(G'',\mathbb{1}_{u'}) = c_F$. Therefore, by Directed Leaf Proportionality, $F_{v'}(G',\mathbb{1}_{u'}) = F_{v'}(G'',\mathbb{1}_{u'}) + a \cdot F_{u'}(G'',\mathbb{1}_{u'}) = a \cdot c_F$. Since RWD also satisfies our axioms (Lemma 26) and $c_{RWD} = 1$, we obtain that indeed $F_{v'}(G',\mathbb{1}_{u'}) = c_F \cdot RWD_{v'}(G',\mathbb{1}_{u'})$.

It remains to prove that $F_{v'}(G \oplus G',\mathbb{1}_u) = F_v(G,\mathbb{1}_u) \cdot RWD^a_{v'}(G',\mathbb{1}_{u'})$. To this end, let $G^* = (V \cup \{v'\}, E)$ be graph G with node v' added. Observe that if we add edge (v,v') to this graph, we obtain broken cactus $G \oplus G'$. Hence, Directed Leaf Proportionality yields

$$F_{v'}(G \oplus G',\mathbb{1}_u) = F_{v'}(G^*,\mathbb{1}_u) + a \cdot F_v(G^*,\mathbb{1}_u).$$

In G^*, node v' is isolated. Hence, $F_{v'}(G^*,\mathbb{1}_u) = 0$. By Locality, $F_v(G^*,\mathbb{1}_u) = F_v(G,\mathbb{1}_u)$. Thus, we get that $F_{v'}(G \oplus G',\mathbb{1}_u) = F_v(G,\mathbb{1}_u) \cdot a = F_v(G,\mathbb{1}_u) \cdot RWD_{v'}(G',\mathbb{1}_{u'})$.

(II) Let us move to the case in which there is at least one cycle in G'. This means that the successors of node u' consists of v' and $n \geq 1$ other nodes. In what follows, using Proposition 32, we "uncycle" graph G' into a collection of broken cacti $G^1, \dots, G^n$ (see Fig. 4.12). Each of these broken cacti will have less edges than G', thus with a series of graph operations and by the inductive assumption, we will prove the induction hypothesis. To this end, let us denote the set of direct successors of u' by $S^1_{u'}(G') = \{v', w_1, \dots, w_n\}$. Also, for convenience, let $\alpha_i = \mu_{G'}(u',w_i)/\deg^+_{u'}(G')$, for every $i \in \{1,\dots,n\}$, and $\alpha_0 = \mu_{G'}(u',v')/\deg^+_{u'}(G')$. Finally, take $w'_1, \dots, w'_n, v_1, \dots, v_n \notin V \cup V'$ and let us denote the broken cacti obtained by the construction from Proposition 32 applied to graph G' by $G^1 = (V^1, E^1), \dots, G^n = (V^n, E^n)$ with starts $w'_1, \dots, w'_n$ and ends $v_1, \dots, v_n$ respectively in such a way that $w_i \in V^i$ for every $i \in \{1,\dots,n\}$. Then, from Proposition 32 we get that

$$P^{v'}_{G',\mathbb{1}_{u'}}(x) = \frac{\alpha_0 x}{1 - \sum_{i=1}^{n} \alpha_i P^{v_i}_{G^i,\mathbb{1}_{w'_i}}(x)}. \tag{4.25}$$

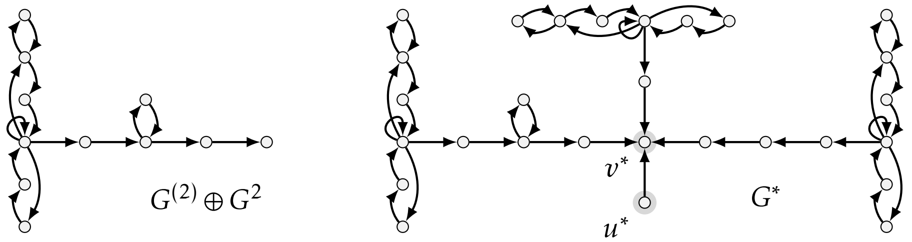


Figure 4.13: An illustration to the second part of case (II) of the proof of Lemma 36 for broken cactus G' from Fig. 4.12.

Next, let us consider pairwise disjoint broken cacti $G^{(1)},\dots,G^{(n)}$ that are isomorphic to G' and disjoint from $G^1,\dots,G^n$ and G. Also, let us denote their starts by $u^{(1)},\dots,u^{(n)}$ and ends by $v^{(1)},\dots,v^{(n)}$, respectively. Then, consider appending graph G^i to $G^{(i)}$, i.e., $G^{(i)} \oplus G^i$. From Proposition 29a (Isomorphism) we obtain $P^{v^{(i)}}_{G^{(i)},\mathbb{1}_{u^{(i)}}}(x) = P^{v'}_{G',\mathbb{1}_{u'}}(x)$. Thus, by Proposition 30, $P^{v_i}_{G^{(i)}\oplus G^i,\mathbb{1}_{u^{(i)}}}(x) = P^{v'}_{G',\mathbb{1}_{u'}}(x)\cdot P^{v_i}_{G^i,\mathbb{1}_{w'_i}}(x)$. Let us multiply both sides by α_i and sum for all $i \in \{1,\dots,n\}$. Then, from Eq. (4.25) we get

$$\sum_{i=1}^{n} \alpha_i P^{v_i}_{G^{(i)}\oplus G^i,\mathbb{1}_{u^{(i)}}}(x) = \frac{\alpha_0 x}{1-\sum_{i=1}^{n}\alpha_i P^{u'}_{G^i,\mathbb{1}_{w'_i}}(x)} \cdot \sum_{i=1}^{n} \alpha_i P^{u'}_{G^i,\mathbb{1}_{w'_i}}(x).$$

For any two formal series P and Q, it holds that $Q\cdot P/(1-P)+Q = Q/(1-P)$. Hence, by adding $\alpha_0 x$ to both sides of the above equation, we obtain that

$$\alpha_0 x + \sum_{i=1}^{n} \alpha_i P^{u'}_{G^{(i)}\oplus G^i,\mathbb{1}_{u^{(i)}}}(x) = \frac{\alpha_0 x}{1-\sum_{i=1}^{n}\alpha_i P^{u'}_{G^i,\mathbb{1}_{w'_i}}(x)} = P^{v'}_{G',\mathbb{1}_{u'}}(x), \tag{4.26}$$

where the second equality comes from Eq. (4.25).

Building upon this, we will now follow Proposition 31 to obtain graph G^* with generating function $P^{v'}_{G',\mathbb{1}_{u'}}(x)$ but constructed from graphs $G^{(1)} \oplus G^1,\dots,G^{(n)} \oplus G^n$ by adding them together and redirecting their ends (see Fig. 4.13). To obtain a graph with generating function $\alpha_0 x$, from the left hand side of Eq. (4.26), let us take new nodes u^*, v^* and denote $G^{(0)} = (\{u^*,v^*\}, \{\!\{(u^*,v^*)\}\!\})$. Now, let

$$(G^*,b^*) = R_{v^{(1)}\to v^*}\left(\dots\left(R_{v^{(n)}\to v^*}\left(G^{(0)} + \sum_{i=1}^{n} G^{(i)} \oplus G^i, \alpha_0 \mathbb{1}_{u^*} + \sum_{i=1}^{n} \alpha_i \mathbb{1}_{u^{(i)}}\right)\right)\dots\right).$$

Since $P^{v^*}_{G^{(0)},\mathbb{1}_{u^*}}(x) = x$, Proposition 31 yields $P^{v^*}_{G^*,b^*}(x) = \alpha_0 x + \sum_{i=1}^{n} \alpha_i P^{v_i}_{G^{(i)}\oplus \hat{G}^i,\mathbb{1}_{u^{(i)}}}(x)$. Hence, from Eq. (4.26) we obtain that $P^{v^*}_{G^*,b^*}(x) = P^{v'}_{G',\mathbb{1}_{u'}}(x)$. Moreover, observe that $b^*(G^*) = \alpha_0 + \dots + \alpha_n = 1 = \mathbb{1}_{u'}(G')$. Thus, from Lemma 34 we get

$$F_{v^*}(G^*,b^*) = F_{v'}(G',\mathbb{1}_{u'}). \tag{4.27}$$

On the other hand, from case (I) we have $F_{v^*}((\{u^*,v^*\},\{\!\{(u^*,v^*)\}\!\}),\mathbb{1}_{u^*}) = c_F \cdot a$. Thus, by Locality and Sink Merging, $F_{v^*}(G^*,b^*) = \alpha_0 c_F \cdot a + \sum_{i=1}^{n} \alpha_i F_{v_i}(G^{(i)} \oplus G^i, \mathbb{1}_{u^{(i)}})$. By Proposition 32, for every $i \in \{1,\dots,n\}$, broken cactus G^i has less edges than G'. Thus, $F_{v^*}(G^*,b^*) = \alpha_0 c_F \cdot a + \sum_{i=1}^{n} \alpha_i F_{v^{(i)}}(G^{(i)},\mathbb{1}_{u^{(i)}}) \cdot RWD^a_{v_i}(G^i,\mathbb{1}_{w'_i})$ from the inductive assumption. Also, for every $i \in \{1,\dots,n\}$, we get $F_{v^{(i)}}(G^{(i)},\mathbb{1}_{u^{(i)}}) = F_{v'}(G',\mathbb{1}_{u'})$ from Lemma 34. Hence, $F_{v^*}(G^*,b^*) = \alpha_0 c_F \cdot a + F_{v'}(G',\mathbb{1}_{u'}) \cdot \sum_{i=1}^{n} \alpha_i RWD^a_{v_i}(G^i,\mathbb{1}_{w'_i})$.

Combining this with Eq. (4.27) yields

$$F_{v'}(G', \mathbb{1}_{u'}) = \alpha_0 c_F \cdot a + F_{v'}(G', \mathbb{1}_{u'}) \cdot \sum_{i=1}^{n} \alpha_i RWD^a_{v_i}(G^i, \mathbb{1}_{w'_i}).$$

Since $\mathbb{1}_{w'_i}(G^i) = 1$, by Eq. (4.3), for every $i \in \{1,\dots,n\}$, $RWD^a_{v_i}(G^i, \mathbb{1}_{w'_i})$ is the probability of visiting node v_i in the random walk with decay factor a starting from node w'_i. Thus, $RWD^a_{v_i}(G^i, \mathbb{1}_{w'_i}) \le 1$. Hence, $\sum_{i=1}^{n} \alpha_i RWD^a_{v_i}(G^i, \mathbb{1}_{w'_i}) < 1$, because $\sum_{i=1}^{n} \alpha_i = 1 - \alpha_0 < 1$. Therefore, we can transform the above equation into

$$F_{v'}(G', \mathbb{1}_{u'}) = c_F \frac{\alpha_0 \cdot a}{1 - \sum_{i=1}^{n} \alpha_i RWD^a_{v_i}(G^i, \mathbb{1}_{w'_i})}. \tag{4.28}$$

Since RWD also satisfies our axioms (Lemma 26) and $c_{RWD} = 1$, we get that

$$RWD^a_{v'}(G', \mathbb{1}_{u'}) = \frac{\alpha_0 \cdot a}{1 - \sum_{i=1}^{n} \alpha_i RWD^a_{v_i}(G^i, \mathbb{1}_{w'_i})}.$$

Hence, $F_{v'}(G', \mathbb{1}_{u'}) = c_F \cdot RWD^a_{v'}(G', \mathbb{1}_{u'})$.

Thus, it remains to prove that $F_{v'}(G \oplus G', \mathbb{1}_u) = F_v(G, \mathbb{1}_u) \cdot RWD^a_{v'}(G, \mathbb{1}_{u'})$. To this end, we will use a similar reasoning, but instead of graphs $G^{(1)} \oplus G^1, \dots, G^{(n)} \oplus G^n$ we will consider graphs $G \oplus G^{(1)} \oplus G^1, \dots, G \oplus G^{(n)} \oplus G^n$. Observe that multiplying both sides of Eq. (4.26) by $P^v_{G,\mathbb{1}_u}(x)$ we get

$$\alpha_0 x \cdot P^v_{G,\mathbb{1}_u}(x) + \sum_{i=1}^{n} \alpha_i P^v_{G,\mathbb{1}_u}(x) \cdot P^{u'}_{G^{(i)} \oplus G^i, \mathbb{1}_{u^{(i)}}}(x) = P^v_{G,\mathbb{1}_u}(x) \cdot P^{v'}_{G',\mathbb{1}_{u'}}(x).$$

Thus, from Proposition 30 we obtain that

$$\alpha_0 x \cdot P^v_{G,\mathbb{1}_u}(x) + \sum_{i=1}^{n} \alpha_i P^{v_i}_{G \oplus G^{(i)} \oplus G^i, \mathbb{1}_u}(x) = P^{v'}_{G \oplus G', \mathbb{1}_u}(x). \tag{4.29}$$

Therefore, we will follow Proposition 31 and construct graph $(G^\star, b^\star)$ based on graphs $G \oplus G^{(1)} \oplus G^1, \dots, G \oplus G^{(n)} \oplus G^n$, and G such that $P^v_{G^\star, b^\star}(x) = P^{v'}_{G \oplus G', \mathbb{1}_u}(x)$. To this end, let us define pairwise disjoint broken cacti $\hat{G}^{(1)}, \dots, \hat{G}^{(n)}$ isomorphic to graphs $G \oplus G^{(1)} \oplus G^1, \dots, G \oplus G^{(n)} \oplus G^n$, respectively, and let us denote their starts as $\hat{u}^{(1)}, \dots, \hat{u}^{(n)}$ and ends as $\hat{v}^{(1)}, \dots, \hat{v}^{(n)}$, respectively. From Proposition 29a (Isomorphism) we obtain

$$P^{\hat{v}^{(i)}}_{\hat{G}^{(i)}, \mathbb{1}_{\hat{u}^{(i)}}}(x) = P^{v_i}_{G \oplus G^{(i)} \oplus G^i, \mathbb{1}_u}(x), \quad \text{for every } i \in \{1, \dots, n\}. \tag{4.30}$$

Also, by $\hat{G}$ we denote the graph resulting from appending a simple broken cactus $G^{(0)} = (\{u^\star, v^\star\}, \{\!\{(u^\star, v^\star)\}\!\})$, for new nodes $u^\star, v^\star$, to graph G, i.e., let $\hat{G} = G \oplus G^{(0)}$ (see Fig. 4.14). From Proposition 30 we get that $P^{v^\star}_{\hat{G},\mathbb{1}_u}(x) = x \cdot P^v_{G,\mathbb{1}_u}(x)$. Combining this with Eq. (4.29) and Eq. (4.30) we obtain

$$\alpha_0 P^{v^\star}_{\hat{G},\mathbb{1}_u}(x) + \sum_{i=1}^{n} \alpha_i P^{\hat{v}^{(i)}}_{\hat{G}^{(i)}, \mathbb{1}_{\hat{u}^{(i)}}}(x) = P^{v'}_{G \oplus G', \mathbb{1}_u}(x). \tag{4.31}$$

Building upon this, let us define

$$(G^\star, b^\star) = R_{v^{(1)} \to v^\star}\left(\dots \left(R_{v^{(n)} \to v^\star}\left(\hat{G} + \sum_{i=1}^{n} \hat{G}^{(i)}, \alpha_0 \mathbb{1}_u + \sum_{i=1}^{n} \alpha_i \mathbb{1}_{\hat{u}^{(i)}}\right)\right) \dots \right).$$

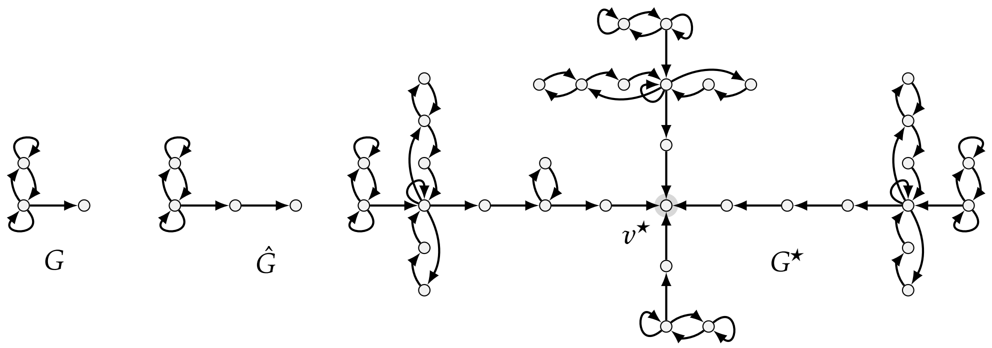


Figure 4.14: An illustration to the third part of case (II) of the proof of Lemma 36 for broken cactus G' from Fig. 4.12 and an example broken cactus, G.

From Proposition 31 and Eq. (4.31) we get that indeed

$$P^{v^\star}_{G^\star,b^\star}(x) = \alpha_0 P^{v^\star}_{\hat{G},\mathbb{1}_u}(x) + \sum_{i=1}^{n} \alpha_i P^{\hat{v}^{(i)}}_{\hat{G}^{(i)},\mathbb{1}_{\hat{u}^{(i)}}}(x) = P^{v'}_{G\oplus G',\mathbb{1}_u}(x).$$

Since $b^\star(G^\star) = \sum_{i=0}^{n} \alpha_i = 1$, from Lemma 34 we get that

$$F_v(G^\star, b^\star) = F_{v'}(G \oplus G', \mathbb{1}_u). \tag{4.32}$$

On the other hand, observe that from case (I) of the proof we can obtain that $F_{v^\star}(\hat{G}, \mathbb{1}_u) = F_v(G, \mathbb{1}_u) \cdot RWD^a_v(G^{(0)}, \mathbb{1}_{u^\star}) = a \cdot F_v(G, \mathbb{1}_u)$. Thus, from Locality and Sink Merging we get

$$F_v(G^\star, b^\star) = \alpha_0 a \cdot F_v(G, \mathbb{1}_u) + \sum_{i=1}^{n} \alpha_i F_{\hat{v}^{(i)}}(\hat{G}^{(i)}, \mathbb{1}_{\hat{u}^{(i)}}). \tag{4.33}$$

From Lemma 34 we get $F_{\hat{v}^{(i)}}(\hat{G}^{(i)}, \mathbb{1}_{\hat{u}^{(i)}}) = F_{v_i}(G \oplus G^{(i)} \oplus G^i, \mathbb{1}_u)$, for every $i \in \{1, \ldots, n\}$. Since graph G^i has less edges than G', from the inductive assumption we get that $F_{\hat{v}^{(i)}}(\hat{G}^{(i)}, \mathbb{1}_{\hat{u}^{(i)}}) = F_{v^{(i)}}(G \oplus G^{(i)}, \mathbb{1}_u) \cdot RWD^a_{v_i}(G^i, w_i)$. As $G^{(i)}$ is isomorphic to G', Lemma 34 yields $F_{\hat{v}^{(i)}}(\hat{G}^{(i)}, \mathbb{1}_{\hat{u}^{(i)}}) = F_{v'}(G \oplus G', \mathbb{1}_u) \cdot RWD^a_{v_i}(G^i, w_i)$. Inserting this into Eq. (4.33), we get $F_v(G^\star, b^\star) = \alpha_0 a \cdot F_v(G, \mathbb{1}_u) + \sum_{i=1}^{n} \alpha_i F_{v'}(G \oplus G', \mathbb{1}_u) \cdot RWD^a_{v_i}(G^i, w_i)$. Thus, from Eq. (4.32) we obtain

$$F_{v'}(G \oplus G', \mathbb{1}_u) = \alpha_0 a \cdot F_v(G, \mathbb{1}_u) + F_{v'}(G \oplus G', \mathbb{1}_u) \cdot \sum_{i=1}^{n} \alpha_i RWD^a_{v_i}(G^i, w_i),$$

which can be transformed into

$$F_{v'}(G \oplus G', \mathbb{1}_u) = \frac{\alpha_0 a F_v(G, \mathbb{1}_u)}{1 - \sum_{i=1}^{n} \alpha_i \cdot RWD^a_{v_i}(G^i, \mathbb{1}_{w_i'})}.$$

Therefore, the thesis follows from Eq. (4.28) □

Now, let us move to arbitrary graphs. In the next lemma, we show that centrality F is equal to RWD (up to a scalar multiplication) for every sink in every graph.

Lemma 37. *If a centrality measure, F, satisfies Random Walk Property, Locality, Sink Merging, and Directed Leaf Proportionality, then there exists $c_F \in \mathbb{R}_{\geq 0}$ such that for every graph $G = (V, E)$, node weights b, and sink $v \in V$, it holds that*

$$F_v(G, b) = c_F \cdot RWD^a_v(G, b).$$

Proof. Consider an arbitrary graph $G = (V,E)$, node weights b, and sink $v \in V$. If $b(G) = 0$, then the thesis follows from Lemma 27. Otherwise, since v is a sink, it is also a limiting node. Thus, from Proposition 33 we get that there exist constants $\alpha_1,\dots,\alpha_n \geq 0$ and a collection of broken cacti $G^1,\dots,G^n$ with starts $u_1,\dots,u_n$ and ends $v_1,\dots,v_n$, respectively, such that

$$P^v_{G,b}(x) = \sum_{i=1}^{n} \alpha_i \cdot P^{v_i}_{G^i,\mathbb{1}_{u_i}}(x).$$

By Proposition 29a (Isomorphism), without loss of generality, we can assume that graphs $G^1,\dots,G^n$ are pairwise disjoint. Next, we construct graph (G',b') by adding together graphs $G^1,\dots,G^n$ and then merging their ends. Formally, let us denote $(G',b') = R_{v_1\to v_n}(\dots(R_{v_{n-1}\to v_n}(G^1+\dots+G^n,\alpha_1\cdot\mathbb{1}_{u_1}+\dots+\alpha_n\mathbb{1}_{u_n}))\dots)$. Then, from Proposition 31 we have that $P^v_{G,b}(x) = P^{v_n}_{G',b'}(x)\cdot\sum_{i=1}^n \alpha_i$. Observe that $b'(G') = \sum_{i=1}^n \alpha_i$. Hence, from Lemma 35 we get that

$$F_v(G,b) = b(G)\cdot F_v(G',b'). \tag{4.34}$$

On the other hand, observe that from Locality and Sink Merging we obtain $F_v(G',b') = c_F\cdot\sum_{i=1}^n \alpha_i\cdot F_{v_i}(G^i,\mathbb{1}_{u_i})$. Since graphs $G^1,\dots,G^n$ are broken cacti, from Lemma 36 we get that $F_{v_i}(G^i,\mathbb{1}_{u_i}) = c_F\cdot RWD^a_{v_i}(G^i,\mathbb{1}_{u_i})$, for every $i\in\{1,\dots,n\}$. Thus, $F_v(G',b') = c_F\cdot\sum_{i=1}^n\alpha_i\cdot RWD^a_{v_i}(G^i,\mathbb{1}_{u_i})$. Combining this with Eq. (4.34) we obtain that

$$F_v(G,b) = c_F\cdot b(G)\cdot\sum_{i=1}^{n}\alpha_i\cdot RWD^a_{v_i}(G^i,\mathbb{1}_{u_i}).$$

Since RWD also satisfies our axioms (Lemma 26) and $c_{RWD} = 1$, we obtain that indeed $F_v(G,b) = c_F\cdot RWD^a_v(G,b)$. □

Finally, we show that centrality of every node in every graph is equal to RWD. First, up to a scalar multiplication, and then, in the exact values.

Lemma 38. *If a centrality measure, F, satisfies Random Walk Property, Locality, Sink Merging, Lack of Self-Impact, and Directed Leaf Proportionality, then there exists a constant, $c_F \in \mathbb{R}_{\geq 0}$, such that for every graph $G = (V,E)$ and node weights b, it holds that*

$$F_v(G,b) = c_F\cdot RWD^a_v(G,b), \quad \text{for every } v\in V.$$

Proof. Consider an arbitrary graph $G=(V,E)$, node weights b, and fix $v\in V$. Then, let us remove outgoing edges of v from graph G. Formally, let $G' = (V, E-\Gamma^+_v(G))$. From Lack of Self-Impact we have that $F_v(G,b) = F_v(G',b)$. Observe that in graph G' node v is a sink. Thus, from Lemma 37 we get that $F_v(G',b) = c_F\cdot RWD^a_v(G',b)$. Since we know that RWD also satisfies Lack of Self-Impact (Lemma 26), we obtain that $RWD^a_v(G',b) = RWD^a_v(G,b)$, from which the thesis follows. □

Lemma 39. *If a centrality measure, F, satisfies Random Walk Property, Locality, Sink Merging, Lack of Self-Impact, Directed Leaf Proportionality, and One-Node Graph, then for every graph $G=(V,E)$ and node weights b, it holds that*

$$F_v(G,b) = RWD^a_v(G,b), \quad \text{for every } v\in V.$$

Proof. Observe that $RWD^a_v((\{v\},\emptyset),[1]) = 1$. Therefore, the thesis follows from One-Node Graph and Lemma 38. □

4.3.2 Personalized Decay Centrality (Theorem 24)

Let us move to the proof of Theorem 24, i.e., that personalized decay centrality is the only centrality measure that satisfies Shortest Paths Property, Locality, Sink Merging, Directed Leaf Proportionality, and One-Node Graph. We begin by showing that personalized decay indeed satisfies all five axioms.

Lemma 40. *For every decay factor $a \in [0,1)$, personalized decay centrality satisfies Shortest Paths Property, Locality, Sink Merging, Directed Leaf Proportionality, and One-Node Graph.*

Proof. Let us take an arbitrary graph $G = (V,E)$, node weights b, and consider the axioms one by one.

- For Shortest Paths Property, take graph $G' = (V',E')$, weights b', and node $v \in V \cap V'$ such that $b(G) = b(G')$ and $|\{u \in V : dist_G(u,v) = k \wedge b(u) = x\}| = |\{u \in V' : dist_{G'}(u,v) = k \wedge b'(u) = x\}|$, for every $k \in \mathbb{N}$ and $x \in \mathbb{R}$. Then, from the definition of personalized decay centrality (Eq. (2.1)) we get that also
$$Y^a_v(G,b) = \sum_{u \in V} b(u) \cdot a^{dist_G(u,v)} = \sum_{u \in V'} b'(u) \cdot a^{dist_{G'}(u,v)} = Y^a_v(G',b).$$
- For Locality let us consider graph $G' = (V',E')$ such that $V \cap V' = \emptyset$, weights b', and node $v \in V$. Observe that in graph $(G+G', b+b')$ there is no path from nodes in V' to v. Thus, for every $u \in V'$, we have $dist(u,v) = \infty$. Therefore, from Eq. (2.1), we get that
$$Y^a_v(G+G', b+b') = \sum_{u \in V \cup V'} b(u) \cdot a^{dist(u,v)} = \sum_{u \in V} b(u) \cdot a^{dist(u,v)} = Y^a_v(G,b).$$
- For Sink Merging, consider two sinks $u,w \in V$ such that $P_u(G) \cap P_w(G) = \emptyset$. Since u and w are sinks, for every $v \in V \setminus \{u,w\}$, there is no path from u or w to v in graph G. Likewise, there is no such path in a graph obtained by redirection of u into w. Also, the distances to v from other nodes do not change, hence from Eq. (2.1) we get
$$Y^a_v(G,b) = \sum_{s \in V \setminus \{u,w\}} b(s) \cdot a^{dist(s,v)} = Y^a_v(R_{u \to w}(G,b)).$$

  It remains to consider node w. Since u and w do not have common predecessors, we get that each predecessor of w in graph $R_{u \to w}(G,b)$ is either a predecessor of u in G or a predecessor of w in G, but not both. The distances from predecessors are not affected, therefore
$$Y^a_w(R_{u \to w}(G,b)) = \sum_{s \in P_u(G) \cup P_w(G)} b(s) \cdot a^{dist(s,v)} = \sum_{s \in P_u(G)} b(s) \cdot a^{dist(s,v)} + \sum_{s \in P_w(G)} b(s) \cdot a^{dist(s,v)} = Y^a_u(G,b) + Y^a_w(G,b).$$
- For Directed Leaf Proportionality, consider sink $u \in V$ and an isolated node $v \in V$. Also, let $G' = (V, E \sqcup \{\!\{(u,v)\}\!\})$. Observe that in graph G', for every node $s \in V \setminus \{v\}$, each walk from s to v must visit node u in its last but one step.

Hence, $dist_{G'}(s,v) = dist_{G'}(s,u) + 1$. Also, no path to u can go through node v, thus $dist_{G'}(s,u) = dist_G(s,u)$. Therefore, from Eq. (2.1) we get

$$Y_v^a(G',b) = a \cdot \left(\sum_{s \in V \setminus \{v\}} b(s) \cdot a^{dist(s,u)} \right) + b(v) \cdot a^0 = a \cdot Y_u^a(G,b) + b(v).$$

Since personalized decay of an isolated node is equal to its weight, the axiom follows.

- Finally, for One-Node Graph assume that $(G,b) = ((\{v\},\emptyset),[1])$. Observe that in such a case, from Eq. (2.1) we get $Y_v^a(G,b) = b(v) = 1$. □

Now we move to the second part of the proof, in which we show that an arbitrary centrality measure F that satisfies Shortest Paths Property, Locality, Sink Merging, and Directed Leaf Proportionality is equal to personalized decay centrality up to a scalar multiplication. Similarly to Section 4.3.1, the decay factor of personalized decay centrality, which centrality measure F is equal to, is given by constant a in Directed Leaf Proportionality. Thus, formally, we show that there exists a constant, $c_F \in \mathbb{R}_{\geq 0}$, such that for every graph $G = (V,E)$, node weights b, and node $v \in V$, it holds that $F_v(G,b) = c_F \cdot Y_v^a(G,b)$. Then, from One-Node Graph we obtain that $c_F = 1$ and $F_v(G,b) = Y_v^a(G,b)$.

We prove that F is equal to personalized decay (up to a scalar multiplication) by considering nodes in arbitrary graphs that are: first, isolated (Lemma 41), and then, arbitrary (Lemma 42).

Lemma 41. *If a centrality measure, F, satisfies Locality and Sink Merging, then there exists a constant, $c_F \in \mathbb{R}_{\geq 0}$, such that for every graph $G = (V,E)$, node weights b, and isolated node $v \in V$, it holds that*

$$F_v(G,b) = c_F \cdot b(v) = c_F \cdot Y_v^a(G,b).$$

Proof. It is easy to check that from the definition of personalized decay centrality indeed $Y_v^a(G,b) = b(v)$. Hence, let us focus on proving that there exists $c_F \in \mathbb{R}_{\geq 0}$ such that $F_v(G,b) = c_F \cdot b(v)$, for every graph (G,b) and isolated node v.

To this end, let us first consider one-node graphs without any edges, i.e., graphs of the form: $((\{v\},\emptyset),[x])$ for some node v and $x \in \mathbb{R}_{\geq 0}$. Consider two such graphs, $((\{u\},\emptyset),[x])$ and $((\{v\},\emptyset),[y])$, for arbitrary $u \neq v$ and $x,y \in \mathbb{R}_{\geq 0}$. Let (G,b) be their sum: $(G,b) = ((\{u,v\},\emptyset),[x,y])$. From Locality we know that their centralities are the same as in the original graphs. In particular,

$$F_u(G,b) + F_v(G,b) = F_u((\{u\},\emptyset),[x]) + F_v((\{v\},\emptyset),[y]). \tag{4.35}$$

Nodes u and v are sinks and do not have any predecessors. Thus, by Sink Merging, redirecting node v into u increases the centrality of u by the centrality of v. Such a redirection results in graph $((\{u\},\emptyset),[x+y])$, so we get:

$$F_u((\{u\},\emptyset),[x+y]) = F_u(R_{v \to u}(G,b)) = F_u(G,b) + F_v(G,b). \tag{4.36}$$

Combining Eq. (4.35) and Eq. (4.36) we have

$$F_u((\{u\},\emptyset),[x+y]) = F_u((\{u\},\emptyset),[x]) + F_v((\{v\},\emptyset),[y]). \tag{4.37}$$

We make the following observations:

(a) $F_v((\{v\},\emptyset),[0]) = 0$, for every v (from Eq. (4.37) with $y = 0$);

(b) $F_v((\{v\},\emptyset),[y]) = F_u((\{u\},\emptyset),[y])$, for every $u \neq v$ and $y \in \mathbb{R}_{\geq 0}$ (from Eq. (4.37) with $x = 0$ and (a));

(c) $F_v((\{v\},\emptyset),[x+y]) = F_v((\{v\},\emptyset),[x]) + F_v((\{v\},\emptyset),[y])$, for every v and $x, y \in \mathbb{R}_{\geq 0}$ (from Eq. (4.37) and (b)).

Note that (b) implies that the centrality of v in the weighted graph $((\{v\},\emptyset),[x])$ depends solely on weight x. In other words, there exists a function $f : \mathbb{R}_{\geq 0} \to \mathbb{R}$ such that $F_v((\{v\},\emptyset),[x]) = f(x)$. Since centralities are non-negative, we know that f is also non-negative, i.e., $f(x) \geq 0$, for every $x \in \mathbb{R}_{\geq 0}$. On the other hand, from (c) we know that f is additive, i.e., $f(x+y) = f(x) + f(y)$, for every $x, y \in \mathbb{R}_{\geq 0}$. Thus, from non-negativity and additivity we obtain that f is linear [23], i.e., $f(x) = c_F \cdot x$ for some $c_F \in \mathbb{R}_{\geq 0}$, for every $x \in \mathbb{R}_{\geq 0}$. As a result, we know that that there exists $c_F \in \mathbb{R}_{\geq 0}$ such that for every node v we have $F_v((\{v\},\emptyset),[x]) = c_F \cdot x$. Then, the thesis follows from Locality. □

Next, let us move to arbitrary nodes in arbitrary graphs and show that their centrality is equal to personalized decay (up to a scalar multiplication).

Lemma 42. *If a centrality measure, F, satisfies Shortest Paths Property, Locality, Sink Merging, and Directed Leaf Proportionality, then there exists a constant, $c_F \in \mathbb{R}_{\geq 0}$, such that for every graph $G = (V,E)$ and node weights b, it holds that*

$$F_v(G,b) = c_F \cdot Y^a_v(G,b), \quad \textit{for every } v \in V.$$

Proof. The proof is by induction on the number of edges. If there are no edges in the graph, then the thesis is implied by Lemma 41. Hence, let us focus on the case in which there is at least one edge in a graph.

Fix $v \in V$. Let us denote the maximal distance from any node in V to v by d, i.e., let $d = \max_{u \in V} dist_G(u,v)$. Then, denote the sets of nodes from which the distance to node v is equal $0,1,\ldots,d$ by $V_0, V_1, \ldots, V_d$, respectively. Formally, $V_0 = \{v\}$ and $V_i = \{u \in P_v(G) : dist(u,v) = i\}$, for every $i \in \{1,\ldots,d\}$. Observe that for every node $u \in V_i$, there is an outgoing edge from u to a node in V_{i-1}. Let us denote it by e_u.

Now, consider the case in which u has also another outgoing edge, namely $e \neq e_u$. Let G' be a graph with edge e removed, i.e., $G' = (V, E - \{\!\{e\}\!\})$. Since in G', for every $i \in \{1,\ldots,d\}$, every node from V_i still has an outgoing edge to one node in V_{i-1}, the distances to v from all of its predecessors does not change, i.e., for every $s \in P_v(G)$ we have $dist_G(s,v) = dist_{G'}(s,v)$. Thus, from Shortest Paths Property we have that $F_v(G,b) = F_v(G',b)$. Graph G' has one edge less than G, therefore from the inductive assumption we get the that $F_v(G,b) = c_F \cdot Y^a_v(G',b)$. Since personalized decay centrality also satisfies our axioms (Lemma 40) and $c_Y = 1$, we obtain the thesis.

