Algebraic Combinatorics
an overview

Gabriel Coutinho

December 10, 2020

These are the course notes of a (under)grad course being offered at UFMG in 2019.1.

Contents

1 Power series and generating functions

1.1 Definition and operations

Given a sequence of numbers $A = (a_k)_{k\geq 0}$, one defines the generating function associated to it by

$$A(x) = \sum_{k\geq 0} a_k x^k.$$

Despite the perhaps misleading name and notation that suggests $A(x)$ is a function, this power series should be seen as a formal object. That is, you should not worry for example whether the infinite sum converges or not. This would be a problem if one would be interested in computing $A(x_0)$ for a real number $x_0 \neq 0$, but that shall never[1] be the case. Instead, our only concern is to know or compute or discover or be able to find all coefficients of a power series in a finite process.

We start with some rules.

(i) Given two power series $A(x) = \sum_{k\geq 0} a_k x^k$ and $B(x) = \sum_{k\geq 0} b_k x^k$, their sum is defined as:

$$A(x) + B(x) = \sum_{k\geq 0} (a_k + b_k) x^k.$$

(ii) Given two power series $A(x)$ and $B(x)$, their product is defined as:

$$A(x)B(x) = \sum_{k\geq 0} \left(\sum_{j=0}^{k} a_j b_{k-j} \right) x^k.$$

Note in particular that $A(x)B(x) = B(x)A(x)$ (a fact that is not necessarily true for all mathematical objects you can multiply — I heard "matrices"?).

Seen as formal objects, one might wonder what kind of mathematical object the set of all formal power series are. They form a "ring", but you need not worry about this for now.

Sometimes, given a power series $A(x)$, it is possible to find its multiplicative inverse, that is, a power series $B(x)$ so that $A(x)B(x) = 1$.

Example 1.1. Say $A(x) = \sum_{k\geq 0} x^k$. Is there a $B(x)$ such that $A(x)B(x) = 1$ (according to the product rule we placed above) ?

Certainly. Note that if you start examining from b_0, it must be that $b_0 = 1$. Next, you will find the only possibility $b_1 = -1$. And surprisingly, this is all you need:

$$\left(\sum_{k\geq 0} x^k \right) (1 - x) = 1.$$

(Note that $B(x) = 1 - x$ is a perfectly valid power series: it is just that $b_k = 0$ for all $k \geq 2$.) □

[1] Ok, "never" is a strong word. Maybe at some point we might be interested in doing this, but then we shall worry about convergence.

Exercise 1.2. Find the inverse of $A(x) = \sum_{k\geq 0}(k+1)x^k$.

Exercise 1.3. Is it true that the inverse of any power series is a finite sum, that is, a polynomial?

Exercise 1.4. What if I had asked for the inverse of $A(x) = \sum_{k\geq 0} kx^k$?

Exercise 1.5. Can you guess now which power series have a multiplicative inverse and which do not?

1.2 Counting - a first example

Now we move forward a bit. Knowing that we can add and multiply and compute inverses of power series, let us see how this can be actually useful.

Example 1.6. A sequence of numbers (a_k) satisfies the recurrence $a_{k+1} = 2a_k + 1$, with $a_0 = 0$. Can we find a "formula" for a_k? Well, define

$$A(x) = \sum_{k\geq 0} a_k x^k.$$

Note that

$$\sum_{k\geq 0} a_{k+1}x^k = \sum_{k\geq 0}(2a_k + 1)x^k,$$

Multiply both sides by x. We obtain

$$A(x) - a_0 = 2xA(x) + x\left(\sum_{k\geq 0} x^k\right),$$

thus, as $a_0 = 0$, and using the multiplicative inverses,

$$A(x) = \frac{x}{(1-x)(1-2x)}.$$

(Here $1/P(x)$ means the multiplicative inverse of the power series $P(x)$.) Now we would like to deal with the expression on the right hand side, and fortunately your second calculus course comes to aid. The partial fraction expansion

$$\frac{1}{(1-x)(1-2x)} = \frac{\alpha}{1-x} + \frac{\beta}{1-2x}$$

has $\alpha = -1$ and $\beta = 2$. Thus

$$A(x) = 2x\sum_{k\geq 0}(2x)^k - x\sum_{k\geq 0} x^k.$$

We want to know who is a_n. Easily we have

$$[x^n]A(x) = 2^n - 1.$$

□

Exercise 1.7. Find a formula for the Fibonacci numbers.

1.3 Derivative

We can also define an operator that maps formal power series to formal power series called "derivative". It behaves just as you would expect in terms of the operations and rules, but here it has absolutely no meaning in terms limits and analysis[2]

The formal derivative of the power series $A(x) = \sum_{k\geq 0} a_k x^k$ is defined as

$$A(x)' = \sum_{k\geq 0} k a_k x^{k-1}.$$

Exercise 1.8. Verify (using the definitions of the operations) that

$$(A(x) + B(x))' = A(x)' + B(x)',$$

and that

$$(A(x)B(x))' = A(x)'B(x) + A(x)B(x)'.$$

Further, verify that if $A(x)' = 0$, then $A(x) = a_0$ for some a_0.

Exercise 1.9. Verify now that

$$\sum_{k\geq 0} k x^{k-1} = \frac{1}{(1-x)^2}$$

in two different ways: using derivatives, or simply making $(1/(1-x))^2$.

Exercise 1.10. Give the following power series

$$\sum_{k\geq 0} \frac{1}{k!} x^k$$

the special name of $\exp(x)$. Prove that if $A(x)' = A(x)$, then $A(x) = \alpha \exp(x)$ for some constant α.

Now let us apply this concept to another counting problem.

Example 1.11. Say we have a sequence (a_k) with $a_{k+1} = 2a_k + k$, $a_0 = 1$. Moving on

$$\sum_{k\geq 0} a_{k+1} x^k = \sum_{k\geq 0} (2a_k + k) x^k,$$

which, after multiplying by x, gives

$$A(x) - 1 = 2xA(x) + \frac{x^2}{(1-x)^2}.$$

[2]You will soon realize that we are living the dream of doing "calculus" without worrying about analysis — and as long as you do not plug-in values, everything is safe.

Rearranging terms, and aiming to a partial fraction expansion, we reach

$$A(x) = \frac{1-2x+2x^2}{(1-x)^2(1-2x)} = \frac{\alpha}{(1-x)^2} + \frac{\beta}{(1-x)} + \frac{\gamma}{(1-2x)}.$$

Now this last equality lives in the realm of rational functions and here we can actually make use of substitution to find the coefficients α, β and γ. Multiplying both sides by $(1-x)^2$ and making $x = 1$, gives $\alpha = -1$. Multiply by $(1-2x)$ and making $x = 1/2$ leads to $\gamma = 2$. Now simply making $x = 0$ and knowing α and γ, we find $\beta = 0$, which therefore takes us to

$$A(x) = \frac{1-2x+2x^2}{(1-x)^2(1-2x)} = \frac{-1}{(1-x)^2} + \frac{2}{(1-2x)}.$$

Thus

$$A(x) = -\sum_{k\geq 0} kx^{k-1} + 2\sum_{k\geq 0}(2x)^k,$$

hence

$$a_n = [x^n]A(x) = 2^{n+1} - (n+1).$$

□

1.4 Binomial theorem

Now for several applications of generating series, we might encounter things that look like

$$(1+rx)^\alpha$$

where r is any number and α is not an integer. Say for now α is a rational number. How to deal with such things? First, let us clarify what these mean.

Example 1.12. Say we want to find the first few term of the series $A(x)$ so that $A(x)^3 = 1 + x$. Note that $a_0^3 = 1$, so $a_0 = 1$. Then $3a_1 = 1$, so $a_1 = 1/3$. Who is a_2 ?

This series $A(x)$ is defined as $(1+x)^{1/3}$. Note that if $B(0) = b_0 = 1$, then $B(x)^{n/m}$ is always well defined, and its first term is always equal to 1. □

We start with the following result, which basically tells us that a formal power series is equal to the its MacLaurin series expansion.

Lemma 1.13. *Let $A(x)$ be a formal power series. Then*

$$A(x) = \sum_{k\geq 0} \frac{A^{(k)}(0)}{k!} x^k.$$

Proof. Follows immediately by induction, noting that

$$A^{(n)}(x) = \sum_{k\geq n} \frac{k!}{(k-n)!} a_k x^{k-n},$$

and thus

$$A^{(n)}(0) = n! \cdot a_n.$$

□

With that in hand, we can prove the Binomial Theorem (in the realm of formal power series).

Theorem 1.14. *For $\alpha \in \mathbb{Q}$ and r any number, we have*

$$(1+rx)^\alpha = \sum_{k\geq 0} \binom{\alpha}{k} (rx)^k,$$

with the understanding that $\binom{\alpha}{k}$, for $\alpha \in \mathbb{Q}$, means

$$\binom{\alpha}{k} = \frac{1}{k!} \cdot \prod_{j=1}^{k} (\alpha + 1 - j)$$

Proof. Say $A(x) = (1+rx)^\alpha$. Then

$$A^{(n)}(x) = \alpha(\alpha-1)\dots(\alpha-n+1)r^n(1+rx)^{\alpha-n}.$$

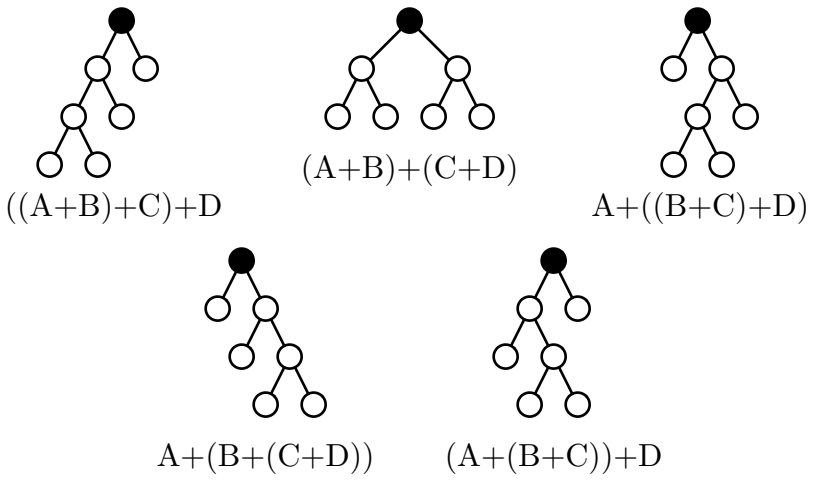
As we discussed above, the first term of $(1+rx)^{\alpha-n}$ is 1, thus

$$A^{(n)}(0) = \alpha(\alpha-1)\dots(\alpha-n+1)r^n.$$

The result now follows applying the lemma above. □

1.5 Catalan Numbers

Let us now bring up the first non-trivial application of formal power series to solve a counting problem. How many rooted complete binary trees (rooted tress, in which each node has 0 or 2 children) are there with n leaves (childless nodes)? Alternatively, into how many ways can you unambiguously parenthesize a sum of n elements? Hopefully, the following picture shall make both questions above clear:



Say C_{n-1} (the shift is for historical reasons) is the quantity of such things. The picture above shows that $C_3 = 5$. It is easy to see that $C_0 = 1$, $C_1 = 1$, $C_2 = 2$.

Exercise 1.15. Find C_4.

Now if you noted a clever way to solve the question above, you are probably ready to write a recurrence relation to these coefficients:

$$C_n = C_0C_{n-1} + C_2C_{n-2} + ... + C_{n-1}C_0.$$

Now let us define

$$A(x) = \sum_{k\geq 0} C_k x^k.$$

What does the recurrence above tells us? Note that

$$A(x)^2 = \sum_{k\geq 0} \left(\sum_{j=0}^{k} C_j C_{k-j} \right) x^k = \sum_{k\geq 0} C_{k+1} x^k.$$

Hence

$$xA(x)^2 = A(x) - C_0 = A(x) - 1.$$

Up to this point, we had no need to deal with things such as $1/x$. In fact, we argued that only power series with a non-zero constant term had an inverse, but nothing prevents us from extending the ring of formal power series to the ring of formal Laurent series, which are things of the form

$$\sum_{k\geq \alpha} a_k x^k,$$

where α is any integer (possibly negative). The set of all such things is a field — in fact, it is the field of of fractions of the ring of formal power series. Thus you are free to write things as $1/A(x)$ for any $A(x) \neq 0$. In particular, you can complete squares in the equation above, and there will be precisely two formal series satisfying it. That is,

$$(2x) \cdot A(x) = 1 \pm \sqrt{1-4x}.$$

From the Binomial Theorem, it follows that

$$2xA(x) = 1 \pm \sum_{k\geq 0} \binom{1/2}{k} (-4)^k x^k$$

Because the left hand side has no constant term, it follows that the solution we are looking for is

$$A(x) = \frac{-1}{2} \sum_{k\geq 1} \binom{1/2}{k} (-4)^k x^{k-1}.$$

Therefore

$$C_{n-1} = \frac{-1}{2}\binom{1/2}{n}(-4)^n = \frac{4^n}{2^{n+1}n!} \cdot 1 \cdot 3 \cdot ... \cdot (2n-3) = \frac{1}{n}\binom{2n-2}{n-1}.$$

You can check now that C_4 is indeed 14.

You are invited to check the wikipedia page about Catalan numbers and learn just so many interesting connections between distinct combinatorial objects, as well as combinatorial derivation of the formula for C_n found above.

`https://en.wikipedia.org/wiki/Catalan_number`

1.6 Composition

Say we have two power series $A(x)$ and $B(x)$. Can we actually define the "composition" $A(B(x))$? Naively, we would say this is

$$\sum_{k\geq 0} a_k \left(\sum_{j\geq 0} b_j x^j\right)^k$$

Can we actually compute the coefficient, of, say, x? This shall be equal to

$$a_1b_1 + a_2(b_0b_1 + b_1b_0) + a_3(3b_1b_0^2) + ...$$

which violates our original assumption that the coefficients should be computable through a finite process (recall that in the realm of formal power series, we are not allowed to plug-in values and verify if the above series converges or not...)

Now, if $b_0 = 0$, then things change. In fact, all coefficients of $A(B(x))$ become computable in a finite process. In particular

$$[x^0]A(B(x)) = a_0 \ , \ [x^1]A(B(x)) = a_1b_1 \ , \ [x^2]A(B(x)) = a_1b_2 + a_2b_1^2 \ ,$$

$$[x^3]A(B(x)) = a_1b_3 + a_2(2b_1b_2) + a_3b_1^3 \ , \text{ and so on.}$$

- The composition $A(B(x))$ is defined if and only if $b_0 = 0$ or $A(x)$ is a polynomial.

Theorem 1.16. *Assume $A(B(x))$ is well defined. Then*

$$(A(B(x)))' = B'(x)A'(B(x)).$$

Proof. This follows immediately from

$$[B(x)^k]' = B'(x) \cdot kB(x)^{k-1},$$

which can be proved by induction using the product formula. □

Theorem 1.17. *Assume $A(0) = 0$. Then there is a series $B(x)$ with $B(0) = 0$ so that*

$$A(B(x)) = B(A(x)) = x.$$

Such function $B(x)$ is called the functional inverse of $A(x)$.

Proof. Given $A(x)$, it is easy to define $B(x)$ so that $A(B(x)) = x$. For instance, $b_1 = 1/a_1$, and the remaining coefficients can be defined recursively. Now let $C(x)$ be constructed so that $B(C(x)) = x$. From $A(B(x)) = x$, substitute now x for $C(x)$, obtaining

$$A(x) = A(B(C(x))) = C(x),$$

as we wanted. □

Recall now that we defined the power series

$$\exp(x) = \sum_{k\geq 0} \frac{1}{k!}x^k.$$

We can also define a power series

$$\log(1+x) = \sum_{k\geq 1} \frac{(-1)^{k+1}}{k}x^k.$$

Note from this definition that

$$\log(1+x)' = 1 - x + x^2 - .. + .. = \frac{1}{1+x}.$$

In fact, if $B(0) = 0$, it now follows that

$$\log(1+B(x))' = B'(x)\frac{1}{1+B(x)}.$$

Exercise 1.18. Prove that

$$\log(\exp(x)) = x.$$

(Hint: take the derivative from the expression on the left.)

Exercise 1.19. Prove that $\log(1+x)^a = a\log(1+x)$, and that $(1+x)^a(1+x)^b = (1+x)^{a+b}$.

1.7 LIFT

In this section, we prove (a weak version of[3]) Lagrange Implicit Function Theorem, which shall prove itself a very useful tool.

Theorem 1.20. *Let $\phi(x)$ and $f(x)$ be formal power series, with $\phi(0) \neq 0$. Assume $A(x)$ satisfies the functional equation*

$$A(x) = x\phi(A(x)).$$

Then, for $n \geq 1$,

$$[x^n]f(A(x)) = \frac{1}{n}[x^{n-1}]f'(x)\phi(x)^n.$$

For this proof, we again assume to be working with formal Laurent series, that is, things of the form $\sum_{k\geq\alpha} a_k x^k$, with $\alpha \in \mathbb{Z}$, possibly a negative number. Note that for any formal Laurent series $A(x)$, it follows that $[x^{-1}]A'(x) = 0$. Let $\operatorname{val} A(x)$ be equal to the smallest index k so that $a_k \neq 0$.

[3]the strong version says that given $\phi(x)$, $A(x)$ exists and is unique. But to prove this, we actually need some analysis...

Proof. First, assume $B(x)$ is a formal power series with $B(0) = 0$, and let $m = \operatorname{val} B(x)$. First we shall see that

$$[x^{-1}]A(x) = \frac{1}{m}[x^{-1}]A(B(x))B'(x).$$

To see this, first observe that we can simply ignore the coefficients of $A(x)$ (due to linearity). Then we can analyse each power separately. For $n \neq -1$, we have

$$[x^{-1}]B(x)^n B'(x) = \frac{1}{n+1}[x^{-1}](B(x)^{n+1})' = 0.$$

For $n = -1$, first let $B(x) = x^m C(x)$, where $C(0) \neq 0$. So we have

$$\frac{1}{m}[x^{-1}]B(x)^{-1}B'(x) = \frac{1}{m}[x^{-1}]\frac{mx^{m-1}C(x) + x^m C'(x)}{x^m C(x)} = 1 + \frac{1}{m}[x^{-1}]\log'(C(x)) = 1.$$

Now let $B(x) = \frac{x}{\phi(x)}$. By hypothesis, $B(A(x)) = x$, and thus $A(B(x)) = x$. Note also that both $\operatorname{val} A(X) = \operatorname{val} B(x) = 1$. Thus

$$\begin{aligned}
[x^n]f(A(x)) &= [x^{-1}]x^{-n-1}f(A(x)) \\
&= [x^{-1}]B(x)^{-n-1}f(A(B(x)))B'(x) \\
&= [x^{-1}]B(x)^{-n-1}f(x)B'(x) \\
&= [x^{-1}]\left(\frac{B(x)^{-n}}{-n}\right)' f(x) \\
&= -[x^{-1}]\left(\frac{B(x)^{-n}}{-n}\right) f'(x) \\
&= \frac{1}{n}[x^{-1}]\phi(x)^n x^{-n} f'(x) \\
&= \frac{1}{n}[x^{n-1}]\phi(x)^n f'(x).
\end{aligned}$$

□

What is this useful for?

Example 1.21. Recall the functional equation we had for the Catalan generating series:

$$xA(x)^2 + 1 = A(x).$$

Make $B(x) = A(x) - 1$. Thus

$$x(B(x) + 1)^2 = B(x).$$

Let $\phi(x) = (1 + x)^2$. This is all set to apply LIFT, as $B(x) = x\phi(B(X))$. We obtain

$$c_n = [x^n]A(x) = [x^n]B(x) = \frac{1}{n}[x^{n-1}](1 + x)^{2n} = \frac{1}{n+1}\binom{2n}{n}.$$

Exercise 1.22. Let $A(x)$ be the Catalan generating series. Find the coefficients of $A(x)^k$ using LIFT.

Exercise 1.23. How many rooted trees with n non-leaves are there so that every node has either 0 or m children?

Corollary 1.24. *Let $A(x) = xB(x)$, with $B(0) \neq 0$. Let $C(x)$ be the compositional inverse of $A(x)$, meaning, $A(C(x)) = C(A(x)) = x$. Then, for $n \geq 1$,*

$$[x^n]C(x) = \frac{1}{n}[x^{n-1}]B(x)^{-n}.$$

Proof. Follow immediately from LIFT, noting that

$$C(x) = xB^{-1}(C(x)),$$

where B^{-1} is the multiplicative inverse of $B(x)$. □

1.8 Application to quicksort analysis

Suppose you are given a list of n distinct integers and you are required to order this list. The quicksort algorithm performs as follows:

(a) Pick the first element of the list, say α.

(b) Partition the remaining of the list into sublists L^- and L^+ which are respectively the elements smaller than and greater than α.

(c) Run the algorithm in L^- and L^+ (recursively).

(d) Return L^- sorted, α and L^+ sorted.

I shall now ask what is the "expected" running time (number of comparisons) of this procedure if the initial list is supposedly random?

Let a_n be the expected number of comparisons needed to sort a list of length L. If this list is truly random, it follows that

$$a_n = (n-1) + \frac{1}{n}\sum_{k=0}^{n-1} a_k + a_{n-1-k} = (n-1) + \frac{2}{n}\sum_{k=0}^{n-1} a_k,$$

with initial value $a_0 = 0$. Now let

$$A(x) = \sum_{k\geq 0} a_k x^k.$$

Exercise 1.25. Prove that

$$xA'(x) = \frac{2x^2}{(1-x)^3} + \frac{2x}{1-x}A(x)$$

Now, this is a differential equation. It has a unique a solution (in the realm of formal power series), because the sequence a_k is uniquely defined, and its recurrence relation is equivalent to this ODE. So we need only guess one solution and verify it works. To find the best guess, treat this is a standard ODE and use your favourite method to solve it. After that, you will (successfully) verify that the solution you found works for the formal power series as well. In fact, we have

$$A(x) = \frac{-2(x + \log(1-x))}{(1-x)^2}.$$

Exercise 1.26. From this equation, show that

$$a_n = -4n + 2(n+1)\sum_{j=1}^{n} \frac{1}{j} \quad \text{(which in particular is } \approx 2n\log n + \Theta(n)).$$

1.9 Exponential generating functions

The exponential generating function is also a formal power series, but now we associate to the sequence $(a_k)_{k\geq 0}$ the series

$$A(x) = \sum_{k\geq 0} \frac{a_k}{k!} x^k.$$

This seems like an artificial addition, but it shall turn out to be quite convenient for some purposes. First, note the following immediate properties:

(i) $xA(x) = \sum_{k\geq 1} k\frac{a_{k-1}}{k!}x^k$.

(ii) $A'(x) = \sum_{k\geq 0} \frac{a_{k+1}}{k!}x^k$.

