Graph Theory *

Yaokun Wu

Department of Mathematics, Shanghai Jiao Tong University,

Shanghai, 200240, P.R. China

Email: ykwu@sjtu.edu.cn

Fall Semester, 2004

*Slides used in class for the course Graph Theory (ACM class, SJTU)

Don't walk in front of me; I may not follow. Don't walk behind me; I may not lead. Just walk beside me and be my friend. – Albert Camus (1913 – 1960)

1 Polyhedra and Farkas' Lemma

Recall:

Theorem 1.1. *Let $A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m$. $Ax = b$ has a solution $x \in \mathbb{R}^n$ if and only if there is no solution $y \in \mathbb{R}^m$ for the following system of linear relations: $y^\top A = 0, y^\top b \neq 0$.*

Proof.

$$\begin{aligned} Ax = b \Leftrightarrow & \quad Span(A, b) = Span(A) \\ \Leftrightarrow & \quad (Span(A, b))^\perp = (Span(A))^\perp \\ \Leftrightarrow & \quad (Span(A))^\perp \setminus (Span(A, b))^\perp = \emptyset \\ \Leftrightarrow & \quad \text{there is no } y \text{ with } y^\top A = 0, y^\top b \neq 0 \end{aligned}$$

□

Another proof: Use Gaussian elimination. [1]

Exercise 1.2. Let $A \in \mathbb{Q}^{m \times n}, b \in \mathbb{R}^m$. $Ax = b$ has a solution $x \in \mathbb{R}^n$ if and only if there is no solution $y \in \mathbb{Q}^m$ for the following system of linear relations: $y^\top A = 0, y^\top b \neq 0$.

[1] Fourier-Motzkin elimination, a counterpart of the Gaussian elimination, can be adopted to produce an algebraic approach to deal with linear inequalities. However, in the following, we will emphasize the geometric aspect of the subject instead.

Theorem 1.1 is a good characterization for a system of linear equations to have a solution as we have certificates in both cases:

$$\begin{cases} x, & \text{if} \quad \text{solvable}, \\ y, & \text{otherwise}. \end{cases}$$

Problems possessing a good characterization $\longleftrightarrow$ $\mathcal{NP} \cap$co-$\mathcal{NP}$

Question:

Given $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$, then

$$\exists x \in \mathbb{R}^m, Ax \leq b$$

if and only if there is no $y \in \mathbb{R}^m$ such that ...... (See Theorem 1.11)

Given $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$, then $Ax = b$ has a nonnegative solution if and only if there is no $y \in \mathbb{R}^m$ such that...... (See Theorem 1.8)

Linear equations $\leftrightarrow$ Theorem 1.1 (the basic theorem of linear algebra) $\leftrightarrow$ Matroid

Linear inequalities $\leftrightarrow$ Farkas' Lemma (the basic theorem of linear programming) $\leftrightarrow$ Oriented Matroid

The theory of oriented matroids provides a broad setting in which to model, describe, and analyze combinatorial properties of geometrical configurations. It is a common generalization of many apparently totally different mathematical objects.

We mathematicians are all a bit crazy. – Lev Davidovich Landau (Jan. 22, 1908 – Apr. 1, 1968)

Several important theorems which are pairwise equivalent: Farkas' Lemma, Nash Equilibrium Theorem (the basic theorem of Game Theory), Von Neumann Minmax theorem, Linear Programming Duality Theorem (the basic theorem of linear programming), Helly Theorem (For any finite set R of half-spaces in d dimensional Euclidean space E, either there is a point in the intersection of all of R, or there is a subset R' of R, having cardinality at most $d+1$, such that the intersection of R' is empty.), Brouwer's fixed point theorem, Kakutani's fixed point theorem, Hahn-Banach Theorem (a basic theorem of Functional Analysis), the equivalence of the dual representations of a polyhedron (implicit representation: intersection of a finite number of half spaces; generator representation: convex combinations of a finite number of vertices and rays)

Using an appealing metaphor of Broyden, one may say that they resemble cities situated on a high plateau; travel between them is not too difficult; the hard part is the initial ascent from the plains below.
– C. Roos, T. Terlaky, Note on a paper of Broyden, *Operations Research Letters* 25 (1999), 183–186.

The equivalence of the dual representations of a polyhedron is very intuitive but not so easy to prove rigorously. Have a try to prove the equivalence of the dual representations of a polytope (which says that a bounded set is the intersection of finitely half spaces if and only if it is the convex hull of finitely many points) so that you can get some feeling of its hardness!

Exercise 1.3. Read [31, §2.2] yourself – especially, you should have more motivation to do so if you fail to prove the result on the dual representations of a polytope. (It is very well-written and the content should be must-know. But unfortunately I will have no class time to cover it and so it is left to you.)

Let $\mathcal{B}$ be a subset of $\mathbb{R}^m$. $Cone(\mathcal{B}) \doteq \{\sum_{b\in\mathcal{B}} t_b b : \ t_b \geq 0, t_b = 0$ with the exception of a finite number of $b\}$. $Conv(\mathcal{B}) \doteq \{\sum_{b\in\mathcal{B}} t_b b : \ t_b \geq 0, t_b = 0$ with the exception of a finite number of $b, \sum_{b\in\mathcal{B}} t_b = 1\}$.

Let $\mathcal{A} \subseteq \mathbb{R}^n$ and $b \in \mathbb{R}^{\mathcal{A}}$. $P(\mathcal{A}, b) \doteq \{x \in \mathbb{R}^n : \ a^\top x \leq b_a, \forall a \in \mathcal{A}\}$. $P^=(\mathcal{A}, b) \doteq \{x \in \mathbb{R}^n : \ x \geq 0, a^\top x = b_a, \forall a \in \mathcal{A}\}$.

A `convex set` is a set which is closed under the operation $Conv$. A `polyhedron` is a set of the form $P(A, b)$ and a `polytope` is a bounded polyhedron. A `cone` is a set of the form $Cone(B)$.

Let $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$.

$A_i \doteq$ the ith row of A.

$A_I \doteq$ the submatrix of A formed by those rows indexed by I.

$Cone(A) \doteq Cone\{((A^\top)_1)^\top, \cdots, ((A^\top)_n)^\top\}$.

$Conv(A) \doteq Conv\{((A^\top)_1)^\top, \cdots, ((A^\top)_n)^\top\}$.

$Span(A) = Im(A) \doteq \{Ax : \ x \in \mathbb{R}^n\}$.

$Ker(A) \doteq \{y \in \mathbb{R}^n : \ Ay = 0\}$.

$P(A, b) \doteq \{x \in \mathbb{R}^n : \ Ax \leq b\}$. $P^=(A, b) \doteq \{x \in \mathbb{R}^n : \ Ax = b, x \geq 0\}$.

$\mathbb{R}^n_+ \doteq \{x \in \mathbb{R}^n : \ x \geq 0\}$.

A subset H of $\mathbb{R}^m$ is called a `hyperplane` ($(m-1)$-dimensional linear manifold) if there exist a nonzero vector $c \in \mathbb{R}^m$ and a $\delta \in \mathbb{R}$ such that $H = \{x \in \mathbb{R}^m : \ c^\top x = \delta\}$. We say that H `separates` z and C if z and C are in different components of $\mathbb{R}^m \setminus H$.

We sometimes use the notations $H_{c,\delta} = \{x : \ c^\top x = \delta\}$, $H^+_{c,\delta} = \{x : \ c^\top x > \delta\}$, and $H^-_{c,\delta} = \{x : \ c^\top x < \delta\}$.

Theorem 1.4. *(Separating Hyperplane Theorem) Let $\mathcal{C}$ be a closed convex set in $\mathbb{R}^m$ and let $z \in \mathbb{R}^m \backslash \mathcal{C}$. Then there exists a hyperplane separating z and $\mathcal{C}$.*

Proof. The case of $\mathcal{C} = \emptyset$ is trivial and so we assume that $\mathcal{C} \neq \emptyset$. Taking into account that $\mathcal{C}$ is closed, there exists $y \in \mathcal{C}$ that is nearest to z. (Indeed, as $\mathcal{C}$ is also convex, such a point y is unique.)

Let $v = z - y$. We assert that

$$\forall x \in \mathcal{C}, (x - y)^\top v \leq 0 \tag{1.1}$$

$$(z - y)^\top v > 0 \tag{1.2}$$

If this is verified, then we have $x^\top v \leq y^\top v < z^\top v$ for any $x \in \mathcal{C}$ and so it is readily seen that the hyperplane $\{u \in \mathbb{R}^m : \ u^\top v = \frac{y^\top v + z^\top v}{2} = \frac{z^\top z - y^\top y}{2}\}$ is what we require.

(1.2) is obvious as $v = z - y \neq 0$. Let us turn to prove (1.1). Put $H^+ = \{x \in \mathbb{R}^m : \ (x - y)^\top v > 0\}$ and $\mathcal{B} = \{w \in \mathbb{R}^m : (w - z)^\top (w - z) < v^\top v\}$. It is intuitively obvious that for any point $x \in H^+$ the line segment between x and y will intersect $\mathcal{B}$. More formally, we check that $\frac{d}{dt}\|z - (y + t(x - y))\|^2 = -2(x - y)^\top v$ when $t = 0$ and so deduce that for $x \in H^+$ there is a nonempty open interval with y as one endpoint and lying in $Conv(x, y) \cap \mathcal{B}$. Since $\mathcal{C}$ is convex and $\mathcal{C} \cap \mathcal{B} = \emptyset$, we conclude that $H^+ \cap \mathcal{C} = \emptyset$, which is just (1.1), as desired. □

Exercise 1.5. A pair of nonempty closed bounded convex sets can be separated by a hyperplane if and only if their intersection is empty.

If the closed convex set $\mathcal{C}$ turns out to be a cone, we can pick up a linear subspace as the separating hyperplane, as the next theorem shows.

Theorem 1.6. *Let $\mathcal{A} \subseteq \mathbb{R}^m$ and $b \in \mathbb{R}^m$. If $Cone(\mathcal{A})$ is closed, then either*

$$b \in Cone(\mathcal{A}) \tag{1.3}$$

or

$$\exists y \in \mathbb{R}^m, y^\top b < 0, \textit{and } -y \in P(\mathcal{A}, 0), \ \textit{namely } y^\top a \geq 0, \forall a \in \mathcal{A} \tag{1.4}$$

but not both.

We need a simple lemma.

Lemma 1.7. *For any $\mathcal{A} \subseteq R^m$, it holds $P(\mathcal{A}, 0) = P(Cone(\mathcal{A}), 0)$.*

Proof. (of Theorem 1.6) $Cone(\mathcal{A})$ is clearly a closed convex set and so we can apply Theorem 1.4 to find that either it holds (1.3) or it holds the following:

$$\exists y \in \mathbb{R}^m, \exists \delta \in \mathbb{R}, y^\top b < \delta, \text{ and } y^\top a \geq \delta, \forall a \in Cone(\mathcal{A}), \tag{1.5}$$

but not both. Lemma 1.7 means that (1.4) is equivalent to

$$\exists y \in \mathbb{R}^m, y^\top b < 0, \text{ and } y^\top a \geq 0, \forall a \in Cone(\mathcal{A}), \tag{1.6}$$

while (1.5) follows from (1.6) trivially. Therefore, to complete the proof we only need to deduce (1.4) from (1.5). First, $0 \in Cone(\mathcal{A})$ tells us that $\delta \leq y^\top 0 = 0$ and so $y^\top b < 0$ follows from $y^\top b < \delta$. We then prove by contradiction that $y^\top a \geq 0, \forall a \in \mathcal{A}$. Otherwise, there is $a \in \mathcal{A}$ such that $y^\top a < 0$. Taking a sufficiently large number $t > 0$, we will have $ta \in Cone(\mathcal{A})$ and $y^\top(ta) < \delta$, contradicting with $y^\top a \geq \delta, \forall a \in Cone(\mathcal{A})$, as asserted in (1.5). This ends the proof. □

If $\mathcal{A}$ is a finite set, Theorem 1.6 takes the following form. (why?)

Theorem 1.8. *(Farkas' Lemma) Given $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$, then either*

$$b \in Cone(A) \tag{1.7}$$

or

$$\exists y \in R^m, y^\top A \geq 0, y^\top b < 0 \tag{1.8}$$

but not both.

Farkas' motivation came not from mathematical economics nor yet pure mathematics but from physics; as a professor of Theoretical Physics he was interested in the problem of mechanical equilibrium and it was these that gave rise to the need for linear inequalities. In this he was continuing the classical work of Fourier and Gauss, though Farkas claims to be the first to appreciate the importance of homogenious linear inequalities to these problems. – C.G. Broyden, A simple algebraic proof of Farkas' Lemma and related theorems, *Optim. Methods Software* 8 (1998), 185–199.

All truths are easy to understand once they are discovered; the point is to discover them. – Galileo Galilei (1564 – 1642); Italian astronomer & physicist.

Recall that for a matrix $B_{m\times(n+1)}$, $Ker(B) = \{x \in \mathbb{R}^{n+1} : Bx = 0\}$ and $Im(B^\top) = \{B^\top y \in \mathbb{R}^{n+1} : \ y \in \mathbb{R}^m\}$ are orthogonal complement to each other in $\mathbb{R}^{n+1}$. That is

$$Ker(B)^\perp = Span(B_1^\top, \cdots B_m^\top). \tag{1.9}$$

By means of linearization, we obtain the next theorem from Theorem 1.8.

Theorem 1.9. *Let $L \leq \mathbb{R}^{n+1}$ be a subspace of $\mathbb{R}^{n+1}$. Then either*

$$\exists y \in L, y \geq 0, y_{n+1} > 0 \tag{1.10}$$

or

$$\exists u \in L^\perp, u \geq 0, u_{n+1} > 0 \tag{1.11}$$

but not both.

Proof. Let $L = Ker(B)$ where $B = \begin{pmatrix} A & -b \end{pmatrix}$, $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$. It follows that $L^\perp = Im(B^\top)$ and so (1.11) is just (1.8). In addition, we prove that (1.10) is equivalent to (1.7):

$$\begin{aligned} b \in Cone(A) \Leftrightarrow & \quad \exists x \in R^n, x \geq 0, Ax - b = 0 \\ \Leftrightarrow & \quad \exists x \in R^n, x \geq 0, \begin{pmatrix} A & -b \end{pmatrix} \begin{pmatrix} x \\ 1 \end{pmatrix} = 0 \\ \Leftrightarrow & \quad \exists y \in L, y \geq 0, y_{n+1} > 0. \end{aligned}$$

By now, we see that this theorem is a consequence of Theorem 1.8. □

Exercise 1.10. Prove that L and $L^{\perp}$ can be taken to be the circuit space and the cocircuit space of a digraph and give the graphical interpretation of Theorem 1.9 in such a case and then provide a graphical proof of the fact.

(Some of you complains that it is hard to find any relevant references for this exercise. So let me indicate below how you can get something of help.)

First, I would like to add the hint that it is the Arc Colouring Lemma (or Minty's Lemma) for which I expected you to give two proofs (using Theorem 1.9 and using first principle)

Then let me tell you that some discussion of the Arc Colouring Lemma and its use in algorithm analysis can be located in the following wonderful book available in our library:

Claude Berge, *Graphs*, North-Holland, Amsterdam, 1985.

(Warning: The terminology used by Berge is different from what I use in class; also, be kindly advised that the relevant materials scatter in the above book and you should not expect to locate them in one minute.)

Finally, be informed that you are more than welcome to submit to me a report on the application of the Arc Colouring Lemma.

Claude Berge (June 5, 1926 – June 30, 2002) is a famous combinatorist. Here is what he looks like: `http://www.ecp6.jussieu.fr/GT04/Berge/Berge.html`.

Gian-Carlo Rota once wrote:" Two Frenchmen have played a major role in the renaissance of combinatorics: Berge and Schützenberger. Berge has been the more prolific writer, and his books have carried the word farther and more efficiently than anyone anywhere. I recall the pleasure of reading the disparate examples in his first book, which made it impossible to forget the material. Soon after reading, I would be one of many who unknotted themselves from the tentacles of the continuum and joined the Rebel Army of the Discrete."

Fortunately, we can find several books written by Berge in our library:

- *Principles of Combinatorics*, Academic Press, New-York, 1971. (Only Chinese translation available; unfortunately, it is not a very good translation.)
- *Graphs and Hypergraphs*, North-Holland, Amsterdam, 1973.
- *Graphs*, North-Holland, Amsterdam, 1985.
- *Hypergraphs : Combinatorics of Finite Sets*, North Holland, Amsterdam, 1989. (Both the English translation and the Chinese translation are available. Note that the Chinese translation is based on the English one.)

Homogenization together with its corresponding inverse operation, dehomogenization are useful tools in polyhedra theory. For $A \in R^{n\times m}$ and $b \in \mathbb{R}^n$, the homogenization of the polyhedron $P(A, -b)$ is the polyhedral cone $P(B, 0)$, where $B = \begin{pmatrix} A & b \\ 0 & 1 \end{pmatrix} \in \mathbb{R}^{(n+1)\times(m+1)}$. See [2, §4.2] for an intuitive geometric insight into this algebraic definition. It will be no surprise that we make use of the special matrix B in the proof of the following theorem when you understand that its special form comes out of the homogenization process and so it is really not as special as you might think at the first sight. Indeed, we will first use delinearization and then use dehomogenization in the process of proving the next theorem.

Theorem 1.11. *Given $A \in \mathbb{R}^{n\times m}$ and $b \in \mathbb{R}^n$, then either*

$$\exists x \in \mathbb{R}^m, Ax \leq b \tag{1.12}$$

or

$$\exists v \in \mathbb{R}^n_+, v^\top A = 0, v^\top b < 0 \tag{1.13}$$

but not both.

Observe that Theorem 1.11 provides us good certificates both when $Ax \leq b$ is solvable and when it is unsolvable.

Proof. Let $B = \begin{pmatrix} A & b \\ 0 & 1 \end{pmatrix} \in \mathbb{R}^{(n+1)\times(m+1)}$, $L = Im(B) = \{y = Bz \in \mathbb{R}^{n+1} : z \in \mathbb{R}^{m+1}\}$ and $L^\perp = Ker(B^\top) = \{u \in \mathbb{R}^{n+1} : u^\top B = 0\}$. In virtue of Theorem 1.9, we can finish the proof by simply interpreting (1.10) and (1.11) for this pair $(L, L^\perp)$:

$$\begin{aligned}
(1.10) \Leftrightarrow &\quad \exists y \in L, y \leq 0, y_{n+1} < 0 \\
\Leftrightarrow &\quad \exists z \in \mathbb{R}^{m+1}, Bz \leq 0, (Bz)_{n+1} < 0 \\
\Leftrightarrow &\quad \exists z \in \mathbb{R}^{m+1}, Bz \leq 0, z_{m+1} = -1 \\
\Leftrightarrow &\quad \exists x \in \mathbb{R}^{m}, Ax - b \leq 0 \\
\Leftrightarrow &\quad (1.12)
\end{aligned}$$

$$\begin{aligned}
(1.11) \Leftrightarrow &\quad \exists u \in \mathbb{R}^{n+1}, u^\top B = 0, u \geq 0, u_{n+1} > 0 \\
\Leftrightarrow &\quad \exists u \in \mathbb{R}^{n+1}_{+}, u^\top B = 0, u_{n+1} > 0 \\
\Leftrightarrow &\quad \exists v \in \mathbb{R}^{n}_{+}, v^\top A = 0, v^\top b = -u_{n+1} < 0 \\
\Leftrightarrow &\quad (1.13)
\end{aligned}$$

□

Exercise 1.12. Use Theorem 1.8 to prove Theorem 1.11 directly. (hint: see [31, §2.3])

The following theorem is obviously a generalization of Theorem 1.8 (setting $V = \{0\}$!). As another example of the use of the homogenization technique, we will show that it is indeed equivalent to Theorem 1.8. Note that this time we will homogenize a polyhedron given by generators rather than defined by linear inequalities.

Theorem 1.13. *Let $V, E \subseteq \mathbb{R}^m$ be finite sets and $b \in \mathbb{R}^m$. Then either*

$$b \in Conv(V) + Cone(E) = P \tag{1.14}$$

or

$$\exists c \in \mathbb{R}^m, \lambda \in \mathbb{R}, \text{ such that } c^\top v \geq \lambda, \forall v \in V, c^\top e \geq 0, \forall e \in E, c^\top b < \lambda \tag{1.15}$$

but not both.

Proof. We set $A = \begin{pmatrix} V & E \\ -\mathbf{1}^\top & 0 \end{pmatrix}$, where $\mathbf{1}$ stands for the all ones column vector of length $\#V$. Note that (1.14) is equivalent to

$$\begin{pmatrix} b \\ -1 \end{pmatrix} \in Cone(A). \tag{1.16}$$

(Hence Cone (A) is a homogenization of P!) Theorem 1.8 now justifies the claim that either (1.16) (and hence (1.14)) is valid or

$$\exists z \in \mathbb{R}^{m+1}, z^\top A \geq 0, z^\top \begin{pmatrix} b \\ -1 \end{pmatrix} < 0 \tag{1.17}$$

but not both. Thus for our purpose we must illustrate the equivalence of (1.15) and (1.17). However, this is immediate after writing $z = \begin{pmatrix} c \\ \lambda \end{pmatrix}$ where $c \in \mathbb{R}^m$ and $\lambda \in \mathbb{R}$. □

Mathematicians are like Frenchmen: whatever you say to them they translate into their own language and forthwith it is something entirely different. – German writer, Johann Wolfgang von Goethe

Wow, Berge and Schützenberger!

We have given quite a few theorems above. Note that sometimes to prove a theorem is just to translate something from one language to another language. In some sense, we only prove Theorem 1.4 and then we are merely doing translations in establishing the other theorems. Here, we use homogenization, linearization and their inverse processes to look at the same thing from different viewpoints and then report what we have seen. To be able to recognize these different formulation of the same thing will be very helpful in applications. Note that the mathematics you need in order to do creative work is not written down in most cases. Sometimes, you complain that no mathematics can be applied to your problem and later find that the mathematics is indeed ready long long ago. This happens because you do not know the exact formulation of the mathematics that matches your need, although you might believe that you have a good understanding of that kind of easy mathematics.

We proceed to introduce some equivalents of Farkas' Lemma. But the proof will not be an obvious translation now. (Perhaps this is because I do not find the correct viewpoint.)

Theorem 1.14. *Let* $c \in \mathbb{R}^m$ *and* P *a polyhedron in* $\mathbb{R}^m$. *If* $\sup\{c^\top x : x \in P\}$ *is finite, then* $\max\{c^\top x : x \in P\}$ *is attained.*

Please think over what the theorem is talking about! It is again a very intuitive one!

Even if you are on the right track, you will get run over, if you just sit there. – Will Rogers, actor, (1879–1935)

Proof. (of Theorem 1.14) Let $\delta = \sup\{c^\top x : x \in P\}$ and $P = P(A, b)$ for some $A \in R^{n\times m}$ and $b \in R^n$. Then the question reduces to demonstrating the existence of an x such that $Ax \leq b$ and $c^\top x \geq \delta$, namely the existence of solution x to the system of linear inequalities $\begin{pmatrix} A \\ -c^\top \end{pmatrix} x \leq \begin{pmatrix} b \\ -\delta \end{pmatrix}$. If such an x does not exist, then Theorem 1.11 says that there is a $v \in R_+^{n+1}$ satisfying

$$v^\top \begin{pmatrix} A \\ -c^\top \end{pmatrix} = 0 \tag{1.18}$$

and

$$v^\top \begin{pmatrix} b \\ -\delta \end{pmatrix} < 0. \tag{1.19}$$

We will show that this is impossible and hence complete the proof. Let $v = \begin{pmatrix} u \\ \lambda \end{pmatrix}$ where $u \in R_+^n$ and $\lambda \in R_+$. Note that for any $x \in P$ we have

$$\begin{aligned} \lambda c^\top x = & \quad u^\top Ax \quad \text{by (1.18)} \\ \leq & \quad u^\top b \quad \text{by } \ x \in P \text{ and } \ u \in R_+^n \end{aligned} \tag{1.20}$$

In addition, due to the fact that δ is finite we conclude that $P \neq \emptyset$. We are now in a place to give

$$\begin{aligned} \lambda\delta = & \quad \lambda \sup\{c^\top x : \ x \in P\} \\ \leq & \quad u^\top b \quad \text{by } \ (1.20) \text{ and } \ P \neq \emptyset \\ < & \quad \lambda\delta \quad \text{by (1.19)} \end{aligned}$$

which is surely a contradiction, as desired. □

Note that unlike Theorem 1.4, Theorem 1.14 does not hold any more when P is only assumed to be closed and convex. Also note that Theorem 1.14 is trivial when P is closed and bounded, and hence when it is a polytope. In general, as an equivalent form of Farkas' Lemma, we have the Decomposition Theorem for Polyhedra: P is a polyhedron if and only if $P = Q + C$ for some polytope Q and some polyhedral cone C. It is easy to see that Theorem 1.14 comes easily from the Decomposition Theorem for Polyhedra.

Exercise 1.15. Prove the Decomposition Theorem for Polyhedra. (This is Exercise 2.12 in [31] and you are invited to prove it either following the hint given in [31] or basing on some theorems I already proved in class. I promise to present a proof at the end of this chapter.)

Exercise 1.16. Give a direct elementary proof of Theorem 1.14. (Here I mean a proof in the spirit of that of Theorem 1.4.) (hint: a polyhedron is defined by a **finite** system of **linear** inequalities.)

Exercise 1.17. Let $A \in \mathbb{R}^{m\times n}, b \in \mathbb{R}^n, c \in \mathbb{R}^m$. If $P(A, c) \neq \emptyset$ and $\sup\{b^\top x : \; x \in P(A, c)\}$ is finite, then b is a nonnegative combination of the columns of $A^\top$.

With Theorem 1.8 and Theorem 1.14 at hand, we are well prepared to prove the celebrated LP duality theorem. I hope that the beautiful geometric idea behind the LP duality theorem as embodied in the proof below will get you excited. You will like the proof here if you have had a try at Exercise 1.17.

Theorem 1.18. *(Duality Theorem of Linear Programming) Let* $A \in \mathbb{R}^{m\times n}, b \in \mathbb{R}^n, c \in \mathbb{R}^m$*. Then both* $\max\{b^\top x : \ Ax \leq c\}$ *and* $\min\{y^\top c : \ y \geq 0, y^\top A = b^\top\}$ *exist and are equal, provided both* $P(A,c)$ *and* $P^=(A^\top, b)$ *are nonempty.*

Proof. (of Theorem 1.18) First note that

$$\sup\{b^\top x : Ax \le c\} \le \inf\{y^\top c : y \ge 0, y^\top A = b^\top\}, \qquad (1.21)$$

because for $x \in P(A,c)$ and $y \in P^{=}(A^\top, b)$ it holds $b^\top x = (y^\top A)x = y^\top(Ax) \le y^\top c$. As both $P(A,c)$ and $P^{=}(A^\top, b)$ are nonempty, $\sup\{b^\top x : Ax \le c\}$ is finite and hence it follows from Theorem 1.14 that there is $x^* \in P(A,c)$ and $\delta \in \mathbb{R}$ such that

$$b^\top x^* = \max\{b^\top x : Ax \le c\} = \delta. \qquad (1.22)$$

Our strategy is to find a $y^* \in P^{=}(A^\top, b)$ satisfying

$$y^{*\top} c = \delta, \qquad (1.23)$$

which together with (1.21) and (1.22) will end the proof.

Assume without loss of generality that $A_i x \leq c_i$, $i = 1, \cdots, k$, are those inequalities in $Ax \leq c$ for which the equality holds at $x = x^*$. Now, by Theorem 1.8 we can assert that either

$$b \in Cone(A_1^\top, \cdots, A_k^\top) \tag{1.24}$$

or

$$\exists y \in \mathbb{R}^n, b^\top y < 0, A_i y \geq 0, i = 1, \cdots, k, \tag{1.25}$$

but not both.

We point out that (1.25) cannot occur. Otherwise, for the corresponding y we can pick up a small enough positive number ϵ such that $z = x^* - \epsilon y$ fulfils $z \in P(A, c)$ and $b^\top z > b^\top x^*$, contradicting (1.22).

Therefore, we arrive at $b \in Cone(A_1^\top, \cdots, A_k^\top)$ (especially, $k \geq 1$!), say

$$\exists \lambda_1, \cdots, \lambda_k \geq 0, b = \sum_{i=1}^{k} \lambda_i A_i^\top. \tag{1.26}$$

Put $y^* = (\lambda_1, \cdots, \lambda_k, 0, \cdots, 0) \in \mathbb{R}^m$. Clearly, (1.26) guarantees that $y^* \in P^=(A^\top, b)$. Moreover, we have

$$\begin{aligned} \delta = & \quad b^\top x^* \quad by \;\; (1.22) \\ = & \quad \textstyle\sum_{i=1}^k \lambda_i A_i x^* \quad \text{by} \;\; (1.26) \\ = & \quad \textstyle\sum_{i=1}^k \lambda_i c_i \\ = & \quad y^{*\top} c, \end{aligned}$$

verifying (1.23) and hence ending the proof. □

A careful checking of our proof above yields (Get your hands dirty with the checking! Do not just take my words.)

Theorem 1.19. *Let $A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^n, c \in \mathbb{R}^m$. Then both* $\max\{b^\top x : \; Ax \leq c\}$ *and* $\min\{y^\top c : \; y \geq 0, y^\top A = b^\top\}$ *exist and are equal, provided $P(A, c)$ is nonempty and* $\sup\{b^\top x : \; Ax \leq c\}$ *is finite.*

Theorem 1.20. *Let $A \in \mathbb{R}^{m\times n}, b \in \mathbb{R}^n, c \in \mathbb{R}^m$. Then both* $\max\{b^\top x : \ x \geq 0, Ax \leq c\}$ *and* $\min\{y^\top c : \ y \geq 0, y^\top A \geq b^\top\}$ *exist and are equal, provided both $P(A, c) \cap R^n_+$ and $P(-A^\top, b) \cap \mathbb{R}^m_+$ are nonempty.*

Proof.

$$\begin{aligned}
& \max\{b^\top x : \ x \geq 0, Ax \leq c\} \\
= \quad & \max\{b^\top x : \begin{pmatrix} A \\ -I \end{pmatrix} x \leq \begin{pmatrix} c \\ 0 \end{pmatrix}\} \\
= \quad & \min\{\begin{pmatrix} y_1^\top & y_2^\top \end{pmatrix}\begin{pmatrix} c \\ 0 \end{pmatrix} : \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \geq 0, \begin{pmatrix} y_1^\top & y_2^\top \end{pmatrix}\begin{pmatrix} A \\ -I \end{pmatrix} = b^\top\} \\
= \quad & \min\{y^\top c : \ y \geq 0, y^\top A \geq b^\top\}
\end{aligned}$$

Here, the first and the last equalities are doing translations and the middle one follows from Theorem 1.18. □

Exercise 1.21. (The Minimax Theorem of Von Neumann on two-person zero-sum game) Let $P_n = \{x \in \mathbb{R}^n_+ : \ \sum_{i=1}^n x_i = 1\}$. Prove that for every $A \in \mathbb{R}^{m\times n}$ there is $x^* \in P_m$ and $y^* \in P_n$ such that $\min_{y\in P_n} {x^*}^\top Ay = \max_{x\in P_m} x^\top Ay^*$.

By and large it is uniformly true that in mathematics there is a time lapse between a mathematical discovery and the moment it becomes useful; and that this lapse can be anything from 30 to 100 years, in some cases even more; and that the whole system seems to function without any direction, without any reference to usefulness, and without any desire to do things which are useful. – John von Neumann (1903–1957).

We summarize the rule of Primal-Dual Correspondence as follows:

$A \cdots\cdots\cdots\cdots\cdots\cdots A^\top$

Objective: Max $b^\top x$ $\cdots\cdots\cdots\cdots\cdots\cdots$ Objective: Min $c^\top y$

jth variable $x_j \geq 0 \cdots$ jth constraint $\sum_i A(i,j)y_i = (A^\top y)_j \geq b_j$

jth variable $x_j \leq 0 \cdots$ jth constraint $\sum_i A(i,j)y_i = (A^\top y)_j \leq b_j$

jth variable x_j free $\cdots$ jth constraint $\sum_i A(i,j)y_i = (A^\top y)_j = b_j$

ith constraint $\sum_j A(i,j)x_j = (Ax)_i \leq c_i$ $\cdots$ ith variable $y_i \geq 0$

ith constraint $\sum_j A(i,j)x_j = (Ax)_i \geq c_i$ $\cdots$ ith variable $y_i \leq 0$

ith constraint $\sum_j A(i,j)x_j = (Ax)_i = c_i$ $\cdots$ ith variable y_i free

Exercise 1.22. Verify the correctness of the above rule, namely prove [31, Exercise 2.25].

The general LP Duality Theorem (See the Primal-Dual Correspondence as above) tells us that to certify that a proposed x^* is an optimal solution to the primal problem is to test that it is feasible and to find a feasible solution y^* to the dual problem and verify $b^\top x^* = c^\top y^*$.

When it is inconvenient to directly compute $b^\top x$ or $c^\top y$, we often appeal to the next theorem.

Theorem 1.23. *(Complementary Slackness Theorem) For a general form of a linear programming problem (P) and its dual (D) as illustrated above, two points x^* and y^* are simultaneously optimal solutions for (P) and (D) if and only if the following two statements hold:*

(i) x^* is feasible for (P) and y^* is feasible for (D);

(ii) $y^{*\top}(Ax^* - c) = 0$ (symmetrically, $(y^{*\top}A - b^\top)x^* = 0$).

Note that Theorem 1.23 says that if a variable is slack then the dual constraint is tight, or in the contrapositive, if a constraint is loose, then the dual variable is zero. This is why it is named the Complementary Slackness Theorem.

Exercise 1.24. Prove Theorem 1.23.

The mathematical sciences particularly exhibit order, symmetry, and limitation; and these are the greatest forms of the beautiful.
– Aristotle (384-322 BC)

Let $A \in \mathbb{R}^{m\times n}$ and $c \in \mathbb{R}^m$.

$r^{=}(A,c) \doteq \{i: \ A_i z = c_i, \forall z \in P(A,c)\}$.

For any subset $F \subseteq \mathbb{R}^n$, $r^{=}_F(A,c) \doteq \{i: \ A_i z = c_i, \forall z \in F\}$.

Especially, for a point $z \in \mathbb{R}^n$, $r^{=}_z(A,c) \doteq \{i: \ A_i z = c_i\}$.

A_c ($c^{=}_A$, resp.) is the submatrix of A (c, resp.) obtained by deleting all those rows j such that $j \notin r^{=}(A,c)$; $A_{c,F}$ ($c^{=}_{A,F}$, resp.) is the submatrix of A (b, resp.) obtained by deleting all those rows j such that $j \notin r^{=}_F(A,c)$. We write $A_{c,z}$ ($c^{=}_{A,z}$) for $A_{c,F}$ ($c^{=}_{A,F}$) when F consists of a single point z.

