1 Introduction and Overview

Many problems in combinatorics, number theory, probability theory, reliability theory and statistics can be solved by applying a unifying method, which is known as the *principle of inclusion-exclusion*. The principle of inclusion-exclusion expresses the indicator function of a union of finitely many sets as an alternating sum of indicator functions of their intersections. More precisely, for any finite family of sets $\{A_v\}_{v\in V}$ the principle of inclusion-exclusion states that

$$\chi\left(\bigcup_{v\in V} A_v\right) = \sum_{\substack{I\subseteq V\\ I\neq\emptyset}} (-1)^{|I|-1}\chi\left(\bigcap_{i\in I} A_i\right), \tag{1.1}$$

where $\chi(A)$ denotes the indicator function of A with respect to some universal set Ω, that is, $\chi(A)(\omega) = 1$ if $\omega \in A$, and $\chi(A)(\omega) = 0$ if $\omega \in \Omega\backslash A$. Equivalently, (1.1) can be expressed as $\chi\big(\bigcap_{v\in V} \complement A_v\big) = \sum_{I\subseteq V}(-1)^{|I|}\chi\big(\bigcap_{i\in I} A_i\big)$, where $\complement A_v$ denotes the complement of A_v in Ω and $\bigcap_{i\in\emptyset} A_i = \Omega$. A proof by induction on the number of sets is a common task in undergraduate courses. Usually, the A_v's are measurable with respect to some finite measure μ on a σ-field of subsets of Ω. Integration of the indicator function identity (1.1) with respect to μ then gives

$$\mu\left(\bigcup_{v\in V} A_v\right) = \sum_{\substack{I\subseteq V\\ I\neq\emptyset}} (-1)^{|I|-1}\mu\left(\bigcap_{i\in I} A_i\right), \tag{1.2}$$

which now expresses the measure of a union of finitely many sets as an alternating sum of measures of their intersections. The step leading from (1.1) to (1.2) is referred to as the *method of indicators* [GS96b]. Naturally, two special cases are of particular interest, namely the case where μ is the counting measure on the power set of Ω, and the case where μ is a probability measure on a σ-field of subsets of Ω. These special cases are sometimes attributed to Sylvester (1883) and Poincaré (1896), although the second edition of Montmort's book "Essai d'Analyse sur les Jeux de Hazard", which appeared in 1714, already contains an implicit use of the method, based on a letter by N. Bernoulli in 1710. A first

K. Dohmen: LNM 1826, pp. 1–4, 2003.

explicit description of the method was given by Da Silva (1854). For references to these sources and additional historical notes, we refer to Takács [Tak67].

Since the identities (1.1) and (1.2) contain $2^{|V|}-1$ terms and intersections of up to $|V|$ sets, one often resorts to inequalities like that of Boole [Boo54]:

$$\chi\left(\bigcup_{v\in V} A_v\right) \leq \sum_{i\in V} \chi(A_i)\,. \tag{1.3}$$

A more general result, first discovered by Ch. Jordan [Jor27] and later by Bonferroni [Bon36], states that for any finite family of sets $\{A_v\}_{v\in V}$ and any $r\in\mathbb{N}$,

$$\chi\left(\bigcup_{v\in V} A_v\right) \leq \sum_{\substack{I\subseteq V\\ 0<|I|\leq r}} (-1)^{|I|-1}\chi\left(\bigcap_{i\in I} A_i\right) \quad (r \text{ odd}), \tag{1.4}$$

$$\chi\left(\bigcup_{v\in V} A_v\right) \geq \sum_{\substack{I\subseteq V\\ 0<|I|\leq r}} (-1)^{|I|-1}\chi\left(\bigcap_{i\in I} A_i\right) \quad (r \text{ even}). \tag{1.5}$$

Nowadays, these inequalities are usually referred to as *Bonferroni inequalities*. Again, there is no real restriction in using indicator functions rather than measures, since these inequalities can be integrated with respect to any finite measure μ (e.g., a probability measure) on any σ-field containing the sets A_v, $v\in V$.

Numerous extensions of the classical Bonferroni inequalities (1.4) and (1.5) were established in the second half of the 20th century. An excellent survey of the various results, applications and methods of proof is given in the recent book of Galambos and Simonelli [GS96b]. The following inequalities due to Galambos [Gal75], which are valid for any finite collection of sets $\{A_v\}_{v\in V}$, improve (1.4) and (1.5) by including additional terms based on the $(r+1)$-subsets of V:

$$\chi\left(\bigcup_{v\in V} A_v\right) \leq \sum_{\substack{I\subseteq V\\ 0<|I|\leq r}} (-1)^{|I|-1}\chi\left(\bigcap_{i\in I} A_i\right) - \frac{r+1}{|V|}\sum_{\substack{I\subseteq V\\ |I|=r+1}} \chi\left(\bigcap_{i\in I} A_i\right) \ (r \text{ odd}),$$

$$\chi\left(\bigcup_{v\in V} A_v\right) \geq \sum_{\substack{I\subseteq V\\ 0<|I|\leq r}} (-1)^{|I|-1}\chi\left(\bigcap_{i\in I} A_i\right) + \frac{r+1}{|V|}\sum_{\substack{I\subseteq V\\ |I|=r+1}} \chi\left(\bigcap_{i\in I} A_i\right) \ (r \text{ even}).$$

Inequalities for the measure or indicator function of a union which are valid for any finite collection of sets like the preceding ones are frequently referred to as *Bonferroni-type inequalities* [GS96a, GS96b] or as *inequalities of Bonferroni-Galambos type* [MS85, Mǎr89, TX89]. A new inequality of Bonferroni-Galambos type based on chordal graphs will be established in Chapter 4 of this book.