(iii) $xA'(x) = \sum_{k\geq 1} k\frac{a_k}{k!}x^k$.

With formal ordinary power series, we would take the derivative to extract exponents, and multiply by x to shift coefficients. Here it was pretty much the opposite idea.

1.10 Dearrangements

Example 1.27. A permutation of $(1, ..., n)$ that fixes no element is called a dearrangement. How many of those are there? Say this number is d_n. Note that $d_2 = 1$, and $d_3 = 2$ (but $d_4 \neq 3$...) Well, for any dearrangement on n elements, the last element can be mapped to any of the other $n-1$ possibilities. If n goes to 1, say, then there are two possibilities. If 1 goes to n, then what is left is precisely a dearrangement on $n-2$ elements. If 1 goes somewhere else, then the whole thing is bijection with a dearrangement on $n-1$ elements. So (and already shifting by indices by 1...)

$$d_{n+1} = n(d_n + d_{n-1}).$$

(The factor n stands for all possible places where $n+1$ could go.) Note that if we simply define $d_0 = 1$ and $d_1 = 0$, all is safe. Now we can define the exponential generating functions of this sequence

$$D(x) = \sum_{k\geq 0} \frac{d_k}{k!} x^k.$$

From this, it follows that

$$D'(x) = xD(x) + xD'(x).$$

Thus

$$D'(x) = \frac{x}{1-x} D(x).$$

Again, solving the differential equation, it follows that

$$D(x) = \exp(-x)\frac{1}{1-x} = \left(\sum_{k\geq 0} \frac{(-1)^k}{k!} x^k\right)\left(1 + x + x^2 + ...\right)$$

which gives

$$\frac{d_k}{k!} = \sum_{j=0}^{k} \frac{(-1)^j}{j!}.$$

□

Exercise 1.28. Verify that d_k is the nearest integer to $k!/\mathrm{e}$, for all k.

Exercise 1.29. Find the formula for d_k using ordinary generating functions.

1.11 Partitions and Bell numbers

Example 1.30. The Bell number B_n is defined as the number of partitions of a set of size n. For example: $B_1 = 1$, $B_2 = 2$, and $B_3 = 5$, and $B_4 = 15$. First note that

$$B_n = \sum_{i=1}^{n} \binom{n-1}{i-1} B_{n-i},$$

using the convention $B_0 = 1$. (One way of understanding this recurrence relation is that a partition of $\{1, ..., n\}$ can be determined by first picking the subset containing n and then partitioning the remaining elements.) Now let

$$B(x) = \sum_{k\geq 0} \frac{B_k}{k!} x^k.$$

Thus

$$B'(x) = \sum_{k\geq 0} \frac{B_{k+1}}{k!} x^k,$$

and then

$$B'(x) = \sum_{k\geq 0} \left(\sum_{i=0}^{k} \binom{k}{i} B_{k-i}\right) \frac{x^k}{k!} = \sum_{k\geq 0} \sum_{i=0}^{k} \frac{x^i}{i!} \frac{B_{k-i} x^{k-i}}{(k-i)!},$$

therefore

$$B'(x) = \exp(x)B(x).$$

Again, a differential equation. The general solution $B(x) = \alpha \exp(\exp(x))$ gives $\alpha = \exp(-1)$ as $B(0) = 1$. Finally

$$B(x) = \exp(\exp(x) - 1).$$

We would like to have a nice formula for B_n. In the realm of formal power series, there is not much we can do here. However, by noting that $\mathrm{e}^{(\mathrm{e}^x)}$ has a convergent Taylor series, it follows that

$$B(x) = \frac{1}{\mathrm{e}} \sum_{k\geq 0} \frac{1}{k!} (\mathrm{e}^x)^k = \frac{1}{\mathrm{e}} \sum_{k\geq 0} \sum_{j\geq 0} \frac{1}{k!j!} (kx)^j.$$

Thus

$$B_n = \frac{1}{\mathrm{e}} \sum_{k\geq 0} \frac{k^n}{k!}.$$

This is known as Dobinski's formula. □

1.12 Trees and graphs

Let T_n stand for the number of rooted labelled trees on n vertices. For instance, $T_1 = 1$, $T_2 = 2$, $T_3 = 9$, but $T_4 = 64$ (note that the number of rooted labelled trees on n vertices is equal to n times the number of unrooted labelled trees on n vertices).

Now let us try to associate T_{n+1} with previous values. There are $n + 1$ choices for the root. For each, if we delete the root, we are left with a forest, say with k trees. Say each one has size j_1, ... , j_k. We also permute the n vertices to chose which goes to each of the trees, making sure to discount inner permutations inside each tree. We then multiply by the possible number of rooted trees that can be made in each of the subtrees (note that the root would be the vertex attached to the original root). In the end, we recall to divide by $k!$, to account for permutations of the whole trees of the forest. All together, this gives

$$T_{n+1} = (n+1) \sum_{k=1}^{n} \frac{1}{k!} \sum_{\substack{j_1,\dots,j_k \geq 1 \\ j_1+\dots+j_k=n}} \frac{n!}{j_1! \cdot \dots \cdot j_k!} T_{j_1} \cdot \dots \cdot T_{j_k}$$

Let $T(x) = \sum_{n\geq 1} T_n \frac{x^n}{n!}$. The relation above implies that

$$\frac{1}{x} T(x) = \sum_{n\geq 0} T_{n+1} \frac{x^n}{(n+1)!} = 1 + \sum_{n\geq 1} \sum_{k=1}^{n} \frac{1}{k!} \left(\sum_{\substack{j_1,\dots,j_k \geq 1 \\ j_1+\dots+j_k=n}} \frac{n!}{j_1! \cdot \dots \cdot j_k!} T_{j_1} \cdot \dots \cdot T_{j_k} \right) \frac{x^n}{n!}.$$

Thus

$$\frac{1}{x} T(x) = 1 + \sum_{n\geq 1} \sum_{k=1}^{n} \frac{1}{k!} \sum_{\substack{j_1,\dots,j_k \geq 1 \\ j_1+\dots+j_k=n}} T_{j_1} \frac{x^{j_1}}{j_1!} \cdot T_{j_2} \frac{x^{j_2}}{j_2!} \cdot \dots \cdot T_{j_k} \frac{x^{j_k}}{j_k!}.$$

We can know split the sum for all possible values of k. If $k = 1$, we have simply $T(x)$. If $k = 2$, we are seeing $T(x)^2$. In fact

$$\frac{1}{x}T(x) = 1 + \sum_{k\geq 1} \frac{T(x)^k}{k!} = \exp(T(x)).$$

Just like that, we have arrived at the celebrated functional relation

$$T(x) = x\exp(T(x)).$$

Not only this is a clean nice expression, but it also is ready to be hammered with Lagrange Implicit Function Theorem. Meaning, $\phi(x) = \exp(x)$, $f(x) = x$, and we have

$$\frac{T_n}{n!} = \frac{1}{n}[x^{n-1}]\exp(nx) = \frac{n^{n-1}}{(n-1)!} \implies T_n = n^{n-1}.$$

Naturally, the number of unrooted labelled trees will be n^{n-2}.

In fact, we have just witnessed the classical application of a very general principle. When counting combinatorial structures which are "disconnected" and are somehow made of "connected" substructures satisfying the same property, the exponential generating function comes very handy.

To see another application, let g_n stand for the number of graphs on n vertices (and possibly disconnected) so that each connected component satisfy a certain property. Let c_n be the number of connected graphs on n vertices satisfying the same property. If $C(x) = \sum_{k\geq 1} c_k \frac{x^k}{k!}$, then $C(x)^n/n!$ is the exponential generating for the graphs with precisely n connected components satisfying the given property. If $G(x) = \sum_{n\geq 1} g_n \frac{x^n}{n!}$, then

$$G(x) = \sum_{n\geq 1} \frac{C(x)^n}{n!} = \exp(C(x)) - 1.$$

It is not difficult to apply the reasoning above to find the number of rooted trees, as we did. The property in question is simply nothing, then g_n stands for the number of graphs on n vertices, which we all (should) know to be $g_n = 2^{\binom{n}{2}}$. Thus

$$C(x) = \log\left(1 + \sum_{n\geq 1} 2^{\binom{n}{2}} \frac{x^n}{n!}\right) = \sum_{k\geq 1} \frac{(-1)^{k+1}}{k}\left(\sum_{n\geq 1} 2^{\binom{n}{2}} \frac{x^n}{n!}\right)^k.$$

This might not lead to a nice formula, but it certainly allows for a decent method to compute the number of connected graphs on n vertices (as well as a good way of estimating).

Exercise 1.31. Find the exponential generating function for the number of labelled graphs on n vertices such that each connected component is a regular graph of valency 2?

Exercise 1.32. How many labelled forests of rooted trees on a total of n vertices are there? (This question should be very easy.)

Exercise 1.33. Find the exponential generating function for the numbers of labelled forests of unrooted trees on a total of n vertices.

1.13 Permutations

A permutation on a set of n elements $V = \{1, ..., n\}$ is a bijection from V to itself. Each permutation has a "cycle structure", which are the minimal sets you can partition V so that each class of the partition is invariant under the permutation.

Example 1.34. The permutation on $\{1, 2, 3, 4, 5\}$ that maps each of these elements respectively to $(2, 3, 1, 5, 4)$ contains two cycles. One corresponds to the action of the permutation on $\{1, 2, 3\}$, and the other on $\{4, 5\}$. In fact, this permutation can be represented in the following way:

$$(231)(45)$$

to indicate that 2 goes to 3, which goes to 1, which goes to 2 (each cycle turns around), and that 4 and 5 swap places. Note that the following would represent the same permutation:

$$(54)(123).$$

Permutations can also be represented by matrices (once you chose an ordering for the elements of the set). In fact, the permutation above is given by

$$\begin{pmatrix} 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix},$$

which applied to the vector, say, $\begin{pmatrix} 0 & 1 & 0 & 0 & 0 \end{pmatrix}^T$ gives $\begin{pmatrix} 0 & 0 & 1 & 0 & 0 \end{pmatrix}^T$, meaning that 2 goes to 3.

You can compose permutations, which means that you would multiply the matrices, or simply change the cycles. For instance, $(231) \circ (12)(3)$ means that 1 goes to 2, which then goes to 3. 3 is unaffected, then goes to 1. And 2 goes to 1, which then returns to 2. So

$$(231) \circ (12)(3) = (13)(2).$$

□

All cycles can be written as a composition of cycles of length 2 (and therefore all permutations can be written as a composition of cycles of length 2). Cycles of length 2 are called "transpositions". They are the only cycles which correspond to "symmetric" entries in the matrix representation of a permutation.

Example 1.35. Find a recurrence relation for the coefficients s_n that count the number of permutation whose all cycles have length either 1 or 2.

Exercise 1.36. Using your recurrence relation above, show (by induction?), that s_n is even for all $n > 1$, and that $s_n > \sqrt{n!}$ (if you know group theory, give a group theoretic reason on why s_n is even).

Exercise 1.37. Find the exponential generating function for the sequence s_n.

How many sets of k elements can be defined on a set of k elements? Clearly, only 1. Let $S(x) = \sum_{k\geq 0}(a_k/k!)x^k$ be the exponential generating series of such sequence, that is, with $a_k = 1$, we have $S(x) = \exp(x)$.

How many permutations can be defined? Clearly, $p_k = k!$. If $P(x)$ is the corresponding exponential generating series, it follows that $P(x) = (1-x)^{-1}$.

Now, looking at the cycles expressing a permutation on n elements, those fixed points (cycles of length 1) correspond to a subset of $\{1, ..., n\}$. The cycles of length bigger than one all contain only points being dearranged! It means that we can count the number of permutations p_n by first choosing k elements to be fixed, then multiplying this choice by the number of sets of k elements ($a_k = 1$) and the number of dearrangements on $n-k$ (d_{n-k}). Thus

$$P(x) = \sum_{n\geq 0} \frac{p_n}{n!}x^n = \sum_{n\geq 0}\sum_{k=0}^{n} \left(\binom{n}{k} a_k d_{n-k} \right) \frac{x^n}{n!} = S(x)D(x).$$

Immediately leading to

$$D(x) = \exp(-x)(1-x)^{-1}.$$

Note: no ODEs have been solved this time. Now this example has shown a very interesting principle: if you are counting the number of ways a set can be partitioned into two parts, one with one type of structure, and the other with a different type, then you shall eventually multiply exponential generating series.

Exercise 1.38. Recall now the sequence s_n of those permutations with cycles of length at most 2. Find its exponential generating function again, this time without using the recurrence relation. First, you will need to write the exponential generating series for the numbers q_{2k} which express the number of ways you can split $2k$ elements into k cycles of length 2.

Example 1.39. As a review of the past section, consider now the generating series for the numbers m_k of permutation on k elements with just one cycle. Clearly $m_k = (k-1)!$ (make $m_0 = 0$), and this is not a surprise. Regardless, we have

$$M(x) = \sum_{n\geq 1} \frac{x^n}{n} = -\log(1-x).$$

Now each permutation is made out of blocks, all of which corresponding to permutations of only one cycle. Thus

$$P(x) = M(x) + \frac{1}{2!}M(x)^2 + \frac{1}{3!}M(x)^3 + ... = \exp(M(x)),$$

thus

$$P(x) = \exp(\log(1-x))^{-1} = (1-x)^{-1} = 1 + x + x^2 + ...$$

exactly as we would expect.

Finally, one last exercise.

Exercise 1.40. For even n (only!), let e_n and o_n stand respectively for the number of permutations with all cycle of even and odd length. Let $E(x)$ and $O(x)$ be their exponential generating functions, and $P(x)$, again, the exponential generating function for all permutations (but recall, n is even!). Our goal is to show that $e_n = o_n$.

(a) Verify that $P(x) = (1-x^2)^{-1}$.

(b) Prove that $E(x) = (1-x^2)^{-1/2}$. Use the example above as an inspiration.

(c) Argue that $P(x) = E(x) \cdot O(x)$. Conclude that $O(x) = E(x)$, and thus $e_n = o_n$.

(d) Find a formula for e_n.

(e) Try to find a bijective proof of the formula in (d), comparing a permutation with cycles of even length with 2 distinct partitions of the set into subsets of size 2.

(f) Try to find a bijective proof of the equality above in (c) (this will be hard).

1.14 Bernoulli numbers

The Bernoulli numbers. They are defined by the recurrence $b_0 = 1$ and

$$\sum_{k=0}^{n} \binom{n+1}{k} b_k = 0.$$

How would we use this equation to find the exponential generating function? Well, this product looks very much like what would appear if we took the product of two generating functions. It is almost

$$B(x)\exp(x) = \left(\sum_{k\geq 0} \frac{b_k}{k!}x^k\right)\left(\sum_{j\geq 0} \frac{1}{j!}x^j\right),$$

except that the index seems to appear slightly off.

Exercise 1.41. Find a way to fix this, and prove that

$$B(x) = x(\exp(x)-1)^{-1}.$$

Note that $B(x) + x/2$ is an even function. Deduce from this that $b_k = 0$ for all odd $k \geq 3$.

The Bernoulli numbers are connected to many branches of mathematics. Check

`https://en.wikipedia.org/wiki/Bernoulli_number`.

Now for one interesting application, suppose you would like to find a formula for

$$p_m(n) = \sum_{k=0}^{n-1} k^m.$$

As you might (or might not) remember, when you were learning induction, some of these formulas for fixed m and variable were provided, and you had to prove them by induction. In fact, for fixed m, you probably remember that $p_m(n)$ is a polynomial of degree $m+1$ in n. What we would like to do now is to study $p_m(n)$ for when m and n vary (and perhaps find a formula?). So let

$$P(x,n) = \sum_{m\geq 0} p_m(n)\frac{x^m}{m!} = \sum_{m\geq 0}\left(\sum_{k=0}^{n-1} k^m\right)\frac{x^m}{m!} = \sum_{k=0}^{n-1}\sum_{m\geq 0}\frac{(kx)^m}{m!}.$$

Thus

$$P(x,n) = \frac{\exp(nx)-1}{\exp(x)-1}.$$

As a consequence

$$xP(x,n) = B(x)(\exp(nx)-1).$$

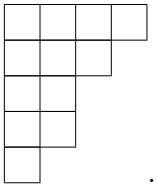
Exercise 1.42. Verify now that

$$\sum_{k=0}^{n-1} k^m = \frac{1}{m+1}\sum_{k=0}^{m}\binom{m+1}{k}B_k n^{m+1-k}.$$

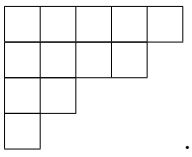
1.15 Integer partitions

Now let us return to studying partitions. This time, partitions of an integer. Let $p(n)$ represent the number of ways of writing n as a sum of positive integers (this time, we wish to disregard the order the summands appear). So $1 = 1$, $2 = 2$ or $2 = 1+1$, $3 = 3$, $3 = 2+1$ and $3 = 1+1+1$. And so on. Can we actually find a formula for $p(n)$? Well if you watched "The man who knew infinity", you would remember now that one of the greatest mathematical minds of all time, S. Ramanujan, devoted a great deal of effort to solve this problem (and even so, not "fully"). We shall be humble and content ourselves with only the very beginning of this beautiful theory. But we will nevertheless show at least one surprising result.

Before we continue, let us first create a new language to deal with partitions. If we have $12 = 4+3+2+2+1$, we will simply say $\lambda = 43221$ is a partition of 12. If we write a partition using what is called a Ferrers diagram, we mean to write each of these numbers as a row of squares, from largest number to smallest:



By reflecting this partition, we obtain what is known as the conjugate partition, called λ^*:



Here $\lambda = 43221$, and $\lambda^* = 5421$.

Exercise 1.43. Find the conjugate parition of $\lambda = 6443211$.

An immediate consequence is that $\lambda^{**} = \lambda$. If we define $p(n; \leq k)$ to mean the number of partitions of n with at most k summands, and $q(n; \leq k)$ the number of partitions of n whose each summand is at most k, then the duality between any partition and its conjugate immediately gives

Corollary 1.44.

$$p(n; \leq k) = q(n; \leq k).$$

Now, consider

$$(1 + x + x^2 + ...)(1 + x^2 + x^4 + ...) \cdot ... \cdot (1 + x^k + x^{2k} + ...) \cdot ...$$

The coefficient of x^n is simply the number of ways to write

$$n = a_1 \cdot 1 + a_2 \cdot 2 + ... + a_n \cdot n,$$

where each a_i is precisely a choice of the a_ith summand in the ith term of the product. Each of theses ways to write n as a sum corresponds uniquely to a partition of n. Thus we have

$$\sum_{n \geq 0} p(n)x^n = \prod_{j \geq 1} \frac{1}{1 - x^j}.$$

Note that here already we could play with these expressions. For instance, limiting the indices of the product, we are basically limiting the possible sizes of the integers we use to build up n. For example:

$$\sum_{n \geq 0} q(n; \leq k)x^n = \prod_{j=1}^{k} \frac{1}{1 - x^j}$$

is the ordinary generating series for the number of ways of writing n as a sum of integers of size at most k, and therefore, also for the number of ways of writing n with at most k parts. Thus

$$\prod_{j=1}^{k} \frac{1}{1 - x^j} - \prod_{j=1}^{k-1} \frac{1}{1 - x^j} = z^k \prod_{j=1}^{k} \frac{1}{1 - x^j}$$

is the generating series for the number of ways of writing n as a sum with precisely k terms (define this number to be $p(n; = k)$.) It follows an easy exercise.

Exercise 1.45. Show that $p(n; = k) = p(n - k; \leq k) = q(n - k; \leq k)$.

Now, finally, let $p_d(n)$ stand for the number of partitions of n with distinct parts, and $p_o(n)$ the number of partitions of n with all parts equal to an odd number . Recall how we constructed the generating function for $p(n)$. Here, we want those a_is to be at most 1. Thus

$$\sum_{n \geq 0} p_d(n)x^n = \prod_{j \geq 1} (1 + x^j).$$

Note that $(1+x^j)(1-x^j) = (1-x^{2j})$. Hence

$$\sum_{n\geq 0} p_d(n)x^n = \prod_{j\geq 1} \frac{(1-x^{2j})}{(1-x^j)} = \prod_{j\geq 1} \frac{1}{1-x^{2j-1}} = \sum_{n\geq 0} p_o(n)x^n.$$

Thus we have $p_d(n) = p_o(n)$, and again we reached a combinatorial equality proved by using power series. It is a very interesting challenge to try to prove this equality by finding a bijection between the set of partitions of n with distinct parts to the set of partitions of n with odd parts (hint: use binary expression of numbers).

Exercise 1.46. Prove that $p(n; \leq 3)$ is the nearest integer to $(n+3)^2/12$. You will have to use the partial fraction decomposition.

Exercise 1.47. Let $P(x)$ be the integer partition generating function. By looking at $P'(x)/P(x)$, show that

$$p(n) = \frac{1}{n}\sum_{i=1}^{n} \sigma_i p(n-i),$$

where σ_i is the sum of the divisors of i. After providing this generating series proof, give a combinatorial proof.

1.16 More variables

We have overlooked this fact across the past pages, but there is nothing preventing you from using more than one variable, and this shall give some advantages when counting certain types of structures.

Example 1.48. We shall start with a simple problem. Into how many ways can you select r integers from 1 to n so that no two are adjacent? Let $a(r,n)$ be the answer to this problem (assume $a(0,0) = 1$). It is not hard to see that

$$a(r,n) = a(r,n-1) + a(r-1,n-2),$$

where the first summand corresponds to the cases where the first chosen integer is larger than 1, and the second to the cases where it is equal to 1. We can now define

$$A(x,y) = \sum_{n,r\geq 0} a(r,n)x^n y^r.$$

From this, it follows that

$$A(x,y) = 1 + x + xy + x(A(x,y)-1) + x^2 yA(x,y),$$

giving

$$A(x,y) = \frac{1+xy}{1-x-x^2y},$$

which shall immediately lead to the answer upon an application of the Binomial Theorem. Now this is definitely not the easiest way to solve this problem (and you are invited to try to find an immediate combinatorial way), but it certainly illustrates the principle.

Example 1.49. The trivial example. Let $b_{n,k} = \binom{n}{k}$. We can write

$$A(x,y) = \sum_{n,k\geq 0} b_{n,k}x^n y^k = \sum_{n\geq 0}\left(\sum_{k=0}^{n}\binom{n}{k}y^k\right)x^n = \sum_{n\geq 0}(1+y)^n x^n = \frac{1}{1-x-xy}.$$

Example 1.50. Recall now the Catalan numbers. They were counting, amongst other things, the number of ways to walk from $(0,0)$ to $(2n,0)$ using steps $(1,1)$ and $(1,-1)$ without ever having a negative coordinate. Imagine now we wish to finish at position (n,k). Let $d_{n,k}$ be the number of such solutions. Clearly $d_{n,k} = 0$ if and only if $n \geq k \geq 0$ and $n+k$ is even. Let

$$D(x,y) = \sum_{n,k\geq 0} d_{n,k}x^n y^k.$$

Clearly $d_{n,n} = 1$ for all n, and $d_{n,k} = d_{n-1,k-1} + d_{n-1,k+1}$. From this, it follows that

$$xyD(x,y) + \frac{x}{y}(D(x,y) - D(x,0)) = D(x,y) - 1.$$

As $D(x,0) = \frac{1}{2x^2}(1-\sqrt{1-4x^2})$, we have

$$D(x,y) = \frac{1-\sqrt{1-4x^2}-2xy}{2x(xy^2+x-y)},$$

from which you can (easily) extract a formula for $d_{n,k}$.