$x \in P(A,c)$ is called an `inner point` of $P(A,c)$ if $A_i x < c_i$ for all $i \notin r^{=}(A,c)$. $x \in P(A,c)$ is called an `interior point` of $P(A,c)$ if $A_i x < c_i$ for all i.

Exercise 1.25. Every nonempty polyhedron has an inner point.

Recall: The following statements are equivalent:

- $x^1, \cdots, x^k \in \mathbb{R}^n$ are affinely independent.
- $x^2 - x^1, \cdots, x^k - x^1$ are linearly independent.
- $\begin{pmatrix} x^1 \\ 1 \end{pmatrix}, \cdots, \begin{pmatrix} x^k \\ 1 \end{pmatrix} \in \mathbb{R}^{n+1}$ are linearly independent.

A polyhedron P is of `dimension` k, denoted by $\dim(P) = k$, if the maximum number of affinely independent points in P is $k+1$. (This does NOT amount to saying that the maximum number of linearly independent points in P is k. In the definition of linearly independence, the vector 0 has a special role while in the definition of affinely independence we have homogenized the space.)

We adopt the convention that if $P = \emptyset$, then $\dim(P) = -1$.

Theorem 1.26. *Assume that* $P(A,c) \neq \emptyset$ *for* $A \in \mathbb{R}^{m\times n}$ *and* $c \in R^m$. *Then it holds* $\dim(P(A,c)) = n - \operatorname{rank}(A_c)$.

Proof. From basic linear algebra we know that there are $n - \operatorname{rank}(A_c) + 1$ affinely independent solutions $x \in \mathbb{R}^n$ of $A_c x = 0$, say $x = x^1, \cdots, x^{n-\operatorname{rank}(A_c)+1}$. Since $P(A,c) \neq \emptyset$, we can take an inner point of P, say $\tilde{x}$ (Exercise 1.25). Now for ϵ sufficiently small, $\tilde{x} + \epsilon x^i$, $i = 1, \cdots, n - \operatorname{rank}(A_c) + 1$, are affinely independent points in $P(A,c)$. This proves $\dim P(A,c) \geq n - \operatorname{rank}(A_c)$.

Conversely, from the definition of the dimension of a polyhedron, $P(A,c)$ contains points $y^1, \cdots, y^{\dim P(A,c)}$ which are affinely independent. But according to the definition of A_c, $x = y^1, \cdots, y^{\dim P(A,c)}$ are affinely independent solutions to $A_c x = b^=_A$. This shows $\dim P(A,c) \leq n - \operatorname{rank}(A_c)$.

Putting the pieces together, we arrive at the conclusion, as desired. □

Theorem 1.26 means that $\operatorname{rank}(A_c)$ is an invariant of the polyhedron $P = P(A,c) \neq \emptyset$, namely it is independent of the particular inequality description of P.

For $b \in \mathbb{R}^n, \delta \in \mathbb{R}$, the inequality $(b, \delta) : \ b^\top x \le \delta$ is called a `valid inequality` for a polyhedron P provided it is satisfied by all points of P.

If (b, δ) is a valid inequality for P, then we refer to $F = \{x \in P : \ b^\top x = \delta\}$ as a `face` of P and we say that (b, δ) represents F. A face F of P is said to be `proper` if $F \neq \emptyset$ and $F \neq P$.

The face F represented by (b, δ) is nonempty if and only if $\max\{b^\top x : \ x \in P\} = \delta$ and in such a case it holds F is the set of optimal solutions to the linear program $\delta = \max\{b^\top x : \ x \in P(A, c)\}$. When $F \neq \emptyset$, we say that (b, δ) `supports` P.

Theorem 1.27. *Let $A \in \mathbb{R}^{m \times n}$ and $c \in \mathbb{R}^m$. If F is a nonempty face of $P(A, c)$, then F is a polyhedron and $F = \{x \in \mathbb{R}^n : \ A_i x = c_i$ for $i \in r_F^=(A, c)$ and $A_i x \le c_i$ for $i \notin r_F^=(A, c)\}$.*[2]

That F is a polyhedron is obvious (why?). The essential point of Theorem 1.27 is the algebraic description of the faces. A polyhedron may have infinitely many supporting hyperplanes while the description of the faces only involves the fixed parameters A and c. It is NOT trivial. It reduces INFINITY to FINITY.

Corollary 1.28. *The number of distinct faces of a polyhedron is finite.*

(why?)

[2]Some of you indicated to me that this is a circular definition. Please note that this is not any definition of F. Instead, this is an assertion on F.

Proof. (of Theorem 1.27) Suppose F is the set of optimal solutions to the linear program $\delta = \max\{b^\top x : x \in P(A, c)\}$. We write

$$I^* = r_F^=(A, c). \tag{1.27}$$

Let x^* be a point of F such that

$$r_{x^*}^=(A, c) = I^*. \tag{1.28}$$

(the existence of such an x^* follows exactly in the same way as one proves the existence of an inner point of a nonempty polyhedron. Indeed, after we confirm that F is a polyhedron, x^* turns out to be an inner point of F.)

As in the proof of Theorem 1.18, corresponding to this x^*, there is $y^* \in P^=(A^\top, b)$ satisfying $y^{*\top} c = \delta$. Observe that $y^{*\top} A = b^\top$. [3]

Let F' be the polyhedron $\{x \in R^n : A_i x = c_i$ for $i \in I^*$ and $A_i x \leq c_i$ for $i \notin I^*\}$.

We are ready to demonstrate $F' = F = \{x : b^\top x = \delta, x \in P(A, c)\}$, which is what we want. First, because of Eq. (1.27), it is trivially true that $F \subseteq F'$. $F' \subseteq F$ can be seen from the following calculation: for any $x \in F'$, $b^\top x = \sum_{i \in I^*} y_i^* A_i x = \sum_{i \in I^*} y_i^* c_i = y^{*\top} c = \delta$. (The first equality comes from our comment on the footnote 3 and the third equality is due to Eq. (1.27)) ☐

(Notice that the calculation appeared at the end of the proof is also the key step in our proof of the LP duality Theorem!)

[3]Recall that y^* consists of the nonnegative coefficients appeared in the expression of $b^\top$ as a conical combination of those A_i for $i \in r_{x^*}^=(A, c)$, which is just I^*, as a result of Eq. (1.28).

A face F of a polyhedron P is a `facet` of P provided $\dim(F) = \dim(P) - 1$.

Exercise 1.29. If F is a facet of $P(A, c)$, then F is represented by some inequality (A_i, c_i).

Let $A \in \mathbb{R}^{m\times n}$ and $c \in \mathbb{R}^m$. If $P(A, c)$ has a face $F \neq \emptyset$, then Theorem 1.26 together with Theorem 1.27 guarantees that

$$\dim(F) = n - \operatorname{rank}(A_{c,F}) \geq n - \operatorname{rank}(A). \qquad (1.29)$$

(how?)

Indeed, we can say something more.

Theorem 1.30. *Suppose* $P(A, c) \neq \emptyset$ *and* $\operatorname{rank}(A) = n - k$. *Then each minimal nonempty face of* P *under inclusion has dimension* k.

Proof. Let F be a minimal nonempty face of P under inclusion.

If $\dim(F) = 0$, then Eq. (1.29) forces $k = 0$ and hence the result follows. So, we assume $\dim(F) > 0$ in the sequel.

Let $\tilde{x}$ be an inner point of F. (this means that $\tilde{x}$ is outside of any facet of F.) Since $\dim(F) > 0$, there exists some other point $y \in F$. Consider the line joining $\tilde{x}$ and y, that is $z(\lambda) = \tilde{x} + \lambda(y - \tilde{x})$ where $\lambda \in R$. Suppose that the line intersects the hyperplane $H_{A_i,c_i} = \{x : \ A_i x = c_i\}$ for some $i \notin r_F^{=}(A, c)$. Let $\lambda^* = \min\{|\lambda^i| : \ i \notin r_F^{=}(A, c), z(\lambda^i) \text{ lies in } H_{A_i,c_i}\}$ and assume that $\lambda^* = |\lambda^k|$. Notice that $\lambda^* \neq 0$ as $\tilde{x}$ is an inner point of F. Thus we find that $F' = F \cap H_{A_i,c_i} \neq \emptyset$ is a face of P of smaller dimension than F (why?), which is a contradiction.

Therefore, the line does not intersect H_{A_i,c_i} for all $i \notin r_F^{=}(A, c)$. But this means that $A\tilde{x} + \lambda A(y - \tilde{x}) \leq c$ for all $\lambda \in R$. This instead gives $A(y - \tilde{x}) = 0$ for all $y \in F$. Consequently, $F = \{y : \ Ay = A\tilde{x}\}$. Since $\operatorname{rank}(A) = n - k$, Theorem 1.26 yields $\dim(F) = k$. □

Exercise 1.31. A set F is a minimal nonempty face of $P(A, c) \neq \emptyset$ if and only if $F = \{x : \ A'x = c'\}$ for some subsystem $A'x \leq c'$ of $Ax \leq c$. (hint: Theorems 1.27 and 1.30)

Exercise 1.31 comes from [19], in which the authors also proved the famous Hoffman-Kruskal Theorem.

The set of all faces of a polyhedron P, denoted by $\mathcal{F}(P)$, ordered by set inclusion is called the `face lattice` of P.

The face lattice $\mathcal{F}(P)$ satisfies the Jordan-Dedekind chain property, i.e., every maximal chain: $F = F_1 \subset F_2 \subset \cdots \subset F_k = F'$ between any two ordered elements $F \subset F'$ has the same length.

Let $P = P(A, c)$. We associate with (A, c) the lattice $\mathcal{F}(A, c)$.

We prove that $\mathcal{F}(A, c) = \mathcal{F}(P)$. We have seen lots of geometric—-algebraic correspondences!

(to be finished)

Let P be a polyhedron and $z \in P$. z is an `extreme point` of P if there do not exist $z^1 \neq z^2 \in P$ such that $z = \frac{1}{2}(z^1 + z^2.)$ z is a `vertex` of P if and only if it is a zero-dimensional face of P

Theorem 1.32. *Let $P = P(A, c) \neq \emptyset$ for some $A \in \mathbb{R}^{m \times n}$ and $c \in \mathbb{R}^m$. The following assertions are equivalent:*

1. *z is an extreme point of P.*

2. *z is a vertex of P.*

3. $\operatorname{rank}(A_{c,z}) = n$.

Proof. By Theorem 1.27, z is a zero-dimensional face of P iff $\operatorname{rank}(A_{c,z}) = n$ (and z is not a zero-dimensional face of P iff $\operatorname{rank}(A_{c,z}) < n$).

Suppose z is a zero-dimensional face of P and so $\operatorname{rank}(A_{c,z}) = n$. We choose $I \subseteq r_z(A, c)$ such that A_I is an $n \times n$ invertible matrix. Now, $z = \frac{1}{2}z^1 + \frac{1}{2}z^2 \Rightarrow c_I = A_I z = \frac{1}{2}A_I z^1 + \frac{1}{2}A_I z^2 \leq c_I \Rightarrow A_I z^1 = A_I z^2 = A_I z = c_I \Rightarrow z = z^1 = z^2 = A_I^{-1} c_I$.

Conversely, if $\operatorname{rank} A_{c,z} < n$, then we can pick up a $y \neq 0$ such that $A_{c,z} y = 0$ and thus for sufficiently small $\epsilon \neq 0$, both $z^1 = z + \epsilon$ and $z^2 = z - \epsilon$ fall in P and $z = \frac{1}{2}(z^1 + z^2)$. $\square$

Corollary 1.33. *For* $A \in \mathbb{R}^{m\times n}$ *and* $c \in \mathbb{R}^m$, $P(A,c) \neq \emptyset$ *has a vertex if and only if* $\operatorname{rank}(A) = n$.

Exercise 1.34. Let $A \in \mathbb{R}^{m\times n}$ and $c \in \mathbb{R}^n$. Then $P(A,c)$ has at most $\binom{m}{n}$ vertices. (This means that if you find a lot of vertices in a polyhedron, then you will need many inequalities to describe it. In other words, the number of vertices is some kind of measure of the complexity of a polyhedron.)

Recall Corollary 1.28. You are encouraged to give tighter estimations for the number of faces of a polyhedron. Note that McMullen [4] was able to prove that $P(A,c)$ has at most $f(m,n) = \binom{m-\lfloor\frac{n+1}{2}\rfloor}{m-n} + \binom{m-\lfloor\frac{n+2}{2}\rfloor}{m-n}$ vertices. (Can you?)

Corollary 1.35. *If a polyhedron* P *has a vertex and* $\max\{b^\top x : x \in P\} = \delta$, *then there is vertex* x^* *of* P *such that* $b^\top x^* = \delta$.

Proof. Immediate from Theorem 1.30. □

It is often true that a vertex has some combinatorial meaning and thus this corollary will be of much help to us (not in this course, perhaps).

[4] P. McMullen, The maximum numbers of faces of a convex polytope, *Mathematika* 17 (1970), 179–184.

Consider a polyhedron in the standard LP form. Then, a basic feasible solution to the LP problem and an extreme point of the polyhedron are equivalent; the former is algebraic and the latter is geometric.

The SIMPLEX METHOD for solving a linear programming problem is to proceed from one basic feasible solution (an extreme point of the feasible region) to an adjacent one, in such a way as to continuously decrease the value of the objective function until a minimizer is reached.

In contrast, INTERIOR-POINT algorithms will move in the interior of the feasible region and reduce the value of the objective function, hoping to by-pass many extreme points on the boundary of the region.

Theorem 1.36. *(Caratheodory's Theorem)*

1. *Given a set S, for any point p in $Conv(S)$ there is a subset T satisfying $p \in Conv(T)$, $|T| = dim(sapn(S)) + 1$, and the points of T are affinely independent.*

2. *Given a set S, for any point p in $Cone(S)$, there is a subset T of S with p in $Cone(T)$, with $|T| = dim(span(S))$, and the points of T are linearly independent.*

Proof. Claim 1 is a lifted version of claim 2 and so we only consider the latter one.

Suppose T is a subset of S such that $p \in Cone(T)$ and the size $|T|$ is as small as possible. Let

$$p = \sum_{s \in T} c_s s, c_s > 0. \tag{1.30}$$

(why?) To finish the proof, it suffices to show that $|T| \leq dim(span(S))$ (why?). If this is not the case, then T is linearly dependent and so for some coefficients d_s which are not all zeros,

$$\sum_{s \in T} d_s c_s s = 0. \tag{1.31}$$

(why?) We can assume that $\max_{s \in T} d_s = d_{s_*} > 0$ for some $s_* \in T$ (why?). Rewrite Eq. (1.31) as

$$c_{s_*} s_* = \sum_{\substack{s \in T \\ s \neq s_*}} -\frac{d_s}{d_{s_*}} c_s s.$$

Now substitute this expression for $c_{s_*} s_*$ into Eq. (1.30). This eliminates the appearance of s_* in the sum, and keeps all other coefficients nonnegative, considering that $c_s - \frac{d_s}{d_{s_*}} c_s \geq 0$ for each $s \in T$. This contradicts the choice of T and thus we are done. □

Exercise 1.37. Prove claim 1 of Theorem 1.36.

Exercise 1.38. Let $S \subseteq \mathbb{R}^d$ and u an interior point of $Conv(S)$. Prove that we can choose $2d$ points $v_1, \cdots, v_{2d} \in S$ such that u lies in the interior of $Conv(v_1, \cdots, v_{2d})$. (The interior point here should be understood as a topological concept. But you can check that this coincides with its algebraic counterpart which we defined previously.)

Say what you know, do what you must, come what may. – Sofia Kovalevskaia (1850–1891), Motto on her paper "On the Problem of the Rotation of a Solid Body about a Fixed Point."

An interesting book about Kovalevskaia: *A Convergence of Lives: Sofia Kovalevskaia: Scientist, Writer, Revolutionary (Lives of Women in Science)*, Rutgers University Press, 1993.

We define the `polar` of $\mathcal{C} \subseteq R^m$ to be $\mathcal{C}^{\maltese} = P(\mathcal{C}, 0) = \{y \in \mathbb{R}^m : \; y^\top x \leq 0, \forall x \in C\}$.

Example 1.39. Let P be the set of all positive semi-definite matrices in $\mathbb{R}^{m \times m}$. Prove that P is a convex cone and $P^{\maltese} = -P$. (Here, you should identify $\mathbb{R}^{m \times m}$ with $\mathbb{R}^{m^2}$.)

Proof. If A and B are both members of P, we can write $B = QQ^\top$ and henceforth $TrAB^\top = TrAQQ^\top = TrQ^\top AQ \geq 0$, implying $-P \subseteq P^{\maltese}$. For the reverse direction, take any $-A \in P^{\maltese}$. For any $x \in R^m$, construct the matrix $B = xx^\top \in P$. According to the definition of $P^{\maltese}$, we have $x^\top Ax = TrAxx^\top = TrAB^\top \geq 0$. This verifies $A \in P$ and the proof is ended. $\square$

Exercise 1.40. The set $P = \{\begin{pmatrix} t \\ x \end{pmatrix} \in \mathbb{R}^{m+1} : \; t \in \mathbb{R}, x \in \mathbb{R}^m, t \geq |x|\}$ is a convex cone, called the second-order cone. It holds $P^{\maltese} = -P$.

Theorem 1.41. *For any $\mathcal{A} \subseteq \mathbb{R}^m$, it holds $P(\mathcal{A}, 0)^{\maltese} = Cone(\mathcal{A})$ if and only if $Cone(\mathcal{A})$ is closed.*

Proof. Since $P(\mathcal{A}, 0)^{\maltese}$ is closed, the necessity part of the claim is obvious. We proceed to establish the sufficiency part by proving $Cone(\mathcal{A}) \subseteq P(\mathcal{A}, 0)^{\maltese}$ and $\mathbb{R}^m \setminus Cone(\mathcal{A}) \subseteq \mathbb{R}^m \setminus P(\mathcal{A}, 0)^{\maltese}$.

Let a be any element of $Cone(\mathcal{A})$, say $a = \sum_{i=1}^{n} \gamma_i A^i$, $\gamma_i > 0$, $A^i \in \mathcal{A}$. Then for any $x \in P(\mathcal{A}, 0)$, we have $a^\top x = \sum_{i=1}^{n} \gamma_i A^{i\top} x \leq 0$. This means $a \in P(\mathcal{A}, 0)^{\maltese}$.

Now consider a $b \notin Cone(\mathcal{A})$. By Theorem 1.6, there is a dual certificate y such that

$$y^\top b > 0 \tag{1.32}$$

and

$$y^\top a \leq 0 \tag{1.33}$$

for any $a \in Cone(\mathcal{A})$. Especially, Eq. (1.33) gives $y^\top A \leq 0$ for any $A \in \mathcal{A}$. But this just amounts to saying $y \in P(\mathcal{A}, 0)$. Therefore, by further noting Eq. (1.32), we deduce that $b \notin P(\mathcal{A}, 0)^{\maltese}$, finishing the proof. □

Note that Theorem 1.41 is an oriented version of the dual relation Eq. (1.9). Also not that Theorem 1.41 is a converse of Lemma 1.7.

Exercise 1.42. Deduce Theorem 1.6 from Theorem 1.41.

Theorem 1.43. *Let* $\mathcal{C} = Cone(\mathcal{A})$ *for some* $\mathcal{A} \subseteq \mathbb{R}^m$*. Then* $\mathcal{C}^{\maltese\maltese} = \mathcal{C}$ *if and only if* $\mathcal{C}$ *is closed.*

Proof. Theorem 1.41 says that $\mathcal{C}$ is closed if and only if $P(\mathcal{A}, 0)^{\maltese} = \mathcal{C}$. But Lemma 1.7 asserts that $P(\mathcal{A}, 0) = P(\mathcal{C}, 0)$, which is just $\mathcal{C}^{\maltese}$ by definition. Combining these two facts we obtain the result, as desired. □

Exercise 1.44. Prove the following duality relation for $\cap$ and $+$ with respect to the taking polar operation $^{\maltese}$: For any two closed cones A and B, it holds

$$\begin{cases} (A \cap B)^{\maltese} = A^{\maltese} + B^{\maltese}; \\ (A + B)^{\maltese} = A^{\maltese} \cap B^{\maltese}. \end{cases} \tag{1.34}$$

(hint: Use Theorem 1.43.)

The `zero/nonzero pattern matrix` of a matrix A, denoted by $pat(A)$, is a $(0, 1)$ matrix of the same size with A whose i, j-entry is zero if and only if the corresponding entry of A is zero. Let J_n be the $n \times n$ matrix of all ones.

Exercise 1.45. [5]Let A and B be two $(0, 1)$ matrices such that $A + B = J_n$. If $X \in \mathbb{R}^{n \times n}$ has the property that $TrXY^{\top} \leq 0$ for each positive semi-definite matrix Y with $pat(Y) \leq B$, then $X = V - U$ where U is positive semi-definite and $pat(V) \leq A$. (hint: Exercise 1.44 and Example 1.39.)

[5]This fact is used in establishing the fascinating theory of the Lovász number of a graph; see, for example, the excellent survey paper of Donald E. Knuth [22].

Besides Theorem 1.43, it seems that the combination of Lemma 1.7 and Theorem 1.41 asserts more, namely $P^{\maltese\maltese} = P$ for any $P = P(\mathcal{A}, 0)$. Indeed, these "two" results are talking about the same thing from different viewpoints. As a relevant result, let us mention the statement of Krein-Milman Theorem from functional analysis: In a locally convex topological vector space, every compact convex set is the closed convex hull of its extreme points. (I do not mean that Krein-Milman Theorem is too hard to prove in this course. It is a very intuitive result and holds in a very general SPACE. It is the definition of that SPACE that you need to open some textbook on functional analysis. Equipped with the definition, you are invited to prove the result yourselves.)

For a treatment of the material in this chapter which does not avoid the use of some "advanced" terminology and does show you many beautiful figures and why those "advanced mathematics" are indeed "obvious" and "elementary", please go to

Marcel Berger (1927–), Geometry, Vol. 3: Convex Polytopes, Regular Polyhedra, Areas and Volumns, Cedic/Fernand Nathan, 1979. (A wonderful Chinese translation is available in our Minhang library)

ZHOU Kexi, a translator of the above 5-volume book of Berger, was formerly a professor of mathematics in East China Normal University (ECNU) and is now a famous translator of French literature and an editor of Shanghai Translation Publishing House. He will give a public talk entitled From Mathematics to Literature Translation during the anniversary of ECNU this year.

Results dealing with polyhedra often follow easily from those on polyhedral cones but they only add details rather than improve insights.

Let $P \subseteq \mathbb{R}^m$ be a polyhedron. The `dual` of P is $P^* = \{y \in \mathbb{R}^m : y^\top x \leq 1, \forall x \in P\}$.

Exercise 1.46. A polyhedron P contains the origin if and only if $P^{**} = P$.

A `completely positive matrix` is a matrix which can be expressed as a sum of several nonnegative symmetric rank one matrices. We denote by CP_n the set of all n by n completely positive matrices.

Exercise 1.47. $A \in CP_n$ if and only if there is a nonnegative matrix B of n columns such that $A = B^\top B$.

Exercise 1.48. Show that for $n > 1$ there exist positive definite matrices in $\mathbb{R}_+^{n\times n} \setminus CP_n$.

Let $COP_n = \{A \in \mathbb{R}^{n\times n} : \ x^\top Ax \geq 0, \forall x \in \mathbb{R}_+^n\}$ and call the members of COP_n `copositive matrices`.

Exercise 1.49. Both COP_n and CP_n are closed convex cones in $\mathbb{R}^{n\times n}$. Furthermore, let $\mathcal{C}_n = \{-A : \ A \in COP_n\}$, then $\mathcal{C}_n^{\maltese} = CP_n$ and $\mathcal{C}_n = CP_n^{\maltese}$. (hint: $\mathcal{C}_n = CP_n^{\maltese}$ is easy. Then you may appeal to Theorem 1.43.)

For the role of copositive matrices in the study of computer aided geometric design (CAGD), please go to this accessible little book written for students in high school and university who love mathematics: CHANG, Gengzhe, The Mathematics of Surfaces, Hunan Education Press, 1995.

For some relevant material on Exercise 1.49, you can consult the classic text of Marshall Hall, Jr. (1910-1990), Combinatorial Theory, whose second expanded edition appeared in 1986 and whose first edition appeared in 1967.

Note that Donald E. Knuth, the master of The Art of Computer Programming, earned a Ph.D in mathematics under the supervision of Marshall Hall, Jr., at the California Institute of Technology in 1963. You may also notice with possibly a bit surprise that the title of the PhD dissertation of Knuth is "Finite Semifields and Projective Planes". (Oh, but where is art and where is computer?)

Science is what we understand well enough to explain to a computer. Art is everything else we do. – D.E. Knuth (January 10,1938–)

The combination of the following two theorems is often called the Minkowski-Weyl's Theorem.

Theorem 1.50. *(Minkowski's Theorem) Let* $P = P(A, c)$*, where* $A \in \mathbb{R}^{n\times m}, c \in \mathbb{R}^n$*. Then* $P = Conv(\mathcal{A}) + Cone(\mathcal{B})$ *for some finite set* $\mathcal{A}, \mathcal{B} \subseteq \mathbb{R}^m$*.*

Theorem 1.51. *(Weyl's Theorem) Let* $P = Conv(\mathcal{A}) + Cone(\mathcal{B})$ *for some finite set* $\mathcal{A}, \mathcal{B} \subseteq \mathbb{R}^m$*. Then there are* $A \in \mathbb{R}^{n\times m}, c \in \mathbb{R}^n$ *such that* $P = P(A, c)$*.*

Corollary 1.52. P *is a polytope (bounded polyhedron) if and only if it is the convex hull of a finite number of points (consisting of its vertices).*

Corollary 1.52 will be of frequent use later; check the proof of Theorems 3.5, 5.3, 6.1, 6.19, etc.. Rockafellar once made a remark on it:

This classical result is an outstanding example of a fact which is completely obvious to geometric intuition, but which wields important algebraic content and is not trivial to prove.

Rather than deal with such awkwardness, it's better to move to cones.

Theorem 1.53. *(Minkowski-Weyl's Theorem for conical polyhedron) For a subset P of $\mathbb{R}^m$, the following statements are equivalent:*

1. $P = P(A, 0)$ *for some matrix A of m columns;*

2. $P = Cone(B)$ *for some matrix B of m rows.*

Exercise 1.54. Use dehomogenization to deduce Minkowski-Weyl's theorem, namely Theorems 1.50 and 1.51, from Theorem 1.53.

Perhaps the philosophically most relevant feature of modern science is the emergence of abstract symbolic structure as the hard core of objectivity behind –as Eddington puts it–the colorful tale of the subjective storyteller mind. The combinatorics of aggregates and complexes deals with some of the simplest such structures imaginable. It is gratifying that combinatorial mathematics is so closely related to the philosophically important problems of individuation and probability, and that it accounts for some of the most fundamental phenomena in inorganic and organic nature. – Hermann Klaus Hugo Weyl (1885–1955), whose Ph.D advisor is David Hilbert.

The mathematical education of the young physicist [Albert Einstein] was not very solid, which I am in a good position to evaluate since he obtained it from me in Zurich some time ago. – Hermann Minkowski (1864–1909), a lifelong friend of David Hilbert.

We call A a `representation matrix` of $P(A,0)$ and B a `generating matrix` of $Cone(B)$. A pair (A,B) of real matrices is said to be a `double description pair` or simply a DD pair if it holds $P(A,0) = Cone(B)$.

Exercise 1.55. (A,B) is a DD pair if and only if $(B^\top, A^\top)$ is a DD pair. (hint: Lemma 1.7 and Theorem 1.41)

Exercise 1.55 means that the two implications of Theorem 1.53 are equivalent (why?). So Theorem 1.53 comes from

Theorem 1.56. *(Weyl's Theorem for conical polyhedron) For each finite set $\mathcal{B} \subseteq \mathbb{R}^m$, there is positive integer n and $A \in \mathbb{R}^{n\times m}$ such that $Cone(\mathcal{B}) = P(A,0)$.*

If there is a problem you can't solve, then there is an easier problem you can solve: find it.– George Pólya (1887–1985)

Theorem 1.56 follows readily from

Theorem 1.57. *Let* $\mathcal{C} = \{v_1, \cdots, v_n\} \subseteq \mathbb{R}^m$ *and* $Span(\mathcal{C}) = \mathbb{R}^m$*. If* $b \notin Cone(\mathcal{C})$*, then there are* $m-1$ *linearly independent vectors from* $\mathcal{C}$*, say* $v_1, \cdots, v_{m-1}$*, such that we can take* $y \in Span(v_1, \cdots, v_{m-1})^{\perp}$ *such that* $b^{\top}y < 0$ *and* $x^{\top}y \geq 0$ *for all* $x \in Cone(\mathcal{C})$*.*

An equivalent formulation of Theorem 1.57 is

Theorem 1.58. *Let* $\mathcal{C} = \{v_1, \cdots, v_n\} \subseteq \mathbb{R}^m$ *and* $Span(\mathcal{C}) = \mathbb{R}^m$*. If* $b \notin Cone(\mathcal{C})$*, then there is a vector* y *such that* $b^{\top}y < 0$*,* $v_i^{\top}y \geq 0$*,* $i = 1, \cdots, n$*, and there are at least* $m-1$ *linearly independent vectors, say* $v_1, \cdots, v_{m-1}$*, such that* $v_i^{\top}y = 0$*,* $i = 1, \cdots, m-1$*.*

Exercise 1.59. Prove Theorem 1.57 (or Theorem 1.58). (Consider the nearest hyperplane spanned by $m-1$ vectors from $\mathcal{C}$ to b.)

Recall what we have gone through:

Theorem 1.4 $\dashrightarrow$ Theorem 1.6 $\dashrightarrow$ Theorem 1.57

A mathematician who can only generalize is like a monkey who can only climb up a tree, and a mathematician who can specialize is like a monkey who can only climb down a tree. In fact neither the up nor down monkey is a viable creature. A real monkey must find food and escape his enemies and so he must incessantly climb up and down, up and down. A real mathematician must be able to generalize and specialize. – Donald D. Spencer, a mathematics professor at Princeton, whose Ph.D. advisors are John Littlewood and G.H. Hardy.

We have now finished the proof of the Minkowski-Weyl's Theorem, namely the equivalence of the implicit representation and generator representation for a polyhedron.

The last step of our approach is to verify Weyl's Theorem for conical polyhedron. Many books adopt the approach of proving Minkowski's Theorem first. To do so, one often need to prepare some kinds of Krein-Milman Theorem as we mentioned before (often appeared in some "elementary" form).

Note that Minkowski's Theorem can be proved constructively by a simple strongly polynomial algorithm, while no such proof for Weyl's Theorem is known and every known proof essentially relies on some proof of Farkas' Lemma.

Minkowski-Weyl's Theorem, also known as the Representation Theorem for polyhedra, or Decomposition Theorem of Polyhedra, is the basic theorem of convexity theory. Indeed, it is equivalent to the both easy-to-understand and easy-to-prove Farkas Lemma.

Here is a very nice small book in Chinese:

SHI, Shuzhong, Convexity, Hunan Education Press, 1998.

It addresses the topic of convexity from some other viewpoints and restricts to the essential and intuitive two dimensional case. Prof. Shi includes in his small book lots of wonderful comments, not only those having direct connection with the discussion of convexity and its application in economics. Note that this book is in the same series as the book written by Prof. Chang.

Additional homework:

Exercise 1.60. Work out [31, 2.11,2.12,2.13,2.14,2.22,2.28] and then try to organize your results together with some possible extensions you deduce yourself or learn from some other sources into a report on polyhedral theory and its applications.

Exercise 1.61. An interval of $\{1, 2, \cdots, n\} = [n]$ is a set of the form $\{a, a+1, \cdots, b\}$, $1 \leq a \leq b \leq n$. Let $x_1, \cdots, x_n$ be positive numbers. Then there exists a subset $S \subseteq [n]$ such that

$$\sum_{i \in S} \frac{1}{x_i} \leq n, \tag{1.35}$$

and for every interval I of $[n]$ with $I \cap S = \emptyset$,

$$\sum_{i \in I} x_i < 1. \tag{1.36}$$

(hint: Use Theorem 1.35, [31, 2.22]; See S.H. Schanuel, A combinatorial problem of Shields and Pearcy, *Proc. Amer. Math. Soc.* 65 (1977), 185–186.)

Exercise 1.62. (W. Blaschke, 1916) The intersection of half-spaces $\{x : \ x^\top a \leq d\}$ and $\{x : \ x^\top a \geq c\}$ such that $d \geq c$ is called a slab of thickness $\frac{d-c}{|a|}$. Prove that if a bounded convex set in R^n is contained in no slab of thickness less than t, then, for every positive ϵ, this set contains a ball of radius $\frac{t-\epsilon}{n+1}$. (hint: Chapter 17 of Vašek Chvátal, Linear Programming, W.H. Freeman and Company, New York, 1983.)

Exercise 1.63. A polyhedron $P \subseteq \mathbb{R}^m$ is of `blocking type` if $P \subseteq \mathbb{R}^m_+$ and if $y \geq x \in P$ implies $y \in P$. For a polyhedron $P \subseteq \mathbb{R}^m$, define its `blocking polyhedron` $B(P)$ by $B(P) = \{z \in \mathbb{R}^m_+ : z^\top x \geq 1, \forall x \in P\}$. Prove that for a polyhedron P of blocking type, it holds

(1) $B(P)$ is again a polyhedron of blocking type;

(2) $B(B(P)) = P$.

Polyhedra, linear inequalities and linear programming can be seen as three views of the same concept. Polyhedra represent a geometrical point of view, linear inequalities the algebraic point of view, and linear programming the optimization point of view.

Should I refuse a good dinner simply because I do not understand the processes of digestion? — Oliver Heaviside (1850-1925)

All the following books, except the last one, are available in our library and I strongly recommend you have a look into some parts of them and wish you enjoy the reading.

Further readings:

- Alexander Barvinok, A Course in Convexity, American Mathematical Society, 2002.
- George L. Nemhauser, Laurence A. Wolsey, Integer and Combinatorial Optimization (Wiley Interscience Series in Discrete Mathematics and Optimization), John Wiley & Sons Inc, 1988.
- Christos H. Papadimitriou, Kenneth Steiglitz, Combinatorial Optimization : Algorithms and Complexity, Dover Publications, 1998.
- Bernhard Korte, Jens Vygen, Combinatorial Optimization: Theory and Algorithms (Algorithms and Combinatorics , Vol. 21), 2nd ed., Springer, 2002.
- Alexander Schrijver, Combinatorial Optimization: Polyhedra and Efficiency (3 volume-set) (Algorithms and Combinatorics, Vol. 24), Springer, 2003.
- Jean H. Gallier, Geometric Methods and Applications: For Computer Science and Engineering, Springer, 2000.

