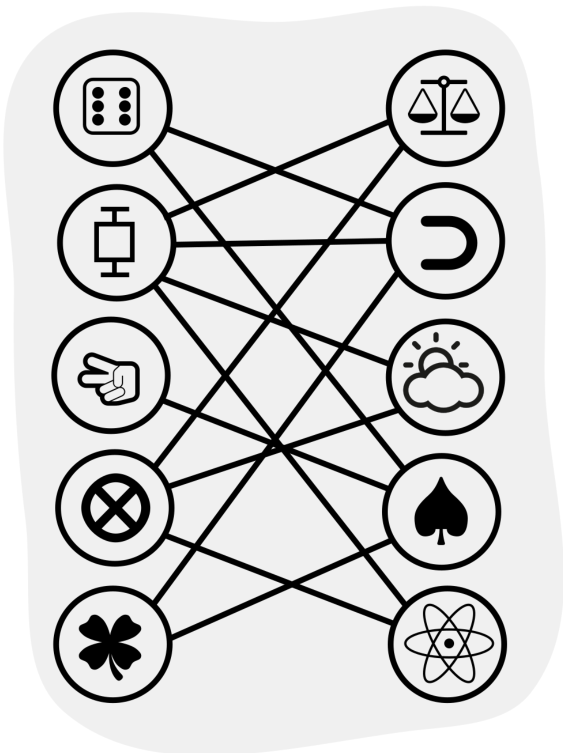
COUPLINGS AND MATCHINGS

On the equivalence of Strassen's theorem and some combinatorial theorems

V.T. KOPERBERG

Bachelor's thesis, November 2016

Supervisors: Dr. L. Avena and Dr. S. Taati



Mathematisch Instituut, Universiteit Leiden

CONTENTS

1 INTRODUCTION

Some mathematical theorems represent ideas that are so fundamental, that the same idea is discovered again and again in different forms. A sign that two different theorems might represent the same idea is that they are both a direct consequence of each other, i. e. there exist simple proofs of each of these theorems that use the other theorem. In the literature this is sometimes confusingly referred to as the 'equivalence' of two theorems.

One such fundamental idea can be found in different branches of combinatorics and optimisation theory. The best known formulation of this idea is *Hall's marriage theorem* (1935). In his original paper Hall [3, pp. 26,30] already mentioned the similarity between his theorem and a theorem by König from 1916. Since then numerous other theorems have also been found to closely resemble Hall's marriage theorem. These include *Menger's theorem* (1927), the *König minimax theorem* (1931), the *Birkhoff-von Neumann theorem* (1946), *Dilworth's theorem* (1950) and the *max-flow min-cut theorem* by Ford and Fulkerson (1956). For a more extensive discussion on the equivalence of these theorems the reader is referred to [7].

In this thesis we will argue that *Strassen's theorem* (1965) also belongs to this collection of similar theorems. For our argument we will consider a simplified version of Strassen's theorem. We will derive Hall's marriage theorem from this version of Strassen's theorem and show that this version of Strassen's theorem is a corollary of the max-flow min-cut theorem. These two derivations are sufficient to show the equivalence between Strassen's theorem and both the marriage theorem and the max-flow min-cut theorem, as the max-flow min-cut theorem can be derived from Hall's theorem [7, pp. 89, 96]. This is a remarkable result, since Strassen's theorem is a theorem from probability theory instead of combinatorics. As such the original proof made use of analytical tools instead of the combinatorial methods employed in the proofs of the other mentioned theorems [8].

In chapter 2 the simplified version of Strassen's theorem will be given. In chapter 3 we will investigate the similarities between Strassen's theorem and Hall's marriage theorem. With this goal in mind we will introduce the 'subforest lemma', which shall be proven by mimicking the proof of the marriage theorem. We show that this lemma is equivalent to the simplified version of Strassen's theorem and that we can deduce Hall's marriage theorem from this lemma. The subforest lemma and the results relating it to both Strassen's theorem and the marriage theorem are considered to be the innovative content of this thesis. In chapter 4 we will relate Strassen's theorem to network flow theory. Here we give a proof of the simplified version of Strassen's theorem that employs the max-flow min-cut theorem, showing that this version of Strassen's theorem can be seen as a direct consequence of this theorem from flow theory.

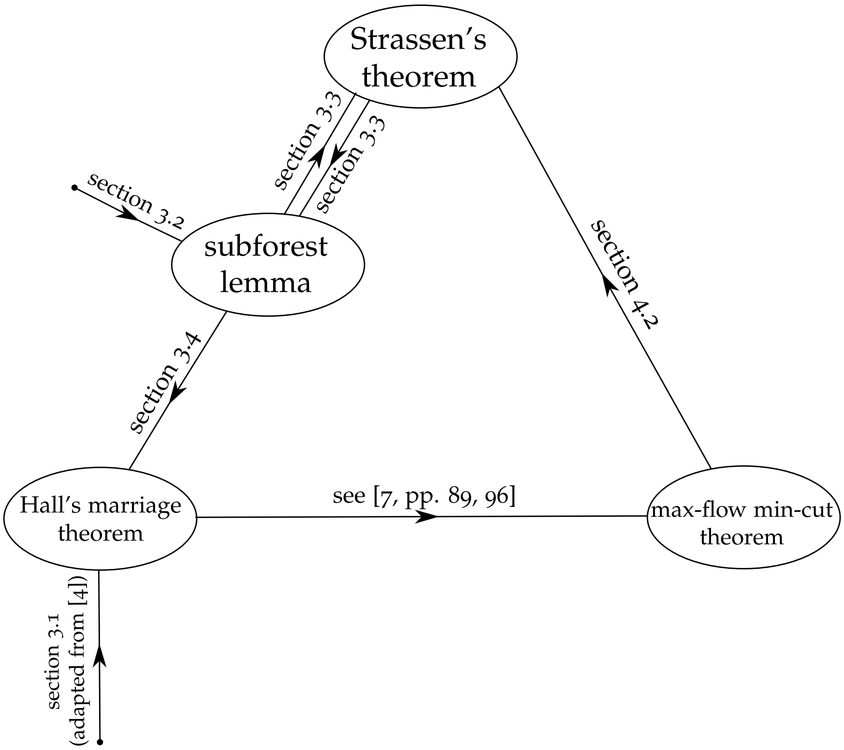


Figure 1: *An overview of the theorems covered in this thesis. The arrows indicate the proofs of these theorems. Arrows from one theorem to another represent proofs that utilise other theorems, thus demonstrating a relation between two theorems. The other arrows indicate independent proofs. The novelty of this thesis is represented by the subforest lemma and its relation with Strassen's theorem and Hall's marriage theorem.*

2 STRASSEN'S THEOREM

The main subject in this thesis will be Strassen's theorem. This theorem gives a necessary and sufficient condition for the existence of a coupling of probability measures that has some nice properties. We will first give a short example of such a coupling after which we will give the simplified version of Strassen's theorem that will be discussed in this thesis. A discussion of the original theorem can be found in [5].

2.1 COUPLINGS

Couplings are a proof technique used in probability theory, that can be used to compare different probability measures.

coupling

Definition 2.1. Let A and B be finite sets, $\mathbf{P}$ a probability measure on A and $\mathbf{P}'$ a probability measure on B. Then a *coupling* of $\mathbf{P}$ and $\mathbf{P}'$ is a probability measure $\hat{\mathbf{P}}$ on $A \times B$, such that the marginal distributions of $\hat{\mathbf{P}}$ correspond to $\mathbf{P}$ and $\mathbf{P}'$. That is, for all $U \subseteq A$ and $W \subseteq B$

$$\mathbf{P}(U) = \hat{\mathbf{P}}(U \times B) \text{ and } \mathbf{P}'(W) = \hat{\mathbf{P}}(A \times W).$$

Example 2.1. Consider two biased coins X and Y for which the probabilities of throwing heads are equal to p_X and p_Y with $p_X < p_Y$. If both coins are tossed independently it is possible that coin X shows heads while Y shows tails, even thought the probability that Y shows heads is larger. Through means of coupling it is possible to throw both coins dependently, such that coin Y will always show heads if X does so.

Identify heads with 1 and tails with 0. Let $A = B = \{1, 0\}$ and let $\mathbf{P}$ and $\mathbf{P}'$ be the laws of X and Y respectively. Define the probability measure $\hat{\mathbf{P}}$ on $A \times B$ by taking for all $(a, b) \in A \times B$

$$\hat{\mathbf{P}}(a,b) = \begin{cases} 1 - p_Y & \text{if } (a,b) = (0,0), \\ p_Y - p_X & \text{if } (a,b) = (0,1), \\ 0 & \text{if } (a,b) = (1,0), \\ p_X & \text{if } (a,b) = (1,1). \end{cases}$$

Then $\hat{\mathbf{P}}$ is a coupling of $\mathbf{P}$ and $\mathbf{P}'$ such that $\hat{\mathbf{P}}(\{(a,b) \in A \times B \colon a \leqslant b\}) = 1$.

This means that the random variable $(\hat{X}, \hat{Y})$, of which the law is $\hat{\mathbf{P}}$, satisfies $\hat{X} \leqslant \hat{Y}$. Also, $\hat{X}$ and $\hat{Y}$ can be considered dependently thrown coins of which the probabilities of throwing heads are p_X and p_Y respectively.