Exercise 1.51. Let $d_{n,k}$ now be the number of paths from $(0,0)$ to $(2n,2n)$ using precisely k steps of type $(1,1)$ made above the line $y=0$, and $n-k$ steps of type $(1,1)$ made below this line. Let

$$P(x,y) = \sum_{n\geq k\geq 0} d_{n,k}x^k y^{n-k}.$$

As usual, let $D(x)$ be the generating series for the Catalan numbers.

(a) Show that

$$P(x,y) = \frac{1}{1-xD(x)-yD(y)}.$$

(b) Using $D(x) = 1 + xD(x)^2$, prove that $d_{n,k}$ is constant for all k, $0 \leq k \leq n$. Conclude that $d_{n,k} = C_n$ for all k.

If you try to find a combinatorial bijective proof of this, you will be tempted to make a simple reflection of the negative part of the path, but this won't work. You will have to be more clever!

Example 1.52. Let us talk about permutations again. Consider the exponential generating series for the numbers m_k of permutation on k elements with just one cycle. As we saw, $m_k = (k-1)!$ (make $m_0 = 0$). But now, we shall use a second variable, whose exponent counts the number of cycles, instead of their size. We have

$$M(x,y) = \sum_{n\geq 1} y\frac{x^n}{n} = -y\log(1-x).$$

Now each permutation is made out of blocks, all of which corresponding to permutations of only one cycle. Thus

$$P(x,y) = M(x,y) + \frac{1}{2!}M(x,y)^2 + \frac{1}{3!}M(x,y)^3 + ... = \exp(M(x,y)),$$

thus

$$P(x,y) = \exp(\log(1-x))^{-y} = (1-x)^{-y}.$$

Note that $s_{n,k} = n![x^n y^k]P(x,y)$ is counting the number of permutation on n elements with precisely k cycles — a number which by itself could be of interest. With this information, we can compute, for instance, the expected number of cycles in a permutation. That is,

$$\sum_{k\geq 0} k\frac{s_{n,k}}{n!} = [x^n]\frac{\mathrm{d}}{\mathrm{d}y}P(x,y)|_{y=1},$$

which will be equal to

$$1 + \frac{1}{2} + ... + \frac{1}{n}.$$

Exercise 1.53. What is the expected number of parts in a random partition of the set $\{1, ..., n\}$?

This last example about permutations just reminded me of a cute puzzle (although it is related to permutations, it has nothing to do with generating series.)

There are 100 prisoners. Their names are placed in 100 wooden boxes, one name in each, and the boxes are lined up in a table in a room. Each prisoner must come alone into the room and their goal is to find their name. To achieve that, each prisoner is allowed to come into the room, look in at most 50 boxes, and leave, without changing anything in the room, and without making any sort of communication with the other prisoners. The rules of the game are simple: if all prisoners find their own names, they all walk free. If at least one of them doesn't, they all die. What should be their strategy? (Believe me, there is a good strategy which guarantees more than 25% chance of survival).

1.17 References

Here is the set of references used to write the past few pages.

For the formalism of power series:

(a) Ivan Niven. Formal power series. *The American Mathematical Monthly*, 76(8):871–889, 1969.

I've used a set of unpublished course notes by Kevin Purbhoo (as far as I know most of those are based on Goulden and Jackson's book.)

(b) Kevin Purbhoo. *Unpublished notes of an enumeration course.* CO630 - University of Waterloo, 2011

(c) Ian P Goulden and David M Jackson. *Combinatorial enumeration.* Courier Corporation, 2004

The application to quicksort analysis and some of the examples and exercises are from Peter Cameron's book.

(d) Peter J Cameron. *Combinatorics: topics, techniques, algorithms.* Cambridge University Press, 1994

For the deeper stuff on partitions, I used Aigner's textbook.

(e) Martin Aigner. *A course in enumeration*, volume 238. Springer Science & Business Media, 2007

van Lint and Wilson's book is very fun to read, and I have extensively consulted their chapter 14 for guidance and to copy some exercises:

(f) J H Van Lint and R M Wilson. *A course in combinatorics*. Cambridge university press, 2001

The book by Wilf seems to be a standard reference, and I used some of his examples.

(g) Herbert S Wilf. *generatingfunctionology*. AK Peters/CRC Press, 2005

2 The adjacency matrix of a graph

In this section, we shall introduce the basic theory of symmetric matrices, including a result generally overlooked in a first or second linear algebra course. We shall define the adjacency matrix of a graph, and then make connections between the algebraic properties of this matrix and the combinatorial properties of the graph.

2.1 Symmetric matrices

We shall work over the vector space $\mathbb{R}^n$. If $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$, then $\langle \mathbf{v}, \mathbf{u} \rangle = \mathbf{v}^T\mathbf{u}$ is an inner product (meaning, it is a positive-definite commutative bilinear form). A linear operator $\mathbf{M} : \mathbb{R}^n \to \mathbb{R}^n$ is self-adjoint if $\langle \mathbf{Mv}, \mathbf{u} \rangle = \langle \mathbf{v}, \mathbf{Mu} \rangle$ for all $\mathbf{u}$ and $\mathbf{v}$, and, because $\mathbf{M}$ can (and will) be seen as a square matrix, it follows that $\mathbf{M}$ is a self-adjoint operator if and only if $\mathbf{M} = \mathbf{M}^T$, that is, $\mathbf{M}$ is a symmetric matrix. Symmetric matrices enjoy two key important properties: they are diagonalizable by orthogonal eigenvectors, and all of their eigenvalues are real. We start proving both properties.

Lemma 2.1. *The eigenvalues of a real symmetric matrix are real numbers.*

Proof. Let $\mathbf{Mu} = \lambda\mathbf{u}$, with $\mathbf{u} \neq \mathbf{0}$. Some of these things could be complex numbers, so we can take the conjugate on both sides, recovering

$$\mathbf{M}\overline{\mathbf{u}} = \overline{\lambda}\overline{\mathbf{u}}.$$

Thus $\overline{\mathbf{u}}$ is an eigenvector with eigenvalue $\overline{\lambda}$. Thus

$$\lambda\mathbf{u}^T\overline{\mathbf{u}} = (\mathbf{Mu})^T\overline{\mathbf{u}} = \mathbf{u}^T(\mathbf{M}\overline{\mathbf{u}}) = \overline{\lambda}\mathbf{u}^T\overline{\mathbf{u}}.$$

Because $\mathbf{u}^T\overline{\mathbf{u}} \neq 0$ if $\mathbf{u} \neq 0$, then $\lambda = \overline{\lambda}$. □

Now simply assume whenever we are dealing with a symmetric matrix, its eigenvalues are real, and any eigenvector can be assumed to be real.

Lemma 2.2. *Let* $\mathbf{M}$ *be a real symmetric matrix, and assume* $\mathbf{u}$ *and* $\mathbf{v}$ *are eigenvectors associated to different eigenvalues. Then* $\mathbf{v}^T\mathbf{u} = 0$*, that is, they are orthogonal.*

Proof. Say $\mathbf{Mu} = \lambda\mathbf{u}$ and $\mathbf{Mv} = \mu\mathbf{v}$, with $\lambda \neq \mu$. It follows that

$$\lambda(\mathbf{v}^T\mathbf{u}) = \mathbf{v}^T\mathbf{Mu} = (\mathbf{v}^T\mathbf{Mu})^T = \mathbf{u}^T\mathbf{M}^T\mathbf{v} = \mathbf{u}^T\mathbf{Mv} = \mu(\mathbf{u}^T\mathbf{v}) = \mu(\mathbf{v}^T\mathbf{u}).$$

As $\lambda \neq \mu$, it must be that $\mathbf{v}^T\mathbf{u} = 0$. □

The lemma above already implies that if $\mathbf{M}$ is diagonalizable, then it is diagonalizable with orthogonal eigenvectors — as, in fact, we eigenvectors corresponding to distinct eigenvalues are orthogonal, and inside each eigenspace we can always find an orthogonal basis. We move forward.

A subspace U of $\mathbb{R}^n$ is said to be $\mathbf{M}$-invariant if, for all $\mathbf{u} \in U$, $\mathbf{Mu} \in U$. This is a key fundamental concept in linear algebra, and several results are proven by noting that certain subspaces are invariant for certain operator.

Lemma 2.3. *Let* $\mathbf{M}$ *be a real symmetric matrix. If* U *is* $\mathbf{M}$*-invariant, then* $U^\perp$ *is also* $\mathbf{M}$*-invariant.*

Proof. Note that $\mathbf{v} \in U^\perp$, by definition, if $\mathbf{v}^T\mathbf{u} = 0$ for all $\mathbf{u} \in U$. For all $\mathbf{u} \in U$ and $\mathbf{v} \in U^\perp$, note that

$$(\mathbf{Mv})^T\mathbf{u} = \mathbf{v}^T\mathbf{Mu} = \mathbf{v}^T(\mathbf{Mu}) = 0,$$

because $\mathbf{u} \in U$, U is $\mathbf{M}$-invariant, and so $\mathbf{Mu} \in U$, and $\mathbf{v} \in U^\perp$. Thus $\mathbf{Mv} \in U^\perp$, as we wanted. □

Let λ be such that $\det(\lambda\mathbf{I} - \mathbf{M}) = 0$. Then $\lambda\mathbf{I} - \mathbf{M}$ is singular, and therefore it contains at least one non-zero vector in its kernel. This is saying that all square matrices $\mathbf{M}$ contain at least one eigenvector for each root of $\phi_{\mathbf{M}}(x) = \det(x\mathbf{I} - \mathbf{M})$. As $\mathbf{M}$ is symmetric, we now know that all possible roots of $\phi_{\mathbf{M}}$ are real.

Lemma 2.4. *Let* U *be an* $\mathbf{M}$*-invariant subspace. Then there is one eigenvector of* $\mathbf{M}$ *in* U*.*

Proof. Let $\mathbf{P}$ be a matrix whose columns form an orthonormal basis for U. As U is $\mathbf{M}$-invariant, it follows that there is a matrix $\mathbf{N}$ so that

$$\mathbf{MP} = \mathbf{PN}.$$

(Stop now and think carefully why this equality is true.) In particular, $\mathbf{N} = \mathbf{P}^T\mathbf{MP}$, so $\mathbf{N}$ is symmetric. Let $\mathbf{u}$ be one eigenvector of $\mathbf{N}$ with eigenvalue λ. Then

$$\mathbf{MPu} = \mathbf{PNu} = \lambda\mathbf{Pu},$$

and, moreover $\mathbf{Pu} =\neq \mathbf{0}$, as the columns of $\mathbf{P}$ are linearly independent. Thus $\mathbf{Pu}$ is an eigenvector for $\mathbf{M}$ in U. □

These four lemmas above are all you need to prove the following result by induction as an exercise.

Theorem 2.5. *Let* $\mathbf{M}$ *be a real symmetric matrix. Then* $\mathbf{M}$ *is diagonalizable by set of orthogonal eigenvectors, all of them corresponding to real eigenvalues.*

Exercise 2.6. Write the proof of this theorem as an exercise.

Corollary 2.7. *Let* $\mathbf{v}_1, ..., \mathbf{v}_n$ *be an orthonormal basis of eigenvectors for* $\mathbf{M}$*, each corresponding to an eigenvalue* $\lambda_1, ..., \lambda_n$ *(these are not necessarily distinct). Let* $\mathbf{P}$ *be the matrix whose* i*th column is* $\mathbf{v}_i$*, and* Λ *the diagonal matrix whose* i*th diagonal element is* λ_i*. Then*

$$\mathbf{P}^T\mathbf{MP} = \mathbf{\Lambda},$$

and

$$\mathbf{M} = \lambda_1(\mathbf{v}_1\mathbf{v}_1^T) + ... + \lambda_n(\mathbf{v}_n\mathbf{v}_n^T).$$

Proof. A linear operator is defined and determined by its action on a basis. The first equality follows from the fact that both sides act equally on the canonical basis of $\mathbb{R}^n$. The second follows from

$$\mathbf{M} = \mathbf{P\Lambda P}^T,$$

and, by definition of matrix product, $\mathbf{M} = \mathbf{v}_1(\lambda_1\mathbf{v}_1^T) + ... + \mathbf{v}_n(\lambda_n\mathbf{v}_n^T)$. □

You should recall right now that, because $\mathbf{v}_i$ is normalized, then $\mathbf{P}_i = \mathbf{v}_i\mathbf{v}_i^T$ is the matrix that represents the orthogonal projection onto the line spanned by $\mathbf{v}_i$, that is, $\mathbf{P}_i$ is a projection as $\mathbf{P}_i^2 = \mathbf{P}_i$, and it is an orthogonal projection as $\mathbf{P}_i$ is symmetric. Note that $\mathbf{P}_i\mathbf{P}_j = \mathbf{0}$ whenever $i \neq j$, and so any sum of the $\mathbf{P}_i$s for distinct indices will correspond to the orthogonal projection onto the space spanned by the $\mathbf{v}_i$s of the same indices. In particular $\sum_{i=1}^{n} \mathbf{P}_i = \mathbf{I}$.

Exercise 2.8. Assume $\mathbf{P}_i$s are orthogonal projections. Show that $\mathbf{P}_1 + \mathbf{P}_2$ is an orthogonal projection if and only if $\mathbf{P}_1\mathbf{P}_2 = \mathbf{0}$.

Show now that $\mathbf{P}_1 + ... + \mathbf{P}_k$ is an orthogonal projection if and only if $\mathbf{P}_i\mathbf{P}_j = \mathbf{0}$ for $i \neq j$.

Say $\mathbf{M}$ is an $n \times n$ symmetric matrix with distinct eigenvalues $\theta_0, ..., \theta_d$. When we write the second equation from the statement of Corollary 2.7, we can collect the terms corresponding to equal eigenvalues, and have

$$\mathbf{M} = \sum_{r=0}^{d} \theta_r \mathbf{E}_r, \tag{1}$$

where, according to the discussion above, each $\mathbf{E}_r$ corresponds to the orthogonal projection onto the θ_r eigenspace. Equation (1) is usually referred to the as the *spectral decomposition* of the matrix $\mathbf{M}$.

Exercise 2.9. Find the spectral decomposition of

$$\mathbf{M} = \begin{pmatrix} 1+\sqrt{2} & 0 & 1-\sqrt{2} & 0 \\ 0 & 1+\sqrt{2} & 0 & 1-\sqrt{2} \\ 1-\sqrt{2} & 0 & 1+\sqrt{2} & 0 \\ 0 & 1-\sqrt{2} & 0 & 1+\sqrt{2} \end{pmatrix}$$

Hint: do not try to compute the characteristic polynomial. It is easier to simply try to look and guess which are the eigenvectors and eigenvalues.

Note that the $\mathbf{E}_r$ are symmetric matrices satisfying $\mathbf{E}_r\mathbf{E}_s = \delta_{rs}\mathbf{E}_r$, and $\sum_{r=0}^{d} \mathbf{E}_r = \mathbf{I}$.

Exercise 2.10. Prove (or at least convince yourself) that for any polynomial $p(x)$, it follows that

$$p(\mathbf{M}) = \sum_{r=0}^{d} p(\theta_r)\mathbf{E}_r.$$

Exercise 2.11. Let $\mathbf{M}$ be a symmetric matrix, with spectral decomposition as in (1).
(A) What is the minimal polynomial of $\mathbf{M}$? (B) Prove that for each $\mathbf{E}_r$, there is a polynomial p_r of degree d so that $p_r(\mathbf{M}) = \mathbf{E}_r$. Describe this polynomial as explicitly as you can.

Exercise 2.12. Prove that two symmetric matrices $\mathbf{M}$ and $\mathbf{N}$ commute if and only if they can be simultaneously diagonalized by the same set of orthonormal eigenvectors. Is it true that if $\mathbf{M}$ and $\mathbf{N}$ commute, then there is always a polynomial p so that $p(\mathbf{M}) = \mathbf{N}$? Characterize what else you need to observe to guarantee that such polynomial exists.

Exercise 2.13. Let $\mathbf{A}$ and $\mathbf{B}$ be matrices (not necessarily squared shaped), so that both products $\mathbf{AB}$ and $\mathbf{BA}$ are defined. Prove that

$$\operatorname{tr} \mathbf{AB} = \operatorname{tr} \mathbf{BA},$$

and conclude that if $\mathbf{M}$ is a symmetric matrix with eigenvalues $\lambda_1, ..., \lambda_n$, then $\operatorname{tr} \mathbf{M}$ is equal to $\lambda_1 + ... + \lambda_n$. How about $\operatorname{tr} \mathbf{M}^2$?

2.2 The adjacency matrix of a graph

Given a graph G on a vertex set V, one can always define an arbitrary ordering to the vertices, that is, let $V = \{a_1, ..., a_n\}$, and encode the graph as a symmetric 01-matrix as follows. *The adjacency matrix* $\mathbf{A}$ of G is defined as $\mathbf{A}_{ij} = 1$ of $a_i \sim a_j$, and $\mathbf{A}_{ij} = 0$ otherwise (including the diagonal elements).

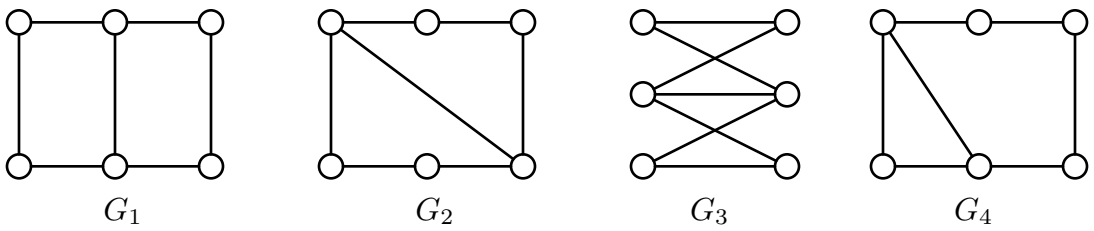
The field of spectral graph theory concerns itself with the main problem of relating spectral properties of matrices that encode adjacency in a graph (such as $\mathbf{A}$) with the combinatorial properties of the graph. We shall see several examples of such relations.

Exercise 2.14. Let G be a graph, suppose the vertices V are ordered, and let $\mathbf{A}$ be the corresponding adjacency matrix of G. Suppose you reorder the vertices by means of a permutation. Let $\mathbf{P}$ be the 01 matrix representing this permutation. Show that the new adjacency matrix obtained from this re-ordering is $\mathbf{PAP}^T$. Conclude that the eigenvalues are the same, and the only change in the eigenvectors is a permutation of its entries.

Because of this exercise, we shall simply ignore the underlying ordering, and speak of "the" adjacency matrix of G.

Given two graphs $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$ on the same number of vertices, a very natural question is whether or not they encode the same combinatorial structure, which can be translated as: is there a function $f : V_1 \to V_2$ that maps edges to edges and non-edges to non-edges? Such a function, if it exists, is called a *graph isomorphism*. You can think of an isomorphism like this: draw both graphs in the plane, and try to move the vertices of one of them (without creating or destroying edges) so that the two drawings look exactly the same.

Example 2.15. Graphs G_1, G_2 and G_3 are all isomorphic, but G_4 is "different".



G_1 G_2 G_3 G_4

Two isomorphic graphs can always be seen as graphs on the same vertex set, and the isomorphism is a re-ordering that preserves adjacency and non-adjacency. Thus:

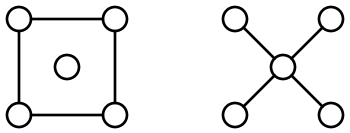
Theorem 2.16. *Let G and H be isomorphic graphs. Order their vertex sets from 1 to n, and let $\mathbf{P}$ be the permutation matrix that corresponds to the isomorphism from G to H. Then*

$$\mathbf{P}\mathbf{A}(G)\mathbf{P}^T = \mathbf{A}(H).$$

As a consequence, $\mathbf{A}(G)$ and $\mathbf{A}(H)$ have the same eigenvalues. □

Exercise 2.17. Order the vertices of G_1 and G_2 equally in terms of their geometric position. Then find the matrix $\mathbf{P}$ so that $\mathbf{P}\mathbf{A}(G_1)\mathbf{P} = \mathbf{A}(G_2)$. Compute the eigenvalues of G_1 and G_4 (using a software?) and conclude that they cannot be isomorphic.

One of the motivations of the development of spectral graph theory was the hope that two graphs would be isomorphic if and only if they had the same eigenvalues. Such a claim would immediately provide an efficient polynomial time algorithm to decide whether two graphs are isomorphic (and yet no such algorithm is known to this day). Two graphs with the same eigenvalues are called *cospectral graphs*. The following pair of graphs are the smallest known cases of cospectral but (clearly) non-isomorphic graphs. They have spectrum $2, 0^{(3)}, -2$.



G_1 G_2

This example also shows that the spectrum of a graph does not determine whether the graph is connected or not. This immediately raises the general question: what graph properties can be determined from the spectrum?

A *walk* of length r in a graph G is a sequence of $r+1$ (possibly repeated) vertices $a_0, ..., a_r$ with the property that $a_i \sim a_j$. A walk is *closed* if $v_0 = v_r$.

Lemma 2.18. *The number of distinct walks of length r from a to b in G is precisely equal to $(\mathbf{A}^r)_{ab}$.*

Exercise 2.19. Verify this result on at least 3 different graphs checking powers $r = 1, 2, 3$ for each. Then, sketch a proof by induction of this result.

Corollary 2.20. *If G has diameter D, then it must have at least $D+1$ distinct eigenvalues.*

Proof. Let

$$\mathbf{A}(G) = \sum_{r=0}^{d} \theta_r \mathbf{E}_r$$

be the spectral decomposition of $\mathbf{A}(G)$. Let W be the subspace of $\mathrm{Sym}_n(\mathbb{R})$ generated by $\{\mathbf{A}^0, \mathbf{A}, \mathbf{A}^2, ...\}$. As we saw in the past section, all powers of $\mathbf{A}$ are a linear combinations of the $\mathbf{E}_r$s, and each $\mathbf{E}_r$ is a polynomial in $\mathbf{A}$. Moreover, the matrices $\mathbf{E}_r$ are pairwise orthogonal, thus they are all linearly independent. As a consequence, $\dim W = d + 1$, and $\{\mathbf{E}_0, ..., \mathbf{E}_d\}$ form a basis for W. Now observe that if $r \leq D$, then at least one entry of $\mathbf{A}^r$ is non-zero "for the first time", meaning that it was equal to 0 for all smaller powers of $\mathbf{A}$. Thus $\{\mathbf{A}^0, \mathbf{A}, \mathbf{A}^2, ..., \mathbf{A}^D\}$ form a linearly independent set in W, and $D \leq d$. □

Let us now return to the problem of deciding what can be determined by the spectrum of a graph alone. Clearly the number of vertices in a graph is determined by the spectrum. An immediate consequence of the Lemma 2.18 is that the number of edges is also determined by the spectrum.