The main part of this work deals with improved Bonferroni inequalities and improved inclusion-exclusion identities that require the collection of sets to satisfy some structural restrictions. Examples of such well-structured collections of sets arise in some problems of statistical inference [NW92, NW97], reliability

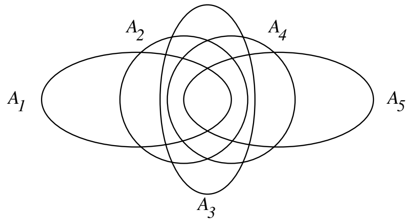

Figure 1.1: A Venn diagram of five sets.

theory [Doh98, Doh03, GNW02] and chromatic graph theory [Doh99a, Doh99d]. We shall mainly be interested in inclusion-exclusion identities of the form

$$\chi\left(\bigcup_{v\in V} A_v\right) = \sum_{I\in\mathcal{S}}(-1)^{|I|-1}\chi\left(\bigcap_{i\in I} A_i\right)$$

and inequalities of type

$$\chi\left(\bigcup_{v\in V} A_v\right) \leq \sum_{\substack{I\in\mathcal{S}\\ |I|\leq r}}(-1)^{|I|-1}\chi\left(\bigcap_{i\in I} A_i\right) \quad (r \text{ odd}), \tag{1.6}$$

$$\chi\left(\bigcup_{v\in V} A_v\right) \geq \sum_{\substack{I\in\mathcal{S}\\ |I|\leq r}}(-1)^{|I|-1}\chi\left(\bigcap_{i\in I} A_i\right) \quad (r \text{ even}), \tag{1.7}$$

where $\mathcal{S}$ is a restricted set of non-empty subsets of V, and where (1.6) and (1.7) are at least as sharp as their classical counterparts (1.4) and (1.5). A first straightforward approach is to account only for non-empty subsets I of V satisfying $\bigcap_{i\in I} A_i \neq \emptyset$. In fact, Lozinskii [Loz92] demonstrates that a skillful implementation of this approach leads to a reduction of the average running time of the standard inclusion-exclusion algorithm for counting satisfying assignments of propositional formulae in conjunctive normal form. In the present work, however, we are interested in more subtle improvements that arise from logical dependencies of the sets involved. Consider for instance the five sets A_1–A_5, whose Venn diagram is shown in Figure 1.1. The classical inclusion-exclusion identity for the indicator function of the union of these sets gives a sum of $2^5 - 1 = 31$ terms, many of which are equal with opposite sign. After cancelling out, we are left with

$$\begin{aligned}\chi(A_1\cup A_2\cup A_3\cup A_4\cup A_5) &= \chi(A_1)+\chi(A_2)+\chi(A_3)+\chi(A_4)+\chi(A_5)\\ &\quad -\chi(A_1\cap A_2)-\chi(A_2\cap A_3)-\chi(A_3\cap A_4)-\chi(A_4\cap A_5)\,,\end{aligned}$$

which contains only 9 terms. Our purpose is to predict such cancellations in advance and thus to obtain shorter inclusion-exclusion identities and sharper Bonferroni inequalities for the indicator function of a union.

The book is organized as follows: In Chapter 2 we review some basic notions of graph theory, lattice theory and combinatorial topology. Readers with an appropriate background in these disciplines may want to skip this chapter.

Then, in Chapter 3, we give an introduction to the upcoming theory of abstract tubes, which was initiated by Naiman and Wynn [NW92, NW97], and which serves as a framework for establishing improved inclusion-exclusion identities as well as improved Bonferroni inequalities which are provably at least as sharp as their classical counterparts. Some appealing geometric examples of abstract tubes demonstrate the applicability of this beautyful new theory.

In Chapter 4 the main results of abstract tube theory in Chapter 3 are applied in establishing improved inclusion-exclusion identities and Bonferroni inequalities associated with closure and kernel operators. Several results from the literature such as the semilattice sieve of Narushima [Nar74, Nar82] and the tree sieve of Naiman and Wynn [NW92] are thus rediscovered in a unified way. We also provide some elementary alternative proofs (not using abstract tubes) as well as a generalization of one of the identities. One of these alternative proofs uses Zeilberger's abstract lace expansion [Zei97], which turns out to be a valuable tool in establishing new inclusion-exclusion expansions. Inspired by our results on closure operators and abstract tubes, we derive a new generally-valid inequality of Bonferroni-Galambos type, where the selection of intersections in the estimates is determined by a chordal graph. By varying this graph, several well-known and new Bonferroni-type inequalities are obtained in a unified way.

In Chapter 5 we deduce some recursive schemes from our inclusion-exclusion results in Chapter 4, which will then be used in Chapter 6 in deriving Shier's recursive algorithm and semilattice expression for system reliability [Shi88, Shi91].

In Chapter 6 our results of the preceding chapters are applied to reliability analysis of coherent binary systems such as communications networks, k-out-of-n systems and consecutive k-out-of-n systems. We then turn our attention to reliability covering problems and identify a comprehensive class of hypergraphs for which the coverage probability can be computed in polynomial time.

In Chapter 7, which is devoted to applications to combinatorics and related topics, we give an abstract tube generalization of Narushima's principle of inclusion-exclusion on partition lattices [Nar74, Nar77] and of Whitney's broken circuit theorem on the chromatic polynomial of a graph. The results on the chromatic polynomial are then extended to a new two-variable polynomial generalizing the chromatic polynomial, the independence polynomial, and the matching polynomial. After generalizing our inclusion-exclusion identity for kernel operators to a result on sums over partially ordered sets, similar results as for the chromatic polynomial are obtained for the Tutte polynomial, the characteristic polynomial and the β invariant of a matroid, the Euler characteristic of an abstract simplicial complex and the Möbius function of a finite lattice.