Das Prinzip von Einschluss und Ausschluss

Kord Eickmeyer

OWO-Vorlesung, 7. Oktober 2019

1 Mengen

Der Begriff der *Menge* (also einer Ansammlung von Objekten) wird als bekannt vorausgesetzt. Wir benötigen in dieser Vorlesung nur endliche Mengen, die wir wie üblich mit geschweiften Klammern schreiben, z.B. $\{a, b, c, d\}$. Mengen bezeichnen wir in der Regel mit lateinischen Großbuchstaben (etwa M, S, A_1, $\ldots$), Elemente mit Kleinbuchstaben (es sei denn, die Elemente sind selbst Mengen).

Folgende Schreibweisen sollten bekannt sein:

Schreibweise	Bedeutung
$\emptyset$	die leere Menge
$x \in A$	„x ist Element der Menge A“
$A \subseteq B$	„Die Menge A ist eine Teilmenge der Menge B“, d.h. wann immer $x \in A$ gilt, gilt auch $x \in B$. Für jede Menge A gilt $A \subseteq A$.
$A \cap B$	die Schnittmenge der Mengen A und B, z.B. $\{a, b, c\} \cap \{b, c, d\} = \{b, c\}$.
$A \cup B$	die Vereinigung der Mengen A und B, z.B. $\{a, b, c\} \cup \{b, c, d\} = \{a, b, c, d\}$
$A \setminus B$	die Menge aller Elemente, die in A aber nicht in B liegen
$\|A\|$	die Anzahl der Elemente der Menge A.

Aus der Schule ist vielleicht die Summenschreibweise $\sum_{i=1}^{n} a_i$ für $a_1 + a_2 + \cdots + a_n$ bekannt. Wir gehen mit dieser Schreibweise häufig recht locker um, sofern wir davon ausgehen, dass der Leser versteht, was wir meinen (eine gefährliche aber verbreitete Annahme). Die Summe aus dem Beispiel könnte also auch als

$$\sum_{i \in \{1,2,\ldots,n\}} a_i \qquad \text{oder gar} \qquad \sum_i a_i \tag{1.1}$$

geschrieben werden. Entsprechend schreiben wir auch

$$\bigcup_{i=1}^{n} A_i \quad \text{für} \quad A_i \cup A_2 \cup \cdots \cup A_n. \tag{1.2}$$

Jetzt wird es etwas komplizierter: Wenn $I \subseteq \{1, \dots, n\}$ ist, schreiben wir

$$\bigcup_{i \in I} A_i \quad \text{bzw.} \quad \bigcup_I A_i \tag{1.3}$$

für die Vereinigung aller Mengen A_i, für die $i \in I$ gilt, also z.B.

$$\bigcup_{i \in \{1,3,5\}} A_i = A_1 \cup A_3 \cup A_5. \tag{1.4}$$

Was ist dann aber $\bigcup_\emptyset A_i$? Da wir gar keine Mengen vereinigt haben, bietet sich als natürliche Definition

$$\bigcup_\emptyset A_i := \emptyset \tag{1.5}$$

an (schließlich ist $\emptyset$ die „kleinste" mögliche Menge). Analog definieren wir

$$\bigcap_{i \in I} A_i \quad \text{bzw.} \quad \bigcap_I A_i \tag{1.6}$$

als Durchschnitt der Mengen A_i, für die $i \in I$ gilt. Hier gibt es keine natürliche Wahl von $\bigcap_\emptyset A_i$, da diese Menge im Prinzip „alle möglichen Elemente" enthalten müsste (wir haben ja nichts „weggeschnitten"). Wenn aber $A_1, \dots, A_n \subseteq S$ für eine im Kontext gegebene Menge S gilt, setzen wir

$$\bigcap_\emptyset A_i := S. \tag{1.7}$$

Dies ist eine willkürliche Festlegung, die sich für uns als sinnvoll erweisen wird.