Corollary 2.21. *Let G be a graph on n vertices, with m edges, and let $\lambda_1, ..., \lambda_n$ the eigenvalues of $\mathbf{A}(G)$. Then*

$$\lambda_1^2 + ... + \lambda_n^2 = 2m.$$

Proof. Both sides are equal to $\operatorname{tr} \mathbf{A}^2$. □

Exercise 2.22. Find a formula for the number of triangles (cycles of length 3) found as subgraphs of G that depends only on the eigenvalues of G. Explain why the number of cycles of length 4 is not determined by the spectrum alone (as you witnessed in the example above).

Exercise 2.23. Does the spectrum alone determines the length of the shortest odd cycle of a graph? Explain.

Exercise 2.24. If G has n vertices, prove that all eigenvalues of lie in the interval $(-n, n)$.

Exercise 2.25. Let G be a k-regular graph (that is, all eigenvalues have k neighbours). Prove that k is an eigenvalue for G by describing a corresponding eigenvector.

Let $\mathbf{J}$ stand for the matrix whose all entries are equal to 1. If G is a graph, let $\overline{G}$ stand for the complement graph of G, that is, the graph whose edges are precisely the non-edges of G. Then, clearly,

$$\mathbf{A}(\overline{G}) = \mathbf{J} - \mathbf{A}(G) - \mathbf{I}.$$

As immediate consequence of the past exercise, we have:

Lemma 2.26. *Let G be a k-regular graph, with eigenvalues $k = \lambda_1, ..., \lambda_n$. Then the eigenvalues of $\overline{G}$ are*

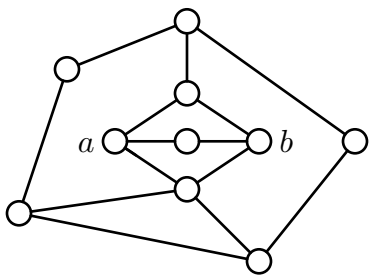
$$n - k - 1,\ -\lambda_2 - 1,\ ...,\ -\lambda_n - 1.$$

Proof. The all 1s vector $\mathbf{1}$ is an eigenvector of G. Let $\mathbf{v}_2, ..., \mathbf{v}_n$ complete a basis of orthogonal eigenvectors. Then

$$(\mathbf{J} - \mathbf{A}(G) - \mathbf{I})\mathbf{1} = (n - k - 1)\mathbf{1} \quad \text{and} \quad (\mathbf{J} - \mathbf{A}(G) - \mathbf{I})\mathbf{v}_i = -\lambda_i - 1,$$

as $\mathbf{J}\mathbf{v}_i = \mathbf{0}$ because $\mathbf{1}$ and $\mathbf{v}_i$ are orthogonal. □

Exercise 2.27. Assume G contains a pair of vertices a and b so that the neighbourhood of a is equal to neighbourhood of b (the rest of the graph can be anything). For example:



(a) Prove that 0 is an eigenvalue of this graph (Hint: look at a and b and try to produce one eigenvector for 0). If the example looks too complicate, forget about the 5-cycle and focus only on a, b and their neighbours.

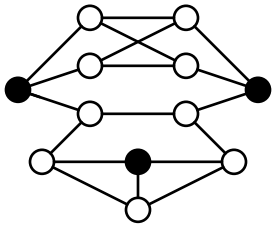
(b) What could you say if a and b shared the same neighbourhood, but were also neighbours themselves?

Exercise 2.28. Assume $G = (V, E)$ is a k-regular graph which contains a subset of vertices $U \subseteq V$ satisfying the following properties:

(a) No two vertices in U are neighbours.

(b) Any vertex in $V \backslash U$ contains exactly one neighbour in U.

Prove that if such U exists, then -1 is an eigenvalue of the graph. (Hint: recall G is assumed to be k-regular, and, again, try to produce one eigenvector. Try first in the example below, where the dark vertices are the vertices in U.)



In this next section, we shall see that two important properties about a graph can be determined from its spectrum alone: whether the graph is regular, and whether the graph is bipartite.

2.3 Perron-Frobenius (a special case)

Let $\mathbf{M}$ be a real $n \times n$ matrix with nonnegative entries. For example, the adjacency matrix of a graph. This matrix is called primitive if, for some integer k, $\mathbf{M}^k > 0$, and it is called irreducible if for all indices i and j, there is an integer k so that $(\mathbf{M}^k)_{ij} > 0$. All primitive matrices are irreducible, but the converse is not necessarily true.

Example 2.29. Consider

$$\begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \quad , \quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad , \quad \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Verify that the first is primitive, the second and third are both irreducible, but not primitive, and the fourth is neither.

Exercise 2.30. Prove that if $\mathbf{M}$ is irreducible, then $\mathbf{I} + \mathbf{M}$ is primitive.

Exercise 2.31. Let G be a graph. Show that

(a) $\mathbf{A}(G)$ is irreducible if and only if G is connected.

(b) $\mathbf{A}(G)$ is not primitive if G is bipartite.

Over the next few results, we shall actually see, amongst other things, that $\mathbf{A}(G)$ is irreducible but not primitive if and only if G is connected and bipartite. Results below are known as the Perron-Frobenius theory. This theory applies generally to matrices which are assumed to be irreducible and nothing else. We shall however add the hypothesis that the matrices are also symmetric, for the proofs become simpler and more meaningful, and our matrices will almost always be symmetric anyway.

Our first observation.

Lemma 2.32. *Let* $\mathbf{M}$ *be a nonnegative symmetric matrix,* $\mathbf{M} \neq \boldsymbol{0}$*. If* λ *is the largest eigenvalue of* $\mathbf{M}$*, then* $\lambda > 0$*.*

Proof. Follows immediately from $\operatorname{tr} \mathbf{M} \geq 0$. □

For any vector $\mathbf{u} \in \mathbb{R}^n$, and symmetric matrix $\mathbf{M}$, define

$$R_{\mathbf{M}}(\mathbf{u}) = \frac{\mathbf{u}^T\mathbf{M}\mathbf{u}}{\mathbf{u}^T\mathbf{u}}.$$

This is known as the Rayleigh quotient of $\mathbf{u}$ with respect to $\mathbf{M}$. Note that $R_{\mathbf{M}}(\alpha\mathbf{u}) = R_{\mathbf{M}}(\mathbf{u})$ for all $\alpha \neq 0$, so we shall typically assume $\mathbf{u}$ has been normalized. In a sense, this is a measurement of how much $\mathbf{M}$ displaces $\mathbf{u}$, also proportional to how much $\mathbf{M}$ stretches or shrinks $\mathbf{u}$. Therefore one should expect that this is maximum when $\mathbf{u}$ is an eigenvector of $\mathbf{M}$, corresponding to a large eigenvalue.

Lemma 2.33. *If* $\mathbf{u}$ *is eigenvector of* $\mathbf{M}$ *with eigenvalue* θ*, then* $R_{\mathbf{M}}(\mathbf{u}) = \theta$*. If* λ *is the largest eigenvalue of* $\mathbf{M}$*, then, for all* $\mathbf{v} \in \mathbb{R}^n$*,* $R_{\mathbf{M}}(\mathbf{v}) \leq \lambda$*. Equality holds for some* $\mathbf{v}$ *only if* $\mathbf{v}$ *is eigenvector for* λ*.*

Proof. Only the second and third assertions deserve a proof. Let $\mathbf{M} = \sum_{r=0}^{d} \theta_r \mathbf{E}_r$ be the spectral decomposition of $\mathbf{M}$. Assume λ_0 is the largest eigenvalue, and that $\mathbf{v}$ is a normalized vector. Then

$$\begin{aligned} R_{\mathbf{M}}(\mathbf{v}) = \mathbf{v}^T\mathbf{M}\mathbf{v} &= \theta_0(\mathbf{v}^T E_0\mathbf{v}) + \theta_1(\mathbf{v}^T E_1\mathbf{v}) + ... + \theta_d(\mathbf{v}^T E_d\mathbf{v}) \\ &\leq \theta_0((\mathbf{v}^T E_0\mathbf{v}) + (\mathbf{v}^T E_1\mathbf{v}) + ... + (\mathbf{v}^T E_d\mathbf{v}) = \theta_0. \end{aligned}$$

Equality holds if and only if $(\mathbf{v}^T E_r\mathbf{v}) = 0$ for all $r > 0$, which is the same as saying that $\mathbf{v}$ belongs to the θ_0 eigenspace. □

Lemma 2.34. *Let* $\mathbf{M}$ *be symmetric, non-negative and irreducible, with largest eigenvalue* λ*. There is a corresponding eigenvector* $\mathbf{u}$ *to* λ *so that* $\mathbf{u} > \boldsymbol{0}$*.*

Proof. Let $\mathbf{v}$ be a normal eigenvector for λ, and define $\mathbf{u}$ to be made from $\mathbf{v}$ by taking the absolute value at each entry (also denoted by $\mathbf{u} = |\mathbf{v}|$). Note that $\mathbf{u}$ is still normal, and, moreover

$$\lambda = R_{\mathbf{M}}(\mathbf{v}) = |R_{\mathbf{M}}(\mathbf{v})| \leq R_{\mathbf{M}}(\mathbf{u}) \leq \lambda.$$

(Second equality follows from $\lambda > 0$. First inequality from is simply the triangle inequality. Second follows from Lemma 2.33.)

Hence $R_{\mathbf{M}}(\mathbf{u}) = \lambda$, and $\mathbf{u}$ is an eigenvector for λ, with $\mathbf{u} \geq \mathbf{0}$. To see that $\mathbf{u} > \mathbf{0}$, note that as $\mathbf{M}$ is irreducible, it follows from Exercise 2.30 that $\mathbf{I} + \mathbf{M}$ is primitive, and so there is a k so that $(\mathbf{I}+\mathbf{M})^k > \mathbf{0}$. The vector $\mathbf{u}$ is also eigenvector for this matrix (with eigenvalue $(1 + \lambda)^k$, but

$$\mathbf{0} < (\mathbf{I} + \mathbf{M})^k \mathbf{u} = (1 + \lambda)^k \mathbf{u},$$

implying $\mathbf{u} > \mathbf{0}$. □

Lemma 2.35. *The largest eigenvalue λ of a symmetric, non-negative and irreducible matrix is simple.*

Proof. From the proof of the past lemma, we know that no eigenvector for λ contains an entry equal to 0. No subspace of dimension larger than 1 can be such that all of its non-zero vectors have no non-zero entries. □

And finally:

Lemma 2.36. *Let $\mathbf{M}$ be symmetric, non-negative and irreducible. Let λ be its largest eigenvalue. Let μ be any other eigenvalue. Then $\lambda \geq |\mu|$, and, moreover, if $-\lambda$ is an eigenvalue, then $\mathbf{M}^2$ is not irreducible.*

Proof. Let $\mathbf{v}$ be an eigenvector for μ. As $\mathbf{v}$ is orthogonal to the positive eigenvector corresponding to λ, at least one entry of $\mathbf{v}$ is negative. Thus

$$|\mu| = |R_{\mathbf{M}}(\mathbf{v})| < R_{\mathbf{M}}(|\mathbf{v}|) \leq \lambda.$$

Now note that λ^2 is the largest eigenvalue of $\mathbf{M}^2$ (which is, still, symmetric and non-negative). If $-\lambda$ is eigenvalue of $\mathbf{M}$, then the eigenspace of λ^2 in $\mathbf{M}^2$ is at least 2-dimensional, thus $\mathbf{M}^2$ cannot be irreducible. □

It is quite surprising at first sight that the hypothesis on $\mathbf{M}$ being symmetric can be dropped entirely from the results above. The geometric intuition remains the same: a nonnegative irreducible matrix acts in the nonnegative orthant and there it encounters a unique direction which is an eigenvector. The proofs of these results are not hard per se, but I didn't feel they would add much to this notes. You are however invited to check any reference on spectral graph theory or non-negative matrix theory to find your favourite version of these results.

Now, to the applications.

Theorem 2.37. *Let $\mathbf{A}$ be the adjacency matrix of a connected graph G, and $\lambda_1 \geq ... \geq \lambda_n$ its spectrum.*

(a) G is k-regular if and only if $(1/n)(\lambda_1^2 + ... + \lambda_n^2) = \lambda_1$, and, in this case, $k = \lambda_1$.

(b) G is bipartite if and only if $\lambda_1 = -\lambda_n$. If this is the case, then for all λ_i, $-\lambda_i$ is also an eigenvalue.

Proof.

(a) Let $\mathbf{1}$ be the all 1s vector. The equality is equivalent to

$$R_{\mathbf{A}}(\mathbf{1}) = \lambda_1,$$

which, as we saw, is equivalent to $\mathbf{1}$ being an eigenvector of λ_1. This vector is eigenvector if and only if all row sums of $\mathbf{A}$ are equal, or, equivalently, all vertices have the same degree, which is going to be precisely equal to the eigenvalue λ_1.

(b) If G is bipartite, its adjacency matrix can always be written as

$$\mathbf{A} = \begin{pmatrix} \mathbf{0} & \mathbf{B} \\ \mathbf{B}^T & \mathbf{0} \end{pmatrix}.$$

If $\begin{pmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \end{pmatrix}$ is eigenvector for λ_i, then it is easy to see that $\begin{pmatrix} \mathbf{v}_1 \\ -\mathbf{v}_2 \end{pmatrix}$ is eigenvector for $-\lambda_i$. On the other hand, if $-\lambda_1$ is eigenvalue, then, from Lemma 2.36, it follows that $\mathbf{A}^2$ is not irreducible. Thus there are at least two vertices you can never walk from one to another with an even number of steps. Therefore there can be no odd cycles in this graph.

□

Corollary 2.38. *Let λ be the largest eigenvalue of $\mathbf{A}(G)$. Let Δ be the largest degree of G, and let ∂ be its average degree. Then*

$$\partial \leq \lambda \leq \Delta.$$

Proof. The first inequality follows from the fact that

$$\partial = R_{\mathbf{A}}(\mathbf{1}) \leq \lambda.$$

(Note in particular that this implies $\lambda \geq \delta$, where δ is the smallest degree of G). For the second, we have $\mathbf{A1} \leq \Delta\mathbf{1}$, and with $\mathbf{v}$ eigenvector for λ, we can multiply by $\mathbf{v}^T$ on the left. As $\mathbf{v} > \mathbf{0}$, the sign is preserved, and

$$\lambda \mathbf{v}^T\mathbf{1} = \mathbf{v}^T\mathbf{A1} \leq \Delta\mathbf{v}^T\mathbf{1},$$

so $\theta \leq \Delta$. □

Exercise 2.39. Prove that $\lambda \geq \sqrt{\Delta}$. (Hint: look at $\mathbf{A}^2$ and the proof above).

2.4 Eigenvalues of some classes of graphs

Consider the following classes of graphs:

(a) K_n - complete graphs on n vertices.

(b) $K_{n,m}$ - complete bipartite graphs with n vertices on one side, and m vertices on the other (in particular if $n = 1$, these are the stars).

(c) C_n - cycle graphs on n vertices.

(d) P_n - path graphs on n vertices.

Our goal here is to determine the eigenvalues (and eigenvectors) of these classes.

(a) This is easy. $\mathbf{A}(K_n) = \mathbf{J} - \mathbf{I}$. The eigenvalues of $\mathbf{J}$ are n (simple, with eigenvector $\mathbf{1}$) and 0 (all others). Thus the spectrum of K_n is $n-1$ and -1.

(b) Write

$$\mathbf{A} = \begin{pmatrix} \mathbf{0} & \mathbf{J}_{a,b} \\ \mathbf{J}_{b,a} & \mathbf{0} \end{pmatrix}.$$

There are $b-1$ vectors in the kernel of $\mathbf{J}_{a,b}$ and $a-1$ vectors in the kernel of $\mathbf{J}_{b,a}$. Each corresponding to an eigenvector for the eigenvalue 0 of $\mathbf{A}$. The two eigenvectors remaining are

$$\begin{pmatrix} \sqrt{b}\mathbf{1} \\ \sqrt{a}\mathbf{1} \end{pmatrix} \text{ and } \begin{pmatrix} \sqrt{b}\mathbf{1} \\ -\sqrt{a}\mathbf{1} \end{pmatrix},$$

corresponding to the eigenvalues $\sqrt{ab}$ and $-\sqrt{ab}$ respectively.

(c) This one is trickier. $\mathbf{A}(C_n)$ is the sum of two permutation matrices corresponding to the cycle $(123...n)$ and its inverse, say $\mathbf{P}$ and $\mathbf{P}^{-1}$. An eigenvector for a cyclic matrix can be easily built from an n-root of unity ω:

$$\mathbf{P}\begin{pmatrix} 1 \\ \omega \\ \vdots \\ \omega^{n-1} \end{pmatrix} = \begin{pmatrix} \omega^{n-1} \\ 1 \\ \vdots \\ \omega^{n-2} \end{pmatrix} = \omega^{n-1}\begin{pmatrix} 1 \\ \omega \\ \vdots \\ \omega^{n-1} \end{pmatrix} \text{ and } \mathbf{P}^{-1}\begin{pmatrix} 1 \\ \omega \\ \vdots \\ \omega^{n-1} \end{pmatrix} = \begin{pmatrix} \omega \\ \omega^2 \\ \vdots \\ 1 \end{pmatrix} = \omega\begin{pmatrix} 1 \\ \omega \\ \vdots \\ \omega^{n-1} \end{pmatrix},$$

thus the eigenvalues are $\omega^{n-1} = \omega^{-1}$ and ω, hence the eigenvalues of $\mathbf{A}(C_n) = \mathbf{P} + \mathbf{P}^{-1}$ are $\omega^{-1} + \omega$ for all nth roots of unity, that is, $\omega = \mathrm{e}^{2\pi \mathrm{i}(k/n)}$, $k = 0, ..., n-1$. Thus the eigenvalues of C_n are

$$2\cos\left(2\pi\frac{k}{n}\right) \quad \text{for } k = 0, ..., n-1.$$

Note that 2 is always the largest (and simple) eigenvalue, and that -2 is an eigenvalue if and only if n is even. All other eigenvalues have multiplicity 2.

(d) We provide one way of finding this now. The other will come later as an exercise. Look at the cycle C_{2n+2}. Let ω be a $(2n+2)$th root of unity. Then

$$\begin{pmatrix} 1 \\ \omega \\ \vdots \\ \omega^{2n+1} \end{pmatrix} \text{ and } \begin{pmatrix} 1 \\ \omega^{-1} \\ \vdots \\ \omega^{-(2n+1)} \end{pmatrix}$$

are both eigenvalues of $\mathbf{A}(C_{2n+2})$ for $\omega+\omega^{-1}$, and so is any linear combination of them. In particular

$$\begin{pmatrix} 1 \\ \omega \\ \vdots \\ \omega^{2n+1} \end{pmatrix} - \begin{pmatrix} 1 \\ \omega^{-1} \\ \vdots \\ \omega^{-(2n+1)} \end{pmatrix} = \begin{pmatrix} 0 \\ \omega-\omega^{-1} \\ \vdots \\ \omega^{2n+1}+\omega^{-2n-1} \end{pmatrix}.$$

Note that there will be another 0 at position $n+2$, corresponding to $\omega^{n+1}-\omega^{-n-1} = -1-(-1)=0$. The n non-zero entries (only when $\omega \neq 1$) from positions 2 to $n+1$ are part of an eigenvector of C_{2n+2} which do not get interfered by the rest of the graph (those 0s at positions 1 and $n+2$ "disconnect" the eigenvector). Hence this part of the eigenvector is also an eigenvector for P_n (subgraph of C_{2n+2} from positions 2 to $n+1$). Therefore the spectrum of $\mathbf{A}(P_n)$ is

$$\omega+\omega^{-1} = 2\cos\left(\pi\frac{k}{n+1}\right) \quad \text{for } k=1,...,n.$$

2.5 Strongly regular graphs

A graph G on n vertices, not equal to K_n, is called "strongly-regular" if it satisfies the properties

(a) G is k-regular, for some k.

(b) Any two neighbours of G share precisely a common neighbours.

(c) Any two non-neighbours of G share precisely c common neighbours.

Exercise 2.40. What is the diameter of G?

Exercise 2.41. Let $\mathbf{A}=\mathbf{A}(G)$. Explain why

$$\mathbf{A}^2 = k\mathbf{I}+a\mathbf{A}+c(\mathbf{J}-\mathbf{I}-\mathbf{A}).$$

Exercise 2.42. Prove that $\mathbf{A}^3$ can be written as a polynomial of degree at most 2 computed at $\mathbf{A}$. Conclude that $\mathbf{A}$ has 3 distinct eigenvalues, and find a formula for these eigenvalues in terms of k, a and c (recall that k must be one of them).

Exercise 2.43. Find a formula for n that depends uniquely on k, a and c.

Exercise 2.44. Find a formula for the multiplicities of the three eigenvalues of $\mathbf{A}$.

Exercise 2.45. Prove that there is no strongly regular graph with $a=c=1$ (Hint: the multiplicities you found above must be integers!!)

2.6 Graph isomorphism

Perhaps one of the nicest and most relevant applications of basic spectral graph theory is a polynomial-time algorithm to decide whether two graphs with only simple eigenvalues are isomorphic or not. At first one wonders if graphs have usually simple eigenvalues or not, and the answer is yes! This is no trivial result though, and was only settled in 2014 by Terence Tao and Van Vu. The consequence is that Graph Isomorphism is in P for almost all graphs. In this section, we will see how to construct such an algorithm.

Now we shall assume throughout this section that all graphs being treated have simple eigenvalues, that is, the multiplicity of all eigenvalues is equal to 1. Our typical notation will be that a symmetric matrix $\mathbf{A}$ is diagonalized as $\mathbf{A} = \mathbf{PDP}^T$.

Lemma 2.46. *Let* $\mathbf{A}$ *and* $\mathbf{B}$ *be symmetric matrices with the same simple eigenvalues, with corresponding diagonalizations*

$$\mathbf{A} = \mathbf{UDU}^T \quad \textit{and} \quad \mathbf{B} = \mathbf{VDV}^T.$$

There is a permutation matrix $\mathbf{P}$ *so that* $\mathbf{PAP}^T = \mathbf{B}$ *if and only if there is a diagonal matrix* $\mathbf{E}$*, whose entries are* ± 1*, so that* $\mathbf{PU} = \mathbf{VE}$*.*

Before continuing, recall that $\mathbf{U}^T = \mathbf{U}^{-1}$, $\mathbf{V}^T = \mathbf{V}^{-1}$ and $\mathbf{P}^T = \mathbf{P}^{-1}$, because all these matrices are orthogonal matrices.