Hence, let us assume that for every $i \in \{1,\ldots,d\}$ every node $u \in V_i$ has exactly one outgoing edge and it goes to a node in V_{i-1}. Now, consider the case in which v has an outgoing edge, $e \in \Gamma^+_v(G)$. Let G' be a graph with edge e removed, i.e., $G' = (V, E - \{\!\{e\}\!\})$. Again, the distances to v from all of its predecessors does not change, i.e., $dist_G(s,v) = dist_{G'}(s,v)$. Thus, from Shortest Paths Property we have that $F_v(G,b) = F_v(G',b)$. Graph G' has one edge less than G, therefore from the inductive assumption, the fact that personalized decay centrality also satisfies our axioms (Lemma 40), and $c_Y = 1$, we get the thesis.

Thus, let us assume that v is a sink. Now, consider the case in which v has more than one incoming edge, i.e., $|\Gamma^-_v(G)| > 1$. Observe that all of the direct predecessors

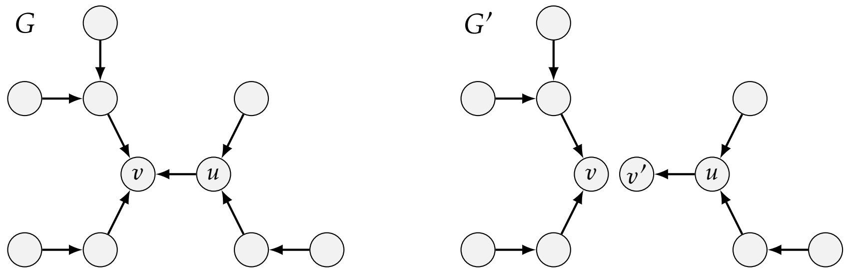


Figure 4.15: An illustration to the proof of Lemma 42 for an example graph, G.

of v are in set V_1. Therefore, each of them has exactly one edge and it goes to node v. Hence, also $|P^1_v(G)| > 1$. Let us denote one of the direct predecessors of v by $u \in P^1_v(G)$. Next, let us split node v into two nodes: node v', with only one incoming edge (u,v') and zero weight, and node v, with all of the other original incoming edges of v and the original weight of v (see Fig. 4.15 for an illustration). Formally, let $G' = (V \cup \{v'\}, E - \{\!\{(u,v)\}\!\} \sqcup \{\!\{(u,v')\}\!\})$, $b'_V = b$, and $b'(v') = 0$. Observe that indeed $(G,b) = R_{v' \to v}(G',b')$.

Moreover, let us show that v and v' do not have a common predecessor. Assume otherwise and take a common predecessor, $s \in P_v(G') \cap P_{v'}(G')$, such that its distance to node v is minimal. Node s has exactly one outgoing edge. Let us denote it by (s,s'). Observe that s' has to also be a common predecessor of v and v', but $dist_{G'}(s',v) < dist_{G'}(s,v)$. Thus, we arrive at a contradiction. Since v and v' are sinks without commons predecessors, from Sink Merging we get that

$$F_v(G,b) = F_v(G',b') + F_{v'}(G',b'). \tag{4.38}$$

Observe that v' and its predecessors and their outgoing edges constitute a connected component in G'. Thus, let us denote the graph with this connected component removed, i.e., let $G'' = (V \setminus P_{v'}(G') \setminus \{v'\}, E - \{\!\{e_w : w \in P_{v'}(G')\}\!\})$. Also, let us denote the graph in which we keep this connected component and remove all other nodes and their edges, i.e., let $G^* = (\{v'\} \cup P_{v'}(G'), \{\!\{e_w : w \in P_{v'}(G') \setminus \{u\}\}\!\} \sqcup \{\!\{(u,v')\}\!\})$. Observe that indeed G'' and G^* are disjoint and $G'' + G^* = G'$. From Locality we obtain that $F_v(G',b') = F_v(G'', b'_{V \setminus P_{v'}(G')})$ and $F_{v'}(G',b') = F_{v'}(G^*, b'_{\{v'\} \cup P_{v'}(G')})$. Since both G'' and G^* has less edges than G, from the inductive assumption and the fact that personalized decay centrality also satisfies our axioms (Lemma 40) we get $F_v(G',b') = c_F \cdot Y^a_v(G',b')$ and $F_{v'}(G',b') = c_F \cdot Y^a_{v'}(G',b')$. Hence, the thesis follows from Eq. (4.38).

Thus, let us assume that v has at most one incoming edge. Now, consider the case in which v has exactly one such edge, namely (u,v). Consider graph with (u,v) removed, i.e., let $G' = (V, E - \{\!\{(u,v)\}\!\})$. From the inductive assumption we get that $F_u(G',b) = c_F \cdot Y^a_u(G',b)$ and $F_v(G',b) = c_F \cdot Y^a_v(G',b)$. Observe that from Directed Leaf Proportionality we have that

$$F_v(G,b) = a \cdot F_u(G',b) + F_v(G',b) = c_F \cdot \Big(a \cdot Y^a_u(G',b) + Y^a_v(G',b)\Big).$$

Since personalized decay centrality also satisfies Directed Leaf Proportionality (Lemma 40), we obtain the thesis.

Finally, let us assume that v does not have any incoming edges. In such a case, v is isolated and the thesis follows from Lemma 41. □

Finally, we add One-Node Graph and show that personalized decay centrality is uniquely characterized.

Lemma 43. *If a centrality measure, F, satisfies Shortest Paths Property, Locality, Sink Merging, Directed Leaf Proportionality, and One-Node Graph, then for every graph $G = (V, E)$ and node weights b, it holds that*

$$F_v(G, b) = Y_v^a(G, b), \quad \text{for every } v \in V.$$

Proof. Observe that $Y_v^a((\{v\}, \emptyset), [1]) = 1$. Therefore, the thesis follows from One-Node Graph and Lemma 42. □

4.3.3 PageRank (Theorem 25)

Finally, let us prove that Random Walk Property, Locality, Sink Merging, Edge Swap, Directed Leaf Proportionality, and One-Node Graph uniquely characterize PageRank. Let us start by showing that PageRank indeed satisfies all six axioms.

Lemma 44. *For every decay factor $a \in [0, 1)$ PageRank satisfies Random Walk Property, Locality, Sink Merging, Edge Swap, Directed Leaf Proportionality, and One-Node Graph.*

Proof. Let us take an arbitrary graph $G = (V, E)$, node weights b, and consider the axioms one by one.

- For Random Walk Property, consider graph $G' = (V', E')$, node weights b', and node $v \in V \cap V'$ such that $b(G) = b(G')$ and $p^1_{G,b}(v, t, k) = p^1_{G',b'}(v, t, k)$, for every $t, k \in \mathbb{N}$. Observe that from Eq. (4.1) we get

  $$p^a_{G,b}(v, t, k) = a^t \cdot p^1_{G,b}(v, t, k) = a^t \cdot p^1_{G',b'}(v, t, k) = p^a_{G',b'}(v, t, k),$$

  for every $t, k \in \mathbb{N}$. Hence, the axiom follows from Eq. (4.2).

- Locality follows from Lemma 6.

- For Sink Merging, consider two sinks $u, w \in V$ such that $P_u(G) \cap P_w(G) = \emptyset$. Observe that nodes u and w are out-twins, hence from Node Redirect we obtain that $PR^a_v(R_{u\to w}(G, b)) = PR^a_v(G, b)$, for every $v \in V \setminus \{u, w\}$, and also that $PR^a_w(R_{u\to w}(G, b)) = PR^a_u(G, b) + PR^a_w(G, b)$. Thus, Sink Merging is satisfied.

- The fact that PageRank satisfies Edge Swap is proven in Section 3.2.1.

- For Directed Leaf Proportionality, consider a sink, $u \in V$, and an isolated node, $v \in V$. Also, denote $G' = (V, E \sqcup \{\!\{(u, v)\}\!\})$. Observe that from the PageRank recursive equation (Eq. (2.5)) for node v and graph (G', b) we obtain that $PR^a_v(G', b) = a \cdot PR^a_u(G', b) + b(v)$. Since u is a sink in G, it is not a successor of itself in G'. Thus, from Edge Deletion we get that $PR^a_u(G', b) = PR^a_u(G, b)$. In turn, from Baseline we get that $b(v) = PR_v(G, b)$. Combining all three equations we get that PageRank satisfies Directed Leaf Proportionality.

- Finally, for One-Node Graph assume that $(G, b) = ((\{v\}, \emptyset), [1])$. Since PageRank satisfies Baseline, we have $PR^a_v(G, b) = b(v) = 1$.

□

Now, let us move to the second part of the proof, in which we show that if a centrality measure F satisfies Random Walk Property, Locality, Sink Merging, Edge Swap, and Directed Leaf Proportionality, then F is equal to PageRank up to a scalar multiplication. Formally, we show that there exists constant $c_F \in \mathbb{R}_{\geq}$ such that for every graph $G = (V, E)$, node weights b, and node $v \in V$, it holds that $F_v(G, b) = c_F \cdot PR^a_v(G, b)$, where a is a constant from Directed Leaf Proportionality.

Consider Lemma 37, in which we proved that centrality measure satisfying Random Walk Property, Locality, Sink Merging, and Directed Leaf Proportionality is equal to RWD (up to a scalar multiplication) for every sink. Observe that since it does not require Lack-of Self-Impact, it applies to F as well. On the other hand, from Proposition 22 we know that for every sink RWD is equal to PageRank. Thus, we already know that F is equal to PageRank (up to a scalar multiplication) for all sinks. As we will show, if a graph does not have any cycles, then the centrality of every node is not affected by the removal of outgoing edges of this node. Hence, the remainder of the proof is very similar to the end of the proof in Section 3.2.2 where we considered all possible graphs knowing that the thesis holds for all graphs with no cycles (Lemma 13).

Lemma 45. *If a centrality measure, F, satisfies Random Walk Property, Locality, Sink Merging, Edge Swap, and Directed Leaf Proportionality, then there exists a constant, $c_F \in \mathbb{R}_{\geq 0}$, such that for every graph $G = (V, E)$ and node weights b, it holds that*

$$F_v(G, b) = c_F \cdot PR^a_v(G, b) \quad \textit{for every } v \in V.$$

Proof. The proof is similar to the proof of Lemma 13 in Section 3.2.2. In the same manner, we use induction on the number of cycles in G. The induction basis is proven differently than in Lemma 13, but the induction step is analogous.

First, assume that there are no cycles in G. Fix $v \in V$ and observe that since there is no cycle, v can be visited at most once by the random walk, i.e., we have $p^1_{G,b}(v, t, k) = 0$, for every $t \in \mathbb{N}$ and $k > 1$. Thus, removing outgoing edges of v does not affect the probabilities with which it is visited. Formally, let us denote $G' = (V, E - \Gamma^+_v(G))$. In this way, we obtain $p^1_{G,b}(v, t, k) = p^1_{G',b}(v, t, k)$, for every $t, k \in \mathbb{N}$. Hence, from Random Walk Property we get $F_v(G, b) = F_v(G', b)$. Since v is a sink in G', from Lemma 37 we get $F_v(G', b) = c_F \cdot RWD^a_v(G', b)$. Moreover, from Proposition 22 we get that for sinks both RWD and PageRank are equal. Therefore, $RWD^a_v(G', b) = PR^a_v(G', b)$. Combining the equations we get that $F_v(G, b) = c_F \cdot PR^a_v(G', b)$. Since PageRank also satisfies our axioms (Lemma 44) and $c_{PR} = 1$, the induction basis follows.

Now, let us assume that G has at least one cycle. In such a case, fix node w that belongs to at least one cycle and let x_w be its PageRank, i.e., $x_w = PR^a_w(G, b)$. Consider graph (G', b') obtained from (G, b) by adding two-node graph consisted of node s with weight x_w, node t with weight 0, and edges from s to t in the number equal to $\deg^+_w(G)$. Formally, let $(G^{st}, b^{st}) = ((\{s, t\}, \deg^+_w(G) \cdot \{\!\{(s, t)\}\!\}), [x_w, 0])$ and $(G', b') = (G + G^{st}, b + b^{st})$. Then,

$$PR^a_w(G', b') = x_w \quad \text{and} \quad PR^a_s(G', b') = x_w, \tag{4.39}$$

where the first equation holds from Locality (which we know that PageRank satisfies from Lemma 44) and the second from PageRank recursive equation (Eq. (2.5)). Moreover, nodes w and s have the same number of outgoing edges in graph (G', b'), i.e., $\deg^+_w(G)$. Therefore, if we swap the ends of all of their outgoing edges, then from Edge Swap this operation will not affect PageRank of any node. Formally, let $G'' = (V \cup \{s, t\}, E - \Gamma^+_u(G) \sqcup \{\!\{(w, t), (s, w') : (w, w') \in \Gamma^+_u(G)\}\!\})$. From the fact that PageRank satisfies Edge Swap we get that

$$PR^a_v(G'', b') = PR^a_v(G', b'), \quad \text{for every } v \in V. \tag{4.40}$$

Observe that in graph (G'', b') all of the outgoing edges of w go to node t. Hence, graph (G'', b') has less cycles than graph (G, b) (every cycle in the former graph is also a cycle in the later one, but the former graph does not contain cycles with w). Hence, from the inductive assumption we know that

$$F_v(G'', b') = c_F \cdot PR_v^a(G'', b'), \quad \text{for every } v \in V. \tag{4.41}$$

Thus, combining Eqs. (4.39)–(4.41) we obtain that

$$F_w(G'', b') = c_F \cdot PR_w^a(G'', b') = c_F \cdot x_w = c_F \cdot PR_s^a(G'', b') = F_s(G'', b').$$

Thus, nodes w and s have equal centralities and equal number of outgoing edges in graph (G'', b'). Therefore, again from Edge Swap, this time for centrality F, we get that

$$F_v(G', b') = F_v(G'', b'), \quad \text{for every } v \in V.$$

Combining this with Eq. (4.40) and Eq. (4.41) yields $F_v(G', b') = c_F \cdot PR_v(G', b')$. Hence, the thesis follows from Locality. □

Finally, let us assume that F satisfies One-Node Graph as well and prove that in such a case PageRank is uniquely characterized.

Lemma 46. *If a centrality measure, F, satisfies Random Walk Property, Locality, Sink Merging, Edge Swap, Directed Leaf Proportionality, and One-Node Graph, then for every graph $G = (V, E)$ and node weights b, it holds that*

$$F_v(G, b) = PR_v^a(G, b), \quad \textit{for every } v \in V.$$

Proof. Observe that $PR_v^a((\{v\}, \emptyset), [1]) = 1$. Therefore, the thesis follows from One-Node Graph and Lemma 45. □

4.4 Comparison of PageRank and RWD

Our axiomatic characterizations highlight two differences between random walk decay centrality and PageRank. In this section, we focus on these two differences and show how they affect the behaviour of these centrality measures.

4.4.1 Strategy-Proofness (with Respect to Outgoing Edges)

In many settings, outgoing edges are subject to node's decision or manipulations. Examples include the Twitter social network (where outgoing edges represent the accounts that are followed by a user) and the World Wide Web (where outgoing edges represent the links to other websites). Consequently, in such settings, Lack of Self-Impact implies a property of *strategy-proofness* of centrality measures—if outgoing edges do not affect the centrality of a node, then the node has no incentive to manipulate its outgoing connections.

Interestingly, PageRank does not satisfy Lack of Self-Impact. In the following example we show how, by adding outgoing edges, a node can increase its centrality and position in the ranking according to PageRank, but not according to random walk decay centrality.

Example 5. *Let us consider graph G from Fig. 4.16. Graph G consists of two 4-cycles, $(u_1, u_2, u_3, u_4, u_1)$, $(v_1, v_2, v_3, v_4, v_1)$. The two cycles are connected via 3 edges: (v_4, u_4), (u_3, v_3), and (u_2, v_2). Due to the edges connecting both cycles, the nodes v_2, v_3, and v_4 are visited more often by the random walk, and are thus ranked first by both PageRank and random walk decay centrality. Node u_1, that will be of our interest, is ranked 5th according to both measures.*

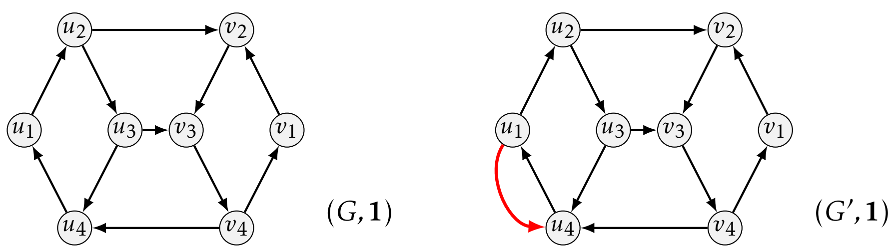


Figure 4.16: Two graphs considered in Example 5, $G = (V,E)$ and $G' = (V,E')$, with unit weights. Graph G' is obtained from graph G by adding (u_1,u_4).

Fig. 4.16 also depicts G', which is obtained from G by adding edge (u_1,u_4). Since this is an outgoing edge for u_1, adding it does not affect random walk decay centrality of u_1. In contrast, this edge has a significant impact on PageRank of u_1. The reason lies in the fact that the random walk will now visit u_1 much more often—whenever the random walk reaches node u_1, with probability $1/2$ it will go back to u_4, from which the only outgoing edge goes to u_1. As a result, both u_1 and u_4 top the ranking according to PageRank. The centralities of all nodes for $a = 0.9$ are as follows:

node v	$PR_v^{0.9}(G,\mathbf{1})$	$PR_v^{0.9}(G',\mathbf{1})$	$RWD_v^{0.9}(G,\mathbf{1})$	$RWD_v^{0.9}(G',\mathbf{1})$
v_3	13.89 *(1st)*	11.88 *(3rd)*	6.02 *(1st)*	5.67 *(1st)*
v_4	13.50 *(2nd)*	11.69 *(4th)*	5.85 *(2nd)*	5.58 *(2nd)*
v_2	11.67 *(3rd)*	9.93 *(5th)*	5.5 *(3rd)*	5.07 *(4th)*
u_2	9.57 *(4th)*	7.33	5.16 *(4th)*	4.50
$\boldsymbol{u_1}$	**9.52 *(5th)***	**14.08 *(2nd)***	**5.13 *(5th)***	**5.13 *(5th)***
u_4	9.46	14.53 *(1st)*	5.10	5.30 *(3rd)*
v_1	7.08	6.26	4.03	3.71
u_3	5.30	2.45	4.3	3.16

This concludes Example 5.

In Example 5, a node improved its PageRank by adding an edge to its direct predecessor. In the next section, we will discuss how incoming edges affect both centrality measures.

4.4.2 Diversity (of Incoming Edges)

PageRank is a feedback centrality and as such the centrality it assigns to nodes depends solely on the centrality of their direct predecessors. This means that it does not look on any the other aspects like the relative placement of its predecessors in the topology of the network. In our axiomatic characterization, this property is captured by Edge Swap, which implies that an incoming edge from a node with the lowest centrality in a densely connected part of the graph could be as profitable as an incoming edge from a node with the highest centrality in a different, less densely connected part.

Random walk decay centrality does not satisfy Edge Swap. In fact, it is more profitable to have an incoming edge from a diverse set of nodes. We demonstrate this point with the following example.

Example 6. *Let us consider graph G from Fig. 4.17. Observe that this graph consists of three more densely connected parts, so called* communities: $\{u_1,u_2,u_3\}$, $\{v_1,v_2,v_3\}$, *and* $\{w_1,w_2,w_3,w_4\}$. *These communities are connected through nodes u_1,v_1,w_1 which*

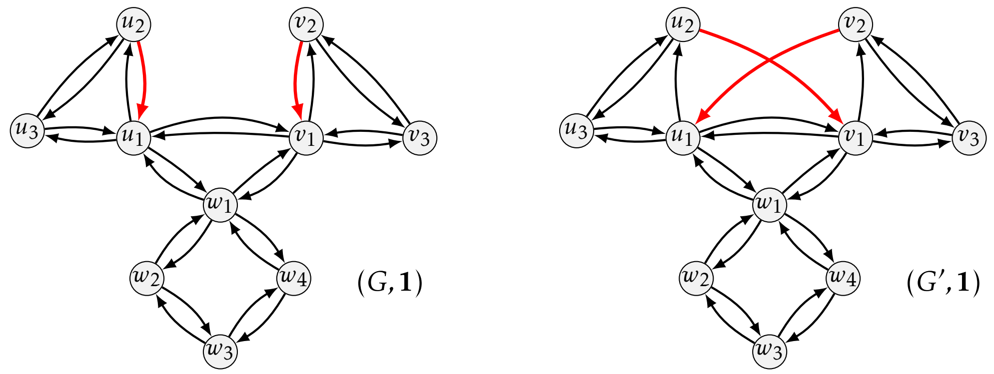


Figure 4.17: Two graphs considered in Example 6, $G = (V, E)$ and $G' = (V, E')$, with unit node weights. Graph G' is obtained from graph G by replacing edges (u_2, u_1) and (v_2, v_1) with edges (u_2, v_1) and (v_2, u_1).

form a 3-clique. Since w_1 belongs to the biggest community, both its random walk decay centrality and its PageRank are the highest. The nodes u_1 and v_1 have the second highest values, with symmetrical positions in the graph.

Fig. 4.17 also depicts the graph G', which is obtained from G by rewiring the two highlighted (red) edges. Specifically, the edges (u_2, u_1) and (v_2, v_1) are replaced by (u_2, v_1) and (v_2, u_1); as a result, the two new edges connect two communities. Since u_2 and v_2 both have two edges and clearly the same centralities in graph G, from Edge Swap we know that PageRank of every node in G' is the same as in G. In contrast, the centralities of both nodes u_1 and v_1 increase according to random walk decay centrality. This is because, according to this centrality, an edge from a different community is more profitable than an edge from your own community. In our example, the random walk that starts from nodes v_2 and v_3 reaches node u_1 faster in graph G'. As a result, in G', random walk decay centralities of u_1 and also v_1 are higher than random walk decay centrality of node w_1. The centralities of all nodes for $a = 0.9$ are:

node v	$PR_v^{0.9}(G, \mathbf{1})$	$PR_v^{0.9}(G', \mathbf{1})$	$RWD_v^{0.9}(G, \mathbf{1})$	$RWD_v^{0.9}(G', \mathbf{1})$
w_1	14.87 *(1st)*	14.87 *(1st)*	6.11 *(1st)*	6.11 *(3rd)*
$\boldsymbol{u_1}$	**14.71** *(**2nd**)*	**14.71** *(**2nd**)*	**5.75** *(**2nd**)*	**6.17** *(**1st**)*
$\boldsymbol{v_1}$	**14.71** *(**2nd**)*	**14.71** *(**2nd**)*	**5.75** *(**2nd**)*	**6.17** *(**1st**)*
w_3	8.26 *(4th)*	8.26 *(4th)*	3.68	3.68
w_2	8.06 *(5th)*	8.06 *(5th)*	3.96 *(4th)*	3.96
w_4	8.06 *(5th)*	8.06 *(5th)*	3.96 *(4th)*	3.96
u_2	7.83	7.83	3.71	4.21 *(4th)*
v_2	7.83	7.83	3.71	4.21 *(4th)*
u_3	7.83	7.83	3.71	3.98
v_3	7.83	7.83	3.71	3.98

This concludes Example 6.

Example 6 shows that random walk decay centrality increases when incoming edges become more diverse. As such, it avoids putting at the top of the ranking several nodes from the same community, which often happens in PageRank [4, 88].

4.4.3 Real-World Example

Let us conclude this section with an example on a real-world network.

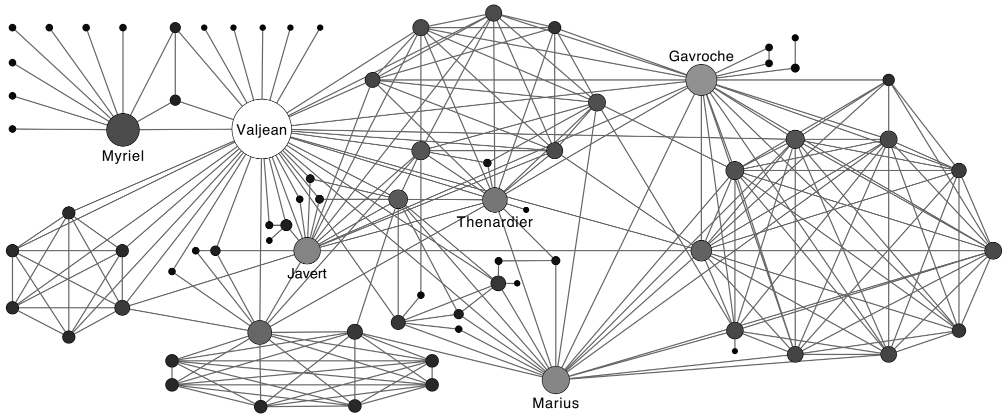


Figure 4.18: The graph of character co-occurrence in *Les Misérables* novel [50]. The size of nodes corresponds to their PageRank and their lightness to RWD.

Example 7. *Fig. 4.18 depicts the graph of character co-occurrence in Les Misérables novel by Victor Hugo [50]. Each node represents a character from the novel and an edge between two nodes means that corresponding characters have met at least once. Since the graph is undirected, we treat each undirected edge as a pair of edges in both directions. The values of random walk decay centrality and PageRank as well as the rankings based on them for five nodes with the highest PageRank are presented in the following table:*

Character	*PageRank ($a = 0.9$)*		*RWD ($a = 0.9$)*	
Jean Valjean	*1.*	58.28	*1.*	30.89
Bishop Myriel	*2.*	30.53	*16.*	10.43
Gavroche	*3.*	28.79	*2.*	18.36
Marius Pontmercy	*4.*	24.94	*3.*	17.20
Inspector Javert	*5.*	24.14	*4.*	16.94

Jean Valjean, the main protagonist of the novel, is indicated as the most important node by both PageRank and RWD. Now, for PageRank, the second most important node is Bishop Myriel. He is a side character, important in the first out of five volumes of the book, but not appearing later on. However, when we look closely at the graph we notice two things:

- *First, most of the characters that interact with Myriel interact only with him. Thus, he has a lot of outgoing edges to nodes that themselves have edges only to him, which greatly increase PageRank as we discussed in Example 5.*
- *Second, the three remaining characters interacting with Myriel is Valjean and two of his neighbours. Hence, incoming edges of Myriel are not diversified as discussed in Example 6.*

For these reasons, random walk decay centrality gives Bishop Myriel much less importance than PageRank. In fact, according to random walk decay centrality, he is ranked as 16th, giving place to other characters like the iconic child of the street Gavroche, one of the main protagonists Marius Pontmercy, or Inspector Javert—the main antagonist of the novel. This concludes Example 7.

Chapter 5

An Axiom System for Feedback Centralities

The aim of this chapter is to propose a coherent axiomatization of all four considered feedback centralities: eigenvector centrality, Katz centrality, Seeley index, and PageRank. We approach this goal by building upon our axiomatization of PageRank from Chapter 3. Specificaly, we propose a consistent characterization of these four feedback centralities in the form of a system of seven axioms. *Locality*, *Edge Deletion*, *Node Combination* are general axioms satisfied by all four centralities. *Edge Compensation* and *Edge Multiplication* concern modification of one node and its incident edges. Finally, *Cycle* and *Baseline* specify centralities in simple borderline graphs. We show that each of four feedback centralities in question is uniquely characterized by a subset of 5 axioms: 3 general ones, one one-node-modification axiom, and one borderline axiom. Our axiomatic characterizations are summarized in Tab. 5.1.

The chapter is structured as follows. We begin by introducing some additional notation in Section 5.1. Next, in Section 5.2, we present the axioms which constitute our axiom system. Finally, in Section 5.3, we prove that particular sets of axioms uniquely characterize respective centrality measures.

The content of this chapter is an extended version of the paper published in the proceedings of the IJCAI-21 conference [82].

5.1 Additional Notation

In this section, we build on the notation introduced in Section 2 to introduce basic concepts used in this chapter.

In previous chapters we have considered multigraphs with node weights. A multigraph with node weights, G, can alternatively be viewed as a simple graph

Centrality	General axioms	Node-modification axiom	Borderline axiom
Eigenvector	LOC, ED, NC	Edge Compensation	Cycle
Katz	LOC, ED, NC	Edge Compensation	Baseline
Seeley index	LOC, ED, NC	Edge Multiplication	Cycle
PageRank	LOC, ED, NC	Edge Multiplication	Baseline

Table 5.1: Our axiomatic characterizations of eigenvector centrality, Katz centrality, Seeley index, and PageRank. General axioms are Locality (LOC), Edge Deletion (ED), and Node Combination (NC).

with a set of non-repeating edges, each of which is assigned a weight, μ_G, that can take only positive natural values. In this chapter, we generalize this notion and allow for arbitrary positive real weights on edges. Formally, a *graph* is now a pair, $G = (V, E)$, where V is a set of nodes and $E \subseteq V \times V$ is a set of edges. Additionally, the weights of a graph are given by $\theta = (b, \mu)$, where $b : V \to \mathbb{R}_{\geq 0}$ are node weights and $\mu : E \to \mathbb{R}_{>0}$ are edge weights. A *weighted graph* is then a pair (G, θ) and by $\mathcal{G}$ we denote the set of all possible such weighted graphs. Also, like before, to denote small weighted graphs, we will use the following simplified notation:

$$(G,\theta) = \Big((\{v_1,\dots,v_n\},\{e_1,\dots,e_m\}),([b_1,\dots,b_n],[\mu_1,\dots,\mu_m])\Big)$$

which means $G = (\{v_1,\dots,v_n\},\{e_1,\dots,e_m\})$ and $\theta = (b,\mu)$ such that $b(v_i) = b_i$, for every $i \in \{1,\dots,n\}$, and $\mu(e_j) = \mu_j$, for every $j \in \{1,\dots,m\}$.

Moreover, let $\tilde{\mu} : V \times V \to \mathbb{R}_{\geq 0}$ be the extension of edge weights μ such that $\tilde{\mu}(u,v) = \mu(u,v)$, if $(u,v) \in E$, and $\tilde{\mu}(u,v) = 0$, otherwise. Furthermore, we extend our abbreviate notation, which was introduced in Section 2 for node weights, so that it applies also to edge weights. More in detail, for a subset of edges $M \subseteq E$, by μ_M we will understand edge weights μ with the domain restricted to M and by μ_{-M} the domain restricted to $E \setminus M$. If M contains one element, i.e., $M = \{e\}$, we will skip parentheses and simply write μ_e and μ_{-e}. Also, for a constant, $x \in \mathbb{R}_{>0}$, we define $x \cdot \mu$ as follows: $(x \cdot \mu)(e) = x \cdot \mu(e)$, for every $e \in E$. Moreover for every two edge weights with possibly different domains, $\mu : E \to \mathbb{R}_{>0}$ and $\mu' : E' \to \mathbb{R}_{>0}$, we define $\mu + \mu' : E \cup E' \to \mathbb{R}_{>0}$ as $(\mu + \mu')(e) = \tilde{\mu}(e) + \tilde{\mu}'(e)$, for every $e \in E \cup E'$. For example, $(\mu_{-e} + 2\mu_e)$ are edge weights obtained from μ by doubling the weight of edge e. Finally, for two weights $\theta = (b,\mu)$ and $\theta' = (b',\mu')$ their sum is defined as $\theta + \theta' = (b + b', \mu + \mu')$.

If not stated otherwise, all definition from Section 2 still hold. Those that were based on μ_G, e.g., definitions of centrality measures, now use $\tilde{\mu}$ instead of μ_G. The only definition that has to be changed is the definition of *out-degree*. Previously, it counted the number of outgoing edges. Since, the number of outgoing edges is no longer directly connected to the total weight of these edges, we define the out-degree as the total weight of outgoing edges, i.e., $\deg^+_v(G,\theta) = \sum_{e \in \Gamma^+_v(G)} \mu(e)$.

For any $x \in \mathbb{R}_{>0}$, graph (G,θ) is *x-out-regular* if the out-degree of every node in the graph equals x, i.e., $\deg^+_v(G,\theta) = x$, for every $v \in V$. Note that in an x-out-regular graph, the principle eigenvalue is always $\lambda = x$. A weighted graph is *out-regular* if it is x-out-regular for some $x \in \mathbb{R}_{>0}$.

Next, let us define *proportional combining*, i.e., a generalization of node redirection that allows for combining any two nodes, not only out-twins. For centrality measure, F, weighted graph, $(G,\theta) = ((V,E),(b,\mu))$, and two nodes, $u, w \in V$, proportional combining of node u into w results in a graph $C^F_{u\to w}(G,\theta)$ that is obtained in two steps:

- scaling weights of outgoing edges of u and w proportionally to the centralities of their starts, i.e., multiplying weights of outgoing edges of u by $F_u(G,\theta)/(F_u(G,\theta) + F_w(G,\theta))$ and by $F_w(G,\theta)/(F_u(G,\theta) + F_w(G,\theta))$ outgoing edges of w (if, in the result of this scaling, an edge would have zero weight, then it is removed); and
- merging node u into node w, i.e., deleting node u, transferring its incoming and outgoing edges to node w and adding the weight of node u to node w.

See Fig. 5.1 for an illustration. Observe that for every centrality measure F, if nodes u and w are out-twins, then proportional combining indeed reduces to redirection as defined in Section 2.

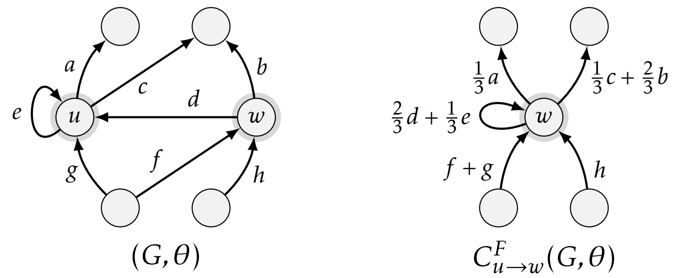


Figure 5.1: An example of a weighted graph, (G,θ), and the corresponding graph obtained through proportional combining of node u into w, i.e., $C^F_{u\to w}(G)$, assuming that $F_u(G) = 1$ and $F_w(G) = 2$. The weight of each edge is shown.

To give an intuition behind proportional combining, consider a network in which each node represents an Internet domain and for two domains, A and B, the weight of an edge from A to B is equal to the average number of links on page in domain A that point to a page in domain B. If there are no such links at all, then there is no edge. In such a network, if the centrality of two domains, A and B, is equal (or proportional) to the number of pages in these domains, then transferring all pages from domain A to domain B (preserving their links and backlinks), would result in proportional combining of node A into B.

Finally, let us introduce a class of cycle graphs. We will call a graph a *cycle graph* if it consists of exactly one directed cycle and there are no edges outside of the cycle. Formally, it is of the form $G = (\{v_1,\dots,v_k\},\{(v_1,v_2),\dots,(v_{k-1},v_k),(v_k,v_1)\})$.

5.2 Axioms

In this section, we present seven axioms used in our axiomatic characterization.

All centrality measures considered in this chapter, except for PageRank, are defined only for a subclass of all graphs. Specifically, eigenvector centrality is defined on $\mathcal{G}^{EV}$, sums of disjoint strongly connected graphs with the same principal eigenvalues, Seeley index is defined on $\mathcal{G}^{SI}$, sums of disjoint strongly connected graphs, and for any decay factor $a \in \mathbb{R}_{\geq 0}$, Katz centrality is defined on $\mathcal{G}^{K(a)}$, graphs for which $\lambda > 1/a$ (for details see Section 2.2.2). Therefore, we will consider restricted versions of our axioms for them. Specifically, an axiom restricted to class $\mathcal{G}^*$ is obtained by adding an assumption that all graphs appearing in the axiom statement belong to $\mathcal{G}^*$. In this way, we obtain a weaker version of the axiom.