Simplicity, intuitive appeal and universality of applications make teaching convexity (and writing a book on convexity) a rather gratifying experience. – Alexander Barvinok (1963–)

Everybody who wants to be graded in this course is expected to submit to me some notes, solutions to exercises, comments, or anything else having to do with this course, after the long National Day Holiday.

Be glad of life because it gives you the chance to love, to work, to play, and to look up at the stars. – Henry Van Dyke

HAPPY MOON FESTIVAL!

2 Integer Programming, Totally Unimodular Matrices and Hoffman-Kruskal Theorem

The first rule of discovery is to have brains and good luck. The second rule is to sit tight and wait till you get a bright idea. – George Pólya

Let $c \in \mathbb{R}^m, b \in \mathbb{R}^n, h \in \mathbb{R}^p, A \in \mathbb{R}^{m \times n}, G \in \mathbb{R}^{m \times p}$.

A `linear mixed-integer programming problem` (MIP) is a problem of the form

$$\max\{b^\top x + h^\top y : \; Ax + Gy \leq c, x \in \mathbb{Z}_+^n, y \in \mathbb{R}_+^p\}.$$

A `linear (pure) integer programming problem` (IP)

$$\max\{b^\top x : \; Ax \leq c, x \in \mathbb{Z}_+^n\},$$

is a special case of MIP in which there are no continuous variables.

A `linear programming problem` (LP)

$$\max\{h^\top y : \; Gy \leq c, y \in \mathbb{R}_+^p\},$$

is a special case of MIP in which there are no integer variables.

LP is solvable in polynomial time while IP is $\mathcal{NP}$-hard in general. So, it is of great interest to know for which polyhedra an IP collapses to an LP. (Even when an IP cannot be reduced to an LP, it is still useful to relax the IP to some LP and get bounds for the original IP problem.)

Define a polyhedron P to be `integral` if for each vector b for which

$$\max_{x \in P} b^\top x$$

is finite, the maximum is attained by some integer vector. In other word, when discussing on an integral polyhedron, there is no difference between LP (optimization problem with continuous variables) and IP (combinatorial optimization problem).

Let P be a convex set. The `integer hull` of P is the convex hull of the set of integer points in P. The attention on integer points corresponds to our interest of solving IP utilizing what we know about tackling LP.

Exercise 2.1. A polyhedron P is integral if and only if the integer hull of P is P itself.

Exercise 2.2. A polyhedron is integral if and only if each of its minimal faces contains integral points. Especially, for $c \in \mathbb{R}^m$ and $A \in \mathbb{R}^{m \times n}$ with $\operatorname{rank} A = n$, $P(A, c)$ is integral if and only if all vertices of $P(A, c)$ are integer vectors. (In the process of proving Theorem 2.6 we essentially reprove this trivial result.)

A `unimodular matrix` is a square integer matrix whose determinant has magnitude 1. Under multiplication, the set of unimodular matrices of order n forms a group which corresponds to the group of all linear transformations which preserve $\mathbb{Z}^n$, the integer lattice in $\mathbb{R}^n$.

Exercise 2.3. Each unimodular matrix is a product of some unimodular elementary matrices. (For the definition of elementary matrices, please look into this very beautifully written textbook: Nathan Jacobson, Basic Algebra I. Nathan Jacobson (1910–1999) was the president of the American Mathematical Society from 1971 to 1973. To follow our proof of Theorem 4.5, you should know what is an elementary operation–an operation by an elementary matrix.)

A matrix is `totally unimodular` (TU) if each of its nonsingular square submatrices is unimodular.

Be aware that not every totally unimodular matrix is unimodular and not every unimodular matrix is totally unimodular!

Lemma 2.4. *Let A be an $n \times m$ TU matrix. Then each of $-A, A^\top, \begin{pmatrix} A & I_n \end{pmatrix}, \begin{pmatrix} A \\ I_m \end{pmatrix}, \begin{pmatrix} A \\ -A \end{pmatrix}$ and $\begin{pmatrix} A & -A \end{pmatrix}$ is also a TU matrix.*

Exercise 2.5. Prove Lemma 2.4.

Theorem 2.6. *Let A be a totally unimodular $m \times n$ matrix and let $c \in \mathbb{Z}^m$. Then the polyhedron $P = P(A, c) = \{x \in \mathbb{R}^n : \ Ax \leq c\}$ is an integral polyhedron.*

Proof. Let $b \in \mathbb{R}^n$ and suppose $x = x^*$ is an optimum solution of

$$\max_{x \in P} b^\top x.$$

Choose $d_1, d_2 \in \mathbb{Z}^n$ such that $d_1 \leq x^* \leq d_2$. Then the polyhedron $Q = P(B, \alpha) \subseteq P$ is nonempty, where $B = \begin{pmatrix} A \\ -I_n \\ I_n \end{pmatrix}$ and $\alpha = \begin{pmatrix} c \\ -d_1 \\ d_2 \end{pmatrix}$. Because B is of full column rank, Corollary 1.33 says that Q has a vertex. (Indeed, you can see that Q is a bounded polyhedron and hence Corollary 1.52 further implies that Q is the convex hull of its vertices.) Now, by appealing to Corollary 1.35 we find that there is a vertex of Q and hence a point lying in P, say y, such that $b^\top y = \max_{x \in P} b^\top x$. We can end the proof by demonstrating that $y \in \mathbb{Z}^n$. Note that $B_{\alpha,y} y = \alpha^{=}_{B,y}$ while Theorem 1.32 tells us that the integer matrix $B_{\alpha,y}$ has full column rank. So there are $B' \in \mathbb{Z}^{n \times n}$ and $\alpha' \in \mathbb{Z}^n$ such that $B'y = \alpha'$ and $\det B' \neq 0$. However, it is easy to see from Lemma 2.4 that B is TU and thus $\det B' = \pm 1$. This guarantees that $y = B'^{-1}\alpha'$ is an integer vector, as desired. $\square$

Exercise 2.7. Let A be an $m \times n$ TU matrix. Let $c \in \mathbb{Z}^m$ and $u \in \mathbb{Z}^n$. Then each of the following polyhedra is integral:

(1) $P(A, c)$;

(2) $P(A, c) \cap \mathbb{R}^n_+$;

(3) $P^=(A, c)$;

(4) $P^=(A, c) \cap \{x : \ x \leq u\}$.

(hint: Lemmas 2.4 and 2.6)

Example 2.8. [6](Communicated by Bert Gerards) Let $A = \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix}$. Then A is not totally unimodular. But for every $c \in \mathbb{Z}^2$ such that $P(A,c) = \{x : Ax \leq c\} \neq \emptyset$, $P(A,c)$ has exactly one vertex and that vertex is integral. This means the converse of Theorem 2.6 is not true.

We mention that it was proved in [19] that the following two assertions are equivalent for $A \in \mathbb{Z}^{m\times n}$:

- For each $c \in \mathbb{Z}^m$ the polyhedron $P(A,c)$ is integral.
- For every set S of rows of A, let B be the submatrix of A formed by S and $|S| = s$, then $\mathcal{D}_s(B)$ is either 0 or 1. Here we use the symbol $\mathcal{D}_s(B)$ to refer to the greatest common divisor of all $s \times s$ minors of B; see page 106.

Exercise 2.2 says that a polyhedron with vertices is integral if and only if all its vertices are integral. In addition, Theorem 1.32 asserts that a point is a vertex of a polyhedron if and only if it is an extreme point of the polyhedron. These two facts will make it easier to extract information from the assertion that a polyhedron is integral. Note that for any $A \in \mathbb{Z}^{m\times n}$ and any $c \in \mathbb{R}^m$, although $P(A,c)$ may have no vertices, the polyhedron $P(A,c) \cap \mathbb{R}^n_+ = P(\begin{pmatrix} A \\ -I_n \end{pmatrix}, \begin{pmatrix} c \\ 0 \end{pmatrix})$ always has. Our next result says that the converse of Theorem 2.6 is true when we restrict to those matrices having $-I_n$ as a submatrix.

[6]Alexander Schrijver comments that any integer non-totally-unimodular matrix with determinant 1 can act as such an example.

Theorem 2.9. *(Hoffman-Kruskal Theorem, 1956) Let $A \in \mathbb{Z}^{m\times n}$. Then A is totally unimodular if and only if the polyhedron $P(A,c)\cap \mathbb{R}^n_+ = \{x:\ Ax \le c, x \ge 0\}$ is integral for each $c \in \mathbb{Z}^m$.*

Proof. The necessity is immediate from Lemma 2.4 and Theorem 2.6. (why?)

To derive the sufficiency, it is clear that we only need to show that each nonsingular square submatrix of A has determinant ± 1. This in turn reduces to the task of establishing that each $m \times m$ nonsingular submatrix of $B = \begin{pmatrix} A & I_m \end{pmatrix}$ has determinant ± 1 (why?).

As a preparation, we first claim that for any choice of $c \in \mathbb{Z}^n$, each vertex (if any) of the polyhedron $P^=(B,c)$, say $z = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}$, where $z_1 \in \mathbb{R}^n_+$ and $z_2 \in \mathbb{R}^m_+$, is integral. In virtue of our remarks preceding the statement of Theorem 2.9, our assumption is that all extreme points of $P(A,c)\cap\mathbb{R}^n_+$ are integral and z is an extreme point of $P^=(B,c)$. We will turn to prove $z_1 \in \mathbb{Z}^n$, from which $z_2 = c - Az_1 \in \mathbb{Z}^m$ follows and hence our claim that $z \in \mathbb{Z}^{m+n}$ is verified. Observe that $c - Az_1 = z_2 \ge 0$. This means $z_1 \in P(A,c)\cap\mathbb{R}^n_+$. Consequently, due to our assumption, it is enough to prove that z_1 is an extreme point of $P(A,c)\cap\mathbb{R}^n_+$. Suppose, otherwise, that there are two other points $v, w \in P(A,c)\cap\mathbb{R}^n_+$ such that $z_1 = \frac{1}{2}(v+w)$. Then $z_2 = c - Az_1 = \frac{1}{2}(c-Av) + \frac{1}{2}(c-Aw)$. Hence, $z = \frac{1}{2}\begin{pmatrix} v \\ c-Av \end{pmatrix} + \frac{1}{2}\begin{pmatrix} w \\ c-Aw \end{pmatrix}$, which contradicts the assumption that z is an extreme point of $P^=(B,b)$. This proves our claim.

We are ready to verify that each $m \times m$ nonsingular submatrix C of B has determinant ± 1. This can be achieved by showing that $C^{-1} \in \mathbb{Z}^{m \times m}$ (why?). The latter instead is equivalent to the fact that $C^{-1}v \in \mathbb{Z}^m$ for each $v \in \mathbb{Z}^m$ (indeed, it is enough to check m such vectors which form a basis of $\mathbb{Z}^m$) For any $v \in \mathbb{Z}^m$, we can take

$$u \in \mathbb{Z}^m \tag{2.1}$$

such that

$$z = u + C^{-1}v > 0. \tag{2.2}$$

Let $c = Cz = Cu + v \in \mathbb{Z}^m$. Without loss of generality, assume that C consists of the first m columns of B, namely $B = \begin{pmatrix} C & D \end{pmatrix}$. Correspondingly, put $z' = \begin{pmatrix} z \\ 0 \end{pmatrix} \in \mathbb{R}^n_+$. We can check that z' is a vertex of $P^=(B, c)$ (Exercise 2.10) and hence our former claim gives $z \in \mathbb{Z}^m$. Noting Eqs. (2.1) and (2.2), we finish the proof finally. □

Exercise 2.10. Do the checking left to you in the proof of Theorem 2.9, namely z' is a vertex of $P^=(B, c)$.

Exercise 2.11. (Integral version of Theorem 1.20) An integral matrix A is totally unimodular if and only if for all integral vectors b and c both optima in the LP duality below are attained by integral vectors (if they are finite):

$$\max\{b^\top x : \ Ax \le c, x \ge 0\} = \min\{y^\top c : \ y^\top A \ge b^\top, y \ge 0\}.$$

Alan J. Hoffman is a fellow of International Business Machines Corporation since 1978 and a member of National Academy of Sciences since 1982. His principal areas of research are linear algebra, linear programming, and combinatorics. Many of his classic papers are collected into [28]. Especially, you can find there the reprinted paper [19] in which our Theorem 2.9 was born.

It fits the theorem into a larger context that enables us to know it better. – A.J. Hoffman (May 30, 1924–)

The applicability of polyhedral methods in combinatorial optimization often comes down to our ability to prove the integrality of various kinds of polyhedra. In view of Theorem 2.9, it is important to be able to efficiently detect whether or not a given matrix is a TUM.

Exercise 2.12. [31, Exercise 8.8] Let A be a totally unimodular matrix. Show that the columns of A can be split into two classes such that the sum of the columns in one class, minus the sum of the columns in the other class, gives a vector with entries $0, 1$, and -1 only.

Polyhedral combinatorics is the application of the theory of linear systems and linear algebra to combinatorial problems. Even though many of its results are of a pure 'combinatorial' flavour, its roots are in the development of algorithms for combinatorial problems. In fact, it is possible to divide the development of polyhedral combinatorics into three periods, demarcated by major developments in the theory of algorithms. ... In fact, the 'first period' of polyhedral combinatorics mentioned in the introduction was largely a period of study of polyhedra defined by totally unimodular matrices. – W.R. Pulleyblank [29]

We have the following characterization of total unimodularity (For a proof please go to Theorem 2.7 of Chapter III.1, G. Nemhauser and L. Wolsey, Integer and Combinatorial Optimization, Wiley, 1999).

Theorem 2.13. *An $m \times n$ matrix A is a TUM if and only if for every $J \subseteq \{1, 2, \cdots, n\}$ there exists a partition J_1, J_2 of J such that $|\sum_{j \in J_1} A(i,j) - \sum_{j \in J_2} A(i,j)| \leq 1$ for $i = 1, \cdots, m$.*

Unfortunately, Theorem 2.13 only guarantees the existence of a good certificate of showing that a matrix is not unimodular but does not provide an easy means of showing that a matrix is unimodular. This defect is remedied by Seymour's decomposition theorem for regular matroids[7], proved in his paper

P.D. Seymour, Decomposition of regular matroids, *Journal of Combinatorial Theory* B 28 (1980) 305–359,

which is a deep and beautiful result in mathematics and yields a polynomial-time test for total unimodularity. The work of Seymour implies that every TUM can be built by "gluing together" "network matrices" and two fixed 5×5 $(0, \pm 1)$ matrices.

We can only touch on a bit of network matrices in the sequel.

[7]Please refer to page 91 for regular matroid.

Paul Seymour was awarded the 2003 Ostrowski Prize. The following is copied from `http://www.math.uwaterloo.ca/PM_Dept/News/information.shtml`:

Paul Seymour was born in 1950 in England, obtained his D.Phil. from the University of Oxford in 1975, and is currently a professor of mathematics at Princeton University. He received the George Polya Prize in 1983 and the Fulkerson Prize in 1979 and 1994, the second time jointly with Neil Robertson and Robin Thomas. In 1994 he gave a plenary lecture at the International Congress of Mathematicians.

Paul Seymour has enriched mathematics with a number of spectacular results. His work is known not only by all discrete mathematicians but also by most theoretical computer scientists.

For instance, Seymour gave a precise characterization of totally unimodular matrices, a result which is one of the deepest in the theory of matroids. Furthermore, recently Seymour and his student Chudnovsky combined their work with that of Seymour and his close collaborators Robertson and Thomas in order to prove the strong perfect graph conjecture of Berge. Berge's conjecture had stood since 1961 and was one of the most important open problems in graph theory.

The following is copied from SIAM webpage `http://www.siam.org/prizes/an04booklet.htm`:

The 2004 George Polya Prize is awarded jointly to Neil Robertson and Paul Seymour for their magnificent proof of the Wagner Conjecture in the theory of graph minors. Given any infinite sequence of finite graphs, the conjecture states that there are always two graphs in the sequence such that the first is a minor of the second[8]. The Robertson-Seymour proof of this simply stated result is a true tour-de-force spanning twenty important research papers and providing a structural characterization of finite graphs that has deep and far-reaching consequences.

Note that A.C. Yao won the 1987 George Polya Prize.

Here is a paper telling you why you should have interest in graph minors if you think that you have some interest in algorithms: D. Bienstock, M.A. Langston, Algorithmic Implications of the Graph Minor Theorem, a chapter of *Handbook of Operations Research and Management Science*, available at `http://www.cs.utk.edu/~langston/courses/cs594-fall2003/BL.pdf`

[8]For a definition of graph minor, see page 173.

Seymour has wonderful students. We mention here two of them.

Popular Science names Maria Chudnovsky to its 'Brilliant 10' list in 2004. The following is copied from `http://www.dailyprincetonian.com/archives/2004/09/27/news/10856.shtml`:

What merited Chudnovsky such an honor? Along with her adviser, Paul Seymour, Chudnovsky unlocked a major, mathematical 40-year-old riddle known as the perfect graph conjecture. The conjecture explains why some organizational problems, like constructing a cell phone network using the fewest transmitters or assigning teachers to various classrooms, are harder than others. Popular Science credits Chudnovsky and her team as the first to prove this longstanding theory. Russian-born Chudnovsky came to Princeton in 2000 to pursue graduate work after her education in Israel. "I wanted to do my Ph.D. at a very good place and I was interested in combinatorics," Chudnovsky added, who received her doctorate last year. "At college in Israel, I had a very good combinatorics professor, and I liked math in high school, so graph theory was a natural choice for me." "It's like solving a crossword puzzle all day long," she said.

Joshua Greene becomes a Ph.D. student of Seymour since this year. Please find here the Chinese translation of a paper written by him when he was an undergraduate: `http://www.math.sjtu.edu.cn/teacher/wuyk/lovasz.pdf`

To introduce network matrix, we begin with a result proved by Poincaré in 1900.

Theorem 2.14. *Let A be a $(0, \pm 1)$ matrix where each column has at most one 1 and at most one -1. Then A is totally unimodular.*

Proof. Since the assumed property satisfied by A is hereditary, we only need to prove that $\det A \in \{0, \pm 1\}$ when A is of size $k \times k$. When $k = 1$, the assertion is trivial. So, suppose $k \geq 2$ and proceed by induction on k. If A has a column having at most one nonzero, then expanding the determinant along this column we have that $\det A \in \{0, \pm 1\}$ according to our inductive assumption. Otherwise, each column has both a 1 and a -1, then the sum of the rows of A is 0 and hence $\det A = 0$. □

The first result in the combinatorial theory of polytopes is the formula connecting the number of vertices, edges and faces of a three dimensional polytope obtained by Descartes and later, independently, by Euler in 1736. (It is now called Euler's formula.) Poincaré generalized Euler's formula to convex polytopes of any dimension (He did much deeper generalization than what I describe here!) which was a fundamental result of combinatorial topology. These simple and beautiful works lead to rich results in later algebraic topology and modern graph theory.

Exercise 2.15. Let $A = \begin{pmatrix} A_1 \\ A_2 \end{pmatrix}$ be a $(0, \pm 1)$ matrix where each column has at most two nonzero elements and for each column of A which contains two nonzero elements the corresponding column sums of A_1 and A_2 are equal. Prove that A is totally unimodular.

If nature were not beautiful, it would not be worth knowing, and if nature were not worth knowing, life would not be worth living. – Henri Poincaré (April 29, 1854 – July 17, 1912)

A `digraph` Γ is specified by two sets $V(\Gamma)$ and $E(\Gamma)$, called the `vertex set` and the `arc set` of Γ, respectively, and two maps $\mathtt{i}_\Gamma$ and $\mathtt{t}_\Gamma$ both from $E(\Gamma)$ to $V(\Gamma)$, called the `tail operator` and the `head operator`, respectively. Here `i` and `t` refer to initial and terminal, respectively, and you can think of an arc $e \in E(\Gamma)$ as going from $\mathtt{i}_\Gamma(e)$ to $\mathtt{t}_\Gamma(e)$. When there in no confusion, we sometimes write $e = (u, v)$ if e goes from u to v (there may be several of such arcs in general!).The matrix representation M_f of a map f from S to T is a matrix whose rows are indexed by S and whose columns indexed by T and whose (s, t)-entry is 1 if $f(s) = t$ and 0 otherwise. The `incidence matrix` of Γ is $M_{\mathtt{i}_\Gamma} - M_{\mathtt{t}_\Gamma}$ and the `adjacency matrix` of Γ is $M_{\mathtt{i}_\Gamma}^\top M_{\mathtt{t}_\Gamma}$. We often just write $\Gamma = (V, E)$ and assume implicitly that there is some (oriented) incidence relation between V and E.

The forthcoming corollary implies many graph theoretical results, including the well-known max-flow min-cut theorem (Please find its formulation in any graph theory textbook and imagine how can the deduction from Corollary 2.16 be possible).

Corollary 2.16. *The incidence matrix of a digraph is totally unimodular.*

Proof. Take transpose and then use Theorem 2.14. □

A $(0,1)$ matrix is called an `interval matrix` if in each column the 1's appear consecutively. Interval matrices appear as the clique-vertex matrix of interval graphs. Interval graphs arise in linear scheduling problems, the problem of register allocation in computer, analysis of DNA chains, and so on.

Corollary 2.17. *Each interval matrix is totally unimodular.*

Proof. For any $t \times t$ submatrix B of an interval matrix, consider MB, where

$$M = \begin{pmatrix} 1 & -1 & 0 & \cdots & \cdots & 0 & 0 \\ 0 & 1 & -1 & \cdots & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & & \vdots & \vdots \\ \vdots & \vdots & \vdots & & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \cdots & 1 & -1 \\ 0 & 0 & 0 & \cdots & \cdots & 0 & 1 \end{pmatrix}_{t \times t},$$

and then apply Theorem 2.14. $\square$

A digraph is `connected` provided its underlying undirected graph is connected. A `spanning tree` of a connected digraph $\Gamma = (V, E)$ is a digraph (V, E') such that $E' \subseteq E$ and for any proper subset E'' of E', (V, E'') fails to be connected. Let Γ be a connected digraph and Γ' a spanning tree of Γ. The `network matrix` [9] arising from Γ and Γ' is the matrix M whose rows are indexed by $E \setminus E'$ and whose columns are indexed by E', where for $e = (u, v) \in E$ and $e' \in E'$,

$$M(e,e') = \begin{cases} 1, & \text{if the } u\text{–}v \text{ path in } \Gamma' \text{ uses } e' \text{ in the forward direction;} \\ -1, & \text{if the } u\text{–}v \text{ path in } \Gamma' \text{ uses } e' \text{ in the backward direction;} \\ 0, & \text{otherwise.} \end{cases}$$

The above definition describes a network matrix row by row. We can also specify it column by column as follows. Take $e' \in E'$. Since Γ' is a spanning tree of Γ, $(V, E' \setminus \{e'\})$ has exactly two connected components, say, one on the vertex set V_0 and the other on the vertex set V_1 and e' goes from V_0 to V_1. Then associate with each vertex in V_0 a potential 0 and each in V_1 a potential 1. Now you see that for each arc $e \in E \setminus E'$, $M(e, e')$ is just the potential difference across the arc e, namely the potential of $\mathtt{t}(e)$ minus that of $\mathtt{i}(e)$.

The following important result was proved by Tutte in 1965.

Theorem 2.18. *Network matrices are totally unimodular.*

[9] In the literature, it is the transpose of the matrix defined here which is called NETWORK MATRIX. In proving Theorem 2.18 we need to consider the incidence matrix of a digraph. According to our definition of the incidence matrix of a digraph, to use the current definition of network matrix will save me the trouble of writing the symbol $\top$ in the proof. But I find that I have to refer to the transpose later in the exercise and thus I bring to me additional trouble in the course of intending to save a trouble. I decide to change my definition of incidence matrix together with the definition of network matrix in this notes when I have more free time.

We remark that the concept of network matrix also appears in Norman Biggs' algebraic potential theory on graphs. This theory connects several respectable subjects of mathematics to graph theory. Note that many persons, like Henry Whitehead, believe that combinatorics is just "slums of topology" and is quite different from "real mathematics". Yes, the work of Seymour is great and most of them are far away from the so-called "real mathematics". But, we can also find in graph theory many traditional real mathematics, like potential theory and analytic functions in graph theory, which we address in our seminar recently:

L. Lovász: Discrete Analytic Functions: a survey [2000], `http://research.microsoft.com/users/lovasz/analytic.pdf`

N. Biggs, Algebraic potential theory on graphs, Bull. London Math. Soc. 29 (1997), 641–682.

R. Bacher, P. de la Harpe, T. Nagnibeda, The lattice of integral flows and the lattice of integral cuts on a finite graph, Bull. Soc. Math. France, 125 (1997), 167–198.

To prove Theorem 2.18, we still need an additional lemma.

Lemma 2.19. *Let N be the incidence matrix of a connected digraph Γ and let $\widetilde{N}$ be obtained from N by deleting one of its columns. Then $\widetilde{N}$ has full column rank. Moreover, for a set J of rows of $\widetilde{N}$,* $\det \widetilde{N}_J = \pm 1$ *if the arcs corresponding to J form a spanning tree of Γ and* $\det \widetilde{N}_J = 0$ *otherwise.*

Exercise 2.20. Prove Lemma 2.19.

We mention in passing that by using the Binet-Cauchy formula, we can deduce from the combination of Corollary 2.16 and Lemma 2.19 the very useful Matrix-Tree Theorem; please refer to any standard textbook of graph theory. Note that Biggs gave another proof of the Matrix-Tree Theorem using his potential theory on graphs.

Proof. (of Theorem 2.18) Let M be a network matrix arising from the connected digraph $\Gamma = (V, E)$ and its spanning tree $\Gamma' = (V, E')$. Let N be the incidence matrix of Γ and $\widetilde{N}$ be obtained from N by deleting a single column. Similarly, let L be the incidence matrix of Γ' and $\widetilde{L}$ be obtained from N by deleting a single column which corresponds to the same vertex as that column being deleted from N.

By Lemma 2.19 we have $\widetilde{L}$ is unimodular and thus so is $\widetilde{L}^{-1}$. Meanwhile, Theorem 2.14 asserts that the matrix $\begin{pmatrix} \widetilde{N} \\ \widetilde{L} \end{pmatrix}$ is totally unimodular. It then follows that each $(n-1) \times (n-1)$ nonsingular submatrix of $\begin{pmatrix} \widetilde{N} \\ \widetilde{L} \end{pmatrix} \widetilde{L}^{-1} = \begin{pmatrix} \widetilde{N}\widetilde{L}^{-1} \\ I_{n-1} \end{pmatrix}$ is unimodular, where n is the number of vertices of Γ. This implies that $\widetilde{N}\widetilde{L}^{-1}$ is totally unimodular.

Carefully checking the definition of L, M, and N, we find that $ML = N$. This gives us $M\widetilde{L} = \widetilde{N}$ and hence the unimodular matrix $\widetilde{N}\widetilde{L}^{-1}$ is nothing but M. This completes the proof. □

Exercise 2.21. The transpose of an interval matrix is a network matrix.

Exercise 2.22. The transpose of the incidence matrix of a bipartite graph is a network matrix.

A `graph` G is specified by two sets $V(G)$ and $E(G)$, called the `vertex set` and the `edge set` of G, respectively, and one map ∂_G from $E(G)$ to $\{\{u,v\} : \ u,v \in V(G)\}$, called the `boundary operator`. You can think of an edge $e \in E(\Gamma)$ as connecting the pair of (possibly equal) vertices in $\partial_G(e)$. For each $e \in E(G)$, we call the points in $\partial_G(e)$ the `endpoints` of e. The `incidence matrix` [10] of a graph G is the matrix M whose rows are indexed by $E(G)$ and arcs by $V(G)$ and $M(e,v) =$ the number of times of v appeared in $\partial_G(e)$. Sometimes, we simply identify $e \in E(G)$ with $\partial_G(e)$.

A `bipartite graph` is a graph whose vertices can be colored either red or black such that each edge connects two vertices of different colors. We often write $G = (V_1, V_2, E)$ to mean that G is a bipartite graph with vertex set $V_1 \cup V_2$ and edge set E and V_1 and V_2 are disjoint independent sets of G (V_1 and V_2 are the two colored sets). A `cycle` of a graph is one of its subgraphs such that the column sums of whose incidence matrix are all even. A `circuit` is a cycle which is minimum under inclusion. The `length` of a cycle is the number of its arcs. An odd cycle is a cycle whose length is odd.

[10]For any finite set X and any $Y \subseteq X$, the `incidence vector` $\chi^Y \in \mathbb{R}^X$ of Y over X is defined by putting $\chi^Y(x) = 1$ if $x \in Y$ and $\chi^Y(x) = 0$ otherwise. Each row of the incidence matrix of a loopless graph is just the transpose of the incidence vector of the boundary of corresponding edge.

Exercise 2.23. A graph is bipartite if and only if it has no odd circuits.

Let N_k be the matrix obtained from I_k by shifting each row cyclically 1 column to the right.

Exercise 2.24. (1) The eigenvalues of N_k are all those k^{th} roots of unity.

(2) $\det(I_k + N_k) = \begin{cases} 2 & if\ k \text{ is odd;} \\ 0 & if\ k \text{ is even.} \end{cases}$

Exercise 2.22 together with Theorem 2.18 says that the incidence matrix of a bipartite graph is totally unimodular. We can prove it directly as well as prove its converse.

Theorem 2.25. *A graph is bipartite if and only if its incidence matrix is totally unimodular.*

Proof. If a graph G contains an odd circuit, say of length k, then it corresponds to a submatrix I_k+N_k of the incidence matrix of G. By Exercise 2.24, the incidence matrix of G is not totally unimodular.

Conversely, let G be bipartite and we proceed to show that the transpose of its incidence matrix, say A, is TUM. We write $A = \begin{pmatrix} A_1 \\ A_2 \end{pmatrix}$, where the rows of A_1 correspond to those red vertices and the rows of A_2 black vertices. Clearly, Exercise 2.15 applies to give the result, as desired. □

I am extremely busy this week and so I ask Miss Deng come here to represent me. Please first have a look at some remarks below.

I wish to challenge you to excel rather than simply survive.

It is an achievement of you if you can finish the first reading of the book [31] of Schrijver when you are a sophomore. Doing so, you have a new understanding of graph theory, which may be very useful for your later study and research; more importantly, you have much more confidence in yourself, improve your mathematical maturity and find some more enjoyment in your life.

Note that the classroom time for our course is only about 20 hours and I myself know little of graph theory. This means that the help I can provide is very limited. I am here mainly to grade you.

Since you have another course on algorithm this semester and will have a course on computability and complexity theory next semester, I can happily leave the teaching of these materials in [31] to other professors. But to finish the remaining part of the book is still an impossible task for me. You see, I have chosen to focus on the basics of the polyhedral method in graph theory.

To make the best use of classroom time, I will first of all present the full proofs of some representative theorems which might be not so easy for you to have the patience to read through yourself as a beginner, especially considering that they do not look like those parts of graph theory which you are familiar with when you are in high school (after all, you are now in university and you should like to experience many surprises and challenges.).

We have finished Farkas' Lemma and Hoffman-Kruskal Theorem and we will further prove Edmonds' Matching Polytope Theorem (both for bipartite graphs and nonbipartite graphs) and the Matroid Polytope Theorem. (We just prove four theorems in a semester and you should agree with me that this is a piece of cake for your ACM class students.) After finishing these four theorems, we will talk about some interesting applications. I am not sure how much time will be left for this final part of our course and I have not determined the topics to be discussed yet.

If you want to feel easy and meaningful with sitting in my talks, you should have a goal, keep yourself interested and work hard. Not to fail the course is a reasonable goal. A better goal is to develop your ability and develop your courage when facing with unknowns. You will be happier if you have the belief that struggling in our course not only helps you realize the first goal but also helps you attain the second goal. Surely, a very concrete goal is to understand those four theorem I just mentioned, at least remember them and know what they are talking about (and how they can be used)! I have been trying to motivate your interest in this course, especially I am emphasizing the beauty and surprise of the existing connections among graph theory, geometry, algebra and optimization. A better way to achieve it may be to give lots of "applications". Due to the lack of class hours, I have to suggest that you become self-motivated by reading the various applications described in [31] (Have you found the map of our Xuhui district in section 5.3 of [31]?). By now, you are ready to look into [31, Chapter 8] to see how the polyhedral methods can be used to deduce almost all theorems in many popular Chinese graph theory textbooks for undergraduates or even graduates. I also hope that the various contents marked with "Application"s in [31] will interest you. No pains, no gains. It is your responsibility to learn the material (if you want to be graded in this course) and most of this learning must take place outside the classroom. No one should expect to pass a serious course by just sitting in every talk and keeping in terrible silence always.

Let me remind you that there are 4 talks tomorrow afternoon. You are welcome to attend. At least, you will find then that how I enjoy a talk of which I hardly catch up with a sentence. Learning is a process. When you were a new-born baby, I guess, you also could not understand what your parents spoke to you, right?

We should continue with the matching polytope theorem. But, since I have no time to prepare the notes right now, we have to digress a little and invite Miss Deng to talk about the basic concepts of matroids (used when coming to the Matroid Polytope Theorem) Please welcome.

Polyhedral combinatorics provides a means of obtaining min-max theorems which can be used to prove optimality of a solution to a problem. According to [29], it involves a four step process:

- Linearize the problem by representing the feasible objects by vectors (usually their incidence vectors)
- Consider the convex hull of these vectors.
- Obtain a linear system sufficient to define the convex hull.
- Apply LP duality theorem to this linear system in order to obtain a min-max theorem.

For the fourth step. we already prepare the LP duality theorem in Chapter 1 (We will see that LP duality Theorem and its variants are often crucial in completing the third step as well). But, generally, the major difficulties encountered are with the third step. Note that the vectors arising from those feasible objects are often integral vectors, even $(0,1)$ vectors, and so the convex hull turns out to be an integral hull. We have been armed with the Hoffman-Kruskal Theorem now. We will still try to cover some results on Totally Dual Integral (TDI) system and the description of the polytopes arising from matroids. We will illustrate the use of this results through some interesting applications.

Since Miss Deng talked about some basics of matroids, it seems better to finish the matroid polytope theorem first and discuss the matching polytopes (which is an important example to demonstrate the polyhedral methods) after that. Also, I will defer the discussion of TDI to some later parts together with some other applications. You should convince yourself that what you learn here is very useful to solve graph theory (optimization) problems. As I mentioned before, we have very limited class time and you can have a look at corresponding parts of the textbook yourself to find that we are really not too far away from the right way.