Proof. We have $\mathbf{PAP}^T = \mathbf{B}$ if and only if

$$\mathbf{P}(\mathbf{UDU}^T)\mathbf{P}^T = \mathbf{VDV}^T, \text{ or equivalently } \mathbf{V}^T\mathbf{PUD} = \mathbf{DV}^T\mathbf{PU}.$$

Let $\mathbf{E} = \mathbf{V}^T\mathbf{PU}$. Because all entries of $\mathbf{D}$ are distinct, it is enlightening to verify that $\mathbf{E}$ must be diagonal. Not only that, $\mathbf{E}^2 = 1$, so $\mathbf{E}$ contains only ± 1s. The other direction is immediate. □

This is already enough to tell us something quite strong. Recall that an automorphism of G is a permutation of $V(G)$ that preserves adjacency and non-adjacency.

Theorem 2.47. *If* G *is a graph and* $\mathbf{A}(G)$ *has simple eigenvalues, then any automorphism of* G *has order* 2*.*

Proof. Let $\mathbf{P}$ be the permutation matrix representing the automorphism. Thus $\mathbf{PAP}^T = \mathring{A}$, and by the corollary above, it follows that there is a ± 1 diagonal matrix $\mathbf{E}$ so that

$$\mathbf{PU} = \mathbf{UE}.$$

Hence $\mathbf{P}^2 = (\mathbf{UEU}^T)^2 = \mathbf{I}$. □

Combinatorially, this is saying that every automorphism of a graph with simple eigenvalues is splitting the vertices into some being fixed and some being swapped. Whenever you find a graph with a different type of automorphism, you already know now that at least one of its eigenvalues is not simple.

Exercise 2.48. Prove that if $\mathbf{P}$ and $\mathbf{Q}$ represent automorphisms of a graph with simple eigenvalues, then $\mathbf{PQ} = \mathbf{QP}$.

We return to the main problem of this section which is to determine, given two graphs G and H with adjacency matrices $\mathbf{A}$ and $\mathbf{B}$ having both with the same simple eigenvalues, whether there is a permutation matrix $\mathbf{P}$ so that $\mathbf{PAP}^T = \mathbf{B}$. Henceforth, assume

$$\mathbf{A} = \mathbf{UDU}^T \quad \text{and} \quad \mathbf{B} = \mathbf{VDV}^T.$$

Again, due to Lemma 2.46, the existence of such $\mathbf{P}$ is equivalent to determining whether there is a ± 1 diagonal matrix $\mathbf{E}$ so that $\mathbf{U}$ and $\mathbf{VE}$ have the same rows (they shall appear in different order, but adjusting this ordering is precisely what a candidate $\mathbf{P}$ does when multiplying $\mathbf{U}$ from the left). The isomorphism problem hence becomes that of determining whether $\mathbf{U}$ and $\mathbf{VE}$ have the same rows. We can hence permute the rows of each freely.

Exercise 2.49. Show that if anyone of the eigenvectors of $\mathbf{A}$ has entries with distinct absolute values, the problem becomes very easy.

Here is how we shall decide whether such $\mathbf{E}$ exists. The rows of $\mathbf{U}$ and $\mathbf{V}$ are indexed by a set V (of size n). You could think of V as the labels of the vertices of both graphs. Each partition of V corresponds to a partition of the row set of these matrices. We will try to find partitions of V satisfying certain special properties. If we succeed, then it will be possible to efficiently solve for $\mathbf{E}$.

First, a definition. If $\mathbf{u}$ and $\mathbf{v}$ are vectors in $\mathbb{R}^n$, let $\mathbf{u} \circ \mathbf{v}$ denote the entry-wise product of these vectors, that is, the vector whose entries are obtained by multiplying the corresponding entries of $\mathbf{u}$ and $\mathbf{v}$.

If $C \subseteq V$, let $\mathbf{U}(C)$ denote the submatrix of $\mathbf{U}$ which contains only the rows indexed by C. Same for $\mathbf{V}(C)$. For any matrix $\mathbf{M}$ with n columns and $k \in [n] = \{1, ..., n\}$, let $\mathbf{M}_k$ be the k-th column of $\mathbf{M}$.

We now describe a method to decide whether $\mathbf{E}$ exists.

(a) First, partition V so that for each class C of the partition, the entries of $\mathbf{U}(C)_1$ have the same absolute value. Refine this partition according to $\mathbf{U}(C)_2$. Repeat until $\mathbf{U}(C)_n$. This will be the coarsest partition which, for any of its classes and any of the columns of $\mathbf{U}$, the absolute values of the entries of the column corresponding to the class is constant. Each class of the partition determines a row vector containing the absolute values of each column.

(b) Repeat the procedure for $\mathbf{V}$ (creating another partition of course). Now compare these rows vectors determined by each class of both partitions. Either there is one row vector amongst the classes of $\mathbf{U}$ but not amongst those of $\mathbf{V}$ — in which case $\mathbf{E}$ cannot exist; or we move forward.

(c) Now, refine the partition in $\mathbf{U}$ so that each column in each class has either only positive entries only, or negative entries only, or 0 entries only, or non-zero entries and the same number of positive and negative entries.

(d) Look now to the parts which had columns with mixed signs. Consider all products of the form $\mathbf{U}(C)_i \circ \mathbf{U}(C)_j$. If the number of + and - is distinct, refine the partition so that they become equal on both new parts.

(e) Repeat this for all subsets $S = \{s_1, ..., s_k\} \subseteq [n]$ so that for all parts C of the partition, the products

$$\mathbf{U}(C)_{s_1} \circ ... \circ \mathbf{U}(C)_{s_k}$$

contain entries so that either all entries are positive, or all entries are negative, or all entries are 0, or they are non-zero and the number of positive entries is equal to the number of negative entries.

(f) If all columns in a part have the same sign or are 0, let's say this is a column of type 1. If they display k distinct sign patterns, then let us say they are of type k. The key observation now is that there cannot be two parts of the same type with the same size, otherwise their union would have been a part that would have not been partitioned.

(g) Upon performing the same procedure in $\mathbf{V}$, we can now match parts of each partition. Two matched parts of type 1 determine the only possible sign choice for $\mathbf{E}$ (and thus a unique candidate permutation $\mathbf{P} = \mathbf{VEU}^T$). If there are no parts of type 1, check those of type 2. Two matched parts of type 2 determine two possible choices. We need only check all possibilities any given parts yields. And so on for types of larger index. The largest possible k so that a part is of type k is $\log n$, thus there are at most n choices to be checked, in the worst case.

2.7 References

Here is the set of references used to write the past few pages.

I used Chapter 8 of Godsil and Royle to write about the spectral decomposition of a symmetric matrix. This was also my reference for the basics and some exercises on the adjacency matrix, and also for strongly regular graphs.

(a) Chris Godsil and Gordon Royle. *Algebraic Graph Theory*. Springer-Verlag, New York, 2001.

Exercise 2.28 comes from Chan and Godsil "Symmetry and Eigenvectors".

I looked extensively for a nice intuitive proof of Perron-Frobenius in its full form, but the best I could do relied on using fixed point theorems. I then came up with the simplified version assuming matrices in question are symmetric. A good reference is Brouwer and Haemers, Chapter 2.

(b) Andries E Brouwer and Willem H Haemers. *Spectra of Graphs*. Universitext. Springer, New York, 2012

I also used the reference above for the spectrum of paths and cycles.

It is surprisingly hard to find a good reference for graph isomorphism (but this is no excuse for the poor job I made in describing the algorithm). The published paper by Babai, Grigoryev and Mount proves a stronger result, but relies on more group theory that I wanted to use. Cvetkovic, Rowlinson and Simic (Eigenspaces of Graphs) develop an interesting machinery to deal with the problem, but it also seemed too much for one lecture only. A manuscript (litereally) of the original result by Leighton and Miller is available at Miller's website, and despite its poor quality, it was probably the best source I could find (if you are able to decipher it all, please let me know). I should also refer to Spielman's 2018 lecture on the topic (available at his website), though he focus on the related problem of determining the automorphism group of the graph.

3 Graph polynomials

Significant part of the algebraic graph theory of graphs revolves around studying polynomials whose definition is based on the graph. Coefficients or evaluations of such polynomials typically count things associated to the graph, but algebraic properties of them and of their roots also tend to bring interesting considerations about the graph.

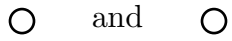
One motivation to define polynomials for graphs is the hope that a given polynomial would be efficiently computable and at the same time completely identify the graph up to isomorphism. No such polynomial is known in general (otherwise graph isomorphism would be an easier problem). Another motivation possibly come (historically as well) from the famous Reconstruction Conjecture. We start our section with a brief introduction to this conjecture.

3.1 Reconstruction — an interlude

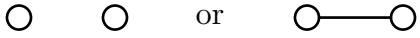
Given a graph G on n vertices, the set of n subgraphs obtained from G upon deleting each one of its vertices is called the *deck of* G. If G and its deck are presented with labelled vertices, then there is not much to ask or wonder. A completely more interesting question rises with one simply erases (or arbitrarily mixes up) the labels — we shall hence assume all graphs in this section are of such form.

Conjecture 1 (Kelly-Ulam)**.** *For any graph G on $n > 2$ vertices, G is completely determined by its deck.*

The hypothesis on $n > 2$ is necessary because the two subgraphs



could have been obtained from either of the following graphs,



but these seem to be only known case of such phenomenon. Several graph theorists have worked on this conjecture for the past decades, and yet a complete answer seems to be far from being found. Partial results usually have two flavours: either one determines that graphs belonging to a certain class are reconstructible (from its deck), or one determines which properties or invariants of a graph are reconstructible. For the remainder of this section, we will mostly focus on the second type of question. But in this brief interlude, we prove the following results.

Let $\nu(H, G)$ denote the number of subgraphs of G isomorphic to H. It is not surprising that this parameter is reconstructible.

Lemma 3.1 (Kelly)**.** *For any graphs G and H,*

$$(|V(G)| - |V(H)|)\ \nu(H, G) = \sum_{a \in V(G)} \nu(H, G \backslash a)$$

Proof. The result is trivial if $|V(H)| \geq |V(G)|$. Assume otherwise. We shall count the number of pairs (H', a) where H' is a copy of H in G, $a \in V(G)$ but $a \notin V(H')$. By choosing H' first, there are $(|V(G)| - |V(H)|)\ \nu(H, G)$ such pairs. By choosing a first, the number of copies of H not using a is precisely $\nu(H, G\backslash a)$. The result thus follows. □

Corollary 3.2. *If G has more than two vertices, the parameter $|E(G)|$ is reconstructible from the deck of G.*

Corollary 3.3. *The degree sequence of G (that is, the sequence of numbers listing the degrees of the vertices of G) is reconstructible.*

Exercise 3.4. Using Kelly's lemma, prove both corollaries above.

Theorem 3.5. *If G is a regular graph on more than 2 vertices, then G is reconstructible.*

Proof. From the degree sequence, decide whether G is regular. If it is, examine any of the graphs in its deck, and add a missing vertex so that it becomes regular. This graph will be equal to G. □

3.2 Walks

For any graph G, define $\phi_G(x)$ to be

$$\phi_G(x) = \det(x\mathbf{I} - \mathbf{A}).$$

The characteristic polynomial of a graph and of its subgraphs interplay nicely with walk counts and eigenvectors of the graph. Over the next few results, we shall make this relationship clearer.

Lemma 3.6. *If G is disconnected, and G_1 and G_2 are disjoint subgraphs of G with $G_1 \cup G_2 = G$, then*

$$\phi_G = \phi_{G_1} \cdot \phi_{G_2}.$$

This above is immediate from the block expansion of a determinant.

We now write a generating function whose coefficients are matrices:

$$W_G(x) = \sum_{k \geq 0} (\mathbf{A})^k x^k.$$

This is known as the walk generating function of G — the ij entry of the coefficient multiplying x^k counts the number of walks of length k from i to j. Rules for formal power series apply (existence of multiplicative inverses, substitutions, Laurent power series, etc.), and so we have

$$W_G(x) = \frac{1}{(\mathbf{I} - x\mathbf{A})}.$$

Notice that we are working with matrices whose coefficients are over $\mathbb{R}((x))$, but that shall mean no harm. In fact, properties about the determinant that you can prove exploring its

Laplace expansion still hold true, in particular, for any $\mathbf{M}$ matrix with coefficients which are power series in x,

$$\mathbf{M} \cdot \operatorname{adj}(\mathbf{M}) = \det(\mathbf{M})\mathbf{I}. \tag{2}$$

Recall now that $\operatorname{adj}(\mathbf{M})$ is the matrix defined as

$$(\operatorname{adj} \mathbf{M})_{ij} = (-1)^{i+j} \det \mathbf{M}[j,i],$$

where $\mathbf{M}[j,i]$ stands for the matrix $\mathbf{M}$ removed of row j and column i.

Specifically, we are interested in what happens when $\mathbf{M} = (\mathbf{I} - x\mathbf{A})$. Equation (2) becomes

$$W_G(x) = \frac{\operatorname{adj}(\mathbf{I} - x\mathbf{A})}{\det(\mathbf{I} - x\mathbf{A})} = \frac{\operatorname{adj}(\mathbf{I} - x\mathbf{A})}{\det(\mathbf{I} - x\mathbf{A})}. \tag{3}$$

Corollary 3.7. *The generating function for the number of closed walks around a vertex a in the variable x is*

$$W_G(x)_{aa} = \frac{\phi_{G\setminus a}(x^{-1})}{x \cdot \phi_G(x^{-1})}.$$

Proof. Follows immediately from

$$W_G(x) = \frac{\operatorname{adj}(\mathbf{I} - x\mathbf{A})}{\det(\mathbf{I} - x\mathbf{A})} = \frac{x^{n-1} \operatorname{adj}(x^{-1}\mathbf{I} - \mathbf{A})}{x^n \det(x^{-1}\mathbf{I} - \mathbf{A})},$$

and the definition of the adjugate. □

We would also appreciate to have an expression for $W_G(x)_{ab}$. For that, we make use of an old trick due to Jacobi to arrive at an expression. For any matrix $\mathbf{M}$ with rows and columns indexed by a set V, let $\mathbf{M}_D$ stand for the submatrix with rows and columns indexed by $D \subseteq V$. The following theorem is the correct generalization of Corollary 3.7.

Theorem 3.8. *Let D be a subset of $V(G)$ (assume without loss of generality that the rows and columns indexed by D are the first). Then*

$$\det[W_G(x)]_D = \frac{1}{x^{|D|}} \frac{\phi_{G\setminus D}(x^{-1})}{\phi_G(x^{-1})}.$$

Proof. Let $\mathbf{C}$ be the matrix obtained from $\mathbf{I}$ upon replacing its first $|D|$ columns by the first $|D|$ columns of $\operatorname{adj}(\mathbf{I} - x\mathbf{A})$. Hence

$$(\mathbf{I} - x\mathbf{A}) \cdot \mathbf{C} = \begin{pmatrix} \det(\mathbf{I} - x\mathbf{A})\mathbf{I}_{|D|} & ? \\ \mathbf{0} & (\mathbf{I} - x\mathbf{A})_{\overline{D}} \end{pmatrix}.$$

Note that

$$\det \mathbf{C} = \det \operatorname{adj}(\mathbf{I} - x\mathbf{A})_D = \det[W_G(x)]_D \cdot (\det(\mathbf{I} - x\mathbf{A})^{|D|}).$$

Thus

$$\det[W_G(x)]_D = \frac{\det[(\mathbf{I} - x\mathbf{A})_{\overline{D}}]}{\det(\mathbf{I} - x\mathbf{A})} = \frac{x^{n-|D|} \det(x^{-1}\mathbf{I} - \mathbf{A})_{\overline{D}}}{x^n \det(x^{-1}\mathbf{I} - \mathbf{A})},$$

which yields the result. □

If $D = \{a, b\}$, then

$$W_G(x)_{aa}W_G(x)_{bb} - W_G(x)_{ab}^2 = \frac{1}{x^2}\frac{\phi_{G\setminus ab}(x^{-1})}{\phi_G(x^{-1})},$$

therefore

$$W_G(x)_{ab} = \frac{1}{x}\frac{\sqrt{\phi_{G\setminus a}(x^{-1})\phi_{G\setminus b}(x^{-1}) - \phi_G(x^{-1})\phi_{G\setminus ab}(x^{-1})}}{\phi_G(x^{-1})}.$$

Notice in particular, from Equation (3), and replacing $y = x^{-1}$, that

$$\sqrt{\phi_{G\setminus a}(y)\phi_{G\setminus b}(y) - \phi_G(y)\phi_{G\setminus ab}(y)} = \operatorname{adj}(y\mathbf{I} - \mathbf{A})_{ab},$$

which is a polynomial (meaning: a power series with finite terms), and therefore the term inside the square root must be a perfect square (a fact that is not at all immediate at first sight).

Exercise 3.9. Let $\mathcal{P}_{ab}$ be the set of all paths from a to b. Prove that

$$\sqrt{\phi_{G\setminus a}(y)\phi_{G\setminus b}(y) - \phi_G(y)\phi_{G\setminus ab}(y)} = \sum_{P\in\mathcal{P}_{ab}} \phi_{G\setminus P}(y).$$

Hints:

(i) This will be a proof by induction.

(ii) Define $N_G(y)_{ab}$ to be the generating function for the walks that start at a, never return to it, and end at b. Find a relation between W_{ab}, N_{ab} and W_{aa}.

(iii) Find a relation between N_{ab} and W_{cb} (in $G\setminus a$), where c runs over the neighbours of a.

(iv) Apply induction.

3.3 Spectral decomposition

Say $\mathbf{A} = \sum_{r=0}^{d} \theta_r \mathbf{E}_r$. From the walk generating function $W_G(x) = \sum_{k\geq 0}(\mathbf{A})^k x^k = (\mathbf{I} - x\mathbf{A})^{-1}$, we have

$$W_G(x) = \sum_{r=0}^{d} \frac{1}{1 - x\theta_r} E_r. \tag{4}$$

Thus,

$$W_G(x^{-1}) = \sum_{r=0}^{d} \frac{x}{x - \theta_r} E_r.$$

If we focus on the diagonal entries, we have

$$\frac{x\phi_{G\setminus a}(x)}{\phi_G(x)} = W_G(x^{-1})_{aa} = \sum_{r=0}^{d} \frac{x}{x - \theta_r}(E_r)_{aa},$$

thus, multiplying both sides by $(x-\theta_r)$ and evaluating at $x=\theta_r$, yields

$$(E_r)_{aa} = \left.\frac{(x-\theta_r)\phi_{G\backslash a}(x)}{\phi_G(x)}\right|_{x=\theta_r}$$

For off-diagonal entries, we obtain

$$(E_r)_{ab} = \left.\frac{(x-\theta_r)\sqrt{\phi_{G\backslash a}(x)\phi_{G\backslash b}(x) - \phi_G(x)\phi_{G\backslash ab}(x)}}{\phi_G(x)}\right|_{x=\theta_r}.$$

Exercise 3.10. Show that if θ_r is an eigenvalue of $\mathbf{A}(G)$ with multiplicity m_r, then, for any $a \in V(G)$, its multiplicity in $\mathbf{A}(G\backslash a)$ is at least $m_r - 1$. Prove that equality holds if and only if there is at least one eigenvector for θ_r whose entry corresponding to a is non-zero.

Exercise 3.11. The goal of this exercise is to show that for any two matrices $\mathbf{M}$ and $\mathbf{N}$ so that $\mathbf{MN}$ and $\mathbf{NM}$ are defined, the following identity holds

$$\det(\mathbf{I}-\mathbf{MN}) = \det(\mathbf{I}-\mathbf{NM}).$$

To achieve this, find the two matrices that make both products below true, and finish the exercise.

$$\begin{pmatrix}\mathbf{I} & -\mathbf{M}\\ \mathbf{N} & \mathbf{I}\end{pmatrix} = \begin{pmatrix}\mathbf{I} & -\mathbf{M}\\ \mathbf{0} & \mathbf{I}\end{pmatrix}\begin{pmatrix}? & ?\\ ? & ?\end{pmatrix} \quad \text{and} \quad \begin{pmatrix}\mathbf{I} & -\mathbf{M}\\ \mathbf{N} & \mathbf{I}\end{pmatrix} = \begin{pmatrix}\mathbf{I} & \mathbf{0}\\ \mathbf{N} & \mathbf{I}\end{pmatrix}\begin{pmatrix}? & ?\\ ? & ?\end{pmatrix}$$

Exercise 3.12. Let $w(x)$ be the generating function whose coefficient of x^k count the total amount of all walks in the graph of length k. The goal of this exercise is to show that

$$w(x) = \frac{1}{x}\left(\frac{(-1)^n\phi_{\overline{G}}(-1-x^{-1})}{\phi_G(x^{-1})} - 1\right).$$

Recall that $\mathbf{A}(\overline{G}) = \mathbf{J}-\mathbf{I}-\mathbf{A}(G)$. You will use that $w(x) = \mathbf{1}^T W_G(x)\mathbf{1}$, that $\mathbf{J} = \mathbf{1}\mathbf{1}^T$, and finally the past exercise.

3.4 Reconstructing

In this section, we will show that the characteristic polynomial is reconstructible from the deck of the graph — that is, if the conjecture is false, then any counterexamples will have to be graphs with the same spectrum.

We would like to be able to reduce $\phi_G(x)$ somehow to an expression depending on the vertex-deleted subgraphs of G. Our best chance is then to look at Corollary 3.7, and take the trace in Equation (4). First, answer the exercise.

Exercise 3.13. Explain why $\operatorname{tr} E_r = m_r$, the multiplicity of θ_r as an eigenvalue.

Now we shall have

$$\frac{1}{\phi_G(x)}\left(\sum_{a\in V(G)}\phi_{G\backslash a}(x)\right) = \operatorname{tr}\left[x^{-1}W_G(x^{-1})\right] = \sum_{r=0}^{d}\frac{m_r}{x-\theta_r}.$$

Hence

$$\sum_{a\in V(G)} \phi_{G\backslash a}(x) = \sum_{r=0}^{d} \left(m_r(x-\theta_r)^{m_r-1}\right) \prod_{s\neq r}(x-\theta_s)^{m_s} = \phi_G(x)'.$$

This shows that we need only the characteristic polynomial of the graphs in the deck of G to recover the characteristic polynomial of G, except for its constant term. This actually will prove itself a considerably harder task, to which we devote the remaining of this subsection.

We start by actually finding a combinatorial expansion for the coefficients of $\phi(x)$, which in its own self is interesting and relevant. A *sesquivalent* subgraph H of G is a subgraph satisfying

(i) $|V(H)| = |V(G)|$.