2.2 STRASSEN'S THEOREM FOR FINITE SETS

In this thesis, we will consider a simplified version of Strassen's theorem. For this simplified version we will use the following notation.

$\mathcal{N}_R$

Notation. Let A and B be sets and $R \subseteq A \times B$ a relation. Then for each $U \subseteq A$ the set of *neighbours* of U in R is denoted by

$$\mathcal{N}_R(U) = \{b \in B : (U \times \{b\}) \cap R \neq \varnothing\}.$$

Similarly, for any subset $W \subseteq B$ we let $\mathcal{N}_R(W)$ denote the set

$$\mathcal{N}_R(W) = \{a \in A : (\{a\} \times W) \cap R \neq \varnothing\}.$$

This notation is intentionally chosen to resemble the notation used for the graph-theoretical concept of neighbouring vertices, as will be given in definition 3.4. This is done to stress the similarity between the two concepts. This similarity will be clarified in section 3.2.

We will now state the main theorem of this thesis.

Theorem 2.1 (Strassen's theorem for finite sets)**.** *Let A and B be finite sets, $\mathbf{P}$ and $\mathbf{P}'$ probability measures on A and B respectively and $R \subseteq A \times B$ a relation. Then there exists a coupling $\hat{\mathbf{P}}$ of $\mathbf{P}$ and $\mathbf{P}'$ that satisfies $\hat{\mathbf{P}}(R) = 1$ if and only if*

$$\mathbf{P}(U) \leqslant \mathbf{P}'(\mathcal{N}_R(U)), \quad \textit{for all } U \subseteq A. \tag{1}$$

Strassen criterion

We will call condition (1) the *Strassen criterion*.

The setting in theorem 2.1 is completely symmetrical with respect to sets A and B. The Strassen criterion however, seems to break this symmetry. The following lemma shows that the symmetry is, in fact, still preserved.

Lemma 2.2 (Symmetry in Strassen's theorem for finite sets)**.** *The Strassen criterion*

$$\mathbf{P}(U) \leqslant \mathbf{P}'(\mathcal{N}_R(U)), \quad \textit{for all } U \subseteq A,$$

holds if and only if

$$\mathbf{P}'(W) \leqslant \mathbf{P}(\mathcal{N}_R(W)), \quad \textit{for all } W \subseteq B.$$

Proof. We will only prove the 'only if' part, since the 'if' part can be proven similarly. Assume that

$$\mathbf{P}(U) \leqslant \mathbf{P}'(\mathcal{N}_R(U)), \quad \text{for all } U \subseteq A.$$

Then for all $W \subseteq B$ we have that $\mathcal{N}_R(A \setminus \mathcal{N}_R(W)) \subseteq B \setminus W$, thus that

$$\begin{aligned}\mathbf{P}'(W) &= 1 - \mathbf{P}'(B \setminus W)\\ &\leqslant 1 - \mathbf{P}'(\mathcal{N}_R(A \setminus \mathcal{N}_R(W)))\\ &\leqslant 1 - \mathbf{P}(A \setminus \mathcal{N}_R(W))\\ &= \mathbf{P}(\mathcal{N}_R(W)).\end{aligned}$$

□

3 BIPARTITE GRAPHS

In this chapter we will show the similarity between Strassen's theorem and Hall's marriage theorem. For this we will use a lemma that we will call the 'subforest lemma'. We will first state and prove Hall's marriage theorem and then prove the subforest lemma in a manner similar to the proof of the marriage theorem. This alone indicates some connection between the marriage theorem and this lemma. We will continue by demonstrating the equivalence of Strassen's theorem for finite sets and the subforest lemma At the end of this chapter the subforest lemma will be used to derive Hall's marriage theorem.

It is important to stress that equivalences between theorems should not be confused with the formal notion of logical equivalence. Indeed, any two mathematical theorems are logically equivalent, since all theorems represent tautological statements. The equivalences discussed in this thesis represent an informal notion of similarity between two theorems. As mentioned in the introduction, this form of equivalence between two theorems is obtained when both theorems can be derived from each other.

3.1 HALL'S MARRIAGE THEOREM

Hall's marriage theorem is a theorem from graph theory. More specifically it covers matchings in bipartite graphs. Before we can give the theorem we will require some graph-theoretical definitions and notation. We have tried to keep our notation consistent with [1]. All graphs are assumed to be finite, undirected, simple graphs. That is, a graph is a pair $G = (V, E)$ of sets such that $E \subseteq [V]^2$. Here $[V]^2$ denotes the set of subsets of V containing two elements.

Definition 3.1. Let $G = (V, E)$ be a graph. If there exists a partitioning $\{A, B\}$ of the vertex set V, such that each edge in E has one endpoint in A and one endpoint in B then G is called a *bipartite* graph. The partition $\{A, B\}$ is called the *bipartition* of G. *bipartite graph*

Definition 3.2. Let $G = (V, E)$ be a graph. Then a subset of the edges $M \subseteq E$ is called a *matching*, if all endpoints of edges in M are distinct. *matching*

Definition 3.3. Let M be a matching and $U \subseteq V$ a subset of the vertices. Then M *saturates* U if each vertex in U is an endpoint of an edge in M. *saturating*

An example of the above definitions is given in figure 2.

neighbour $\mathcal{N}_G(U)$

Definition 3.4. Let $G = (V, E)$ be a graph. Two vertices $x, y \in V$ are called *neighbours* if $\{x, y\}$ is an edge in E. Let $U \subseteq V$. Then the set of neighbours in $V \setminus U$ of vertices in U is denoted by $\mathcal{N}_G(U)$.

subgraph

Definition 3.5. Let $G = (V, E)$ and $G' = (V', E')$ be graphs such that $V' \subseteq V$ and $E' \subseteq E$. Then G' is called a *subgraph* of G.

induced subgraph $G[V']$

Definition 3.6. Let $G = (V, E)$ be a graph and $G' = (V', E')$ a subgraph of G. If E' contains all edges in E with both endpoints in V', then G' is called an *induced subgraph* of G and denoted by $G' = G[V']$.

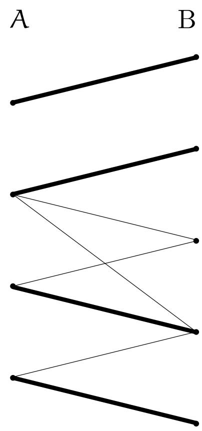


Figure 2: *An example of a bipartite graph with bipartition* $\{A, B\}$ *and a matching saturating* A*. The bold lines give the edges in the matching.*

Theorem 3.1 (Hall's marriage theorem)**.** *Let* G *be a bipartite graph with bipartition* $\{A, B\}$*. Then* G *contains a matching saturating* A *if and only if*

$$|U| \leqslant |\mathcal{N}_G(U)|, \quad \textit{for all } U \subseteq A. \tag{2}$$

marriage condition

Condition (2) is called the *marriage condition*.

The following proof is an adaptation of the proof of the marriage theorem by Halmos and Vaughan [4]. A graphical representation of this proof can be found in figure 3.

Proof of theorem 3.1. The necessity of the marriage condition is evident. Suppose that the marriage condition holds. We will prove the sufficiency by induction to $|A|$, the cardinality of A. For $|A| = 1$ a matching of A is given by taking any of the edges in E. Assume that $|A| \geqslant 2$ and that for each bipartite graph with bipartition $\{A', B'\}$ such that $|A'| < |A|$ the marriage condition is sufficient for the existence of a matching saturating A'. We will distinguish two cases.

(i) Suppose that there exists a non-empty strict subset $U \subsetneq A$ such that

$$|U| = |\mathcal{N}_G(U)|.$$

Then, by the induction hypothesis, the induced subgraph $G_1 = G[U \cup \mathcal{N}_G(U)]$ contains a matching M_1 saturating U. For each $W \subseteq A \setminus U$ we have that

$$\begin{aligned}|W| &= |W \cup U| - |U| \\ &\leqslant |\mathcal{N}_G(W \cup U)| - |\mathcal{N}_G(U)| \\ &= |\mathcal{N}_G(W) \setminus \mathcal{N}_G(U)|.\end{aligned}$$

This means that the induced subgraph $G_2 = G[(A \setminus U) \cup (B \setminus \mathcal{N}_G(U))]$ also satisfies the marriage condition. Therefore, by the induction hypothesis, there exists a matching M_2 in G_2 that saturates $A \setminus U$. Thus $M_1 \cup M_2$ is a matching in G that saturates A.

(ii) Suppose instead that for each non-empty strict subset $U \subsetneq A$ we have that $|U| < |\mathcal{N}_G(U)|$. Let $b \in B$ and let $a \in \mathcal{N}_G(\{b\})$ be a vertex neighbouring b. Then for each non-empty $W \subseteq A \setminus \{a\}$ we have that

$$\begin{aligned}|W| &\leqslant |\mathcal{N}_G(W)| - 1 \\ &\leqslant |\mathcal{N}_G(W) \setminus \{b\}|.\end{aligned}$$

It follows that the induced subgraph $G'[V \setminus \{a, b\}]$ satisfies the marriage condition. By the induction hypotheses, there exists a matching M in G' saturating $A \setminus a$. Taking $M \cup \{a, b\}$ gives a matching in G that saturates A.