3 Matroids

3.1 Matroids and greedy algorithms

3.2 Examples of matroids

Matroids were introduced in 1935 by Hassler Whitney, a Wolf prize recipient in 1982, as a unified abstraction of the notion of linear independence of the sets of columns of a matrix and the sets of cycles of a graph. This is a natural continuation of his work in his Ph.D dissertation entitled "The Coloring of Graphs" written under Birkhoff's supervision in Harvard University. At the same time, Van der Waerden discovered independently in 1937 these structures when axiomatizing the notion of algebraic independence. Since then, matroids have been recognized as an important structure in combinatorial optimization and have been used as a framework for approaching various combinatorial problems.

In last seventies, Gian-Carlo Rota, one author of the first book on matroid theory, mounted a campaign to try to change the name of matroid to "geometry", an abbreviation of "combinatorial geometry", which points towards its connections with geometric configurations in a non-metric environment. Today, both geometry and matroid are in use, and many other terms used by successive discovers of this concept have perished, although the latter predominates.

It is asserted that **'Perhaps the two most fundamental well-solved models in combinatorial optimization are the optimal matching problem and the optimal matroid intersection problem'**; see the following nice-written expository paper:

W.H. Cunningham, Matching, matroids, and extensions, Math. Program., B, 91 (2002), 515–542.

We will only touch on a bit of matroids and the main goal in this chapter is the matching polytope theorem. We leave the discussion of the Edmonds' matroid intersection theorem to some later part of this course (after we have learned what is TDI).

What's in a name? That which we call a rose by any other name would smell as sweet. – William Shakespeare, Romeo and Juliet

One of you wrote an email to me to comment on a proof Miss Deng presented in class last week. This is very good and the better for me will be to get your comments on my own talks.

From that email, I find that Miss Deng has introduced the development of matroid theory based on this online paper: Matroids: The Value of Abstraction, available at `http://www.ams.org/new-in-math/cover/matroids1.html`

Especially, Deng cited the following from the above paper:

This major avenue of investigation, which was already raised by Whitney, is whether or not any matroid is representable over some field F. This question is still an unsolved problem.

Please note that this assertion is not accurate. Please go to any standard textbook on matroid theory and find examples of matroids which can not be represented over any field. Indeed, the long-standing question here should be what is a characterization of those matroids which can be represented over some field. A relevant question is to find good characterizations when the field is some fixed one. Thanks to the hard work of some mathematicians, we have now known nice excluded-minor characterizations for matroids representable over $GF(2)$, $GF(3)$ and $GF(4)$, respectively. Another relevant question is to determine those matroids representable over every field, which we call `regular matroids`. Tutte proved in 1965 that a matroid is regular if and only if it is representable over $GF(2)$ and over some field of characteristic other than two. It is Seymour who came out of a satisfying characterization of regular matroids. As Oxley [11] writes, "Probably the most important theorem ever proved in matroid theory is a deep and important structure theorem for regular matroids due to Seymour. Not only does this theorem provides a beautiful decomposition of regular matroids, but it also has profound implications for combinatorial optimization in that it leads to a polynomial-time algorithm to determine whether a real matrix is totally unimodular."

[11] James Oxley, What is a matroid?, Cubo 5 (2003), 179–218. (Here is a revision: `http://www.math.lsu.edu/~oxley/survey4.pdf`)

In the late sixties of last century, R.T. Rockafellar introduced oriented matroids based on his insight on the combinatorial essence of the LP duality theorem. As we remarked at the beginning of this semester, matroids correspond to linear equations while oriented matroids linear inequalities. There are much fun with oriented matroids but we have no ability to go that far in this course.

In his Ph.D thesis, Whitney must aim at attacking the The Four Color Conjecture (4CC), which is now Four Color Theorem (though some persons still regard it as 4CC as the current proofs all rely on computer programming.) Observe that Whitney and Tutte have coauthored a paper on 4CC: "Kempe chains and the four colour problem", in: Studies in Graph Theory, Part II (ed. D.R. Fulkerson), Math. Assoc. of America, 1975, 378-413. This webpage describes a new proof of the Four Color Theorem and a four-coloring algorithm found by Neil Robertson, Daniel P. Sanders, Paul Seymour and Robin Thomas: `http://www.math.gatech.edu/%7Ethomas/FC/fourcolor.html#Why`

While it has sometimes been said that the four color problem is an isolated problem in mathematics, we have found that just the opposite is the case. The four color problem . . . is central to the intersection of algebra, topology, and statistical mechanics. – An algebraic approach to the planar colouring problem, by Louis Kauffman and H. Saleur, in Comm. Math. Phys. 152 (1993), 565-590.

In the review by G.L. Alexanderson for the book of Robin Wilson, "Four Colors Suffice: How the Map Problem Was Solved", we read the following paragraphs, in which you find the familiar names of Minkowski and Whitney.

Hermann Minkowski heard of the problem, and thinking that he could polish it off in a topology class at Göttingen, he proceeded to outline a proof. It, like so many others, turned out to be defective. In the first half of the twentieth century, those working on it most assiduously were Americans: G.D. Birkhoff, Oswald Veblen, Philip Franklin, Hassler Whitney, among others. It was Birkhoff's contribution that led to the mention in the preface of coloring maps on a honeymoon. Arthur Bernhart at Oklahoma completed Birkhoff's work on maps with rings of six countries. Shortly after Bernhart was married his wife saw Mrs. G. D. Birkhoff at a meeting and asked her, "Did your husband make you draw maps for him to color on your honeymoon, as mine did?" Garrett Birkhoff, G.D. Birkhoff's son, when asked about the story said he thought it is probably apocryphal.

G.D. Birkhoff, we are told, once remarked that almost every great mathematician had worked on the four-color problem at one time or another. We assume he referred to all great mathematicians who lived beyond a certain date. Birkhoff thought that he could solve the problem by investigating the properties of chromatic polynomials. But his work in that direction did not pay off.

3.3 Matroid polytope theorem

We first have a warm-up of what was covered in last two sections. A `matroid` $M = (X, \mathcal{I})$ is an ordered pair consisting of a finite set X and a nonempty family $\mathcal{I}$ of so-called `independent` subsets of X which satisfy the following axioms:

(i) A subset of an independent set is independent;

(ii) For any $S \subseteq X$, all inclusionwise maximal members of $\{I \in \mathcal{I} : I \subseteq S\}$ have the same cardinality, called the `rank` of S and denoted by $r_M(S)$ or simply $r(S)$. A `basis` of $S \subseteq X$ is a set $I \subseteq A$ such that $|I| = r(I) = r(S)$.

(3.1)

A great many combinatorial optimization problems can be formulated as optimization problems over independence systems: let $c = (c_x : x \in X)$ be a vector of real (rational) costs, namely a weight function over X; find $I^* \in \mathcal{I}$ which maximizes $\sum_{x \in I} c_x$ over all $I \in \mathcal{I}$. The most obvious method of attack for such a problem is the so-called `greedy algorithm`. Sort the elements of X into the order $x_1, x_2, \cdots, x_n$ in such a way that $c_{x_i} \geq c_{x_j}$ whenever $i \leq j$. Then construct I^* as follows: Initially, put $I_0^* = \emptyset$; For i going from 1 to n, let $I_i^* = I_{i-1}^* \cup \{x_i\}$ if $I_{i-1}^* \cup x_i \in \mathcal{I}$ and $c_{x_i} > 0$ while let $I_i^* = I_{i-1}^*$ otherwise; Finally take $I^* = I_n^*$.

This naive algorithm does work and this fact is an important characteristic of the matroid structure.

Theorem 3.1. *Let $M = (X, \mathcal{I})$ be an ordered pair consisting of a finite set X and a nonempty family $\mathcal{I}$ of subsets of X which satisfies* (3.1) *(i). Then M is a matroid if and only if for any weight function c over X, the greedy algorithm leads to a set $I^* \in \mathcal{I}$ of maximum weight among all $I \in \mathcal{I}$.*

Let $M = (X, \mathcal{I})$ be a matroid. The `matroid polytope` $P(M)$ of M is, by definition, the convex hull of the incidence vectors of the independent sets of M. So, $P(M)$ is a polytope in $\mathbb{R}^X$. For each $z \in \mathbb{R}^X$ and $U \subseteq X$, we write $z(U)$ [12] for the sum of entries in U, which can be thought of as the weight of U assigned by the weight function z.

Exercise 3.2. Each vector $z \in P(M)$ satisfies

$$z(x) \geq 0, \forall x \in X \tag{3.2}$$

and

$$z(U) \leq r_M(U), \forall U \subseteq X. \tag{3.3}$$

Exercise 3.3. The polyhedron P in $\mathbb{R}^X$ specified by the system of inequalities (3.2) and (3.3) is bounded and hence a polytope.

Exercise 3.4. If $z \in \mathbb{Z}^X$ satisfies formulae (3.2) and (3.3), then z is the incidence vector of some $I \in \mathcal{I}(M)$.

The following theorem was proved by Jack Edmonds in his paper: Matroids and the greedy algorithm, Mathematical Programming, 1 (1971), 127–136.

Theorem 3.5. *(Matroid Polytope Theorem) For any matroid $M = (X, \mathcal{I})$, $P(M)$ is the polyhedron P in $\mathbb{R}^X$ defined by formulae* (3.2) *and* (3.3).

Exercise 3.2 says that $P(M) \subseteq P$. Exercise 3.3 along with Corollary 1.52 means that to prove $P \subseteq P(M)$ it is enough to prove that each vertex of P is the incidence vector of some independence set of M. Exercise 3.4 further reduces this task to demonstrating that each vertex of P is integral (Indeed, we will directly prove that each vertex of P is an incidence vector of some independent set of M and so the result of Exercise 3.4 is not necessary for our proof of the Matroid Polytope Theorem.)

[12] I find that this notation will cause confusion later and so this notation will not be used then. I will change the use of this notation in this part when I find the time.

Proof. (of Theorem 3.5) According to our preceding analysis, it suffices to show that for each weight function $c \in \mathbb{R}^X$ there exists an independent set I^* whose incidence vector z^* is the optimum solution to the primal problem PP of maximizing $c^\top z$ for $z \in P$. We apply the greedy algorithm with respect to the current weight function c to get the I^* and proceed to verify that this is really what we want. To prove the optimality of z^*, our strategy is **to find a feasible solution y^* to the dual problem and check that they have the same objective value**.

The dual linear program to maximizing $c^\top z$ subject to formulae (3.2) and (3.3) is

$$DP : \min \sum_{U \subseteq X} y_U r_M(U)$$

subject to $y_U \geq 0, \forall U \subseteq X$, and $\sum_{x \in U} y_U \geq c(x), \forall x \in X$. (why? In applications, it is very important to be able to recognize and formulate the dual of a given LP in combinatorial languages.)

Recall that the greedy algorithm first sorted the elements of X into an order such that $c_{x_1} \geq c_{x_2} \geq \cdots \geq c_{x_n}$. Let k be the largest i for which $c_{x_i} > 0$ and for $1 \leq i \leq k$ define $A^i = \{x_1, \cdots, x_i\}$. Note that $I^* \subseteq \{x_1, \cdots, x_k\}$. Let $y^* \in \mathbb{R}^{2^X}$ be defined by

$$y^*_U = \begin{cases} c_{x_i} - c_{x_{i+1}}, & \text{if } U = A^i \text{ for } 1 \leq i \leq k; \\ c_{x_k}, & \text{if } U = A^k; \\ 0, & \text{otherwise.} \end{cases}$$

We can check that y^* is feasible to DP. (how?) So, we can end the proof by showing that z^* and y^* have the same objective value (for PP and DP, respectively.):

$c^\top z^* = c(I^*) = \sum_{x \in I^*} c(x) = \sum_{i=1}^{k} c(x_i)(r_M(A^i) - r_M(A^{i-1})) = c(x_k) r_M(A^k) + \sum_{i=1}^{k-1} (c(x_i) - c(x_{i+1})) r_M(A^i) = \sum_{U \subseteq X} y^*_U r_M(U)$.

(The second inequality is due to the rule of greedy algorithm; The third equality comes from summation by parts, an analogue to the integration by parts and sometimes called Abel's lemma–you should be familiar with it if you have had a good calculus course.) □

Edmonds' proof of the Matroid Polytope Theorem has become one of the fundamental methods of polyhedral combinatorics. It illustrates the usefulness of looking at a dual problem.

There are several additional remarks about the proof. First, the last step of the proof can be substituted by citing the Complementary Slackness Theorem (Theorem 1.23) (The current argument indeed suggests to you a way to solving Exercise 1.24.) Second, checking the proof of Theorem 3.5, you will find that it provides a polyhedral combinatorial method of establishing the forward direction of Theorem 3.1 (A simple pure combinatorial proof of Theorem 3.1 was presented last week.). Finally, note that if we consider an integer valued weight function c, then the optimal dual solution y^* constructed in the proof is also integral. This means that the system given by (3.2) and (3.3) is TDI (We will elaborate on this concept later).

You can find in our library the following book dedicated to Jack Edmonds "in appreciation of his ground breaking work that laid the foundations for a broad variety of subsequent results achieved in combinatorial optimization".

Combinatorial Optimization – Eureka, You Shrink! Papers Dedicated to Jack Edmonds, 5th International Workshop, Aussois, France, March 5-9, 2001, Revised Papers Series : Lecture Notes in Computer Science, Vol. 2570, Jünger, Michael; Reinelt, Gerhard; Rinaldi, Giovanni (Eds.) 2003.

To know why the book has such a strange title, please read its preface: `http://www.springeronline.com/sgw/cda/pageitems/document/cda_downloaddocument/0,10900,0-0-45-89155-0,00.pdf`

Anyone who cannot cope with mathematics is not fully human. At best he is a tolerable subhuman who has learned to wear shoes, bathe and not make messes in the house. – Robert A. Heinlein

4 Integral Polyhedra, Total Dual Integrality and Edmonds-Giles Theorem

My mother, who taught kindergarten and first grade before her marriage, said that I was the stubbornest child she had ever known. I would say that my stubbornness has been to a great extent responsible for whatever success I have had in mathematics. But then it is a common trait among mathematicians. – Julia Robinson

We know that a polyhedron is integral if and only if

- each of its minimal faces contains integral points,

which clearly implies

- each of its supporting hyperplanes contains integral points.

It certainly follows from the latter one that

- each of its rational supporting hyperplanes (to be defined shortly) contains integral points,

and thus

- the inner products of each integral vector with the points in the polyhedron either have no finite supremum or achieve the supremum at an integer value.

These four items look weaker and weaker (Try to produce your examples to get a picture of the hierarchy, please.) The main topic of this chapter is to show that in some cases they may be just equivalent. Undoubtedly, this type of facts are very useful in checking the integrality of a polyhedron and hence obtaining combinatorial min-max results.

A system of inequalities $Ax \leq c$ is called a `rational system` provided both A and c are rational. Since we are interested in integral polyhedra, it is natural to expect that by moving into the field of rational systems we can find a bunch of good results. On the other hand, note that most systems of linear equalities arising from combinatorial applications are rational.

A rational system of linear inequalities $Ax \leq c$, say $A \in \mathbb{Q}^{m\times n}$ and $c \in \mathbb{Q}^m$, is called `totally dual integral` (TDI) provided, for all $b \in \mathbb{Z}^n$ such that $\max\{b^\top x : \ Ax \leq c\}$ is finite, the dual $\min\{y^\top c : y^\top A = b^\top, y \in \mathbb{R}^m_+\}$ has an integral optimal solution y.

Inspecting our proof of Theorem 1.18 and noting the statement of Theorem 1.19, we can easily deduce that $Ax \leq c$ is TDI if each $b \in \mathbb{Z}^n$ with $\sup\{b^\top x : \ x \in P(A, c)\}$ being finite can be expressed as a non-negative integer combination of the columns of $A^\top$. (You are supposed to think over my assertion and write down in your notes the underlying argument.) Please compare this definition of TDI with Exercise 1.17, which is the essential part of our (geometric) proof of the LP Duality Theorem (Theorem 1.18). We mention that the latter definition of TDI can be formulated using the concept of `Hilbert basis`, of which we will not run into details.

The world is a dangerous place, not because of those who do evil, but because of those who look on and do nothing. – Albert Einstein (1879–1955)

Note that the definition of TDI does not pertain to polyhedra but to their linear inequality descriptions. This property is even sensitive to scaling as the next result reveals.

Theorem 4.1. *For any rational system $Ax \leq c$, where $A \in \mathbb{Q}^{m \times n}$ and $c \in \mathbb{Q}^m$, there exists a positive integer q such that $\frac{1}{q}Ax \leq \frac{c}{q}$ is TDI.*

Proof. For any $b \in \mathbb{Z}^n$, $P^{=}(A^\top, b)$, if nonempty, is a polyhedron having vertices. Since both A and b are rational, as a result of Exercise 1.31 and Cramer's rule, these vertices are rational as well. This says that there is a positive integer q such that the entries of all vertices are integer multiples of $\frac{1}{q}$. Therefore, all vertices of $P^{=}(\frac{1}{q}A^\top, b)$ are integral. Correspondingly, the dual system $\frac{1}{q}Ax \leq \frac{c}{q}$, which defines the same polyhedron as $Ax \leq c$, is TDI. $\square$

After getting Theorem 4.1, you might be wondering what is the point of the concept of TDI. Let us now introduce the result mentioned in the title of this chapter from which you can have some feeling that TDI is a property deserving some attention.

Theorem 4.2. *(Edmonds-Giles Theorem) If $Ax \leq c$ is a TDI system and c is integral, the polyhedron $P(A, c)$ is integral.*

Theorem 4.2 first appeared in [9]. It is because of this theorem Edmonds and Giles were led to formulate the concept of TDI system in the same paper.

Note that according to a book review of [28], Hoffman asserts that the concept of TDI system appeared, several years earlier before the concept was given a name and developed into a theory by Edmonds and Giles, in his paper

A.J. Hoffman, A generalization of max flow–min cut, Math. Prog., 6 (1974) 352–359,

(I have not been able to have access to either the 2003 book [28] or the 1974 paper of Hoffman.)

This idea of **using dual integrality to establish primal integrality** was also used by Fulkerson (1971) and Lehman (1979) independently.

Actually, we will prove a result stronger than Theorem 4.2. To get it, we have to make some preparations.

A subgroup of $\mathbb{R}^m$ is a `lattice` if it can be generated by linearly independent vectors. (Each abelian group is a $\mathbb{Z}$-module and so here we are talking about generated over $\mathbb{Z}$ though we are referring to the linear independence over $\mathbb{R}$. Do not confuse it.)

Exercise 4.3. For any positive integer i, let $x_i = (i, \lfloor \sqrt{2}i \rfloor) \in \mathbb{Z}^2$. Then the additive group generated by $\{x_i : \ i = 1, 2, \cdots\}$ is not a lattice.

Exercise 4.4. If $a_1, \cdots, a_n \in \mathbb{Q}^m$, then the additive group generated by them forms a lattice. (hint: use Hermite normal form)

The next theorem gives necessary and sufficient conditions for being an element of the lattice generated by the columns of a rational matrix A. It can also be viewed as an integer analogue of Theorems 1.1 and 1.8 (Farkas' Lemma).

Theorem 4.5. *Let $A \in \mathbb{Q}^{m\times n}$ and $b \in \mathbb{R}^m$. Then there is $x \in \mathbb{Z}^n$ such that $Ax = b$ if and only if for each $y \in \mathbb{Q}^m$ which satisfies $A^\top y \in \mathbb{Z}^n$ it holds $b^\top y \in \mathbb{Z}$.*

What I am going to tell you about is what we teach our physics students in the third or fourth year of graduate school... It is my task to convince you not to turn away because you don't understand it. You see my physics students don't understand it... That is because I don't understand it. Nobody does. – Richard Phillips Feynman (1918–1988)

Proof. (of Theorem 4.5) The forward implication follows at once and so our task below is to prove the converse.

First note that there exists $\widetilde{x} \in \mathbb{R}^n$ such that

$$A\widetilde{x} = b. \tag{4.1}$$

Otherwise, by Exercise 1.2 there would be a $\widetilde{y} \in \mathbb{Q}^m$ such that $\widetilde{y}^\top A = 0 \in \mathbb{Z}^n$ and $\widetilde{y}^\top b \neq 0$. By scaling if necessary, we can further require $\widetilde{y}^\top b \notin \mathbb{Z}$. This is impossible in view of our assumption.

Without loss of generality, assume that $A = \begin{pmatrix} A_1 \\ A_2 \end{pmatrix}$, $b = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$, where A_1 and b_1 have $r = \text{rank}(A) = \text{rank}(A_1)$ rows. This means that there is $P \in \mathbb{R}^{(m-r)\times r}$ such that $PA_1 = A_2$. Now Eq. (4.1) implies $b_2 = A_2\widetilde{x} = PA_1\widetilde{x} = Pb_1$. Thus, if we can prove the existence of an $x \in \mathbb{Z}^n$ such that $A_1 x = b_1$, then we have $A_2 x = PA_1 x = Pb_1 = b_2$ and hence $Ax = b$ follows.

To finish the proof, it remains to show the existence of $x \in \mathbb{Z}^n$ such that $A_1 x = b_1$. Note that both sides of the equivalence to be proved are invariant under elementary column operations and A_1 is of full row rank. So, by multiplying on the right suitable elementary matrices, we may assume that $A_1 = \begin{pmatrix} B & 0 \end{pmatrix}$ where $B \in \mathbb{Q}^{r\times r}$ is nonsingular (Recall the Hermite normal form or Smith normal form you met in your linear algebra course.). Clearly, for $x = \begin{pmatrix} B^{-1}b_1 \\ 0 \end{pmatrix}$, we have $A_1 x = b_1$. Is x a member of $\mathbb{Z}^n$? Yes, this is guaranteed by our assumption as each row of $\begin{pmatrix} B^{-1} & 0 \end{pmatrix}$ is rational and we have the following two equalities:

$$B^{-1}b_1 = \begin{pmatrix} B^{-1} & 0 \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} B^{-1} & 0 \end{pmatrix} b,$$

$$\begin{pmatrix} B^{-1} & 0 \end{pmatrix} A = \begin{pmatrix} B^{-1} & 0 \end{pmatrix} \begin{pmatrix} B & 0 \\ * & * \end{pmatrix} = \begin{pmatrix} I & 0 \end{pmatrix}.$$

□

The following easy fact in elementary number theory is a special case of Theorem 4.5 (why?).

Corollary 4.6. *Let $a_1, \cdots, a_n \in \mathbb{Z}$ and $g = \gcd(a_1, \cdots, a_n)$. Then x can be expressed as an integer linear combination of $a_1, \cdots, a_n$ if and only if $\frac{x}{g} \in \mathbb{Z}$.*

We also point out that Theorem 4.5 can be deduced from the so-called Kronecker's Approximation Theorem (why?):

Theorem 4.7. *Let $A \in \mathbb{R}^{m \times n}$ and let $b \in \mathbb{R}^n$. Then the following are equivalent:*

1. *For each $\epsilon > 0$, there is an $x \in \mathbb{Z}^n$ with $\begin{pmatrix} -\epsilon \\ \vdots \\ -\epsilon \end{pmatrix} < Ax - b < \begin{pmatrix} \epsilon \\ \vdots \\ \epsilon \end{pmatrix}$;*
2. *For each $y \in \mathbb{R}^m$, if $y^\top A \in \mathbb{Z}^n$ then $y^\top b \in \mathbb{Z}$.*

God made the integers; all else is the work of Man. – Leopold Kronecker (1823 – 1891)

Exercise 4.8. Let $A \in \mathbb{Q}^{m\times n}$ and $x \in \mathbb{R}^n$. If $Ax \in \mathbb{Q}^m$, then there is $y \in \mathbb{Q}^n$ such that $Ax = Ay$. If we further know that $x \in \mathbb{R}^n_+$, then the required y can be taken from $\mathbb{Q}^n_+$ accordingly. (hint: For the first assertion, mimic the corresponding argument for Theorem 4.5; for the second, apply the first assertion and note that $Ker(A)$ can be spanned over $\mathbb{R}$ by some rational vectors and that $\mathbb{Q}$ is dense in $\mathbb{R}$.)

Exercise 4.9. Is Theorem 4.5 still true when we relax the requirement $A \in \mathbb{Q}^{m\times n}$ to be $A \in \mathbb{R}^{m\times n}$?

Exercise 4.10. Let $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$. Then there is $x \in \mathbb{Z}^n$ such that $Ax = b$ if and only if for each $y \in \mathbb{R}^m$ which satisfies $A^\top y \in \mathbb{Z}^n$ it holds $b^\top y \in \mathbb{Z}$.

Both of the next two exercises are generalizations of Corollary 4.6.

Exercise 4.11. Let $A \in \mathbb{Z}^{m\times n}$ and $b \in \mathbb{Z}^m$. There is $x \in \mathbb{Z}^n$ such that $Ax = b$ if and only if for each natural number M the congruence $Ax \equiv b \pmod{M}$ has an integral solution x.

Recall that the kth `determinant factor` of an integer matrix A, recorded as $\mathcal{D}_k(A)$, is the greatest common divisor of all $k \times k$ minors of it.

Exercise 4.12. Let $A \in \mathbb{Z}^{m\times n}$, $b \in \mathbb{Z}^m$ and $B = \begin{pmatrix} A & b \end{pmatrix}$. If $\text{rank}(A) = m$, then there is $x \in \mathbb{Z}^n$ such that $Ax = b$ if and only if $\mathcal{D}_m(A)$ divides $\mathcal{D}_m(B)$.

To those who do not know mathematics it is difficult to get across a real feeling as to the beauty, the deepest beauty of nature. If you want to learn about nature, to appreciate nature, it is necessary to understand the language that she speaks in. – Richard Phillips Feynman (1918-1988)

A polyhedron is said to be `rational` if it can be defined by a rational system.

Theorem 4.13. *A rational polyhedron P is integral if and only if each rational supporting hyperplane of it contains integral vectors.*

Proof. The necessity is obvious as the intersection of P with a supporting hyperplane of it is an nonempty face of P.

To prove the sufficiency, suppose $P = P(A, c)$ for $A \in \mathbb{Q}^{m \times n}$ and $c \in \mathbb{Q}^m$, and there are integral points in each rational supporting hyperplane of P. We need to show that each minimal nonempty face F of P contains integral points. By Exercise 1.31, we assume that $F = \{x : \ A'x = c'\}$ where $A'x \le c'$ is a subsystem of $Ax \le c$. Thus, the problem turns out to be showing that there exists integral solution x to the equation $A'x = c'$.

But how to achieve it?

Yes, we can have a try with our Theorem 4.5. You should open your brains and be ready to pick up some work actively. Do not just be used to giving ear to my voice. It should be an easy job to finish the remaining proof.

My brain is open. – Paul Erdös (Mar. 26, 1913 – Sep. 20, 1996)

One more point: the author makes abundant and unblushing use of the words "clearly", "obviously", "evidently" etc. They are not used to blur the picture. On the contrary, they test the reader's understanding, for if he does not agree that the omitted reasoning is clear, obvious, and evident, he had better turn back a few pages and make a fresh start. – Lars V. Ahlfors (1907 – 1996), the preface to the 1st edition of his classic undergraduate textbook "Complex Analysis".

(continuation of the proof of Theorem 4.13)

Based on the assumption that there is no integral solution x to $A'x = c'$, we now try to reach a contradiction by finding a rational supporting hyperplane H of P which includes no integral vectors. Assume that $A' \in \mathbb{Q}^{\ell \times k}$. By means of Theorem 4.5, there is a $y \in \mathbb{Q}^{\ell}$ such that $A'^{\top} y = a \in \mathbb{Z}^k$ and $c'^{\top} y = \delta \notin \mathbb{Z}$. Since A' and c' are rational, we can assume additionally that $y \in \mathbb{Q}^{\ell}_{+}$ (why?). We then let $H = H_{a,\delta}$ and proceed to check that it is what we want. This breaks into three parts.

First, because a and δ are rational, H is rational by definition. Second, for each $x \in P$, we have

$$\begin{aligned} a^{\top} x &= y^{\top} A' x \\ &\leq y^{\top} c' && \text{by } y \geq 0 \text{ and } A'x \leq c' \\ &= \delta. \end{aligned}$$

This means that P lies in one side of H and $P \cap H \supseteq F$ is not empty. Henceforth, H is a supporting hyperplane of P. Third, since a is integral whereas δ is not, H contains no integral points. □

Exercise 4.14. Let $A \in \mathbb{Q}^{m \times n}$ and $c \in \mathbb{R}^m$. Then $P(A, c)$ is integral if and only if for each $b \in \mathbb{Q}^n$ with $\max\{b^{\top} x : \ x \in P(A, c)\} = \delta < \infty$ it holds $H_{b,\delta} \cap \mathbb{Z}^n \neq \emptyset$. (hint: follow the proof of Theorem 4.13.)

Exercise 4.15. Theorem 4.5 is the key to get Theorem 4.13 and Exercise 4.13. Can you develop any criterion about the integrality of a polyhedron based on Exercise 4.10, a variant of Theorem 4.5?

I cannot teach you how to produce original ideas. I guess, genius is not a product of any education system (though good environment does help a genius become a real master.). I only try to teach you how to imitate and try to share with you some understanding of some easy mathematics together with some common experience of living with mathematics.

Although Shimura had a whimsical streak – even today he retains host fondness for Zen jokes – he was far more conservative and conventional than his intellectual partner. Shimura would rise at dawn and immediately get down to work, whereas his colleague would often still be awake at this time, having worked through the night. Visitors to his apartment would often find Taniyama asleep in the middle of the afternoon. ... While Shimura was fastidious, Taniyama was sloppy to the point of laziness. Surprisingly this was a trait that Shimura admired: "He was gifted with the special capability of making many mistakes, mostly in the right direction. I envied him for this and tried in vain to imitate him, but found it quite difficult to make good mistakes. – Simon Singh, Fermat's Last Theorem.

The Japanese genius Yutaka Taniyama (1927–1958) killed himself in despair, while the German industrialist Paul Wolfskehl claimed Fermat had saved him from suicide. For more stories, read the book "Fermat's Last Theorem".

His approach to mathematics is as unique as his life. He has invented a new kind of art: the art of raising problems. Paul Erdös says that mathematics is eternal because it has an infinity of problems; and in his view, the more elementary a problem is, the better. He also invented his system of offering prizes for problems. His lectures with the title "My favorite problems in combinatorics" (or equivalent) always draw large audiences. Many try to imitate him in problem raising, but few can master this art: his problems may appear ad hoc or random at the first sight, especially for those not closely acquainted with the field, but after a few months or years of tirelessly pursuing one problem after the other, they suddenly connect up and form whole new theories - as if Erdös had those theories and metatheorems in his mind right away, and gave us only their corollaries. – Laszlo Lovász, Microsoft Research, a winner of John von Neumann Medal, Gödel Prize, Wolf Prize, Knuth Prize, Fulkerson Prize, among many other honors.

Theorem 4.16. *Let $A \in \mathbb{Q}^{m\times n}$ and $c \in \mathbb{R}^m$. Then $P(A, c)$ is integral if and only if for all $b \in \mathbb{Z}^n$ we have $\max\{b^\top x : x \in P(A, c)\}$ is an integer as long as it is finite.*

Proof. The forward direction is trivial. Conversely, in virtue of Exercise 4.13, it suffices to prove that the hyperplane $H_{b,\delta}$ contains integral points for any $b \in \mathbb{Q}^n$ such that $\max_{x\in P(A,c)} b^\top x = \delta < \infty$. By scaling, we can only consider the case that $b \in \mathbb{Z}^n$ and $\mathcal{D}_1(b) = 1$ ($b = 0$ is trivial). Now our assumption gives $\delta \in \mathbb{Z}$. At this stage, we can appeal to Corollary 4.6 and then finish the proof. □

Proof. (of Theorem 4.2) As in the proof of Theorem 4.16, only the backward implication needs to be considered. Now take $b \in \mathbb{Z}^n$ such that $\max\{b^\top x : x \in P(A, c)\} = \delta < \infty$. We prove that $\delta \in \mathbb{Z}$ and hence Theorem 4.16 gives the result. Indeed, LP duality theorem yields $\delta = \min_{y\in P^=(A^\top,b)} c^\top y$. But $Ax \le c$ is a TDI system and thus there is $y \in \mathbb{Z}^n \cap P^=(A^\top, b)$ such that $\delta = c^\top y$. It then follows from the integrality of c that $\delta \in \mathbb{Z}$, as required. □

You enter the first room of the mansion and it's completely dark. You stumble around bumping into the furniture but gradually you learn where each piece of furniture is. Finally, after six months or so, you find the light switch, you turn it on, and suddenly it's all illuminated. You can see exactly where you were. Then you move into the next room and spend another six months in the dark. So each of these breakthroughs, while sometimes they're momentary, sometimes over a period of a day or two, they are the culmination of, and couldn't exist without, the many months of stumbling around in the dark that precede them. – Andrew Wiles

Theorem 4.2 helps to motivate our interest in TDI. We will see that a converse to Theorem 4.2 is also true. This is reflected in the next theorem, which describes a further connection between TDI representations and integral polyhedra and hence strengthens our interest in TDI. It is interesting that our proof for this encouraging result also applies to prove the disappointing fact of Theorem 4.1.

Theorem 4.17. [13] *[16] A rational polyhedron is integral if and only if it is the set of feasible solutions to some TDI system $Ax \leq c$ where c is integral.*

Proof. The "if" part is just Theorem 4.2. So our goal here is the "only if" part.

Consider a rational polyhedron P. We can assume that $P = \{x \in \mathbb{R}^n : Mx \leq b\}$ where $M \in \mathbb{Z}^{m \times n}$. Let $\mathcal{L} = \{\ell \in \mathbb{Z}^n : \ell^\top = y^\top M, y \in \mathbb{R}^m, 0 \leq y_i \leq 1, \forall i\}$; in words, $\mathcal{L}$ is the set of all integral vectors whose transpose can be written as a nonnegative combination of the rows of M where no multiplier is greater than 1. As $\mathcal{L}$ is the set of integral points in a bounded region, $|\mathcal{L}|$ is a positive integer, say k. For each $\ell \in \mathcal{L}$, we put $t(\ell) = \max\{\ell^\top x : x \in P\}$.

Take the system $Ax \leq c$ consisting of the k inequalities $\ell^\top x \leq t(\ell)$ for all $\ell \in \mathcal{L}$. Since the transpose of each row of M appears in $\mathcal{L}$, the system $Ax \leq c$ defines the same polyhedron P as $Mx \leq b$. We assert that $Ax \leq c$ is TDI (and hence confirms Theorem 4.1 once more). Taking into account that P is integral and thus for each $\ell \in \mathcal{L}$, $t(\ell)$ must be an integer, we know that c is integral and so our theorem is a consequence of this assertion.