2 Das Prinzip von Einschluss und Ausschluss

Wir wollen folgenden Satz beweisen:

Satz 2.1. *Sei S eine endliche Menge und $A_1, \dots, A_n \subseteq S$. Dann gilt*

$$\left| S \setminus \bigcup_{i=1}^n A_i \right| = \sum_{I \subseteq \{1,\dots,n\}} (-1)^{|I|} \left| \bigcap_I A_i \right|. \tag{2.1}$$

Um den Satz und seinen Beweis etwas übersichtlicher aufschreiben zu können definieren wir

$$A_I := \bigcap_{i \in I} A_i \quad \text{und} \quad [n] := \{1, \dots, n\}. \tag{2.2}$$

Die Gleichung aus dem Satz lautet damit

$$\left| S \setminus \bigcup_{i=1}^n A_i \right| = \sum_{I \subseteq [n]} (-1)^{|I|} \left| A_I \right|. \tag{2.3}$$

Beweis. Wir nehmen an, der Satz wäre falsch und es gäbe Mengen S und $A_1, \ldots, A_n \subseteq S$, für die die Formel falsch ist. Wir können unser Gegenbeispiel so wählen, dass n minimal ist, dass also für alle Mengensysteme T und $B_1, \ldots, B_m \subseteq T$ mit $m < n$ die Formel gilt.

Nun muss offenbar $n > 1$ gelten, da im Fall $n = 1$

$$\begin{aligned} |S \setminus A_1| &= |S| - |A_1| \\ &= (-1)^{|\emptyset|} |A_\emptyset| + (-1)^{|\{1\}|} \left|A_{\{1\}}\right| \end{aligned} \tag{2.4}$$

gilt; hier ist die Formel also richtig. Außerdem gilt

$$S \setminus \bigcup_{i=1}^{n} A_i = \Big(S \setminus \bigcup_{i=1}^{n-1} A_i\Big) \setminus A_n. \tag{2.5}$$

Die Anzahl der Elemente der ersten Menge auf der rechten Seite kennen wir, es gilt nämlich

$$\left|S \setminus \bigcup_{i=1}^{n-1} A_i\right| = \sum_{I \subseteq [n-1]} (-1)^{|I|} |A_I|. \tag{2.6}$$

(Genau deswegen haben wir unser Gegenbeispiel so gewählt, dass n minimal ist.) Hiervon müssen wir noch die Anzahl der Elemente abziehen, die in A_n liegen, aber in keiner der Mengen $A_1, \ldots, A_{n-1}$. Dafür setzen wir

$$\begin{aligned} T &:= A_n \\ B_i &:= A_i \cap A_n \quad \text{für } i = 1, \ldots, n-1 \end{aligned} \tag{2.7}$$

Wir suchen also die Anzahl der Elemente der Menge

$$T \setminus \bigcup_{i=1}^{n-1} B_i \tag{2.8}$$

und können wieder unsere Formel verwenden:

$$\left|T \setminus \bigcup_{i=1}^{n-1} B_i\right| = \sum_{I \subseteq [n-1]} (-1)^{|I|} |B_I|. \tag{2.9}$$

Jetzt bauen wir die beiden Ergebnisse zusammen:

$$
\begin{aligned}
\left|S \setminus \bigcup_{i=1}^{n} A_i\right| &= \left|\Big(S \setminus \bigcup_{i=1}^{n-1} A_i\Big) \setminus A_n\right| \\
&= \left|S \setminus \bigcup_{i=1}^{n-1} A_i\right| - \left|T \setminus \bigcup_{i=1}^{n-1} B_i\right| \\
&= \sum_{I\subseteq[n-1]} (-1)^{|I|} \left|A_I\right| - \sum_{I\subseteq[n-1]} (-1)^{|I|} \left|B_I\right|. \\
&= \sum_{I\subseteq[n-1]} (-1)^{|I|} \left|A_I\right| - \sum_{I\subseteq[n-1]} (-1)^{|I|} \left|\bigcap_{i\in I}(A_i \cap A_n)\right|. \qquad (2.10) \\
&= \sum_{I\subseteq[n-1]} (-1)^{|I|} \left|A_I\right| - \sum_{I\subseteq[n-1]} (-1)^{|I|} \left|A_n \cap A_I\right|. \\
&= \sum_{I\subseteq[n-1]} (-1)^{|I|} \left|A_I\right| - \sum_{I\subseteq[n-1]} (-1)^{|I|} \left|A_{I\cup\{n\}}\right|. \\
&= \sum_{I\subseteq[n-1]} (-1)^{|I|} \left|A_I\right| + \sum_{I\subseteq[n-1]} (-1)^{|I\cup\{n\}|} \left|A_{I\cup\{n\}}\right|. \\
&= \sum_{I\subseteq[n]} (-1)^{|I|} \left|A_I\right|,
\end{aligned}
$$

was zu beweisen war. □

3 Was bedeutet das?

Was wir gerade gemacht haben ist typisch für Mathematik: Wir haben mehrere kleine Schritte gemacht, von denen jeder einzelne leicht nachvollziehbar ist. Trotzdem geht sehr schnell die Übersicht verloren. Die wollen wir uns jetzt zurückerobern.

Kleine Beispiele

Zunächst mal die Formel:

$$\left|S \setminus \bigcup_{i=1}^{n} A_i\right| = \sum_{I\subseteq[n]} (-1)^{|I|} \left|A_I\right|. \qquad (3.1)$$

Was heißt das für kleine Werte von n?

n	$\lvert S \setminus \bigcup A_i\rvert$
0	$\lvert S\rvert$
1	$\lvert S\rvert - \lvert A_1\rvert$
2	$\lvert S\rvert - \lvert A_1\rvert - \lvert A_2\rvert + \lvert A_1 \cap A_2\rvert$
3	$\lvert S\rvert - \lvert A_1\rvert - \lvert A_2\rvert - \lvert A_3\rvert$
	$+ \lvert A_1 \cap A_2\rvert + \lvert A_1 \cap A_3\rvert + \lvert A_2 \cap A_3\rvert - \lvert A_1 \cap A_2 \cap A_3\rvert$

Mit Mengendiagrammen können wir uns das recht leicht klarmachen. Hilfreich kann auch ein Beispiel sein:

Beispiel 3.1. In einem Jahrgang mit 100 Studierenden sind 45 Darmstädter, 53 hatte Mathe-LK und 55 hatten Physik-LK. Von den Darmstädtern hatten 28 Mathe-LK und 32 Physik-LK. 35 der Studierenden hatten sowohl Mathe- als auch Physik-LK, und 20 der Darmstädter hatten sowohl Mathe- als auch Physik-LK.

Wieviele Studierende kamen weder aus Darmstadt, noch hatten sie Mathe- oder Physik-LK?

Ohne unsere Formel können wir die Aufgabe mit einem Mengendiagramm lösen, in das wir der Reihe nach die bekannten Größen eintragen. Mit der Formel erhalten wir

$$100 - 45 - 53 - 55 + 28 + 32 + 35 - 20 = 22. \tag{3.2}$$

Eine andere Sichtweise

Oft hilft es auch, eine andere Sichtweise auf das Problem zu suchen.