Again, most of our axioms are invariance axioms. They identify simple graph operations that do not affect centralities of all or most nodes in a graph. The last two axioms serve as a borderline: they specify centralities in elementary graphs.

We begin with three axioms that are satisfied by all four feedback centralities in question. The first two of them are Locality from Section 4.2 and Edge Deletion from Section 3.1. We recall them below for reader's convenience.

> ***Locality*** *(centrality of a node depends only on the connected component of this node): For every two disjoint graphs $G = (V,E)$, $G' = (V',E')$ and weights θ, θ', it holds that*
>
> $$F_v(G,\theta) = F_v((G+G',\theta+\theta')), \quad \textit{for every } v \in V.$$
>
> ***Edge Deletion*** *(removing an edge from a graph does not affect centralities of nodes which are not successors of the start of this edge): For every graph*

$G = (V,E)$, *weights* $\theta = (b,\mu)$, *and edge* $(u,w) \in E$, *it holds that*

$$F_v(G,\theta) = F_v((V,E \setminus \{(u,w)\}),(b,\mu_{-(u,w)})), \quad \textit{for every } v \in V \setminus S_u(G).$$

The final axiom satisfied by all four considered feedback centralities is inspired by Node Redirect axiom from Section 3.1. However, it allows for combination of nodes that are not necessarily out-twins.

> ***Node Combination*** *(proportional combining of nodes with equal out-degrees and equal out-degrees of successors sums up their centralities and does not affect the centrality of other nodes): For every graph* $G = (V,E)$, *weights* θ, *and nodes* $u,w \in V$ *such that* $\deg^+_u(G,\theta) = \deg^+_w(G,\theta) = \deg^+_s(G,\theta)$ *for every* $s \in S_u(G) \cup S_w(G)$, *it holds that*
>
> $$F_v(G,\theta) = F_v(C^F_{u\to w}(G,\theta)), \quad \textit{for every } v \in V \setminus \{u,w\}$$
>
> *and* $F_u(G,\theta) + F_w(G,\theta) = F_w(C^F_{u\to w}(G,\theta))$.

Assume two nodes $u,w \in V$ and their successors have the same out-degree, but possibly different centralities. Node Combination states that in a graph obtained from proportional combining of u into w, the centrality of w is the sum of centralities of both nodes and centralities of other nodes do not change. This property is characteristic to feedback centralities which associate a benefit from an incoming edge with the importance of a node this edge comes from. We note that PageRank, Seeley index, and Katz centrality also satisfy a relaxed version of the axiom without the assumption about equal out-degrees of successors. The assumption, however, is necessary for eigenvector centrality.

In our next two axioms we consider a modification of one node: its weight and weights of its incident edges. See Fig. 5.2 for an illustration. The first of these axioms is Edge Multiplication from Section 3.1 adapted to the weighted graphs setting

> ***Edge Multiplication*** *(multiplying the weights of the outgoing edges of a node by a constant does not affect the centrality of any node): For every graph* $G = (V,E)$, *weights* $\theta = (b,\mu)$, *node* $u \in V$, *and constant* $x \in \mathbb{R}_{>0}$, *it holds that*
>
> $$F_v(G,\theta) = F_v(G,(b,\mu_{-\Gamma^+_u(G)} + x \cdot \mu_{\Gamma^+_u(G)})), \quad \textit{for every } v \in V.$$

Edge Multiplication is satisfied by both PageRank and Seeley index. However, it is not satisfied by eigenvector and Katz centralities, as it increase the importance of modified edges x times. For them, we propose a similar axiom.

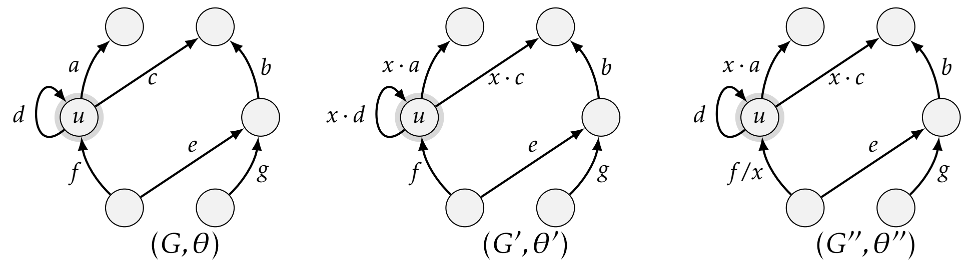


Figure 5.2: An example weighted graph, (G,θ), and the graph obtained from (G,θ) considered in Edge Multiplication, (G',θ'), and Edge Compensation, (G'',θ'').

Edge Compensation *(multiplying the weights of the outgoing edges of a node by x and dividing its weight and the weights of its incoming edges by x divides the centrality of this node by x and does not affect the centrality of other nodes): For every graph* $G = (V, E)$*, weights* $\theta = (b, \mu)$*, node* $u \in V$*, and constant* $x \in \mathbb{R}_{>0}$ *it holds that*

$$F_v(G, \theta) = F_v(G, (b', \mu')), \quad \textit{for every } v \in V \setminus \{u\}$$

and also $F_u(G,\theta)/x = F_u(G,(b',\mu'))$*, where we have that* $b' = b_{-u} + b_u/x$ *and* $\mu' = \mu_{-\Gamma_u^{\pm}(G)\setminus\{(u,u)\}} + \mu_{\Gamma_u^{-}(G)\setminus\{(u,u)\}}/x + \mu_{\Gamma_u^{+}(G)\setminus\{(u,u)\}} \cdot x$.

To provide an intuition for both axioms let us consider a webpage A, on which we can find links to other webpages, e.g., B, but to each such webpage there are exactly two links on A. Now, Edge Multiplication states that if from each such pair of links on A we would remove exactly one, then it does not affect the importance of B (or any other webpage), because each webpage still receives the same share of links from A. For Edge Compensation the situation is different. Such removal of the half of the links on A does not affect the importance of B as long as A itself receives twice as many links from webpages that already have links to A (and the weight, or basic importance, of A also doubles). We can say that the loss of links going from A to B is compensated for B by the increase in links going to A.

Finally, the last two axioms concern simple borderline cases. See Fig. 5.3 for an illustration. The first of them is Baseline introduced in Section 3.1.

Baseline *(the centrality of an isolated node is equal to its weight): For every graph* $G = (V, E)$*, weights* $\theta = (b, \mu)$*, and isolated node* $v \in V$*, it holds that*

$$F_v(G, \theta) = b(v).$$

Baseline is satisfied by PageRank and Katz centrality. However, since there are no isolated nodes in strongly connected graphs, it does not make sense when we restrict the class of graphs to sums of disjoint strongly connected graphs. In such a case we propose the following borderline axiom.

Cycle *(the centrality of a node in an out-regular cycle graph is an average weight of all nodes): For every graph* $G = (V, E)$ *and weights* $\theta = (b, \mu)$ *such that* (G, θ) *is an out-regular cycle graph, it holds that*

$$F_v(G, \theta) = b(G)/|V|, \quad \textit{for every } v \in V.$$

The axiom concerns the simplest strongly connected graph: a cycle. Specifically, if weights of all edges in a cycle graph are equal, then centralities of all

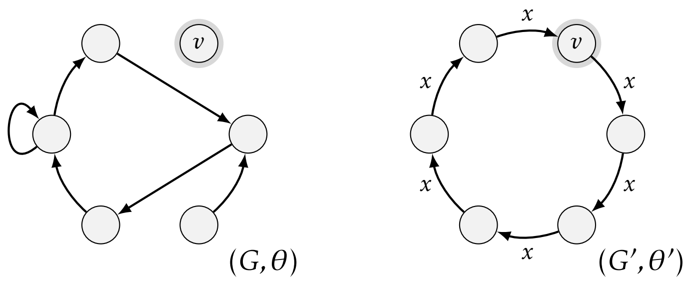


Figure 5.3: Graphs considered in Baseline, (G, θ), and Cycle, (G', θ'), axioms.

nodes are also equal. Moreover, Cycle normalizes the sum of centralities to be equal to the sum of node weights. It is satisfied by both eigenvector centrality and Seeley index.

As we will show, these seven axioms are enough to obtain the axiomatizations of all four feedback centralities.

Theorem 47. *A centrality measure defined on $\mathcal{G}^{SI}$ satisfies Locality, Edge Deletion, Node Combination, Edge Multiplication, and Cycle if and only if it is Seeley index.*

Theorem 48. *A centrality measure defined on $\mathcal{G}^{EV}$ satisfies Locality, Edge Deletion, Node Combination, Edge Compensation, and Cycle if and only if it is eigenvector centrality.*

Theorem 49. *A centrality measure defined on $\mathcal{G}$ satisfies Locality, Edge Deletion, Node Combination, Edge Multiplication, and Baseline if and only if it is PageRank.*

Theorem 50. *For every $a \in \mathbb{R}_{\geq 0}$, a centrality measure defined on $\mathcal{G}^{K(a)}$ satisfies Locality, Edge Deletion, Node Combination, Edge Compensation, and Baseline if and only if it is Katz centrality.*

5.3 Proofs of Uniqueness

In this section, we provide formal proofs of Theorems 47–50, i.e., we show that each of the four feedback centralities: eigenvector centrality, Katz centrality, Seeley index, and PageRank, is uniquely characterized by the respective subset of axioms. Since the proofs for Seeley index and eigenvector centrality have similar structure, much the same as the proofs for PageRank and Katz centrality, first we focus on the two former centralities, i.e., Theorems 47 and 48, and then on the latter two centralities, i.e., Theorems 49 and 50.

5.3.1 Seeley Index and Eigenvector Centrality (Theorems 47 and 48)

In this section, we prove that Seeley index is uniquely characterized by Locality, Edge Deletion, Node Combination, Edge Multiplication, and Cycle and eigenvector centrality by Locality, Edge Deletion, Node Combination, Edge Compensation, and Cycle. More in detail, we begin with Lemmas 51–53 devoted to showing that both centrality measures satisfy respective axioms. Then, we move to the proof that our axioms imply Seeley index and eigenvector centralities in Lemmas 54–58.

First, in the following lemma we prove that for every out-regular graph both centrality measures assign the same value to every node.

Lemma 51. *For every graph $G = (V, E)$ and weights $\theta = (b, \mu)$ such that (G, θ) is out-regular and $(G, \theta) \in \mathcal{G}^{SI}$ it holds that*

$$EV_v(G, \theta) = SI_v(G, \theta), \quad \textit{for every } v \in V.$$

Proof. Consider an arbitrary graph $G = (V, E)$ and weights $\theta = (b, \mu)$ such that graph (G, θ) is x-out-regular and $(G, \theta) \in \mathcal{G}^{SI}$. Observe that (G, θ) is a sum of disjoint x-out-regular strongly connected parts. Thus, the principal eigenvalue of each part is x. Hence, (G, θ) is a sum of disjoint strongly connected graphs with principal eigenvalue x, i.e., $(G, \theta) \in \mathcal{G}^{EV}$. If $b(G) = 0$, then $SI_v(G, \theta) = 0 = EV_v(G, \theta)$, for every $v \in V$. Thus, assume that $b(G) > 0$.

Now, since principal eigenvalue of (G,θ) is $\lambda = x$, from Eq. (2.6) and Eq. (2.12) we get that

$$w^{1/\lambda}_{G,\theta}(v,t) = \sum_{\omega\in\Omega_t(G):\omega(t)=v} \frac{b(\omega(0))}{b(G)} \cdot \prod_{i=0}^{t-1} \frac{\mu(\omega(i),\omega(i+1))}{x} = p^1_{G,\theta}(v,t),$$

for every $v \in V$. Thus, Eq. (2.8) and Eq. (2.14) yield $EV_v(G,\theta) = SI_v(G,\theta)$. □

In the next two lemmas, we show that Seeley index and eigenvector centrality satisfy the corresponding sets of axioms. We begin with Seeley index.

Lemma 52. *Seeley index defined on $\mathcal{G}^{SI}$ by Eq.* (2.8) *satisfies Locality, Edge Deletion, Node Combination, Edge Multiplication, and Cycle.*

Proof. Let us take an arbitrary graph $G = (V,E)$ and weights $\theta = (b,\mu)$ such that $(G,\theta) \in \mathcal{G}^{SI}$, and consider the axioms one by one.

- For Locality, let us consider graph $G' = (V',E')$ such that $V \cap V' = \emptyset$, weights $\theta' = (b',\mu')$, and an arbitrary node $v \in V$. Observe that if we have $b(G) = 0$, then $SI_v(G,\theta) = 0 = SI_v(G+G',\theta+\theta')$. Hence, assume that $b(G) > 0$. Then, observe that in graph $(G+G',\theta+\theta')$ any walk that starts at one of the nodes in V' cannot visit nodes in V and vice versa. Thus, for every $t \in \mathbb{N}$, we have that $\{\omega \in \Omega_t(G+G') : \omega(t) = v\} = \{\omega \in \Omega_t(G) : \omega(t) = v\}$. Moreover, weights of edges in E and weights and out-degrees of nodes in V are the same in both graphs (G,θ) and $(G+G',\theta+\theta')$. Therefore, from Eq. (2.6) we obtain that $p^1_{G+G',\theta+\theta'}(v,t)/(b(G)+b'(G')) = p^1_{G,\theta}(v,t)/b(G)$. Hence, Locality follows from Eq. (2.8).

- For Edge Deletion, consider edge $(u,w) \in E$, graph $G' = (V, E \setminus \{(u,w)\})$, and weights $\theta' = (b,\mu_{-(u,w)})$. Fix an arbitrary node $v \in V \setminus S_u(G)$. Observe that for $G \in \mathcal{G}^{SI}$ such a node exists only if nodes u and v belong to different connected components, i.e., there exist graphs $(G_v,\theta_v) = ((V_v,E_v),(b_{V_v},\mu_{E_v}))$ and $(G_u,\theta_u) = ((V_u,E_u),(b_{V_u},\mu_{E_u}))$ such that $V_v \cap V_u = \emptyset$ and $G_v + G_u = G$. Since Seeley index satisfies Locality, we get that $SI_v(G,\theta) = SI_v(G_v,\theta_v)$. Now, let us take graph $G'_u = (V_u, E_u \setminus \{(u,w)\})$ and weights $\theta'_u = (b_{V_u},\mu_{E_u\setminus\{(u,w)\}})$. Observe that $(G_v+G'_u,\theta_v+\theta'_u) = (G',\theta')$. Thus, again from Locality we get that $SI_v(G_v,\theta_v) = SI_v(G',\theta')$ and Edge Deletion follows.

- For Node Combination, take $u,w \in V$ such that $\deg^+_u(G) = \deg^+_w(G) = \deg^+_s(G)$, for every $s \in S_u(G) \cup S_w(G)$. Let $(G',\theta') = ((V',E'),(b',\mu')) = C^{SI}_{u\to w}(G,\theta)$. Since Seeley index satisfies Locality, without loss of generality let us assume that G is strongly connected. Then, from Theorem 2 we know that Seeley index can be equivalently defined as the solution to Seeley index recursive equation (Eq. (2.4)) and normalization condition $\sum_{v\in V} SI_v(G) = b(G)$. Observe that proportional combining does not affect the sum of node weights in a graph, i.e., $b(G) = b'(G')$. Therefore, it suffices to show that $(x_v)_{v\in V\setminus\{u\}}$ defined as $x_v = SI_v(G,\theta)$, for every $v \in V \setminus \{u,w\}$, and $x_w = SI_u(G,\theta) + SI_w(G,\theta)$ satisfies Seeley index recursive equation (Eq. (2.4)), for every $v \in V \setminus \{u\}$ and graph (G',θ').

  To this end, fix $v \in V$ and observe that from Seeley index recursive equation (Eq. (2.4)) for graph (G,θ) we have

$$SI_v(G,\theta) = \frac{\tilde{\mu}(u,v)SI_u(G,\theta)}{\deg^+_u(G,\theta)} + \frac{\tilde{\mu}(w,v)SI_w(G,\theta)}{\deg^+_w(G,\theta)} + \sum_{s\in P^1_v(G)\setminus\{u,w\}} \frac{\mu(s,v)SI_s(G,\theta)}{\deg^+_s(G,\theta)}. \tag{5.1}$$

Recall that $\deg^+_u(G,\theta) = \deg^+_w(G,\theta)$. Also, proportional combining does not affect out-degrees, so $\deg^+_s(G,\theta) = \deg^+_s(G',\theta')$, for every $s \in V \setminus \{u\}$. Now, assume that $v \in V \setminus \{u,w\}$. Then, by the definition of proportional combining, $\mu'(w,v) = (SI_u(G,\theta)\tilde{\mu}(u,v) + SI_w(G,\theta)\tilde{\mu}(w,v))/(SI_u(G,\theta) + SI_w(G,\theta))$. Thus, Eq. (5.1) can be transformed into

$$SI_v(G,\theta) = \frac{\tilde{\mu}'(w,v)x_w}{\deg^+_w(G',\theta')} + \sum_{s \in P^1_v(G)\setminus\{u,w\}} \frac{\mu(s,v)SI_s(G,\theta)}{\deg^+_s(G',\theta')}. \tag{5.2}$$

The incoming edges of v that come from nodes other then u and w are unaffected by proportional combining. Thus, $P^1_v(G) \setminus \{u,w\} = P^1_v(G') \setminus \{u,w\}$ and $\mu(s,v) = \mu'(s,v)$, for every $s \in P^1_v(G) \setminus \{u,w\}$. Hence, from Eq. (5.2) we obtain

$$x_v = \sum_{s \in P^1_v(G')} \frac{\mu'(s,v)}{\deg^+_s(G',\theta')} \cdot x_s,$$

which is exactly the Seeley index recursive equation for graph (G',θ') and node v.

Therefore, it remains to consider node w. To this end, let us add sidewise Eq. (5.1) for $v = u$ and $v = w$. From the definition of proportional combining

$$\mu'(w,w) = \frac{SI_u(G,\theta)(\tilde{\mu}(u,u) + \tilde{\mu}(u,w)) + SI_w(G,\theta)(\tilde{\mu}(w,u) + \tilde{\mu}(w,w))}{SI_u(G,\theta) + SI_w(G,\theta)}.$$

Observe that the other incoming edges to u and w are simply combined, i.e., $P^1_w(G') \setminus \{w\} = P^1_u(G) \cup P^1_w(G) \setminus \{u,w\}$ and $\mu'(s,w) = \tilde{\mu}(s,u) + \tilde{\mu}(s,w)$, for every $s \in P^1_w(G') \setminus \{w\}$. Again, $\deg^+_u(G,\theta) = \deg^+_w(G,\theta)$ and $\deg^+_s(G,\theta) = \deg^+_s(G',\theta')$, for every $s \in V \setminus \{u\}$. Thus, from Eq. (5.1) we obtain

$$x_w = SI_u(G,\theta) + SI_w(G,\theta) = \frac{\tilde{\mu}'(w,w)x_w}{\deg^+_w(G',\theta')} + \sum_{s \in P^1_w(G')\setminus\{u,w\}} \frac{\mu'(s,v)}{\deg^+_s(G',\theta')} x_s,$$

which is the Seeley index recursive equation for graph (G',θ') and node w.

- For Edge Multiplication, take arbitrary nodes $u,v \in V$, constant $x \in \mathbb{R}_{>0}$, and weights $\theta' = (b, \mu_{-\Gamma^+_u(G)} + x \cdot \mu_{\Gamma^+_u(G)})$. If $b(G) = 0$, then $SI_v(G,\theta) = 0 = SI_v(G,\theta')$. Thus, assume that $b(G) > 0$. Observe that for every $t \in \mathbb{N}$ and walk $\omega \in \Omega_t(G)$ such that $\omega(t) = v$, we have

$$\frac{b(\omega(0))}{b(G)} \cdot \prod_{i=0}^{t-1} \frac{\mu(\omega(i),\omega(i+1))}{\deg^+_{\omega(i)}(G,\theta)} = \frac{b(\omega(0))}{b(G)} \cdot \prod_{i=0}^{t-1} \frac{\mu'(\omega(i),\omega(i+1))}{\deg^+_{\omega(i)}(G,\theta')}$$

  It holds because for every $i \in \{0,\dots,t-1\}$, if $\omega(i) = u$, then both the numerator and the denominator are multiplied by x, and if $\omega(i) \neq u$, both the numerator and the denominator are unaffected. Thus, Eq. (2.6) yields $p^1_{G,\theta}(v,t) = p^1_{G,\theta'}(v,t)$. Summing for all $t \in \mathbb{N}$, we get that $SI_v(G,\theta) = SI_v(G,\theta')$ from Eq. (2.8). Hence, Edge Multiplication is satisfied.

- Finally, for Cycle, observe that if graph (G,θ) is a cycle graph, then every node $v \in V$ has exactly one direct predecessor, say u, i.e., $P^1_v(G) = \{u\}$. Hence, from Eq. (2.4) we get that $SI_v(G,\theta) = SI_u(G,\theta)$. Thus, all nodes have equal centralities. From Theorem 2 we get that $b(G) = \sum_{v \in V} SI_v(G,\theta) = |V| \cdot SI_v(G)$. Hence, $SI_v(G,\theta) = b(G)/|V|$.

□

Now, let us focus on eigenvector centrality and prove that it satisfies the corresponding axioms.

Lemma 53. *Eigenvector centrality defined on $\mathcal{G}^{EV}$ by Eq.* (2.14) *satisfies Locality, Edge Deletion, Node Combination, Edge Compensation, and Cycle.*

Proof. Let us take an arbitrary graph, $G = (V,E)$, and weights, $\theta = (b,\mu)$ such that $(G,\theta) \in \mathcal{G}^{EV}$, and consider the axioms one by one.

- For Locality, let us consider graph $G' = (V',E')$ and weights $\theta' = (b',\mu')$ such that $V \cap V' = \emptyset$ and $(G',\theta'),(G+G',\theta+\theta') \in \mathcal{G}^{EV}$. Fix $v \in V$. If $b(G) = 0$, then trivially $EV_v(G,\theta) = 0 = EV_v(G+G',\theta+\theta')$. Thus, assume $b(G) > 0$. Observe that in graph $(G+G',\theta+\theta')$ any walk that starts at one of the nodes in V' cannot visit nodes in V and vice versa. Thus, for every $t \in \mathbb{N}$, we have that $\{\omega \in \Omega_t(G+G') : \omega(t) = v\} = \{\omega \in \Omega_t(G) : \omega(t) = v\}$. Moreover, weights of edges in E and weights and out-degrees of nodes in V are the same in both graphs (G,θ) and $(G+G',\theta+\theta')$. Therefore, from Eq. (2.6) we obtain that $w^{1/\lambda}_{G+G',\theta+\theta'}(v,t)/(b(G)+b'(G')) = w^{1/\lambda}_{G,\theta}(v,t)/b(G)$. Since $(G+G',\theta+\theta') \in \mathcal{G}^{EV}$, we know that principal eigenvalue λ is the same in both (G,θ) and $(G+G',\theta+\theta')$. Hence, Locality follows from Eq. (2.8).

- For Edge Deletion, consider edge $(u,w) \in E$, graph $G' = (V, E \setminus \{(u,w)\})$, and weights $\theta' = (b, \mu_{-(u,w)})$. Fix an arbitrary node $v \in V \setminus S_u(G)$. Observe that for $G \in \mathcal{G}^{EV}$ such a node exists only if u and v belong to different strongly connected components, i.e., there exist graphs $(G_v,\theta_v) = ((V_v,E_v),(b_{V_v},\mu_{E_v}))$ and $(G_u,\theta_u) = ((V_u,E_u),(b_{V_u},\mu_{E_u}))$ such that $V_v \cap V_u = \emptyset$ and $G_v + G_u = G$. Since eigenvector centrality satisfies Locality, we get that $EV_v(G,\theta) = EV_v(G_v,\theta_v)$. Now, take graph $G'_u = (V_u, E_u \setminus \{(u,w)\})$ and weights $\theta'_u = (b_{V_u}, \mu_{E_u \setminus \{(u,w)\}})$. Observe that $(G_v + G'_u, \theta_v + \theta'_u) = (G',\theta')$. Thus, again from Locality we get that $EV_v(G_v,\theta_v) = EV_v(G',\theta')$ and Edge Deletion follows.

- For Node Combination, take $u,w \in V$ such that $\deg^+_u(G) = \deg^+_w(G) = \deg^+_s(G)$, for every $s \in S_u(G) \cup S_w(G)$. Let $(G',\theta') = C^{EV}_{u\to w}(G,\theta)$.

  First, assume each node in G is a successor of u or w, i.e., $V = S_u(G) \cup S_w(G)$. This means that (G,θ) is out-regular. Thus, from Lemma 51 we obtain that $EV_v(G,\theta) = SI_v(G,\theta)$, for every $v \in V$. Since proportional combining preserves out-regularity, (G',θ') is also out-regular. Thus, by Lemma 51, also $EV_v(G',\theta') = SI_v(G',\theta')$. Seeley index satisfies Node Combination (Lemma 52), hence $EV_v(G',\theta') = SI_v(G',\theta') = SI_v(G,\theta) = EV_v(G,\theta)$, for every $v \in V \setminus \{u,w\}$, and $EV_w(G',\theta') = SI_w(G',\theta') = SI_u(G,\theta) + SI_w(G,\theta) = EV_u(G,\theta) + EV_w(G,\theta)$.

  Now, assume that there is at least one node in V that is not a successor of u and not a successor of w. Since $(G,\theta) \in \mathcal{G}^{EV}$, it is a sum of disjoint strongly connected graphs. Hence, it can be expressed as a sum of two disjoint graphs, i.e., $(\hat{G},\hat{\theta}) = ((\hat{V},\hat{E}),(b_{\hat{V}},\mu_{\hat{E}}))$ and $(\tilde{G},\tilde{\theta}) = ((\tilde{V},\tilde{E}),(b_{\tilde{V}},\mu_{\tilde{E}}))$, such that we have $\hat{V} = S_u(G) \cup S_w(G)$, $\hat{V} \cap \tilde{V} = \emptyset$, and $\hat{G} + \tilde{G} = G$. Also, let $(\hat{G}',\hat{\theta}') = C^F_{u\to w}(\hat{G},\hat{\theta})$. Now, in graph $(\hat{G},\hat{\theta})$ all nodes are successors of either u or w. Thus, from the previous case we get that $EV_v(\hat{G},\hat{\theta}) = EV_v(\hat{G}',\hat{\theta}')$, for every $v \in V \setminus \{u,w\}$, and also, for node w, we have $EV_w(\hat{G},\hat{\theta}) = EV_u(\hat{G}',\hat{\theta}') + EV_w(\hat{G}',\hat{\theta}')$. Moreover, observe that $\hat{G}' + \tilde{G} = G'$. Hence, Node Combination follows from Locality.

- For Edge Compensation, take arbitrary $u \in V$ and $x \in \mathbb{R}_{>0}$, and let $\theta' = (b',\mu')$, where $b' = b_{-u} + b_u/x$ and $\mu' = \mu_{-\Gamma^{\pm}_u(G)\setminus\{(u,u)\}} + \mu_{\Gamma^+_u(G)\setminus\{(u,u)\}} \cdot x + \mu_{\Gamma^-_u(G)\setminus\{(u,u)\}}/x$.

Fix $v \in V \setminus \{u\}$. If $b(G) = 0$, then trivially $EV(G,\theta) = 0 = EV(G,\theta')$. Thus, assume that $b(G) > 0$. We will show that for every $t \in \mathbb{N}$ and walk $\omega \in \Omega_t(G)$ that ends at v, i.e., $\omega(t) = v$, it holds that

$$b(\omega(0)) \cdot \prod_{i=1}^{t} \frac{1}{\lambda} \mu(\omega(i-1), \omega(i)) = b'(\omega(0)) \cdot \prod_{i=1}^{t} \frac{1}{\lambda} \mu'(\omega(i-1), \omega(i)). \quad (5.3)$$

To this end, observe that since $\omega(t) \neq u$, for every step $i \in \{0, \ldots, t-1\}$ in which the walk arrives at node u, i.e., $\omega(i-1) \neq u$ and $\omega(i) = u$ (if the walk starts in u, we treat step 0 as arrival as well), there exist exactly one step $j > i$ in which the walk departs from u, i.e., $\omega(k) = u$, for every $k \in \{i, \ldots, j-1\}$, and $\omega(j) \neq u$. Now, in (G,θ'), the factor for step i (arrival) decreases by x, i.e., $\mu'(\omega(i-1), \omega(i)) = \mu(\omega(i-1), \omega(i))/x$ (or $b'(i) = b(i)/x$ if $i = 0$), but at the same time it holds that the factor for step j (departure) increases by x, i.e., $\mu'(\omega(j-1), \omega(j)) = \mu(\omega(j-1), \omega(j)) \cdot x$. Since there is equal number of arrivals and departures from u, Eq. (5.3) holds. Thus, from Eq. (2.12) and Eq. (5.3) we have that $w^{1/\lambda}_{G,\theta}(v,t)/b(G) = w^{1/\lambda}_{G,\theta'}(v,t)/b'(G)$ and from Eq. (2.14) we get $EV_v(G,\theta) = EV_v(G,\theta')$.

By similar reasoning as in Eq. (5.3), for node u we can obtain that for every $t \in \mathbb{N}$ and $\omega \in \Omega_t(G)$ that ends at u, i.e., $\omega(t) = u$, it holds that

$$b(\omega(0)) \cdot \prod_{i=1}^{t-1} \frac{1}{\lambda} \mu(\omega(i-1), \omega(i))/x = b'(\omega(0)) \cdot \prod_{i=1}^{t-1} \frac{1}{\lambda} \mu'(\omega(i-1), \omega(i)).$$

Here, there is a departure step for every but last arrival at u. However, since $\omega(t) = u$, for the last arrival, there is no departure. Hence, for weights θ' the product is divided by x one more time than it is multiplied by x. Thus, from Eq. (2.12) we get that $w^{1/\lambda}_{G,\theta'}(v,t)/b'(G) = w^{1/\lambda}_{G,\theta}(v,t)/b(G)/x$. Hence, from Eq. (2.14) we get that $EV_u(G,\theta') = EV_u(G,\theta)/x$.

- Finally, for Cycle, assume that (G,θ) is out-regular cycle graph. Then, from Lemmas 51 and 52, for every $v \in V$, we have $EV_v(G,\theta) = SI_v(G,\theta) = b(G)/|V|$.

□

Now, let us move to the second part of the proof in which we show that the corresponding sets of axioms imply Seeley index and eigenvector centrality. To this end, we first prove a simple property of Seeley index and eigenvector centrality that multiplying the node weights by a constant results in the multiplication of centrality of every node by the same constant (Lemma 54). Next, we consider an arbitrary centrality measure, F, that satisfies Locality, Edge Deletion, Node Combination, and Cycle, i.e., common axioms for Seeley index and eigenvector centrality. We prove that F gives the same values as Seeley index and eigenvector centrality for every node in strongly connected out-regular graphs in which proportion of weights of any two edges is rational (Lemma 55), and then, in any strongly connected out-regular graphs (Lemma 56). Next, we show that if we assume that F satisfies also Edge Multiplication, we obtain that F is equal to Seeley index for all graphs in $\mathcal{G}^{SI}$ (Lemma 57). Similarly, if F satisfies also Edge Compensation, then F is equal to eigenvector centrality for all graphs in $\mathcal{G}^{EV}$ (Lemma 58).

We begin with a simple lemma that captures a useful property of Seeley index and eigenvector centrality: multiplying node weights by a constant, multiplies centralities by the same constant.

Lemma 54. *For every graph $G = (V,E)$ and weights $\theta = (b,\mu)$ such that $(G,\theta) \in \mathcal{G}^{SI}$ it holds that*

$$SI_v(G,(x \cdot b,\mu)) = x \cdot SI_v(G,\theta), \quad \textit{for every } v \in V$$

and $EV_v(G,(x \cdot b,\mu)) = x \cdot EV_v(G,\theta)$, for every $v \in V$, if additionally $(G,\theta) \in \mathcal{G}^{EV}$.

Proof. Consider an arbitrary graph $G = (V,E)$ and weights $\theta = (b,\mu)$ such that $(G,\theta) \in \mathcal{G}^{SI}$. Let $\theta' = (x \cdot b,\mu)$. Fix node $v \in V$. If $b(G) = 0$, then observe that $SI_v(G,\theta) = 0 = SI_v(G,\theta')$ and the same is true for eigenvector centrality. Thus, let us assume that $b(G) > 0$. Then, observe that for every walk $\omega \in \Omega_t(G)$ such that $\omega(t) = v$, we have that

$$\frac{b(\omega(0))}{b(G)} \cdot \prod_{i=0}^{t-1} \frac{\mu(\omega(i),\omega(i+1))}{\deg^+_{\omega(i)}(G,\theta)} = \frac{x \cdot b(\omega(0))}{x \cdot b(G)} \cdot \prod_{i=0}^{t-1} \frac{\mu(\omega(i),\omega(i+1))}{\deg^+_{\omega(i)}(G,\theta')}.$$

Thus, from Eq. (2.6) we have that $p^1_{G,\theta}(v,t) = p^1_{G,\theta'}(v,t)$. Hence, from Eq. (2.8) we obtain that $SI(G,\theta) \cdot x = SI(G,\theta')$. The same result can be obtained for eigenvector centrality from Eq. (2.12) and Eq. (2.14). □

In the following lemma we focus on strongly connected out-regular graph with positive sum of node weights, i.e, $b(G) > 0$, in which the proportions of weights of any two edges is rational. We prove that if a centrality measure satisfies four of our axioms, then it is equal to Seeley index and eigenvector centrality for each node in each such graph.

Lemma 55. *If a centrality measure, F, defined on $\mathcal{G}^{SI}$ (or $\mathcal{G}^{EV}$) satisfies Locality, Edge Deletion, Node Combination, and Cycle, then for every $\lambda > 0$ and every strongly connected graph $G = (V,E)$ and weights $\theta = (b,\mu)$ such that $(G,\theta) \in \mathcal{G}^{SI}$, (G,θ) is λ-out regular, $b(G) \neq 0$, and $\mu(e)/\mu(e') \in \mathbb{Q}$, for every $e,e' \in E$, it holds that*

$$F_v(G,\theta) = SI_v(G,\theta) = EV_v(G,\theta), \quad \textit{for every } v \in V. \tag{5.4}$$

Proof. First, let us assume that $b(G) = 1$. We will relax this assumption at the end of the proof.