□

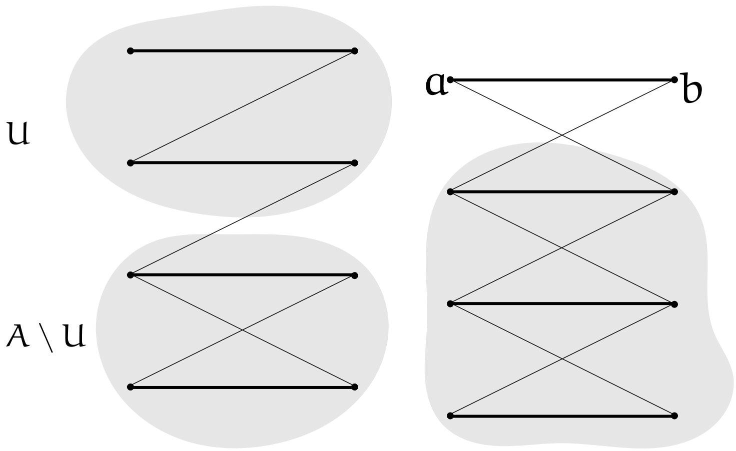
3.2 THE SUBFOREST LEMMA

The setting in Strassen's theorem naturally allows for the construction of the bipartite graph $G = (A \cup B, E)$ with vertices $A \cup B$ and edges $E = \{\{a, b\} \in [A \cup B]^2 \colon (a, b) \in R\}$. This allows for the usage of graph-theoretical concepts. Bipartite graphs generated in this manner from two finite sets A, B and a relation R between them, will be denoted by $G(A, B; R)$. This construction also shows the similarity between the two concepts of neighbour, since for $U \subseteq A$ we have that $\mathcal{N}_R(U) = \mathcal{N}_{G(A,B;R)}(U)$. $G(A, B; R)$

As the name suggest, the subforest lemma concerns a certain type of graph called a 'forest'. The definition of a forest can be given as follows.

Definition 3.7. A graph is called a *cycle* if it is connected and each vertex has exactly two neighbours. *cycle*

Definition 3.8. A graph is called a *forest* if it does not contain any cycles as subgraphs. *forest*



(a) *Construction of a matching in a graph that belongs to case (i).*

(b) *Construction of a matching in a graph that belongs to case (ii).*

Figure 3: *These illustrations clarify the proof of theorem 3.1 and show how in both cases a matching is constructed. The gray areas give the different subgraphs to which the induction hypothesis is applied. The bold lines indicate the edges in the matching.*

For convenience we will also use the following definition in the proof of the subforest lemma.

rescaled restriction $\mathbf{P}_U$

Definition 3.9. Let A be a finite set and $\mathbf{P}$ a probability measure on A. For a non-$\mathbf{P}$-null subset $U \subseteq A$ we can define the *rescaled restriction* of $\mathbf{P}$ to U as the probability measure $\mathbf{P}_U$ on U given by

$$\mathbf{P}_U(V) = \frac{\mathbf{P}(V)}{\mathbf{P}(U)}, \quad \text{for all } V \subseteq U.$$

We can now state and prove the subforest lemma, which will be valuable in showing the similarity between Strassen's theorem and the marriage theorem.

Lemma 3.2 (Subforest lemma). *Let A and B be finite sets, $\mathbf{P}$ and $\mathbf{P}'$ probability measures on A and B respectively. Let $R \subseteq A \times B$ be a relation such that*

$$\mathbf{P}(U) \leqslant \mathbf{P}'(\mathcal{N}_R(U)), \quad \textit{for all } U \subseteq A.$$

Then there exists a $T \subseteq R$ such that

(i) $\mathbf{P}(U) \leqslant \mathbf{P}'(\mathcal{N}_T(U))$ *for all* $U \subseteq A$,

(ii) *the graph* $G(A, B; T)$ *is a forest.*

Proof. We will use induction to $|A \cup B|$. For $|A \cup B| = 2$ the graph $G(A, B; R)$ is a forest, thus then the statement holds. Now assume that $|A \cup B| \geqslant 3$ and that the statement holds if $|A \cup B|$ is smaller. Note that if either $|A| = 1$ or $|B| = 1$, then the graph $G(A, B; R)$ is a forest. We can therefore assume that both $|A| \geqslant 2$ and $|B| \geqslant 2$.

Also note that we can assume for all $x \in A \cup B$ that either $\mathbf{P}(\{x\} \cap A) > 0$ or that $\mathbf{P}'(\{x\} \cap B) > 0$. Otherwise, there exists an $x \in A \cup B$ with $\mathbf{P}(\{x\} \cap A) = 0$ and $\mathbf{P}'(\{x\} \cap B) = 0$. By lemma 2.2, we can assume without loss of generality that $x \in A$. The relation $R \setminus (\{x\} \times B)$ satisfies the Strassen criterion with respect to the rescaled restriction $\mathbf{P}_{A \setminus \{x\}}$ and $\mathbf{P}'$. Since $|A \setminus \{x\} \cup B| < |A \cup B|$, by the induction hypothesis, there exists a $T \subseteq R \setminus (\{x\} \times B)$ satisfying the Strassen criterion with respect to $\mathbf{P}_{A \setminus \{x\}}$ and $\mathbf{P}'$. such that $G(A \setminus \{x\}, B; T)$ is a forest. Note that the graph $G(A, B; T)$ is also a forest and that T satisfies the Strassen criterion with respect to $\mathbf{P}$ and $\mathbf{P}'$, which would conclude the proof.

With these assumptions in mind we will distinguish two cases:

(i) Assume that there exists a non-empty strict subset $U \subsetneq A$ such that

$$\mathbf{P}(U) = \mathbf{P}'(\mathcal{N}_R(U))$$

or a non-empty strict subset $W \subsetneq B$ such that

$$\mathbf{P}'(W) = \mathbf{P}(\mathcal{N}_R(W)).$$

Without loss of generality, we can assume that there exists an $U \subset A$ such that

$$\mathbf{P}(U) = \mathbf{P}'(\mathcal{N}_R(U)).$$

Consider the relation $R \cap (U \times \mathcal{N}_R(U))$. Since R satisfies the Strassen criterion with respect to $\mathbf{P}$ and $\mathbf{P}'$, we have for all $V \subseteq U$ that

$$\begin{aligned}
\mathbf{P}_U(V) &= \frac{\mathbf{P}(V)}{\mathbf{P}(U)} \\
&\leqslant \frac{\mathbf{P}'(\mathcal{N}_R(V))}{\mathbf{P}(U)} \\
&= \frac{\mathbf{P}'(\mathcal{N}_R(V))}{\mathbf{P}'(\mathcal{N}_R(U))} \\
&= \mathbf{P}'_{\mathcal{N}_R(U)}(\mathcal{N}_R(V)) \\
&= \mathbf{P}'_{\mathcal{N}_R(U)}(\mathcal{N}_{R \cap (U \times \mathcal{N}_R(U))}(V)).
\end{aligned}$$

Thus $R \cap (U \times \mathcal{N}_R(U))$ satisfies the Strassen criterion with respect to $\mathbf{P}_U$ and $\mathbf{P}'_{\mathcal{N}_R(U)}$. By our induction hypothesis, there exists a $T_1 \subseteq R \cap (U \times \mathcal{N}_R(U))$ that satisfies the Strassen criterion with respect to $\mathbf{P}_U$ and $\mathbf{P}'_{\mathcal{N}_R(U)}$ such that $G(U, \mathcal{N}_R(U); T_1)$ is a forest. Note that for the complements $U^C = A \setminus U$ and $\mathcal{N}_R(U)^C = B \setminus \mathcal{N}_R(U)$ we also have that $\mathbf{P}(U^C) = \mathbf{P}'(\mathcal{N}_R(U)^C)$. Therefore, it follows similarly that there exists a relation $T_2 \subseteq R \cap \left(U^C \times \mathcal{N}_R(U)^C\right)$ that satisfies the Strassen criterion with respect to $\mathbf{P}_{U^C}$ and $\mathbf{P}'_{\mathcal{N}_R(U)^C}$ and such that $G(U^C, \mathcal{N}_R(U)^C; T_2)$ is a forest. Note that the relation $T_1 \cup T_2 \subseteq R$ satisfies the Strassen criterion with respect to $\mathbf{P}$ and $\mathbf{P}'$. Indeed, for each $V \subseteq A$

$$\begin{aligned}\mathbf{P}(V) &= \mathbf{P}(V \cap U) + \mathbf{P}(V \cap U^C) \\ &= \mathbf{P}_U(V \cap U)\mathbf{P}(U) + \mathbf{P}_{U^C}(V \cap U^C)\mathbf{P}(U^C) \\ &\leqslant \mathbf{P}'_{\mathcal{N}_R(U)}(\mathcal{N}_{T_1}(V \cap U))\mathbf{P}(U) + \\ &\quad \mathbf{P}'_{\mathcal{N}_R(U)^C}(\mathcal{N}_{T_2}(V \cap U^C))\mathbf{P}(U^C) \\ &= \frac{\mathbf{P}'(\mathcal{N}_{T_1}(V \cap U))}{\mathbf{P}'(\mathcal{N}_R(U))}\mathbf{P}(U) + \frac{\mathbf{P}'(\mathcal{N}_{T_2}(V \cap U^C))}{\mathbf{P}'(\mathcal{N}_R(U)^C)}\mathbf{P}(U^C) \\ &= \mathbf{P}'(\mathcal{N}_{T_1}(V \cap U)) + \mathbf{P}'(\mathcal{N}_{T_2}(V \cap U^C)) \\ &= \mathbf{P}'(\mathcal{N}_{T_1 \cup T_2}(V)).\end{aligned}$$

Furthermore, the graph $G(A, B; T_1 \cup T_2)$ is a forest. Thus the statement also holds in this case. Note that the forest $G(A, B; T_1 \cup T_2)$ contains at least two connected components, since none of the vertices in U is connected to a vertex in U^C.