Wir hatten die Mengen $A_I = \bigcap_I A_i$ definiert, die genau die Elemente enthalten, die in allen Mengen A_i mit $i \in I$ liegen, anders geschrieben

$$A_I = \{x \in S \mid x \in A_i \text{ wenn } i \in I\}. \tag{3.3}$$

Ob ein $x \in A_I$ gilt hängt dabei nicht davon ab, ob x in einer Menge A_j liegt, deren Index j nicht in I enthalten ist. Das können wir ändern. Wir setzen

$$B_I = \{x \in S \mid x \in A_i \textit{ genau dann,} \text{ wenn } i \in I\}. \tag{3.4}$$

In unserem Mengendiagramm oben sind die Mengen B_I genau die einzelnen Regionen des Diagramms. Und die gesuchte Zahl ist die Anzahl der Elemente in

$$B_\emptyset = S \setminus \bigcup_i A_i. \tag{3.5}$$

Wir wollen also $|B_\emptyset|$ aus den Zahlen $|A_I|$ für $I \subseteq [n]$ ausrechnen. Als wir das Mengendiagram der Reihe nach ausgefüllt haben, haben wir nebenbei auch alle anderen $|B_I|$ ausgerechnet.

Dabei haben wir nichts anderes getan, als ein bestimmtes lineares Gleichungssystem zu lösen. Es gilt nämlich:

$$|A_I| = \sum_{I \subseteq J} |B_J| \tag{3.6}$$

(Achtung, die Summe läuft über alle J, für die I eine Teilmenge von J ist.) Wir vereinfachen wieder unsere Notation ein wenig, indem wir $a_I := |A_I|$

und $b_I := |B_I|$ schreiben; außerdem lassen wir bei den Indizes die Mengenklammern und Kommas weg und schreiben z.B. b_{12} statt $b_{\{1,2\}}$. Damit liefert die obige Gleichung im Fall $n = 3$:

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} b_\emptyset \\ b_1 \\ b_2 \\ b_3 \\ b_{12} \\ b_{13} \\ b_{23} \\ b_{123} \end{pmatrix} = \begin{pmatrix} a_\emptyset \\ a_1 \\ a_2 \\ a_3 \\ a_{12} \\ a_{13} \\ a_{23} \\ a_{123} \end{pmatrix} \tag{3.7}$$

Wenn wir dieses Gleichungssystem mit dem Gauß-Verfahren lösen, erhalten wir

$$\begin{pmatrix} b_\emptyset \\ b_1 \\ b_2 \\ b_3 \\ b_{12} \\ b_{13} \\ b_{23} \\ b_{123} \end{pmatrix} = \begin{pmatrix} 1 & -1 & -1 & -1 & 1 & 1 & 1 & -1 \\ 0 & 1 & 0 & 0 & -1 & -1 & 0 & 1 \\ 0 & 0 & 1 & 0 & -1 & 0 & -1 & 1 \\ 0 & 0 & 0 & 1 & 0 & -1 & -1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} a_\emptyset \\ a_1 \\ a_2 \\ a_3 \\ a_{12} \\ a_{13} \\ a_{23} \\ a_{123} \end{pmatrix}, \tag{3.8}$$

und die erste Zeile der Matrix ist genau die Formel aus Satz 2.1. Die anderen Zeilen der Lösungsmatrix lassen sich ähnlich interpretieren. Übung: Übersetzen Sie die anderen Zeilen zurück in die Mengenschreibweise und finden Sie ein allgemeines Muster.

4 Warum machen wir das?

Satz 2.1 können wir verwenden, um kompliziertere kombinatorische Fragestellungen zu lösen, etwa die folgende:

Frage 4.1. Wieviele Möglichkeiten gibt es, n Personen auf k Räume zu verteilen, wenn kein Raum leer bleiben soll?

Für $k > n$ ist die Antwort offensichtlich 0, da ja kein Raum leer sein soll. Für $n = k$ kennen Sie die Antwort wahrscheinlich schon aus der Schule: $n! = 1 \cdot 2 \cdot 3 \cdot \cdots \cdot n$, da hier in jedem Raum genau eine Person sein muss.