[13]Please refer to III.1, Proposition 1.7 of the book by Nemhauser & Wolsey, Integer and Combinatorial Optimization, Wiley, 1999, for a shorter proof.

The remainder of the proof is to verify our claim above. Let w be an integral vector such that

$$\max\{w^\top x : \ x \in P\} \tag{4.2}$$

has a finite optimal objective value, say $\lambda < \infty$. We have to construct an integral optimal solution to

$$\min\{z^\top c : \ z \in P^=(A, w) = P\}. \tag{4.3}$$

To this end, we consider the following problem:

$$\min\{y^\top b : y \in P^=(M, w)\}. \tag{4.4}$$

As $P = P(A, c) = P(M, b)$, we know that both (4.3) and (4.4) are dual problems to (4.2) and so, by the LP duality theorem, they share the same optimal objective value λ. Assume that $y^* \in \mathbb{R}^m$ achieves the minimum in (4.4)[14]. Basing on y^*, we try to generate a $z^* \in \mathbb{Z}^k$ which attains the minimum in (4.4).

[14] It seems that we need not worry about the case that y^* is integral; if so, the system $Mx \leq b$ itself is TDI. This argument is TOTALLY wrong. We are given a w and it may happen that we have good luck to meet with an integral y^*. But to prove $Mx \leq b$ is TDI, what we have to do is to get an integral solution to the dual problem whether or not we have the good luck. Our claim is that for another system describing the same polyhedron, namely $Ax \leq c$, we can prove the existence of integral solution to its dual problem whether or not both the integral testing vector w and our luck are good enough that y^* comes out integral. In one word, we are checking the TDI property of one system and you have no freedom to choose the system according to the inputs.

Observe that A is built from M by adding some more rows. Without loss of generality, let $A = \begin{pmatrix} M \\ N \end{pmatrix}$. Correspondingly, let $c = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$ and write each feasible solutions to (4.3) as $z = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}$ where the dual variables of c_1 and z_1 correspond to rows of M and c_2 and z_2 to N. It is worth noting that

$$c_1 \leq b. \tag{4.5}$$

Use $\lfloor y^* \rfloor$ to denote the integral vector obtained from y^* by rounding down each component of y^* to the nearest integer below it. To get the required integral vector z^* from y^*, it is very natural to first try to setting

$$z_1^* = \lfloor y^* \rfloor. \tag{4.6}$$

To finish the construction of z^*, namely to determine z_2^*, we have to take care of two things:

(i) z^* is integral and is feasible to (4.3);

(ii) $z^{*\top} c$ attains the optimal objective value of (4.3), namely λ.

The fact that z^* is feasible to (4.3) amounts to saying $z_2^* \geq 0$ and

$$z^{*\top} A - z_1^{*\top} M = w^\top - z_1^{*\top} M. \tag{4.7}$$

The LHS of (4.7) is nothing but $z_2^{*\top} N$; while we have

$$\begin{aligned} \text{RHS of (4.7)} = & \quad w^\top - \lfloor y^* \rfloor M \quad \text{(by (4.6))} \\ = & \quad y^{*\top} M - \lfloor y^* \rfloor M \quad \text{(by } y^* \text{ is feasible to (4.4)).} \end{aligned}$$

Therefore, to get (i), it suffices to require

$$z_2^* \geq 0, z_2^* \in \mathbb{Z}^{k-m}, \text{and } z_2^{*\top} N = y^{*\top} M - \lfloor y^* \rfloor M. \tag{4.8}$$

Requirement (ii) is that $z^{*\top} c = \lambda$, which is equivalent to

$$\begin{aligned} z_2^{*\top} c_2 = & \quad \lambda - z_1^{*\top} c_1 \\ = & \quad \lambda - \lfloor y^* \rfloor^\top c_1 \quad \text{(by (4.6)).} \end{aligned} \tag{4.9}$$

We mention that for those z^* already satisfying (4.8), (4.9) is equivalent to

$$z_2^{*\top} c_2 \leq \lambda - \lfloor y^* \rfloor^\top c_1, \tag{4.10}$$

since λ is the minimum objective value for (4.3).

Note that, at this point, we have reduced our task to finding a z_2^* for which both (4.8) and (4.9) ((4.10)) hold.

First consider the case that $y^* = \lfloor y^* \rfloor$. It is obvious that $z_2^* = 0$ fulfils (4.8) and so the corresponding z^* is feasible to (4.3). Since λ is the optimal objective value to (4.3), we get

$$0 = z_2^{*\top} c_2 \geq \lambda - z_1^{*\top} c_1 = \lambda - \lfloor y^* \rfloor^\top c_1 \tag{4.11}$$

Moreover, as λ is also the optimal objective value to (4.4), we have

$$0 = \lambda - y^{*\top} b. \tag{4.12}$$

But it follows from (4.5) and $y^* \geq 0$ that

$$\lambda - y^{*\top} c_1 \geq \lambda - y^{*\top} b. \tag{4.13}$$

Thus, combining (4.11), (4.12) and (4.13), we arrive at

$$0 = z_2^{*\top} c_2 \geq \lambda - \lfloor y^* \rfloor^\top c_1 = \lambda - y^{*\top} c_1 \geq 0,$$

and so the equalities have to hold throughout. This implies the validity of (4.9), as desired.

It remains to consider the case that $y^* - \lfloor y^* \rfloor \neq 0$.

We will show that in $Nx \leq c_2$ there is an inequality

$$\overline{w}^\top x \leq t(\overline{w}), \tag{4.14}$$

where

$$\overline{w}^\top = y^{*\top} M - \lfloor y^* \rfloor M = w^\top - \lfloor y^* \rfloor M, \tag{4.15}$$

and

$$t(\overline{w}) \leq \lambda - \lfloor y^* \rfloor^\top c_1. \tag{4.16}$$

If so, setting the variable of z_2^* corresponding to the inequality (4.14) to be one and all other variables zeros, we get a solution to both (4.8) and (4.10).

We do the verification in two parts.

On the one hand, each component of $y^* - \lfloor y^* \rfloor$ is a nonnegative number less than 1 (This guarantees that (4.14) does not appear in $Mx \leq c_1$.). On the other hand, $M^\top y^* = w$, $\lfloor y^* \rfloor$ and M are all integral and hence $M^\top(y^* - \lfloor y^* \rfloor)$ is integral. Consequently, $\overline{w} = M^\top(y^* - \lfloor y^* \rfloor) \in \mathcal{L}$. This says that (4.14) does appear in $Nx \leq c_2$.

For any $y \in P^=(M, \overline{w})$, we have $y + \lfloor y^* \rfloor \in P^=(M, w)$ and thus, as y^* is an optimal solution to (4.4), $y^\top b = (y + \lfloor y^* \rfloor)^\top b - \lfloor y^* \rfloor^\top b \geq y^{*\top} b - \lfloor y^* \rfloor^\top b = \lambda - \lfloor y^* \rfloor^\top b$, with equality holding when $y = y^* - \lfloor y^* \rfloor$. But $y^* - \lfloor y^* \rfloor \in P^=(M, \overline{w}^\top)$. So we find that

$$t(\overline{w}) = \lambda - \lfloor y^* \rfloor^\top b. \tag{4.17}$$

Just as (4.13), in virtue of (4.5) and $\lfloor y^* \rfloor \geq 0$, (4.16) is immediate from (4.17)[15]. □

Mathematics is not a deductive science – that's a cliche. When you try to prove a theorem, you don't just list the hypotheses, and then start to reason. What you do is trial and error, experimentation, guesswork. – Paul R. Halmos, in: I Want to be a Mathematician, Washington: MAA Spectrum, 1985.

The most wonderful gift I have received this year is a copy of the Chinese translation of "I Want to be a Mathematician". I do like it very much.

[15]Because of the equivalence between (4.9) and (4.10), we know that indeed it holds equality in (4.16).

The importance of Theorem 4.17 lies in the fact that it guarantees that we can ALWAYS try the following procedure to deduce the integrality of a given rational polyhedron provided it is integral:

- Find an approximate defining system $Ax \leq b$ with A and b integral;
- Prove $Ax \leq b$ is TDI;
- Using Theorem 4.17, conclude that $P(A, b)$ is integral.

We close this chapter with a result on the connection between TUM and TDI.

Theorem 4.18. *$A \in \mathbb{Z}^{m \times n}$ is totally unimodular if and only if the system $Ax \le c$, $x \ge 0$, is TDI for each $c \in \mathbb{Z}^m$.*

Proof. If A is TUM, then so is $A^\top$. Consequently, for each $b \in \mathbb{Z}^n$, we deduce from Theorem 2.6 that $\{y \in \mathbb{R}^m : \ A^\top y \ge b, y \ge 0\}$ is an integral polyhedron. But for any $c \in \mathbb{R}^m$, $\min\{y^\top c : \ A^\top y \ge b, y \ge 0\}$ is just the dual program to $\max\{b^\top x : \ Ax \le c, x \ge 0\}$ and so we get that when $c \in \mathbb{Z}^m$ the rational system $Ax \le c$, $x \ge 0$ is TDI.

For the reverse direction, suppose $Ax \le c$, $x \ge 0$, is TDI for each $c \in \mathbb{Z}^m$. By Theorem 4.2, this shows that $P\left(\begin{pmatrix} A \\ -I_n \end{pmatrix}, c\right) = P(A, c) \cap \mathbb{R}^n_+$ is integral for each $c \in \mathbb{Z}^m$. An application of Theorem 2.9 then completes the proof. $\square$

Compare with Exercises 2.11.

5 Matroid Intersection Theorem

...some general ideas which tend to play a role of systematizing and unifying concepts. The first of these is the notion of a good characterization, formulated by J. Edmonds. ... Good characterization reflects a deep underlying logical duality of the property and, as the reader may convince himself by comparing good and "non-good" characterizations occuring throughout this book, often amounts to "the" solution of the problem.... Another idea which has proved very fruitful is that combinatorial optimization problems can generally be formulated as linear programming problems with integrality constraints....Last but not least we mention the use of linear algebra. This ranges from application of matrix calculus to the introduction of homology and cohomology groups. A common background of many applications of linear algebra is Matroid Theory, which is now a flourishing branch of combinatorics itself. – L. Lovász, Preface to *Combinatorial Problems and Exercises*, North-Holland, 1979.

Matroid theory took a great step further with the Matroid Intersection and Matroid Union Theorems. They are closely related and have been fruitful in affording a unified context for many min-max relations in graph theory.

In 1965, Jack Edmonds and Delbert Fulkerson [6] proved that the union of two matroids is again a matroid and, more importantly, derived the expression for its rank function, which can be regarded as a max-min relation as rank function is the objection value function for an max problem.

Theorem 5.1. *(Matroid Union Theorem) If* $M_i = (X, \mathcal{I}_i)$ *are matroids with rank functions* r_i *for* $i = 1, \cdots, k$*, then* $(X, \cup_{i=1}^k \mathcal{I}_i)$ *is a matroid with rank function* $r(X) = \min_{U \subseteq X}(\sum_{i=1}^k r_i(U) + |X \setminus U|)$*.*

Unlike union, for two matroids $M_1 = (X, \mathcal{I}_1)$ and $M_2 = (X, \mathcal{I}_2)$, $\mathcal{I}_1 \cap \mathcal{I}_2$ is usually only a family of downwards-monotone subsets (called an abstract simplicial complex) and it may hold that $(X, \mathcal{I}_1 \cap \mathcal{I}_2)$ is not a matroid (In matroid theory, we define a matroid to be the intersection of two matroids in some other ways). However, there is an elegant and very useful max-min relation for the size of a common independent set in two matroids on the same ground set, as was shown again by Jack Edmonds [8] in 1970.

Theorem 5.2. *(Matroid Intersection Theorem) Let* $M_1 = (X, \mathcal{I}_1)$ *and* $M_2 = (X, \mathcal{I}_2)$ *be two matroids with rank functions* r_1 *and* r_2*, respectively. Then* $\max_{I \in \mathcal{I}_1 \cap \mathcal{I}_2} |I| = \min_{U \subseteq X}(r_1(U) + r_2(X \setminus U))$*.*

We will first prove Theorem 5.2 and then deduce Theorem 5.1 as a corollary.

For both Theorem 5.1 and Theorem 5.2, the direction that max $\leq$ min is obvious and so every layman will find it satisfying to be able to recognize that weak connection. (Here is the straightforward proof for max $\leq$ min in Theorem 5.2: For any $I \in \mathcal{I}_1 \cap \mathcal{I}_2$ and $U \subseteq X$, we have $|I| = |I \cap U| + |I \cap (X \setminus U)| \leq r_1(U) + r_2(X \setminus U)$, since $I \cap U \in \mathcal{I}_1$ and $I \cap (X \setminus U) \in \mathcal{I}_2$.) But the other direction in both theorems is really nontrivial and surprising. Similarly, most people will be able to prove the weak LP duality theorem and thus they have the ability to be excited by the LP duality theorem even though they cannot work out its proof themselves. As with many other results in mathematics, this is why these results are especially attractive.

We have learned the Edmonds-Giles Theorem (Theorem 4.2) in Chapter 4; Giles is a Ph.D. student of Edmonds in University of Waterloo. Even earlier, we proved the Edmonds' Matroid Polytope Theorem (Theorem 3.5) in Chapter 3. The bulk of this chapter is to establish the results of Edmonds on matroid union and matroid intersection with the help of his Matroid Polytope Theorem. Later, we will shift to the topic of matching theory and the main theorem there will be Edmonds' Matching Polytope Theorem. Recall the comment of Cunningham, another student of Edmonds, which we cite on page 90. Many important results in combinatorial optimization come from Edmonds and his outstanding students.

This is the house that Jack[16] *built. Ralph*[17] *prepared a lot. There were many independent contractors who did beautiful work; some putting on splendid additions. Martin*[18]*, Laci*[19]*, and Lex*[20] *rewired the place. The work continues. But this is the house that Jack built.* – Jon Lee, Preface to *A First Course in Combinatorial Optimization*, Cambridge Texts in Applied Mathematics, Cambridge University Press, 2004.

[16] Jack Edmonds

[17] Ralph Edward Gomory

[18] Martin Grötschel

[19] László Lovász

[20] Alexander Schrijver

After getting his master degree with a thesis on embedding graphs into surfaces[21], Edmonds joined NBS in 1959. During his career in NBS, Edmonds discovered his famous matching algorithm. It is before the talk to announce his characterization of matchings and the matching algorithm in a summer workshop in 1963, Edmonds shouted "Eureka–You shrink!". To understand why this result put Edmonds in storm then, please go to his seminal paper in which this discovery is described: Paths, trees, and flowers, *Canad. J. Math.* 17 (1965), 449–467. Note that you can find the above paper of Edmonds in the book [15] which is available in our Minhang library.

[21]I got a copy of this book in this summer: C.P. Bonnington, C.H.C. Little, The Foundations of Topological Graph Theory. I find that there is a whole chapter in the book describing the work of Jack Edmonds presented in his master thesis.

I wish that you will also like to read this web material prepared by C. Witzgall: `http://nvl.nist.gov/pub/nistpubs/sp958-lide/140-144.pdf`

We cite here the comments of Witzgall on the paper "Paths, trees, and flowers": "Why was it a breakthrough? The answer is that all good graph-theoretical algorithms known at the time addressed "unimodular" problems such as the "Shortest Path" and "Network Flow" problems, the rigorous proof for the latter having been given by Edmonds with collaboration by Richard M. Karp. These are problems that could be formulated as integrality-preserving linear programs, which by themselves did not create good algorithms but indicated the potential for such. Edmonds' matching algorithm was the very first instance of a good algorithm for a problem outside that mode. "

Please compare with the comments of Pulleyblank, a student of Edmonds, on page 72.

I hit it lucky by putting a poetic title on a paper that was mathematically a hit. – Jack Edmonds (1935–)

How happy it is if you can produce some good math! For me, it is also enough comfort to be able to understand some good math.

Jack Edmonds was awarded the 1985 John von Neumann Theory Prize. Below is the citation of the award copied from `http://www.informs.org/Prizes/vonNeumannDetails.html`

Jack Edmonds has been one of the creators of the field of combinatorial optimization and polyhedral combinatorics. His 1965 paper "Paths, Trees and Flowers" was one of the first papers to suggest the possibility of establishing a mathematical theory of efficient combinatorial algorithms. In that paper and in the subsequent paper "Maximum Matching and a Polyhedron with 0-1 Vertices" Edmonds gave remarkable polynomial-time algorithms for the construction of maximum matchings. Even more importantly these papers showed how a good characterization of the polyhedron associated with a combinatorial optimization problem could lead via the duality theory of linear programming, to the construction of an efficient algorithm for the solution of that problem.

During the 1960's, while working at the National Bureau of Standards, Edmonds explored the relationship between matroids and optimization. His beautiful theory of matroid partition and intersection remains one of the deepest results in the area. This work further illustrates the deep interconnections between combinatorial minmax theorems, polyhedral structure, duality theory and efficient algorithms. During this period he also made notable advances in the theory of network flow algorithms.

In 1969 Edmonds moved to the University of Waterloo, where in collaboration with his many outstanding students, he has continued to explore combinatorial optimization problems and the associated polyhedra. His work during the past fifteen years has revolved around the theories of submodular functions, total dual integrality and oriented matroids.

Throughout his career Edmonds has expended immeasurable time and effort assisting young researchers. Through his influence many outstanding young mathematicians have been drawn to the field of theoretical operations research.

To prove Theorem 5.2, we linearize the problem by considering the relationship between the polytope spanned by the incidence vectors of those elements of $\mathcal{I}_1 \cap \mathcal{I}_2$, denoted $P(\mathcal{I}_1 \cap \mathcal{I}_2)$ hereafter, and the two matroid polytopes $P(M_1)$ and $P(M_2)$. (There are various clever "combinatorial" proofs for Theorem 5.2; but, rather than presenting polished proofs, we want to emphasize the basic idea of polyhedral combinatorics.)

Theorem 3.5 says that $P(M_1)$, the polytope spanned by the incidence vectors of $\mathcal{I}_1$, is defined by the system

$$\begin{cases} z(x) \geq 0, & \text{for } x \in X; \\ \sum_{x \in U} z(x) \leq r_1(U), & \text{for } U \subseteq X, \end{cases}$$

and $P(M_2)$, the polytope spanned by the incidence vectors of $\mathcal{I}_2$, is defined by the system

$$\begin{cases} z(x) \geq 0, & \text{for } x \in X; \\ \sum_{x \in U} z(x) \leq r_2(U), & \text{for } U \subseteq X, \end{cases}$$

respectively. Then we know that $P(\mathcal{I}_1 \cap \mathcal{I}_2) \subseteq P(M_1) \cap P(M_2)$, the latter being a polytope defined by

$$\begin{cases} z(x) \geq 0, & \text{for } x \in X; \\ \sum_{x \in U} z(x) \leq r_1(U), & \text{for } U \subseteq X; \\ \sum_{x \in U} z(x) \leq r_2(U), & \text{for } U \subseteq X. \end{cases} \tag{5.1}$$

It turns out that the intersection of two matroid polytopes gives exactly the convex hull of the common independent sets. We will illustrate later that this polytope characterization is equivalent to Theorem 5.2, one formulated in geometric form and the other combinatorial form.

Theorem 5.3. *$P(M_1) \cap P(M_2) = P(\mathcal{I}_1 \cap \mathcal{I}_2)$.*

Proof. (of Theorem 5.3) It is helpful to review the proof of Theorem 3.5 in order to get a good understanding of the subsequent arguments. We will use several tricks already introduced there and so we will feel free to skip corresponding explanations.

As in proving Theorem 3.5, we only need to show that each vertex of $P(M_1) \cap P(M_2)$ is integral.

Since (5.1) is a rational system, we can check the integrality of $P(M_1) \cap P(M_2)$ using Theorem 4.16; that is, for any $b \in \mathbb{Z}^X$, we try to see if $\max\{b^\top z : \ z \text{ satisfies } (5.1)\}$ is always an integer. To do this, in view of the LP duality theorem, we turn to the dual problem and hope that it might be easier to deal with. The next step is then to write out the dual linear programming problem according to the rule given on page 29:

$$DP: \quad \min \sum_{U \subseteq X} (y_1(U) r_1(U) + y_2(U) r_2(U))$$

subject to

$$\begin{cases} \sum_{x \in U \subseteq X} (y_1(U) + y_2(U)) \geq b(x), & \text{for } x \in X; \\ y_1(U) \geq 0, & \text{for } U \subseteq X; \\ y_2(U) \geq 0, & \text{for } U \subseteq X. \end{cases} \tag{5.2}$$

Note that the columns of the constraint matrix of (5.2) are indexed by two disjoint copies of 2^X, which we refer to as $2^X \cup 2^X$, where the first 2^X corresponds to variables of y_1 and the second corresponds to variables of y_2. To prove that the above optimization problem has integer optimal objective value for each integral b, it is enough to show that the optimal value can be achieved at an integral vector $y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$ (this amounts to saying that the vertices of the polytope defined by the rational system (5.2) are all integral). Note that the foregoing analysis is just to remind you once again the idea hidden behind the concept of TDI; indeed, we are just to prove that (5.1) is TDI and we can omit the above illustrations by directly citing Theorem 4.2.

Schrijver [32, Vol. B, Theorem 41.12] presents a short proof for the fact that (5.1) is box-TDI, which is stronger than the assertion that (5.1) is TDI. We adopt a variant of his argument below in which you can find some familiar arguments used in proving Theorem 3.5.

By Theorem 2.6, our task will be completed if we can demonstrate that the constraint matrix of (5.2) is totally unimodular. Unfortunately, the truth is that the matrix can fail to be TU. Does it mean that Theorem 2.6 is useless for our purpose now? No, it is always too soon for despair. After all, our task is to find integral vectors from optimal solutions to (5.2) with integral b and so we may focus on part of those optimal solutions for which we can have some other algebraic descriptions such that Theorem 2.6 becomes usable (when the face of the polytope is minimal, namely a vertex, those special optimal solutions come out to be all optimal solutions as there is only one optimal solution).

For any $R_1 \cup R_2 \subseteq 2^X \cup 2^X$ and any $y \in \mathbb{R}^{2^X \cup 2^X}$, we can shorten y to be $\overline{y}$ by deleting all positions of y lying outside of $R_1 \cup R_2$. Note that under the assumption that there does exist optimal solutions to (5.2) which takes value 0 outside of $R_1 \cup R_2$, y^* is such an optimal solutions if and only if its shortened vector $\overline{y^*}$ is an optimal solutions to

$$\min \sum_{U \in R_1} y_1(U) r_1(U) + \sum_{U \in R_2} y_2(U) r_2(U)$$

subject to

$$\begin{cases} \sum_{x \in U \in R_1} y_1(U) + \sum_{x \in U \in R_2} y_2(U)) \geq b(x), & \text{for } x \in X; \\ y_1(U) \geq 0, & \text{for } U \in R_1; \\ y_2(U) \geq 0, & \text{for } U \in R_2. \end{cases} \tag{5.3}$$

Observe that the constraint matrix to (5.3) consists of only those columns of the constraint matrix of (5.2) which are indexed by $R_1 \cup R_2$. We can expect that such a new matrix may become a TUM and so we turn to see if we can say something about the support of an optimal solution to (5.2) (If the corresponding face is minimal, we are discussing what a vertex looks alike!).

We continue to analyze what special constraint we can put on some special optimal solutions to DP. That is to say, based on the existence of optimal solutions, we will establish the existence of some special optimal solutions.

Given any $b \in \mathcal{R}^X$, we suppose that the minimum of DP is attained at $y = \begin{pmatrix} y_1^* \\ y_2^* \end{pmatrix}$. Clearly, y_1^* can be replaced by any optimal solution to

$$\min \sum_{U \subseteq X} y_1(U) r_1(U)$$

subject to

$$\begin{cases} \sum_{x \in U \subseteq X} y_1(U) \geq b(x) - \sum_{x \in U \subseteq X} y_2^*(U), & \text{for } x \in X; \\ y_1(U) \geq 0, & \text{for } U \subseteq X. \end{cases} \tag{5.4}$$

This is the dual to the optimum independent set problem in the matroid M_1 with weight function $w(x) = b(x) - \sum_{x \in U \subseteq X} y_2^*(U)$– such type of a problem we have encountered on page 197. As with the proof of Theorem 3.5, we can solve DP1 by appealing to the greedy algorithm and hence we are allowed to modify our y_1^* (that is, choose the output of the greedy algorithm as a new y_1^*) so that there are $A^1 \subseteq A^2 \subseteq \cdots \subseteq A^k \subseteq X$ and $y_1^*(U) = 0$ whenever $U \notin \{A^1, \cdots, A^k\}$. Furthermore, we use greedy algorithm once more and thus assume that y_2^* only takes nonzero values on a chain $B^1 \subseteq B^2 \subseteq \cdots \subseteq B^\ell$. (This analysis also tells us that each vertex of the polytope defined by (5.2) has a support consisting of two chains in 2^X–we do not use this fact in the ensuing proof.)

Summing up, we have known that for any $b \in \mathcal{R}^X$, there are two chains $\mathcal{F}_1$ and $\mathcal{F}_2$ of subsets of X such that each optimum solution to

$$\min \sum_{U \subseteq X} (y_1(U) r_1(U) + y_2(U) r_2(U))$$

subject to

$$\begin{cases} \sum_{x \in U \subseteq X} (y_1(U) + y_2(U)) \geq b(x), & \text{for } x \in X; \\ y_1(U) = 0, & \text{for } U \in 2^X \setminus \mathcal{F}_1; \\ y_2(U) = 0, & \text{for } U \in 2^X \setminus \mathcal{F}_2; \\ y_1(U) \geq 0, & \text{for } U \in \mathcal{F}_1; \\ y_2(U) \geq 0, & \text{for } U \in \mathcal{F}_2. \end{cases} \tag{5.5}$$

is also an optimal solution to DP.

So, for $b \in \mathbb{Z}^X$, the existence of integral optimal solution to DP comes from the existence of integral optimal solution to (5.3) where $R_1 = \mathcal{F}_1$ and $R_1 = \mathcal{F}_2$ are two chains. Therefore, thanks to Theorem 2.6, we will be done if we can prove now that the constraint matrix of (5.3) is TU. Observe that this matrix in discussion can be written as $C = \mathcal{M}(X, \mathcal{F}_1, \mathcal{F}_2) = \begin{pmatrix} C_1 & C_2 \end{pmatrix}$, where for $i = 1, 2$, C_i is the point-set incidence matrix for X and $\mathcal{F}_i$, namely for $x \in X$ and $U \in \mathcal{F}_i$, $C_i(x, U) = \begin{cases} 1, & \text{if } x \in U; \\ 0, & \text{otherwise.} \end{cases}$

Taking any square submatrix D of C, we want to deduce that $\det D \in \{0, \pm 1\}$. We may assume that $D = \mathcal{M}(X', \mathcal{F}'_1|_{X'}, \mathcal{F}'_2|_{X'})$, where $X' \subseteq X, \mathcal{F}'_1 \subseteq \mathcal{F}_1$ and $\mathcal{F}'_2 \subseteq \mathcal{F}_2$. Suppose $\mathcal{F}'_1 = \{A^1 \subseteq A^2 \subseteq \cdots \subseteq A^k\}$ and $\mathcal{F}'_2 = \{B^1 \subseteq B^2 \subseteq \cdots \subseteq B^\ell\}$. For $i = 1, \cdots, k-1$, we subtract the column of C_1 corresponding to A^i from that corresponding to A^{i+1}; similarly, for $i = 1, \cdots, \ell - 1$, we subtract the column of C_2 corresponding to B^i from that corresponding to B^{i+1}. This results in a $(0, 1)$ matrix $D = \begin{pmatrix} D_1 & D_2 \end{pmatrix}$ where each row of D_i, $i = 1, 2$, has at most one 1. Theorem 2.14 asserts that $\begin{pmatrix} D_1 & -D_2 \end{pmatrix}^\top$ has determinant 0 or ± 1 and so $\det D \in \{0, \pm 1\}$ follows, as desired. $\square$

We have been discussing the pair of dual linear programming problems associated with two matroids $M_i = (X, \mathcal{I}_i)$, $i = 1, 2$:

$$PP: \quad \max \sum_{x \in X} b(x) z(x)$$

subject to

$$\begin{cases} z(x) \geq 0, & \text{for } x \in X; \\ \sum_{x \in U} z(x) \leq r_1(U), & \text{for } U \subseteq X; \\ \sum_{x \in U} z(x) \leq r_2(U), & \text{for } U \subseteq X. \end{cases} \tag{5.6}$$

$$DP: \quad \min \sum_{U \subseteq X} (y_1(U) r_1(U) + y_2(U) r_2(U))$$

subject to

$$\begin{cases} \sum_{x \in U \subseteq X} (y_1(U) + y_2(U)) \geq b(x), & \text{for } x \in X; \\ y_1(U) \geq 0, & \text{for } U \subseteq X; \\ y_2(U) \geq 0, & \text{for } U \subseteq X. \end{cases} \tag{5.7}$$

It is worth mentioning a side result of our arguments in proving Theorem 5.3.

Corollary 5.4. *If b is integral, then both* (5.6) *and* (5.7) *have integral optimal solutions.*

Proof. Our proof of Theorem 5.3 shows that both (5.6) and (5.7) are TDI and so the result follows.

Indeed, we even have obvious clue that the vertices of the two polytopes are integral. Theorem 5.3 says that (5.6) is a description of $P(\mathcal{I}_1 \cap \mathcal{I}_2)$, which is surely an integral polytope as it is the convex hull of some integral vectors by definition. The remark at the end of page 129 says that each vertex of the polytope corresponding to (5.7) has special support and then the arguments on page 130 apply to yield that each of them is integral. □

We are all in the gutter, but some of us are looking at the stars. – Oscar Wilde (Irish poet, novelist, dramatist and critic, 1854–1900)

Proof. (of Theorem 5.2) Set $b \in \mathbb{R}^X$ to be the all ones vector in (5.6) and its dual (5.7). Since a linear functional restricted on a polytope achieves its maximal value at a vertex, it follows from Theorem 5.3 that $\max_{I \in \mathcal{I}_1 \cap \mathcal{I}_2} |I|$ is just the optimal objective value for (5.6). To finish the proof, it suffices to apply LP duality theorem and prove that $\min_{U \subseteq X}(r_1(U) + r_2(X \setminus U))$ is the optimal objective value for (5.7). As with the proof of Theorem 5.3, our strategy is to start from any optimal solution $y^* = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$ to (5.7) and try to move to some special optimal solution of it – this time, we want to deduce the existence of an optimal solution y^* such that there is $U \subseteq X$ such that $y_1^*(V) = \begin{cases} 1, & \text{if } V = U; \\ 0, & \text{otherwise,} \end{cases}$ and $y_2^*(V) = \begin{cases} 1, & \text{if } V = X \setminus U; \\ 0, & \text{otherwise.} \end{cases}$ By Corollary 5.4, we can require that y^* is integral; further, by inspecting the form of (5.7), we assume that y^* is a $(0,1)$ vector.

For each $x \in X$, we have a constraint $\sum_{x \in U \subseteq X}(y_1(U) + y_2(U)) \geq b(x) = 1$. This means that a $(0,1)$ vector y is a feasible solution to (5.7) as long as $V_1 \cup V_2 = X$, where $V_i = \cup_{y_i(U)=1} U$, $i = 1, 2$. In other words, in the process of finding a suitable new optimal solution, we can adjust the distribution of 1s in the $(0,1)$ vector y^* arbitrarily just not to decrease the objective value and to keep $V_1 \cup V_2$ unchanged.

It is evident that the rank function r of a matroid satisfies the so-called `submodular inequality`, namely $r(A \cup B) + r(A \cap B) \leq r(A) + r(B)$. In particular, this implies that it holds for a rank function r that $r(\cup_i S_i) \leq \sum_i r(S_i)$. Therefore, for $i = 1, 2$, since $r_i(\emptyset) = 0$, by resetting $y_i^*(V_i) = 1$ and $y_i^*(U) = 0$ for all $U \neq V_i$ we will arrive at a new optimal solution, still denoted by y^*. Moreover, we can still choose a (possibly new) $(0,1)$ vector y^* as the optimal solution by changing the only nonzero entry of y_2^* from V_2 to $V_2 \setminus V_1$, considering that $r_2(V_2 \setminus V_1) \leq r_2(V_2)$. This optimal solution is in the form of what we require, thus ending the proof. □

A `partition matroid` is a matroid induced by a partition of the ground set in which a set is independent if and only if it has at most one element from each block of the partition. It is an easy exercise to check that partition matroid thus defined is really a matroid. Partition matroids forms a subclass of `transversal matroids` [22].

Consider a graph $G = (V, E)$. For any $F \subseteq E$, we write $\partial(F) = \cup_{e \in F} \partial(e)$. For any $U \subseteq V$, we let $\delta(U) = \{e \in E : \partial_G(e) \cap U \neq \emptyset, \partial_G(e) \cap V \setminus U \neq \emptyset\}$, $\gamma(U) = \{e \in E : \partial_G(e) \subseteq U\}$. We also define the `open neighbor set` of U to be $N(U) = \partial(\delta(U)) \setminus U$. When $U = \{v\}$, we write $\delta(v)$ for $\delta(\{v\})$ and $\gamma(v)$ for $\gamma(\{v\})$, and so forth. δ is usually called the `coboundary operator` and ∂ the `boundary operator`. A `matching` of G is a subset $M \subseteq E$ such that $|M \cap \delta(v)| \leq 1$ and $|M \cap \gamma(v)| = 0$ for each $v \in V$. Each element from $\gamma(v)$ is called a *loop* at $v \in V$. G is `loopless` if there is no loop in G. An `independent set` of G is a subset U of V satisfying $\gamma(U) = \emptyset$. A `vertex cover` is a subset $U \subseteq V$ such that $\partial(e) \cap U \neq \emptyset$ for each $e \in E$. An `edge cover` of G is a subset F of E with $\partial(F) = V$.

---

[22]Both transversal matroids and partition matroids should have been introduced in §3.2. See §10.3 of [31].

Here is our first attempt to specialize Theorem 5.2 to find interesting facts. Recall our quotation from Spencer on page 53.

Example 5.5. Let $G = (V_1, V_2, E)$ be a bipartite graph with color classes V_1 and V_2. Consider $M_i = (E, \mathcal{I}_i)$, where $\mathcal{I}_i = \{F \subseteq E : \deg_F(v) \leq 1, \forall v \in V_i\}$, $i = 1, 2$. Note that both M_1 and M_2 are partition matroids, where M_i corresponds to the partition $E = \cup_{v \in V_i} \delta(v)$ for $i = 1, 2$. An application of the Matroid Intersection Theorem to this two matroids gives the König's matching theorem [23][23]: The maximum cardinality of a matching in a bipartite graph is equal to the minimum cardinality of a vertex cover. For two very short direct proofs of König's matching theorem, we refer to that due to De Caen included in [31, §3.2] and that of Rizzi presented in [32, Vol. A, §16.2].