The second equality of Eq. (5.4), i.e., that $SI_v(G,\theta) = EV_v(G,\theta)$, follows from Lemma 51. Thus, let us focus on proving that $F_v(G,\theta) = SI_v(G,\theta)$, for every $v \in V$. To this end, let us first define *impact* of an edge. For every strongly connected weighted graph $(G,\theta) = ((V,E),(b,\mu))$ and edge $(u,v) \in E$ let the impact of (u,v) be equal to $I_{G,\theta}(u,v) = SI_u(G,\theta) \cdot \mu(u,v)/\deg^+_u(G,\theta)$ (see Fig. 5.4 for illustration). Intuitively, impact measures the amount of centrality that node u transfers to node v. Indeed, from Seeley index recursive equation (Eq. (2.4)) we see that the centrality of a node is equal to both the sum of impacts of its outgoing edges and the sum of impacts of its incoming edges, i.e.,

$$\sum_{e \in \Gamma^-_v(G)} I_{G,\theta}(e) = SI_v(G,\theta) = \sum_{e \in \Gamma^+_v(G)} I_{G,\theta}(e). \tag{5.5}$$

Another property that we will use in the proof is that proportional combining preserves the impact of edges. More in detail, for any $(u,v),(u',v') \in E$ such that $v' \notin \{u,u'\}$ and graph $(G',\theta') = C^{SI}_{u' \to u}(G,\theta)$ we have

$$I_{G',\theta'}(u,v) = \begin{cases} I_{G,\theta}(u,v), & \text{if } v \neq v', \\ I_{G,\theta}(u,v) + I_{G,\theta}(u',v'), & \text{otherwise.} \end{cases} \tag{5.6}$$

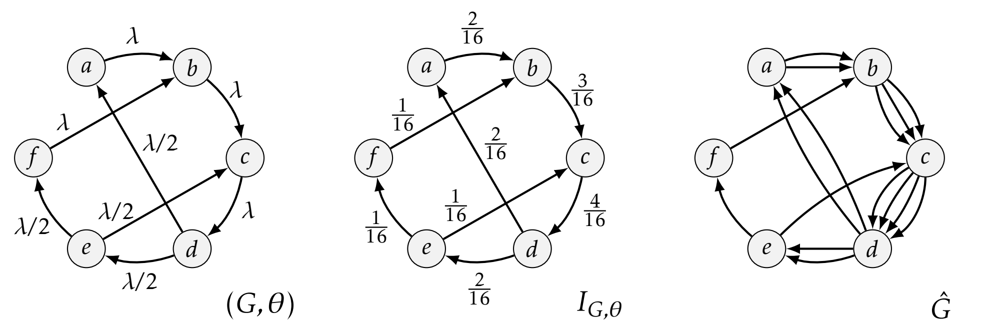


Figure 5.4: An illustration to the proof of Lemma 55. The leftmost graph, (G,θ), is a λ-out-regular graph with the weight of each edge shown. The middle graph is graph (G,θ) as well, but with impact of each edge shown instead of its weight. The rightmost graph, $\hat{G}$, is an unweighted multigraph obtained from (G,θ).

This comes from the fact, that the node resulting from the combination has the weight of its outgoing edges decreased by $SI_u(G,\theta)/(SI_u(G,\theta)+SI_{u'}(G,\theta))$, but at the same time its centrality increases by the inverse value. Hence, the impact of its outgoing edges is unaffected.

From Theorem 2 we know that Seeley index of nodes in V can be equivalently defined as the solution to the system of Seeley index recursive equations (Eq. (2.4)) and normalization equation $\sum_{v\in V} SI_v(G,\theta) = b(G)$. Observe that since proportions of the weights of edges are rational, i.e., $\mu(e)/\mu(e') \in \mathbb{Q}$, for every $e,e' \in E$, then also the coefficients in the system of equations, i.e., $\mu(u,v)/\deg^+_u(G)$, are rational (they are reciprocal of $\deg^+_u(G)/\mu(u,v)$, which are the sums of proportions $\mu(e)/\mu(u,v)$ for all $e \in \Gamma^+_u(G)$). If also $b(G) = 1$, then all coefficients are rational and the solution, i.e., $SI_v(G,\theta)$, for every $v \in V$, is rational. Moreover, since both $SI_v(G,\theta)$, for every $v \in V$, and $\mu(u,v)/\deg^+_u(G)$, for every $(u,v) \in E$, are rational, the impact of every edge is rational as well, i.e., $I_{G,\theta}(e) \in \mathbb{Q}$, for every $e \in E$. Building upon this, we will consider a walk on graph G that follows each edge the number of times that is proportional to its impact. Next, we will construct a cycle graph based on this walk and by proportional combining of its nodes transform it into the original graph G. Hence, based on Cycle and Node Combination, we will establish centrality F of each node.

To this end, observe that since the impact of each edge is rational, there exist $N \in \mathbb{N}$ such that for every $e \in E$, the product $N \cdot I_{G,\theta}(e)$ is an integer. Building upon this, let us define an auxiliary unweighted multigraph, $\hat{G} = (V,\hat{E})$. Its nodes are the nodes of graph G and its multiset of edges, $\hat{E} = (E,m)$, consists of edges in E with the multiplicity of each edge $e \in E$ equal to $m(e) = N \cdot I_{G,\theta}(e)$ (see Fig. 5.4 for illustration). Observe that

$$|\hat{E}| = \sum_{e\in E} m(e) = N \cdot \sum_{e\in E} I_{G,\theta}(e) = N \cdot \sum_{v\in V} SI_v(G,\theta) = N \cdot b(G) = N.$$

Now, from Eq. (5.5) we get that in $\hat{G}$ every node has equal number of incoming and outgoing edges (when accounted for their multiplicity). Hence, from Euler theorem for directed graphs $\hat{G}$ is an Euler multigraph. This means that there exists an Euler walk $\varepsilon = (\varepsilon(0),\varepsilon(1),\dots,\varepsilon(N))$ of length N in which $\varepsilon(0) = \varepsilon(N)$ and each edge is followed exactly once, i.e,

$$\Big|\{i : \varepsilon(i) = u \ \wedge\ \varepsilon(i+1) = v\}\Big| = m(u,v), \quad \text{for every } (u,v) \in E.$$

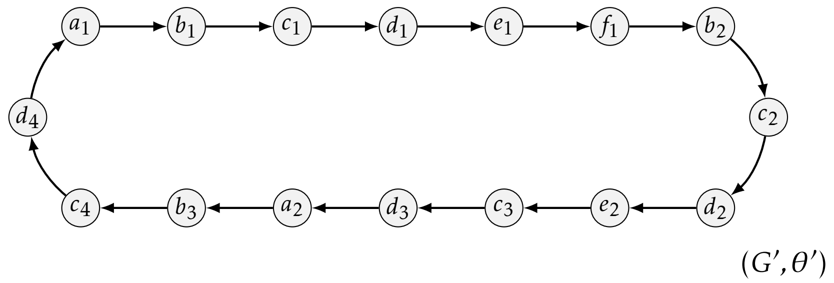


Figure 5.5: Cycle graph, (G',θ'), corresponding to an example Euler cycle on multigraph $\hat{G}$ from Fig. 5.4. By proportional combining of nodes that are labeled with the same letter, we can obtain graph (G,θ) from Fig. 5.4.

For each node $v \in V$ we denote the set of indices on which walk ε visits node v, i.e., let $E_v = \{i \in \{1,\dots,N\} : \varepsilon(i) = v\}$. Observe that the number of visits at v, is equal to the in-degree (or the out-degree as it is equal) of v in multigraph $\hat{G}$. This, in turn, is equal to N times the sum of impacts of its incoming edges in (G,θ), which, by Eq. (5.5), is its centrality, i.e.,

$$|E_v| = \sum_{e\in\Gamma_v^-(G)} m(e) = N \cdot \sum_{e\in\Gamma_v^-(G)} I_{G,\theta}(e) = N \cdot SI_v(G,\theta). \tag{5.7}$$

Next, based on Euler walk ε, let us construct a λ-out-regular cycle graph. To this end, let us consider a set of N pairwise-distinct nodes $V' = \{v_1,\dots,v_N\}$ that will correspond to consecutive steps of ε. For later convenience, let us take them in such a way that some of them are equal to particular nodes from V. More in detail, for every $v \in V$ let node with the index of the first step in which walk ε visits node v, i.e., $\min(E_v)$, be equal to node v, i.e., $v_{\min(E_v)} = v$. Now, the graph is given by $G' = (V',E')$ with weights $\theta = (b',\mu')$, where $E' = \{(v_1,v_2),\dots,(v_{N-1},v_N),(v_N,v_1)\}$, node weights are weights of a particular node visited by ε divided by the total number of its visits, i.e., $b'(v_i) = b(\varepsilon(i))/|E_{\varepsilon(i)}|$, for every $i \in \{1,\dots,N\}$, and $\mu'(e) = \lambda$, for every $e \in E'$. See Fig. 5.5 for illustration. Since (G',θ') is indeed a λ-out-regular cycle graph, from Cycle axiom we get that $F_v(G',\theta') = b'(G')/N$, for every $v \in V'$. Observe that $b'(G') = \sum_{v\in V'} b'(v) = \sum_{v\in V} |E_v| \cdot b(v)/|E_v| = b(G) = 1$, for every $v \in V'$. Therefore,

$$F_v(G',\theta') = 1/N, \quad \text{for every } v \in V. \tag{5.8}$$

Now, we sequentially combine nodes in G' that correspond to the same node in walk ε to obtain a graph isomorphic to G. More in detail, for every $v \in V$, let us sequentially combine every node in $\{v_i : i \in E_v\} \setminus \{v\}$ into v (recall that v is also v_i with i being the minimal index in E_v). By $G'' = (V'',E'')$ and $\theta'' = (b'',\mu'')$ let us denote the graph and weights resulting from conducting this sequential combining for all nodes $v \in V$. As a result, we obtain

$$F_v(G'',\theta'') = \sum_{i\in E_v} F_{v_i}(G',\theta') = |E_v| \cdot 1/N = SI_v(G,\theta), \quad \text{for every } v \in V, \tag{5.9}$$

where the consecutive equalities come from Node Combination, Eq. (5.7), and Eq. (5.8). Hence, to prove that $F_v(G,\theta) = SI_v(G,\theta)$ it remains to prove that $(G,\theta) = (G'',\theta'')$.

To this end, observe that indeed $V'' = V$, because all other nodes in V' have been combined into one of the nodes in V. As for edges, observe that for any edge $(u,v) \in E''$ there exists $i \in \{1,\dots,N\}$ such that in the construction of graph G'' node

v_{i-1} was combined into u (or $u = v_{i-1}$) and node v_i was combined into v (or $v = v_i$). As a result, $(\varepsilon(i-1), \varepsilon(i)) = (u,v)$, hence $(u,v) \in E$. Converse reasoning is analogous. For node weights, we have $b''(v) = \sum_{v_i : i \in EV_v} b'(v_i) = b(v)$, for every $v \in V$. Finally, for edge weights, fix $(u,v) \in E''$ and recall that proportional combining preserves the impact of edges (Eq. (5.6)). Thus, the impact of edge $(u,v) \in E''$ is the sum of impacts of edges $(v_{i-1}, v_i) \in E'$ such that v_{i-1} has been combined into u (or $v_{i-1} = u$) and v_i into v (or $v_i = v$). There are exactly $m(u,v)$ of such edges and the impact of every edge in graph (G', θ') is equal to $1/N$, thus

$$I_{G'',\theta''}(u,v) = m(u,v)/N = I_{G,\theta}(u,v).$$

Since (u,v) in both (G'', θ'') and (G, θ) has the same impact, u has the same Seeley index in both graphs (Eq. (5.9)), and both graphs are λ-out-regular (combining nodes preserves out-regularity), we get that $\mu''(u,v) = \mu(u,v)$, for every $(u,v) \in E$. Thus, indeed, $(G'', \theta'') = (G, \theta)$ and from Eq. (5.9) we get $F_v(G,\theta) = SI_v(G,\theta)$.

It remains to relax our first assumption that $b(G) = 1$. To this end, we will show that for any $x > 0$ it holds that $F_v(G,(x \cdot b, \mu)) = x \cdot F_v(G,\theta)$. Combined with Lemma 54 this will prove the thesis. Consider graph $(G', \theta'_x) = ((V', E'), (x \cdot b', \mu'))$, i.e., graph (G', θ') with node weights scaled by x. Then, from Cycle axiom we get that $F_v(G', \theta'_x) = x/N$, for every $v \in V'$. Next, we consider sequential proportional combining of nodes that we used to obtain graph (G, θ) from (G', θ') but starting from graph (G', θ'_x) instead of (G', θ'). Observe that after each step of the process, we obtain the graph from the corresponding step of the original process but with node weights and centralities scaled by x. At the end of the process, we obtain graph $(G,(x \cdot b, \mu))$ and Node Combination we obtain that indeed $F_v(G,(x \cdot b, \mu) = x \cdot F_v(G,\theta)$, for every $v \in V$. □

Now, let us move to arbitrary strongly connected out-regular graphs.

Lemma 56. *If a centrality measure, F, defined on $\mathcal{G}^{SI}$ (or $\mathcal{G}^{EV}$) satisfies Locality, Edge Deletion, Node Combination, and Cycle, then for every $\lambda > 0$ and every strongly connected graph $G = (V,E)$ and weights $\theta = (b,\mu)$ such that $(G,\theta) \in \mathcal{G}^{SI}$ and (G,θ) is λ-out-regular, it holds that*

$$F_v(G,\theta) = SI_v(G,\theta) = EV_v(G,\theta), \quad \textit{for every } v \in V.$$

Proof. First, let us restrict attention to graphs with unit node weights multiplied by a constant. Formally, we prove the following claim:

> For every $c, \lambda > 0$, strongly connected graph $G = (V,E)$, node $\hat{v} \in V$, and weights $\theta = (c \cdot \mathbb{1}_{\hat{v}}, \mu)$ such that $(G,\theta) \in \mathcal{G}^{SI}$ and (G,θ) is λ-out-regular, it holds that $F_v(G,\theta) = SI_v(G,\theta) = EV_v(G,\theta)$, for every $v \in V$. (∗)

To this end, for every such graph, let us distinguish one outgoing edge of each node $u \in V$, denoted by e_u, so that: (1) there exist a walk ω that begins with edge e_u, ends at $\hat{v}$ and does not visit u again before reaching $\hat{v}$, (2) among edges satisfying condition (1) the number of other outgoing edges of u with weights that are not a rational multiple of the weight of e_u, i.e., $|\{e \in \Gamma^+_u(G) : \mu(e)/\mu(e_u) \notin \mathbb{Q}\}|$, is minimal (since graph is strongly connected, such edge e_u always exists). By $k_{G,\theta}$ let us denote the sum of the numbers of not-rationally-proportional edges for all nodes, i.e., let

$$k_{G,\theta} = \sum_{u \in V} |\{e \in \Gamma^+_u(G) : \mu(e)/\mu(e_u) \notin \mathbb{Q}\}|.$$

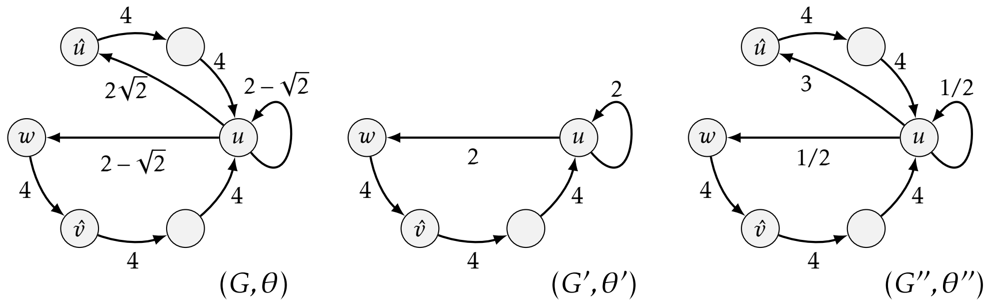


Figure 5.6: An illustration to the proof of Lemma 56. The proportion of weights of edges $(u,\hat{u})$ and $e_u = (u,w)$ in 4-out-regular graph (G,θ) is not rational. Graph (G',θ') is graph (G,θ) with edge $(u,\hat{u})$ removed along with a part of a graph that is then disconnected from $\hat{v}$. Graph (G'',θ'') is graph (G,θ) with edge weights adjusted so that the proportion of weights of $(u,\hat{u})$ and (u,w) is now rational. Intuitively, since $3 > 2\sqrt{2}$, it is possible to take graphs (G',θ') and (G'',θ'') and "combine" them (with their node weights properly scaled) to obtain graph (G,θ).

We will prove the thesis for every strongly connected λ-out-regular graph with unit node weights multiplied by a constant using induction on k_G.

To this end, observe that if $k_{G,\theta} = 0$, then for every node $u \in V$, the weight of each of its outgoing edge, $e \in \Gamma_u^+(G)$, can be written as $\mu(e_u) \cdot q_e$ for some $q_e \in \mathbb{Q}$. Thus,

$$\lambda = \deg_u^+(G,\theta) = \sum_{e \in \Gamma_u^+(G)} \mu(e) = \mu(e_u) \cdot \left(\sum_{e \in \Gamma_u^+(G)} q_e \right).$$

Hence, $\lambda/\mu(e_u) \in \mathbb{Q}$ which implies that also $\lambda/\mu(e) \in \mathbb{Q}$, for every $u \in V$ and every $e \in \Gamma_u^+(G)$. Thus, for every edges $e, e' \in E$, we have $\mu(e)/\mu(e') = (\mu(e)/\lambda) \cdot (\lambda/\mu(e') \in \mathbb{Q}$. As a result, the thesis follows from Lemma 55.

Therefore, let us focus on the case in which $k_{G,\theta} > 0$. Then, there exists a node $u \in V$ and its outgoing edge, $(u,\hat{u}) \in E$, such that $\mu(u,\hat{u})/\mu(e_u) \notin \mathbb{Q}$. In what follows, we will construct two additional graphs: (G',θ'), in which edge $(u,\hat{u})$ is removed (possibly along with a number of nodes), and (G'',θ''), in which the weights of outgoing edges of u are adjusted so that the weight of $(u,\hat{u})$ is a rational multiple of the weight of e_u. See Fig. 5.6 for an illustration. Next, we will construct graph (G,θ) from the combination of (G',θ') and (G'',θ'') and since both $k_{G',\theta'}$ and $k_{G'',\theta''}$ are smaller than $k_{G,\theta}$, this, together with the inductive assumption, will lead us to the induction hypothesis.

Let us begin with graph (G',θ'). Since removing just $(u,\hat{u})$ can result in a graph that is not strongly connected, we remove $(u,\hat{u})$ and all the nodes that would not be in the same strongly connected component of the graph as node $\hat{v}$. Formally, let $G_{-(u,\hat{u})} = (V, E \setminus \{(u,\hat{u})\})$ be a graph with just $(u,\hat{u})$ removed. Observe that since in G there exists a walk that begins with e_u, ends at $\hat{v}$, and does not pass through u before it reaches $\hat{v}$, then after removal of $(u,\hat{u}) \neq e_u$, it is still possible to reach $\hat{v}$ from any other node, i.e., $P_{G_{-(u,\hat{u})}}(\hat{v}) = P_G(\hat{v}) = V$. Thus, we remove exactly these nodes that cannot be reached from $\hat{v}$ without edge $(u,\hat{u})$. Hence, let $V' = S_{G_{-(u,\hat{u})}}(\hat{v})$. Since we remove outgoing edge of u, then for sure $u \in V'$. Moreover, for every node $w \in V' \setminus \{u\}$, its successors in $G_{-(u,\hat{u})}$ are also successors of $\hat{v}$. Thus, we preserve all of the outgoing edges of all nodes in $V' \setminus \{u\}$, i.e., let $E' = \{(s,t) \in E : s \in V'\} \setminus \{(u,\hat{u})\}$. Building upon this, let us define graph $G' = (V', E')$ and weights $\theta' = (c \cdot \mathbb{1}_{\hat{v}}, \mu')$ such that weights of outgoing edges of u are scaled, so that the graph is still λ-out-regular, i.e., $\mu'(e) = \mu(e) \cdot \lambda/(\lambda - \mu(u,\hat{u}))$, for every $e \in \Gamma^+_u(G')$, and the weights of the remaining edges remain unchanged, i.e, $\mu'(e) = \mu(e)$, for every $e \in E' \setminus \Gamma^+_u(G')$. Since $(u,\hat{u})$ is not an edge in G' and the proportions of weights between the remaining outgoing edges of u are unchanged, we obtain that $k_{G',\theta'} < k_{G,\theta}$. Hence, by the inductive assumption,

$$F_v(G',\theta') = SI_v(G',\theta') = EV_v(G',\theta'), \quad \text{for every } v \in V. \tag{5.10}$$

Now, let us construct graph $(G'',\theta'') = ((V,E),(c \cdot \mathbb{1}_{\hat{v}}, \mu''))$ in which weight of edge $(u,\hat{u})$ is scaled by $x > 1$ and the weights of the remaining outgoing edges of u are scaled by $y < 1$ in such a way that: (1) proportion $(x \cdot \mu(u,\hat{u}))/(y \cdot \mu(e_u))$ is now rational, and (2) the sum of the weights of the outgoing edges of u is still equal to λ so that (G'',θ'') is still λ-out-regular. To this end, take any $q \in \mathbb{Q}$ such that $q > \mu(u,\hat{u})/\mu(e_u)$. The new edge weights are given by:

$$\mu''(u,\hat{u}) = q \cdot \frac{\mu(e_u) \cdot \lambda}{\lambda - \mu(u,\hat{u}) + q\mu(e_u)}, \quad \mu''(e) = \frac{\mu(e) \cdot \lambda}{\lambda - \mu(u,\hat{u}) + q\mu(e_u)},$$

for every $e \in \Gamma^+_u(G) \setminus \{(u,\hat{u})\}$, and $\mu''(e) = \mu(e)$, for every $e \in E \setminus \Gamma^+_u(G)$. Observe that indeed $\mu''(u,\hat{u})/\mu''(e_u) = q \in \mathbb{Q}$ and that the sum of the weights of the outgoing edges of u is equal to λ. Thus, (G'',θ'') is still λ-our-regular. Since $q > \mu(u,\hat{u})/\mu(e_u)$, it can be calculated that $x = \mu''(u,\hat{u})/\mu(u,\hat{u}) > 1$, i.e., that $\mu''(u,\hat{u}) > \mu(u,\hat{u})$:

$$\begin{aligned} &q\mu(e_u) > \mu(u,\hat{u}) \\ &q\mu(e_u)(\lambda - \mu(u,\hat{u})) > \mu(u,\hat{u}))(\lambda - \mu(u,\hat{u})) \\ &q\mu(e_u)\lambda > \mu(u,\hat{u})(\lambda - \mu(u,\hat{u}) + q\mu(e_u)) \\ &q\mu(e_u)\frac{\lambda}{\lambda - \mu(u,\hat{u}) + q\mu(e_u)} > \mu(u,\hat{u}). \end{aligned}$$

As $\mu''(u,\hat{u})/\mu''(e_u) \in \mathbb{Q}$ and the proportions of other edge weights do not change, it means that $k_{G'',\theta''} < k_{G,\theta}$. Hence, by the inductive assumption,

$$F_v(G'',\theta'') = SI_v(G'',\theta'') = EV_v(G'',\theta''), \quad \text{for every } v \in V. \tag{5.11}$$

In what follows, we will show that through a combination of graphs (G',θ') and (G'',θ'') one can obtain graph (G,θ). To this end, we will follow four steps:

1. we scale the node weights of graphs (G',θ') and (G'',θ''),
2. we construct graph $(G^\dagger,\theta^\dagger)$ that is isomorphic to (G',θ'), but has disjoint set of nodes to (G'',θ''),
3. we sum graphs $(G^\dagger,\theta^\dagger)$ and (G'',θ''),
4. in this sum we proportionally combine corresponding nodes from $(G^\dagger,\theta^\dagger)$ into (G'',θ'').

First, for an arbitrary $p \in \mathbb{R}_{>0}$, let us denote $(G',\theta'_p) = ((V',E'),(pc \cdot \mathbb{1}_{\hat{v}}, \mu'))$, which is graph (G',θ') with node weights scaled by p. Next, let us consider graph $(G^\dagger,\theta^\dagger_p)$ that is isomorphic to (G',θ'_p). Formally, take graph $(G^\dagger,\theta^\dagger_p) = ((V^\dagger,E^\dagger),(pc \cdot \mathbb{1}_{\hat{v}^\dagger}, \mu^\dagger))$, where $V^\dagger = \{v^\dagger : v \in V'\}$ such that $V^\dagger \cap V = \emptyset$, $E^\dagger = \{(s^\dagger,t^\dagger) : (s,t) \in E'\}$, and also $\mu^\dagger(s^\dagger,t^\dagger) = \mu'(s,t)$, for every $(s,t) \in E'$. It is clear that $SI_{v^\dagger}(G^\dagger,\theta^\dagger_p) = SI_v(G',\theta'_p)$, for every $v \in V'$, and that $k_{G^\dagger} = k_{G'}$. Therefore, from the inductive assumption we obtain that $F_{v^\dagger}(G^\dagger,\theta^\dagger_p) = SI_v(G',\theta'_p)$. Hence, from Lemma 54 and Eq. (5.10) we get

$$F_{v^\dagger}(G^\dagger,\theta^\dagger_p) = p \cdot F_v(G',\theta'), \quad \text{for every } v \in V'. \tag{5.12}$$

Since $0 = \tilde{\mu}'(u,\hat{u}) < \mu(u,\hat{u}) < \mu''(u,\hat{u})$, we will combine graphs $(G^\dagger, \theta^\dagger_p)$ and (G'', θ'') in order to obtain our original graph (G,θ). To this end, consider graph $(G^\dagger + G'', \theta^\dagger_p + \theta'')$. Let us sequentially proportionally combine each node $v^\dagger$ into v and denote the obtained graph by $(G^*,\theta^*) = ((V,E),((1+p)c\cdot \mathbb{1}_{\hat{v}}, \mu^*))$. Since graph $(G^\dagger + G'', \theta^\dagger_p + \theta'')$ is out-regular, from Locality, Node Combination, and Eq. (5.12) we get that

$$F_v(G^*,\theta^*) = \begin{cases} p\cdot F_v(G',\theta') + F_v(G'',\theta''), & \text{for every } v \in V', \\ F_v(G'',\theta''), & \text{for every } v \in V\setminus V'. \end{cases} \tag{5.13}$$

In what follows, we prove that if we take the value of $p = \frac{F_u(G'',\theta'')}{F_u(G',\theta')}\left(\frac{\mu''(u,\hat{u})}{\mu(u,\hat{u})} - 1\right)$, then the edge weights in the obtained graph are equal to the edge weights of the original graph (G,θ), i.e., $\mu^* = \mu$. For every $(s,t) \in E \setminus \Gamma^+_u(G^*)$ observe that we have $\mu(s,t) = \mu''(s,t)$ and also $\mu'(s,t) = \mu(s,t)$ if $(s,t)\in E'$. Thus, when we combine both nodes $s^\dagger$ into s and $t^\dagger$ into t in graph $(G^\dagger + G'', \theta^\dagger_p + \theta'')$, the weight of edge (s,t) will be preserved. Hence, $\mu^*(s,t) = \mu(s,t)$. Since $(u,\hat{u}) \notin E'$, it does not have a corresponding edge in graph $G^\dagger$. Hence, from the definition of proportional combining we get

$$\begin{aligned} \mu^*(u,\hat{u}) &= \frac{F_u(G'',\theta'')\cdot\mu''(u,\hat{u})}{p\cdot F_u(G',\theta') + F_u(G'',\theta'')} = \\ &= \frac{F_u(G'',\theta'')\cdot\mu''(u,\hat{u})}{F_u(G'',\theta'')\left(\frac{\mu''(u,\hat{u})}{\mu(u,\hat{u})} - 1\right) + F_u(G'',\theta'')} = \\ &= \mu(u,\hat{u}). \end{aligned}$$

For the other outgoing edges of u, $e, e' \in \Gamma^+_u(G^*)\setminus\{(u,\hat{u})\}$, the proportions of their weights are equal in all the three graphs, i.e., $\mu(e)/\mu(e') = \mu'(e)/\mu'(e') = \mu''(e)/\mu''(e')$. Thus, in (G^*,θ^*), this proportion is also preserved, i.e., $\mu^*(e)/\mu^*(e') = \mu(e)/\mu(e')$. Moreover, observe that $(G^\dagger + G'', \theta^\dagger_p + \theta'')$ is λ-out-regular. Since proportional combining preserves out-regularity, graph (G^*,θ^*) is λ-out-regular as well. Hence, the sum of weights of edges in $\Gamma^+_u(G^*)\setminus\{(u,\hat{u})\}$ is equal to $\lambda - \mu(u,\hat{u})$. Since the sum and the proportions of the weights of these edges are the same in both (G,θ) and (G^*,θ^*), the weights are equal as well. As a result, we obtain that $\mu = \mu^*$, which means that graph $(G^*,\theta^*) = ((V,E),((1+p)c\cdot\mathbb{1}_{\hat{v}},\mu))$ is graph (G,θ) with node weights scaled by $(1+p)$.

Observe that we have chosen $c \in \mathbb{R}_{>0}$ arbitrarily. Thus, let us consider the same operation, but with constant $c/(1+p)$ instead of constant c, i.e., take graphs $(G', \theta'_{p/(1+p)}) = ((V',E'),(pc/(1-p)\cdot\mathbb{1}_{\hat{v}},\mu')$ and $(G'',\theta''_{1/(1+p)}) = ((V,E), c/(1+p)\cdot\mathbb{1}_{\hat{v}},\mu'')$. Then, the obtained graph will be exactly the original graph (G,θ). Hence, analogously to Eq. (5.13) we get that

$$F_v(G,\theta) = \begin{cases} F_v(G',\theta'_{p/(1+p)}) + F_v(G'',\theta''_{1/(1+p)}), & \text{for every } v\in V', \\ F_v(G'',\theta''_{1/(1+p)}), & \text{for every } v \in V\setminus V'. \end{cases}$$

By the inductive assumption and Lemma 54, $F_v(G',\theta'_{p/(1+p)}) = p/(1+p)\cdot F_v(G',\theta')$ and $F_v(G'',\theta''_{1/(1+p)}) = 1/(1+p)\cdot F_v(G'',\theta'')$. Therefore,

$$F_v(G,\theta) = \begin{cases} p/(1+p)\cdot F_v(G',\theta') + 1/(1+p)\cdot F_v(G'',\theta''), & \text{for every } v \in V', \\ 1/(1+p)\cdot F_v(G'',\theta''), & \text{for every } v\in V\setminus V'. \end{cases} \tag{5.14}$$

Observe that, by Eq. (5.10) and Eq. (5.11), the value of $p = \frac{F_u(G'',\theta'')}{F_u(G',\theta')}\left(\frac{\mu''(u,\hat{u})}{\mu(u,\hat{u})} - 1\right)$ does not depend on the choice of centrality measure F. Thus, since Seeley index and eigenvector centrality satisfy our axioms (Lemmas 52 and 53) we know that they also satisfy Eq. (5.14) with the same value of p. Combining Eq. (5.14) with Eq. (5.10) and Eq. (5.11), we get that $F_v(G,\theta) = SI_v(G,\theta) = EV_v(G,\theta)$, for every $v \in V$, from which $(*)$ follows.

It remains to relax the initial assumption that node weights have to be unit node weights multiplied by a constant. To this end, let us consider an arbitrary strongly connected λ-out-regular graph $(G,\theta) = ((V,E),(b,\mu))$ and two cases: the first one (I) in which $b(G) > 0$, and the second one (II) in which $b(G) = 0$.

(I) If $b(G)) > 0$, then let us denote the set of nodes that have positive weights by $V^* = \{v \in V : b(v) > 0\}$. For every $v \in V^*$ let us construct weights $\theta_v = (b(v) \cdot \mathbb{1}_v, \mu)$. Observe that each graph (G,θ_v) is strongly connected and λ-out-regular with unit node weights multiplied by a constant. Thus, from $(*)$, for every $u \in V$, we have $F_u(G,\theta_v) = SI_u(G,\theta_v) = EV_u(G,\theta_v)$.

In order to combine all graphs (G,θ_v) into one graph (G,θ), for each graph (G,θ_v) let us define graph (G',θ'_v) isomorphic to it. More in detail, let $G' = (V',E')$ such that $V \cap V' = \emptyset$ and that $V' = \{u' : u \in V\}$ and $E' = \{(u',w') : (u,w) \in E\}$. Next, let $\theta'_v = (b(v) \cdot \mathbb{1}_{v'}, \mu')$, where $\mu'(u',w') = \mu(u,w)$, for every $(u,w) \in E$. Graph (G',θ'_v) is also λ-out-regular with unit node weights multiplied by a constant. Thus, from $(*)$, we have $F_{u'}(G',\theta'_v) = SI_{u'}(G',\theta'_v) = SI_u(G,\theta_v) = F_u(G,\theta_v)$, for every $u \in V$.

Building upon this, let us consider the following operation: Let us choose one node $v \in V^*$ and take graph G_v and say that at the beginning it is our current graph. Next, for node $u \in V^* \setminus \{v\}$ let us take graph G'_u, add it to the current graph, sequentially combine node w' into w, for all $w \in V$, and say that the resulting graph is now the current graph. Let us perform this operation for all $u \in V^* \setminus \{v\}$ exactly once. Observe that after each such addition of graph G'_u, the nodes, the edges, and the edge weights of the current graph remain unchanged. Only the node weights of the current graph are summed with node weights of just added graph G'_u. Hence, the graph that we obtain after such an operation for all $u \in V^* \setminus \{v\}$ is the original graph G. Now, from Locality and Node Combination we obtain that for every $u \in V$, we have

$$F_u(G,\theta) = \sum_{v \in V^*} F_u(G,\theta_v).$$

Since Seeley index and eigenvector centrality satisfy our axioms (Lemmas 52–53), this equation holds also for them. Thus, $F_u(G,\theta) = SI_u(G,\theta) = EV_u(G,\theta)$, for every $u \in V$.

(II) Consider a strongly connected λ-out-regular graph $(G,\theta) = ((V,E),(b,\mu))$ such that $b(G) = 0$. From Eq. (2.8) and Eq. (2.14) we see that for such a graph Seeley index and eigenvector centrality are equal to zero for every node. We prove that the same is true for centrality F. Assume otherwise, i.e., there exists strongly connected λ-out-regular graph $(G,\theta) = ((V,E),(b,\mu))$ and node $v \in V$ such that $b(G) = 0$ and $F_v(G,\theta) > 0$. Then, for node $v' \notin V$ consider a small λ-out-regular graph with only v' and a loop, i.e., $(G^v,\theta^v) = ((\{v\},\{(v,v)\}),([1],[\lambda]))$. Next, let us add it to graph G, i.e., let $(G',\theta') = (G + G^v, \theta + \theta^v)$. From Cycle axiom we have that $F_{v'}(G^v,\theta^v) = 1$. Thus, by Locality, $F_{v'}(G',\theta') = 1$. Now, let us combine node v' into node v in graph G', i.e., let us take $(G'',\theta'') = C^F_{v' \to v}(G',\theta')$. From Node Combination and Locality we have that

$$F_v(G'',\theta'') = F_{v'}(G',\theta') + F_v(G',\theta') = 1 + F_v(G,\theta).$$

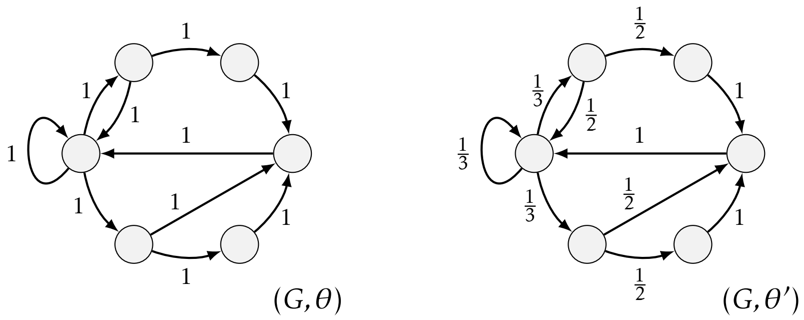


Figure 5.7: An illustration to the proof of Lemma 57. The graph on the left hand side, (G,θ), is a strongly connected graph that is not out-regular. The graph on the right hand side, (G,θ'), is a graph obtained from (G,θ) by dividing the weights of outgoing edges of v by $\deg^+_v(G,\theta)$ for every node v. Note that (G,θ') is now 1-out-regular.