Für beliebige Werte $k < n$ ist die Frage deutlich kniffliger. Wir numerieren die Räume mit Zahlen 1 bis k. Wir bezeichnen mit S die Menge aller Möglichkeiten, wie sich n Personen auf k Räume verteilen; damit gilt $|S| = k^n$. Mit A_m bezeichnen wir die Menge aller Möglichkeiten, bei denen (mindestens) der Raum m leer bleibt; das sind $(k-1)^n$, da sich die n Personen auf die verbleibenden $k - 1$ Räume aufteilen. Für die Menge $A_{\{m_1,m_2\}}$

der Möglichkeiten, bei denen (mindestens) die Räume m_1 und m_2 leer bleiben, erhalten wir $\left|A_{\{m_1,m_2\}} = (k-2)^n\right|$ usw. Allgemein gilt für die Menge A_I aller Möglichkeiten, wie sich n Personen auf die Räume $[k] \setminus I$ verteilen,

$$|A_I| = (k - |I|)^n. \tag{4.1}$$

Damit gilt nach Satz 2.1:

$$\begin{aligned} \left|S \setminus \bigcup_i A_i\right| &= \sum_{I \subseteq [k]} (-1)^{|I|} |A_I| \\ &= \sum_{I \subseteq [k]} (-1)^{|I|} (k - |I|)^n \\ &= \sum_{m=0}^{k} (-1)^m \binom{k}{m} (k-m)^n \end{aligned} \tag{4.2}$$

Insbesondere gilt für $k = n$ wieder

$$n! = \sum_{m=0}^{n} (-1)^m \binom{n}{m} (n-m)^n. \tag{4.3}$$

Beispiel 4.2. Für 4 Personen und 2 Räume erhalten wir

$$(-1)^0 \binom{2}{0} 2^4 + (-1)^1 \binom{2}{1} 1^4 + (-1)^2 \binom{2}{0} 0^4 = 16 - 2 + 0 = 14 \tag{4.4}$$

Möglichkeiten: $123|4, 124|3, 134|2, 234|1, 12|34, 13|24, 14|23$ und jeweils die entsprechenden Möglichkeiten mit linker und rechter Seite vertauscht.

Weitere Anwendungen von Satz 2.1 werden Sie auf dem Übungszettel sehen.

5 Die Möbius'sche Umkehrformel

Jetzt wollen wir Satz 2.1 noch einmal stark verallgemeinern. Dazu schauen wir uns die Formulierung als lineares Gleichungssystem (3.7) an. Die Indizes der Variablen $a_\emptyset$, $a_{\{1\}}$ sind gerade die Teilmengen der Menge $[n]$. Diese fassen wir in der *Potenzmenge* $2^{[n]}$ von $[n]$ zusammen:

$$\begin{aligned} 2^{[n]} &= \{I \mid I \subseteq [n]\} \\ &= \{\emptyset, \{1\}, \{2\}, \ldots, \{n\}, \{1,2\}, \ldots, [n]\} \end{aligned} \tag{5.1}$$

Als wir das Gleichungssystem aufgestellt haben, mussten wir die Variablen auf die Spalten und Zeilen der Matrix aufteilen. Ganz willkürlich haben wir diese Zuteilung nicht durchgeführt: Auf der Menge $2^{[n]}$ ist eine sogenannte *partielle Ordnung* durch die Teilmengenbeziehung $\subseteq$ definiert, und wir haben diese partielle Ordnung zu einer linearen Ordnung erweitert (d.h. wenn $I \subseteq J$, dann steht a_I vor a_J in der Ordnung).

Anstelle der Variablen a_I und b_I können wir auch Funktionen schreiben:

$$f : 2^{[n]} \to \mathbb{R}, I \mapsto b_I \qquad \text{bzw.} \qquad \hat{f} : 2^{[n]} \to \mathbb{R}, I \mapsto a_I, \tag{5.2}$$

wobei die Funktionen f und $\hat{f}$ wie folgt zusammenhängen:

$$\hat{f}(I) = \sum_{I \subseteq J} f(J) \tag{5.3}$$

Diese Situation kommt häufiger vor: Gegeben ist eine Menge P mit einer partiellen Ordnung $\preceq$, d.h. für bestimmte Paare $x, y \in P$ gilt $x \preceq y$, wobei

- $x \preceq x$ für alle $x \in P$,
- wenn $x \preceq y$ und $y \preceq z$ für $x, y, z \in P$, dann auch $x \preceq z$, und
- wenn $x \preceq y$ und $y \preceq x$, dann ist $x = y$.