(ii) Every connected component of H is either an isolated vertex, or an edge, or a cycle.

For each sesquivalent subgraph H of G, let $v(H)$, $e(H)$ and $c(H)$ denote the number of connected components which are, respectively, isolated vertices, edges and cycles.

Theorem 3.14 (Harary, Biggs)**.** *Let G be a simple graph, and $\mathcal{H}$ the set of all sesquivalent subgraphs of G. Then*

$$\phi_G(x) = \sum_{H\in\mathcal{H}} (-1)^{e(H)}(-2)^{c(H)}x^{v(H)}.$$

Proof. Leibniz formula for the determinant gives

$$\phi_G(x) = \det(x\mathbf{I}-\mathbf{A}) = \sum_{\sigma\in S_n} (-1)^{\epsilon(\sigma)} \prod_{i=1}^{n} (x\mathbf{I}-\mathbf{A})_{i\sigma(i)}.$$

(The sum runs over all permutations of $\{1,...,n\}$, and $\epsilon(\sigma)$ is the number of cycles of even length in the decomposition of σ as a product of disjoint cycles.)

Consider the set of all permutations fixing precisely the points belonging to the subset $D \subseteq V(G)$. The sum of the terms corresponding to these permutations will therefore be

$$x^{|D|}(-1)^{n-|D|}\det(\mathbf{A}(G\backslash D)).$$

Each permutation of $V(G)\backslash D$ with fixed points contributes nothing to the determinant of $\mathbf{A}(G\backslash D)$. Those without will contain cycles of length two, or longer. Note that the support of the cycle structure of a permutation is a sesquivalent subgraph of $G\backslash D$. The cycles of length 2 are edges. The longer ones are the cycles of the graph. Each of the longer cycles of σ could have their orders reversed, yielding a permutation corresponding to the same sesquivalent subgraph H. Thus the total number of permutations corresponding to the sesquivalent subgraph H is $2^{c(H)}$.

Say the permutation σ corresponds to sesquivalent subgraph H. The quantity of cycles of odd length in σ has the same parity as $n-|D|$. If this is even, then total number of cycles, which is $e(H)+c(H)$, has the same parity as the number of even cycles, which is $\epsilon(\sigma)$. Otherwise, total number of cycles has opposite parity. Thus, if σ corresponds to the sesquivalent subgraph H with no isolated vertices, then

$$(-1)^{n-|D|}(-1)^{\epsilon(\sigma)} = (-1)^{e(H)+c(H)}.$$

Therefore the sum of the terms corresponding to the permutations fixing the set D will be

$$x^{|D|}(-1)^{n-|D|}\det(\mathbf{A}(G\backslash D)) = x^{|D|}\sum_H (-1)^{e(H)+c(H)}2^{c(H)}.$$

where the sum runs over the sesquivalent subgraphs of $G\backslash D$ with no isolated vertices. Varying the set D over all subsets of G will yield the desired expressions of the theorem. □

The constant term in $\phi_G(x)$, which is $(-1)^n\det(\mathbf{A}(G))$, is, according to the theorem above, equal to

$$\sum_H (-1)^{e(H)}(-2)^{c(H)}$$

where the sum runs over the sesquivalent subgraphs H of G with no isolated vertices.

Recall Kelly's lemma, which is useful to count copies of a subgraph H with $|V(H)| < |V(G)|$.

Lemma 3.15. *For any graphs G and H,*

$$(|V(G)| - |V(H)|)\ \nu(H,G) = \sum_{a\in V(G)} \nu(H, G\backslash a).$$

With a little more work, we have the following. Recall that a graph homomorphism from G_1 to G_2 is a function from $V(G_1)$ to $V(G_2)$ that preserves adjacency (but not necessarily non-adjacency).

Lemma 3.16. *G on n vertices, and H a disconnected graph on n vertices. Then $\nu(H,G)$ is reconstructible.*

Proof. Let H_1 and H_2 be disjoint subgraphs whose union is H. There are $\nu(H_1,G)\nu(H_2,G)$ homomorphisms from H to G which are injective on H_1 and H_2. Several of those however overlay images of vertices from H_1 and H_2. But we can count those. For each F on fewer than n vertices, there are $\nu(F,G)$ copies of F in G, and we can count the number of surjective homomorphisms from H to F which are injective in both H_1 and H_2. We multiply both things, and sum this for all F. We then subtract the total from $\nu(H_1,G)\nu(H_2,G)$ to recover $\nu(H,G)$. □

The result above allows us to compute the sum

$$\sum_H (-1)^{e(H)}(-2)^{c(H)}$$

for all disconnected H. The only thing remaining now to account for are the connected H.

A graph has vertex connectivity 1 if it is connected and contains a vertex whose removal disconnects the graph (a cut-vertex). A block is a maximal subgraph that does not contain a cut-vertex. For example, a tree contains $n-1$ blocks (each corresponding to an edge). The number of blocks in a 1-connected subgraph is the number of cut-vertices added by 1.

Lemma 3.17. *Let H be a 1-connected graph, on n vertices. The number of subgraphs of G with n vertices that contain the same collection of blocks of H is reconstructible.*

Proof. Assume H contains exactly two blocks H_1 and H_2 (thus $|V(H_1)| + |V(H_2)| = n+1$). Consider all homomorphisms from $H_1 \cup H_2$ to G which are injective in both H_1 and H_2. There are $\nu(H_1, G)\nu(H_2, G)$ such homomorphisms. The number of such mappings whose image is contained in a vertex deleted subgraph of G is reconstructible (see lemma above and Kelly's lemma). Thus the number of those whose image is G, obtained from overlaying only one vertex of H_1 with one of H_2, is reconstructible. These will correspond precisely to the spanning subgraphs of G which have H_1 and H_2 as their blocks. Now we can simply apply induction on the number of blocks of H to account for when H has any number of blocks. □

Using both lemmas above, one can show that:

Corollary 3.18. *If G is disconnected, then G is reconstructible. If G is a tree, then G is reconstructible.*

Exercise 3.19. Write the details proving the corollary above.

Corollary 3.20. *The number of Hamilton cycles of G can be reconstructed from the deck.*

Proof. The number of edges of G is reconstructible, so we can count the number of subgraphs of G with precisely n edges. We can also count how many of those are in vertex-deleted subgraphs, thus we can recover how many spanning subgraphs of G have precisely n edges. Out of these, we can count those which are disconnected and those which contain a cut-vertex, because they will contain a unique cycle of length $k < n$. The remaining graphs in the count will be Hamilton cycles. □

Clearly the implicit algorithm in the proof above is extremely inefficient, but there was no hope of providing an efficient algorithm that counts the number of Hamilton cycles in a graph anyway (deciding whether one exists is already itself a hard task).

Theorem 3.21. *The characteristic polynomial of G is reconstructible from the deck.*

Proof. We proved that

$$\phi_G(x)' = \sum_{a \in V(G)} \phi_{G\setminus a}(x).$$

The constant term of $\phi_G(x)$ is

$$\sum_H (-1)^{e(H)}(-2)^{c(H)}$$

where the sum runs over the sesquivalent subgraphs H of G with no isolated vertices. Those which are disconnected can be dealt with Lemma 3.16. Those which are connected correspond precisely to the Hamilton cycles of G, and this number can be reconstructed from Corollary 3.20. □

Recall the we proved that

$$\phi_{G\setminus a}(y)\phi_{G\setminus b}(y) - \phi_G(y)\phi_{G\setminus ab}(y)$$

is a perfect square of a polynomial, say $q_{ab}(y)$. If $\phi_G(y)$ is irreducible over the rationals, that it is easy to show that $\phi_{G\backslash ab}(y)$ is completely determined by $\phi_G(y)$, $\phi_{G\backslash a}(y)$, and $\phi_{G\backslash b}(y)$.

Having the eigenvalues of $G\backslash ab$, we can recover its number of edges. So we know the number of edges in G, $G\backslash a$, $G\backslash b$ and $G\backslash ab$. Hence we can find whether there is an edge between a and b in G. As a consequence:

Theorem 3.22 (Tutte)**.** *If characteristic polynomial of G is irreducible over the rationals, then G itself is reconstructible.* □

We list two open questions related to our work in this chapter.

Problem 3.1. *Can you reconstruct of the characteristic polynomial of the Laplacian matrix from the deck?*

Problem 3.2. *Instead of the deck of G, assume you have access only to the characteristic polynomials of the graphs in the deck. Can you reconstruct $\phi_G(x)$? (It is known that this is possible if you have the characteristic polynomials of the graphs in the deck and their complement.)*

Define now the three-variable polynomial

$$\Phi_G(y,z,x) = \sum_{H\in\mathcal{H}} y^{e(H)} z^{c(H)} x^{v(H)}.$$

Note that $\phi_G(x) = \Phi_G(-1,-2,x)$.

Exercise 3.23. Prove that

$$\frac{\partial}{\partial x}\Phi_G(y,z,x) = \sum_{a\in V(G)} \Phi_{G\backslash a}(y,z,x).$$

Find expressions for

$$\frac{\partial}{\partial y}\Phi_G(y,z,x) \text{ and } \frac{\partial}{\partial z}\Phi_G(y,z,x).$$

Exercise 3.24. Verify that $\Phi_G(y,z,x)$ is reconstructible from the deck of G.

Exercise 3.25. Find a recurrence for Φ assuming G contains a cut-edge (meaning: write Φ_G in terms of Φ for some subgraphs of G.) Try the same exercise assuming G contains a cut-vertex.

Exercise 3.26. Let $\mathcal{C}_a$ be the set of cycles containing a vertex a. Explain why

$$\Phi_G = x\Phi_{G\backslash a} + y\sum_{b\sim a} \Phi_{G\backslash ab} + z\sum_{C\in\mathcal{C}_a} \Phi_{G\backslash C}.$$

Exercise 3.27. Assume all cycles of G have the same length, say c. Find a partial differential equation satisfied by Φ.

3.5 The matching polynomial of a graph

Let $\mathcal{M}(G)$ be the set of all spanning subgraphs of G whose connected components are either isolated vertices or isolated edges. The matching polynomial of a graph is defined as

$$\mu_G(x) = \sum_{M \in \mathcal{M}(G)} (-1)^{e(M)} x^{v(M)}.$$

Note that it is precisely equal to the evaluation $\Phi(-1, 0, x)$ of the polynomial $\Phi(y, z, x)$ defined in the past subsection. In fact,

Theorem 3.28. *Given a graph G,*

$$\mu_G(x) = \phi_G(x)$$

if and only if G is a tree.

Proof. One direction is obvious from the formula of Φ. The other I leave as a challenging exercise. □

Exercise 3.29. Verify that

$$\mu_G(x)' = \sum_{a \in V(G)} \mu_{G \setminus a}(x),$$

and, prove that, if $e = \{u, v\}$ is an edge of G, then

$$\mu_G(x) = \mu_{G \setminus e}(x) - \mu_{G \setminus uv}(x).$$

Exercise 3.30. Find recurrences for $\mu_{P_n}(x)$, $\mu_{K_n}(x)$ and $\mu_{C_n}(x)$ based on the matching polynomials of smaller graphs in each of the families. (Hint: use Exercise 3.26).

The recurrences you found in the past exercise show that matching polynomials in each of those families of graphs form what is a known as a sequence of orthogonal polynomials. We will not get into details of the theory of orthogonal polynomials, but over the next few results we will see some glimpse of it. Given polynomials $p(x)$ and $q(x)$, we define an inner product by

$$\langle p, q \rangle = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \mathrm{e}^{-x^2/2} p(x) q(x) \mathrm{d}x.$$

Do not get scared. Just bear with me. But maybe now it would be a good time to remember that

$$1 = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \mathrm{e}^{-x^2/2} \mathrm{d}x \quad \text{and} \quad 0 = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} x \mathrm{e}^{-x^2/2} \mathrm{d}x.$$

Exercise 3.31. Prove these equalities. Hint: one of them is easy. For the other, write its square, and change variables to polar coordinates.

Lemma 3.32. *Let*

$$M(n) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \mathrm{e}^{-x^2/2} x^n \, dx.$$

The number of perfect matchings in K_n is equal to $M(n)$.

Proof. Integration by parts implies

$$M(n) = \frac{1}{\sqrt{2\pi}}\left[\frac{x^{n+1}}{n+1}\mathrm{e}^{-x^2/2}\right]_{-\infty}^{+\infty} + \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}\frac{x^{n+2}}{n+1}\mathrm{e}^{-x^2/2}\mathrm{d}x.$$

The first term is 0. So it follows that $M(n) = M(n+2)/(n+1)$. As seen above, $M(1) = 0$ and $M(0) = 1$. Hence $M(\text{odd}) = 0$ and

$$M(2m) = (2m-1)!!$$

as we wanted to show. □

Recall that $(-1)^{n/2}\mu_G(0)$ is the number of perfect matchings in G. Denote this number by $\mathrm{pm}(G)$.

Theorem 3.33. *For any G, we have*

$$\mathrm{pm}(\overline{G}) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}\mathrm{e}^{-x^2/2}\mu_G(x)dx$$

sketch. The proof is by induction on the number of edges in G. If G has no edges, this falls precisely in the statement of the lemma. If G has one edge, then both sides satisfy the same recursion given by the second part of Exercise 3.29. □

Exercise 3.34. Prove that

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}\mathrm{e}^{-x^2/2}\mu_{K_n}(x)\mu_{K_m}(x)\mathrm{d}x = \begin{cases} m!, & \text{if } m = n; \\ 0, & \text{otherwise.} \end{cases}$$

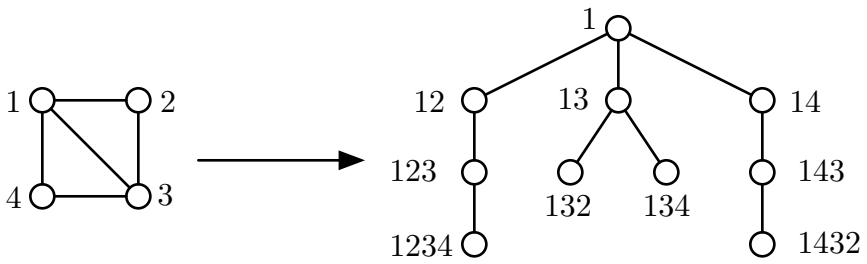
Hint: look at $K_n \cup K_m$, its complement, and the past exercise.

The conclusion from the result above is that the family $\{\mu_{K_n}(x)\}_{n\geq 0}$ is a family of orthogonal polynomials according to the inner product defined in this subsection.

3.6 Real roots

Our goal here is to show that the matching polynomial of any graph has only real roots.

Given a graph G and a vertex u in G, the *path tree* of G with respect to u is a rooted tree whose vertices correspond to the paths of G that start at u, and a the children of a vertex corresponding to path P are those vertices corresponding to paths obtained from one further edge at the end of P. For example



Theorem 3.35. *Let G be a graph, $u \in V(G)$. Let $T = T(G,u)$ be the path tree of G with respect to u. Then*

$$\mu_T(x)\mu_{G\setminus u}(x) = \mu_G(x)\mu_{T\setminus u}(x),$$

and $\mu_G(x)$ divides $\mu_T(x)$.

Proof. If G itself is already a tree, then there is nothing to prove, as $G = T$. We may assume the results holds true for vertex-deleted subgraphs of G. Thus

$$\mu_G(x) = x\mu_{G\setminus u} - \sum_{v\sim u} \mu_{G\setminus uv}(x).$$

Thus, applying induction, we have

$$\frac{\mu_G(x)}{\mu_{G\setminus u}(x)} = x - \sum_{v\sim u} \frac{\mu_{T(G\setminus u,v)\setminus v}(x)}{\mu_{T(G\setminus u,v)}(x)}.$$

Now, $T(G\setminus u, v)$ is isomorphic to the branch of $T(G,u)$ attached to u that starts at the vertex corresponding to the path uv. Thus

$$\frac{\mu_{T(G\setminus u,v)\setminus v}(x)}{\mu_{T(G\setminus u,v)}(x)} = \frac{\mu_{T(G,u)\setminus\{u,uv\}}(x)}{\mu_{T(G,u)\setminus u}(x)}.$$

Therefore

$$\frac{\mu_G(x)}{\mu_{G\setminus u}(x)} = \frac{x\mu_{T(G,u)\setminus u}(x) - \sum_{v\sim u}\mu_{T(G,u)\setminus\{u,uv\}}(x)}{\mu_{T(G,u)\setminus u}(x)} = \frac{\mu_{T(G,u)}(x)}{\mu_{T(G,u)\setminus u}(x)},$$

as wanted. For the second assertion, by induction, it follows that $\mu_{G\setminus u}(x)$ divides $\mu_{T(G\setminus u,v)}(x)$. As $T(G\setminus u, v)$ is a branch of $T(G,u)\setminus u$, it follows that $\mu_{T(G\setminus u,v)}(x)$ divides $\mu_{T(G,u)\setminus u}(x)$, so $\mu_{G\setminus u}(x)$ itself divides $\mu_{T(G,u)\setminus u}(x)$. Hence $\mu_G(x)$ divides $\mu_T(x)$. □

Corollary 3.36. *The roots of $\mu_G(x)$ are real, for any G. Moreover, they are symmetrically distributed around the origin.*

Proof. The polynomial $\mu_G(x)$ divides $\mu_T(x)$, which is equal to $\phi_T(x)$. This is the characteristic polynomial of a symmetric matrix, hence its roots are real. Therefore the roots of $\mu_G(x)$ are real.

The second part follows immediately from the fact that all exponents of x in $\mu_G(x)$ are either all odd or all even. □

Exercise 3.37. Prove that the zeros of $\mu_{G\setminus u}$ interlace those of μ_G. If G is connected, prove that the largest zero of μ_G is simple, and strictly larger than that of $\mu_{G\setminus u}$. Hint: use Theorem 3.35.

We can also bound the largest eigenvalue of μ relatively well.

Exercise 3.38. Show (again) that the largest eigenvalue of a non-negative matrix is upper bounded by its largest row sum.

Exercise 3.39. Extend the result above to argue that the largest eigenvalue of a non-negative matrix $\mathbf{M}$ is upper bounded by the largest row sum of $\mathbf{DMD}^{-1}$ for any positive diagonal matrix $\mathbf{D}$.

Exercise 3.40. Let T_Δ be a tree so that all vertices have degree $\Delta > 2$ or 1. Prove that its largest eigenvalue is upper bounded by $2\sqrt{\Delta - 1}$. Hint: Fix a vertex of degree Δ to call the root, and conjugate $\mathbf{A}(T_\Delta)$ by the diagonal matrix defined as $\mathbf{D}_{aa} = \sqrt{\Delta - 1}^{d(a)}$, where $d(a)$ is the distance from a to the root. Use the exercises above.

Exercise 3.41. Argue that any tree of maximum degree $\Delta > 1$ has its largest eigenvalue small or equal than $2\sqrt{\Delta - 1}$.

Exercise 3.42. Let G be a graph with $\Delta(G) > 1$. Show that the largest root λ of $\mu_G(x)$ satisfies

$$\sqrt{\Delta(G)} \leq \lambda \leq 2\sqrt{\Delta(G) - 1}.$$

(The upper bound should follow easily from the exercises above. The lower bound is your job to find.)

3.7 Number of matchings

The fact that the roots of $\mu_G(x)$ are real brings a combinatorial consequence. A sequence of numbers $(a_i)_{i\geq 0}$ is log-concave if $a_i^2 \geq a_{i-1}a_{i+1}$ for all $i \geq 1$. If the numbers are positive, then this is equivalent to having $(a_{i+1}/a_i)_{i\geq 0}$ non-increasing. Thus, a log-concave sequence of positive numbers is unimodal, meaning, it first increases, then stays constant, then decreases.

The binomial coefficients $\binom{n}{k}$, $k = 0, ..., n$, form a (finite) log-concave sequence. Clearly, if (a_i) and (b_i) are log-concave, so is $(a_i b_i)$.

Lemma 3.43. *If $p(x) = \sum_i a_i x^i$ is a polynomial of degree n with real roots only, then $(a_i/\binom{n}{i})$ form a log-concave sequence.*

Proof. This follows from writing

$$\frac{\mathrm{d}^{n-i-2}}{\mathrm{d}^{n-i-2}x} x^{n-i} \frac{\mathrm{d}^i}{\mathrm{d}^i x} p(x^{-1}) = \frac{1}{2} n! \left(\frac{a_i}{\binom{n}{i}} x^2 + \frac{a_{i+1}}{\binom{n}{i+1}} 2x + \frac{a_{i+2}}{\binom{n}{i+2}} \right).$$

(Fill in the details). □

Let m_k be the number of matchings in G with k edges. Note that

$$\mu_G(x) = \sum_{k\geq 0} (-1)^k m_k x^{n-2k}.$$

Corollary 3.44. *The sequence $(m_k)_{k\geq 0}$ is log-concave (and therefore unimodal).*

Proof. Assume n is even. Then $\mu_G(x) = q(x^2)$. Note that

$$p(x) = \sum_{k\geq 0} m_k x^k = x^{n/2} q(-x^{-1}),$$

which also has real roots. Similar argument for n odd. It follows then from Lemma that $(m_k)_{k\geq 0}$ is a log-concave sequence. □

3.8 Average

In this final section about the matching polynomial, we prove a remarkable result connecting $\mu(x)$ and $\phi(x)$.

Theorem 3.45. *Let G be a graph with m edges. Then*

$$\mu_G(x) = \frac{1}{2^m}\sum_F \phi_F(x),$$

where the sum runs over all 2^m signed graphs F whose underlying edges are exactly those of G.

To be a clear, $\mathbf{A}(F)$ is precisely $\mathbf{A}(G)$, except that certain symmetric off-diagonal entries have been changed to -1.

Proof. We have

$$\frac{1}{2^m}\sum_F \phi_F(x) = \frac{1}{2^m}\sum_F \sum_{\sigma\in S_n} (-1)^{\epsilon(\sigma)} \prod_{i=1}^{n} (x\mathbf{I} - \mathbf{A}(F))_{i\sigma(i)},$$

then

$$\frac{1}{2^m}\sum_F \phi_F(x) = \frac{1}{2^m}\sum_{\sigma\in S_n} (-1)^{\epsilon(\sigma)} \sum_F \prod_{i=1}^{n} (x\mathbf{I} - \mathbf{A}(F))_{i\sigma(i)},$$

Note that if σ contains a cycle with more than two vertices, then

$$\sum_F \prod_{i=1}^{n} (x\mathbf{I} - \mathbf{A}(F))_{i\sigma(i)} = 0,$$

as we can sum over all possible signings of this cycle having the rest constant, and later vary the rest, but the sum over all possible signings of a cycle of length larger than 2 is 0.