Since G'' is a strongly connected λ-out-regular graph with unit node weights, from $(*)$ we get that $F_v(G'',\theta'') = SI_v(G'',\theta'')$. However, we know that the sum of Seeley indices in a graph is the sum of weights in that graph (Theorem 2). Thus, $\sum_{u\in V} SI_u(G'') = 1 + b(G) = 1$ and we arrive at a contradiction, because we get that $SI_v(G'',\theta'') = F_v(G'',\theta'') = 1 + F_v(G,\theta) > 1$. □

We have shown that any centrality measure, F, that satisfies Locality, Edge Deletion, Node Combination, and Cycle for every strongly connected out-regular graph assigns the same centrality to each node as Seeley index and eigenvector centrality. Now, we will move to arbitrary graphs that are not necessarily out-regular. For such graphs Seeley index and eigenvector centrality are not always equal. Therefore, at this point, the proofs for both centrality measures split.

We begin with Seeley index and the following lemma in which we show that if F additionally satisfies Edge Multiplication, then for every node of every graph in $\mathcal{G}^{SI}$ centrality measure F is equal to Seeley index.

Lemma 57. *If a centrality measure, F, defined on $\mathcal{G}^{SI}$ satisfies Locality, Edge Deletion, Node Combination, Edge Multiplication, and Cycle, then for every graph $G = (V,E)$ and weights $\theta = (b,\mu)$ such that $(G,\theta) \in \mathcal{G}^{SI}$, it holds that*

$$F_v(G,\theta) = SI_v(G,\theta), \quad \textit{for every } v \in V.$$

Proof. Fix an arbitrary graph $(G,\theta) = ((V,E),(b,\mu)) \in \mathcal{G}^{SI}$. From Locality, without loss of generality, we can assume that (G,θ) is connected. Since $(G,\theta) \in \mathcal{G}^{SI}$, this means that G is strongly connected. Let us then, consider a modification of (G,θ) in which the weight of each edge is divided by the out-degree of its start. Formally, let $(G,\theta') = ((V,E),(b,\mu'))$, where $\mu'(u,v) = \mu(u,v)/\deg^+_u(G,\theta)$, for every $(u,v) \in E$ (see Fig. 5.7 for an illustration). Observe that graph (G,θ') is 1-out-regular. Hence, from Lemma 56 we get that $F_v(G,\theta') = SI_v(G,\theta')$, for every $v \in V$. Now, graph (G,θ) can be obtained from (G,θ') by multiplying outgoing edges of every node $v \in V$ by $\deg^+_v(G,\theta)$. Since both F and Seeley index satisfy Edge Multiplication (Lemma 52), we get that

$$F_v(G,\theta) = F_v(G,\theta') = SI_v(G,\theta') = SI_v(G,\theta), \quad \text{for every } v \in V.$$

□

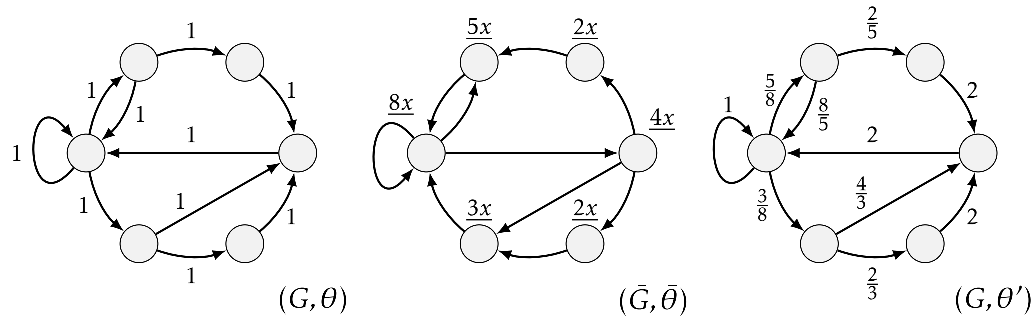


Figure 5.8: An illustration to the proof of Lemma 58. The leftmost graph, (G,θ), is a strongly connected graph that is not out-regular. The middle graph, $(\bar{G},\bar{\theta})$, is an opposite graph to (G,θ), i.e., the direction of every edge is inverted. Underlined values at each node are equal to eigenvector centrality of this node (note that we only need the relative values). The rightmost graph, (G,θ'), is a graph obtained from (G,θ) by dividing the weights of outgoing edges of v by $EV_v(\bar{G},\bar{\theta})$ and multiplying the incoming edges of v by the same value for every node v. Note that graph (G,θ') is now 2-out-regular.

Now, let us move to the analogous lemma for eigenvector centrality, i.e., let us prove that if a centrality measure satisfies Locality, Edge Deletion, Node Combination, Edge Compensation, and Cycle, then it is equal to eigenvector centrality for every node of every graph in $\mathcal{G}^{EV}$.

Lemma 58. *If a centrality measure, F, defined on $\mathcal{G}^{EV}$ satisfies Locality, Edge Deletion, Node Combination, Edge Compensation, and Cycle, then for every graph $G = (V,E)$ and weights $\theta = (b,\mu)$ such that $(G,\theta) \in \mathcal{G}^{EV}$, it holds that*

$$F_v(G,\theta) = EV_v(G,\theta), \quad \textit{for every } v \in V.$$

Proof. Fix an arbitrary graph $(G,\theta) = ((V,E),(b,\mu)) \in \mathcal{G}^{EV}$. From Locality, without loss of generality, we can assume that (G,θ) is connected. Since $(G,\theta) \in \mathcal{G}^{EV}$, this means that (G,θ) is strongly connected. In what follows, we modify its edge weights using Edge Compensation in such a way that the obtained graph is out-regular.

To this end, let us first consider a graph opposite to (G,θ), i.e., a graph in which each edge is in the opposite direction. Formally, let $(\bar{G},\bar{\theta}) = ((V,\bar{E}),(b,\bar{\mu}))$, where $\bar{E} = \{(u,v) : (v,u) \in E\}$ and $\bar{\mu}(u,v) = \mu(v,u)$, for every $(u,v) \in \bar{E}$ (see Fig. 5.8). Now, in graph $(\bar{G},\bar{\theta})$ let us multiply the weights of outgoing edges of node $v \in V$ by $EV_v(\bar{G},\bar{\theta})$ and divide the weights of its incoming edges as well as the weight of v also by $EV_v(\bar{G},\bar{\theta})$. Because eigenvector centrality satisfies Edge Compensation, we know that this operation does not affect the centralities of nodes other than v and divides the centrality of v by $EV_v(\bar{G},\bar{\theta})$, making it equal to 1. Proceeding with this operation for each node $v \in V$, we obtain graph $(\bar{G},\bar{\theta}')$ in which all nodes have eigenvector centrality equal to 1. Formally, $\bar{\theta}' = (b',\bar{\mu}')$, where $b'(v) = b(v)/EV_v(\bar{G},\bar{\theta})$, for every $v \in V$, and $\bar{\mu}'(u,v) = \bar{\mu}(u,v) \cdot EV_u(\bar{G},\bar{\theta})/EV_v(\bar{G},\bar{\theta})$, for every $(u,v) \in \bar{E}$.

If all nodes in a graph have equal centralities, then from eigenvector centrality recursive equation (Eq. (2.2)) we get that the in-degree of each node is equal to principal eigenvalue λ, i.e, $\deg^-_v(\bar{G},\bar{\theta}') = \lambda$, for every $v \in V$. Hence, $(\bar{G},\bar{\theta}')$ is λ-in-regular and the graph opposite to it would be λ-out-regular. Let us define graph (G,θ') as opposite to $(\bar{G},\bar{\theta}')$ but with slightly modified node weights. Formally, let us denote $\theta' = (b'',\mu')$ where $b''(v) = b(v) \cdot EV_v(\bar{G},\bar{\theta})$, for every $v \in V$, and also $\mu'(u,v) = \mu(u,v) \cdot EV_v(\bar{G},\bar{\theta})/EV_u(\bar{G},\bar{\theta})$, for every $(u,v) \in E$. See Fig. 5.8 for an illustration. Graph (G,θ') is λ-out-regular. Hence, from Lemma 56 we get that

$$F_v(G,\theta') = EV_v(G,\theta'), \quad \text{for every } v \in V. \tag{5.15}$$

Graph (G,θ) can be obtained from (G,θ') by multiplying the weights of outgoing edges of v by $EV_v(\bar{G},\bar{\theta})$ and dividing the weights of its incoming edges as well as its own weight also by $EV_v(\bar{G},\bar{\theta})$. Thus, from Edge Compensation for both F and eigenvector centrality and Eq. (5.15), we get that

$$F_v(G,\theta) = F_v(G,\theta')/EV_v(\bar{G},\bar{\theta}) = EV_v(G,\theta')/EV_v(\bar{G},\bar{\theta}) = EV_v(G,\theta), \quad \text{for every } v \in V.$$

□

5.3.2 PageRank and Katz Centrality (Theorems 49 and 50)

Finally, we move to the proofs for PageRank and Katz centrality. First, we show that both centrality measures satisfy their respective sets of axioms (Lemma 59 and 60). Then, we prove that PageRank and Katz centrality are also implied by these sets of axioms (Lemmas 61–74).

We begin with the following lemma, in which we show that PageRank satisfies our axioms. Although Locality, Edge Deletion, Edge Multiplication, and Baseline have been considered in the previous chapters, the model have slightly changed (through addition of edge weights). Therefore, we consider these axioms again.

Lemma 59. *For every decay factor $a \in [0,1)$, PageRank defined by Eq. (2.7) satisfies Locality, Edge Deletion, Node Combination, Edge Multiplication, and Baseline.*

Proof. Let us take an arbitrary graph $G = (V,E)$ and weights $\theta = (b,\mu)$, and consider axioms one by one.

- For Locality, let us consider graph $G' = (V',E')$ such that $V \cap V' = \emptyset$, weights $\theta' = (b',\mu')$, and an arbitrary node $v \in V$. If $b(G) = 0$, then it trivially follows that $PR^a_v(G,\theta) = 0 = PR^a_v(G+G',\theta+\theta')$. Thus, let us assume that $b(G) > 0$. Observe that in graph $(G+G',\theta+\theta')$ walks that start at nodes in V' cannot visit nodes in V and vice versa. Therefore, for every $t \in \mathbb{N}$, we have that $\{\omega \in \Omega_t(G+G') : \omega(t) = v\} = \{\omega \in \Omega_t(G) : \omega(t) = v\}$. By Eq. (2.6), this implies that also $p^a_{G+G',\theta+\theta'}(v,t)/(b(G)+b'(G')) = p^a_{G,\theta}(v,t)/b(G)$, because weights of edges in E and the out-degrees of nodes in V are the same in both graph (G,θ) and graph $(G+G',\theta+\theta')$. Hence, Locality follows from Eq. (2.7).

- For Edge Deletion, consider edge $(u,w) \in E$ and an arbitrary node that is not a successor of u, i.e, $v \in V \setminus S_u(G)$. Let $G' = (V, E \setminus \{(u,w)\})$ be graph G with edge (u,w) removed and $\theta' = (b,\mu_{-(u,w)})$ its weights. If $b(G) = 0$, then observe that $PR^a_v(G,\theta) = 0 = PR^a_u(G',\theta')$. Thus, let us assume that $b(G) > 0$. Observe that since v is not a successor of u in graph G, then a walk on G that has visited node u may not visit v later on. Therefore, the removal of edge (u,w) does not affect the walks of length t that end at node v, i.e., we have that $\{\omega \in \Omega_t(G) : \omega(t) = v\} = \{\omega \in \Omega_t(G') : \omega(t) = v\}$. Moreover, for each walk $\omega \in \Omega_t(G)$ that does not visit u, i.e., $\omega(i) \neq u$, for every $i \in \{0,\dots,t\}$, we have that

$$\frac{b(\omega(0))}{b(G)} \prod_{i=0}^{t-1} a \cdot \frac{\mu(\omega(i),\omega(i+1))}{\deg^+_{\omega(i)}(G,\theta)} = \frac{b(\omega(0))}{b(G)} \prod_{i=0}^{t-1} a \cdot \frac{\mu_{-(u,w)}(\omega(i),\omega(i+1))}{\deg^+_{\omega(i)}(G,\theta')},$$

i.e., it is equal for graphs (G,θ) and (G',θ'). Thus, since $b(G) = b'(G)$, by Eq. (2.6), also $p^a_{G,\theta}(v,t) = p^a_{G',\theta'}(v,t)$. Hence, from Eq. (2.7) we obtain that $PR^a_v(G,\theta) = PR^a_v(G',\theta')$.

- For Node Combination, take $u,w \in V$ such that $\deg^+_u(G) = \deg^+_w(G) = \deg^+_s(G)$, for every $s \in S_u(G) \cup S_w(G)$. Let $(G',\theta') = ((V',E'),(b',\mu')) = C^{PR^a}_{u\to w}(G,\theta)$. Since PageRank is equivalently defined as the solution to system of PageRank recursive equations (Theorem 1), it suffices to show that $(x_v)_{v\in V\setminus\{u\}}$ defined as $x_v = PR^a_v(G,\theta)$ for every $v \in V \setminus \{u,w\}$ and $x_w = PR^a_u(G,\theta) + PR^a_w(G,\theta)$ satisfies PageRank recursive equation (Eq. (2.5)) for every $v \in V \setminus \{u\}$ and graph (G',θ').

  To this end, fix $v \in V$ and observe that from PageRank recursive equation (Eq. (2.5)) for graph (G,θ) we have

$$PR^a_v(G,\theta) = a\left(\sum_{s\in P^1_v(G)} \frac{\mu(s,v)PR^a_s(G,\theta)}{\deg^+_s(G,\theta)}\right) + b(v). \tag{5.16}$$

  Recall that $\deg^+_u(G,\theta) = \deg^+_w(G,\theta)$. Also, proportional combining does not affect out-degrees, so $\deg^+_s(G,\theta) = \deg^+_s(G',\theta')$, for every $s \in V \setminus \{u\}$. First, assume $v \in V \setminus \{u,w\}$. Then, by the definition of the proportional combining, $\mu'(w,v) = (PR^a_u(G,\theta)\tilde{\mu}(u,v) + PR^a_w(G,\theta)\tilde{\mu}(w,v))/(PR^a_u(G,\theta) + PR^a_w(G,\theta))$. Thus, Eq. (5.16) can be transformed into

$$PR^a_v(G,\theta) = a\left(\frac{\tilde{\mu}'(w,v)x_w}{\deg^+_w(G',\theta')} + \sum_{s\in P^1_v(G)\setminus\{u,w\}} \frac{\mu(s,v)PR^a_s(G,\theta)}{\deg^+_s(G',\theta')}\right) + b(v). \tag{5.17}$$

  The incoming edges of v that come from nodes other then u and w are unaffected by proportional combining. Thus, $P^1_v(G) \setminus \{u,w\} = P^1_v(G') \setminus \{u,w\}$ and $\mu(s,v) = \mu'(s,v)$, for every $s \in P^1_v(G) \setminus \{u,w\}$. Also, since $v \notin \{u,w\}$ we have $b(v) = b'(v)$. Hence, from Eq. (5.17) we get that

$$x_v = a\left(\sum_{s\in P^1_v(G')} \frac{\mu'(s,v)}{\deg^+_s(G',\theta')} \cdot x_s\right) + b'(v),$$

  which is the PageRank recursive equation for graph (G',θ') and node v.

  Therefore, it remains to consider node w. To this end, let us sidewise add Eq. (5.16) for $v = u$ and $v = w$. By the definition of proportional combining,

$$\mu'(w,w) = \frac{PR^a_u(G,\theta)(\tilde{\mu}(u,u) + \tilde{\mu}(u,w)) + PR^a_w(G,\theta)(\tilde{\mu}(w,u) + \tilde{\mu}(w,w))}{PR^a_u(G,\theta) + PR^a_w(G,\theta)}.$$

  Other incoming edges to u and w are simply combined, i.e., we have that $P^1_w(G') \setminus \{w\} = P^1_u(G) \cup P^1_w(G) \setminus \{u,w\}$ and $\mu'(s,w) = \tilde{\mu}(s,u) + \tilde{\mu}(s,w)$, for every $s \in P^1_w(G') \setminus \{w\}$. Again, $\deg^+_u(G,\theta) = \deg^+_w(G,\theta)$ and $\deg^+_s(G,\theta) = \deg^+_s(G',\theta')$ for every $s \in V \setminus \{u\}$. Also, $b'(w) = b(u) + b(w)$. Thus, from Eq. (5.16) we get

$$x_w = PR^a_u(G,\theta) + PR^a_w(G,\theta) = a\left(\frac{\tilde{\mu}'(w,w)x_w}{\deg^+_w(G',\theta')} + \sum_{s\in P^1_w(G')\setminus\{u,w\}} \frac{\mu'(s,v)x_s}{\deg^+_s(G',\theta')}\right) + b'(w),$$

  which is the PageRank recursive equation for graph (G',θ') and node w.

- For Edge Multiplication, consider arbitrary nodes $u, v \in V$, constant $x \in \mathbb{R}_{>0}$, and weights $\theta' = (b, \mu_{-\Gamma_u^+(G)} + x \cdot \mu_{\Gamma_u^+(G)})$. If $b(G) = 0$, then observe that $PR_v^a(G,\theta) = 0 = PR_u^a(G,\theta')$. Thus, let us assume that $b(G) > 0$. Then, observe that for every $t \in \mathbb{N}$ and walk $\omega \in \Omega_t(G)$ such that $\omega(t) = v$, we have

$$\frac{b(\omega(0))}{b(G)} \cdot \prod_{i=0}^{t-1} \frac{\mu(\omega(i), \omega(i+1))}{\deg^+_{\omega(i)}(G,\theta)} = \frac{b(\omega(0))}{b(G)} \cdot \prod_{i=0}^{t-1} \frac{\mu'(\omega(i), \omega(i+1))}{\deg^+_{\omega(i)}(G,\theta')}$$

  It holds because for every $i \in \{0, \dots, t-1\}$, if $\omega(i) = u$, then both the numerator and the denominator are multiplied by x, and if $\omega(i) \neq u$, both the numerator and the denominator are unaffected. Thus, from Eq. (2.6) we obtain that $p^a_{G,\theta}(v,t) = p^a_{G,\theta'}(v,t)$. Summing for all $t \in \mathbb{N}$, we get $PR_v^a(G,\theta) = PR_v^a(G,\theta')$ from Eq. (2.7). Hence, Edge Multiplication is satisfied.

- Finally, for Baseline, observe that it follows directly from PageRank recursive equation (Eq. (2.5)).

□

Now, let us move to the analogous lemma for Katz centrality.

Lemma 60. *For every decay factor $a \in \mathbb{R}_{\geq 0}$ Katz centrality defined on $\mathcal{G}^{K(a)}$ by Eq.* (2.13) *satisfies Locality, Edge Deletion, Node Combination, Edge Compensation, and Baseline.*

Proof. Let us take an arbitrary graph $G = (V, E)$ and weights $\theta = (b, \mu)$ such that $(G,\theta) \in \mathcal{G}^{K(a)}$, and consider the axioms one by one.

- For Locality, let us consider graph $G' = (V', E')$ such that $V \cap V' = \emptyset$, weights $\theta' = (b', \mu')$, and an arbitrary node $v \in V$. If $b(G) = 0$, then it trivially follows that $K_v^a(G,\theta) = 0 = K_v^a(G+G', \theta+\theta')$. Thus, let us assume that $b(G) > 0$. Observe that in graph $(G+G', \theta+\theta')$ walks that start at nodes in V' cannot visit nodes in V and vice versa. Therefore, for every $t \in \mathbb{N}$, we have that $\{\omega \in \Omega_t(G+G') : \omega(t) = v\} = \{\omega \in \Omega_t(G) : \omega(t) = v\}$. By Eq. (2.12), this implies that also $w^a_{G+G',\theta+\theta'}(v,t)/(b(G)+b'(G')) = w^a_{G,\theta}(v,t)/b(G)$, because weights of edges in E and the out-degrees of nodes in V are the same in graphs (G,θ) and $(G+G', \theta+\theta')$. Hence, Locality follows from Eq. (2.13).

- For Edge Deletion, consider edge $(u,w) \in E$ and an arbitrary node that is not a successor of u, i.e, $v \in V \setminus S_u(G)$. Let $G' = (V, E \setminus \{(u,w)\})$ be graph G with edge (u,w) removed and $\theta' = (b, \mu_{-(u,w)})$ its weights. If $b(G) = 0$, then observe that $K_v^a(G,\theta) = 0 = K_u^a(G',\theta')$. Thus, let us assume that $b(G) > 0$. Observe that since v is not a successor of u in graph G, then a walk on G that has visited node u may not visit v later on. Therefore, the removal of edge (u,w) does not affect the walks of length t that end at node v, i.e., we have that $\{\omega \in \Omega_t(G) : \omega(t) = v\} = \{\omega \in \Omega_t(G') : \omega(t) = v\}$. Moreover, for each walk $\omega \in \Omega_t(G)$ that does not visit u, i.e., $\omega(i) \neq u$, for every $i \in \{0, \dots, t\}$, we have that

$$\frac{b(\omega(0))}{b(G)} \prod_{i=0}^{t-1} a \cdot \mu(\omega(i), \omega(i+1)) = \frac{b(\omega(0))}{b(G)} \prod_{i=0}^{t-1} a \cdot \mu_{-(u,w)}(\omega(i), \omega(i+1)),$$

  i.e., it is equal for graphs (G,θ) and (G',θ'). Thus, since $b(G) = b'(G)$, by Eq. (2.12), also $w^a_{G,\theta}(v,t) = w^a_{G',\theta'}(v,t)$. Hence, from Eq. (2.13) we obtain that $K_v^a(G,\theta) = K_v^a(G',\theta')$.

- For Node Combination, take $u,w \in V$ such that $\deg^+_u(G) = \deg^+_w(G) = \deg^+_s(G)$, for every $s \in S_u(G) \cup S_w(G)$. Let $(G',\theta') = ((V',E'),(b',\mu')) = C^{K^a}_{u\to w}(G,\theta)$. Since Katz centrality is equivalently defined as the solution to the system of Katz centrality recursive equations (Theorem 3), it suffices to show that $(x_v)_{v\in V\setminus\{u\}}$ defined as $x_v = K^a_v(G,\theta)$, for every $v \in V \setminus \{u,w\}$, and $x_w = K^a_u(G,\theta) + K^a_w(G,\theta)$ satisfies Katz centrality recursive equation (Eq. (2.3)), for every $v \in V \setminus \{u\}$ and graph (G',θ').

  To this end, fix $v \in V$ and observe that from Katz centrality recursive equation (Eq. (2.3)) for graph (G,θ) we have

$$K^a_v(G,\theta) = a\left(\sum_{s\in P^1_v(G)} \mu(s,v)K^a_s(G,\theta)\right) + b(v). \tag{5.18}$$

  First, assume that $v \in V \setminus \{u,w\}$. Then, by the definition of proportional combining, $\mu'(w,v) = (K^a_u(G,\theta)\tilde{\mu}(u,v) + K^a_w(G,\theta)\tilde{\mu}(w,v))/(K^a_u(G,\theta)+K^a_w(G,\theta))$. Thus, Eq. (5.18) can be transformed into

$$K^a_v(G,\theta) = a\left(\tilde{\mu}'(w,v)x_w + \sum_{s\in P^1_v(G)\setminus\{u,w\}} \mu(s,v)K^a_s(G,\theta)\right) + b(v). \tag{5.19}$$

  Incoming edges of v that come from nodes other then u and w are unaffected by proportional combining. Thus, $P^1_v(G) \setminus \{u,w\} = P^1_v(G') \setminus \{u,w\}$ and also $\mu(s,v) = \mu'(s,v)$, for every $s \in P^1_v(G) \setminus \{u,w\}$. Also, since $v \notin \{u,w\}$, we have $b(v) = b'(v)$. Hence, from Eq. (5.19) we get that

$$x_v = a\left(\sum_{s\in P^1_v(G')} \mu'(s,v)\cdot x_s\right) + b'(v),$$

  which is exactly the Katz centrality recursive equation for graph (G',θ') and node v.

  Therefore, it remains to consider node w. To this end, let us sidewise add Eq. (5.18) for $v = u$ and $v = w$. By the definition of proportional combining,

$$\mu'(w,w) = \frac{K^a_u(G,\theta)(\tilde{\mu}(u,u)+\tilde{\mu}(u,w)) + K^a_w(G,\theta)(\tilde{\mu}(w,u)+\tilde{\mu}(w,w))}{K^a_u(G,\theta)+K^a_w(G,\theta)}.$$

  Observe that the other incoming edges to u and w are simply combined together, i.e., $P^1_w(G') \setminus \{w\} = P^1_u(G) \cup P^1_w(G) \setminus \{u,w\}$ and $\mu'(s,w) = \tilde{\mu}(s,u) + \tilde{\mu}(s,w)$, for every $s \in P^1_w(G') \setminus \{w\}$. Also, $b'(w) = b(u) + b(w)$. Thus, from Eq. (5.18) we obtain

$$x_w = K^a_u(G,\theta) + K^a_w(G,\theta) = a\left(\tilde{\mu}'(w,w)x_w + \sum_{s\in P^1_w(G')\setminus\{u,w\}} \mu'(s,v)x_s\right) + b'(w),$$

  which is the Katz centrality recursive equation for graph (G',θ') and node w.

- For Edge Compensation, take arbitrary $u \in V$ and $x \in \mathbb{R}_{>0}$, and let $\theta' = (b',\mu')$ such that $b' = b_{-u} + b_u/x$ and $\mu' = \mu_{-\Gamma^{\pm}_u(G)\setminus\{(u,u)\}} + \mu_{\Gamma^+_u(G)\setminus\{(u,u)\}}\cdot x + \mu_{\Gamma^-_u(G)\setminus\{(u,u)\}}/x$. Fix $v \in V\setminus\{u\}$. If $b(G) = 0$, then it holds that $K^a_v(G,\theta') = 0 = K^a_v(G,\theta)$, for every $v \in V$. Thus, let us assume that $b(G) > 0$.

We will show that for every $t \in \mathbb{N}$ and walk $\omega \in \Omega_t(G)$ that ends at v, i.e., $\omega(t) = v$, it holds that

$$b(\omega(0)) \cdot \prod_{i=1}^{t} a\mu(\omega(i-1), \omega(i)) = b'(\omega(0)) \cdot \prod_{i=1}^{t} a\mu'(\omega(i-1), \omega(i)). \tag{5.20}$$

To this end, observe that since $\omega(t) \neq u$, for every step $i \in \{0, \ldots, t-1\}$ in which the walk arrives at node u, i.e., $\omega(i-1) \neq u$ and $\omega(i) = u$ (if the walk starts in u, we treat step 0 as arrival as well), there exist exactly one step $j > i$ in which the walk departs from u, i.e., $\omega(k) = u$, for every $k \in \{i, \ldots, j-1\}$, and $\omega(j) \neq u$. Now, in (G, θ'), the factor for step i (arrival) decreases by x, i.e., $\mu'(\omega(i-1), \omega(i)) = \mu(\omega(i-1), \omega(i))/x$ (or $b'(i) = b(i)/x$ if $i = 0$), but at the same time it holds that the factor for step j (departure) increases by x, i.e., $\mu'(\omega(j-1), \omega(j)) = \mu(\omega(j-1), \omega(j)) \cdot x$. Since there is equal number of arrivals and departures from u, Eq. (5.20) holds. Thus, from Eq. (2.12) and Eq. (5.20) we have that $w^a_{G,\theta}(v,t)/b(G) = w^a_{G,\theta'}(v,t)/b'(G)$ and from Eq. (2.13) we get $K^a_v(G,\theta) = K^a_v(G,\theta')$.

By similar reasoning as in Eq. (5.20), for node u we can obtain that for every $t \in \mathbb{N}$ and $\omega \in \Omega_t(G)$ that ends at u, i.e., $\omega(t) = u$, it holds that

$$b(\omega(0)) \cdot \prod_{i=1}^{t-1} \frac{1}{\lambda} \mu(\omega(i-1), \omega(i))/x = b'(\omega(0)) \cdot \prod_{i=1}^{t-1} \frac{1}{\lambda} \mu'(\omega(i-1), \omega(i)).$$

Here, there is a departure step for every but last arrival at u. However, since $\omega(t) = u$, for the last arrival, there is no departure. Hence, for weights θ' the product is divided by x one more time than it is multiplied by x. Thus, from Eq. (2.12) we get that $w^a_{G,\theta'}(v,t)/b'(G) = w^a_{G,\theta}(v,t)/b(G)/x$. Hence, from Eq. (2.13) we get that $K^a_u(G,\theta') = K^a_u(G,\theta)/x$.

- Finally, for Baseline observe that it follows directly from Katz centrality recursive equation (Eq. (2.3)).

□

Now, let us move to the proof that our axioms imply PageRank and Katz centrality. In its main part, we will focus on *semi-out-regular* graphs which generalize out-regular graphs: all nodes except for sinks have equal out-degree. Most of the following lemmas will consider such graphs only.

Definition 7. *Graph* $(G,\theta) = ((V,E),(b,\mu))$ *is* semi-out-regular *if there exists constant* $r \in \mathbb{R}_{>0}$ *such that for every* $v \in V$ *it holds that* $\deg^+_v(G,\theta) = r$ *or* $\deg^+_v(G,\theta) = 0$.

Let F be a centrality measure that satisfies Locality, Edge Deletion, Node Combination, Edge Multiplication or Edge Compensation, and Baseline. First we show several simple properties stemming from our axioms that are useful in later proofs:

- proportional combining of any two nodes in semi-out-regular graphs preserves centralities in a graph (Lemmas 61 and 62),
- centrality of a node is equal to its weight if it is a source (Lemma 63) and greater or equal to its weight in any other case (Lemma 64),
- centrality is linear with respect to node weights (Lemma 65), and
- there exists constant a_F such that $F_v((\{u,v\},\{(u,v)\},([x,0],[1])) = a_F \cdot x$, for every $x \in \mathbb{R}_{\geq 0}$.

Then, we define a profit function $p_F(x,y,z)$ as the centrality profit a node obtains from an incoming edge with weight y that starts at a node with centrality x and out-degree z in the smallest graph with such situation possible: a graph that consists of three nodes and two edges. We show that if F satisfies Edge Multiplication, then $p_F(x,y,z) = a_F \cdot x \cdot y/z$, as in PageRank (Lemma 67), and if F satisfies Edge Compensation, $p_F(x,y,z) = a_F \cdot x \cdot y$, as in Katz centrality (Lemma 68). Furthermore, we show that in every semi-out-regular graph (G,θ) the profit of every node v from every edge (u,v) is always equal to $p_F(F_u(G,\theta), \mu(u,v), \deg^+_u(G,\theta))$. In other words, F satisfies recursive equation

$$F_v(G,\theta) = b(v) + \sum_{u \in P^1_v(G)} p_F(F_u(G,\theta), \mu(u,v), \deg^+_u(G,\theta)).$$

First, we prove it only for nodes without loops (Lemma 69) and then for arbitrary nodes (Lemma 70). This allows us to show that if centrality F satisfies Edge Multiplication, then it is equal to PageRank for semi-out-regular graphs (Lemma 71) and all graphs (Lemma 72) and, similarly, if F satisfies Edge Compensation, it is equal to Katz centrality for all semi-out-regular graphs (Lemma 73) and arbitrary graphs (Lemma 74).

In the following two lemmas we show that Locality, Edge Deletion, and Node Combination imply that proportional combining of nodes in graphs that are semi-out-regular preserves centralities of nodes. We start by showing that, with Locality and Edge Deletion, Node Combination can be strengthened: We relax the condition that the out-degrees of the combined nodes must be equal to the out-degrees of all of their successors and require only that they are equal to the out-degrees of these of their successors that are not sinks.

Lemma 61. *If a centrality measure, F, defined on $\mathcal{G}$ (or $\mathcal{G}^{K(a)}$) satisfies Locality, Edge Deletion, and Node Combination, then for every graph $G = (V,E)$, weights $\theta = (b,\mu)$, and nodes $u,w \in V$ such that $(G,\theta) \in \mathcal{G}$ (or $(G,\theta) \in \mathcal{G}^{K(a)}$), $\deg^+_u(G,\theta) = \deg^+_w(G,\theta)$, and for every $s \in S_u(G) \cup S_w(G)$ we have $\deg^+_s(G,\theta) = \deg^+_u(G,\theta)$ or $\deg^+_s(G,\theta) = 0$, it holds that*

$$F_v(C^F_{u \to w}(G,\theta)) = F_v(G,\theta), \quad \text{for every } v \in V \setminus \{u,w\}$$

and $F_w(C^F_{u \to w}(G,\theta)) = F_u(G,\theta) + F_w(G,\theta)$.

Proof. Consider arbitrary nodes $u,w \in V$ such that $\deg^+_u(G,\theta) = \deg^+_w(G,\theta)$ and for every $s \in S_u(G) \cup S_w(G)$, either $\deg^+_s(G,\theta) = \deg^+_u(G,\theta)$ or $\deg^+_s(G,\theta) = 0$. Let $(G',\theta') = C^F_{w \to u}(G,\theta)$.

Observe that if $\deg^+_u(G,\theta) = \deg^+_w(G,\theta) = 0$, then there are no successors of u and w and the thesis follows directly from Node Combination. Thus, assume otherwise. Let us denote $r = \deg^+_u(G,\theta) = \deg^+_w(G,\theta)$.

We will construct an out-regular graph based on (G,θ). To this end, consider the set of all sinks in V, i.e., let $V_s = \{v \in V : \Gamma^+_v(G) = \emptyset\}$. Let us add a new node, $t \notin V$, to the original graph with a loop and an edge from each node in V_s. Formally, let $\bar{G} = (V \cup \{t\}, E \cup \bar{E})$, where $\bar{E} = \{(v,t) : v \in V_s \cup \{t\}\}$. Furthermore, let us define the weights of additional node and edges by $\bar{\theta} = (\bar{b}, \bar{\mu})$, where $\bar{b}(t) = 0$ and $\bar{\mu}(e) = r$, for every $e \in \bar{E}$. In this way, graph $(\bar{G}, \theta + \bar{\theta})$ is indeed r-out-regular. Observe that since nodes in V_s do not have outgoing edges in graph G, they do not have successors in V, i.e., $S_v(\bar{G}) \cap V = \emptyset$, for every $v \in V_s$. Thus, from Edge Deletion and Locality we have that

$$F_v(\bar{G}, \theta + \bar{\theta}) = F_v(G,\theta), \quad \text{for every } v \in V. \tag{5.21}$$

Now, let us perform the same operation on graph (G',θ'). To this end, let us denote $G' = (V',E')$. Since u and w are not sinks, V_s is still the set of all sinks in a graph, i.e., $V_s = \{v \in V' : \Gamma^+_v(G') = \emptyset\}$. Thus, consider $\bar{G}' = (V' \cup \{t\}, E' \cup \bar{E})$. Since (G',θ') is still semi-out-regular (proportional combining preserves out-degrees), we again get that $(\bar{G}',\theta'+\bar{\theta})$ is r-out-regular graph. Moreover, as before, from Edge Deletion and Locality we have that

$$F_v(\bar{G}',\theta'+\bar{\theta}) = F_v(G',\theta'), \quad \text{for every } v \in V'. \tag{5.22}$$

Finally, observe that since $u,w \notin V_s$, graph $(\bar{G}',\theta'+\bar{\theta})$ is also the graph resulting from combining u into w in graph $(\bar{G},\theta+\bar{\theta})$, i.e., $(\bar{G}',\theta'+\bar{\theta}) = C^F_{u\to w}(\bar{G},\theta+\bar{\theta})$. Graph $(\bar{G},\theta+\bar{\theta})$ is out-regular, thus all successors of nodes u and w have equal out-degrees. Hence, from Node Combination we have

$$F_v(\bar{G}',\theta'+\bar{\theta}) = \begin{cases} F_v(\bar{G},\theta+\bar{\theta}), & \text{for every } v \in V \setminus \{u,w\} \\ F_u(\bar{G},\theta+\bar{\theta}) + F_w(\bar{G},\theta+\bar{\theta}), & \text{for } v = w. \end{cases}$$

Combining this with Eq. (5.21) and Eq. (5.22) yields the thesis. □

From this we easily get that proportional combining of nodes in semi-out-regular graphs preserves centrality.