The `term rank` of a matrix A is the maximum number of nonzero entries of A, no two in a line (row or column). We denote this basic combinatorial invariant of a matrix A by $\rho(A)$. It is direct that $\rho(A) \geq \mathrm{rank}(A)$.

Exercise 5.6. For any matrix A, $\rho(A)$ is equal to the minimum number of lines needed to cover all its nonzero entries.

Exercise 5.7. Let A be a real matrix such that $A + A^\top = J_n - I_n$. Show that $\rho(A) \geq n - 1$. (Indeed, a stronger claim holds, namely $\mathrm{rank}(A) \geq n - 1$,[24] which can be proved by considering the negative definite subspace of $J - I$.)

[23] This theorem is named after D. König (1884–1944), a professor at Budapest. The first graph theory textbook was written by this König in 1936.

[24] This fact has connection with a loop switching problem posed in Bell Laboratories; see [36, Chapter 9]. For relevant research on the Satisfiability Problem, see `http://cs-svr1.swan.ac.uk/~csoliver/papers.html`

Ryser obtained formulas of maximum and minimum term ranks for some important matrix classes; see his classic textbook, Combinatorial Mathematics, included in the Carus Mathematical Monograph series (Our library has its Chinese translation).

For some interesting results about rank, term rank, their relation with other graph parameters and a connection to communication complexity (as pointed out by Lovász), see the following papers and their references:

Donniell E. Fishkind, Andrei Kotlov, Rank, term rank, and chromatic number, Discrete Mathematics 250 (2002), 253–257.

Noam Nisan, Avi Wigderson, On rank vs. communication complexity, Combinatorica 15 (1995), 557–565.

Ambition is the last refuge of the failure. – Oscar Wilde

In Example 5.5, we apply Theorem 5.2 on two partition matroids associated with a bipartite graph. Let us try to put Theorem 5.2 into a more complex environment (Imagine that you are encoding in a more complex way which makes it harder to decode – you want to produce a nontrivial theorem. Anyway, this comparison is not very accurate. There are many meaningless correct results for which the most brilliant brains in this world find it too big a challenge. A deep result is not only difficult to establish, but also beautiful and useful.).

Let $G = (V_1, V_2, E)$ be a bipartite graph and $M_i = (V_i, \mathcal{I}_i)$, $i = 1, 2$, be two matroids. We say that a matching M of G is `independent` with respect to M_1 and M_2, provided $\partial(M) \cap V_i$ is independent in M_i, $i = 1, 2$. A `discrete matroid` is a matroid with all subsets of its ground set being independent.

For any given mapping $f : V_1 \to V_2$, we define a bipartite graph $G_f = (V_1, V_2, E)$ by putting $E = \{e_v : v \in V_1\}$ where $\partial(e_v) = \{v, f(v)\}$. We have been considering the partition matroid on E corresponding to the partition $E = \cup_{v \in V_2} \delta(v)$. Since for G_f there is a one-to-one correspondence between V_1 and E, that matroid is evidently isomorphic to the matroid on V_1 corresponding to the partition $V_1 = \cup_{v \in V_2} f^{-1}(v)$. With this transition of viewpoint, we are led to

Theorem 5.8. *Let $M_1 = (V_1, \mathcal{I}_1)$ be a matroid with rank function r_1 and let $M_2 = (V_2, 2^{V_2})$ be a discrete matroid. Given any mapping $f : V_1 \to V_2$, the maximum size of an independent matching of G_f with respect to M_1 and M_2 is* $\min_{U \subseteq V_2}(r_1(N(U)) + |V_2 \setminus U|)$.

Proof. Consider the partition matroid $M_3 = (V_1, \mathcal{I}_3)$ corresponding to the partition $V_1 = \cup_{v \in V_2} f^{-1}(v)$. , Clearly, the rank function r_3 of M_3 is given by $r_3(W) = |f(W)|$. Now note that the maximum size of an independent matching is nothing but $\max_{I \in \mathcal{I}_1 \cap \mathcal{I}_3} |I|$[25]. So, in view of Theorem 5.2, it is equal to $\min_{W \subseteq V_1}(r_1(W) + r_3(V_1 \setminus W)) = \min_{W \subseteq V_1}(r_1(W) + |f(V_1 \setminus W)|)$. Therefore, our remaining task is to establish the equality

$$\min_{W \subseteq V_1} \alpha(W) = \min_{U \subseteq V_2} \beta(U), \tag{5.8}$$

where $\alpha(W) = r_1(W) + |f(V_1 \setminus W)|$ and $\beta(U) = r_1(f^{-1}(U)) + |V_2 \setminus U|$. For any $U \subseteq V_2$, it is straightforward that $\alpha(W) = \beta(U)$ holds for $W = f^{-1}(U)$. Conversely, for any $W \subseteq V_1$, by putting $U = V_2 \setminus f(V_1 \setminus W)$, we have $\alpha(W) = r_1(W) + |f(V_1 \setminus W)| = r_1(W) + |V_2 \setminus U| \geq r_1(f^{-1}(U)) + |V_2 \setminus U| = \beta(U)$. This validates (5.8), as desired. □

[25] We have this quantity in Theorem 5.2 and we encode it into the quantity about independent matching to reveal some "new" fact. In presenting the proof, this process is reversed.

A result having close connection with Theorem 5.8 is

Theorem 5.9. *Let $M_1 = (V_1, \mathcal{I}_1)$ be a matroid with rank function r_1 and let $M_2 = (V_2, 2^{V_2})$ be a discrete matroid. Given any mapping $f : V_1 \to V_2$, let $\mathcal{I}_2 = \{S : S = \partial(M) \cap V_2, M$ is an independent matching of G_f with respect to M_1 and $M_2\}$. Then $(V_2, \mathcal{I}_2)$ is a matroid with rank function r_2 such that $r_2(U)$ is the maximum size of an independent matching with respect to M_1 and M_2 the intersection of whose boundary with V_2 lies in U.*

Generalizing Theorem 5.8[26], we now state

Theorem 5.10. *[3] There exists an independent matching of size k in a bipartite graph $G = (V_1, V_2, E)$ with respect to two matroids $M_i = (V_i, \mathcal{I}_i)$, $i = 1, 2$, if and only if $r_1(N(U)) + r_2(U) \geq k$ [27] holds for all $U \subseteq V_1$, where r_i is the rank function of M_i, $i = 1, 2$.*

Exercise 5.11. Prove Theorem 5.9.

Exercise 5.12. Prove Theorem 5.10.

[26] Recall our quotation from Spencer on page 53 once again.

[27] Since M_1 and M_2 play symmetric role, we know that $r_1(N(U)) + r_2(U)$ can be replaced by $r_1(U) + r_2(N(U))$. Especially, when returning to Theorem 5.8, we should have $\min_{U \subseteq V_2} r_1(N(U)) + |U| = \min_{U \subseteq V_2} r_1(U) + |N(U)|$.

Given k matroids $M_i = (X_i, \mathcal{I}_i)$, $i = 1, \cdots, k$, we define their `direct sum`, denoted $M_1 + \cdots + M_k$, to be the matroid $(X, \mathcal{I})$, where $X = \cup_{i=1}^{k} X_i \times \{i\}$, and $S \subseteq X$ is a member of $\mathcal{I}$ if and only if $S \cap (X_i \times \{i\}) \in \mathcal{I}_i \times \{i\}$ for $i = 1, \cdots, k$. Clearly, $M_1 + \cdots + M_k$ is a matroid and its rank function is given by $r(S) = \sum_{i=1}^{k} r_i(S_i)$, where $S_i = S \cap X_i \times \{i\})$ and r_i is the rank function of M_i, $i = 1, \cdots, k$.

Proof. (of Theorem 5.1) Consider the matroid $M_1 + \cdots + M_k$ whose ground set is a disjoint union of k copies of X, denoted $X_1 \cup \cdots \cup X_k$, where $X_i = X \times \{i\}, i = 1, \cdots, k$. Let $f : X_1 \cup \cdots \cup X_k \to X$ be a mapping such that $f(x, i) = x$. Now we apply Theorem 5.8 and Theorem 5.9 for the resulting G_f and the result follows. □

For any graph G with k connected components and n vertices, we put $\xi(G) = n - k$. This number is the dimension of the cut space of the graph as well as the size of any maximal independent set in the corresponding graphic matroid.

Exercise 5.13. A connected graph G contains k edge-disjoint spanning trees if and only if for any subgraph H of G, $k(\xi(G)-\xi(H)) \leq |E(G)|-|E(H)|$.

Exercise 5.14. A connected graph G may be split up into at most k trees if and only if for any subgraph H of G, $k\xi(H) \geq |E(H)|$.

In passing, I firmly believe that research should be offset by a certain amount of teaching, if only as a change from the agony of research. The trouble, however, I freely admit, is that in practice you get either no teaching, or else far too much. – J.E. Littlewood, (1885 -1977)
Life is good for only two things, discovering mathematics and teaching mathematics. – Siméon Poisson (June 21, 1781 – April 25, 1840)

Example 5.15. The Shannon switching game is an abstract game for two players, a Spanner and a Cutter, invented by Claude Shannon. Given a finite graph G with two special nodes, F and M. The Cutter deletes edges of G and the Spanner seizes them, one per move. The Spanner aims to seize a set S which connects F and M in G, and the Cutter aims to prevent this. The Cutter moves first. It is easy to see that the Spanner has a winning strategy provided he can find in the graph G two disjoint subtrees of G on the same vertex set $V \subseteq V(G)$ such that $\{F, M\} \subseteq V$. The hard side is that if there are no such two trees in G, then the Cutter has a winning strategy. Please refer to `http://www.math.psu.edu/melvin/logic/shannon.pdf` for an elementary (and hence hard) proof of this claim.

A. Lehman, a Fulkerson prize winner, generalized the Shannon switching game on graphs to a game on matroids. Let $M = (X, \mathcal{I})$ be a matroid with rank function r and take $x \in X$. Recall that the span of $S \subseteq X$ in M is $Span(S) = \{x \in X : r(S) = r(S \cup \{x\})\}$. As before, the Cutter and the Spanner alternatively move, deletes one element or seizes one element each time. Now the Spanner aims to seize a set which spans x and the Cutter acts as his opponent. When we take M to be a graphic matroid this is just the game devised by Shannon.

Lehman proved in his paper "A solution of the Shannon switching game. J. Soc. Indust. Appl. Math, 12 (1964), 687–725" that the Spanner has a winning strategy in the generalized Shannon switching game if and only if there are two disjoint subsets X_1, X_2 of $X \setminus \{x\}$ such that $x \in Span(X_1) = Span(X_2)$. It is not difficult to generalize the strategy used in the case of graphic matroid to formulate a winning strategy in the generalized game. But the necessity of the condition is much harder. It can be proved by the Matroid Union Theorem.

For another generalization, the Shannon switching game on digraphs, together with its connection with the theory of oriented matroid, see `http://www.ecp6.jussieu.fr/pageperso/las_vergnas/papers/conj_E.html`.

6 Matching Polytope Theorem

For any graph G, we define its `matching polytope` $P_{\text{matching}}(G)$ to be the convex hull of the incidence vectors of all matchings of G over $E(G)$. A matching M of a graph G is `perfect` provided $\partial(M) = V(G)$[28]. Similarly, the `perfect matching polytope` $P_{\text{perfect matching}}(G)$ is the convex hull of the incidence vectors of all perfect matchings of G over $E(G)$.

The main theorem here will be Edmonds' Matching Polytope Theorem – its deduction and applications will be wonderful examples of polyhedral method, as summarized on page 88 and page 117. Since this is the last main goal of this course and we do have some time left now, we will not be urgent to step on the highway to it. We will deviate a bit often to include more relevant results. In particular, we will start from the much easier matching polytope theorem for bipartite graphs and show that such an innocent-looking result already has many unexpected applications.

[28] In the physics community, perfect matching is sometimes known as "dimer cover".

6.1 Bipartite graphs

These seemingly disparate questions are, in fact, closely related: they all concern sets of independent edges, called matchings, in bipartite graphs, and are answered by the same basic theorem in various guises, attributed to Hall, König and Egerváry. This theorem, which we call Hall's marriage theorem, is a prime example of several results we shall present in this chapter giving necessary and sufficient conditions for the existence of certain objects; in which case **the beauty of the theorem is that a condition whose necessity is obvious is shown to be also sufficient.** *In the natural formulation of our results we shall have two functions, sat f and g, clearly satisfying $f \leq g$, and we shall show that* $\max f = \min g$. *The results of this chapter are closely interrelated, and so the order they are proved in is a matter of taste; to emphasize this, some results will be given several proofs.* – Béla Bollobás, Modern Graph Theory, Springer-Verlag, 2002, page 67.

When G is bipartite, a linear inequality description of $P_{\text{matching}}(G)$ is an immediate consequence of what we learned from Chapter 2.

Theorem 6.1. *For a bipartite graph G, $P_{\text{matching}}(G)$ is the polytope defined by*

$$\begin{cases} x \in \mathbb{R}_+^{E(G)}; \\ \sum_{e\in\delta(v)} x(e) \leq 1, \forall v \in V(G). \end{cases} \tag{6.1}$$

Proof. Let Q be the polytope defined by (6.1). It is easy to see that $P_{\text{matching}}(G) \subseteq Q$. Further note that each integral solution to (6.1) must be the incidence vector of some matching of G. Thus, to prove the reverse inclusion $Q \subseteq P_{\text{matching}}(G)$, it suffices to show that Q is integral [29]. But this fact is clearly guaranteed by Theorem 2.6 and Theorem 2.25. That is all the proof. □

Exercise 6.2. For a bipartite graph G, $P_{\text{perfect matching}}(G)$ is the polytope defined by

$$\begin{cases} x \in \mathbb{R}_+^{E(G)}; \\ \sum_{e\in\delta(v)} x(e) = 1, \forall v \in V(G). \end{cases} \tag{6.2}$$

Exercise 6.3. Theorem 6.1 is no longer true when the bipartiteness condition is deleted. (hint: Consider an odd cycle with each edge weighted $\frac{1}{2}$.)

Exercise 6.4. The `edge cover polytope` $P_{\text{edge cover}}(G)$ of a graph G is the convex hull of the incidence vectors of the edge covers of G. For any bipartite graph G, give a linear inequality description of $P_{\text{edge cover}}(G)$.

[29] Recall that we have used the same type of arguments when proving Theorems 5.3 and 3.5. Also note that Corollary 1.52 is implicitly used.

We now come to

Theorem 6.5. *(Egerváry's Theorem [10]) Let $G = (V, E)$ be a bipartite graph and let $w: E \to \mathbb{R}_+$ be a weight function. Then the maximum weight of a matching is equal to the minimum value of $\sum_{v \in V} y(v)$, where y ranges over all functions $y: V \to \mathbb{R}_+$ satisfying $y(u) + y(v) \geq w(e)$ for each edge e with $\partial(e) = \{u, v\}$. Furthermore, if w is integer-valued, then we can take also y integer-valued.*

Proof. This is a standard application of polyhedral methods and the key step has been finished in Theorem 6.1. Note that the 'furthermore' part follows from Theorem 4.18 and Theorem 2.25. [30] □

When w assigns constant weight 1 to each edge of G, Theorem 6.5 turns out to be just König's matching theorem (Example 5.5). This is very natural, as in both cases we have been proving the integrality of the same polyhedron. In Example 5.5, if we do not use Theorem 5.2 but use directly its equivalent Theorem 5.3, we will also be led to Egerváry's Theorem. König's matching theorem, or its equivalent, Exercise 5.6, is often called König-Egerváry Theorem.

König's matching theorem can be used to deduce the existence of Haar measure on locally compact topological groups; see [26, Theorem 2.5.6]. For its application in establishing Strassen's Monotone Coupling Theorem in measure theory, see [26, Theorem 2.5.2], or
`http://galton.uchicago.edu/~lalley/Courses/388/Matching.pdf`

[30] If, unfortunately, you still have no idea of the proof, please refer to [31, Corollary 3.5b, Theorem 3.6] and be advised that you have to work a bit harder from now on. Note that [31, Theorem 3.6] is proved there without reference to the concept of TDI.

The marriage theorem, proved in 1935 by Philip Hall (1904 – 1982), answers the following question, known as the marriage problem [31]: if we have a finite set of girls each of whom knows several boys, under what conditions can we marry off these girls in such a way that each girl marries a boy she knows (We shall assume that polygamy is not allowed). The Hall's marriage theorem asserts that such a marriage is possible if and only if the Hall's condition is satisfied, namely, for each set G of girls, the set B of their boyfriends has a size at least as large as G (We assume that if x knows y, then y is a friend of x.).

Exercise 6.6. Use König-Egerváry Theorem (Example 5.5 or Theorem 6.5) to prove Hall's marriage theorem.

For a digraph Γ and any $S \subseteq V(\Gamma)$, define $N^+(S)$ to be $\{u : \exists v \in S, e \in E(\Gamma), \mathtt{t}(e) = u, \mathtt{i}(e) = v\}$.

Exercise 6.7. Let D be a digraph. Prove that there exist pairwise disjoint cycles in D such that each vertex of D lies in exactly one of the cycles if and only if $|N^+(S)| \geq |S|$ for all $S \subseteq V(D)$.

Hall's Marriage Theorem provides a necessary and sufficient criterion for a bipartite graph to have a perfect matching but checking Hall's condition will be computationally infeasible (This does not mean that Hall's Theorem is of little use.). The so-called Hungarian method [32] developed by Kuhn, using ideas of Egerváry, enables us find a maximum matching in any bipartite graph efficiently (and hence determine the existence of perfect matching efficiently.); please go to [31, §3.4] if it is not covered in your algorithm course yet. Also note that for general graphs, we will have Edmonds' matching algorithm, as mentioned already on page 124.

[31] My original plan for this course is to talk about Stable Marriage based on the book of Knuth and some other materials. I hope that I have a chance to talk about it in an optional course in the future. To make it possible, there should be enough students who are willing to sit in my talks......

[32] A very nice expository article on it is: András Frank, On Kuhn's Hungarian Method - a tribute from Hungary, TR-2004-14, EGRES Technical Report, `http://www.cs.elte.hu/egres/`

EGRES is a famous research group. The following is copied from its homepage `http://www.cs.elte.hu/egres/`:

Eugene Egerváry (in Hungarian, [33] Egerváry Jenö) can certainly be considered as one of the founding fathers of what is called today Combinatorial Optimization. Extending earlier works of Dénes König, his classic paper (On combinatorial properties of matrices, Mat. Lapok, 38, 1931. 16-28. (Hungarian with German summary)) describes the duality theorem for the weighted bipartite matching problem, which is called today the assignment problem, proves the integrality result, and develops the underlying idea of the first primal-dual type algorithm that is called throughout the literature the Hungarian Method, a name introduced by H. Kuhn who actually used Egerváry's ideas to develop an efficient algorithm. It is remarkable that this algorithm turned out to be strongly polynomial.

These ideas of Egerváry have been the prototype of a huge body of later research in areas like network flows, linear programming, matroids optimization, or matching theory. Even in as sophisticated frameworks as the submodular flows the original ideas of Egerváry could be extended.

Given the significance of the Hungarian Method, here in Budapest, the city of Egerváry, several researchers, seniors, post docs and Ph.D. students, felt obliged to establish a research group called Egerváry Research or, in short, EGRES. This is only an informal group not belonging to any one institution and its members actually work for several institutes. We started to run a series of EGRES technical reports to present our results. ...

Our main goal is to follow the tradition laid down by Egerváry and work on combinatorial optimization problems where nice characterizations, min-max theorems and/or polynomial time algorithms can be given. ... One of our basic interests is to explore further the borderline of NP-complete and polynomially solvable problem classes in combinatorial optimization.

[33] Like our Chinese, Hungarian use their family names as the first names.

For some light reading on Hall's Theorem, let me direct you to `http://episte.math.ntu.edu.tw/articles/mm/mm_10_1_09/index.html`

Note that all mathematics appeared in the above paper written by Prof. Chang, a former head of Dept. of Math. in Chiao Tung Univ., can be found in our textbook [31]. Indeed, that web page tells the story in such a way to entice you into a serious study of some stuff like [31].

If you like such reading material as the online paper of Chang, perhaps you should try to search often yourselves in the internet. Or you can go to our library and take an issue of American Mathematical Monthly. Unfortunately, we have little such Chinese sources in China mainland.

A widely adopted proof of Hall's Marriage Theorem is that given in:

[34]P.R. Halmos, H.E. Vaughan, The marriage problem, Amer. J. Math. 72 (1950), 214–215.

There are lots of other proofs, especially another common approach is to deduce it from the Max-flow Min-cut Theorem of Ford and Fulkerson [35]. Here are two 'new' proofs:

Thierry Coquand, A syntactical proof of the marriage lemma, Theoretical Computer Science, 290 (2003), 1107–1113.

Zhiwei Sun, Hall's theorem revisited, Proc. Amer. Math. Soc., 129 (2001), 3129–3131.

There are many directions along which peoples are trying to generalize this beautiful marriage theorem–this is no wonder as we obtain Hall's Theorem in a specializing process. Especially, I very much like the papers collected here[36]:

`http://www.math.technion.ac.il/~ra/isr.html`

I think it is said that Gauss had ten different proofs for the law of quadratic reciprocity. Any good theorem should have several proofs, the more the better. For two reasons: usually, different proofs have different strengths and weaknesses, and they generalise in different directions – they are not just repetitions of each other. – Sir Michael Atiyah

---

[34]This paper is collected into [15].

[35]I understand that it must have been introduced in your algorithm course. In [31], this is proved by polyhedral method in Corollary 8.4b and algorithmically in Corollary 4.5c. I am very sad that we are not allowed to go into this theorem according to the current situation. Many curious combinatorial theorems can be proved by devising some digraphs with flow capacity constraints and then using the existence of integer flows as guaranteed by the Max-flow Min-cut Theorem.

[36]Please let me know if you also find them interesting and want me print out a copy of any paper there for you.

A nonnegative matrix is called `doubly stochastic` [37] if each of its row sum and each of its column sum is equal to 1. It is obvious that doubly stochastic matrices are all square matrices. A `permutation matrix` is a doubly stochastic matrix of 0's and 1's.

Exercise 6.8. (i) (Birkhoff–Von Neumann Theorem, 1946, 1953) For each doubly stochastic matrix M, there are permutation matrices $P_1, \cdots, P_m$ and nonnegative reals $\lambda_1, \cdots, \lambda_m$, such that $\sum_{i=1}^{m} \lambda_i = 1$ and $M = \sum_{i=1}^{m} \lambda_i P_i$. (hint: Go to Exercise 6.2 by considering the complete bipartite graph whose partite sets are the rows of M and the columns of M, respectively, and whose edge connecting i and j receives weight $M(i,j)$. Note that an element of $\mathbb{R}^{E(G)}$ is just a weight function of $E(G)$.)

(ii) Prove that each doubly stochastic matrix of order n is a convex combination of at most $n^2 - 2n + 2$ permutation matrices. (hint: Theorem 1.36.)

The emphasis on mathematical methods seems to be shifted more towards combinatorics and set theory - and away from the algorithm of differential equations which dominates mathematical physics. – J. von Neumann and O. Morgenstern, Theory of Games and Economic Behavior, Princeton University Press, 1944.

[37] Here is an interesting paper about this common-looking concept: J.D. Louck, Doubly stochastic matrices in quantum mechanics, Foundations of Physics, 27 (1997), 1085–1104.

Exercise 6.9. (Dilworth's Decomposition Theorem) [38] Let P be a finite partially ordered set. If there is no antichain on $m+1$ elements in P, then P is a union of m chains.

Exercise 6.10. (Alon & Tarsi) Prove that if $\frac{|E(H)|}{|V(H)|} \leq d$ for each subgraph H of a graph G, then G has an orientation in which every vertex has out-degree at most d.

Let $\Delta(G)$ denote the maximum degree of a graph G. Let $\chi_e(G)$ stand for the minimum number of matchings required to cover all edges of a graph G. $\chi_e(G)$ is the `edge colouring number` of G.

Exercise 6.11. (König's edge coloring theorem) For every bipartite graph G, $\chi_e(G) = \Delta(G)$.

Exercise 6.12. Let $A_1, \cdots, A_n$ be finite sets. Show that if $\sum_{1 \leq i < j \leq n} \frac{|A_i \cap A_j|}{|A_i| \cdot |A_j|} < 1$, then the sets $A_1, \cdots, A_n$ have a system of distinct representatives.

It is difficult to decide where to stop in any treatment of the theorems of König and P. Hall. Entire books can be written on their ramifications.
– [26], page 40.

[38] Using Tychonov's Theorem on product topology, you will find that the finiteness condition here can be removed.

At the beginning of this course, I told you how to play the Rubik's Cube (or magic cube, invented by Ernö Rubik) without opening your eyes. Since you have suffered a lot in this course, I am afraid that I should try to make you a bit happier than usual and thus let us turn to two more games. Note that the hidden serious background of my Rubik's Cube story is synchronizing automata and symbolic dynamics. This time, the underlying mathematics is the structure of matching polytope of complete bipartite graphs.

Magic Square [39]

The enumeration of magic squares is a classical problem in combinatorics dating at least to McMahon in the early 20th century.

McMahon obtained an explicit formula for the number $h(n)$ of 3×3 magic squares of weight n. Much later, Richard Stanley proved that for every $m \geq 2$, the number of $m \times m$ magic squares of weight n is a polynomial function of n and that the degree of the polynomial is $(m-1)^2$. This implies that the function is determined by $(m-1)^2$ values, by the Lagrange interpolation formula. You are encouraged to produce a proof of Stanley's Theorem with possible help from

`http://galton.uchicago.edu/~lalley/Courses/388/HW1.pdf`.

Here we shall present a proof of MacMahon's formula of which the key step is to investigate the perfect matching polytope of the complete bipartite graph $K_{3,3}$.

[39] written according to `http://galton.uchicago.edu/~lalley/Courses/388/Matching.pdf`

We have defined on page 85 the matrix

$$N = N_3 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}.$$

We further let

$$T_3 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, T_2 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}, T_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}.$$

I, N, N^2, T_1, T_2 and T_3 correspond to the vertices of the perfect matching polytope of $K_{3,3}$.

Let A be any doubly stochastic matrix of order 3. The Birkhoff-Von Neumann Theorem says that A is a convex combination of I, N, N^2, T_1, T_2 and T_3. With the help of the Caratheodory's Theorem (Theorem 1.36), we assert in Exercise 6.8 (ii) that A can be expressed as a convex combination of at most five such matrices among all six 3×3 permutation matrices. Can we say any more about the expression?

To enumerate magic square of order 3 and weight n, we are actually considering the polyhedral cone spanned by I, N, N^2, T_1, T_2 and T_3–these six vectors are the extreme rays of the resulting cone! We will show that there is a standard representation for each vector in this cone and thus we transform the problem to a problem of enumerating the number of standard representations.

Lemma 6.13. *N, N^2, T_1, T_2 and T_3 are linearly independent.*

Proof. Suppose $\sum_{i=1}^{3} a_i T_i + b_1 N + b_2 N^2 = 0$. Looking at the (i,i) position, we deduce that $a_i = 0$, $i = 1, 2, 3$. But the nonzero positions of N and N^2 are disjoint and hence $b_1 = b_2 = 0$ also follows. □

Lemma 6.14. *$\sum_{i=1}^{3} a_i T_i + \sum_{i=1}^{3} b_i N^i = \sum_{i=1}^{3} a_i' T_i + \sum_{i=0}^{2} b_i' N^i$ if and only if there is a constant c such that $a_i - a_i' = b_i' - b_i = c$ for $i = 1, 2, 3$.*

Proof. Follows from $\sum_{i=1}^{3} T_i = \sum_{i=1}^{3} N^i$ and Lemma 6.13. □

Lemma 6.15. *Each $R \in \mathbb{R}_+^{3\times 3}$ of constant line sum n has a unique representation as $R = \sum_{i=1}^{3} b_i N^i + \sum_{i=1}^{3} a_i T_i$ such that $a_i, b_i \in \mathbb{R}_+$, $\sum_{i=0}^{2} b_i + \sum_{i=1}^{3} a_i = n$ and $b_1 b_2 b_3 = 0$.*

Proof. The existence of such an expression is a consequence of $\sum_{i=1}^{3} T_i = \sum_{i=1}^{3} N^i$ and Birkhoff-Von Neumann Theorem. Lemma 6.19 gives the uniqueness. □

The next lemma takes care of the integrality.

Lemma 6.16. *Every* 3×3 *magic square* R *of weight* n *has a unique representation as* $R = \sum_{i=1}^{3} b_i N^i + \sum_{i=1}^{3} a_i T_i$ *such that* $a_i, b_i \in \mathbb{Z}_+$, $\sum_{i=0}^{2} b_i + \sum_{i=1}^{3} a_i = n$ *and* $b_1 b_2 b_3 = 0$.

Proof. By checking the proof of Lemma 6.16, we only need to show the existence of the expression of R as a nonnegative integer linear combination of I, N, N^2, T_1, T_2 and T_3. Consider the bipartite graph G on the partite sets $\{x_1, x_2, x_3\}$ and $\{y_1, y_2, y_3\}$ such that there are $R(i,j)$ edges connecting x_i and y_j. Our task is then translated into the assertion that $E(G)$ is a disjoint union of n perfect matchings (1-factors).[40]

In G there are totally $3n$ vertices and each vertex is incident with n edges. This means that the minimum vertex cover of G has size 3. Consequently, by König's matching theorem (Example 5.5) we know that G has a perfect matching (1-factor). Removing this perfect matching from G yields an $(n-1)$-regular graph, which, by the same argument as before, still has a perfect matching. Continuing in this fashion, we arrive at the required conclusion. □

[40]Our following arguments indeed show that each regular bipartite graph can be decomposed into disjoint 1-factors. Have you already established it in solving Exercise 6.13?

Theorem 6.17. *(MacMahon's formula)* $h(n) = 3\binom{n+3}{4} + \binom{n+2}{2}$.

Proof. By Lemma 6.16 and the principle of inclusion-exclusion[41],

$$h(n) = \sum_{S\neq\emptyset, S\subseteq\{1,2,3\}} (-1)^{|S|-1} h_S(n), \tag{6.3}$$

where $h_S(n)$ denotes the number of six-tuples $(m_1, \cdots, m_6)$ of nonnegative integers summing to n such that $m_i = 0$ for $i \in S$. $h_1(n) = h_2(n) = h_3(n)$ is the number of ways of placing n identical balls into 5 urns, which is clearly $\binom{n+4}{4}$. To see it, you can just think of a bijection between such way of putting balls into urns and the way of selecting 4 points from a line of $n+4$ points[42]. Another method is to use generating function. Let $H_i(z) = \sum_{n=0}^{\infty} h_i(n)z^n$. Then $H_i(n) = (\sum_{n=0}^{\infty} z^n)^5 = \frac{1}{(1-z)^5} = \sum_{n=0}^{\infty} \binom{n+4}{4} z^n$. Comparing the coefficients of z^n again gives the formula for $h_i(n)$. Similarly, we have $h_{12}(n) = h_{23}(n) = h_{31}(n) = \binom{n+3}{3}$ and $h_{123}(n) = \binom{n+2}{2}$. Plugging these into Eq. (6.3) yields $h(n) = 3\binom{n+4}{4} - 3\binom{n+3}{3} + \binom{n+2}{2} = 3\binom{n+3}{4} + \binom{n+2}{2}$.[43] This is McMahon's result. □

A generating function is a clothesline on which we hang up a sequence of numbers for display. – Herbert S. Wilf, Generatingfunctionology, 2nd ed., Academic Press, Inc., 1994. [44]

[41]This well-known principle, though being a simplest among the so-called 'sieve methods', can lead you to some wonderland; see, M. Loebl, On the inclusion-exclusion principle (Lecture text for Spring School on Combinatorics 2000), available at: `http://kam.mff.cuni.cz/~kamserie/kamser-search.cgi?yearlist+Y-kamser-2000+Y-kamser-2001`

[42]Gian-Carlo Rota described combinatorics as "putting different colored marbles in different colored boxes, seeing how many ways you can divide them" and lifted it from a barely respectable obscurity to one of the most active areas of mathematics today.

[43]I hope that you know how to prove the formula $\binom{a}{b} = \binom{a-1}{b-1} + \binom{a-1}{b}$ combinatorially, namely using balls and boxes.

[44]A Chinese translation of it was published by Tsinghua University Press last year!

The Best Card Trick: First, have a look at the following quotes; then, let me explain the trick to you; but for a more lively discussion of this parlor trick, I still recommend the paper of Michael Kleber mentioned below; After reading that paper, please go to [36, Example 5.1] for a very clear mathematical analysis (Do not read it first as you do not like to know the final result before you watch a football game.).

Come one, come all, to see the best card trick there is – you have to see it to believe it. But fear not: I'll tell you how it's done and a little bit about the Birkhoff-Von Neumann Theorem, Hall's Marriage Theorem and Fitch Cheney, the first guy to ever get a math PhD from MIT, too.
– poster of the talk "Many Cheerful Facts" by Michael Kleber

'The Best Card Trick' is the title of a wonderful article by Michael Kleber (Brandeis univ) which appeared in the Mathematical Intelligencer, vol. 24, No1, 2002. The trick (due to William Fitch Cheney, Jr., who received the first MIT PhD in 1927) is this: choose any 5 cards from a 52 pack of cards, and pass them to my assistant, who-having looked at them-passes a certain 4 of them to me. I can then tell you the card that my assistant still holds. The question is: how is it done? In fact the trick can be done with a 124-pack of cards, still choosing 5. In general it can be done chooing 'n' cards from an $(n!+n-1)$-pack of cards. That paper-in which MK presents Elwyn Berlekamp's brilliant solution-is available at:

`http://people.brandeis.edu/~kleber/Papers/card.pdf`

Ian Stewart also wrote about this trick (with acknowledgement to MK) in the Jan. 4th 2003 issue of the New Scientist.
I have just introduced this topic in a new 3rd year course of mine (Challenging Mathematical Puzzles and Problems), and have prepared a related Maple worksheet (there's a html version for anyone who doesn't have Maple) which is available in this corner of my web site:

`http://www.spd.dcu.ie/johnbcos/challenging.htm`

It's a simple (but captivating) way to see some elementary Number Theory in action.
John Cosgrave (A post at Number Theory List)

Theorem 6.18. *(Mendelsohn-Dulmage Theorem [27]) Let V_1 and V_2 be disjoint sets. For $i = 1, 2$, a set $E \subseteq V_1 \times V_2$ is called V_i-admissible if each point of S_i appears in at most one pair of points in E. Let $E_1, E_2 \subseteq V_1 \times V_2$ and assume that E_i is V_i-admissible for $i = 1, 2$. Then there is $E \subseteq E_1 \cup E_2$ which are both V_1-admissible and V_2-admissible and covers all nodes of V_2 covered by E_1 and all nodes of V_1 covered by E_2.*

Note that Theorem 6.18 is a generalization of the Schröder-Bernstein theorem you will encounter in the course Real Analysis and Topology (if you major in MATH) and the course Logic (if you major in CS). There are still much deeper generalizations!