(ii) Now assume that for each non-empty strict subset $U \subsetneq A$

$$\mathbf{P}(U) < \mathbf{P}'(\mathcal{N}_R(U))$$

and that for each non-empty strict subset $W \subsetneq B$

$$\mathbf{P}'(W) < \mathbf{P}(\mathcal{N}_R(W)).$$

The steps taken in this case are illustrated in figure 4. Let ε be the minimal mass of a singleton in A or B, i.e.

$$\varepsilon = \min_{x \in A \cup B} \max\{\mathbf{P}(A \cap \{x\}), \mathbf{P}'(B \cap \{x\})\}.$$

Let $x \in A \cup B$ be given such that $\max\{\mathbf{P}(A \cap \{x\}), \mathbf{P}'(B \cap \{x\})\} = \varepsilon$. By lemma 2.2, we can assume without loss of generality that $x \in A$. So, $\mathbf{P}(\{x\}) = \varepsilon$. Let $b \in \mathcal{N}_R(\{x\})$. Let $\mathcal{U} = \{U \subseteq A : x \notin U \text{ and } U \cap \mathcal{N}_R(b) \neq \varnothing\}$. Take

$$\delta = \min_{U \in \mathcal{U}} \mathbf{P}'(\mathcal{N}_R(U)) - \mathbf{P}(U).$$

Note that $A \setminus \{x\} \in \mathcal{U}$, therefore $\delta \leqslant \mathbf{P}(\{x\})$. Let $U \in \mathcal{U}$ be given such that $\mathbf{P}'(\mathcal{N}_R(U)) - \mathbf{P}(U) = \delta$. We can now add a new element $\tilde{x}$ to A to obtain the set $\tilde{A} = A \cup \{\tilde{x}\}$. Let $\tilde{R} = R \cup \{(\tilde{x}, b)\}$. Define the probability measure $\tilde{\mathbf{P}}$ on $\tilde{A}$ by

$$\tilde{\mathbf{P}}(V) = \begin{cases} \mathbf{P}(V \cap A) + \delta & \text{if } \tilde{x} \in V, x \notin V, \\ \mathbf{P}(V) - \delta & \text{if } \tilde{x} \notin V, x \in V, \\ \mathbf{P}(V \cap A) & \text{otherwise.} \end{cases}$$

Let $V \subseteq \tilde{A}$ be such that $\tilde{x} \in V$ and $x \notin V$. If $b \notin \mathcal{N}_{\tilde{R}}(V \setminus \{\tilde{x}\})$, then

$$\begin{aligned}\tilde{\mathbf{P}}(V) &= \mathbf{P}(V \setminus \{\tilde{x}\}) + \delta \\ &\leqslant \mathbf{P}'(\mathcal{N}_R(V \setminus \{\tilde{x}\})) + \delta \\ &= \mathbf{P}'(\mathcal{N}_{\tilde{R}}(V)) - \mathbf{P}'(\{b\}) + \delta \\ &\leqslant \mathbf{P}'(\mathcal{N}_{\tilde{R}}(V)).\end{aligned}$$

Else if $b \in \mathcal{N}_{\tilde{R}}(V \setminus \{\tilde{x}\})$, then we have that $V \setminus \{\tilde{x}\} \in \mathcal{U}$, thus that

$$\begin{aligned}\tilde{\mathbf{P}}(V) &= \mathbf{P}(V \setminus \{\tilde{x}\}) + \delta \\ &\leqslant \mathbf{P}'(\mathcal{N}_R(V \setminus \{\tilde{x}\})) \\ &= \mathbf{P}'(\mathcal{N}_{\tilde{R}}(V)).\end{aligned}$$

It follows that $\tilde{R}$ satisfies the Strassen criterion with respect to $\tilde{\mathbf{P}}$ and $\mathbf{P}'$. We also have that $\tilde{\mathbf{P}}(U \cup \{\tilde{x}\}) = \mathbf{P}'(\mathcal{N}_{\tilde{R}}(U \cup \{\tilde{x}\}))$.

If we have that $U \neq A \setminus \{x\}$, then we have that $|U \cup \{\tilde{x}\} \cup \mathcal{N}_{\tilde{R}}(U \cup \{\tilde{x}\})| < |A \cup B|$. It follows in the same manner as in case (i), that there exists a $\tilde{T} \subseteq \tilde{R}$ that satisfies the Strassen criterion with respect to $\tilde{\mathbf{P}}$ and $\mathbf{P}'$ and such that $G(\tilde{A}, B; \tilde{T})$ is a forest with x and b in distinct components. The relation $T = (\tilde{T} \cup \{(x, b)\}) \setminus \{(\tilde{x}, b)\}$ satisfies the Strassen criterion with respect to $\mathbf{P}$ and $\mathbf{P}'$ and $G(A, B; T)$ is a forest.

If instead $U = A \setminus \{x\}$, then we have that $\delta = \varepsilon$. This follows since in this case $\mathcal{N}_R(A \setminus \{x\}) = B$. Otherwise there would exist a $y \in B$ with $\mathcal{N}_R(\{y\}) \subseteq \{x\}$, which contradicts the assumptions that $\mathbf{P}'(\{y\}) \geqslant \varepsilon$ and that $\mathbf{P}'(\{y\}) < \mathbf{P}(\mathcal{N}_R(\{y\}))$. Define the probability measure $\tilde{\mathbf{P}}'$ on B as

$$\tilde{\mathbf{P}}' = \frac{\mathbf{P}' - \varepsilon\delta_b}{1 - \varepsilon},$$

where δ_b denotes the Dirac measure on b. The relation $R \setminus (\{x\} \times B)$ satisfies the Strassen criterion with respect to $\mathbf{P}_{A \setminus \{x\}}$ and $\tilde{\mathbf{P}}'$. This can be shown as follows. Since we have that $\mathbf{P}(\{x\}) = \delta$, for each $V \subseteq A \setminus \{x\}$ with $b \in \mathcal{N}_R(V)$

$$\begin{aligned}\mathbf{P}_{A \setminus \{x\}}(V) &= \frac{\mathbf{P}(V)}{1 - \delta} \\ &= \frac{\tilde{\mathbf{P}}(V \cup \{\tilde{x}\}) - \delta}{1 - \delta} \\ &\leqslant \frac{\mathbf{P}'(\mathcal{N}_{\tilde{R}}(V \cup \{\tilde{x}\})) - \delta}{1 - \delta} \\ &= \frac{\mathbf{P}'(\mathcal{N}_{\tilde{R}}(V)) - \delta}{1 - \delta} \\ &= \tilde{\mathbf{P}}'(\mathcal{N}_{\tilde{R}}(V)) \\ &= \tilde{\mathbf{P}}'(\mathcal{N}_{R \setminus (\{x\} \times B)}(V))\end{aligned}$$

For each $V \subseteq A \setminus \{x\}$ with $b \notin \mathcal{N}_R(V)$, we have

$$\begin{aligned}\mathbf{P}_{A\setminus\{x\}}(V) &= \frac{\mathbf{P}(V)}{1-\delta} \\ &\leqslant \frac{\mathbf{P}'(\mathcal{N}_{\tilde{R}}(V))}{1-\delta} \\ &= \tilde{\mathbf{P}}'(\mathcal{N}_{\tilde{R}}(V)) \\ &= \tilde{\mathbf{P}}'(\mathcal{N}_{R\setminus(\{x\}\times B)}(V)).\end{aligned}$$

Therefore, by the induction hypotheses, there exists a $T^* \subseteq R \setminus (\{x\} \times B)$ that satisfies the Strassen criterion with respect to $\mathbf{P}_{A\setminus\{x\}}$ and $\tilde{\mathbf{P}}'$ such that $G(A \setminus \{x\}, B; T^*)$ is a forest. It is easily seen that the graph $G(A, B; T^* \cup \{(x, b)\})$ is a forest. The relation $T^* \cup \{(x, b)\}$ satisfies the Strassen criterion with respect to $\mathbf{P}$ and $\mathbf{P}'$, since for each $V \subseteq A$ with $x \in V$ we have