Lemma 62. *If a centrality measure, F, defined on $\mathcal{G}$ (or $\mathcal{G}^{K(a)}$) satisfies Locality, Edge Deletion, and Node Combination, then for every graph $G = (V,E)$, weights $\theta = (b,\mu)$, and nodes $u,w \in V$ such that $(G,\theta) \in \mathcal{G}$ (or $(G,\theta) \in \mathcal{G}^{K(a)}$) and (G,θ) is semi-out-regular, it holds that*

$$F_v(C^F_{u\to w}(G,\theta)) = F_v(G,\theta), \quad \textit{for every } v \in V \setminus \{u,w\}$$

and $F_w(C^F_{u\to w}(G,\theta)) = F_u(G,\theta) + F_w(G,\theta)$.

Proof. Observe that in semi-out-regular graph all successors of either u or w that are not sinks have out-degrees equal to the out-degree of u. Thus, the thesis follows from Lemma 61. □

In the next three lemmas we will focus on node weights and their relation to centrality. First, we prove a simple property that the centrality of a source node is equal to its weight.

Lemma 63. *If a centrality measure, F, defined on $\mathcal{G}$ (or $\mathcal{G}^{K(a)}$) satisfies Edge Deletion and Baseline, then for every graph $G = (V,E)$, weights $\theta = (b,\mu)$, and node $v \in V$ such that $(G,\theta) \in \mathcal{G}$ (or $(G,\theta) \in \mathcal{G}^{K(a)}$) and $\Gamma^-_v(G) = \emptyset$ we have*

$$F_v(G,\theta) = b(v).$$

Proof. Take an arbitrary graph $G = (V,E)$, weights $\theta = (b,\mu)$, and node $v \in V$ such that $(G,\theta) \in \mathcal{G}$ (or $(G,\theta) \in \mathcal{G}^{K(a)}$) and $\Gamma^-_v(G) = \emptyset$. Observe that in graph G node v is not a successor of any node, i.e., for every $u \in V$, it holds that $v \notin S_u(G)$. Thus, in graph $G' = (V,\emptyset)$, i.e., graph G with all edges removed, from Edge Deletion we have $F_v(G',(b,\mu_\emptyset)) = F_v(G,\theta)$. In graph G' node v is isolated, thus the thesis follows from Baseline. □

Now, let us take an arbitrary node, not necessarily a source, and prove that its centrality in semi-out-regular graph is greater or equal to its node weight.

Lemma 64. *If a centrality measure, F, defined on $\mathcal{G}$ (or $\mathcal{G}^{K(a)}$) satisfies Edge Deletion, Node Combination, and Baseline, then for every graph $G = (V,E)$, weights $\theta = (b,\mu)$, and node $v \in V$ such that $(G,\theta) \in \mathcal{G}$ (or $(G,\theta) \in \mathcal{G}^{K(a)}$) and (G,θ) is semi-out-regular, it holds that*

$$F_v(G,\theta) \geq b(v).$$

Proof. Take an arbitrary graph $G = (V,E)$, weights $\theta = (b,\mu)$, and node $v \in V$ such that $(G,\theta) \in \mathcal{G}$ (or $(G,\theta) \in \mathcal{G}^{K(a)}$) and (G,θ) is semi-out-regular. Let us consider graph G' with an additional node v' with exactly the same set of outgoing edges as v in G, but without any incoming edges. Also, let us transfer the whole weight of node v into v'. Formally, let $G' = (V \cup \{v'\}, E')$, where $E' = E \cup \{(v',u) : (v,u) \in \Gamma_v^+(G)\}$. Also, let $\theta' = (b',\mu')$, where $b' = b_{-v} + b(v) \cdot \mathbb{1}_{v'}$ while $\mu'_E = \mu_E$ and $\mu'(v',u) = \mu(v,u)$, for every $(v,u) \in \Gamma_v^+(G)$. Since v' and v are out-twins, we have that $C^F_{v' \to v}(G',\theta') = (G,\theta)$. Observe that (G',θ') is also semi-out-regular. Thus, from Lemma 62 we have $F_v(G,\theta) = F_v(G',\theta') + F_{v'}(G',\theta')$. Now, from Lemma 63 we get that $F_{v'}(G',\theta') = b(v)$. Hence, from the fact that centrality is always non-negative we obtain $F_v(G,\theta) \geq b(v)$. □

In the following lemma we show that centrality of each node of semi-out-regular graph is linear with respect to the function of node weights.

Lemma 65. *If a centrality measure, F, defined on $\mathcal{G}$ (or $\mathcal{G}^{K(a)}$) satisfies Locality, Edge Deletion, Node Combination, and Baseline, then for every graph $G = (V,E)$, weights $\theta = (b,\mu)$, and node $v \in V$ such that $(G,\theta) \in \mathcal{G}$ (or $(G,\theta) \in \mathcal{G}^{K(a)}$) and (G,θ) is semi-out-regular, it holds that*

(a) $F_v(G,(b+b',\mu)) = F_v(G,\theta) + F_v(G,(b',\mu))$, *for every node weights* $b' : V \to \mathbb{R}_{\geq 0}$,

(b) $F_v(G,(x \cdot b,\mu)) = x \cdot F_v(G,\theta)$, *for every* $x \in \mathbb{R}_{\geq 0}$.

Proof. For (a), take four semi-out-regular graphs: (G,θ), $(G,(b',\mu))$, $(G,(b+b',\mu))$, and the fourth one, isomorphic to them, but with disjoint set of nodes and uniform node weights, i.e., let $(\hat{G},\hat{\theta}) = ((\hat{V},\hat{E}),(\mathbf{1},\hat{\mu}))$, where $\hat{V} = \{\hat{v} : v \in V\}$ and $\hat{V} \cap V = \emptyset$, $\hat{E} = \{(\hat{u},\hat{v}) : (u,v) \in E\}$, and $\hat{\mu}(\hat{u},\hat{v}) = \mu(u,v)$, for every $(u,v) \in E$.

Now, using graph $(\hat{G},\hat{\theta})$ we will combine together graphs (G,θ) and $(G,(b',\mu))$. To this end, consider the sum $(G+\hat{G},\theta+\hat{\theta})$. From Locality we know that the centrality of each $v \in V \cup \hat{V}$ is the same as in the original graphs. Then, let us sequentially combine node v into node $\hat{v}$, for every $v \in V$. Observe that, by Lemma 64, all nodes in $(\hat{G},\hat{\theta})$ have positive centrality. Moreover, graphs (G,θ) and $(\hat{G},\hat{\theta})$ are isomorphic (when not accounting for node weights). Thus, in result of such operation we obtain graph that has the same nodes, edges, and edge weights as $(\hat{G},\hat{\theta})$, but with node weights that are the sum of node weights in both graphs, i.e., we obtain graph $(\hat{G}',\hat{\theta}') = ((\hat{V},\hat{E}),(\hat{b}',\hat{\mu}))$, where $\hat{b}'(\hat{v}) = b(v)+1$, for every $v \in V$. From Locality and Lemma 62 we get

$$F_{\hat{v}}(\hat{G}',\hat{\theta}') = F_v(G,\theta) + F_{\hat{v}}(\hat{G},\hat{\theta}), \quad \text{for every } v \in V. \tag{5.23}$$

Next, let us consider sum of graphs $(\hat{G}',\hat{\theta}')$ and $(G,(b',\mu))$ and this time let us sequentially combine node $\hat{v}$ into node v, for every $v \in V$. Observe that, by Eq. (5.23), centralities of all nodes in $\hat{G}'$ are still positive. Moreover, again, both graphs are isomorphic (when not accounting for node weights). Thus, as before, we obtain a graph with the same nodes, edges, and edge weights, but with node weights summed, i.e., $(G'',\theta'') = ((V,E),(b+b'+\mathbf{1},\mu))$. Thus, from Locality, Lemma 62, and Eq. (5.23), for every $v \in V$, we have

$$F_v(G'',\theta'') = F_v(G,(b',\mu)) + F_{\hat{v}}(\hat{G}',\hat{\theta}') = F_v(G,(b',\mu)) + F_v(G,\theta) + F_{\hat{v}}(\hat{G},\hat{\theta}). \tag{5.24}$$

On the other hand, as we will demonstrate, graph (G'', θ'') can be also obtained when combining graphs $(G, (b+b', \mu))$ and $(\hat{G}, \hat{\theta})$. To this end, consider the sum $(G+\hat{G}, (b+b', \mu)+\hat{\theta})$ and in this graph let us sequentially combine node $\hat{v}$ into v, for every $v \in V$. Observe that in this way we also obtain graph (G'', θ''). Thus, from Locality and Lemma 62 we get $F_v(G'', \theta'') = F_v(G, (b+b', \mu)) + F_{\hat{v}}(\hat{G}, \hat{\theta})$, for every $v \in V$. Combining this with Eq. (5.24) yields

$$F_v(G, (b+b', \mu)) = F_v(G, (b', \mu)) + F_v(G, \theta), \quad \text{for every } v \in V.$$

For (b), consider function $f(x) = F_v(G, (x \cdot b, \mu))$. From (a) we know that function f is additive, i.e., $f(x+y) = f(x) + f(y)$, for every $x, y \in \mathbb{R}_{\geq 0}$. From the definition of centrality measure we know that it is also non-negative, i.e., we have $f(x) \geq 0$, for every $x \in \mathbb{R}_{\geq 0}$. This implies that the function is of the form $f(x) = x \cdot r$, for some $r \in \mathbb{R}_{\geq 0}$ [23]. Since, $f(1) = F_v(G, \theta)$, we get that $F_v(G, (x \cdot b, \mu)) = x \cdot F_v(G, \theta)$, for every $v \in V$. □

PageRank and Katz centrality are both parameterized by a decay factor a. In the next lemma, we pinpoint this value. Specifically, for every centrality measure F that satisfies Locality, Edge Deletion, Node Combination, and Baseline, we define its decay factor, a_F, as the centrality of node v in graph $((\{u,v\}, \{(u,v)\}), ([1,0],[1]))$.

Lemma 66. *If a centrality measure, F, defined on $\mathcal{G}$ (or $\mathcal{G}^{K(a)}$) satisfies Locality, Edge Deletion, Node Combination, and Baseline, then there exists a constant $a_F \in \mathbb{R}_{\geq 0}$ such that for every $x \in \mathbb{R}_{\geq 0}$ and nodes u, v, it holds that*

$$F_v((\{u,v\}, \{(u,v)\}), ([x,0],[1])) = a_F \cdot x.$$

Proof. Let $(G, \theta) = ((\{u,v\}, \{(u,v)\}), ([1,0],[1]))$ and $a_F = F_v(G, \theta)$. Since (G, θ) is semi-out-regular, the thesis follows directly from Lemma 65b. □

Having established basic properties that follow from our axioms, we are now ready to move to the part of the proof devoted to the notion of node's *profit* from an edge, i.e., the part of the centrality of a node that it gets from the particular incoming edge. First, let us formally define the profit function as a centrality of a node in a simple graph with three nodes and one or two edges.

Definition 8. *The* profit function *of centrality F is a function such that for every $x \in \mathbb{R}_{\geq 0}$ and $y, z \in \mathbb{R}_{>0}$ such that $y \leq z$ returns the value $p_F(x,y,z) = F_v(G, \theta)$, where*

$$(G, \theta) = \begin{cases} ((\{u,v,w\}, \{(u,v),(u,w)\}), ([x,0,0],[y,z-y])), & \text{if } y < z, \\ ((\{u,v,w\}, \{(u,v)\}), ([x,0,0],[y])), & \text{otherwise.} \end{cases}$$

In the following two lemmas, we will take F that satisfies Locality, Edge Deletion, Node Combination, and Baseline and prove that its profit function is equal to that of PageRank if it satisfies also Edge Multiplication or to this of Katz centrality if it satisfies Edge Compensation. Let us start with Edge Multiplication and PageRank.

Lemma 67. *If a centrality measure, F, defined on $\mathcal{G}$ satisfies Locality, Edge Deletion, Node Combination, Edge Multiplication, and Baseline, then for every $x, y, z \in \mathbb{R}_{\geq 0}$ such that $z \geq y$, it holds that*

$$p_F(x,y,z) = p_{PR^{a_F}}(x,y,z) = a_F \cdot x \cdot y / z.$$

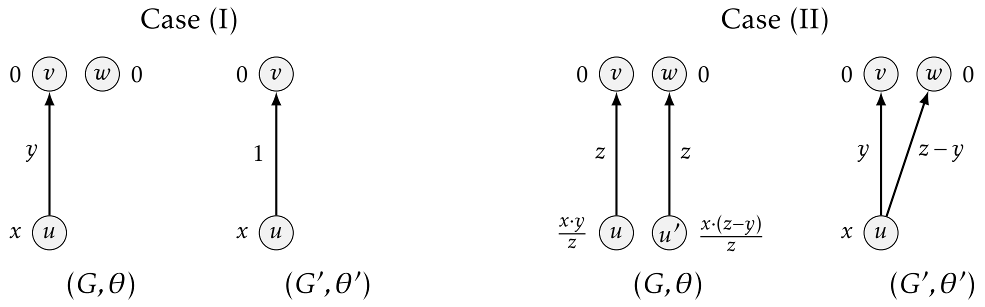


Figure 5.9: Graphs considered in the proof of Lemmas 67 and 68. The weight of each node and edge is shown.

Proof. It is easy to check that from PageRank recursive equation (Eq. (2.5)) we get $p_{PR^{a_F}}(x,y,z) = a_F \cdot x \cdot y/z$. Thus, let us focus on proving that $p_F(x,y,z) = a_F \cdot x \cdot y/z$ as well. To this end, we will consider two cases: the first one (I) in which $y = z$, i.e., a graph with only one edge, and the second one (II) in which $y < z$, i.e., a graph with two edges. We illustrate consecutive graphs in Fig. 5.9.

(I) In the case where $y = z$, let us consider graph from Definition 8 of the form $(G,\theta) = ((\{u,v,w\},\{(u,v)\}),([x,0,0],[y]))$. If we remove node w and change the weight of (u,v) to 1, we obtain $(G',\theta') = ((\{u,v\},\{(u,v)\}),([x,0],[1]))$. From Lemma 66 we have $F_v(G',\theta') = a_F \cdot x$. Hence, from Locality and Edge Multiplication we obtain

$$F_v(G,\theta) = F_v(G',\theta') = a_F \cdot x = PR_v^{a_F}(G,\theta). \tag{5.25}$$

(II) In the case of $y < z$, we begin with two pairs of nodes connected by an edge, i.e., let $(G,\theta) = ((\{u,u',v,w\},\{(u,v),(u',w)\}),([x\cdot y/z, x\cdot(z-y)/z,0,0],[z,z]))$. Observe that from Locality and Eq. (5.25) we have that $F_v(G,\theta) = a_F \cdot x \cdot y/z$.

Now, from Lemma 63 we get that $F_u(G,\theta) = x \cdot y/z$ and $F_{u'}(G,\theta) = x \cdot (z-y)/z$. Therefore, if we now proportionally combine node u' into u in G we obtain graph $(G',\theta') = C^F_{u'\to u}(G,\theta) = ((\{u,v,w\},\{(u,v),(u,w)\}),([x,0,0],[y,z-y]))$, which is a graph from Definition 8. Since (G,θ) is semi-out-regular, from Lemma 62 we get that $p_F(x,y,z) = F_v(G',\theta') = F_v(G,\theta) = a_F \cdot x \cdot y/z$. This concludes the proof. □

Now, let us move to Edge Compensation and Katz centrality.

Lemma 68. *If a centrality measure, F, defined on $\mathcal{G}^{K(a)}$ satisfies Locality, Edge Deletion, Node Combination, Edge Compensation, and Baseline, then for every $x,y,z \geq 0$ such that $z \geq y$, it holds that*

$$p_F(x,y,z) = p_{K^{a_F}}(x,y,z) = a_F \cdot x \cdot y.$$

Proof. The proof follows in a similar fashion to the proof of Lemma 67. The fact that $p_{K^{a_F}}(x,y,z) = a_F \cdot x \cdot y$ comes directly from Katz centrality recursive equation (Eq. (2.3)). Thus, let us prove that $p_F(x,y,z) = a_F \cdot x \cdot y$ as well. To this end, we consider two cases: the first one (I) in which $y = z$, and the second one (II) in which $y < z$. As before, the graphs used in the proof are depicted in Fig. 5.9.

(I) For $y = z$, a graph from Definition 8 is $(G,\theta) = ((\{u,v,w\},\{(u,v)\}),([x,0,0],[y]))$. By removing node w and changing the weight of edge (u,v) to 1, we obtain graph $(G',\theta') = ((\{u,v\},\{(u,v)\}),([x,0],[1]))$. Lemma 66 yields $F_v(G',\theta') = a_F \cdot x$. Since F satisfies Locality and Edge Compensation, we get

$$F_v(G,\theta) = y \cdot F_v(G',\theta') = a_F \cdot x \cdot y = K_v^{a_F}(G,\theta). \tag{5.26}$$

(II) When $y < z$, we consider a graph with two pairs of nodes connected by an edge, i.e., let $(G,\theta) = ((\{u,u',v,w\},\{(u,v),(u',w)\}),([x\cdot y/z, x\cdot(z-y)/z, 0, 0],[z,z]))$. From Eq. (5.26) and Locality we have that $F_v(G,\theta) = a_F \cdot x \cdot y$.

From Lemma 63 we get $F_u(G,\theta) = x \cdot y/z$ and $F_{u'}(G,\theta) = x\cdot(z-y)/z$. Thus, by combining u' into u we obtain $(G',\theta') = ((\{u,v,w\},\{(u,v),(u,w)\}),([x,0,0],[y,z-y]))$, which is a graph from Definition 8. Since (G,θ) is semi-out-regular graph, from Lemma 62 we obtain that $p_F(x,y,z) = F_v(G',\theta') = F_v(G,\theta) = a_F \cdot x \cdot y$. □

In the next two lemmas, we will prove that if a centrality measure F satisfies Locality, Edge Deletion, Node Combination, Edge Multiplication or Edge Compensation, and Baseline, then the centrality of a node in a semi-out-regular graph is equal to the weight of this node plus the sum of profits its receives from its incoming edges. In other words, we will show that F satisfies recursive equation

$$F_v(G,\theta) = b(v) + \sum_{u\in P_v^1(G)} p_F(F_u(G,\theta), \mu(u,v), \deg_u^+(G,\theta)).$$

First, we assume that the node in question does not have a loop.

Lemma 69. *If a centrality measure, F, defined on $\mathcal{G}$ (or $\mathcal{G}^{K(a)}$) satisfies Locality, Edge Deletion, Node Combination, Edge Compensation or Edge Multiplication, and Baseline, then for every graph $G = (V,E)$, weights $\theta = (b,\mu)$, and node $v \in V$ such that (G,θ) is semi-out-regular, $(G,\theta) \in \mathcal{G}$ (or $(G,\theta) \in \mathcal{G}^{K(a)}$), and $(v,v) \notin E$, it holds that*

$$F_v(G,\theta) = b(v) + \sum_{u\in P_v^1(G)} p_F(F_u(G,\theta), \mu(u,v), \deg_u^+(G,\theta)).$$

Proof. We will prove the thesis by induction on the number of incoming edges of node in question. Consider an arbitrary graph $G = (V,E)$, weights $\theta = (b,\mu)$, and node $v \in V$ such that (G,θ) is semi-out-regular, $(G,\theta) \in \mathcal{G}$ (or $(G,\theta) \in \mathcal{G}^{K(a)}$), and $(v,v) \notin E$. If node v does not have any incoming edges, then the thesis follows from Lemma 63. Therefore, we will focus on the case in which it has at least one incoming edge. Let us denote one of them as (u,v), where $u \neq v$. In what follows, through a series of graph operation we will show that the centrality of node v can be expressed as a sum of its profit from edge (u,v), i.e., $p_F(F_u(G,\theta), \mu(u,v), \deg_u^+(G,\theta))$, and the rest that is known due to the inductive assumption.

If $F_u(G,\theta) = 0$, then the profit of node v from edge (u,v) is equal to zero, i.e., $p_F(F_u(G,\theta), \mu(u,v), \deg_u^+(G,\theta)) = 0$. Thus, we will show that the centrality of node v is equal to the sum of its profits from other edges plus its weight. To this end, let $(G^\uparrow, \theta^\uparrow) = ((\{u',v'\},\{(u',v')\}),([1,0],[\deg_u^+(G,\theta)]))$ be a small two-node graph, which we add to our original graph (G,θ). Then, observe that from Locality we have that $F_u(G+G^\uparrow, \theta+\theta^\uparrow) = F_u(G,\theta) = 0$ and from Lemma 63 we have $F_{u'}(G+G^\uparrow,\theta+\theta^\uparrow) = 1$. Thus, when in this summed graph we combine node u into u', we get that the original outgoing edges of node u, including edge (u,v), are removed. Formally, let us denote the obtained graph as $(G',\theta') = C^F_{u\to u'}(G+G^\uparrow,\theta+\theta^\uparrow)$. Since, $(G+G^\uparrow,\theta+\theta^\uparrow)$ is semi-out-regular, from Lemma 62 we have $F_v(G',\theta') = F_v(G+G^\uparrow,\theta+\theta^\uparrow)$ and by Locality, $F_v(G+G^\uparrow,\theta+\theta^\uparrow) = F_v(G,\theta)$. Since in graph (G',θ') node v has one incoming edge less, the thesis follows from the inductive assumption.

Hence, let us focus on the case in which $F_u(G,\theta) > 0$. Consider graph (G',θ') in which we split node u into two nodes: u', with all of its original outgoing edges and no incoming edges, and u, with all of its original incoming edges and only one new outgoing edge (u,v'). See Fig. 5.10 for an illustration. Additionally, let us give node u' weight equal to $F_u(G,\theta)$. Formally, let $G' = (V',E')$ be a graph

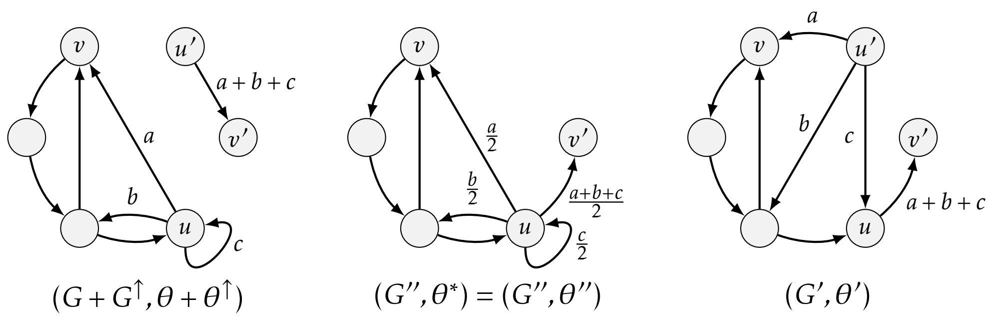


Figure 5.10: An illustration to the first part of the proof of Lemma 69 for an example graph, (G,θ). The weights of the outgoing edges of u and u' are shown.

in which $V' = V \cup \{u',v'\}$ and $E' = E \setminus \Gamma_u^+(G) \cup \{(u',w) : (u,w) \in \Gamma_u^+(G)\} \cup \{(u,v')\}$, while $\theta' = (b',\mu')$, where $b'_V = b_V$, $b'(v') = 0$, $b'(u') = F_u(G,\theta)$ and $\mu'_{-\Gamma_u^+(G)} = \mu_{-\Gamma_u^+(G)}$, $\mu'(u,v') = \deg_u^+(G,\theta)$, and $\mu'(u',w) = \mu(u,w)$, for every $(u,w) \in \Gamma_u^+(G)$.

Now, let us combine node u' into node u. The graph that we obtain is not exactly our original graph G. More in detail, it is graph $G'' = (V \cup \{v'\}, E \cup \{(u,v')\})$ with weights $\theta'' = (b'',\mu'')$, where $b''(v') = 0$, $b''(u) = b(u) + F_u(G)$ and $b''(w) = b(w)$, for every $w \in V \setminus \{u\}$. From Lemma 63 we have that $F_{u'}(G',\theta') = F_u(G,\theta)$, hence we know that also

$$\mu''(u,v') = \frac{\deg_u^+(G,\theta)\cdot F_u(G',\theta')}{F_u(G,\theta)+F_u(G',\theta')},\ \mu''(u,w) = \frac{\mu(u,w)\cdot F_u(G,\theta)}{F_u(G,\theta)+F_u(G',\theta')} \text{ for every } (u,w)\in\Gamma_u^+(G),$$

and $\mu''(e) = \mu(e)$, for every $e \in E \setminus \Gamma_u^+(G)$ (if $F_u(G',\theta') = 0$, the edge (u,v') that is supposed to have zero weight is removed). Fig. 5.10 illustrates the case in which $F_u(G,\theta) = F_u(G',\theta')$. We will prove that it is always the case.

To this end, consider adding to our original graph, (G,θ), a simple graph consisting of nodes u' and v' and edge between them, and then combining node u' into node u. Furthermore, let u' have a weight equal to centrality $F_u(G',\theta')$. Formally, let $(G^{\uparrow},\theta^{\uparrow}) = ((\{u',v'\},\{(u',v')\}),([F_u(G',\theta'),0],[\deg_u^+(G,\theta)]))$. Consequently, let us combine node u' into u in graph $(G+G^{\uparrow},\theta+\theta^{\uparrow})$. As a result, we also obtain graph G'' with possibly different weights, i.e., let $(G'',\theta^*) = C^F_{u'\to u}(G+G^{\uparrow},\theta+\theta^{\uparrow})$. Let $\theta^* = (b^*,\mu^*)$. Observe that $b^*(v') = 0$, $b^*(u) = b(u) + F_u(G',\theta')$, and $b^*(w) = b(w)$, for every $w \in V \setminus \{u\}$. As for edge weights, observe that from Lemma 63 we get $F_{u'}(G+G^{\uparrow},\theta+\theta^{\uparrow}) = F_u(G',\theta')$ and from Locality $F_u(G+G^{\uparrow},\theta+\theta^{\uparrow}) = F_u(G,\theta)$. Thus,

$$\mu^*(u,v') = \frac{\deg_u^+(G,\theta)\cdot F_u(G',\theta')}{F_u(G,\theta)+F_u(G',\theta')},\ \mu^*(u,w) = \frac{\mu(u,w)\cdot F_u(G,\theta)}{F_u(G,\theta)+F_u(G',\theta')} \text{ for every } (u,w)\in\Gamma_u^+(G),$$

and $\mu^*(e) = \mu(e)$, for every $e \in E \setminus \Gamma_u^+(G)$ (if $F_u(G',\theta') = 0$, the edge (u,v') that is supposed to have zero weight is removed). Therefore, $\mu^* = \mu''$.

This means that the only possible difference between θ^* and θ'' is the weight of node u, i.e., $b^*(u) = b(u) + F_u(G',\theta')$ and $b''(u) = b(u) + F_u(G,\theta)$. However, the centrality of node u in graph G'' with both weights is the same: by Lemma 62, from combining u' into u in $(G+G^{\uparrow},\theta+\theta^{\uparrow})$, we get $F_u(G'',\theta^*) = F_u(G',\theta') + F_u(G,\theta)$ and, from combining u' into u in (G',θ'), we have that $F_u(G'',\theta'') = F_u(G,\theta) + F_u(G',\theta')$. Let us prove that this implies that also $b^*(u) = b''(u)$. For assume otherwise. Without loss of generality, let us assume that $b^*(u) > b''(u)$. Then, consider weights $\theta^* - \theta'' = (b^* - b'', \mu'')$, i.e., the same edge weights as in θ^* and θ'', but with the node weight of each node being a difference between its node weight in θ^* and θ''. Since

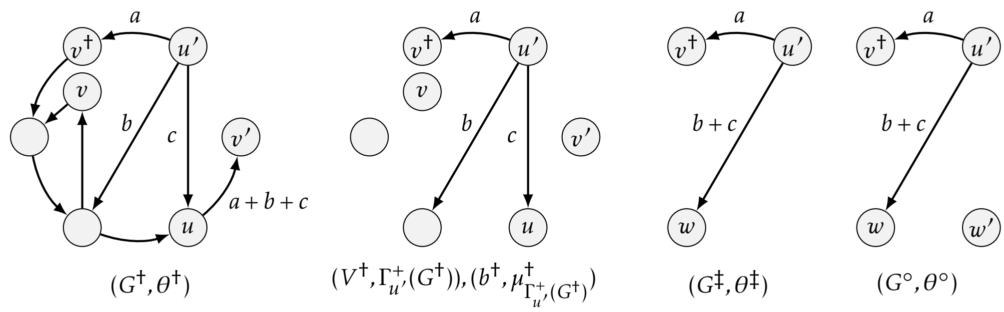


Figure 5.11: An illustration to the second part of the proof of Lemma 69 for graph (G, θ) from Fig. 5.10.

$b^*(u) - b''(u) > 0$, we know that the weight of u under this new weights is positive, i.e., $(b^* - b'')(u) > 0$. Thus, from Lemma 64 we have that $F_u(G'', \theta^* - \theta'') > 0$. However, from Lemma 65a we have that $F_u(G'', \theta'') + F_u(G'', \theta^* - \theta'') = F_u(G'', \theta^*)$. Since $F_u(G'', \theta'') = F_u(G'', \theta^*)$ we arrive at a contradiction. Thus, indeed, $b^*(u) = b''(u)$ which implies that $(G'', \theta^*) = (G'', \theta'')$. Therefore, from Lemma 62 we have that

$$F_w(G, \theta) = F_w(G'', \theta^*) = F_w(G'', \theta'') = F_w(G', \theta'), \quad \text{for every } w \in V. \tag{5.27}$$

In the remainder of the proof, we will split node v into two nodes with the same set of outgoing edges as v: node $v^\dagger$ with one incoming edge, $(u', v^\dagger)$, and zero weight, and node v with the original weight of node v and all of its incoming edges except for (u', v) (see Fig. 5.11 for an illustration). As we will show, in this new graph, centrality of $v^\dagger$ is equal to profit of node v from edge (u, v), i.e., $P_F(F_u(G, \theta), \mu(u, v), \deg^+_u(G, \theta))$, and, by the inductive assumption, centrality of v is equal to the sum of profits of all other incoming edges of v and its weight. Then, by Node Combination, this will prove the induction hypothesis.

Formally, let us consider graph $G^\dagger = (V^\dagger, E^\dagger)$, where we have $V^\dagger = V' \cup \{v^\dagger\}$ and $E^\dagger = E' \setminus \{(u', v)\} \cup \{(u', v^\dagger)\} \cup \{(v^\dagger, w) : (v, w) \in \Gamma^+_v(G')\}$. Moreover, let $\theta^\dagger = (b^\dagger, \mu^\dagger)$, where $b^\dagger_{V'} = b'_{V'}$ and $b^\dagger(v^\dagger) = 0$, while $\mu^\dagger(u', v^\dagger) = \mu'(u', v)$, $\mu^\dagger(v^\dagger, w) = \mu'(v, w)$, for every $(v, w) \in \Gamma^+_v(G)$, and also $\mu^\dagger(e)_{E' \setminus \{(u', v)\}} = \mu'(e)_{E' \setminus \{(u', v)\}}$. Clearly, if we combine node $v^\dagger$ into node v in graph $(G^\dagger, \theta^\dagger)$ we obtain graph (G', θ'). Thus, from Lemma 62 we have $F_v(G', \theta') = F_v(G^\dagger, \theta^\dagger) + F_{v^\dagger}(G^\dagger, \theta^\dagger)$ and from Eq. (5.27) we have $F_v(G, \theta) = F_v(G^\dagger, \theta^\dagger) + F_{v^\dagger}(G^\dagger, \theta^\dagger)$. Observe that node v has less incoming edges in graph $G^\dagger$ than it had in graph G. Hence, by the inductive assumption,

$$F_v(G, \theta) = F_{v^\dagger}(G^\dagger, \theta^\dagger) + b^\dagger(v) + \sum_{w \in P^1_v(G) \setminus \{u\}} p_F(F_w(G^\dagger, \theta^\dagger), \mu^\dagger(w, v), \deg^+_w(G^\dagger, \theta^\dagger)).$$

Now, from Lemma 62 and Eq. (5.27) we obtain $F_w(G^\dagger, \theta^\dagger) = F_w(G, \theta)$, for every $w \in V \setminus \{u, v\}$. Also, we have that $\mu^\dagger(w, v) = \mu(w, v)$ and $\deg^+_w(G^\dagger, \theta^\dagger) = \deg^+_w(G, \theta)$, for every $w \in P^1_v(G) \setminus \{u\}$, and $b^\dagger(v) = b(v)$ as well. Therefore, it remains to prove that $F_{v^\dagger}(G^\dagger, \theta^\dagger) = p_F(F_u(G, \theta), \mu(u, v), \deg^+_u(G, \theta))$.

To this end, observe that the only predecessor of node $v^\dagger$ in graph $G^\dagger$ is node u'. Thus, let us remove all edges that are not outgoing edges of u'. From Edge Deletion we know that the centrality of node $v^\dagger$ is unchanged by this operation, i.e., $F_{v^\dagger}(G^\dagger, \theta^\dagger) = F_{v^\dagger}((V^\dagger, \Gamma^+_{u'}(G^\dagger)), (b^\dagger, \mu^\dagger_{\Gamma^+_{u'}(G^\dagger)}))$.

If u' has one outgoing edge, i.e., $\Gamma^+_{u'}(G^\dagger) = \{(u', v^\dagger)\}$, then, indeed, from Locality and Definition 8 we get that $F_{v^\dagger}(G^\dagger, \theta^\dagger) = p_F(F_u(G, \theta), \mu(u, v), \deg^+_u(G, \theta))$. Thus, let us assume otherwise. Then, observe that in graph $(V^\dagger, \Gamma^+_{u'}(G^\dagger))$ every node except for node u' does not have any outgoing edges. Therefore, let us add a new isolated node with weight 1 to this graph, namely, $w \notin V$, and then sequentially combine all the nodes that are not u' nor $v^\dagger$ into node w. Formally, let us denote graph $(G^\ddagger,\theta^\ddagger) = ((\{u',v^\dagger,w\},\{(u',v^\dagger),(u',w)\}),([F_u(G,\theta),0,1+x],[\mu(u,v),\deg^+_u(G,\theta)-\mu(u,v)]))$, where $x = \sum_{w'\in V^\dagger\setminus\{u',v^\dagger\}} b^\dagger(w')$. Then, from Locality axiom and Lemma 62 we obtain that $F_{v^\dagger}((V^\dagger,\Gamma^+_{u'}(G^\dagger)),(b^\dagger,\mu^\dagger_{\Gamma^+_{u'}(G^\dagger)}) = F_{v^\dagger}(G^\ddagger,\theta^\ddagger)$.