Thus the only permutations that contribute are those with transpositions and fixed points only, and for those the signing is irrelevant. The sum over all such permutations coincides with the matching polynomial of the graph. Therefore

$$\frac{1}{2^m}\sum_F \phi_F(x) = \frac{1}{2^m}\sum_F \sum_{M\in\mathcal{M}(G)} (-1)^{e(M)} x^{v(M)} = \frac{1}{2^m} 2^m \mu_G(x),$$

as we wished. □

Exercise 3.46. Let G be a graph, and F be obtained from G upon signing some of the edges. What exactly can be said about

$$\phi_G(x) + \phi_F(x) \quad ?$$

Exercise 3.47. Assume G is a graph with the property that every cycle of G contains at least one edge that belongs to no other cycle. Show how to compute μ_G efficiently.

3.9 Tutte polynomial - a quick tour

The past sections focused on the characteristic polynomial (of the adjacency matrix) and the matching polynomial, and some of their variations. In this section, we briefly introduce a new class of graph polynomials — this one intimately related to certain counting questions in a graph.

The concepts of deletion and contraction of an edge will be important. We will denote by $G\backslash e$ the graph where the edge e was simply removed, and this operation will be called *edge deletion*. By G/e we will mean the graph where the edge e was removed, and its incident vertices were identified, and this operation shall be named *edge contraction*. Note that if edge contraction is allowed, we also need to allow multi-edges and loops, as the contraction of any edge in a triangle creates a pair of multi-edges, and the contraction of any edge in a pair of multi-edges creates a loop.

For a graph G, we denote by $T_G(x, y)$ its Tutte polynomial, which is recursively defined by he following relations:

(i) If there is no edge, $T_{E_n}(x, y) = 1$.

(ii) If e is an edge, $T_G = \begin{cases} x \cdot T_{G\backslash e} & \text{if } e \text{ is a bridge} \\ y \cdot T_{G\backslash e} & \text{if } e \text{ is a loop} \\ T_{G\backslash e} + T_{G/e} & \text{if } e \text{ is neither a bridge nor a loop} \end{cases}$

Suppose A is a subset of $E(G)$. We define:

- $\kappa(A)$ is the number of connected components of the spanning subgraph of G which contains only the edges in A, i.e., the number of connected components of $G\backslash(E - A)$. Note that $\kappa(E)$ is the number of connected components of G, and that $\kappa(A) \geq \kappa(E)$ for all $A \subseteq E$.
- $r(A)$ is the number of edges of a maximal subset of A which contains no cycle. Because for each connected component in the subgraph defined by A we can form a tree, it is true in general that $r(A) = |V(G)| - \kappa(A)$.
- $n(A) = |A| - r(A)$, known as the nullity of A.

With these definitions, one can write the following formula for $T_G(x, y)$.

Theorem 3.48.

$$T_G(x, y) = \sum_{A \subseteq E} (x - 1)^{r(E)-r(A)} (y - 1)^{n(A)}$$

Exercise 3.49. Prove by induction (or at least convince yourself) that this theorem is true.

Exercise 3.50. What is $T(1, 1)$ counting?

Now let us show how to use this polynomial to count colourings. Let $P_G(x)$ be a function that, for x integers, returns the number of proper colourings of G with x colours. This is a well defined function, but we will see that it admits a very natural extension to the real numbers which is a polynomial of finite degree, called the chromatic polynomial of G.

Suppose the edge e is originally adjacent to u and v. Looking at a proper colouring of $G\backslash e$ with k colours, either the colours of u and v are the same, or they are different. Thus, there is a bijection from the former case to the colourings of G/e with k colours, and another bijection from the latter case to the proper colourings of G. As $p_G(k)$ denotes the number of proper k-colourings of G, we have

$$p_G(k) = p_{G\backslash e}(k) - p_{G/e}(k).$$

Of course, if G contains a loop then $p_G(k) = 0$ for all k, and if G contains no edges, meaning $G = E_n$, then $p_G(k) = k^n$. We can also simply realize that if e is a bridge, then

$$p_G(k) = \frac{k-1}{k} p_{G\backslash e}(k).$$

According to this recursive definition, we may as well define $p_G(x)$ as a polynomial, which will have degree at most n.

Theorem 3.51.

$$p_G(x) = x^{\kappa(E)}(-1)^{r(E)}T(-(x-1), 0)$$

Proof. It is enough to verify that the polynomial on the right satisfies the recursive definition of $p_G(x)$. If G has no edges (meaning $E = \emptyset$), then it is immediate. If G contains a loop, then because $y = 0$, it is also immediate. Now fix $e \in E$. If e is a bridge, then

$$\begin{aligned}
&\frac{x-1}{x}\big(x^{\kappa(E)+1}(-1)^{r(E)-1}T_{G\backslash e}(-(x-1), 0)\big) = \\
=&(x-1)\big(x^{\kappa(E)}(-1)^{r(E)-1}T_{G\backslash e}(-(x-1), 0)\big) = \\
=&x^{\kappa(E)}(-1)^{r(E)}T_G(-(x-1), 0).
\end{aligned}$$

If e is not a bridge, then

$$\begin{aligned}
&\big(x^{\kappa(E)}(-1)^{r(E)}T_{G\backslash e}(-(x-1), 0)\big) - \big(x^{\kappa(E)}(-1)^{r(E)-1}T_{G/e}(-(x-1), 0)\big) = \\
=&\big(x^{\kappa(E)}(-1)^{r(E)}T_{G\backslash e}(-(x-1), 0)\big) + \big(x^{\kappa(E)}(-1)^{r(E)}T_{G/e}(-(x-1), 0)\big) = \\
=&x^{\kappa(E)}(-1)^{r(E)}T_G(-(x-1), 0)
\end{aligned}$$

Either way, we have seen that both polynomials in the statement satisfy the same recursive definition, so they are equal. □

One immediately observes that having the Tutte polynomial allows for an immediate computation of the chromatic number of the graph, so computing the Tutte polynomial is NP-hard.

It is however a quite important invariant. The following two exercises display how ubiquitous this edge deletion/contraction formula is for computing certain graph parameters.

3.9.1 Reliability

Suppose there is a fixed probability p such that each edge of a graph G will be removed with this probability. Let $R_G(p)$ denote the probability that the number of connected components of G does not increase. This is a measure of how reliable a network is. When e is not a bridge, its removal does not change the number of connected components. Thus, if it is removed, what happens with probability p, the chance that the number of connected components of G increases after the whole procedure is precisely the chance that the number of connected components of $G\backslash e$ increases. If e is not removed, G will have the same structure with respect to edge-connectivity of G/e. Therefore:

$$R_G(p) = p \cdot R_{G\backslash e}(p) + (1-p) \cdot R_{G/e}(p)$$

Exercise 3.52. Prove that

$$R_G(p) = p^{n(E)}(1-p)^{r(E)} T_G\left(1, \frac{1}{p}\right).$$

3.9.2 Flows

If this is too strange, research about flows from abelian groups before.

Suppose now that G has an arbitrary orientation. Let H be a finite abelian group and let $\overrightarrow{e}$ be a directed edge. An H-flow of G is a function from the arcs to H so that the sum of the elements that enter a vertex is equal to the sum of what leaves it.

Let e_{uv} denote the vertex of G/e which is the identification of the neighbours u and v of e. We consider an H-flow on G/e and we look at the H-function on $G\backslash e$ which attributes the same values on the edges. When e_{uv} is split back to u and v, either both vertices remain with an excess of 0, or one of them keeps an excess which is the inverse of the excess of the other. In the former case, there is a bijection between some H-flows of G/e and the H-flows of $G\backslash e$. In the latter case, there is a bijection of the other H-flows of G/e and the H-flows of G, where obviously $\overrightarrow{e}$ will receive the non-zero excess of the vertices u and v.

Let $F_G(H)$ denote the number of H-flows on a given orientation of a graph, hence:

$$F_G(H) = F_{G/e}(H) - F_{G\backslash e}(H)$$

Note that the domain of the function is a set of groups, but we will see that this function depends only on the size of the group.

Exercise 3.53. If $|H| = q$, prove that

$$F_G(q) = (-1)^{n(E)} T_G(0, 1-q)$$

3.9.3 Reconstruction

Recall that

$$T_G(x,y) = \sum_{A \subseteq E} (x-1)^{r(E)-r(A)}(y-1)^{n(A)},$$

and recall Lemmas 3.16 and 3.17.

Exercise 3.54. Prove that the Tutte polynomial of a graph is reconstructible from the deck.

The final observation is that if the reconstruction conjecture is false, then the counterexamples will be graphs with the same matching polynomial, same characteristic polynomial, and same Tutte polynomial (and therefore similar chromatic numbers, number of spanning trees, etc.)

3.10 References

Here is the set of references used to write the past few pages.

The main reference for this section is the book of Godsil, wherein references for all the results about the characteristic and matching polynomials can be found (Chapters 1, 2, 4 and 6).

(a) Chris D Godsil. *Algebraic Combinatorics*. Chapman & Hall, New York, 1993

The section on Tutte's polynomial comes mostly from

(b) T. Brylawski and J. Oxley. The Tutte polynomial and its applications. In N. White, editor, *Encyclopedia of Mathematics and its Applications*, volume 40, chapter Matroid Applications. Cambridge University Press, Cambridge, 1992

(c) Chris Godsil and Gordon Royle. *Algebraic Graph Theory*. Springer-Verlag, New York, 2001

Elias Hagos proved that the characteristic polynomial is reconstructible from the characteristic polynomials of the graphs in the deck and their complements (if they are correctly paired up).

(d) Elias M Hagos. The characteristic polynomial of a graph is reconstructible from the characteristic polynomials of its vertex-deleted subgraphs and their complements. *The Electronic Journal of Combinatorics*, 7(1):12, 2000

4 Eigenvalues and the structure of graphs

4.1 Rayleigh quotients and Interlacing

Given a symmetric matrix $\mathbf{M}$, we recall the definition of the Rayleigh quotient of $\mathbf{M}$ with respect to a non-zero vector $\mathbf{v}$:

$$R_{\mathbf{M}}(\mathbf{v}) = \frac{\mathbf{v}^T\mathbf{M}\mathbf{v}}{\mathbf{v}^T\mathbf{v}}.$$

We will always assume vectors whose Rayleigh quotient is being taken are non-zero. As we have seen, if $\mathbf{v}$ is an eigenvector with corresponding eigenvalue θ, then

$$R_{\mathbf{M}}(\mathbf{v}) = \theta.$$

We also saw that if $\lambda_1 \geq ... \geq \lambda_n$ are the eigenvalues of $\mathbf{M}$ with corresponding eigenprojectors E_rs, and assuming $\mathbf{v}$ is normalized, then

$$R_{\mathbf{M}}(\mathbf{v}) = \mathbf{v}^T\left(\sum_{r=1}^{n}\lambda_r E_r\right)\mathbf{v} = \sum_{r=1}^{n}\lambda_r(\mathbf{v}^T E_r\mathbf{v}) \leq \lambda_1\left(\sum_{r=0}^{d}\mathbf{v}^T E_r\mathbf{v}\right) = \lambda_1$$

for all vectors $\mathbf{v}$, and equality holds if and only if $\mathbf{v}$ belongs to the λ_1 eigenspace.

Lemma 4.1. *Let* $\mathbf{M}$ *be a symmetric matrix, with largest eigenvalue* λ_1 *and smallest eigenvalue* λ_n*. Then*

$$\lambda_1 = \max_{\mathbf{v}\in\mathbb{R}^n} R_{\mathbf{M}}(\mathbf{v}) \quad \textit{and} \quad \lambda_n = \min_{\mathbf{v}\in\mathbb{R}^n} R_{\mathbf{M}}(\mathbf{v}).$$

□

Examining more carefully how we bounded the Rayleigh quotient, it is not hard to see that all eigenvalues can be defined as a max or min of the Rayleigh quotient over certain subspaces. Let L_r denote the orthogonal complement to the sum of the eigenlines corresponding to the largest eigenvalues all the way to λ_r, that is

$$L_r = \text{null } (E_1 + E_1 + ... + E_{r-1}).$$

Likewise, define S_r to correspond to the orthogonal complement to the sum of the eigenlines corresponding to the smallest eigenvalues all the way to λ_{r+1}, that is

$$S_r = \text{null } (E_{r+1} + E_{r+2} + ... + E_n).$$

It follows immediately that

$$\lambda_r = \max_{\mathbf{v}\in L_r} R_{\mathbf{M}}(\mathbf{v}) = \min_{\mathbf{v}\in S_r} R_{\mathbf{M}}(\mathbf{v}).$$

The expression of λ_r can be made with the subspaces L_r and S_r implicitly defined, via a min-max formula.

Lemma 4.2 (Courant–Fischer–Weyl min-max principle)**.** *Let* $\mathbf{M}$ *be a symmetric matrix, with eigenvalues* $\lambda_1 \geq ... \geq \lambda_n$*. Then*

$$\lambda_k = \min_{\substack{\textit{subspace } U \\ \dim U = n-k+1}} \max_{\mathbf{v}\in U} R_{\mathbf{M}}(\mathbf{v}) = \max_{\substack{\textit{subspace } U \\ \dim U = k}} \min_{\mathbf{v}\in U} R_{\mathbf{M}}(\mathbf{v}).$$

Proof. I will show the first equality only, as the second is analogous. Note that we have already seen that there is a subspace U of dimension $n-k+1$ so that

$$\lambda_k = \max_{\mathbf{v}\in U} R_{\mathbf{M}}(\mathbf{u}),$$

this subspace is simply the orthogonal complement of the sum of the eigenlines corresponding to the largest $k-1$ eigenvalues. The result will now follow if we verify that, for all subspaces U of dimension $n-k+1$, we have

$$\lambda_k \leq \max_{\mathbf{v}\in U} R_{\mathbf{M}}(\mathbf{u}).$$

To see this, let U be a subspace of dimension $n-k+1$, and let V be the sum of the eigenlines corresponding to the largest k eigenvalues. As $\dim U + \dim V$ exceeds n, it follows that $U \cap V \neq \emptyset$. Let $\mathbf{v}$ belong to this intersection. Then

$$R_{\mathbf{M}}(\mathbf{v}) \geq \lambda_k \sum_{r=1}^{k} \mathbf{v}^T E_r \mathbf{v} \geq \lambda_k,$$

as we wanted. □

Exercise 4.3. We've seen that $\Delta \geq \lambda_1$, the largest eigenvalue of $\mathbf{A}$. If the $(d_1, ..., d_n)$ is the degree sequence in decreasing order, then you can now show that $d_i \geq \lambda_i$.

Such min-max formula provides an alternative and meaningful definition of eigenvalues. For graph theory, it is hard to find interesting applications of this formula by itself. We can use it however to prove a strong result.

Theorem 4.4 (Cauchy's Interlacing)**.** *Let $\mathbf{A}$ be a symmetric $n\times n$ matrix and $\mathbf{S}$ be an $n\times m$ matrix satisfying $\mathbf{S}^T\mathbf{S} = \mathbf{I}$. Let $\mathbf{B} = \mathbf{S}^T\mathbf{A}\mathbf{S}$. Let $\theta_1 \geq ... \geq \theta_n$ be the eigenvalues of $\mathbf{A}$ and $\lambda_1 \geq ... \geq \lambda_m$ be those of $\mathbf{B}$. Then*

(a) For all k with $1 \leq k \leq m$,

$$\theta_{n-(m-k)} \leq \lambda_k \leq \theta_k$$

(b) If equality holds in either of the inequalities above for some λ_k eigenvalue of $\mathbf{B}$, then there is a λ_k-eigenvector $\mathbf{v}$ of $\mathbf{B}$ so that $\mathbf{S}\mathbf{v}$ is an eigenvector for λ_k in $\mathbf{A}$.

(c) Let $\mathbf{v}_1, ..., \mathbf{v}_m$ be an orthogonal basis of eigenvectors of $\mathbf{B}$, with $\mathbf{v}_i$ corresponding to λ_i. If for some $\ell \in \{1, ..., m\}$ we have that $\lambda_k = \theta_k$ for all $k = 1, ..., \ell$ (or $\lambda_k = \theta_{n-(m-k)}$ for all $k = \ell, ..., n$), then $\mathbf{S}\mathbf{v}_k$ is an θ_k eigenvector for $\mathbf{A}$ for $k = 1, ..., \ell$ (respectively for $k = \ell, ..., n$).

(d) If there is an $\ell \in \{1, ..., m\}$ so that $\lambda_k = \theta_k$ for all $k = 1, ..., \ell$, and $\lambda_k = \theta_{n-(m-k)}$ for all $k = \ell+1, ..., n$, then $\mathbf{S}\mathbf{B} = \mathbf{A}\mathbf{S}$. In this case, interlacing is called tight.

Proof. Let $\mathbf{u}_1, ..., \mathbf{u}_n$ be the eigenvectors of $\mathbf{A}$ corresponding to the θ_ks. The key thing now is to observe that, for all k, the subspace

$$\langle \mathbf{v}_1, ..., \mathbf{v}_k \rangle \cap \langle \mathbf{S}^T\mathbf{u}_1, ..., \mathbf{S}^T\mathbf{u}_{k-1} \rangle^{\perp}$$

contains at least one vector. Let $\mathbf{w}$ be such vector, which, in particular, implies $\mathbf{Sw} \in \langle \mathbf{u}_1, ..., \mathbf{u}_{k-1} \rangle^\perp$. Then, by Lemma 4.2, we have

$$\theta_k \geq \frac{(\mathbf{Sw})^T \mathbf{A}(\mathbf{Sw})}{(\mathbf{Sw})^T(\mathbf{Sw})} \geq \frac{\mathbf{w}^T\mathbf{Bw}}{\mathbf{w}^T\mathbf{w}} \geq \lambda_k.$$

If $\theta_k = \lambda_k$, then $\mathbf{w}$ and $\mathbf{Sw}$ are eigenvectors for $\mathbf{B}$ and $\mathbf{A}$ respectively. Item (iii) follows easily by induction. Finally, with tight interlacing, we can guarantee that $\mathbf{Sv}_1, ..., \mathbf{Sv}_m$ are all eigenvectors for $\mathbf{A}$ with the same eigenvalues they have in $\mathbf{B}$. Therefore $\mathbf{SBv}_k = \mathbf{ASv}_k$ for all k, and as the set of eigenvectors form a basis, the two matrices are equal. □

The basic principle for applying interlacing is to carefully chose the matrix $\mathbf{S}$.

Exercise 4.5. Let $\mathbf{A}$ be a $n \times n$ symmetric matrix, with eigenvalues $\theta_1 \geq ... \geq \theta_n$. Let $\mathbf{B}$ be a principal submatrix, size $(n-1) \times (n-1)$, with eigenvalues $\lambda_1 \geq ... \geq \lambda_{n-1}$. Show that, for all k,

$$\theta_{k+1} \leq \lambda_k \leq \theta_k.$$

(This is actually the reason why the result above is called interlacing.)

The stability number of a graph is the size of the largest subset of vertices which contains no edge inside of it — known as an independent or stable set.

Exercise 4.6. Let $\alpha(G)$ be the stability number of a graph. Then

$$\alpha(G) \leq |\{k : \theta_k \geq 0\}| \quad \text{and} \quad \alpha(G) \leq |\{k : \theta_k \leq 0\}|.$$

This follows easily if you note that an independent set corresponds to a block of 0s in $\mathbf{A}(G)$. Write the details.

Consider now the Petersen graph. By using interlacing, we will show two interesting facts about it. The Petersen graph has eigenvalues $3, 1^{(5)}, -2^{(4)}$. If its incidence matrix is $\mathbf{N}$, then the adjacency matrix is $\mathbf{NN}^T - 3\mathbf{I}$, and the adjacency matrix of its line graph is $\mathbf{N}^T\mathbf{N} - 2\mathbf{I}$.

Exercise 4.7. Find the spectrum of its line graph.

If the Petersen graph contains a Hamilton cycle, then its line graph contains an induced cycle C_{10}. This means that we can delete 5 vertices of its line graph, and find C_{10}. The eigenvalues of C_{10} are

$$2, \pm\left(\frac{1 \pm \sqrt{5}}{2}\right), -2$$

whereas 2 and -2 are simple, and the others each have multiplicity 2.

Exercise 4.8. Use interlacing now to show that the Petersen graph does not have a Hamilton cycle.

Finally, an application of the method of finding a vector in an intersection of subspace by looking at the dimension. The graph K_{10} has 45 edges, and therefore it would be possible for the edges of K_{10} to be partitioned into copies of the Petersen graph. However, this is not possible. To see that, assume K_{10} contains already two disjoint copies of Pete, say G and H. The eigenspace corresponding to 1 in G has dimension 5, and the same for that of 1 in H. As both eigenspaces are orthogonal to the line spanned by $\mathbf{1}$, then there must be at least one vector, say $\mathbf{w}$, who is simultaneously an eigenvector for $\mathbf{A}(G)$ and $\mathbf{A}(H)$. Thus

$$\mathbf{A}(\mathbf{J}-\mathbf{I}-\mathbf{A}(G)-\mathbf{A}(H))\mathbf{w} = -3\mathbf{w}.$$

This means that the complement of G and H in K_{10} has eigenvalue -3, and therefore cannot be isomorphic to the Petersen graph.

4.2 Partitions - cliques, cocliques, colourings

Consider a partition of the vertex set of a graph G, with characteristic matrix $\mathbf{P}$ (meaning, rows of $\mathbf{P}$ are vertices, and columns are parts of the partition, with 1s and 0s indicating whether a vertex belongs or not to a part).

A partition of the vertex set of a graph with characteristic matrix $\mathbf{P}$ is called equitable with respect to $\mathbf{A}(G)$ if the column space of $\mathbf{P}$ is $\mathbf{A}(G)$-invariant, that is, if there is a matrix $\mathbf{B}$ so that

$$\mathbf{AP} = \mathbf{PB}.$$

Combinatorially, this means that the number of neighbours a vertex a has in a class C of the partition is determined uniquely by the class that contains a. In other words, any two vertices in a given class have the same number of neighbours in any other given class (including their own).

All graphs contain at least one equitable partition of its vertex set: that in which all classes are singletons. If the graph is regular, then the partition that contains only one class is also equitable.

Given any partition with characteristic matrix $\mathbf{P}$, it is always possible to scale each column of $\mathbf{P}$ so that it becomes a normal vector. If $\mathbf{S}$ is the matrix obtained in this manner, it is immediate to verify that

$$\mathbf{S}^T\mathbf{S} = \mathbf{I},$$

and therefore $\mathbf{S}$ is suitable to be used as in Theorem 4.4.

For a first example of this property, we derive another bound to the independence number of a graph.

Corollary 4.9 (Ratio bound for independent sets)**.** *Let G be k-regular on n vertices, with smallest eigenvalue θ_n. Then*

$$\alpha(G) \leq \frac{n(-\theta_n)}{k-\theta_n}.$$

If equality holds, then the partition of the vertex set into any maximum independent set and its complement is equitable, and in particular, there is a τ eigenvector which is constant in each class of this partition.