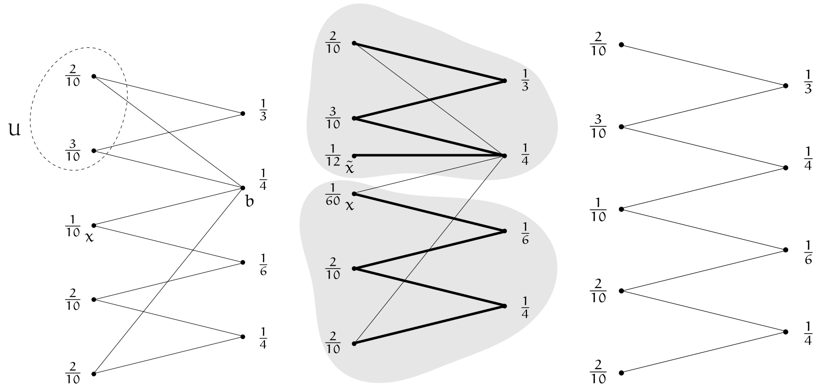
$$\begin{aligned}\mathbf{P}(V) &= \mathbf{P}(V \setminus \{x\}) + \delta \\ &= \mathbf{P}_{A\setminus\{x\}}(V \setminus \{x\})(1-\delta) + \delta \\ &\leqslant \tilde{\mathbf{P}}'(\mathcal{N}_{T^*}(V \setminus \{x\}))(1-\delta) + \delta \\ &= \tilde{\mathbf{P}}'(\mathcal{N}_{T}(V \setminus \{x\}))(1-\delta) + \delta \\ &\leqslant \tilde{\mathbf{P}}'(\mathcal{N}_{T}(V))(1-\delta) + \delta \\ &= \mathbf{P}'(\mathcal{N}_{T}(V)).\end{aligned}$$

For each $V \subseteq A$ with $x \notin V$, we have that

$$\begin{aligned}\mathbf{P}(V) &= \mathbf{P}_{A\setminus\{x\}}(V)(1-\delta) \\ &\leqslant \tilde{\mathbf{P}}'(\mathcal{N}_{T^*}(V))(1-\delta) \\ &= \tilde{\mathbf{P}}'(\mathcal{N}_{T}(V))(1-\delta) \\ &\leqslant \mathbf{P}'(\mathcal{N}_{T}(V)).\end{aligned}$$

In both cases we have shown the existence of a subrelation that induces a forest while preserving the Strassen condition, thus completing the proof. □

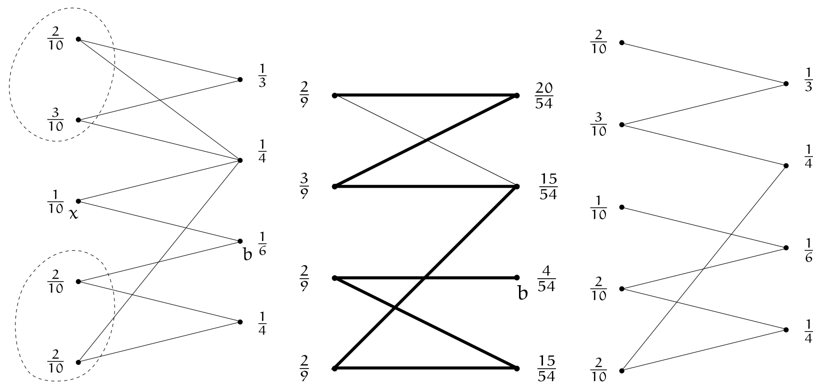
The similarities between this proof and the proof of theorem 3.1 are likely to need little explanation. In both proofs the same distinction is made between the case where there exists a 'critical subset', case (i), and the other case where no such subset exists. However, in the case that there does not exist a critical subset the proof of lemma 3.2 requires some more work than was required in the proof of theorem 3.1. In case (ii) of the proof of theorem 3.1 any vertex can be removed without violating the marriage condition, while in the proof of lemma 3.2 we then have to chose the vertex with minimal mass in order to make the proof work. A clarification of case (ii) of the proof of lemma 3.2 can be found in figure 4.



(a) *Determine the set* U, *which is indicated by the dotted circle.*

(b) *Split the mass of* x *between* x *and* $\tilde{x}$ *to create a critical subset. Then apply the induction hypothesis to the two subgraphs indicated by the gray areas.*

(c) *In the forests obtained by the induction hypothesis merge* x *and* $\tilde{x}$ *to obtain the final subforest.*



(d) *Determine the set* U. *In this case* U *includes all vertices in* A *except for* x.

(e) *Remove all mass from* x *and remove an equal amount of mass from* b. *Rescale the mass of all vertices and apply the induction hypothesis to the induced subgraph without vertex* x.

(f) *Add the edge* $\{x, b\}$ *to the forests obtained by the induction hypothesis to obtain the final subforest.*

Figure 4: Two examples of case (ii) of the proof of lemma 3.2 in section 3.2. Figures 4a to 4c give an example where $U \neq A \setminus \{x\}$ and figures 4d to 4f show an example where $U = A \setminus \{x\}$.

3.3 strassen's theorem and the subforest lemma

We are now equipped to give a proof of Strassen's theorem for finite sets. In this proof we will use the subforest lemma, thus deriving Strassen's theorem for finite sets from this lemma.

Proof of theorem 2.1 (via the subforest lemma). The necessity of the Strassen criterion can be easily demonstrated. Suppose that $\hat{\mathbf{P}}$ is a coupling of $\mathbf{P}$ and $\mathbf{P}'$ with $\hat{\mathbf{P}}(R) = 1$. Then for each $U \subseteq A$ we have that

$$\begin{aligned}\mathbf{P}(U) &= \hat{\mathbf{P}}(U \times B)\\ &= \hat{\mathbf{P}}(U \times \mathcal{N}_R(U))\\ &\leqslant \hat{\mathbf{P}}(A \times \mathcal{N}_R(U))\\ &= \mathbf{P}'(\mathcal{N}_R(U)).\end{aligned}$$

We will now prove the sufficiency of the Strassen criterion. Assume that the Strassen criterion holds. We will apply induction to $|R|$, the cardinality of R. Indeed if $|R| = 1$, then a coupling is given by assigning all mass to the one element of R. Assume that $|R| \geqslant 2$ and that a coupling with all mass on a relation exists for all relations with cardinality less than $|R|$ that satisfy the Strassen criterion.

By lemma 3.2, there exists a $T \subseteq R$ that satisfies the Strassen criterion with respect to $\mathbf{P}$ and $\mathbf{P}'$ such that $G(A, B; T)$ is a forest. It follows that there exists an $x \in A \cup B$ such that $|\mathcal{N}_T(\{x\})| = 1$. By lemma 2.2, we can assume without loss of generality that $x \in A$. Let $y \in \mathcal{N}_T(\{x\})$ be the unique neighbour of x. Note that it follows from the Strassen criterion that $\mathbf{P}(\{x\}) \leqslant \mathbf{P}'(\{y\})$. If $\mathbf{P}(\{x\}) = 1$, we also have that $\mathbf{P}'(\{y\}) = 1$ and a coupling $\hat{\mathbf{P}}$ that satisfies $\hat{\mathbf{P}}(R) = 1$ is given by $\hat{\mathbf{P}}(\{(x, y)\}) = 1$, which would complete the proof. Thus we can assume that $\mathbf{P}(\{x\}) < 1$. Take $\varepsilon = \mathbf{P}(\{x\})$ and define the probability measures $\tilde{\mathbf{P}}$ and $\tilde{\mathbf{P}}'$ on A and B respectively as

$$\tilde{\mathbf{P}} = \frac{\mathbf{P} - \varepsilon\delta_x}{1 - \varepsilon} \quad \text{and} \quad \tilde{\mathbf{P}}' = \frac{\mathbf{P}' - \varepsilon\delta_y}{1 - \varepsilon},$$

where δ_x and δ_y denote the Dirac measures on x and y respectively. The relation $T \setminus \{(x, y)\}$ satisfies the Strassen criterion with respect to $\tilde{\mathbf{P}}$ and $\tilde{\mathbf{P}}'$. This can be shown as follows. Note that $\tilde{\mathbf{P}}(\{x\}) = 0$. Let $U \subseteq A$ be a subset. If $y \in \mathcal{N}_{T\setminus\{(x,y)\}}(U)$, then we have that

$$\begin{aligned}\tilde{\mathbf{P}}(U) &= \tilde{\mathbf{P}}(U \cup \{x\})\\ &= \frac{\mathbf{P}(U \cup \{x\}) - \varepsilon}{1 - \varepsilon}\\ &\leqslant \frac{\mathbf{P}'(\mathcal{N}_T(U \cup \{x\})) - \varepsilon}{1 - \varepsilon}\\ &= \frac{\mathbf{P}'(\mathcal{N}_{T\setminus\{(x,y)\}}(U)) - \varepsilon}{1 - \varepsilon}\\ &= \tilde{\mathbf{P}}'(\mathcal{N}_{T\setminus\{(x,y)\}}(U)).\end{aligned}$$

Otherwise, if $y \notin \mathcal{N}_{T\setminus\{(x,y)\}}(U)$, then

$$\begin{aligned}\tilde{\mathbf{P}}(U) &= \tilde{\mathbf{P}}(U \setminus \{x\}) \\ &= \frac{\mathbf{P}(U \setminus \{x\})}{1-\varepsilon} \\ &\leqslant \frac{\mathbf{P}'(\mathcal{N}_T(U \setminus \{x\}))}{1-\varepsilon} \\ &= \frac{\mathbf{P}'(\mathcal{N}_{T\setminus\{(x,y)\}}(U))}{1-\varepsilon} \\ &= \tilde{\mathbf{P}}'(\mathcal{N}_{T\setminus\{(x,y)\}}(U)).\end{aligned}$$

In both cases the Strassen criterion holds. Since we have $|T \setminus \{(x,y)\}| < |R|$, by the induction hypotheses, there exists a coupling $\hat{\mathbf{P}}^*$ of $\tilde{\mathbf{P}}$ and $\tilde{\mathbf{P}}'$ such that $\hat{\mathbf{P}}^*(T \setminus \{(x,y)\}) = 1$.