Finally, the only difference between graph $(G^\ddagger,\theta^\ddagger)$ and the graph from Definition 8 is the weight of node w. Therefore, let us split node w into two nodes, w and w', such that w has incoming edge (u',w) and zero weight whereas w' has the weight of original node w and no incoming edges. Formally, let us define graph $(G^\circ,\theta^\circ) = ((\{u',v^\dagger,w,w'\},\{(u',v^\dagger),(u',w)\})([F_u(G,\theta),0,0,1+x],[\mu(u,v),\deg^+_u(G,\theta)-\mu(u,v)]))$. Clearly, if we combine w' into w in this graph, then we obtain graph $(G^\ddagger,\theta^\ddagger)$. Thus, from Lemma 62 we obtain that $F_{v^\dagger}(G^\ddagger,\theta^\ddagger) = F_{v^\dagger}(G^\circ,\theta^\circ)$. On the other hand, without node w' graph (G°,θ°) is a graph from Definition 8. Therefore, from Locality we obtain that $F_{v^\dagger}(G^\circ,\theta^\circ) = p_F(F_u(G,\theta),\mu(u,v),\deg^+_u(G,\theta))$. This concludes the proof. □

In the next lemma we relax the assumption that node v does not have a loop.

Lemma 70. *If a centrality measure, F, defined on $\mathcal{G}$ (or $\mathcal{G}^{K(a)}$) satisfies Locality, Edge Deletion, Node Combination, Edge Compensation or Edge Multiplication, and Baseline, then for every graph $G = (V,E)$, weights $\theta = (b,\mu)$, and node $v \in V$ such that (G,θ) is semi-out-regular and $(G,\theta) \in \mathcal{G}$ (or $(G,\theta) \in \mathcal{G}^{K(a)}$), it holds that*

$$F_v(G,\theta) = b(v) + \sum_{u\in P^1_v(G)} p_F(F_u(G,\theta),\mu(u,v),\deg^+_u(G,\theta)).$$

Proof. First, we focus on a graph with only two nodes and (at most) two edges: one connecting the nodes and a loop around the start (see Fig. 5.12 for an illustration). Formally, let $(G,\theta) = ((\{v,w\},\{(v,v),(v,w)\}),([b_v,0],[y,z-y]))$, for some $b_v \in \mathbb{R}_{\geq 0}$ and $y,z \in \mathbb{R}_{>0}$ such that $z \geq y$ (in the case of $z = y$, we consider a graph with only one edge, (v,v), but the proof is the same). Also, let us denote $x = F_v(G,\theta)$. We want to prove that $x = b_v + p_F(x,y,z)$. Observe that if $b_v = 0$, then from Lemma 65b we know that $x = F_v(G,\theta) = 0$. Also, for both Katz centrality and PageRank profit function for $x = 0$ is equal to 0. Hence, the thesis follows from Lemma 67 or 68 (depending on the satisfied axiom).

Thus, let us assume otherwise, i.e., that $b_v > 0$. By Lemma 64, this means that also $x = F_v(G,\theta) > 0$. Let us consider graph such as in Definition 8 with the weight of the source equal to $p\cdot x$, for an arbitrary constant $p \in \mathbb{R}_{\geq 0}$ (see Fig. 5.12). Formally, take $(G',\theta'_p) = ((\{v',u,w'\},\{(v',u),(v',w')\}),([p\cdot x,0,0],[y,z-y]))$. Now, let us add both graphs together to obtain $(G+G',\theta+\theta'_p)$. From Lemma 63 we have that $F_{v'}(G+G',\theta+\theta'_p) = p\cdot x$ and from Locality $F_v(G+G',\theta+\theta'_p) = x$. Thus, when we proportionally combine v' into v and then w' into w we obtain graph $(G'',\theta''_p) = ((\{v,u,w\},\{(v,v),(v,u),(v,w)\}),([b_v+p\cdot x,0,0],[y/(1+p),y\cdot p/(1+p),z-y]$. Observe that graph $(G+G',\theta+\theta'_p)$ is semi-out-regular. Thus, from Lemma 62 we get that

$$F_v(G'',\theta''_p) = F_{v'}(G',\theta'_p) + F_v(G,\theta) = (1+p)\cdot x. \tag{5.28}$$

In what follows, we will show that there exists $p \in \mathbb{R}_{\geq 0}$ such that graph (G'',θ''_p) can be obtained also in another way. To this end, let us consider yet another graph, $(G^*,\theta^*) = ((\{v,v',u,w\},\{(v,v'),(v,w)(v',u),(v',w)\}),([x,0,0,0],[y,z-y,y,z-y]))$. See Fig. 5.12 for an illustration. Observe that v' is not a successor of itself in G^*. Hence,

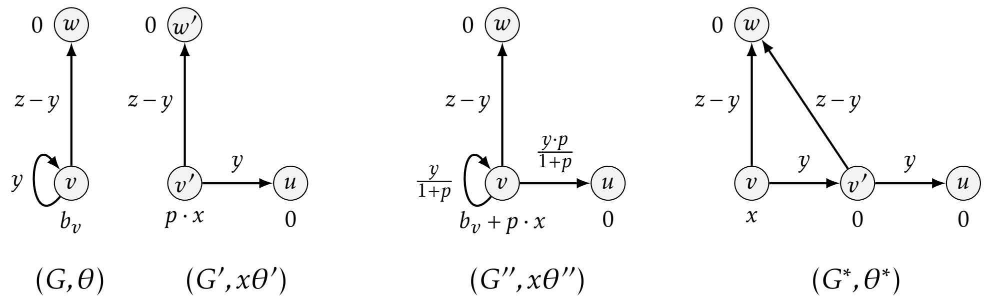


Figure 5.12: Graphs considered in the first part of the proof of Lemma 67.

if we remove node u and edges $(v',u),(v',w)$ from graph G^*, then, by Edge Deletion and Locality, the centrality of v' will not change. What remains is a graph from Definition 8. Thus, $F_{v'}(G^*,\theta^*) = p_F(x,y,z)$. To get a particular value of $p_F(x,y,z)$, let us consider two cases depending on an axiom that centrality F satisfies: (I) Edge Multiplication or (II) Edge Compensation.

(I) If centrality F satisfies Edge Multiplication, then from Lemma 67 we obtain that $F_{v'}(G^*,\theta^*) = a_F \cdot x \cdot y/z$. Let us denote the constant $a_F \cdot y/z$ as $\bar{p}$ (then we have that simply $F_{v'}(G^*,\theta^*) = \bar{p} \cdot x$). From Lemma 63 we get $F_v(G^*,\theta^*) = x = F_v(G,\theta)$. Therefore, when in graph (G^*,θ^*) we proportionally combine node v' into node v, we obtain graph G'' once again, but now with possibly different weights, i.e, $(G'',\theta'') = ((\{v,u,w\},\{(v,v),(v,u),(v,w)\}),([x,0,0],[y/(1+\bar{p}),y\cdot\bar{p}/(1+\bar{p}),z-y]))$. Moreover, from Lemma 62 we get that

$$F_v(G'',\theta'') = F_v(G^*,\theta^*) + F_{v'}(G^*,\theta^*) = (1+\bar{p})\cdot x.$$

Thus, by Eq. (5.28), if we take $p = \bar{p}$, we get $F_v(G'',\theta'') = F_v(G'',\theta''_{\bar{p}})$. Observe also that edge weights in both θ'' and $\theta''_{\bar{p}}$ are the same. This means that (G'',θ'') and $(G'',\theta''_{\bar{p}})$ are two graphs with the same nodes, edges, and edge weights, and in both graphs only v has a positive node weight. By Lemma 65b, this means that the proportion of the weights of v is equal to the proportion of its centralities in both graphs. However, as we noted above the centrality of v is equal in (G'',θ'') and $(G,\theta''_{\bar{p}})$. Hence, the weight of v is also equal in both graphs and we get $\theta'' = \theta''_{\bar{p}}$. Now, looking at the weight of node v as defined in θ'' and $\theta''_{\bar{p}}$ we get $x = b_v + \bar{p}\cdot x$. Since we took $\bar{p} = a_F \cdot y/z = p_F(x,y,z)/x$, we get that

$$F_v(G,\theta) = x = b_v + p_F(x,y,z),$$

which concludes this part of the proof. Observe that we also obtain that the centrality of v is a linear function of its weight, i.e., $F_v(G,\theta) = b_v/(1 - a_F\cdot y/z)$.

(II) If F satisfies Edge Compensation instead of Edge Multiplication, then from Lemma 68 we get that $F_{v'}(G'',\theta'') = a_F \cdot y \cdot x$. Also, if instead of $\bar{p} = a_F\cdot y/z$ we take $\bar{p} = a_F \cdot y$, then the proof follows analogously to the proof of case (I). Furthermore, as in case (I), we obtain that the centrality of node v is a linear function of its weight, i.e., $F_v(G,\theta) = b_v/(1 - a_F\cdot y)$.

As a result, from both cases (I) and (II) we can conclude that for every $y,z \in \mathbb{R}_{>0}$ such that $z \geq y$, there exists a constant $c_{F,y,z}$ such that for every $b_v \in \mathbb{R}_{\geq 0}$, it holds that

$$\begin{cases} F_v((\{v,w\},\{(v,v),(v,w)\}),([b_v,0],[y,z-y])) = b_v \cdot c_{F,y,z}, & \text{if } z > y, \\ F_v((\{v,w\},\{(v,v)\}),([b_v,0],[y])) = b_v \cdot c_{F,y,z}, & \text{otherwise} \end{cases} \tag{5.29}$$

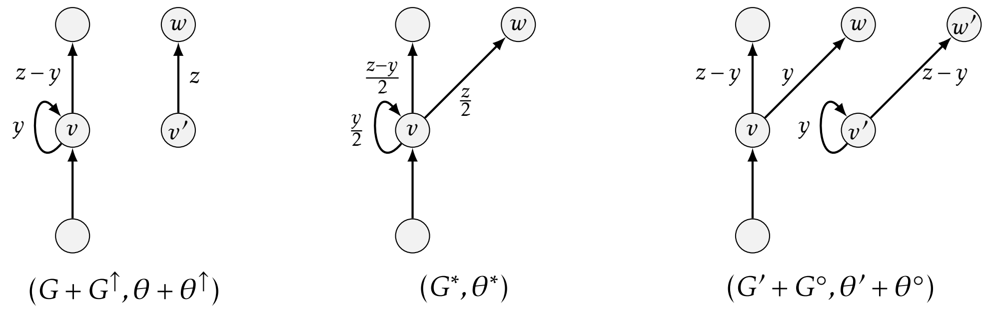


Figure 5.13: An illustration to the second part of the proof of Lemma 70 for an example graph, (G,θ). The weights of outgoing edges of nodes v and v' are shown.

(as noted at the beginning of the proof, if $z = y$, we consider a graph with only one edge, (v,v), but the proof is the same).

In the remainder of the proof, let us consider an arbitrary semi-out-regular graph $(G,\theta) = ((V,E),(b,\mu))$ and its arbitrary node $v \in V$. If v does not have a loop, then the thesis follows from Lemma 69. Hence, let assume otherwise, i.e., $(v,v) \in E$. Let us denote $y = \mu(v,v)$ and $z = \deg^+_v(G,\theta)$. Now, let us consider graph $(G',\theta') = ((V',E'),(b',\mu'))$ that is a modification of (G,θ) in which node v does not have a loop but it has a new outgoing edge to a new node, $w \notin V$, and the weight of node v is increased by $p_F(F_v(G,\theta),y,z)$ (see Fig. 5.13 for an illustration). Formally, let $V' = V \cup \{w\}$, $E' = E \setminus \{(v,v)\} \cup \{v,w\}$, and $b' = b + p_F(F_v(G,\theta),y,z) \cdot \mathbb{1}_v$ while also $\mu'_{-(v,w)} = \mu_{-(v,v)}$ and $\mu'(v,w) = \mu(v,v)$. In this way, $\deg^+_v(G',\theta') = \deg^+_v(G,\theta)$ and graph (G',θ') is still semi-out-regular. Moreover, since $(v,v) \notin E'$, from Lemma 69 we obtain that $F_v(G',\theta') = b'(v) + \sum_{u \in P^1_v(G)\setminus\{v\}} p_F(F_u(G',\theta'),\mu'(u,v),\deg^+_u(G',\theta'))$.

Observe that we have $\mu'(u,v) = \mu(u,v)$ and $\deg^+_u(G',\theta') = \deg^+_u(G,\theta)$, for every $u \in P^1_v(G) \setminus \{v\}$. Also, $b'(v) = b(v) + p_F(F_v(G,\theta),y,z)$. Hence, we get that

$$F_v(G',\theta') = b(v) + p_F(F_v(G,\theta),y,z) + \sum_{u \in P^1_v(G)\setminus\{v\}} p_F(F_u(G',\theta'),\mu(u,v),\deg^+_u(G,\theta)). \tag{5.30}$$

Thus, in what follows, through a series of graph operations, we will show that $F_u(G,\theta) = F_u(G',\theta')$, for every $u \in V$. Combined with the above equation this will yield the thesis.

To this end, we parametrize both graphs (G,θ) and (G',θ') by adding arbitrary weights $r,s \in \mathbb{R}_{\geq 0}$ to node v in both of them, respectively. Formally, let us denote $(G,\theta_r) = (G,(b{+}r{\cdot}\mathbb{1}_v,\mu))$ and $(G',\theta'_s) = (G',(b'{+}s{\cdot}\mathbb{1}_v,\mu'))$. To both of them we will add a small two-node graph with which we will proportionally combine their nodes in order to obtain the same graph (G^*,θ^*) in both cases. We will start with graph (G,θ_r). Let $v' \notin V'$ and denote graph $(G^\uparrow,\theta^\uparrow_r) = ((\{v',w\},\{(v',w)\}),([F_v(G,\theta_r),0],[z]))$. In the sum, $(G + G^\uparrow,\theta_r + \theta^\uparrow_r)$, let us proportionally combine node v' into v, i.e., let $(G^*,\theta^*_r) = C^F_{v'\to v}(G + G^\uparrow,\theta_r + \theta^\uparrow_r)$. Let us denote $\theta^*_r = (b^{r*},\mu^{r*})$. Observe that from Locality we obtain that $F_v(G + G^\uparrow,\theta_r + \theta^\uparrow_r) = F_v(G,\theta_r)$ and from Lemma 63 we have that $F_{v'}(G + G^\uparrow,\theta_x + \theta^\uparrow_r) = F_v(G,\theta_r)$ as well. This has two implications: first, since $(G + G^\uparrow,\theta_r + \theta^\uparrow_r)$ is semi-out-regular, from Lemma 62 we get that

$$F_u(G^*,\theta^*_r) = \begin{cases} F_u(G,\theta_r), & \text{for every } u \in V \setminus \{v\}, \\ 2 \cdot F_v(G,\theta_r), & \text{for } u = v; \end{cases} \tag{5.31}$$

second, this means that the weights of outgoing edges of v and v' are divided by two in graph (G^*,θ^*_r), i.e, $\mu^{r*}_{\Gamma^+_v(G)} = \mu_{\Gamma^+_v(G)}/2$ and $\mu^{r*}(v,w) = z/2$, and the weights of the remaining edges are unchanged, $\mu^{r*}_{-\Gamma^+_v(G)} = \mu_{-\Gamma^+_v(G)}$. See Fig. 5.13 for an illustration.

Now, let us move to graph (G',θ'_s). Here, take node $w' \notin V' \cup \{v'\}$ and consider adding graph $(G^\circ,\theta^\circ_s) = ((\{v',w'\},\{(v',v'),(v',w')\}),([b^\circ_s,0],[y,z-y]))$, where $b^\circ_s = F_v(G',\theta'_s)/c_{F,y,z}$. By Eq. (5.29), this gives us

$$F_{v'}(G^\circ,\theta^\circ_s) = F_v(G',\theta'_s). \tag{5.32}$$

In the sum of both graphs let us combine w' into w and also v' into v. Observe that as a result we will again obtain graph G^*, possibly with different weights. Thus, let us denote $(G^*,\theta^\dagger_s) = C^F_{v'\to v}(C^F_{w'\to w}(G'+G^\circ,\theta'_s+\theta^\circ_s))$. Denote $\theta^\dagger_s = (b^{s\dagger},\mu^{s\dagger})$. From Locality we have that $F_v(G'+G^\circ,\theta'_s+\theta^\circ_s) = F_v(G',\theta'_s)$ and from Locality and Eq. (5.32) we get that also $F_{v'}(G'+G^\circ,\theta'_s+\theta^\circ_r) = F_v(G',\theta'_s)$. Again, this has two implications: first, from Lemma 62 we get that

$$F_u(G^*,\theta^\dagger_s) = \begin{cases} F_u(G',\theta'_s), & \text{for every } u \in V \setminus \{v\}, \\ 2\cdot F_v(G',\theta'_s), & \text{for } u = v; \end{cases} \tag{5.33}$$

second, as before, the weights of outgoing edges of v and v' are divided by two, i.e, $\mu^{s\dagger}_{\Gamma^+_v(G)} = \mu_{\Gamma^+_v(G)}/2$ and $\mu^{s\dagger}(v,w) = z/2$, and the weights of the remaining edges are unchanged, $\mu^{s\dagger}_{-\Gamma^+_v(G)} = \mu_{-\Gamma^+_v(G)}$. Hence, $\mu^{s\dagger} = \mu^{r*}$, for every $r,s \in \mathbb{R}_{\geq 0}$. Therefore, the only possible difference between graphs (G^*,θ^*_r) and $(G^\dagger,\theta^\dagger_s)$, for any $r,s \in \mathbb{R}_{\geq 0}$, may be in the node weights. However, observe that for every $r,s \in \mathbb{R}_{\geq 0}$, we have $b^{r*}(u) = b(u) = b^{s\dagger}(u)$, for every $u \in V \setminus \{v\}$, and also $b^{r*}(w) = 0 = b^{s\dagger}(w)$. Thus, the only difference can be in the weight of node v. In what follows, we will show that in fact, for $s = 0$ and $r = 0$, it holds that $b^{0*}(v) = b^{0\dagger}(v)$. To this end, let us assume otherwise, i.e., that either (I) $b^{0*}(v) > b^{0\dagger}(v)$, or (II) $b^{0*}(v) < b^{0\dagger}(v)$.

(I) Assume that $b^{0*}(v) > b^{0\dagger}(v)$. The weight $b^{s\dagger}(v)$ is the sum of weights of v in (G',θ'_s) and v' in (G°,θ°_s). Thus, $b^{s\dagger}(v) = b'(v)+s+b^\circ_s = b'(v)+s+F_v(G',\theta'_s)/c_{F,y,z}$. Also, from Lemma 65 (a and b) we know that $F_v(G',\theta'_s) = F_v(G',\theta') + s\cdot F_v(G',(\mathbb{1}_v,\mu'))$. Both facts imply that $b^{s\dagger}(v)$ is a linear function of s. Hence, there exists $s > 0$ such that $b^{s\dagger}(v) = b^{0*}(v)$. Thus, for such an s it holds that $(G^*,\theta^*_0) = (G^*,\theta^\dagger_s)$. This implies that $F_v(G^*,\theta^*_0) = F_v(G^*,\theta^\dagger_s)$. Hence, from Eq. (5.31) and Eq. (5.33) we get that $F_v(G,\theta) = F_v(G',\theta'_s)$. Thus, looking again at node weights, we get that

$$b(v) + F_v(G,\theta) = b^{0*}(v) = b^{s\dagger}(v) = b(v) + p_F(F_v(G,\theta),y,z) + s + b^\circ_s. \tag{5.34}$$

From the first part of the proof we know that $F_{v'}(G^\circ,\theta^\circ_s) = p_F(F_{v'}(G^\circ,\theta^\circ_s),y,z) + b^\circ_s$. From Eq. (5.32) we have that $F_{v'}(G^\circ,\theta^\circ_s) = F_v(G',\theta'_s) = F_v(G,\theta)$. Therefore, we get $F_v(G,\theta) = p_F(F_v(G,\theta),y,z)+b^\circ_s$. Subtracting this from Eq. (5.34) yields $b(v) = b(v)+s$, which means that $s = 0$—a contradiction.

(II) Now, assume that $b^{0*}(v) < b^{0\dagger}(v)$. Observe that $b^{r*}(v) = b(v) + r + F_v(G,\theta_r)$. Also, from Lemma 65 (a and b) we get that $F_v(G,\theta_r) = F_v(G,\theta) + r\cdot F_v(G,(\mathbb{1}_v,\mu))$. Combining both facts, we get that $b^{r*}(v)$ is a linear function of r. Hence, there exists $r > 0$ such that $b^{0\dagger}(v) = b^{r*}(v)$. Thus, for such an r it holds that $(G^*,\theta^*_r) = (G^*,\theta^\dagger_0)$. This implies that $F_v(G^*,\theta^*_r) = F_v(G^*,\theta^\dagger_0)$ and, by Eq. (5.31) and Eq. (5.33), also that $F_v(G,\theta_r) = F_v(G',\theta'_0)$. Thus, looking again at the node weights, we get

$$b(v) + r + F_v(G,\theta_r) = b^{r*}(v) = b^{0\dagger}(v) = b(v) + p_F(F_v(G,\theta),y,z) + b^\circ_0. \tag{5.35}$$

From the first part of the proof we know that $F_{v'}(G^\circ,\theta^\circ_0) = p_F(F_{v'}(G^\circ,\theta^\circ_0),y,z) + b^\circ_0$. From Eq. (5.32) we have that $F_{v'}(G^\circ,\theta^\circ_0) = F_v(G',\theta'_0) = F_v(G,\theta_r)$. Therefore, we obtain $F_v(G,\theta_r) = p_F(F_v(G,\theta_r),y,z) + b^\circ_0$. Subtracting this from Eq. (5.35) yields

$$r = p_F(F_v(G,\theta),y,z) - p_F(F_v(G,\theta_r),y,z). \tag{5.36}$$

Recall that from Lemma 65 (a and b) we have $F_v(G,\theta_r) = F_v(G,\theta) + r \cdot F_v(G,(\mathbb{1}_v,\mu))$. Since $r > 0$, by Lemma 64, this means that $F_v(G,\theta_r) > F_v(G,\theta)$. Hence, by Lemma 67 or 68 (depending on satisfied axiom), also $p_F(F_v(G,\theta_r),y,z) > p_F(F_v(G,\theta),y,z)$. Thus, $p_F(F_v(G,\theta),y,z) - p_F(F_v(G,\theta_r),y,z) < 0$, which contradicts Eq. (5.36)

Therefore, indeed, it holds that $b^{0*}(v) = b^{0\dagger}(v)$. As a result, it also holds that $(G^*,\theta_0^*) = (G^*,\theta_0^\dagger)$. This implies that $F_u(G^*,\theta_0^*) = F_u(G^*,\theta_0^\dagger)$ for every $u \in V$. Thus, by Eq. (5.31) and Eq. (5.33), we obtain that $F_u(G,\theta) = F_u(G',\theta')$, for every $u \in V$, which yields the thesis by Eq. (5.30). □

Now we are ready to move to the final part of the proof. Having established that centrality measure F satisfies recursive equation, we will prove that it is equal to PageRank or Katz centrality, depending on the axiom satisfied. Let us first focus on PageRank. We begin by considering semi-out-regular graphs only.

Lemma 71. *If a centrality measure, F, defined on $\mathcal{G}$ satisfies Locality, Edge Deletion, Node Combination, Edge Multiplication, and Baseline then for every graph $G = (V,E)$ and weights $\theta = (b,\mu)$ such that (G,θ) is semi-out-regular, it holds that*

$$F_v(G,\theta) = PR_v^{a_F}(G,\theta), \quad \textit{for every } v \in V.$$

Proof. From Lemma 67 and Lemma 70 we get that for every semi-out-regular graph $(G,\theta) = ((V,E),(b,\mu)) \in \mathcal{G}$ and node $v \in V$, we have

$$F_v(G,\theta) = b(v) + \sum_{u \in P_v^1(G)} a_F \cdot \frac{\mu(u,v)}{\deg_u^+(G,\theta)} \cdot F_u(G,\theta).$$

Hence, centrality F satisfies PageRank recursive equation (Eq. (2.5)) with decay parameter a_F. The system of such equations has a unique solution, therefore $F_v(G,\theta) = PR_v^{a_F}(G,\theta)$, for every semi-out-regular graph (G,θ) and node $v \in V$. □

Now, let us relax the assumption that the graph has to be semi-out-regular.

Lemma 72. *If a centrality measure, F, defined on $\mathcal{G}$ satisfies Locality, Edge Deletion, Node Combination, Edge Multiplication, and Baseline then for every graph $G = (V,E)$ and weights $\theta = (b,\mu)$, it holds that*

$$F_v(G,\theta) = PR_v^{a_F}(G,\theta), \quad \textit{for every } v \in V.$$

Proof. Take an arbitrary $(G,\theta) = ((V,E),(b,\mu))$ and divide the weight of each edge by the out-degree of its start, i.e., let $\theta' = (b,\mu')$, where $\mu'(u,v) = \mu(u,v)/\deg_u^+(G,\theta)$, for every $(u,v) \in E$. From Edge Multiplication we have that $F_v(G',\theta) = F_v(G,\theta)$, for every $v \in V$. Observe that $(G,\theta)'$ is semi-out-regular. Thus, from Lemma 71 we know that $F_v(G,\theta') = PR_v^{a_F}(G,\theta')$, for every $v \in V$. Since PageRank also satisfies Edge Multiplication (Lemma 59), we get that $F_v(G,\theta) = PR_v^{a_F}(G,\theta)$. □

Now, let us prove analogous results for Katz centrality. As before, we begin with the assumption that the graph is semi-out-regular.

Lemma 73. *If a centrality measure, F, defined on $\mathcal{G}^{K(a)}$ satisfies Locality, Edge Deletion, Node Combination, Edge Compensation, and Baseline then for every graph $G = (V,E)$ and weight $\theta = (b,\mu)$ such that graph (G,θ) is semi-out-regular and $(G,\theta) \in \mathcal{G}^{K(a)}$, it holds that*

$$F_v(G,\theta) = K_v^{a_F}(G,\theta), \quad \textit{for every } v \in V.$$

Proof. From Lemma 68 and Lemma 70 we obtain that for every semi-out-regular graph $(G,\theta) = ((V,E),(b,\mu)) \in \mathcal{G}^{K(a)}$ and every $v \in V$, we have

$$F_v(G,\theta) = b(v) + \sum_{u \in P^1_v(G)} a_F \cdot \mu(u,v) \cdot F_u(G,\theta).$$

This means that centrality F satisfies Katz recursive equation (Eq. (2.3)) with decay parameter a_F. Since the system of Katz recursive equations has a unique solution, we obtain that $F_v(G,\theta) = K^{a_F}_v(G,\theta)$, for every semi-out-regular graph (G,θ) and node $v \in V$. □

Now, once again, let us relax the assumption that the graph has to be semi-out-regular.

Lemma 74. *If a centrality measure, F, defined on $\mathcal{G}^{K(a)}$ satisfies Locality, Edge Deletion, Node Combination, Edge Compensation, and Baseline, then for every graph $G = (V,E)$ and weights $\theta = (b,\mu)$ such that $(G,\theta) \in \mathcal{G}^{K(a)}$, it holds that*

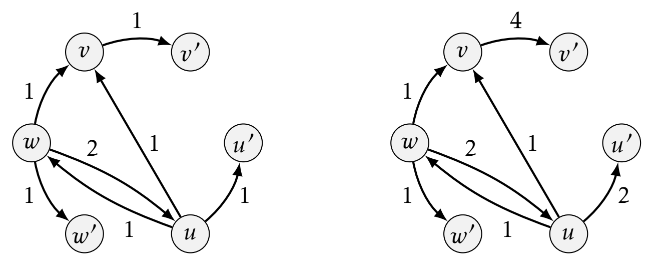
$$F_v(G,\theta) = K^{a_F}_v(G,\theta), \quad \textit{for every } v \in V.$$

Proof. We will say that node $v \in V$ that does not have any outgoing edges and has exactly one incoming edge is a *leaf*. Let $V^L(G) = \{v \in V : \Gamma^+_v(G) = \emptyset \wedge |\Gamma^-_v(G)| = 1\}$ be the set of all leafs in graph G. The *parent* of leaf v is a node, $p(v)$, that has an outgoing edge to v, i.e., $(p(v),v) \in E$. We will denote the set of all parents of leafs by $V^P(G) = \{p(v) : v \in V^L(G)\}$. For every parent, $u \in V^P(G)$, let us arbitrarily choose one of its leafs and denote it by $l(u)$ and also let us denote the set of these chosen leafs by $\hat{V}^L(G) = \{l(u) : u \in V^P(G)\}$. Moreover, let us denote the set of all sinks in G as $V^S(G)$. Note that $V^L(G) \subseteq V^S(G)$.

Our goal is to transform an arbitrary graph into a semi-out-regular one. Intuitively, the out-degree of nodes that are parents of leafs can be arbitrarily increased by multiplying the weight of the incoming edge of the corresponding leaf. By Edge Compensation, such an operation changes only the centrality of the leaf, but does not change the centrality of other nodes. Hence, if all the nodes are either sinks or parents of leafs, we are able to easily transform the graph into a semi-out-regular one. Therefore, the main obstacle are the nodes that are neither sinks nor parents of leafs. We will call such nodes *ordinary* and denote the set of all ordinary nodes in G by $V^O(G) = V \setminus (V^S(G) \cup V^P(G))$. We will prove the thesis by induction on their number, i.e., $|V^O(G)|$.

If $|V^O(G)| = 0$, then each node is either a sink or a parent of a leaf, i.e., we have $V^S(G) \cup V^P(G) = V$. If $V^P(G) = \emptyset$, then the graph consists only of isolated nodes, which means it is semi-out-regular. Thus, the thesis follows from Lemma 73. Assume otherwise, i.e., that $V^P(G) \neq \emptyset$. Let us denote the maximal out-degree of all nodes by $x = \max_{v \in V} \deg^+_v(G,\theta)$ (since $V^P(G) \neq \emptyset$, we know that $x > 0$). In order to transform graph (G,θ) into a semi-out-regular graph, we will scale the weights of edges from parents to leafs in such a way that all parents have out-degree x. More in detail, for every $v \in V^P(G)$ such that $\deg^+_v(G,\theta) < x$, let us consider one of its leafs, $l(v)$, and multiply the weights of $l(v)$ and edge $(v,l(v))$ by a constant such that the new weight of $(v,l(v))$ is equal to $x - (\deg^+_v(G,\theta) - \mu(v,l(v)))$. In this way, in the new graph, (G,θ'), node v will have out-degree equal to x. See Fig. 5.14 for an illustration. Formally, let $\theta' = (b',\mu')$ where

$$b'(v) = b(v) \cdot \frac{x - (\deg^+_{p(v)}(G,\theta) - \mu(p(v),v))}{\mu(p(v),v)}, \quad \text{for every } v \in \hat{V}^L(G)$$



(G,θ) (G,θ')

Figure 5.14: An illustration to the basis of the induction in the proof of Lemma 74. In an example graph, (G,θ), all nodes are either leafs, like u', v', and w', or parents of leafs, like u, v, and w. Multiplying the weights of edges between leafs and their parents we can obtain a semi-out-regular graph, (G,θ').

and $b'(v) = b(v)$, for every $v \in V \setminus \hat{V}^L(G)$, while for every edge $(u,v) \in E$, we set $\mu'(u,v) = x - (\deg^+_u(G,\theta) - \mu(u,v))$, if $v = l(u)$, and $\mu'(u,v) = \mu(u,v)$, otherwise. From Edge Compensation we know that

$$F_v(G',\theta) = \begin{cases} F_v(G,\theta) \cdot \left(1 + \frac{x - \deg^+_{p(v)}(G,\theta)}{\mu(p(v),v)}\right), & \text{for every } v \in \hat{V}^L(G), \\ F_v(G,\theta), & \text{for every } v \in V \setminus \hat{V}^L(G). \end{cases} \tag{5.37}$$

Observe that in graph (G,θ') every node is either a sink and does not have any outgoing edges, or its out-degree is equal to x. Thus, graph (G,θ') is semi-out-regular. Hence, from Lemma 73 we know that $F_v(G,\theta') = K^{a_F}_v(G,\theta')$. Since Katz centrality satisfies Edge Compensation (Lemma 60), we get the thesis from Eq. (5.37).

Let us move to the case in which $|V^O(G)| > 0$. Then, let us take node $v \in V^O(G)$ such that the number of its successors that are ordinary and different from v, i.e., $|S_v(G) \cap V^O(G) \setminus \{v\}|$, is minimal. Observe that for every successor of v, i.e., node $u \in S_v(G)$, we have that $S_u(G) \subseteq S_v(G)$. We can prove that this implies that if u is ordinary, then $S_u(G) = S_v(G)$. For assume otherwise, i.e., that u is ordinary and $S_u(G) \subset S_v(G)$. This means that $|S_u(G) \cap V^O(G) \setminus \{u\}| < |S_v(G) \cap V^O(G) \setminus \{v\}|$, because on the right hand side we count each node that we count on the left hand side and also node u. However, we assumed that for v the number of its successors that are ordinary and different from v is minimal—a contradiction. As a result, we obtain two cases: the first one (I) in which $v \notin S_v(G)$ and there are no ordinary successors of v, i.e., $S_v(G) \cap V^O = \emptyset$; and the second one (II) in which $v \in S_v(G)$ and all ordinary successors of v belong to the same strongly connected component as v, i.e., for every $u, w \in S_v(G) \cap V^O$ we have $u \in S_w(G) \cap P_w(G)$.

(I) Let us begin with the case in which $v \notin S_v(G)$ and all successors of v are either sinks or parents of some leafs, i.e., $S_v(G) \subseteq V^S(G) \cup V^P(G)$. In what follows, we prove that $F_u(G,\theta) = K^{a_F}_u(G,\theta)$, for every $u \in V$, by transforming (G,θ) to a graph with smaller $|V^O(G)|$ and using the inductive assumption. To this end, we follow two steps. First, we increase the out-degrees of v and some successors of v so that all successors of v are sinks or have equal out-degree to v (if v does not have successors that are not sinks we omit this step). Second, using Lemma 61 we add a leaf to node v to decrease $|V^O(G)|$. See Fig. 5.15 for an illustration.