A very short proof of Theorem 6.18 is possible. But a proof of it (indeed some generalization) using polyhedral combinatorics can be found in [18].

S. Kundu, E.L. Lawler, A matroid generalization of a theorem of Mendelsohn and Dulmage, Discrete Mathematics, 4 (1973), 150–163.

D. Gale, L.S. Shapley, College admissions and stability of marriage, Amer. Math. Monthly, 69 (1962), 9–15.

A.J. Hoffman, H.W. Kuhn, On systems of distinct representatives, Linear Inequalities and Related Systems, (Eds., H.W. Kuhn, A.W. Tucker) Princeton University Press, N.J., 1956, pp. 199–206.

J.S. Pym, The linking of sets in graphs, J. London Math. Soc., 44 (1969), 542–550.

Let $G = (V, E)$ be a digraph and let $X_0 \subseteq X \subseteq V$, $Y_0 \subseteq Y \subseteq V$ be such that X_0 is linked to some subset of Y in G and some subset of X is linked to Y_0. Then X^* is linked to Y^* for some X^*, Y^* with $X_0 \subseteq X^* \subseteq X$ and $Y_0 \subseteq Y^* \subseteq Y$. (Gammoids are matroids!)

Tamás Fleiner, A fixed-point approach to stable matchings and some applications, Mathematics of Operations Research, 28 (2003), 103–126.

We have met many Hungarians in this course: D. König, E. Egerváry, John von Neumann, George Pólya, Béla Bollobás, Ernö Rubik, L. Lovász, Paul Erdös, Paul Halmos, Alfred Haar, Tamás Fleiner...

Clearly, the Hungarian educational system has been the most successful for pure mathematics; it's a model that ought to be studied very carefully because it works. — Donald Knuth, Two Year College Mathematics Journal, Vol. 13

They are already here...they are called Hungarians! – Enrico Fermi, when asked if he believed in extraterrestrials.

6.2 General graphs

You are strongly advised to have a look at the preface and the table of contents (or still more!) of [26].

Concerning the importance of the Matching Polytope Theorem, besides the comment of Cunningham on page 90, we cite the following:

Such a set of linear inequalities was found by Edmonds. His result allows us to obtain various minimax theorems in matching theory as special cases of the Duality Theorem of linear programming. This approach initiated the study of other combinatorially defined polyhedra and has led to a whole new branch of combinatorial mathematics – polyhedral combinatorics. – [26, page xx]

Let $\mathcal{P}_{\text{odd}}(V)$ denote the collection of odd subsets of a set V.

Theorem 6.19. *(Edmonds' Perfect Matching Polytope Theorem [7]) For any graph G, $P_{perfect\ matching}(G)$ is fully determined by the following set of inequalities*

$$\begin{cases} \text{(i) } x \in \mathbb{R}_+^{E(G)}; \\ \text{(ii) } \sum_{e \in \delta(v)} x(e) = 1, \sum_{e \in \gamma(v)} x(e) = 0, \forall v \in V(G); \\ \text{(iii) } \sum_{e \in \delta(U)} x(e) \geq 1, \forall U \in \mathcal{P}_{odd}(V(G)), |U| \geq 3. \end{cases} \tag{6.4}$$

By taking U to be the whole vertex set, (6.4)(iii) coincides with the trivial observation that a graph with an odd number of vertices cannot have any perfect matching.

Here is an equivalent formulation of Theorem 6.19.

Theorem 6.20. *(Edmonds' Perfect Matching Polytope Theorem [7]) For any graph G, $P_{perfect\ matching}(G)$ is fully determined by the following set of inequalities*

$$\begin{cases} \text{(i) } x \in \mathbb{R}_+^{E(G)}; \\ \text{(ii) } \sum_{e \in \delta(v)} x(e) = 1, \forall v \in V(G); \\ \text{(iii) } \sum_{e \in \gamma(U)} x(e) \leq \lfloor \frac{|U|}{2} \rfloor, \forall U \in \mathcal{P}_{odd}(V(G)). \end{cases} \tag{6.5}$$

Theorem 6.19 ⇢ **Theorem 6.20** (6.4 ⇢ (6.5)):

Proof. For odd cardinality U, $\sum_{e \in \gamma(U)} x(e) = \frac{1}{2}(\sum_{v \in U} \sum_{e \in \delta(v)} x(e) - \sum_{e \in \delta(U)} x(e)) \leq \frac{1}{2}(|U| - 1) = \lfloor \frac{1}{2}|U| \rfloor$. □

Theorem 6.20 ⇢ **Theorem 6.19** (6.5 ⇢ (6.4)):

Proof. $\sum_{e \in \gamma(v)} x(e) \leq \lfloor \frac{1}{2}|\{v\}| \rfloor = 0$; For odd cardinality U, $\sum_{e \in \delta(U)} x(e) = \sum_{v \in U} \sum_{e \in \delta(v)} x(e) - 2 \sum_{e \in \gamma(U)} x(e) \geq |U| - 2\lfloor \frac{1}{2}|U| \rfloor \geq 1$. □

The third constraint in (6.5), a system of inequalities for odd cardinality sets, is often called the `blossom inequalities`, as they were originally formulated when Edmonds discovered his blossom algorithm for finding a maximum weight matching in a general edge-weighted graph.

The advantage of the form of Theorem 6.20 is that it is self-refining, as can be seen from the next immediate consequence of it.

Theorem 6.21. *(Edmonds' Matching Polytope Theorem [7]) For any graph* G, $P_{matching}(G)$ *is fully determined by the following set of inequalities*

$$\begin{cases} x \in \mathbb{R}_+^{E(G)}; \\ \sum_{e \in \delta(v)} x(e) \leq 1, \forall v \in V(G); \\ \sum_{e \in \gamma(U)} x(e) \leq \lfloor \frac{1}{2}|U| \rfloor, \forall U \in \mathcal{P}_{odd}(V(G)). \end{cases} \tag{6.6}$$

We postpone the proof of Edmonds' Perfect Matching Polytope Theorem until later and go into some of its applications first.

The polytope consisting of the feasible solutions to (6.4)(i) and (ii) is called the `fractional perfect matching polytope` of G, denoted $P^*_{\text{perfect matching}}(G)$. Exercise 6.2 says that $P_{\text{perfect matching}}(G) = P^*_{\text{perfect matching}}(G)$ when G is bipartite. Conversely, strengthening Exercise 6.3, we have

Exercise 6.22. [45] $P_{\text{perfect matching}}(G) \subsetneq P^*_{\text{perfect matching}}(G)$ whenever G is nonbipartite.

Let us be grateful to people who make us happy; they are the charming gardeners who make our souls blossom. – Marcel Proust (10 July 1871 – 18 November 1922) French novelist, author of *Remembrance of Things Past*

[45] Compare with Exercise 2.25.

The `degree` of a vertex v in a graph G is $\deg_G(v) = |\delta(v)| + 2|\gamma(v)|$. A `cubic` graph is a graph of constant degree 3. A graph G is `2-edge connected` if $|\delta(U)| \geq 2$ for all nonempty proper subset U of $V(G)$.

Theorem 6.23. *(Petersen's Theorem[46]) Each 2-edge connected cubic graph has a perfect matching.*

Exercise 6.24. The edge colouring number of any 2-edge connected cubic graph is three.

According to Luis Goddyn[47], the only known proof of the following generalization of Theorem 6.23 relies on the foregoing description of the perfect matching polytope.

Theorem 6.25. *Each edge in a 2-edge connected cubic graph G belongs to a perfect matching of G.*

Exercise 6.26. In Theorem 6.25, the 2-edge connectedness condition cannot be weakened to be 1-edge connected.

Although it can be argued that such famous names as Euler, Kirchhoff and Tait can be found in the historical shadows of matching theory, we shall take as the two principle "founders" of the decipline the Dane, Julius Petersen and the Hungarian, Dénes König. Although their interests certainly overlapped, it is perhaps helpful to identify Petersen with the earliest study of regular graphs (that is, graphs having the same degree at each point) and König with bipartite graph. – [26, page xi]

---

[46]Indeed, Petersen proves a more general result: Every cubic graph with at most three bridges has a perfect matching.

[47]see the notes of a Graph Theory course offered by him available at `http://www.math.sfu.ca/~goddyn/Courses/820/polytopes.pdf`

Proof. (of Theorem 6.25) Let G be 2-edge connected and cubic. Let $\overline{x} \in \mathbb{R}^{E(G)}$ be the constant weight function which assigns weight $\frac{1}{3}$ to each edge of G. To prove that $e \in E(G)$ appears in a perfect matching, it suffices to find $x \in P_{\text{perfect matching}}(G)$ such that $x(e) \neq 0$. Thus, our conclusion will follow provided that we can verify that $\overline{x} \in P_{\text{perfect matching}}(G)$. By Edmonds' Perfect Matching Polytope Theorem, we need to look at those three kinds of constraints in (6.4).

- $\overline{x}(e) = \frac{1}{3} > 0, \forall e \in E(G)$.
- The fact that G is 2-connected and cubic implies that G has no loops. Thus, for any $v \in V(G)$, it holds $\sum_{e \in \delta(v)} \overline{x}(e) = \frac{1}{3} + \frac{1}{3} + \frac{1}{3} = 1$ and $\sum_{e \in \gamma(v)} \overline{x}(e) = \sum_{e \in \emptyset} \overline{x}(e) = 0$.
- Let $U \subsetneq V(G)$ be any odd cardinality set. G is 2-connected means that $|\delta(U)| \geq 2$; while $|U|$ being odd gives that $|\delta(U)| = 3|U| - 2|\gamma(U)|$ is odd. So we get $|\delta(U)| \geq 3$ and hence $\sum_{e \in \delta(U)} \overline{x}(e) = \frac{|\delta(U)|}{3} \geq 1$.

□

Would you like to take a closer look at the above proof? Does it just establish Theorem 6.25 as stated?

Can one learn mathematics by reading it? I am inclined to say no. Reading has an edge over listening because reading is more active – but not much. Reading with pencil and paper on the side is very much better – it is a big step in the right direction. The very best way to read a book, however, with, to be sure, pencil and paper on the side, is to keep the pencil busy on the paper and throw the book away. – Paul Halmos

Everyone who is incapable of learning has taken to teaching. – Oscar Wilde (1854 – 1900)

Take A to be the matrix whose columns are the incidence vectors of all perfect matchings of G, say $M_1, \cdots, M_n$ in that order. If we get a little closer look at the proof of Theorem 6.25, we will see that $\overline{x} = Ax$ for some $x \in \mathbb{R}^n_+$. Therefore, we conclude from Exercise 4.8 that there is a $y \in \mathbb{Q}^n_+$ satisfying $\overline{x} = Ay$. Multiplying a suitable integer on both sides, we know that there is $z \in \mathbb{Z}^n$ such that Az takes a constant positive integer value. That is, if we take $z(1)$ copies of the matching M_1, $z(2)$ copies of the matching M_2, and so on, and make a big list

$$\underbrace{M_1, M_1, \cdots, M_1}_{z(1)}, \underbrace{M_2, M_2, \cdots, M_2}_{z(2)}, \cdots, \underbrace{M_n, M_n, \cdots, M_n}_{z(n)},$$

then each edge of G will appear in the same number of the matchings in the list. In other words, the edges of any 2-edge connected cubic graph can be `uniformly covered` by perfect matchings.

Fulkerson conjectured that six perfect matchings are always enough to constitute a uniform covering, namely a uniform matching covering of multiplicity two is always possible.

Conjecture 6.27. [48] *(The Fulkerson Matching Double Cover Conjecture [13]) For every 2-edge connected cubic graph G, one can always find a list of six perfect matchings in G such that every edge in G appears in exactly two of them.*

Delbert Ray Fulkerson Prize: The Fulkerson Prize for outstanding papers in the area of discrete mathematics is sponsored jointly by the Mathematical Programming Society (MPS) and the American Mathematical Society (AMS). Up to three awards of $1500 each are presented at each (triennial) International Symposium of the MPS. See a list of the former outstanding papers winning this prize at: `http://www.ams.org/prizes/fulkerson-prize.html`

[48]It is asserted in [33] that Berge also made this conjecture independently.

A conjecture weaker than Conjecture 6.27 is as follows:

Conjecture 6.28. *(Berge 1979) For every* 2*-edge connected cubic graph* G*, there are five perfect matchings whose union covers* $E(G)$*.*

The following equivalent form of 4CT, obtained by Tait in 1878, says that Conjecture 6.27 is true for planar graphs.

Theorem 6.29. *(Four Color Theorem) Each planar* 2*-edge connected cubic graph can be decomposed into three perfect matchings.*

Proof. I have discovered a truly remarkable demonstration of this proposition but this margin is too narrow to contain. Please look back to page 92 and then visit the relevant web sites, which has so large a room to contain many more materials. □

The `Petersen graph` has as vertices the 2-subsets of a 5-set[49] and a pair of 2-subsets form an edge if and only if they are disjoint (as sets). Petersen graph is especially important in graph theory. It is depicted on the covers of both the journals *Journal of Graph Theory* and *Discrete Mathematics*. There is even a book totally devoted to this graph. To deny a "proof" of Tait for the Four Color Conjecture, Petersen constructed this famous graph and showed that it is cubic and 2-edge connected but cannot be decomposed into three perfect matchings. Tutte's Four Flow Conjecture [50] says that the Petersen graph is essentially the 'only' such graph and the verification of this beautiful conjecture will be a great breakthrough in graph theory.

Exercise 6.30. Verify the result of Petersen mentioned above.

Despite the apparent simplicity of cubic graphs, many graph problems are surprisingly no easier to solve when restricted to them. Cubic graphs often seem to be the simplest class of graphs for which a problem remains as difficult as it does on a general graph. – R. Greenlaw, R. Petreschi, Cubic graphs, *ACM Computing Surveys* 27 (1995), 471–495.

[49]Note that $5 = 2 \times 2 + 1$! In the same manner, we can construct generally the so-called `odd graphs`. Have a guess of its definition.

[50]Go to GTM 173, written by Diestel, for a rigorous formulation of it.

This page is copied from

`http://www.emba.uvm.edu/~archdeac/problems/fulker.htm`:

A perfect matching or 1-factor in a graph is a collection of edges which together are incident with every vertex exactly once. A 1-factorization is a partition of the edges into 1-factors. If the graph is regular of degree d, then a 1-factorization is equivalent to a proper edge d-coloring. A cubic graph is one that is regular of degree three.

Not every cubic graph can be edge-partitioned into perfect matchings. The Petersen graph is a counterexample. The following was conjectured (I believe independently) by Berge and by Fulkerson. I've heard it called the Berge-Fulkerson conjecture, but more frequently just **Fulkerson's Conjecture** (who I believe was the first to publish it).

Conjecture: Every bridgeless cubic graph has a collection of six perfect matchings together contain every edge exactly twice.

The Petersen graph is not a counterexample (so the conjecture must be true!); the six 1-factors form such a double cover. Note that any 3-edge-colorable cubic graph (Class 1) trivially satisfies the conjecture.

The problem can also be phrased as a fractional edge coloring. That is, each edge receives two "half colors" so that no half color is repeated at a vertex. An extension of this would allow 1-factors of arbitrary rational weight. Another variation would be to allow adding or subtracting 1-factors. I believe that these variations may be known.

Mike Albertson asks if the conjecture is easier for toroidal graphs. He can show that there are 7 matchings together containing every edge exactly twice. He asks if there are 3 matchings that collectively cover all but two edges in a bridgeless toroidal cubic graph.

Archdeacon, Bonnington, and Siran (unpublished work) have examined many non-3-edge-colorable graphs of small order and have verified that they satisify the conjecture. In my opinion this is one of the most important open problems in the field.

A dual problem is called the Cycle double cover conjecture.

This page is copied from

http://www.emba.uvm.edu/~archdeac/problems/cyclecov.htm:

A cycle in a graph is a simple closed walk. The following **Double Cover Conjecture** is one of the most famous problems in graph theory. It is due independently to Szekeres [35] and Seymour [34].

Conjecture: Every bridgeless graph has a collection of cycles which together contain every edge exactly twice.

The conjecture is almost easy. Form G_2 from G by replacing each edge with two parallel edges. Then G_2 has every vertex of even degree. It follows easily from induction that G_2 has an edge partition into cycles. However, some of these cycles may be of length two and hence do not correspond to cycles in a double cover of G.

A stumbling block to inductive proofs has been found in many different contexts. Namely, suppose that each edge e is assigned a weight w(e) = 1 or 2 so that at each vertex the sum of the weights is even. Can we find a cycle cover so that each edge e is used w(e) times? No. A counterexample is formed from the Petersen graph (of course) by assigning weights 2 on a perfect matching and weights 1 on two disjoint 5-cycles.

There are several variations of a topological nature. For example there is the Circular Embedding Conjecture.

Conjecture: Every 2-connected graph has an embedding in some surface such that each face is bounded by a simple cycle.

The face boundaries form a cycle double cover. The two conjectures are equivalent for cubic graphs, but the second is stronger for noncubic graphs. An even stronger conjecture asserts that the faces of the circular embedding can be properly 5-colored. Likewise one could require the embedding to be in an orientable surface.

A stronger conjecture due to Goddyn [17] allows you to fix one cycle in the cover.

Conjecture: For every bridgeless G and every cycle C of G, there is a cycle double cover of G containing C.

Determining the edge colouring number of a graph is NP-complete, but with matching techniques one can determine a fractional version of it in polynomial time. Let G be a graph. The `fractional edge colouring number` $\chi_e^*(G)$ is defined as

$$\min\{\sum_{M\in\mathcal{M}}\lambda(M):\ \lambda\in\mathbb{R}_+^{\mathcal{M}},\sum_{M\in\mathcal{M}}\lambda(M)\chi^M=\chi^{E(G)}\},\qquad(6.7)$$

where $\mathcal{M}=\mathcal{M}(G)$ denotes the collection of all matchings in G.

Clearly, if we require that $\lambda\in\mathbb{Z}_+^{\mathcal{M}}$, then this would define the edge colouring number $\chi_e(G)$ of G. This leads to $\chi_e^*(G)\le\chi_e(G)$.

Exercise 6.31. Let G be the Petersen graph. Show that $\chi_e^*(G)=3$ and $\chi_e(G)=4$.

Exercise 6.32. Prove that $\chi_e^*(G)$ can be defined as

$$\max\{\sum_{e\in E(G)}x(e):\ x\in\mathbb{R}^{E(G)},\sum_{e\in M}x(e)\le 1,\forall\ \ M\in\mathcal{M}(G)\}.\qquad(6.8)$$

It is trivial to see that the edge colouring number of a subgraph is not greater than that of the original graph. This also holds for fractional edge colouring number.

Exercise 6.33. Let $G=(V,E)$ be a graph. Let $E'\subseteq E$ and $V'\subseteq V$. Let $G_1=(V',\gamma(V'))$ and $G_2=(V,E')$. Prove that $\chi_e^*(G)\ge\max(\chi_e^*(G_1),\chi_e^*(G_2))$. (hint: $\chi_e^*(G)\ge\chi_e^*(G_1)$ is direct from (6.7) while $\chi_e^*(G)\ge\chi_e^*(G_2)$ is due to (6.8).)

For any natural number k, let G_k be the graph obtained from G by replacing each edge by k parallel edges of the same endpoints. The next exercise provides an alternate means to prove that the fractional colouring number is monotone decreasing under the taking subgraph operation.

Exercise 6.34. $\chi_e^*(G)=\min_k\frac{\chi_e(G_k)}{k}$. (hint: All vertices of the polytope defined by (6.7) are rational.)

We are now at a place to give another min-max relation.

Theorem 6.35. *The fractional edge colouring number $\chi_e^*(G)$ of a loopless graph G satisfies*

$$\chi_e^*(G) = \max\{\Delta(G), \max_{U\subseteq V(G), |U|\geq 3} \frac{|\gamma(U)|}{\lfloor\frac{|U|}{2}\rfloor}\}. \tag{6.9}$$

Proof. Let μ denote the maximum in Eq. (6.9). We may assume $\mu > 0$.

The inequality $\chi_e^*(G) \geq \mu$ can be seen as follows. Assume that λ attains the minimum in (6.7) and $\deg_G(v) = \Delta(G)$. We have

- $\chi_e^*(G) = \sum_M \lambda(M) \geq \sum_M \lambda(M)|M \cap \delta(v)| = \sum_{e\in\delta(v)} \sum_{e\in M} \lambda(M) = \sum_{e\in\delta(v)} 1 = \Delta(G)$;
- Moreover, for each $U \subseteq V(G)$ with $|U| \geq 3$, $\chi_e^*(G) = \sum_M \lambda(M) \geq \sum_M \lambda(M) \frac{|M\cap\gamma(U)|}{\lfloor\frac{|U|}{2}\rfloor} = \frac{1}{\lfloor\frac{|U|}{2}\rfloor} \sum_{e\in\gamma(U)} \sum_{e\in M} \lambda(M) = \frac{|\gamma(U)|}{\lfloor\frac{|U|}{2}\rfloor}$.

To show the converse inequality $\chi_e^*(G) \leq \mu$, we need to find a feasible solution λ to (6.7) with objective value $\sum_{M\in\mathcal{M}} \lambda(M) = \mu$. We start from the all-$\frac{1}{\mu}$ vector $x \in \mathbb{R}^{E(G)}$. By Theorem 6.21, the following facts say that $x \in P_{\text{matching}}(G)$:

- $\forall e \in E(G)$, $x(e) = \frac{1}{\mu} > 0$;
- $\mu \geq \Delta(G)$ implies $\sum_{e\in\delta(v)} x(e) \leq 1$.
- Let $U \in \mathcal{P}_{\text{odd}}(V(G))$. If U is a singleton set, the assumption that G is loopless gives that $\sum_{e\in\gamma(U)} x(e) \leq 0 = \lfloor\frac{1}{2}|U|\rfloor$. Otherwise, we have $|U| \geq 3$. It then follows from $\mu \geq \max_{U\subseteq V(G),|U|\geq 3} \frac{|\gamma(U)|}{\lfloor\frac{|U|}{2}\rfloor}$ that $\sum_{e\in\gamma(U)} x(e) = \frac{|\gamma(U)|}{\mu} \leq \lfloor\frac{1}{2}|U|\rfloor$.

Henceforth, there is $\lambda \in \mathbb{R}_+^{\mathcal{M}}$ such that $x = \sum_{M\in\mathcal{M}} \frac{\lambda(M)}{\mu}\chi^M$ with $\sum_{M\in\mathcal{M}} \frac{\lambda(M)}{\mu} = 1$. This means that $\sum_M \lambda(M)\chi^M = \mu x = \chi^{E(G)}$ and $\sum_M \lambda(M) = \mu$, completing the proof. $\square$

Exercise 6.36. For a k-regular loopless graph G, $\chi_e^*(G) = k$ if and only if $|\delta(U)| \geq k$ for each $U \in \mathcal{P}_{\text{odd}}(V(G))$.

http://chern.nankai.edu.cn/

I think they don't need too many mathematicians. China is a large country, so naturally it has a lot of talent, particularly in the smaller places. For instance, there's the International Olympiad for the high school students, and China generally does very well. In order to achieve well in competitions like this, the students need training, and as a result other topics could be ignored. Now the parents in China want their children to know more English, go into business, and make more money. And these exams don't give money. One year I think they just did less of this training, and China immediately dropped. What do you do for a country with 1.2 billion people? It means that the standard of living cannot be very high, if you have any social justice. – **Shiing-Shen Chern** (26 Oct. 1911 – 3 Dec. 2004)

But the subject was difficult, so after a number of times, people didn't come anymore. I think I was essentially the only one who stayed till the end. I think I stayed till the end because I followed the subject. Not only that, I was writing a thesis applying the methods to another problem, so the seminar was of great importance to me. – **Shiing-Shen Chern** (26 Oct. 1911 – 3 Dec. 2004)

We need to cultivate feelings for it, at the beginning, you would feel dull, but I feel it is beautiful because it is full of logical things, it is a thing that I go after for a whole life. – **Shiing-Shen Chern** (26 Oct. 1911 – 3 Dec. 2004)

Every good mathematician has to be a problem solver. If you are not a problem solver, you only have vague ideas, how can you make a good contribution? You solve some problems, you use some concepts, and the merit of mathematical contributions, you probably have to wait. You can only see it in the future. – **Shiing-Shen Chern** (26 Oct. 1911 – 3 Dec. 2004)

Before coming back to the proof of Theorem 6.19, we introduce the `contraction operation` for graphs.[51] Let G be a graph and $U \subseteq V(G)$. We write $\overline{U}$ for $V(G) \setminus U$. By contracting U into a point v_U we get the graph $H = G/U$ with vertex set $\overline{U} \cup \{v_U\}$ and edge set $\gamma_G(\overline{U}) \cup \delta_G(U)$, where the incidence relation is given by

$$\partial_H(e) = \begin{cases} \partial_G(e), & \text{if } e \in \gamma_G(\overline{U}); \\ (\partial_G(e) \cap \overline{U}) \cup \{v_U\}, & \text{if } e \in \delta_G(U). \end{cases}$$

A graph H is a `minor` of G, written $G \succ H$, if it is a subgraph of a graph obtained from G by a sequence of contraction operations.[52]

Recall that, as mentioned on page 75, Robertson and Seymour prove the Graph Minor Theorem as the culminating result of a long series of articles which take over 500 pages. This result is said to "dwarf any other result in graph theory and may doubtless be counted among the deepest theorems that mathematics has to offer"[53].

The learning process is not purely cognitive. It is simplistic to suppose that people remember what they are told, and understand things that are explained to them clearly. Much more commonly, people remember what interests them, and understand the things that they enjoy understanding. ... If in considering complex processes we limit our observations and judgments to things that we can analyze and measure exactly, this leads not to scientific exactitude but to tunnel vision. ... – Edwin E. Moise[54]

---

[51] Contraction and deletion are two fundamental operations which correspond to two fundamental containment relations between graphs.

[52] $G \succ H$ if and only if H is obtained from G by a sequence of deletion and contraction operations.

[53] Reinhard Diestel, Graph Theory, Springer-Verlag, 2000.

[54] In 1952, Moise proved that every compact 3-dimensional topological manifold can be triangulated and that any two such triangulations are combinatorially equivalent. This proved the main conjecture of Poincaré in dimension 3. Milnor gave counterexamples to Poincaré's conjecture when the dimension is greater than 5.

We are ready to present a direct proof of Theorem 6.19. Later, we will show that (6.4) is TDI and hence produce an alternate proof of it using what we learned from Chapter 4. We write $\mathcal{PM}(G)$ for the set of all perfect matchings of a graph G.

Proof. (of Theorem 6.19)[55] Let $Q(G)$ be the polytope given by (6.4).

Because the incidence vector of each perfect matching of G satisfies the constraints in (6.4), we are able to announce that $P_{\text{perfect matching}}(G) \subseteq Q(G)$. If the converse inclusion does not hold, we can take a counterexample G with $|V(G)| + |E(G)|$ as small as possible. This means that there is a vertex x of the polytope $Q(G)$ satisfying $x \notin P_{\text{perfect matching}}(G)$ and no such graph can have a smaller number of edges and vertices. We will make a series of observations first and then try to deduce a contradiction from them and hence end the proof.

Our observations are about the existence of a special vertex subset W. Having such a W, we will go to the minors G/W and $G/\overline{W}$, which are both smaller than G and thus cannot be counterexamples, and correspondingly we will project x to be an $x' \in Q(G/W) = P_{\text{perfect matching}}(G/W)$ and an $x'' \in Q(G/\overline{W}) = P_{\text{perfect matching}}(G/\overline{W})$. Combining the expressions of x' and x'' as convex combinations of vertices in corresponding matching polytopes, we can deduce that $x \in P_{\text{perfect matching}}(G)$, yielding the required contradiction.

[55] [32, Theorem 25.1]

1. $|V(G)|$ is even.[56]

2. G has no isolated vertices, namely a vertex connected with no other points. [57]

3. G is not 2-regular.[58]

4. For each $e \in E(G)$, $0 < x(e) < 1$. [59]

5. G is loopless.[60]

6. It holds $\deg_G(v) \geq 2$ for each $v \in V(G)$.[61]

7. $|E(G)| > |V(G)|$.[62]

8. There is $W \in \mathcal{P}_{\text{odd}}(V(G))$ with

$$\sum_{e \in \delta(W)} x(e) = 1 \tag{6.10}$$

such that $3 \leq \min(|W|, |\overline{W}|)$, where $\overline{W} = V(G) \setminus W$. [63].

---

[56]Otherwise, take $U = V$ in (6.4)(iii), contradicting the fact $x \in Q(G) \neq \emptyset$.

[57]x cannot satisfy (6.4)(ii) at an isolated vertex.

[58]Otherwise, it follows from item 2 that G is bipartite and hence Exercise 6.2 guarantees $Q(G) = P_{\text{perfect matching}}(G)$.

[59]If $x(e) = 0$, delete e to get a smaller counterexample; if $x(e) = 1$, then (6.4)(ii) shows that e has to appear in every matching of G and we can delete e and its end vertices to get a smaller counterexample.

[60]Let e be a loop. By (6.4) (i) and (ii), we have $x(e) = 0$, in contradiction with item 4.

[61]It follows from items 2 and 4 and (6.4)(ii).

[62]Follows from items 3 and 6.

[63]As x is a vertex of $Q(G)$, a polytope in $\mathbb{R}^{E(G)}$, Theorem 1.32 says that there are $|E(G)|$ linearly independent constraints in (6.4) satisfied by x with equality. But by items 4, they do not include any one from (6.4)(i). By item 5, they do not include any one from the latter part of (6.4)(ii). There are at most $|V(G)|$ linearly independent constraints in the former part of (6.4)(ii) and hence item 7 shows that x must fulfil one equality from (6.4)(iii) that is linearly independent from those equalities in (6.4)(ii), say $\sum_{e \in \delta(W)} x(e) = 1$, $W \subseteq V(G), |W| \geq 3, |W|$ odd. By item 1, $\overline{W}$ has size at least three or is of size exactly one. If it holds the latter case, since $\delta(W) = \delta(\overline{W})$, this equality corresponding to W is nothing but $\sum_{e \in \delta(\overline{W})} x(e) = 1$, which appears in (6.4)(ii) already. This is impossible in view of our assumption.

Note that each perfect matching of G/W contains exactly one element from $\delta(W)$. Therefore, we can write $\mathcal{PM}(G/W) = \cup_{e\in\delta_G(W)}\mathcal{PM}(G/W)_e$, where $\mathcal{PM}(G/W)_e = \{M \in \mathcal{M}(G/W) : M \cap \delta_G(W) = \{e\}\}$. Similarly, we have $\mathcal{PM}(G/\overline{W}) = \cup_{e\in\delta_G(W)}\mathcal{PM}(G/\overline{W})_e$, where $\mathcal{PM}(G/\overline{W})_e = \{M \in \mathcal{PM}(G/\overline{W}) : M \cap \delta_G(W) = \{e\}\}$. To return to the matching polytope of G, a key fact is that for each $e \in \delta(W)$, $M_1 \cup M_2$ becomes a perfect matching of G for each $M_1 \in \mathcal{PM}(G/W)_e$ and $M_2 \in \mathcal{PM}(G/\overline{W})_e$. We will show that x lies in the convex hull of the vertices of $P_{\text{perfect matching}}(G)$ corresponding to the matchings arising in this way, which will then prove the theorem.

Observe that both $E(G/W)$ and $E(G/\overline{W})$ are subsets of $E(G)$. Thus, for each $y \in \mathbb{R}^{E(G)}$ we can define $y_1 \in \mathbb{R}^{E(G/W)}$ by requiring

$$y_1(e) = y(e) \tag{6.11}$$

for $e \in \gamma_G(\overline{W}) \cup \delta_G(W) = E(G/W)$ and define $y_2 \in \mathbb{R}^{E(G/\overline{W})}$ by putting

$$y_2(e) = y(e) \tag{6.12}$$

for $e \in \gamma_G(W) \cup \delta_G(\overline{W}) = E(G/\overline{W})$.

Since x is a feasible solution to (6.4) and the odd set W is chosen to satisfy (6.10), it follows that $x_1 \in Q(G/W)$ and $x_2 \in Q(G/\overline{W})$.[64] Moreover, since $3 \le \min(|W|, |\overline{W}|)$, we know that $|V(G)| + |E(G)| - 2 \ge \max(|V(G/W)| + |E(G/W)|, |V(G/\overline{W})| + |E(G/\overline{W})|)$. Henceforth, as G is the minimum counterexample to the inclusion relation $P_{\text{perfect matching}}(G) \supseteq Q(G)$, we conclude now that $x_1 \in P_{\text{perfect matching}}(G/W)$ and $x_2 \in P_{\text{perfect matching}}(G/\overline{W})$.

[64]You should check the following:

$$\begin{cases} \text{(i) } x_1 \in \mathbb{R}_+^{E(G/W)}; \\ \text{(ii) } \sum_{e\in\delta_{G/W}(v)} x_1(e) = 1, \sum_{e\in\gamma_{G/W}(v)} x_1(e) = 0, \forall v \in V(G/W); \\ \text{(iii) } \sum_{e\in\delta_{G/W}(U)} x_1(e) \ge 1, \forall U \in \mathcal{P}_{\text{odd}}(V(G/W)), |U| \ge 3. \end{cases}$$

$$\begin{cases} \text{(i) } x_2 \in \mathbb{R}_+^{E(G/\overline{W})}; \\ \text{(ii) } \sum_{e\in\delta_{G/\overline{W}}(v)} x_2(e) = 1, \sum_{e\in\gamma_{G/\overline{W}}(v)} x_2(e) = 0, \forall v \in V(G/\overline{W}); \\ \text{(iii) } \sum_{e\in\delta_{G/\overline{W}}(U)} x_2(e) \ge 1, \forall U \in \mathcal{P}_{\text{odd}}(V(G/\overline{W})), |U| \ge 3. \end{cases}$$

The incidence vectors appeared below are all over the ground set $E(G)$. Also remember our definition and notation for the projection of a vector in $\mathbb{R}^{E(G)}$ to $\mathbb{R}^{E(G/W)}$ and to $\mathbb{R}^{E(G/\overline{W})}$, respectively.