We will now show that the probability measure $\hat{\mathbf{P}}$ on $A \times B$ defined as

$$\hat{\mathbf{P}} = (1-\varepsilon)\hat{\mathbf{P}}^* + \varepsilon\delta_{(x,y)}$$

is a coupling of $\mathbf{P}$ and $\mathbf{P}'$ such that $\hat{\mathbf{P}}(T) = 1$. Here $\delta_{(x,y)}$ denotes the Dirac measure on (x,y). Since it is clear that $\hat{\mathbf{P}}$ satisfies $\hat{\mathbf{P}}(T) = 1$, it only remains to show that $\hat{\mathbf{P}}$ is a coupling of $\mathbf{P}$ and $\mathbf{P}'$. Let $U \subseteq A$. If $x \in U$ we have that

$$\begin{aligned}\hat{\mathbf{P}}(U \times B) &= (1-\varepsilon)\hat{\mathbf{P}}^*(U \times B) + \varepsilon\delta_{(x,y)}(U \times B) \\ &= (1-\varepsilon)\hat{\mathbf{P}}^*(U \times B) + \varepsilon \\ &= (1-\varepsilon)\tilde{\mathbf{P}}(U) + \varepsilon \\ &= \mathbf{P}(U).\end{aligned}$$

Otherwise, if $x \notin U$, then we have that

$$\begin{aligned}\hat{\mathbf{P}}(U \times B) &= (1-\varepsilon)\hat{\mathbf{P}}^*(U \times B) + \varepsilon\delta_{(x,y)}(U \times B) \\ &= (1-\varepsilon)\hat{\mathbf{P}}^*(U \times B) \\ &= (1-\varepsilon)\tilde{\mathbf{P}}(U) \\ &= \mathbf{P}(U).\end{aligned}$$

Now let $W \subseteq B$. If $y \in W$, then we have that

$$\begin{aligned}\hat{\mathbf{P}}(A \times W) &= (1-\varepsilon)\hat{\mathbf{P}}^*(A \times W) + \varepsilon\delta_{(x,y)}(A \times W) \\ &= (1-\varepsilon)\hat{\mathbf{P}}^*(A \times W) + \varepsilon \\ &= (1-\varepsilon)\tilde{\mathbf{P}}'(W) + \varepsilon \\ &= \mathbf{P}'(W).\end{aligned}$$

Otherwise, if $y \notin W$ we have that

$$\begin{aligned}\hat{\mathbf{P}}(A \times W) &= (1-\varepsilon)\hat{\mathbf{P}}^*(A \times W) + \varepsilon\delta_{(x,y)}(A \times W) \\ &= (1-\varepsilon)\hat{\mathbf{P}}^*(A \times W) \\ &= (1-\varepsilon)\tilde{\mathbf{P}}'(W) \\ &= \mathbf{P}'(W).\end{aligned}$$

This shows that $\hat{\mathbf{P}}$ is indeed a coupling of $\mathbf{P}$ and $\mathbf{P}'$.

Since T is a subset of R, the probability measure $\hat{\mathbf{P}}$ also satisfies $\hat{\mathbf{P}}(R) = 1$. Thus $\hat{\mathbf{P}}$ is the sought coupling of $\mathbf{P}$ and $\mathbf{P}'$. Therefore, by induction a coupling $\hat{\mathbf{P}}$ with $\hat{\mathbf{P}}(R) = 1$ exists for all relations R and probability measures $\mathbf{P}$ and $\mathbf{P}'$ that satisfy the Strassen criterion. □

The subforest lemma can also be easily deduced from Strassen's theorem, thus showing that the two statements are equivalent.

Proof of lemma 3.2 (via Strassen's theorem). The statement will be proven by induction to $|R|$. If $|R| = 1$, then the graph $G(A, B; R)$ is a forest. Now assume that $|R| \geqslant 2$ and that the statement holds if $|R|$ is smaller.

If $G(A, B; R)$ is not a forest, then it contains a subgraph $C = (V_C, E_C)$ that is a cycle. Let $R_C = \{(a, b) \in R \colon \{a, b\} \in E_C\}$. By theorem 2.1, there exists a coupling $\hat{\mathbf{P}}_1$ of $\mathbf{P}$ and $\mathbf{P}'$ such that $\hat{\mathbf{P}}_1(R) = 1$. Take

$$\varepsilon = \min_{(a,b)\in R_C} \hat{\mathbf{P}}_1(\{(a, b)\})$$

and let $(x, y) \in R_C$ be given such that $\hat{\mathbf{P}}_1(\{(x, y)\}) = \varepsilon$. Since C is a bipartite cycle, we have that $|E_C|$ is even. Therefore, we can partition the edges in E_C into two equally sized sets $\{I_C, J_C\}$ such that each edge in I_C is only adjacent to edges in J_C and vice-versa, each edge in J_C is only adjacent to edges in I_C. Without loss of generality we can assume that $\{x, y\} \in I_C$. Define the probability measure $\hat{\mathbf{P}}_2$ on $A \times B$ by taking for all $(a, b) \in A \times B$

$$\hat{\mathbf{P}}_2(\{(a, b)\}) = \begin{cases} \hat{\mathbf{P}}_1(\{(a, b)\}) - \varepsilon & \text{if } \{a, b\} \in I_C, \\ \hat{\mathbf{P}}_1(\{(a, b)\}) + \varepsilon & \text{if } \{a, b\} \in J_C, \\ \hat{\mathbf{P}}_1(\{(a, b)\}) & \text{otherwise.} \end{cases}$$

Note that each vertex in $G(A, B; R)$ is incident to as many edges in I_C as to edges in J_C. It follows that $\hat{\mathbf{P}}_2(\{a\} \times B) = \hat{\mathbf{P}}_1(\{a\} \times B)$ for each $a \in A$. Therefore, we have for each $U \subseteq A$ that

$$\begin{aligned} \hat{\mathbf{P}}_2(U \times B) &= \hat{\mathbf{P}}_1(U \times B) \\ &= \mathbf{P}(U). \end{aligned}$$

It follows identically that for each $W \subseteq B$ we have $\hat{\mathbf{P}}_2(A \times W) = \mathbf{P}'(W)$. Thus $\hat{\mathbf{P}}_2$ is a coupling of $\mathbf{P}$ and $\mathbf{P}'$. Note that $\hat{\mathbf{P}}_2(\{(x, y)\}) = 0$, thus that $\hat{\mathbf{P}}_2(R \setminus \{(x, y)\}) = 1$. By theorem 2.1 this means that the relation $R \setminus \{(x, y)\}$ satisfies the Strassen criterion with respect to $\mathbf{P}$ and $\mathbf{P}'$. By the induction hypothesis, there exists a $T \subseteq R \setminus \{(x, y)\}$ that satisfies the Strassen criterion with respect to $\mathbf{P}$ and $\mathbf{P}'$ such that $G(A, B; T)$ is a forest. As we have that $T \subseteq R$, this completes the proof. □

3.4 HALL'S MARRIAGE THEOREM AND THE SUBFOREST LEMMA

A slightly different version of the marriage theorem can be derived from the subforest lemma in a similar manner to the way in which we derived Strassen's theorem from this lemma.

Theorem 3.3. *Let* $G = (A \cup B, E)$ *be a bipartite graph with bipartition* $\{A, B\}$. *Then* G *contains a matching saturating* $A \cup B$ *if and only if*

(i) $|A| = |B|$,

(ii) *the marriage condition holds.*

Proof (via the subforest lemma). Assume that $|A| = |B|$ and that the marriage condition is satisfied. We will apply induction to $|E|$ to show that G contains a matching saturating $A \cup B$. Clearly the statement holds for $|E| = 1$. Now assume that $|E| \geqslant 2$ and that the statement holds if $|E|$ is smaller.