Denote the maximal out-degree of a successor of v by $x = \max_{u \in S_v(G)} \deg^+_v(G)$. If $x > 0$, then in the same way as in case of $|V^O(G)| = 0$, we increase the weight of each edge from a parent to a leaf in $S_v(G)$ so that all successors of v that are not

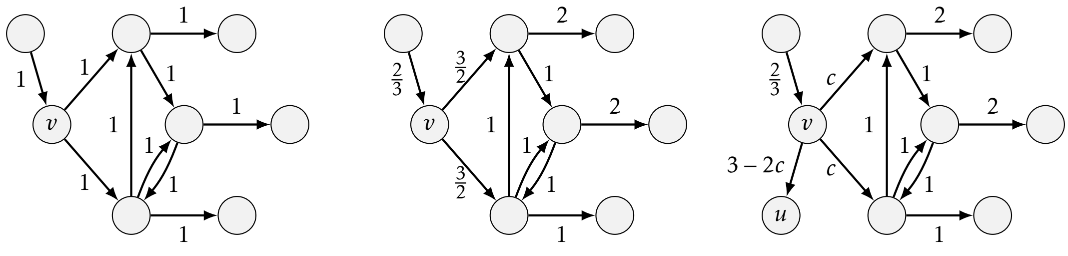


Figure 5.15: An illustration to the case (I) in the proof of Lemma 74. In an example graph, (G,θ), all of the successors of node v are either leafs or parents of leafs. Graph (G,θ'') is obtained from (G,θ) by multiplying the weights of edges from parents to leafs in such a way that the out-degree of each parent is equal to $x = 3$. Also, the weights of the incoming and outgoing edges of v are scaled. In graph (G^*,θ^*) a leaf is added to node v.

sinks have out-degree x. Formally, let $\theta' = (b',\mu')$, where

$$b'(u) = b(u) \cdot \frac{x - (\deg^+_{p(u)}(G,\theta) - \mu(p(u),u))}{\mu(p(u),u))}, \quad \text{if } u \in \hat{V}^L \cap S_v(G),$$

and $b'(u) = b(u)$, otherwise. As for edge weights, for every edge $(u,w) \in E$, we set $\mu'(u,w) = x - \deg^+_u(G,\theta) + \mu(u,w)$, if $w \in \hat{V}^L \cap S_v(G)$, and $\mu'(u,v) = \mu(u,v)$, otherwise. From Edge Compensation we know that

$$F_u(G,\theta') = \begin{cases} F_u(G,\theta) \cdot \left(1 + \frac{x - \deg^+_{p(u)}(G,\theta)}{\mu(p(u),u)}\right), & \text{for every } u \in \hat{V}^L \cap S_v(G), \\ F_u(G,\theta), & \text{for every } u \in V \setminus (\hat{V}^L \cap S_v(G)). \end{cases} \quad (5.38)$$

Next, let us multiply the weights of the outgoing edges of v by $x/\deg^+_v(G,\theta')$ and divide its weight and the weights of its incoming edges by $x/\deg^+_v(G,\theta')$. Formally, let $\theta'' = (b'',\mu'')$ where $b''(v) = b'(v) \cdot \deg^+_v(G,\theta')/x$ and $b''(u) = b'(u)$, for every $u \in V \setminus \{v\}$, while $\mu''(e) = \mu'(e) \cdot \deg^+_v(G,\theta')/x$, if $e \in \Gamma^-_v(G)$, $\mu''(e) = \mu'(e) \cdot x/\deg^+_v(G,\theta')$, if $e \in \Gamma^+_v(G)$, and $\mu''(e) = \mu'(e)$, otherwise. Again, from Edge Compensation we get that

$$F_u(G,\theta'') = \begin{cases} F_u(G,\theta'), & \text{for every } u \in V \setminus \{v\}, \\ F_v(G,\theta') \cdot \frac{\deg^+_v(G,\theta')}{x}, & \text{for } u = v. \end{cases} \quad (5.39)$$

Observe that incoming edges of v does not come from successors of v, because $v \notin S_v(G)$.

So far, we considered a case in which $x > 0$. If $x = 0$, this means that all of the successors of v are sinks. In such a case we can simply set $\theta'' = \theta$.

Now, observe that in (G,θ'') all successors of v that are not sinks have the same out-degree, which is also equal the out-degree of v, i.e., $\deg^+_u(G,\theta'') = x = \deg^+_v(G,\theta'')$, for every $u \in S_v(G) \cap V^P$. We will use this fact to add a leaf to node v using Lemma 61.

To this end, let us consider nodes $u',v' \notin V$ and add to graph G'' a simple graph that consists of nodes u' and v' connected by an edge. Formally, consider graph $(G + G^{\uparrow}, \theta'' + \theta^{\uparrow})$, where $(G^{\uparrow},\theta^{\uparrow}) = ((\{u',v'\},\{(v',u')\}),([1,1],[x]))$. Now, let us combine node v' into v, i.e., let $(G^*,\theta^*) = C^F_{v' \to v}(G + G^{\uparrow}, \theta'' + \theta^{\uparrow})$. From Lemma 63 we get that $F_{v'}(G^{\uparrow},\theta^{\uparrow}) = 1$. Hence, from Lemma 61 and Locality we get that

$$F_u(G^*,\theta^*) = \begin{cases} F_u(G,\theta''), & \text{for every } u \in V \setminus \{v\}, \\ F_v(G,\theta'') + 1, & \text{for } u = v. \end{cases} \quad (5.40)$$

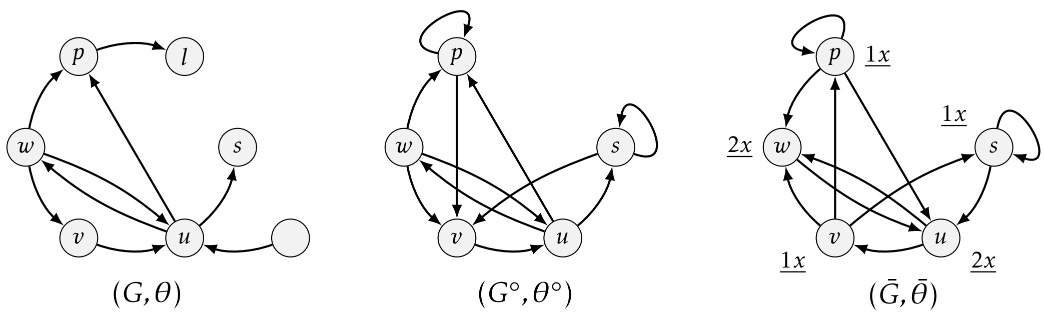


Figure 5.16: An illustration to the first part of the case (II) in the proof of Lemma 74. All edges have weights equal to 1. In an example graph, (G,θ), nodes v, u, and w constitute a strongly connected component, U. Nodes l,p, and s are the successors of v that do not belong to this component and all of them are either sinks or parents of leafs. Graph (G°,θ°) is a strongly connected graph constructed from the nodes in U and their direct successors. The maximal out-degree of a successor of v that is not in U, i.e., x, is equal to 1. Thus, loops around nodes p and s have weight 1. Graph $(\bar{G},\bar{\theta})$ is an opposite graph to (G°,θ°). Eigenvector centrality of every node in this graph is shown next to the node. Note that $\lambda = 2$.

On the other hand, in graph (G^*,θ^*) node v is not ordinary anymore and no ordinary node was added. Thus, by the inductive assumption, $F_u(G^*,\theta^*) = K_u^{a_F}(G^*,\theta^*)$, for every $u \in V$. Since Katz centrality also satisfies our axioms (Lemma 60), from Eqs. (5.38)–(5.40) we obtain that $F_u(G,\theta) = K_u^{a_F}(G,\theta)$, for every $u \in V$.

(II) Now, let us move to the second case in which $v \in S_v(G)$ and all ordinary successors of v belong to the same strongly connected component, i.e., for every $u,w \in S_v(G) \cap V^O(G)$ we have $u \in S_w(G) \cap P_w(G)$. Let us denote the set of all the nodes in this strongly connected components by $U = S_v(G) \cap P_v(G)$. Also, let us denote the set of their outgoing edges by $E^U = \{(u,w) \in E : u \in U\}$ and by U^+ the ends of edges from E^U that are not in U, i.e., let $U^+ = \{w : (u,w) \in E^U\} \setminus U$. Finally, let us denote the maximal out-degree of a successor of v that is not in U by $x = \max_{u \in S_v(G)\setminus U} \deg_u^+(G)$ (if $S_v(G) \setminus U = \emptyset$, let $x = 0$). In what follows, we prove that $F_u(G,\theta) = K_u^{a_F}(G,\theta)$, for every $u \in V$, by transforming (G,θ) to a graph with smaller $|V^O(G)|$ and using the inductive assumption. To this end, we follow three steps. First, we modify the weights of edges E^U to make out-degrees of nodes in U equal λ. Second, we increase the out-degrees of some successors of v so that all successors of v are sinks or have their out-degrees equal to λ (if v does not have successors that are not sinks and are not in U we omit this step). Finally, using Lemma 61 we add a leaf to node v to decrease $|V^O(G)|$.

First, let us consider an auxiliary graph in which all the nodes except for those in U and U^+ are removed. Formally, let $G^* = (U \cup U^+, E^U)$ and $\theta^* = (b_{U\cup U^+}, \mu_{E^U})$. Observe that since all outgoing edges of U remains, nodes in U in graph G^* still constitute a strongly connected component. In order to make the whole graph strongly connected, let us add outgoing edges to the nodes in U^+. More in detail, for each $s \in U^+$ let us add edge (s,s) with weight $\max(x,1)$ and edge (s,v) with weight 1. See Fig. 5.16 for an illustration. Formally, let $E^+ = \{(s,s),(s,v) : s \in U^+\}$ and let $G^\circ = (V^\circ, E^\circ)$, where $V^\circ = U \cup U^+$, $E^\circ = E^U \cup E^+$, and $\theta^\circ = (b^\circ, \mu^\circ)$, where $b^\circ = b_{U\cup U^+}$, $\mu^\circ(s,s) = \max(x,1)$, and $\mu^\circ(s,v) = 1$, for every $s \in U^+$, and $\mu^\circ_{E^U} = \mu_{E^U}$. Observe that G° is indeed strongly connected.

Since G° is strongly connected, we can make it out-regular using the same tech-

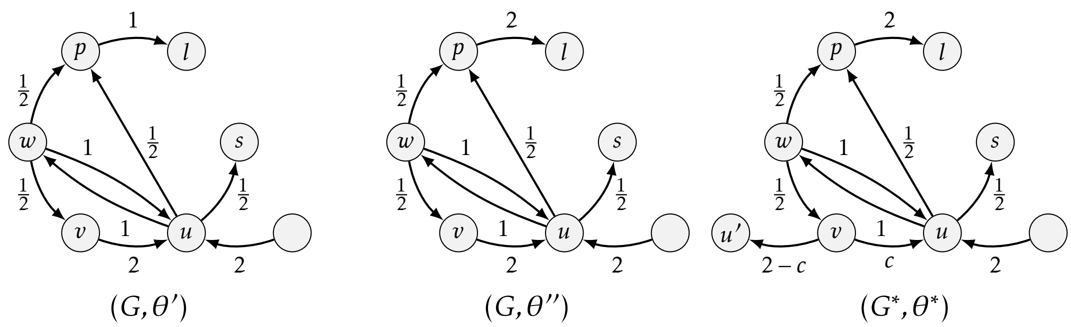


Figure 5.17: An illustration to the second part of the case (II) in the proof of Lemma 74. Graph (G,θ') is obtained from graph (G,θ) from Fig. 5.16 by taking eigenvector centrality of each node in graph $(\bar{G},\bar{\theta})$ and dividing by it the weight of its outgoing edges and multiplying by it its weight and the weights of its incoming edges. Graph (G,θ'') is obtained from (G,θ') by multiplying the weights of edges from parents to leafs so that the out-degrees of parents are equal to 2. In graph (G^*,θ^*) a leaf is added to node v.

nique that we used in Lemma 58. Formally, let us consider the opposite graph to (G°,θ°), i.e., graph $(\bar{G},\bar{\theta}) = ((V^\circ,\bar{E}),(b^\circ,\bar{\mu}))$ such that $\bar{E} = \{(u,w) : (w,u) \in E^\circ\}$ and $\bar{\mu}(u,w) = \mu(w,u)$, for every $(w,u) \in E^\circ$. Now, in graph $(\bar{G},\bar{\theta})$ let us multiply the weights of the outgoing edges of node $u \in V^\circ$ by $EV_u(\bar{G},\bar{\theta})$ and divide the weights of its incoming edges as well as its own weight by $EV_u(\bar{G},\bar{\theta})$. Because eigenvector centrality satisfies Edge Compensation, we know that in this way eigenvector centralities of the remaining nodes in V° do not change and eigenvector centrality of u becomes 1. If we proceed with this operation for all nodes in V° we obtain graph $(\bar{G}',\bar{\theta}')$ in which all nodes have eigenvector centralities equal to 1. Formally, let $(\bar{G}',\bar{\theta}') = ((V^\circ,\bar{E}),(b',\bar{\mu}'))$ where $b'(u) = b^\circ(u)/EV_u(\bar{G},\bar{\theta})$, for every $u \in V^\circ$, and $\bar{\mu}'(u,w) = \bar{\mu}(u,w) \cdot EV_u(\bar{G},\bar{\theta})/EV_w(\bar{G},\bar{\theta})$. Observe that if all the nodes in graph $(\bar{G}',\bar{\theta}')$ have equal eigenvector centrality, then, from eigenvector centrality recursive equation (Eq. (2.2)), we get that the in-degrees of all the nodes are equal, i.e., there exist λ such that $\deg^-_u(\bar{G}',\bar{\theta}') = \lambda$, for every $u \in V$. Moreover, observe that for every $s \in U^+$ we have that $\bar{\mu}'(s,s) = \bar{\mu}(s,s) \cdot EV_u(\bar{G},\bar{\theta})/EV_u(\bar{G},\bar{\theta}) = \mu^\circ(s,s) = \max(x,1)$. Thus, $\lambda > \max(x,1)$.

Now, let us perform this operation on the original graph G, instead of G°, to obtain equal out-degrees of nodes in U (see Fig. 5.17). At the same time, we want to make sure that we will not increase out-degrees of the nodes in U^+ too much (as we cannot decrease their out-degree using leafs, only increase). To this end, let us take an arbitrary constant, $y \in \mathbb{R}_{>0}$, by which we will multiply the incoming edges and divide outgoing edges of all the nodes in V°. Formally, let us define weights $\theta' = (b',\mu')$ in which $b'(u) = b(u) \cdot EV_u(\bar{G},\bar{\theta}) \cdot y$, for every $u \in V^\circ$, and $b'(u) = b(u)$, for every $u \in V \setminus V^\circ$, and

$$\mu'(u,w) = \begin{cases} \mu(u,w) \cdot EV_w(\bar{G},\bar{\theta})/EV_u(\bar{G},\bar{\theta}), & \text{for every } u,w \in V^\circ, \\ \mu(u,w)/(EV_u(\bar{G},\bar{\theta}) \cdot y), & \text{for every } u \in V^\circ, w \notin V^\circ, \\ \mu(u,w) \cdot EV_w(\bar{G},\bar{\theta}) \cdot y, & \text{for every } u \notin V^\circ, w \in V^\circ, \\ \mu(u,w), & \text{for every } u,w \notin V^\circ. \end{cases}$$

Observe that for every $(u,w) \in E^U$, we have $\mu'(u,w) = \bar{\mu}'(w,u)$. Thus, indeed, $\deg^+_u(G',\theta') = \deg^-_u(\bar{G}',\bar{\theta}') = \lambda > x$, for every $u \in U$. Furthermore, from Edge Compensation we have that

$$F_u(G,\theta') = \begin{cases} F_u(G,\theta) \cdot EV_u(\bar{G},\bar{\theta}) \cdot y, & \text{for every } u \in V^\circ, \\ F_u(G,\theta), & \text{for every } u \in V \setminus V^\circ. \end{cases} \quad (5.41)$$

Observe that since for every node $u \in U^+$, the outgoing edges of u go to nodes outside of V°, we have that $\deg_u^+(G,\theta') = \deg_u^+(G,\theta)/(EV_u(\bar{G},\bar{\theta}) \cdot y)$. Thus, let us take such y that the maximal out-degree of node in U^+ is equal to $\max(x,1)$, i.e., let $y = \max(x,1) \cdot \max_{u \in U^+}(\deg_u^+(G,\theta)/EV_u(\bar{G},\bar{\theta}))$. Then, all the successors of v in graph (G,θ') have out-degrees equal to at most $\max(x,1)$. Moreover, those that are in a strongly connected component, i.e., nodes in U, have the out-degrees equal to $\lambda > \max(x,1)$, and those that are not, i.e., nodes in $S_v(G) \setminus U$, are either sinks or parents of leafs and have the out-degrees equal to at most x. If $x > 0$, i.e., there are nodes in $S_v(G) \setminus U$ that are not sinks, then let us increase the out-degrees of parents of leafs in $S_v(G) \setminus U$ to λ by changing the weights of their edges going to a leaf. Formally, let $\theta'' = (b'', \mu'')$, where

$$b''(u) = b'(u) \cdot \frac{\lambda - (\deg_{p(u)}^+(G,\theta') - \mu'(p(u),u))}{\mu'(p(u),u)}, \quad \text{if } u \in \hat{V}^L \cap S_v(G),$$

and $b''(u) = b'(u)$, otherwise. For every $(u,w) \in E$, let $\mu''(u,w) = \lambda - \deg_u^+(G,\theta') + \mu'(u,w)$, if $w \in \hat{V}^L \cap S_v(G)$, and $\mu''(u,v) = \mu'(u,v)$, otherwise. From Edge Compensation we know that

$$F_u(G,\theta'') = \begin{cases} F_u(G,\theta') \cdot \left(1 + \frac{\lambda - \deg_{p(u)}^+(G,\theta')}{\mu'(p(u),u)}\right), & \text{for every } u \in \hat{V}^L \cap S_v(G), \\ F_u(G,\theta'), & \text{for every } u \in V \setminus (\hat{V}^L \cap S_v(G)). \end{cases} \quad (5.42)$$

If $x = 0$, i.e., all successors of v that are not in U are sinks, then let simply $\theta'' = \theta'$. Observe that in (G,θ'') all successors of v are either sinks or have out-degree λ. Hence, we will use Lemma 61 to add a leaf to v in a similar way to how we did it in case (I). Let us consider nodes $u', v' \notin V$ and add a simple two-node graph to (G,θ''), i.e., consider $(G + G^\uparrow, \theta'' + \theta^\uparrow)$, where $(G^\uparrow, \theta^\uparrow) = ((\{u',v'\}, \{(v',u')\}), ([1,1],[x]))$. Now, let us combine node v' into v, i.e., let $(G^*, \theta^*) = C^F_{v' \to v}(G + G^\uparrow, \theta + \theta^\uparrow)$. From Lemma 63 we get that $F_{v'}(G^\uparrow, \theta^\uparrow) = 1$. Hence, from Lemma 61 and Locality we get that

$$F_u(G^*,\theta^*) = \begin{cases} F_u(G,\theta''), & \text{for every } u \in V \setminus \{v\}, \\ F_v(G,\theta'') + 1, & \text{for } u = v. \end{cases} \quad (5.43)$$

On the other hand, in graph G^* node v is not ordinary anymore and no ordinary node was added. Thus, by the inductive assumption, $F_u(G^*,\theta^*) = K_u^{a_F}(G^*,\theta^*)$, for every $u \in V$. Since Katz centrality satisfies our axioms (Lemma 60), from Eqs. (5.41)–(5.43) we get that $F_u(G,\theta) = K_u^{a_F}(G,\theta)$, for every $u \in V$. □

Bibliography

[1] Altman, Alon and Tennenholtz, Moshe. "Ranking systems: the PageRank axioms". In: *Proceedings of the 6th ACM Conference on Electronic Commerce (ACM-EC)*. 2005, pp. 1–8.

[2] Arrow, Kenneth J. "A difficulty in the concept of social welfare". In: *Journal of Political Economy* 58.4 (1950), pp. 328–346.

[3] Ashtiani, Minoo, Salehzadeh-Yazdi, Ali, Razaghi-Moghadam, Zahra, Hennig, Holger, Wolkenhauer, Olaf, Mirzaie, Mehdi, and Jafari, Mohieddin. "A systematic survey of centrality measures for protein-protein interaction networks". In: *BMC Systems Biology* 12.1 (2018), pp. 1–17.

[4] Avrachenkov, Konstantin and Litvak, Nelly. "The effect of new links on Google PageRank". In: *Stochastic Models* 22.2 (2006), pp. 319–331.

[5] Baeza-Yates, Ricardo, Boldi, Paolo, and Castillo, Carlos. "Generalizing pagerank: Damping functions for link-based ranking algorithms". In: *Proceedings of the 29th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR)*. 2006, pp. 308–315.

[6] Batool, Komal and Niazi, Muaz A. "Towards a methodology for validation of centrality measures in complex networks". In: *PloS one* 9.4 (2014), e90283.

[7] Battiston, Stefano, Puliga, Michelangelo, Kaushik, Rahul, Tasca, Paolo, and Caldarelli, Guido. "Debtrank: Too central to fail? financial networks, the fed and systemic risk". In: *Scientific Reports* 2.1 (2012), pp. 1–6.

[8] Bavelas, Alex. "Communication patterns in task-oriented groups". In: *The Journal of the Acoustical Society of America* 22.6 (1950), pp. 725–730.

[9] Bianchini, Monica, Gori, Marco, and Scarselli, Franco. "Inside PageRank". In: *ACM Transactions on Internet Technology (TOIT)* 5.1 (2005), pp. 92–128.

[10] Bloch, Francis, Jackson, Matthew O., and Tebaldi, Pietro. "Centrality Measures in Networks". In: *arXiv preprint arXiv:1608.05845* (2016).

[11] Boldi, Paolo, Luongo, Alessandro, and Vigna, Sebastiano. "Rank monotonicity in centrality measures". In: *Network Science* 5.4 (2017), pp. 529–550.

[12] Boldi, Paolo, Santini, Massimo, and Vigna, Sebastiano. "PageRank as a function of the damping factor". In: *Proceedings of the 14th International Conference on World Wide Web (WWW)*. 2005, pp. 557–566.

[13] Boldi, Paolo and Vigna, Sebastiano. "Axioms for centrality". In: *Internet Mathematics* 10.3-4 (2014), pp. 222–262.

[14] Bolland, John M. "Sorting out centrality: An analysis of the performance of four centrality models in real and simulated networks". In: *Social Networks* 10.3 (1988), pp. 233–253.

[15] Bollen, Johan, Rodriquez, Marko A., and Van de Sompel, Herbert. "Journal status". In: *Scientometrics* 69.3 (2006), pp. 669–687.

[16] Bonacich, Phillip. "Factoring and weighting approaches to status scores and clique identification". In: *Journal of Mathematical Sociology* 2.1 (1972), pp. 113–120.

[17] Borgatti, Stephen P. "Centrality and network flow". In: *Social Networks* 27.1 (2005), pp. 55–71.

[18] Borgatti, Stephen P. and Everett, Martin G. "A graph-theoretic perspective on centrality". In: *Social Networks* 28(4) (2006), pp. 466–484.

[19] Brandes, Ulrik and Fleischer, Daniel. "Centrality measures based on current flow". In: *Proceedings of the 22nd Annual Symposium on Theoretical Aspects of Computer Science (STACS)*. Springer, 2005, pp. 533–544.

[20] Bringmann, Laura F., Elmer, Timon, Epskamp, Sacha, Krause, Robert W., Schoch, David, Wichers, Marieke, Wigman, Johanna TW., and Snippe, Evelien. "What do centrality measures measure in psychological networks?" In: *Journal of Abnormal Psychology* 128.8 (2019), p. 892.

[21] Brink, René van den. "An axiomatization of the Shapley value using a fairness property". In: *International Journal of Game Theory* 30.3 (2002), pp. 309–319.

[22] Broder, Andrei, Kumar, Ravi, Maghoul, Farzin, Raghavan, Prabhakar, Rajagopalan, Sridhar, Stata, Raymie, Tomkins, Andrew, and Wiener, Janet. "Graph structure in the web". In: *Computer networks* 33.1-6 (2000), pp. 309–320.

[23] Cauchy, Augustin LB. *Cours d'analyse de l'École Royale Polytechnique: Analyse algébrique. Pte. 1*. Imprimerie royale, 1821.

[24] Csardi, Gabor and Nepusz, Tamas. "The igraph software package for complex network research". In: *InterJournal* Complex Systems (2006), p. 1695.

[25] Dequiedt, Vianney and Zenou, Yves. "Local and consistent centrality measures in networks". CEPR Discussion Paper No. DP10031. 2014.

[26] Dubey, Pradeep. "On the uniqueness of the Shapley value". In: *International Journal of Game Theory* 4.3 (1975), pp. 131–139.

[27] Durrett, Rick. *Probability: theory and examples*. Vol. 49. Cambridge university press, 2019.

[28] Everett, Martin G. and Borgatti, Stephen P. "Induced, endogenous and exogenous centrality". In: *Social Networks* 32.4 (2010), pp. 339–344.

[29] Freeman, Linton C. "A set of measures of centrality based on betweenness". In: *Sociometry* 40.1 (1977), pp. 35–41.

[30] Freeman, Linton C. "Centrality in social networks: Conceptual clarification". In: *Social Networks* 1.3 (1979), pp. 215–239.

[31] Garg, Manuj. "Axiomatic foundations of centrality in networks". In: *unpublished* (2009).

[32] Gest, Scott D., Graham-Bermann, Sandra A., and Hartup, Willard W. "Peer experience: Common and unique features of number of friendships, social network centrality, and sociometric status". In: *Social Development* 10.1 (2001), pp. 23–40.

[33] Gneiser, Martin, Heidemann, Julia, Klier, Mathias, Landherr, Andrea, and Probst, Florian. "Valuation of online social networks taking into account users' interconnectedness". In: *Information Systems and e-Business Management* 10.1 (2012), pp. 61–84.

[34] Guimera, Roger, Mossa, Stefano, Turtschi, Adrian, and Amaral, LA. Nunes. "The worldwide air transportation network: Anomalous centrality, community structure, and cities' global roles". In: *Proceedings of the National Academy of Sciences* 102.22 (2005), pp. 7794–7799.

[35] Gyongyi, Zoltan, Garcia-Molina, Hector, and Pedersen, Jan. "Combating web spam with trustrank". In: *Proceedings of the 30th International Conference on Very Large Data Bases (VLDB)*. 2004, pp. 576–587.

[36] Hallquist, Michael N., Wright, Aidan GC., and Molenaar, Peter CM. "Problems with centrality measures in psychopathology symptom networks: Why network psychometrics cannot escape psychometric theory". In: *Multivariate Behavioral Research* (2019), pp. 1–25.

[37] Haveliwala, Taher H. "Topic-sensitive PageRank". In: *Proceedings of the 11th International Conference on World Wide Web (WWW)*. ACM, 2002, pp. 517–526.

[38] Heath, Thomas L. *The thirteen books of Euclid's Elements*. Courier Corporation, 1956.

[39] Hossain, Liaquat, Chung, Kon SK., and Murshed, Shahriar TH. "Exploring temporal communication through social networks". In: *IFIP Conference on Human-Computer Interaction*. Springer. 2007, pp. 19–30.

[40] Huberman, Bernardo A., Pirolli, Peter LT., Pitkow, James E., and Lukose, Rajan M. "Strong regularities in world wide web surfing". In: *Science* 280.5360 (1998), pp. 95–97.

[41] Irwin, Michael D. and Hughes, Holly L. "Centrality and the structure of urban interaction: measures, concepts, and applications". In: *Social Forces* 71.1 (1992), pp. 17–51.

[42] Iván, Gábor and Grolmusz, Vince. "When the Web meets the cell: using personalized PageRank for analyzing protein interaction networks". In: *Bioinformatics* 27.3 (2010), pp. 405–407.

[43] Jackson, Matthew O. *Social and Economic Networks*. Vol. 3. Princeton university press, 2008.

[44] Jalili, Mahdi, Salehzadeh-Yazdi, Ali, Asgari, Yazdan, Arab, Seyed S., Yaghmaie, Marjan, Ghavamzadeh, Ardeshir, and Alimoghaddam, Kamran. "CentiServer: a comprehensive resource, web-based application and R package for centrality analysis". In: *PloS one* 10.11 (2015), e0143111.

[45] Jech, Thomas. *Set theory*. Springer Science & Business Media, 2013.

[46] Katz, Leo. "A new status index derived from sociometric analysis". In: *Psychometrika* 18.1 (1953), pp. 39–43.

[47] Kiss, Christine and Bichler, Martin. "Identification of influencers—measuring influence in customer networks". In: *Decision Support Systems* 46.1 (2008), pp. 233–253.

[48] Kitti, Mitri. "Axioms for centrality scoring with principal eigenvectors". In: *Social Choice and Welfare* 46.3 (2016), pp. 639–653.

[49] Kleinberg, Jon M. “Authoritative sources in a hyperlinked environment”. In: *Journal of the ACM (JACM)* 46.5 (1999), pp. 604–632.

[50] Knuth, Donald E. *The Stanford GraphBase: a platform for combinatorial computing*. ACM Press New York, 1993.

[51] Kolmogorov, Andrey Nikolaevich and Bharucha-Reid, Albert T. *Foundations of the theory of probability: Second English Edition*. Courier Dover Publications, 2018.

[52] Koschützki, Dirk, Lehmann, Katharina A., Peeters, Leon, Richter, Stefan, Tenfelde-Podehl, Dagmar, and Zlotowski, Oliver. “Centrality indices”. In: *Network Analysis*. Vol. 3418. Lecture Notes in Computer Science. Springer, 2005, pp. 16–61.

[53] Koschützki, Dirk, Lehmann, Katharina A., Tenfelde-Podehl, Dagmar, and Zlotowski, Oliver. “Advanced centrality concepts”. In: *Network Analysis*. Vol. 3418. Lecture Notes in Computer Science. Springer, 2005, pp. 83–111.

[54] Landherr, Andrea, Friedl, Bettina, and Heidemann, Julia. “A critical review of centrality measures in social networks”. In: *Business & Information Systems Engineering* 2.6 (2010), pp. 371–385.

[55] Lerman, Kristina and Ghosh, Rumi. “Information contagion: An empirical study of the spread of news on digg and twitter social networks”. In: *Proceedings of the International AAAI Conference on Web and Social Media*. Vol. 4. 1. 2010.

[56] Lozares, Carlos, López-Roldán, Pedro, Bolibar, Mireia, and Muntanyola, Dafne. “The structure of global centrality measures”. In: *International Journal of Social Research Methodology* 18.2 (2015), pp. 209–226.

[57] Mises, Richard von and Pollaczek-Geiringer, Hilda. “Praktische Verfahren der Gleichungsauflösung.” In: *Journal of Applied Mathematics and Mechanics/Zeitschrift für Angewandte Mathematik und Mechanik* 9.1 (1929), pp. 58–77.

[58] Nieminen, Juhani. “On the centrality in a directed graph”. In: *Social Science Research* 2.4 (1973), pp. 371–378.

[59] Nieminen, Juhani. “On the centrality in a graph”. In: *Scandinavian Journal of Psychology* 15.1 (1974), pp. 332–336.

[60] Oldham, Stuart, Fulcher, Ben, Parkes, Linden, Arnatkevičiūtė, Aurina, Suo, Chao, and Fornito, Alex. “Consistency and differences between centrality measures across distinct classes of networks”. In: *PloS one* 14.7 (2019), e0220061.

[61] Opsahl, Tore, Agneessens, Filip, and Skvoretz, John. “Node centrality in weighted networks: Generalizing degree and shortest paths”. In: *Social Networks* 32.3 (2010), pp. 245–251.

[62] Page, Lawrence, Brin, Sergey, Motwani, Rajeev, and Winograd, Terry. *The PageRank citation ranking: bringing order to the web*. Tech. rep. Stanford InfoLab, 1999.

[63] Palacios-Huerta, Ignacio and Volij, Oscar. “The measurement of intellectual influence”. In: *Econometrica* 72.3 (2004), pp. 963–977.

[64] Palbom, Anna. “Complexity of the directed spanning cactus problem”. In: *Discrete Applied Mathematics* 146.1 (2005), pp. 81–91.

[65] Peano, Giuseppe. *Arithmetices principia: Nova methodo exposita*. Fratres Bocca, 1889.

[66] Pinski, Gabriel and Narin, Francis. "Citation influence for journal aggregates of scientific publications: Theory, with application to the literature of physics". In: *Information Processing & Management* 12.5 (1976), pp. 297–312.

[67] Radicchi, Filippo. "Who is the best player ever? A complex network analysis of the history of professional tennis". In: *PLOS One* 6.2 (2011), e17249.

[68] Riveros, Cristian and Salas, Jorge. "A family of centrality measures for graph data based on subgraphs". In: *23rd International Conference on Database Theory*. Schloss Dagstuhl-Leibniz-Zentrum für Informatik. 2020.

[69] Sabidussi, Gert. "The centrality index of a graph". In: *Psychometrika* 31.4 (1966), pp. 581–603.

[70] Schoch, David and Brandes, Ulrik. "Re-conceptualizing centrality in social networks". In: *European Journal of Applied Mathematics* 27.6 (2016), pp. 971–985.

[71] Sciarra, Carla, Chiarotti, Guido, Laio, Francesco, and Ridolfi, Luca. "A change of perspective in network centrality". In: *Scientific Reports* 8.1 (2018), pp. 1–9.

[72] Seeley, John R. "The net of reciprocal influence. a problem in treating sociometric data". In: *Canadian Journal of Experimental Psychology* 3 (1949), p. 234.

[73] Shannon, Paul, Markiel, Andrew, Ozier, Owen, Baliga, Nitin S., Wang, Jonathan T., Ramage, Daniel, Amin, Nada, Schwikowski, Benno, and Ideker, Trey. "Cytoscape: a software environment for integrated models of biomolecular interaction networks". In: *Genome Research* 13.11 (2003), pp. 2498–2504.

[74] Shao, Chengcheng, Cui, Pengshuai, Xun, Peng, Peng, Yuxing, and Jiang, Xinwen. "Rank correlation between centrality metrics in complex networks: an empirical study". In: *Open Physics* 16.1 (2018), pp. 1009–1023.

[75] Skibski, Oskar, Michalak, Tomasz P., and Rahwan, Talal. "Axiomatic Characterization of Game-Theoretic Centrality". In: *Journal of Artificial Intelligence Research* 62 (2018), pp. 33–68.

[76] Skibski, Oskar, Rahwan, Talal, Michalak, Tomasz P., and Yokoo, Makoto. "Attachment centrality: Measure for connectivity in networks". In: *Artificial Intelligence* 274 (2019), pp. 151–179.

[77] Skibski, Oskar and Sosnowska, Jadwiga. "Axioms for Distance-Based Centralities". In: *Proceedings of the 32nd AAAI Conference on Artificial Intelligence (AAAI)*. 2018, pp. 1218–1225.

[78] Valente, Thomas W., Coronges, Kathryn, Lakon, Cynthia, and Costenbader, Elizabeth. "How correlated are network centrality measures?" In: *Connections (Toronto, Ont.)* 28.1 (2008), p. 16.

[79] Vigna, Sebastiano. "Spectral ranking". In: *Network Science* 4.4 (2016), pp. 433–445.

[80] Wang, Zhifang, Scaglione, Anna, and Thomas, Robert J. "Electrical centrality measures for electric power grid vulnerability analysis". In: *Proceedings of the 49th IEEE Conference on Decision and Control (CDC)*. 2010, pp. 5792–5797.

[81] Wąs, Tomasz, Rahwan, Talal, and Skibski, Oskar. "Random Walk Decay Centrality". In: *Proceedings of the 33rd AAAI Conference on Artificial Intelligence (AAAI)*. 2019, pp. 2197–2204.

[82] Wąs, Tomasz and Skibski, Oskar. "An axiom system for feedback centralities". In: *Proceedings of the 30th International Joint Conference on Artificial Intelligence (IJCAI)*. 2021, pp. 443–449.

[83] Wąs, Tomasz and Skibski, Oskar. "An Axiomatization of the Eigenvector and Katz Centralities". In: *Proceedings of the 32nd AAAI Conference on Artificial Intelligence (AAAI)*. 2018, pp. 1258–1265.

[84] Wąs, Tomasz and Skibski, Oskar. "Axiomatization of the PageRank centrality". In: *Proceedings of the 27th International Joint Conference on Artificial Intelligence (IJCAI)*. 2018, pp. 3898–3904.

[85] Wąs, Tomasz, Waniek, Marcin, Rahwan, Talal, and Michalak, Tomasz P. "The Manipulability of Centrality Measures-An Axiomatic Approach". In: *Proceedings of the 19th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*. 2020, pp. 1467–1475.

[86] Weng, Jianshu, Lim, Ee-Peng, Jiang, Jing, and He, Qi. "Twitterrank: finding topic-sensitive influential twitterers". In: *Proceedings of the 3rd ACM International Conference on Web Search and Data Mining (WSDM)*. 2010, pp. 261–270.

[87] White, Scott and Smyth, Padhraic. "Algorithms for estimating relative importance in networks". In: *Proceedings of the 9th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM. 2003, pp. 266–275.

[88] Zhirov, Alexander O., Zhirov, Oleg V., and Shepelyansky, Dima L. "Two-dimensional ranking of Wikipedia articles". In: *The European Physical Journal B* 77.4 (2010), pp. 523–531.