Let us suppose that

$$x_1 = \sum_{M_1 \in \mathcal{PM}(G/W)} \alpha_{M_1} \chi_1^{M_1} \tag{6.13}$$

and

$$x_2 = \sum_{M_2 \in \mathcal{PM}(G/\overline{W})} \beta_{M_2} \chi_2^{M_2}, \tag{6.14}$$

where

$$\alpha_{M_1}, \beta_{M_2} \in \mathbb{R}_+ \tag{6.15}$$

and

$$\sum_{M_1 \in \mathcal{PM}(G/W)} \alpha_{M_1} = \sum_{M_2 \in \mathcal{PM}(G/\overline{W})} \beta_{M_2} = 1.\text{[65]} \tag{6.16}$$

Let $e \in \delta_G(W)$. By virtue of the definition of x_1 and x_2, it is immediate from (6.13) and (6.14) that

$$x(e) = \sum_{M_1 \in \mathcal{PM}(G/W)_e} \alpha_{M_1} = \sum_{M_2 \in \mathcal{PM}(G/\overline{W})_e} \beta_{M_2}. \tag{6.17}$$

This then gives

$$x(e) = \frac{x(e)^2}{x(e)} = \sum_{\substack{M_1 \in \mathcal{PM}(G/W)_e \\ M_2 \in \mathcal{PM}(G/\overline{W})_e}} \frac{\alpha_{M_1} \beta_{M_2}}{x(e)}. \tag{6.18}$$

We aim to prove that

$$x = \sum_{e \in \delta_G(W)} \sum_{\substack{M_1 \in \mathcal{PM}(G/W)_e \\ M_2 \in \mathcal{PM}(G/\overline{W})_e}} \frac{\alpha_{M_1} \beta_{M_2}}{x(e)} \chi^{M_1 \cup M_2}. \tag{6.19}$$

Note that (6.10) together with (6.15) and (6.18) says that the RHS of (6.19) is a convex combination of the incidence vectors of the perfect matchings of G and thus we will be done after it is achieved.

[65] Is this relation used in the following arguments? Nowhere? Anything wrong? If you find it a puzzle, try to find a hint at a footnote somewhere else, but surely still in this notes.

Denote the RHS of (6.19) by z.

For $e \in \delta_G(W)$,

$$\begin{aligned} z(e) = & \sum_{f\in\delta_G(W)} \sum_{\substack{M_1\in\mathcal{PM}(G/W)_f \\ M_2\in\mathcal{PM}(G/\overline{W})_f}} \frac{\alpha_{M_1}\beta_{M_2}}{x(f)} \chi^{M_1\cup M_2}(e) \\ = & \sum_{\substack{M_1\in\mathcal{PM}(G/W)_e \\ M_2\in\mathcal{PM}(G/\overline{W})_e}} \frac{\alpha_{M_1}\beta_{M_2}}{x(e)} \\ = & \; x(e). \quad \text{by } (6.18) \end{aligned} \tag{6.20}$$

For $M_1 \in \mathcal{PM}(G/W)$, we write $f(M_1)$ for the unique element in $M_1 \cap \delta_G(W)$. With this convention, for $e \in \gamma_G(\overline{W})$, we can calculate

$$\begin{aligned} z(e) = & \sum_{f\in\delta_G(W)} \sum_{\substack{M_1\in\mathcal{PM}(G/W)_f \\ M_2\in\mathcal{PM}(G/\overline{W})_f}} \frac{\alpha_{M_1}\beta_{M_2}}{x(f)} \chi^{M_1\cup M_2}(e) \\ = & \sum_{e\in M_1\in\mathcal{PM}(G/W)} \alpha_{M_1} \sum_{\substack{M_2\in\mathcal{PM}(G/\overline{W}) \\ f(M_1)\in M_2}} \frac{\beta(M_2)}{x(f(M_1))} \\ = & \sum_{e\in M_1\in\mathcal{PM}(G/W)} \alpha_{M_1} \frac{x_2(f(M_1))}{x(f(M_1))} \quad \text{by } (6.14) \\ = & \sum_{e\in M_1\in\mathcal{PM}(G/W)} \alpha_{M_1} \quad \text{by } (6.12) \\ = & \sum_{M_1\in\mathcal{PM}(G/W)} \alpha_{M_1} \chi_1^{M_1}(e) \\ = & \; x_1(e) \quad \text{by } (6.13) \\ = & \; x(e). \quad \text{by } (6.11) \end{aligned} \tag{6.21}$$

Finally, using $\overline{W}$ in place of W in the above argument yields that

$$z(e) = x(e) \tag{6.22}$$

for $e \in \gamma_G(W)$.

Upon combining (6.20), (6.21) and (6.22), (6.19) is obtained, as expected. □

Exercise 6.37. Derive from Edmonds' matching polytope theorem the linear inequalities determining the convex hull of all symmetric permutation matrices.

Exercise 6.38. Let G be a graph and let $w \in \mathbb{Q}_+^{E(G)}$ be a weighting function. Then the maximum weight of a matching is equal to the minimum value of $\sum_{v\in V(G)} y(v) + \sum_{U\subseteq V(G)} z(U)\lfloor\frac{|U|}{2}\rfloor$, where $y \in \mathbb{Q}_+^{V(G)}$ and $z \in \mathbb{Q}_+^{\mathcal{P}_{\mathrm{odd}(V(G))}}$ satisfy $\sum_{v\in\partial(e)} y(v) + \sum_{U\in\mathcal{P}_{\mathrm{odd}(V(G))}, e\in U} z(U) \geq w(e)$ for each $e \in E(G)$.

Cutting Planes[66]

Let us suppose that we want to determine the convex hull of a set S of integral vectors in $\mathcal{R}^n$. It is often that we can easily find a system of linear inequalities whose integral solutions are exactly the elements of S. For example, when $S = \{\chi^M : M \in \mathcal{P}M(G)\}$ for a graph G, we can take the system as that defining $P^*_{\text{perfect matching}}(G)$, namely (6.4)(i) and (ii).[67] In such a situation, we face the problem of telling integer hull of a given polytope.

Let the given system be $Ax \leq b$. Let $c^\top x \leq d$ be an inequality which is satisfied by all elements of $P(A, b)$[68]. Assume, moreover, that c is integral. Then the inequality $c^\top x \leq \lfloor d \rfloor$ is called a `cut` derived from the original system. Clearly, this cut only cut off some nonintegral points from the given polytope. Also note that the cut of a polytope is irrelevant to the inequality system description of it.

For any polytope P, define $T(P) = \{x : x \text{ satisfies all cuts of } P\}$, which is called the `Gomory-Chvátal truncation` of P. Clearly, the integer hull of P is contained in $T(P)$. It is known that $T(P)$ remains to be a polytope and thus we can still construct its Gomory-Chvátal truncation. A basic result is that we only need to do a finite number of truncation operations to reach the integral hull. The minimum number of rounds required is referred to as the `Chvátal rank` of the polytope P. In view of Exercise 2.1, this parameter measures how far away the polytope is from an integer polyhedron.

---

[66] [26, Box 7D]

[67] Look through the contents covered in this course and collect all such examples!

[68] By Farkas' Lemma, this happens if and only if there is vector $y \geq 0$ such that $c^\top = y^\top A$ and $d = y^\top b$.

The difficulty of applying the above result to deduce an inequality description of the integer hull of a polytope is that we may face with an enormous number of cuts of the original polytopes and its higher order Gomory-Chvátal truncations. But the case with the perfect matching polytope is very interesting. We start from the easy-to-get description of the fractional perfect matching polytope of a graph G, (6.4)(i) and (ii) and generate those cuts in (6.5)(iii). Then Edmonds' Perfect Matching Polytope Theorem tells us that we have been finished. That is, the Chvátal rank of the fractional perfect matching polytope is only one.

Note that Vašek Chvátal was teaching the course "Cutting-plane method and traveling salesman problem" and here is the course website: `http://www.cs.rutgers.edu/~chvatal/611/intro.html`

...it may be interesting to note that each class of integer linear programming problems with bounded rank admits a good characterization.
–Vašek Chvátal [4]

I will be out of Shanghai on Dec. 21 and so the class on that afternoon will be cancelled.

Theorem 6.39. *(Cunningham-Marsh formula) In Exercise 6.38, if* w *comes from* $\mathbb{Z}_+^{E(G)}$*, we can take* y *and* z *integral.*

[69]

Mathematics – this may surprise you or shock you some – is never deductive in its creation. The mathematician at work makes vague guesses, visualizes broad generalizations, and jumps to unwarranted conclusions. He arranges and rearranges his ideas, and he becomes convinced of their truth long before he can write down a logical proof. The conviction is not likely to come early – it usually comes after many attempts, many failures, many discouragements, many false starts. It often happens that months of work result in the proof that the method of attack they were based on cannot possibly work, and the process of guessing, visualizing and conclusion- jumping begins again. . . . The deductive stage, writing the result down, and writing down its Rogers proof are relatively trivial once the real insight arrives; it is more like the draftsman's work, not the architect's. – Paul Halmos

[69] About the puzzle mentioned in the footnote on page 177: Yes, we do not need (6.16). Indeed, we have used (6.17) instead of which (6.16) is a consequence(why?). In this sense, we make use of (6.16) indirectly.

Cunningham-Marsh formula

Edmonds' Matching Polytope Theorem

Theorem 6.40. *(Tutte-Berge formula) For any graph G and integer k, exactly one of the following holds.*

1. *G has a matching M where at most k vertices of G are M-unsaturated.*
2. *There exists $S \subseteq V(G)$ such that odd$(G - S) - |S| > k$. (Here S is called a Tutte barrier.)*

In symbols,

Tutte's 1-factor theorem

Hall's Marriage Theorem → Gallai-Edmonds Structure Theorem [24]

...the idea seemed so obvious to me and so elegant that I fell deeply in love with it. And, like falling in love with a woman, it is only possible if you do not know much about her, so you cannot see her faults. The faults will become apparent later, but after the love is strong enough to hold you to her. – Richard Feynman

7 Miscellaneous

Uncrossing technique[11]

Let $y : 2^V \to \mathbb{Q}_+$ be a nonnegative set function. By the `uncrossing operation` we mean the following modification of y: given two crossing sets X_1 and X_2 with $y(X_1), y(X_2) > 0$, decrease $y(X_1)$ and $y(X_2)$ by $\min\{y(X_1), y(X_2)\}$ and increase $y(X_1 \cap X_2)$ and $y(X_1 \cup X_2)$ by the same amount. If $y(X)$ is defined as the multiplicity of X in a family $\mathcal{F}$, then we speak of `uncrossing` $\mathcal{F}$.

Lemma 7.1. *[11] Let $x_1, x_2, \cdots, x_n$ be nonnegative rational numbers. Suppose that we apply repeatedly the following operation: for some indices $i < j < k < \ell$ where x_j and x_k are positive, decrease x_j and x_k by $\min\{x_j, x_k\}$, and increase x_i and x_ℓ by $\min\{x_j, x_k\}$. Then this operation can be repeated only a finite number of times.*

The usefulness of the uncrossing technique in a combinatorial optimization follows from the nice properties of cross-free families, that are often linked to dual integrality properties. Please review the proof of Theorem 5.3.

Let $(X, \leq)$ be a partially ordered set. If $a \leq b$ and $a \neq b$, then we write $a < b$. We say that two elements $a, b \in X$ are `intersecting` provided a and b are non-comparable and there exists $c \in X$ such that $c < a$, $c < b$. A subset Y of X is `laminar` if Y contains no two intersecting elements.

A 2-matching is a collection of edges such that each vertex is incident to at most two of them. A 2-cover is a collection of vertices such that each vertex belongs to at most two of them.

Theorem 7.2. *(J. Edmonds) The maximum number of edges in a* 2*-matching equals the minimum number of vertices in a* 2*-cover.*

Theorem 7.3. *(J. Edmonds) Let* G *be a* k*-regular* k*-edge-connected graph with an even number of vertices. Then there exists a number* t *such that if we replace every edge by* t *parallel edges the resulting graph has edge colouring number* kt*.*

Theorem 7.4. *(Lucchesi and Younger, 1973) Let* G *be a digraph. Then the maximum number of arc-disjoint dicuts in* G *equals the minimum number of arcs covering all dicuts.*

[14]: Two problems that play a fundamental role in combinatorial optimization are the maximum cardinality matching problem and the matroid intersection problem. This paper surveys matrix rank formulations of these and related problems. We begin by considering the size of a maximum cardinality matching in a bipartite graph, which is a special case of both of the aforementioned problems.

We end this course with a story of **coins and cones**, based on [25]. You must have heard of many sorts of coin-weighing problems. By now, we are concerned with the following one:

A set of coins is said to be of `generic weights` provided for any two subset A and B of it which are of equal weights, for any weight w, the number of coins of weight w in A is the same as that in B. Suppose that there are a set of m coins of generic weights and you can use n weighings to tell if they are of the same weight. How large can m be? Namely, what is the maximum size $f(n)$ of a set of coins of generic weights for which the weight-uniformness can be decided by n weighings?

Exercise 7.5. $f(n) \geq 2^n$.

It was once conjectured that $f(n) = 2^n$ and it has been refuted by an algorithm presented in [25]. Further results can be found in [1].

Our last theorem to prove is as follows.

Theorem 7.6. *[25]* $f(n) \leq \frac{(3^n-1)(n+1)^{\frac{n+1}{2}}}{2}$.

In each weighing, we put a set A of coins on one pan and another set B on the other pan. Clearly, A and B are disjoint. Since we are considering a set of coins of generic weights, to deduce some new information about the weight distribution, we can further assume that $|A| = |B|$[70]. This says that each weighing $\mathcal{W}_i$ corresponds to a pair (A_i, B_i) of disjoint subsets of equal size. A sequence of such weighings can be used to decided the weight-uniformness if and only if when we always have a pan balance in each weighing we can assert that the coins are of uniform weight. The latter is equivalent to the fact that for each proper nonempty subset C of the set of coins, we should have $|C \cap A_i| \neq |C \cap B_i|$ for some weighing $\mathcal{W}_i$.[71]

We are ready to transform our coin-weighing problem into the following more mathematical one:

Consider an $n \times m$ $(0, \pm 1)$ matrix A of zero row sum vector[72]. We say that A is good if for each proper subset $C \neq \emptyset$ of $[m] = \{1, \cdots, m\}$ the row sum vector of $A(\cdot, C)$ is not zero. Let $f(n)$ be the maximum m such that there is a good $n \times m$ matrix A. Estimate $f(n)$.

Let $W = \{0, \pm 1\}^{[n]}$. For each $w \in W$, let λ_w be the number of columns of A which are equal to w. We can check that the matrix A is good if and only if $\sum_{w \in W} \lambda_w w = 0$ while there is no $\lambda \gneqq \mu \gneqq 0$ such that $\sum_{w \in W} \mu_w w = 0$. In addition, the parameter m is nothing but $\sum_{w \in W} \lambda_w$.

Horizontally a peak and vertically a range. (Heng kan cheng ling ce cheng feng) – SU, Shi (1037–1101)

---

[70]This reasoning seems not complete yet. Anyway, we can just use it as an assumption.

[71]Consider the case that the coins in C consist of all coins of a certain weight.

[72]You may imagine that a position occupied by a 1 in the ith row corresponds to an element in A_i and a -1 element of B_i. That is, the ith row is the difference between the characteristic vectors of A_i and of B_i.

Our analysis above leads to the third definition of $f(n)$. You will see that, starting from coins we have finally arrived at a cone.

Consider the polyhedral cone $K \in \mathbb{R}^W$ consisting of all vectors $\lambda \geq 0$ satisfying $\sum_{w \in W} \lambda_w w = 0$. An `indecomposable` element of K is an integral vector of K which can not be broken up into the sum of two nonzero integral vectors of K. The `weight` of a vector x is the sum of its components and denoted $w(x)$. $f(n)$ is the maximum weight of an indecomposable element of the cone K.

This formulation of $f(n)$ tells us that we should try to understand the structure of the cone K.

Theorem 1.53 says that the polyhedral cone K is the set of the conical combinations of a finite number of extreme rays. Suppose $\{tx : t \in \mathbb{R}_+\}$ is an extreme ray. Theorem 1.27 says that we can assume x to be a rational vector. Thus there is a minimum weight integral vector on this extreme ray, which we call a `minimal generator` of K. It is clear that each minimal generator of K is indecomposable. So, it is natural that we will try to estimate the maximum weight of the minimum generators of K.

We use $\overline{x}$ to represent the vector obtained from x by deleting all zeros in x. From Theorems 1.26 and 1.27, we can further deduce

Lemma 7.7. *x is an extreme ray of K if and only if*

- *$\overline{x} > 0$ and $M_x\overline{x} = 0$, where M_x has as columns those $w \in W$ for which $x_w > 0$.*
- *$\operatorname{rank}(M_x) = \ell$, where $\ell + 1$ is the length of $\overline{x}$.*

Let M be the matrix whose columns are all vectors from W. Let L be a full ranked $\ell \times (\ell + 1)$ submatrix of M. Note that $Ly = 0$ has a nontrivial integral solution $y = \left(\det L_1, -\det L_2, \cdots, (-1)^\ell \det L_{\ell+1}\right)^\top$, where L_i is the ℓ by ℓ submatrix obtained from L by deleting the ith column [73]. Let

$$g(L) = \frac{\sum_{i=1}^{\ell+1} |\det L_i|}{\gcd(|\det L_1|, \cdots, |\det L_{\ell+1}|)}. \tag{7.1}$$

Denote $\gamma(n) = \max g(L)$ where L runs over the set of all $\ell \times (\ell+1)$ full ranked submatrices of M, $\ell \leq n$, and $Ly = 0$ has a positive solution. It follows from Lemma 7.7 that

Lemma 7.8. *$\gamma(n)$ is the maximum weight of a minimal generator of K and hence $\gamma(n) \leq f(n)$.*

[73] By adding a copy of some row of L to L we get a singular matrix!

Lemma 7.9. $f(n) \le \frac{3^n-1}{2}\gamma(n)$.

Proof. Observe that for $w \in W$, both w and $-w$ appear as columns of M. Thus, for any indecomposable element $\lambda \neq 0$ of K, either $\lambda_w = 0$ or $\lambda_{-w} = 0$[74]. This says that λ lies in a subspace of dimension at most $\frac{3^n-1}{2}$, the intersection $\mathcal{H}$ of all the hyperplanes $H_i = \{y \in \mathbb{R}^W : \ y_i = 0\}$ for zero coordinates i of λ. Clearly, λ is still an indecomposable element of $K \cap \mathcal{H}$.

By Theorem 1.53, λ is a positive combination of some minimal generators of $K \cap \mathcal{H}$. Moreover, Theorem 1.36 tells us that we need at most $\frac{3^n-1}{2}$ terms in the expression. That is to say, we have $\lambda = \sum_{i=1}^{\frac{3^n-1}{2}} \alpha_i x_i$, where x_i are minimal generators and α_i nonnegative numbers. Since λ is indecomposable, we know that $\lambda - x_i \notin K$ and hence $\alpha_i < 1$ for all i. This gives $w(\lambda) = \sum_{i=1}^{\frac{3^n-1}{2}} \alpha_i w(x_i) < \sum_{i=1}^{\frac{3^n-1}{2}} w(x_i) \le \frac{3^n-1}{2}\gamma(n)$, as required. □

[74]Note that $\lambda = \mu + (\lambda - \mu)$ where $\mu = \min(\lambda_w, \lambda_{-w})(\chi_w + \chi_{-w})$.

Proof. (of Theorem 7.6) In view of Lemma 7.9, we only need to prove $\gamma(n) \leq (n+1)^{\frac{n+1}{2}}$.

Looking back to Eq. (7.1), we find that the quantity $\sum_{i=1}^{\ell+1} |\det L_i|$ can be viewed as the determinant of an $(\ell+1) \times (\ell+1)$ $(0, \pm 1)$ matrix, say L', which is an extension of L by an appropriate $(1, -1)$ row. But $|\det L'| = \det L'$ is the volume of the parallelepiped spanned by its column vectors, which is not larger than the product of their norms[75]. As L' is a $(0, \pm 1)$ matrix, the norm of each of its column is at most $(\ell+1)^{\frac{1}{2}}$ and so $\det L' \leq (\ell+1)^{\frac{\ell+1}{2}}$. The assertion is now immediate from Eq. (7.1). □

An $n \times n$ $(1, -1)$ matrix H is an `Hadamard matrix` of order n if the inner product of any two distinct rows is always 0. Both an Hadamard matrix of order n and the Vandermonde matrix of the n roots of unity have the maximum possible determinant (in absolute value) of any complex matrix with entries in the closed unit disk. Hadamard conjectured that Hadamard matrices of order n exist precisely for $n = 1, 2$ and any multiple of 4. It is easy to show that for any other n, no Hadamard matrix of order n can exist.

To parents who despair because their children are unable to master the first problems in arithmetic I can dedicate my examples. For, in arithmetic, until the seventh grade I was last or nearly last. – Jacques Hadamard (1865 - 1963)

[75] This is known as the Hadamard's inequality.

To say more about $f(n)$, we turn to the concept of Hilbert basis.

Given a polyhedral cone K, let Z denote the set of all integral vectors in it. A finite set of vectors $\{a_1, \cdots, a_t\}$ from Z is called an `integral Hilbert basis` if each integral vector $b \in K$ is a nonnegative integral combination of $a_1, \cdots, a_t$.

Theorem 7.10. *[30, Theorem 16.4] Each rational polyhedral cone P has an integral Hilbert basis. If P is pointed, there is a unique minimal Hilbert basis under inclusion relation.*

Since the cone K under discussion is pointed, Theorem 7.10 guarantees that it has a unique integral Hilbert basis, say H. It is easy to see that H is just the set of all indecomposable elements of K! Correspondingly, we get the fourth definition of $f(n)$: it is the maximum weight of the elements in the unique integral Hilbert basis H of K.

`http://www.math.ucdavis.edu/~maya/papers/`

From the above address, you can find the following interesting papers of Maya Ahmed.

Polyhedral cones of magic cubes and squares, New directions in Computational Geometry, The Goodman-Pollack Festschrift volume, Aronov et al., eds., Springer-Verlag, 2003, 25–41.

How many squares are there, Mr. Franklin?: Constructing and enumerating Franklin squares, Amer.Math.Monthly, Vol.111, 2004, 394–410.

Magic graphs and the faces of the Birkhoff polytope, arXiv:math.CO/0405181

Anyone whose goal is 'something higher' must expect someday to suffer vertigo. What is vertigo? Fear of falling? No, Vertigo is something other than fear of falling. It is the voice of the emptiness below us which tempts and lures us, it is the desire to fall, against which, terrified, we defend ourselves. – Milan Kundera, The Unbearable Lightness of Being

8 – 10 am, January 10, 2005, Zhong Yuan 106

ACM Class Graph Theory Final Exam

Open Book

Name: ____________ Student ID: ____________

Score: ____________

Answer the following questions clearly and completely. The test is worth 120 points total. If you earn x marks, then your final exam score y will be $\min(x, 100)$. But $y - x$ may influence your course grade in some way.

1. (20 marks) Given a digraph G along with a weight function $w \in \mathbb{R}^{E(G)}$, a negative cycle for (G, w) is a cycle C in G with $\sum_{e \in C} w(e) < 0$; a negative flow for (G, w) is a function $f \in \mathbb{R}_+^{E(G)}$ such that $\sum_{e \in E(G)} f(e)w(e) < 0$ and for each vertex $v \in V(G)$, $\sum_{e \in \delta^+(v)} f(e) = \sum_{e \in \delta^-(v)} f(e)$, where $\delta^+(v)$ is the set of outgoing arcs at v and $\delta^-(v)$ the set of incoming arcs at v; and a potential function p for (G, w) is a function $p \in \mathbb{R}^{V(G)}$ such that for each $e \in E(G)$ it holds $p(y) - p(x) \le w(e)$, where x is the initial vertex of e and y the terminal vertex of e.

(i) Use Farkas' Lemma [76] to show that (G, w) possesses either a potential function or a negative flow, but not both;

(ii) Prove that (G, w) has a potential function if and only if it has no negative cycle.

[76] Given $A \in \mathbb{R}^{n \times m}$ and $w \in \mathbb{R}^n$, then either

$$\exists p \in \mathbb{R}^m, Ap \le w$$

or

$$\exists f \in \mathbb{R}_+^n, f^\top A = 0, f^\top w < 0$$

but not both.

2. (10 marks) Prove that each 3×3 nonnegative matrix of constant line sum 1 is a convex combination of at most 5 permutation matrices.

3. (10 marks) For a bipartite graph G, prove that the convex hull of the incidence vectors of all perfect matchings of G over $E(G)$ is the polytope defined by

$$\begin{cases} x \in \mathbb{R}_+^{E(G)}; \\ \sum_{e \in \delta(v)} x(e) = 1, \forall v \in V(G). \end{cases}$$

4. (15 marks) For any graph G with k connected components and n vertices, its rank is defined to be $r(G) = n - k$. Show that for any natural number k, a connected graph G may be split up into at most k trees if and only if for any subgraph H of G, $r(H) \geq \frac{|E(H)|}{k}$.

5. (15 marks) Let G be a graph. Define the fractional edge coloring number $\chi_e^*(G)$ to be $\min\{\sum_{M \in \mathcal{M}} \lambda(M) : \lambda \in \mathbb{R}_+^{\mathcal{M}}, \sum_{M \in \mathcal{M}} \lambda(M)\chi^M = \chi^{E(G)}\}$, where $\mathcal{M}$ denotes the set of all matchings of G. For any natural number k, let G_k be the graph obtained from G by replacing each edge by k parallel edges of the same endpoints. Prove that $\chi_e^*(G) = \lim_{k \to \infty} \frac{\chi_e(G_k)}{k}$. (You are allowed to directly assume that $\lim_{k \to \infty} \frac{\chi_e(G_k)}{k}$ exists, namely you need not worry about proving this preliminary fact.)

6. (20 marks)

(i) Let $P_n = \{x \in \mathbb{R}_+^n : \sum_{i=1}^n x_i = 1\}$. Make use of the LP duality theorem to prove Von Neumann's minmax theorem on two-person zero-sum game: For every $C \in \mathbb{R}^{m \times n}$ we have $\max_{x \in P_m} \min_{y \in P_n} x^\top Cy = \min_{y \in P_n} \max_{x \in P_m} x^\top Cy$.

(ii) Let Π be a problem with a finite set $\mathcal{I}$ of input instances of a fixed size and a finite set of deterministic algorithms $\mathcal{A}$. For input $I \in \mathcal{I}$ and algorithm $A \in \mathcal{A}$, let $C(I, A)$ denote the running time of algorithm A on input I. For probability distribution p over $\mathcal{I}$ and q over $\mathcal{A}$, let I_p denote a random input chosen according to

p and A_q a random algorithm chosen according to q. Prove Yao's minimax principle: For all distributions p over $\mathcal{I}$ and q over $\mathcal{A}$, it holds $\min_{A\in\mathcal{A}} \mathbf{E}[C(I_p, A)] \leq \max_{I\in\mathcal{I}} \mathbf{E}[C(I, A_q)]$.

7. (30 marks) For each part, circle `T` for True or `F` for False. No explanation required.

`T` `F` a. Every polyhedron in $\mathbb{R}_+^{2005}$ has at least one vertex.

`T` `F` b. An assortment of five committees labelled by $\mathcal{Z}_5$ needs to be scheduled, each for a total of one hour. Suppose that for each $i \in \mathcal{Z}_5$, for the reason of having common members committee i and committee $i+1$ cannot meet at the same time. Then the shortest time interval in which all the committees can be scheduled is 3 hours.

`T` `F` c. A polyhedron is integral if and only if it has integral vertices.

`T` `F` d. A totally unimodular matrix must be unimodular.

`T` `F` e. The dual program to

$$P: \max_z \sum_{x\in X} c(x)z(x)$$

subject to $z(x) \geq 0, \forall x \in X$ and $\sum_{x\in U} z(x) \leq r(U), \forall U \subseteq X$, is

$$DP: \min_y \sum_{U\subseteq X} y(U)r(U)$$

subject to $y(U) \geq 0, \forall U \subseteq X$, and $\sum_{x\in U} y(U) \geq c(x), \forall x \in X$.

`T` `F` f. For any graph G, the convex hull of the incidence vectors of all matchings of G over $E(G)$ is fully determined by the following set of inequalities

$$\begin{cases} x \in \mathbb{R}_+^{E(G)}; \\ \sum_{e\in\delta(v)} x(e) \leq 1, \text{ for any non-isolated vertex } v \text{ such that if } v \text{ has only} \\ \text{ one neighbor } u, \text{then} \{u, v\} \text{ is a connected component of } G, \\ \text{and if } v \text{ has exactly two neighbors, then they are not adjacent;} \\ \sum_{e\in\gamma(U)} x(e) \leq \lfloor \frac{|U|}{2} \rfloor, \text{for any } U \in \mathcal{P}_{\text{odd}}(V(G)). \end{cases}$$

I have neither given nor received unauthorized assistance on this assignment.

Pledge: ____________

A professor always included this question on his final exams: "What did you think of this course?" He discontinued the practice after receiving this response: "This was the most complete course I ever took. Anything we didn't go over during class was covered in the final exam." – Adapted from Reader's Digest

References

[1] Noga Alon, Dmitry N. Kozlov, Coins with arbitrary weights, *J. Algorithms* 25 (1997), 162–176.

[2] Achim Bachem, Walter Kern, *Linear Programming Duality: An Introduction to Oriented Matroids,* Universitext 074, Springer, 1992.

[3] R.A. Brualdi, Common transversals and strong exchange systems, *J. Combinatorial Theory* 8 (1970), 307–329.

[4] Vašek Chvátal, Edmonds polytopes and a hierarchy of combinatorial problems, *Discrete Math.* 4 (1973), 305–337.

[5] B.L. Dietrich, A.J. Hoffman, On greedy algorithms, partially ordered sets and submodular functions, *IBM Journal of Research and Development* 47 (2003), 25–30. (Available online at `www.research.ibm.com/journal/rd/471/dietrich.pdf`)

[6] J. Edmonds, D.R. Fulkerson, Transversals and matroid partition, *J. Res. Nat. Bur. Stand.* 69B (1965), 147–153.

[7] Jack Edmonds, Maximum matching and a polyhedron with $0,1$-vertices, *Journal of Research National Bureau of Standards Section B* 69 (1965), 125–130.

[8] J. Edmonds, Submodular functions, matroids and certain polyhedra, *Proc. Int. Conf. on Combinatorics (Calgary)* Gordon and Breach (New York) (1970), 69–87.

[9] J. Edmonds, R. Giles, A min-max relation for submodular functions on graphs, in: *Studies in Integer Programming* (Peter Hammer, et al. eds.) Annals of Discrete Mathematics 1 (1977), 185–204

[10] E. Egerváry, On combinatorial properties of matrices, (in Hungarian) *Mat. Lapok* 38 (1931), 16–28.

[11] András Frank, Tamás Király, Zoltán Király, On the orientation of graphs and hypergraphs, *Discrete Applied Mathematics* 131 (2003), 385–400.

[12] Komei Fukuda, Lecture Notes on Oriented Matroids and Geometric Computation, available at http://roso.epfl.ch/fukuda/lect/polyex_handout/eth_omgc040616.pdf.

[13] D.R. Fulkerson, Blocking and anti-blocking pairs of polyhedra, *Math. Programming* 1 (1971), 168–194.

[14] J.F. Geelen, An algebraic approach to matching problems, Algorithm engineering as a new paradigm (Japanese) (Kyoto, 2000). Sūrikaisekikenkyūsho Kōkyūroku No. 1185, (2001), 63–71.

[15] Ira Gessel, Gian-Carlo Rota, (Eds.) Classic Papers in Combinatorics, Birkhäuser Boston, Inc., Boston, Mass., 1987.

[16] F.R. Giles, W.R. Pulleyblank, Total dual integrality and integer polyhedra, *Linear Algebra and its Applications* 28 (1979), 191–196.

[17] L. Goddyn, Cycle covers of graphs, Ph.D. Thesis, University of Waterloo, Waterloo, Ontario, Canada.

[18] A.J. Hoffman, D. Gale, Two Remarks on the Mendelsohn-Dulmage Theorem, Annals of Discrete Mathematics, 15 (1982), 171–177.

[19] A.J. Hoffman, J. Kruskal, Integral boundary points of convex polyhedra, in: *Linear Inequalities and Related Systems,* K. Kuhn and A. Tucker, Eds., pp. 223–246, Princeton University Press, Princeton, NJ, 1956.

[20] Alan J. Hoffman, Remarks on the Perfect Graph and Pluperfect Graph Theorems, *IBM Research Report RC23012*, 2003.

[21] Garth Isaak, Examples of combinatorial duality, available at http://www.lehigh.edu/~gi02/combdu.pdf

[22] Donal E. Knuth, The sandwich theorem, *The Electronic Journal of Combinatorics* 1 (1994), 1–48.

[23] D. König, Graphs and matrices, (in Hungarian) *Mat. Fiz. Lapok* 38 (1931), 116–119.

[24] Andrei Kotlov, Short proof of the Gallai-Edmonds structure theorem, available at: http://arxiv.org/abs/math/0011204, 2000.

[25] Dmitry N. Kozlov, Van H. Vu, Coins and Cones, *J. Comb. Theory* Ser. A 78 (1997), 1–14.

[26] L. Lovász, M.D. Plummer, Matching Theory, North-Holland, 1986.

[27] N.S. Mendelsohn, A.L. Dulmage, Some generalizations of the problem of distinct representatives, Canadian J. Math. 10 (1958), 230–241.

[28] Charles A. Micchelli (Ed.), Selected Papers of Alan J. Hoffman, With Commentary, World Scientific, 2003.

[29] W.R. Pulleyblank, Polyhedral Combinatorics, in: Mathematical Programming: The State of the Art, (Eds., A. Bachem, M. Grötschel, B. Korte) pp. 312–345, Springer-Verlag, 1983.

[30] Alexander Schrijver, *Theory of Linear and Integer Programming*, John Wiley & Sons, 1998.

[31] Alexander Schrijver, *A Course in Combinatorial Optimization* (Course notes), 2004, available at: http://homepages.cwi.nl/%7Elex/files/dict.pdf.

[32] Alexander Schrijver, Combinatorial Optimization: Polyhedra and Efficiency (3 volume-set) (Algorithms and Combinatorics, Vol. 24), Springer, 2003.

[33] P.D. Seymour, On multi-colourings of cubic graphs, and conjectures of Fulkerson and Tutte, *Proc. London Math. Soc.* 38 (1979), 423–460.

[34] P.D. Seymour, Sums of circuits, in: *Graph Theory and Related Topics*, (J.A. Bondy and U.R.S. Murty Eds.) Academic Press (1979) 341–355.

[35] G. Szekeres, Polyhedral decompositions of cubic graphs, *Bull. Austral. Math. Soc.* 8 (1973), 367–387.

[36] J.H. van Lint, R.M. Wilson, *A Course in Combinatorics*, (Second Edition) Cambridge University Press, 2001.