Let $\mathbf{P}$ and $\mathbf{P}'$ be the uniform distributions on A and B respectively. Let R be the relation induced by E, i. e. $R = \{(a, b) \in A \times B : \{a, b\} \in E\}$. Note that we now have that $G(A, B; R) = G$. Since G satisfies the marriage condition, it follows that R satisfies the Strassen criterion with respect to $\mathbf{P}$ and $\mathbf{P}'$. Therefore, by lemma 3.2, there exists a $T \subseteq R$ satisfying the Strassen condition with respect to $\mathbf{P}$ and $\mathbf{P}'$ such that $G(A, B; T)$ is a forest. Note that $G(A, B; T)$ also satisfies the marriage condition. Since $G(A, B; T)$ is a forest, there exists a vertex $x \in A \cup B$ such that $|\mathcal{N}_{G(A,B;T)}(\{x\})| = 1$. By symmetry we can assume without loss of generality that $x \in A$. Let $y \in \mathcal{N}_{G(A,B;T)}(\{x\})$ be the unique neighbour of x. The induced subgraph $G' = G(A, B; T)[A \cup B \setminus \{x, y\}]$ still satisfies the marriage condition, since for each $U \subseteq A \setminus \{x\}$ we have that

$$\begin{aligned} |U| &= |U \cup \{x\}| - 1 \\ &\leqslant |\mathcal{N}_{G(A,B;T)}(U \cup \{x\})| - 1 \\ &= |\mathcal{N}_{G(A,B;T)}(U) \cup \{y\}| - 1 \\ &\leqslant |\mathcal{N}_{G'}(U)|. \end{aligned}$$

Thus by the induction hypothesis G' contains a matching M saturating $A \cup B \setminus \{x, y\}$. Taking $M \cup \{x, y\}$ gives a matching in G saturating $A \cup B$. □

As the subforest lemma in Strassen's theorem for finite sets are equivalent statements, this proof shows that Hall's marriage theorem is a consequence of Strassen's theorem.

4 NETWORKS

In this chapter we will discuss how Strassen's theorem relates to the max-flow min-cut theorem. Lovász and Plummer [6, pp. 45-46] mention a combinatorial version of Strassen's theorem for finite sets (there called the *supply-demand theorem*) as a corollary of the max-flow min-cut theorem. We will give another proof of Strassen's theorem for finite sets, showing that Strassen's theorem is indeed a corollary of the max-flow min-cut theorem.

4.1 THE MAX-FLOW MIN-CUT THEOREM

Network flow theory is a powerful tool that can be used in many proofs in matching theory. The canonical result in this theory is the max-flow min-cut theorem by Ford and Fulkerson [2] and many result in matching theory can be derived from this theorem. We will first give some necessery concepts used in network flow theory before we state the max-flow min-cut theorem.

directed edges

Definition 4.1. Let $G = (V, E)$ be a graph. We can define the set of *directed edges* of G as the set of ordered pairs

$$\vec{E} = \{(x, y) \in V^2 \colon \{x, y\} \in E\}.$$

The directed edge (x, y) is directed from vertex x to vertex y and the directed edge (y, x) is directed from y to x.

capacity function

Definition 4.2. Let G be a graph. A function $c : \vec{E} \to [0, \infty)$ on the directed edges of G is called a *capacity function*.

For $\vec{F} \subseteq \vec{E}$ it is convenient to write

$$c(\vec{F}) = \sum_{\vec{e} \in \vec{F}} c(\vec{e}).$$

network
source
sink

Definition 4.3. A *network* N is a tuple $N = (G, c, s, t)$, where $G = (V, E)$ is a graph, c is a capacity function and $s, t \in V$ are vertices. The vertex s is called the *source* and the vertex t is called the *sink*.

cut

Definition 4.4. Let $N = (G, c, s, t)$ be a network with $G = (V, E)$. Let $\{S, T\}$ be a partition of V such that $s \in S$ and $t \in T$. Then the set $\vec{E}(S, T)$ of directed edges in $\vec{E}$ directed from a vertex in S to a vertex in T is called a *cut* in N.

flow

Definition 4.5. Let $N = (G, c, s, t)$ be a network with $G = (V, E)$. A function $f : \vec{E} \to \mathbb{R}$ is called a *flow* in N if it satisfies the following three properties:

(i) $f(x,y) = -f(y,x)$ for all $(x,y) \in \vec{E}$,

(ii) $\sum_{y \in \mathcal{N}_G(\{x\})} f(x,y) = 0$ for all $x \in V \setminus \{s,t\}$,

(iii) $f(\vec{e}) \leqslant c(\vec{e})$ for all $\vec{e} \in \vec{E}$.

Again for $\vec{F} \subseteq \vec{E}$ we will write

$$f(\vec{F}) = \sum_{\vec{e} \in \vec{F}} f(\vec{e}).$$

Definition 4.6. Let $N = (G, c, s, t)$ be a network and f a flow in N. Then the value $f(\vec{E}(\{s\}, V \setminus \{s\}))$ is called the *total value* of f and denoted by $|f|$.

total value
$|f|$

Note that for each cut $\vec{E}(S,T)$ the total value of a flow f equals $f(\vec{E}(S,T))$. This follows directly from flow properties (i) and (ii), since

$$\begin{aligned} f(\vec{E}(S,T)) &= f((S \times V) \cap \vec{E}) - f((S \times S) \cap \vec{E}) \\ &= f((\{s\} \times V) \cap \vec{E}) + f(((S \setminus \{s\}) \times V) \cap \vec{E}) \\ &= f(\vec{E}(\{s\}, V \setminus \{s\})). \end{aligned}$$

This immediately gives for all flows f and all cuts $\vec{E}(S,T)$ that

$$|f| \leqslant c(\vec{E}(S,T)). \tag{3}$$

The max-flow min-cut theorem asserts that in each network there exist a flow and a cut for which equation (3) becomes an equality.

Theorem 4.1 (Max-flow min-cut theorem). *In each network, the maximal value of a flow is equal to the minimal capacity of a cut.*

A proof of the max-flow min-cut theorem can be found in [6, p. 45].

4.2 STRASSEN'S THEOREM AND THE MAX-FLOW MIN-CUT THEOREM

In section 3.2 we have seen that the setting in theorem 2.1 can be used to construct an associated bipartite graph. We can also use the setting in Strassen's theorem to construct an associated network. Let A, B be finite sets with probability measures $\mathbf{P}$ and $\mathbf{P}'$ and a relation $R \subseteq A \times B$. We can extend the associated bipartite graph $G(A, B; R)$ by adding two vertices, a source s and a sink t. Edges are drawn between s and each vertex in A and between each vertex in B and t. That is, the graph $G = (V, E)$ is given by

$$V = A \cup B \cup \{s,t\}$$

and

$$E = \{\{a,b\} \in [V]^2 \colon (a,b) \in R\} \cup \{\{s,a\} \in [V]^2 \colon a \in A\} \cup \{\{b,t\} \in [V]^2 \colon b \in B\}.$$

Define the capacity function $c : \vec{E} \to [0,\infty)$ as

$$c(x,y) = \begin{cases} 1 & \text{if } x \in A, y \in B, \\ \mathbf{P}(\{y\}) & \text{if } x = s, \\ \mathbf{P}'(\{x\}) & \text{if } y = t, \\ 0 & \text{otherwise.} \end{cases}$$

$N(A,B;\mathbf{P},\mathbf{P}';R)$ The network $N = (G,c,s,t)$ obtained in this manner is denoted by $N(A,B;\mathbf{P},\mathbf{P}';R)$. See figure 5 for an example of an associated network.

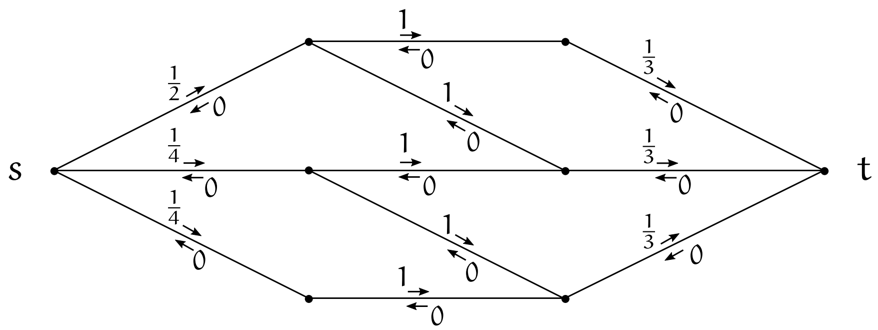


Figure 5: An example of an associated network. The numbers next to the arrows indicate the capacity of each directed edge.

The proof of theorem 2.1 will also use the following two lemmas.

Lemma 4.2. *The network* $N(A,B;\mathbf{P},\mathbf{P}';R) = (G,c,s,t)$ *admits a cut* $\vec{E}(S,T)$ *with capacity* $c(\vec{E}(S,T)) < 1$ *if and only if there exists a* $U \subseteq A$ *such that* $\mathbf{P}(U) > \mathbf{P}(\mathcal{N}_R(U))$.