Nun sind zwei Funktionen $f, \hat{f} : P \to \mathbb{R}$ gegeben, die über die Beziehung

$$\hat{f}(x) = \sum_{y \preceq x} f(x) \tag{5.4}$$

zusammenhängen (beim Prinzip von Einschluss und Ausschluss ist $x \preceq y$ falls $y \subseteq x$; es ist aber üblich, in der allgemeinen Formulierung über die Elemente zu summieren, die *kleiner* als x sind).

Beispiel 5.1. Für jede natürliche Zahl n ist die Menge $[n]$ partiell geordnet, in dem wir festlegen

$$a \preceq b \text{ genau dann, wenn } a \text{ ein Teiler von } b \text{ ist.} \tag{5.5}$$

Die Funktion

$$\varphi(m) = \big|\{k \in [m] \mid k \text{ teilerfremd zu } m\}\big| \tag{5.6}$$

heißt *Euler'sche Phi-Funktion* und spielt in der Zahlentheorie eine wichtige Rolle. Hier ist $\hat{\varphi}(n) = n$. (Warum?)

Wir kennen also $\hat{f}$ und möchten f ausrechnen. Geht das? Wenn ja wie? Die erste Frage können wir mit dem, was wir in der linearen Algebra im ersten Semester lernen werden, leicht beantworten: Wenn wir die partielle Ordnung zu einer totalen Ordnung erweitern (so, wie wir es in Kapitel 3 getan haben), erhalten wir wieder ein lineares Gleichungssystem, dessen Koeffizientenmatrix eine Dreiecksmatrix mit 1en auf der Diagonalen ist, und solche Matrizen sind invertierbar; es ist also möglich, f aus $\hat{f}$ auszurechnen. Genauer gesagt gilt analog zu Gleichung (3.8): Es gibt eine Funktion $\mu : P \times P \to \mathbb{R}$, für die gilt:

$$f(x) = \sum_{y \preceq x} \mu(y, x) \hat{f}(y) \tag{5.7}$$

Die Funktion μ heißt *Möbiusfunktion* der partiell geordneten Menge P; sie hängt nur von der Menge P un der partiellen Ordnung $\preceq$ ab.

Frage 5.2. Beim Prinzip von Einschluss und Ausschluss haben wir die Funktionswerte $\mu(I, \emptyset)$ für alle I bestimmt. Wie lauten sie? Wie lauten die restlichen Funktionswerte?

Die Funktion μ kann natürlich mit Hilfe des Gauß'schen Eliminationsverfahrens berechnet werden, für viele partiell geordnete Mengen P sind jedoch explizite Formeln bekannt (Satz 2.1 ist ein Beispiel dafür). Diese beruhen oft darauf, dass P in einfachere Teile zerlegt werden kann, so dass

1. μ für die einfacheren Teile bekannt ist, und
2. die μ-Funktionen für die einfacheren Teile auf einfache Weise zur μ-Funktion von P kombiniert werden können.

Die Details würden den Rahmen dieser Vorlesung sprengen, das Prinzip, eine komplizierte Struktur (in diesem Fall P) in einfachere Strukturen zu zerlegen, Probleme für diese einfacheren Strukturen zu lösen und die Lösungen zu kombinieren, wird uns aber z.B. in der linearen Algebra noch oft begegnen.

Frage 5.3. Wie sieht die μ-Funktion für die Menge $[n]$ mit der üblichen Ordnung auf den natürlichen Zahlen aus?