Proof. Suppose that there exists a $U \subseteq A$ such that $\mathbf{P}(U) > \mathbf{P}(\mathcal{N}_R(U))$. Then by lemma 2.2 there exists a $W \subseteq B$ such that $\mathbf{P}'(W) > \mathbf{P}(\mathcal{N}_R(W))$. Take

$$T = \{t\} \cup W \cup \mathcal{N}_R(W)$$

and

$$S = \{s\} \cup \mathcal{N}_R(W)^C \cup W^C.$$

Then the cut $\vec{E}(S,T)$ consists of the directed edges (s,a) with $a \in \mathcal{N}_R(W)$, the directed edges (b,t) with $b \in W^C$ and the directed edges (b,a) with $a \in \mathcal{N}_R(W)$ and $b \in W^C$. Note that for each directed edge

(b, a) with $a \in \mathcal{N}_R(W)$ and $b \in W^C$ we have that $c(b, a) = 0$. It follows that for the capacity of this cut we have

$$\begin{aligned} c(\vec{E}(S,T)) &= \mathbf{P}(\mathcal{N}_R(W)) + \mathbf{P}'(W^C) \\ &= \mathbf{P}(\mathcal{N}_R(W)) + 1 - \mathbf{P}'(W) \\ &< 1. \end{aligned}$$

Suppose instead that there exists a cut $\vec{E}(S,T)$ with $c(\vec{E}(S,T)) < 1$. Take $U = A \cap S$. Note that $\mathcal{N}_R(U) \subseteq S$, since otherwise the cut would contain a directed edge with capacity 1. This means that

$$\begin{aligned} \mathbf{P}(U^C) + \mathbf{P}'(\mathcal{N}_R(U)) &\leqslant c(\vec{E}(S,T)) \\ &< 1. \end{aligned}$$

It follows that

$$\mathbf{P}(U) > \mathbf{P}'(\mathcal{N}_R(U)).$$

□

Lemma 4.3. *There exists a coupling* $\hat{\mathbf{P}}$ *of* $\mathbf{P}$ *and* $\mathbf{P}'$ *with* $\hat{\mathbf{P}}(R) = 1$ *if and only if the the maximum total value of a flow in* $N(A, B; \mathbf{P}, \mathbf{P}'; R)$ *equals* 1.

Proof. Suppose that $\hat{\mathbf{P}}$ is such a coupling. Then we can define the function $f : \vec{E} \to \mathbb{R}$ as follows

$$f(x,y) = \begin{cases} \hat{\mathbf{P}}(\{(x,y)\}) & \text{if } x \in A, y \in B, \\ -\hat{\mathbf{P}}(\{(x,y)\}) & \text{if } y \in A, x \in B, \\ \mathbf{P}(\{y\}) & \text{if } x = s, \\ -\mathbf{P}(\{x\}) & \text{if } y = s, \\ \mathbf{P}'(\{x\}) & \text{if } y = t, \\ -\mathbf{P}'(\{y\}) & \text{if } x = t. \end{cases}$$

It is immediately clear from the definition that f satisfies flow properties (i) and (iii). Property (ii) holds since $\hat{\mathbf{P}}$ is a coupling. Indeed we have for $a \in A$ that

$$\begin{aligned} \sum_{y \in \mathcal{N}_G(\{a\})} f(a,y) &= -\mathbf{P}(\{a\}) + \sum_{y \in \mathcal{N}_G(\{a\}) \cap B} \hat{\mathbf{P}}(\{(a,y)\}) \\ &= -\mathbf{P}(\{a\}) + \hat{\mathbf{P}}(\{a\} \times B) \\ &= 0 \end{aligned}$$

and for $b \in B$ that

$$\begin{aligned} \sum_{y \in \mathcal{N}_G(\{b\})} f(b,y) &= \mathbf{P}'(\{b\}) - \sum_{y \in \mathcal{N}_G(\{b\}) \cap A} \hat{\mathbf{P}}(\{(b,y)\}) \\ &= \mathbf{P}'(\{b\}) - \hat{\mathbf{P}}(A \times \{b\}) \\ &= 0. \end{aligned}$$

It follows that f is a flow in $N(A, B; \mathbf{P}, \mathbf{P}'; R)$. It is clear that the network N does not admit a flow with total value larger than 1. Thus, since $|f| = 1$, the maximal value of a flow in N equals 1.

Instead suppose that f is a flow in $N(A, B; \mathbf{P}, \mathbf{P}'; R)$ with $|f| = 1$. Note that $f(s, a) = \mathbf{P}(\{a\})$ for all $a \in A$ and that $f(b, t) = \mathbf{P}'(\{b\})$ for all $b \in B$. Define the measure $\hat{\mathbf{P}}$ on $A \times B$ by

$$\hat{\mathbf{P}}(U \times W) = \sum_{(a,b)\in(U\times W)\cap R} f(a, b), \text{ for all } U \subseteq A, W \subseteq B.$$

Then for each $U \subseteq A$ we have

$$\begin{aligned}\hat{\mathbf{P}}(U \times B) &= \sum_{(a,b)\in(U\times B)\cap R} f(a, b) \\ &= \sum_{a\in U} \sum_{b\in \mathcal{N}_R(\{a\})} f(a, b) \\ &= \sum_{a\in U} \left(-f(a, s) + \sum_{b\in \mathcal{N}_G(\{a\})} f(a, b) \right) \\ &= \sum_{a\in U} f(s, a) \\ &= \mathbf{P}(U).\end{aligned}$$

Similarly, for each $W \subseteq B$ we have $\hat{\mathbf{P}}(A \times W) = \mathbf{P}'(W)$. Thus $\hat{\mathbf{P}}$ is a coupling of $\mathbf{P}$ and $\mathbf{P}'$. Since $E(\{s\} \cup A, \{t\} \cup B)$ is a cut in N, we also have that

$$\begin{aligned}\hat{\mathbf{P}}(R) &= \sum_{(a,b)\in R} f(a, b) \\ &= f(E(\{s\} \cup A, \{t\} \cup B)) \\ &= |f| \\ &= 1.\end{aligned}$$

Therefore, $\hat{\mathbf{P}}$ is the sought coupling. □

We can now give a one-line proof of Strassen's theorem for finite sets.

Proof of theorem 2.1 (via the max-flow min-cut theorem). The statement follows directly from lemma 4.2, lemma 4.3 and theorem 4.1. □

This proof shows that Strassen's theorem for finite sets is a direct result of the max-flow min-cut theorem. Furthermore, it shows how to construct the desired coupling $\hat{\mathbf{P}}$ from a maximal flow in the network $N(A, B; \mathbf{P}, \mathbf{P}'; R)$. This proof therefore reduces the problem of finding $\hat{\mathbf{P}}$ to the problem of finding a maximal flow in $N(A, B; \mathbf{P}, \mathbf{P}'; R)$. Fortunately, there exist numerous tools for finding maximal network flows, e.g. the Edmonds-Karp algorithm [6, pp. 48-49]. These tools can also be applied to find the coupling in Strassen's theorem for finite sets.

4.3 HALL'S MARRIAGE THEOREM AND THE MAX-FLOW MIN-CUT THEOREM

In [7, pp. 89, 96] the max-flow min-cut theorem is derived from Hall's marriage theorem. Together with our derivations of the marriage theorem from Strassen's theorem and of Strassen's theorem from the max-flow min-cut theorem seen in chapter 3 and section 4.2, this shows the equivalence of these three theorems.

The derivation of the max-flow min-cut theorem in [7] makes use of a version of the marriage theorem where each vertex is assigned an integer-valued weight. This weighted version resembles a graph-theoretical restatement of Strassen's theorem for finite sets where only rational-valued probability measures are allowed. Hence this derivation might indicate that it is also possible to derive the max-flow min-cut theorem from Strassen's theorem.

5 CONCLUSION

The simplified version of Strassen's theorem that was discussed in this thesis, theorem 2.1, is equivalent to the class of equivalent combinatorial theorems that includes Hall's marriage theorem and the max-flow min cut theorem. In chapter 3 Hall's marriage theorem was derived from the subforest lemma, which is equivalent to Strassen's theorem. In chapter 4 a proof of Strassen's theorem was given, that shows that Strassen's theorem is a corollary of the max-flow min-cut theorem. Since the max-flow min-cut theorem can be derived from the marriage theorem, the three theorems are equivalent.

Incidently, the derivation of Strassen's theorem from the max-flow min-cut theorem also gives a method for the construction of a coupling that satisfies the requirements given in Strassen's theorem. The proof shows that any method of finding maximal network flows can also be used to find such a coupling.

BIBLIOGRAPHY

[1] Reinhard Diestel. *Graph Theory*. Graduate Texts in Mathematics. Springer-Verlag, 2010.

[2] Lester R. Ford and Delbert R. Fulkerson. 'Maximal flow through a network'. In: *Canadian J. Math.* 8.3 (1956), pp. 399–404.

[3] Philip Hall. 'On representatives of subsets'. In: *J. London Math. Soc.* 10.1 (1935), pp. 26–30.

[4] Paul R. Halmos and Herbert E. Vaughan. 'The Marriage Problem'. In: *Am. J. Math.* 72.1 (1950), pp. 214–215.

[5] Torgny Lindvall. 'On Strassen's theorem on stochastic domination'. In: *Electron. Commun. Prob.* 4 (1999), pp. 51–59.

[6] Lázló Lovász and Michael D. Plummer. *Matching Theory*. North-Holland Mathematics Studies. Elsevier Science, 1986.

[7] Philip F. Reichmeider. *The equivalence of some combinatorial matching theorems*. Polygonal Pub. House, 1984.

[8] Volker Strassen. 'The existence of probability measures with given marginals'. In: *Ann. Math. Stat.* 36 (1965), pp. 423–439.