Proof. Let $\mathbf{P}$ be the characteristic matrix of a partition that contains two class: one is a maximum independent set, and the other is its complement. Let $\mathbf{S}$ be the normalized characteristic matrix. Then

$$\mathbf{S}^T\mathbf{A}\mathbf{S} = \begin{pmatrix} 0 & \frac{\alpha k}{\sqrt{\alpha}\sqrt{n-\alpha}} \\ \frac{\alpha k}{\sqrt{\alpha}\sqrt{n-\alpha}} & \frac{(n-\alpha)k-k\alpha}{n-\alpha} \end{pmatrix} = \begin{pmatrix} 0 & \frac{\sqrt{\alpha}k}{\sqrt{n-\alpha}} \\ \frac{\sqrt{\alpha}k}{\sqrt{n-\alpha}} & k - \frac{k\alpha}{n-\alpha} \end{pmatrix}.$$

Clearly, the eigenvalues are k and $(-k\alpha)/(n-\alpha)$. Due to interlacing, it follows that

$$(-k\alpha)/(n-\alpha) \geq \theta_n,$$

which rearranges to

$$\alpha \leq \frac{n(-\theta_n)}{k-\theta_n}.$$

If you can't compute the eigenvalues easily, you can simply compare the determinant of $\mathbf{S}^T\mathbf{A}\mathbf{S}$ with the product of the largest and smallest eigenvalues of $\mathbf{A}$.

If equality holds, and because the largest eigenvalue of $\mathbf{A}$ and $\mathbf{S}^T\mathbf{A}\mathbf{S}$ are also equal, we have that (iv) in Theorem 4.4 applies. Moreover, (ii) of said theorem implies the assertion about the τ-eigenvector. □

It is quite surprising that this bound is met in several interesting cases, although it is not a good approximation for α in the general case (no such hope exists).

Exercise 4.10. Let δ be the smallest degree of G. If G is any graph (not necessarily regular), with largest eigenvalue θ_1 and smallest eigenvalue θ_n. Show that

$$\alpha \leq \frac{n(-\theta_1\theta_n)}{\delta^2 - \theta_1\theta_n}.$$

Hint: let k be the average degree in the independent set, and proceed as above.

Exercise 4.11. Let G be k-regular on n vertices, with eigenvalues $\theta_1 \geq ... \geq \theta_n$. Assume G contains an induced subgraph H with n' vertices and m' edges. Show that

$$\theta_2 \geq \frac{2m'n - (n')^2k}{n'(n-n')} \geq \theta_n.$$

Characterize what happens if equality holds in either side.

Exercise 4.12. Let $\omega(G)$ be the size of a maximum clique in G, that is, the size of the largest subgraph of G which is isomorphic to a complete graph. Assume G is k-regular. Find an upper bound to ω using the eigenvalues of G.

We now devote our attention to the chromatic number of G. A colouring of $V(G)$ is an assignment of colours to the vertex set of G so that any two neighbours receive different colours. It is always possible to colour a graph with n colours. A graph is 2-colourable if and only if it is bipartite. The chromatic number of a graph $\chi(G)$ is the minimum number of colours necessary to colour the vertices of G.

Just like α and ω, χ is hard to approximate, so any simple formulas using the spectrum of G can only bound, and even so not that well in the general case. However, this is quite significant to the best one could do.

Exercise 4.13. Explain why

$$\alpha \cdot \chi \geq n \quad \text{and also} \quad \omega \leq \chi.$$

The first inequality in the exercise above immediately implies a spectral lower bound to χ in regular graphs, using the upper bound to α. As it turns out, we can ignore the requirement of the graph to be regular.

Theorem 4.14 (Hoffman)**.** *Let G be a graph with chromatic number χ, largest eigenvalue θ_1 and smallest eigenvalue θ_n. Then*

$$\chi(G) \geq 1 - \frac{\theta_1}{\theta_n}.$$

Proof. Let $\mathbf{P}$ be the characteristic matrix of a colouring. To prove this result, it won't be enough to simply scale the columns of $\mathbf{P}$ and proceed with interlacing (in fact, try to do this). Instead, we shall first scale the rows of $\mathbf{P}$. Let $\mathbf{D}$ be a diagonal matrix whose diagonal entries are taken from the Perron eigenvector $\mathbf{v}$ of G. Let $\mathbf{S}$ be the obtained from $\mathbf{DP}$ upon multiplying from the right by a diagonal matrix $\mathbf{E}$ which effects to normalizing its columns. Thus $\mathbf{S}^T\mathbf{S} = \mathbf{I}$, and we proceed with interlacing now. We have $\mathbf{B} = \mathbf{S}^T\mathbf{AS}$ with 0s in the diagonal, as the support of $\mathbf{S}$ corresponds to a colouring of G. Note that $\mathbf{B}$ is $m \times m$ with $m = \chi$. We also note that θ is an eigenvalue of $\mathbf{B}$, because $\mathbf{S}^T\mathbf{AS}(\mathbf{E}^{-1}\mathbf{1}) = \mathbf{S}^T\mathbf{Av} = \theta_1\mathbf{E}^{-1}\mathbf{1}$. Hence, by interlacing,

$$0 = \operatorname{tr}\mathbf{B} = \lambda_1 + \lambda_2 + ... + \lambda_m \geq \theta_1 + (\chi - 1)\theta_n.$$

□

Exercise 4.15. What can you say if equality holds in this bound?

Note that in the last line of the proof, our bound was quite crude. An immediate improvement is to say

Corollary 4.16. *Let G be a graph with chromatic number χ, and eigenvalues $\theta_1 \geq ... \geq \theta_n$. Then*

$$\theta_1 + \theta_n + \theta_{n-1} + ... + \theta_{n-(\chi-2)} \leq 0.$$

□

Exercise 4.17. In this exercise, you will show that if $\theta_2 > 0$, then

$$\chi(G) \geq 1 - \frac{\theta_{n-\chi+1}}{\theta_2}.$$

I will give you a hint. Let $\mathbf{P}$ be the partition matrix of an optimal colouring. Let $\mathbf{v}_1$ be the Perron eigenvector, and $\mathbf{D}$ the diagonal matrix which contains its entries in the diagonal. Consider

$$\ker(\mathbf{P}^T\mathbf{D}) \cap \langle \mathbf{v}_n, ..., \mathbf{v}_{n-\chi+1} \rangle.$$

Prove that this intersection contains a vector, define a diagonal matrix with this vector, and also define $\mathbf{A}' = \mathbf{A} - (\theta_1 - \theta_2)\mathbf{v}_1\mathbf{v}_1^T$. Now proceed as in the proof of Hoffman's theorem.

Exercise 4.18. If m_n is the multiplicity of θ_n, verify now that

$$\chi \geq \min\{1 + m_m, 1 - \frac{\theta_n}{\theta_2}\}.$$

It is quite remarkable that in the results above, while exploring the connection between the eigenvalues of $\mathbf{A}$ and the independence number of G or its chromatic number, nowhere the fact that the entries of $\mathbf{A}$ are restricted to 1s and 0s was used. The only real constraint is that the non-zero entries are restricted to positions corresponding to edges in the graph. In fact, one can vary the entries of $\mathbf{A}$, and as long as the eigenvalue expressions increase (for chromatic number) or decrease (for independence number), better bounds will be obtained. This leads to the interesting topic of applications of semidefinite programming to algebraic graph theory.

There is a famous upper bound for χ:

Theorem 4.19 (Brooks)**.** *Let G be a graph with maximum degree Δ. Then $\chi(G) \leq \Delta$, unless G is a complete graph or an odd cycle, in which cases $\Delta + 1$ colours suffice.* □

This is one of the classical theorems in graph theory. Its proof is certainly not trivial (only purely combinatorial proofs are known, so you will have to research that on your own). I am sure you remember that $\theta_1 \leq \Delta$. It turns out, we can somehow strengthen the statement of Brooks theorem for several graphs.

Theorem 4.20 (Wilf)**.** *If G is a graph with chromatic number χ and largest eigenvalue θ_1, then*

$$\chi \leq 1 + \theta_1.$$

Equality holds if and only if G is an odd cycle or the complete graph.

Proof. Let G' be a subgraph of G which is χ-critical, meaning, the subgraph whose removal of any vertex decreases the chromatic number. In this subgraph, the degree of any vertex is at least $\chi - 1$ (why?). Thus its largest eigenvalue is at least $\chi - 1$. By interlacing, the largest eigenvalue of G is at least $\chi - 1$. □

Exercise 4.21. Finish the proof of the theorem above.

Exercise 4.22. One final exercise here: show that $\theta_n \geq n/2$, for any connected graph.

Problem 4.1. *Find a spectral proof of Brooks theorem.*

4.3 Other eigenvalues

The second largest eigenvalue of $\mathbf{A}^2$ seems to be related to how "random" the graph looks like. This is very vague, I know, reason why it will be better explained looking at the results. Let us assume G is a regular graph of degree k, with eigenvalues $\theta_1 \geq \theta_2 \geq ... \geq \theta_n$. Let $\lambda > 0$ be such that λ^2 is the second largest eigenvalue of A^2, that is, $\lambda = \max\{|\theta_2|, |\theta_n|\}$.

Theorem 4.23. *Let G be k-regular, and λ as before. Let S and T be two subsets of $V(G)$, of respective sizes s and t. Let $e(S,T)$ be the number of edges from S to T. Then*

$$\left|e(S,T) - \frac{kst}{n}\right| \leq \lambda\sqrt{st\left(1-\frac{s}{n}\right)\left(1-\frac{t}{n}\right)} \leq \lambda\sqrt{st}.$$

Proof. Let

$$\mathbf{A} = \sum_{r=1}^{n} \theta_i E_i$$

be the spectral decomposition of $\mathbf{A}$ into 1-dimensional eigenspaces. Let χ_S and χ_T be the characteristic vectors of sets S and T. It follows that

$$e(S,T) = \chi_S^T \mathbf{A} \chi_T = \sum_{r=1}^{n} \theta_i (\chi_S^T E_i \chi_T).$$

Note that $(\chi_S^T E_1 \chi_T) = st/n$ therefore

$$\left|e(S,T) - \frac{kst}{n}\right| = \left|\sum_{r=2}^{n} \theta_i (\chi_S^T E_i \chi_T)\right| \leq \lambda \sum_{r=2}^{n} |\chi_S^T E_i \chi_T|.$$

By Cauchy Schwarz (applied twice),

$$\begin{aligned}\left|e(S,T) - \frac{kst}{n}\right| &\leq \lambda \sum_{r=2}^{n} \sqrt{\chi_S^T E_i \chi_S}\sqrt{\chi_T^T E_i \chi_T} \\ &\leq \sqrt{\sum_{r=2}^{n} \chi_S^T E_i \chi_S}\sqrt{\sum_{r=2}^{n} \chi_T^T E_i \chi_T} \\ &= \sqrt{s - \frac{s^2}{n}}\sqrt{t - \frac{t^2}{n}}.\end{aligned}$$

□

That is, if λ is small compared to k, then between any two subsets of vertices of the graph, the number of edges tends to be the "expected" number, had every edge been put randomly and independently in the graph.

Exercise 4.24. Can you use the result above (or its proof method?) to show the ratio bound for cocliques without going through interlacing?

If $u \in V(G)$, let $N(u)$ denote the neighbourhood of u.

Exercise 4.25. Let T be a subset of $V(G)$ of size t. Show that

$$\sum_{u \in V(G)} \left(|N(u) \cap T| - \frac{kt}{n}\right)^2 \leq \frac{t(n-t)}{n}\lambda^2.$$

From Theorem 4.23, you would indeed expect that a large ratio k/λ implies that the graph "looks" random. If that is indeed the case, the diameter would also be relatively small. We can turn this intuition into a result.

Theorem 4.26. *Let G be a k-regular connected graph on n vertices, $n \geq 2$. Let d be its diameter, $\theta_1 = k$ its largest eigenvalue, and $\lambda = \max\{\theta_2, |\theta_n|\}$. Assume G is not bipartite, thus $\lambda < k$. Then*

$$d \leq \left\lfloor \frac{\log(n-1)}{\log(k/\lambda)} \right\rfloor + 1.$$

Proof. Let m be an integer, with

$$m > \frac{\log(n-1)}{\log(k/\lambda)}.$$

Thus, $k^m > (n-1)\lambda^m$. We will show that this implies that all entries of A^m are positive, therefore $d \leq m$. To see that, we will again apply Cauchy-Schwarz twice. Note that

$$\begin{aligned}
(A^m)_{a,b} = \mathbf{e}_a^T A^m \mathbf{e}_b &= \frac{k^m}{n} + \sum_{r=2}^{n} \theta_r^m (\mathbf{e}_a^T \mathbf{E}_r \mathbf{e}_b) \\
&\geq \frac{k^m}{n} - \lambda^m \sum_{r=2}^{n} |\mathbf{e}_a^T \mathbf{E}_r \mathbf{e}_b| \\
&\geq \frac{k^m}{n} - \lambda^m \sum_{r=2}^{n} \sqrt{\mathbf{e}_a^T \mathbf{E}_r \mathbf{e}_a} \sqrt{\mathbf{e}_b^T \mathbf{E}_r \mathbf{e}_b} \\
&\geq \frac{k^m}{n} - \lambda^m \sqrt{\sum_{r=2}^{n} \mathbf{e}_a^T \mathbf{E}_r \mathbf{e}_a} \sqrt{\sum_{r=2}^{n} \mathbf{e}_b^T \mathbf{E}_r \mathbf{e}_b} \\
&= \frac{k^m}{n} - \lambda^m \left(1 - \frac{1}{n}\right).
\end{aligned}$$

This last term is positive if $k^m > (n-1)\lambda^m$. □

We now proceed to show how to associate the third eigenvalue of a graph to matchings. First, a result due to Tutte whose proof is purely combinatorial, and therefore we skip:

Theorem 4.27. *A graph G has no perfect matching if and only if there is a subset $S \subseteq V(G)$ so that the subgraph of G induced by $V \backslash S$ has more than $|S|$ odd components (that is, a connected component with an odd number of vertices).*

(Note however that one direction of the Theorem is very easy to show).

Again, we will be dealing with regular graphs (for the last time).

Theorem 4.28. *A connected k-regular graph G on n vertices, n even, and eigenvalues $\theta_1 \geq ... \geq \theta_n$, has a perfect matching if*

$$\theta_3 \leq \begin{cases} k - 1 + \frac{3}{k+1} & \text{if } k \text{ is even,} \\ k - 1 + \frac{3}{k+2} & \text{if } k \text{ is odd.} \end{cases}$$

Proof. Assume there is no perfect matching. By (the difficult direction of) Tutte's theorem, there is a set S of size s so that $V\backslash S$ has at least $s+2$ odd components (why not $s+1$ only?). Let $G_1, ..., G_q$ be each of one these, each of size n_i. Then

$$\sum_{i=1}^{q} e(G_i, S) \leq ks.$$

As $s \geq 1$, $e(G_i, S) \geq 1$, this implies $e(G_i, S) < k$ and $n_i > 1$ for at least three values of i. Say $i = 1, 2, 3$, ordered in such way that the largest eigenvalues of $\mathbf{A}(G_i)$, say λ_i, satisfy $\lambda_1 \geq \lambda_2 \geq \lambda_3$. Upon taking the union of these three graphs, we find $\theta_3 \geq \lambda_3$.

We now look at G_3. We have that its average degree is

$$\partial_3 = \frac{2|E(G_3)|}{n_3} = \frac{kn_3 - e(G_3, S)}{n_3} = k - \frac{e(G_3, S)}{n_3}.$$

Note that $e(G_3, S) < k$, and $n_3 > 1$, so $k < n_3$. If k is even, $e(G_3, S)$ is even. If k is odd, then $k \leq n_3 - 2$. Thus

$$\partial_3 \geq \begin{cases} k - \frac{k-2}{k+1} & \text{if } k \text{ is even,} \\ k - \frac{k-1}{k+2} & \text{if } k \text{ is odd.} \end{cases}$$

As G_3 is not regular, because $e(G_3, S)$, we have $\partial_3 < \lambda_3 \leq \theta_3$, as wished. □

As we have seen over several previous results, the hypothesis of a graph being regular comes in very handy when dealing with the eigenvalues of the adjacency matrix. Our goal is to introduce another type of adjacency matrix that shall overcome this necessity.

4.4 Interlude — positive semidefinite matrices

A real matrix $\mathbf{M}$ is positive semidefinite if it satisfies the following properties:

- $\mathbf{M}$ is symmetric.
- $\mathbf{v}^T\mathbf{M}\mathbf{v} \geq 0$ for all $\mathbf{v}$.

If the inequality is strict for all non-zero $\mathbf{v}$, then $\mathbf{M}$ is called positive definite. The only thing we want now is a characterization.

This is probably one of the most famous "exercises" in linear algebra.

Theorem 4.29. *Let $\mathbf{M}$ be a symmetric matrix. The following are equivalent.*

(a) $\mathbf{M}$ is positive semidefinite.

(b) The eigenvalues of $\mathbf{M}$ are non-negative.

(c) There exists a matrix $\mathbf{B}$ so that $\mathbf{M} = \mathbf{B}^T\mathbf{B}$.

(d) For all positive semidefinite matrices $\mathbf{A}$, we have $\langle \mathbf{M}, \mathbf{A} \rangle \geq 0$.

Proof. Assume (a). Let $\mathbf{Mv} = \theta\mathbf{v}$. Then $0 \leq \mathbf{v}^T\mathbf{Mv} = \theta\mathbf{v}^T\mathbf{v}$, thus $\theta \geq 0$. Assume (b). We diagonalize $\mathbf{M}$ as

$$\mathbf{M} = \mathbf{P}^T\mathbf{DP}.$$

As $\mathbf{D} \geq 0$, we have

$$\mathbf{M} = \mathbf{P}^T\sqrt{\mathbf{D}}\sqrt{\mathbf{D}}\mathbf{P} = (\sqrt{\mathbf{D}}\mathbf{P})^T(\sqrt{\mathbf{D}}\mathbf{P}).$$

Assume (c). Then

$$\langle \mathbf{M}, \mathbf{A}\rangle = \operatorname{tr}\mathbf{MA} = \operatorname{tr}\mathbf{B}^T\mathbf{BA} = \operatorname{tr}\mathbf{BAB}^T.$$

As $\mathbf{A}$ is psd, we have $\operatorname{tr}\mathbf{BAB}^T \geq 0$. Finally, assume (d). Take $\mathbf{A} = \mathbf{vv}^T$, which is clearly psd for any $\mathbf{v}$. We have $0 \leq \langle \mathbf{M}, \mathbf{vv}^T\rangle = \mathbf{v}^T\mathbf{Mv}$, as wished. □

Exercise 4.30. Show that $\mathbf{M}$ is positive semidefinite if and only if its principal minors are non-negative (use interlacing?). Recall, a principal minor is a determinant of a square submatrix symmetric about the main diagonal.

4.5 The Laplacian matrix

Let G be a graph, and define $\mathbf{D}(G)$ to be the diagonal matrix whose entries correspond to the degrees of the vertices of G. Define the Laplacian matrix of G by

$$\mathbf{L} = \mathbf{L}(G) = \mathbf{D}(G) - \mathbf{A}(G).$$

Theorem 4.31. *The Laplacian matrix is positive semidefinite. Moreover, the multiplicity of* 0 *as an eigenvalue of* $\mathbf{L}$ *is equal to the number of connected components of* G*.*

Proof. To see this, assume G has been oriented, meaning, each edge has been assigned a direction, thus becoming an arc. Let $\mathbf{N}$ be the corresponding vertex by arc incidence matrix, so that an entry is 0 if the arc does not touch the vertex, $+1$ if the vertex is the head of the arc, and -1 if it is the tail. It is immediate to see that

$$\mathbf{L} = \mathbf{NN}^T.$$

(Note that this does not depend on the choice for the orientation.)

Following, $\mathbf{N}^T\mathbf{v} = 0$ if and only if $\mathbf{Lv} = 0$. It is immediate to see that $\mathbf{N}^T\mathbf{v} = 0$ if and only if $\mathbf{v}$ is constant on each connected component of G, whence the result follows (and describes essentially the unique eigenvector for 0 in a connected graph — the constant vector). □

Exercise 4.32. Assume G is regular, and let $\theta_1 \geq ... \geq \theta_n$ be the eigenvalues of $\mathbf{A}(G)$, with corresponding eigenbasis $\{\mathbf{v}_1, ..., \mathbf{v}_n\}$. Find an expression of the eigenvalues of $\mathbf{L}(G)$, and find a corresponding eigenbasis.

Exercise 4.33. Let $0 = \lambda_1 \leq ... \leq \lambda_n$ be the eigenvalues of $\mathbf{L}(G)$. Find the eigenvalues of $\mathbf{L}(\overline{G})$. Use this exercises to find the eigenvalues of $\mathbf{L}(K_{n,m})$ (this is the complete bipartite graph with n vertices on one side and m on the other).

As we have seen, $\mathbf{L}(G)$ is positive semidefinite. It follows that $\mathbf{x}^T\mathbf{L}\mathbf{x} \geq 0$ for all $\mathbf{x} \in \mathbb{R}^n$. We can moreover find a useful and meaningful expression for this. As $\mathbf{L} = \mathbf{N}\mathbf{N}^T$ where $\mathbf{N}$ is the incidence matrix of an orientation of the graph, it follows that

$$\mathbf{x}^T\mathbf{L}\mathbf{x} = (\mathbf{N}^T\mathbf{x})^T(\mathbf{N}^T\mathbf{x}) = \sum_{uv\in E(G)} (\mathbf{x}_u - \mathbf{x}_v)^2.$$

Exercise 4.34. Assume G is connected, on n vertices, and let λ_2 be its second smallest Laplacian eigenvalue. We certainly know (from the minimax principle for eigenvalues) that

$$\lambda_2 = \min_{\mathbf{v}\perp\mathbf{1}} \frac{\sum_{ab\in E(G)}(\mathbf{v}_a - \mathbf{v}_b)^2}{\sum_{a\in V(G)} \mathbf{v}_a^2}.$$

What I want you to show is that

$$\lambda_2 = \min_{\mathbf{v}\neq\alpha\mathbf{1}} \frac{n\sum_{ab\in E(G)}(\mathbf{v}_a - \mathbf{v}_b)^2}{\sum_{a<b}(\mathbf{v}_a - \mathbf{v}_b)^2}.$$

(The minimum is simply being taken over all vectors which are just not constant.)

Also, and without much difficulty now, prove that

$$\lambda_n = \max_{\mathbf{v}\neq\alpha\mathbf{1}} \frac{n\sum_{ab\in E(G)}(\mathbf{v}_a - \mathbf{v}_b)^2}{\sum_{a<b}(\mathbf{v}_a - \mathbf{v}_b)^2}.$$

Exercise 4.35. Revisit the first few subsections of this section and prove analogous results using the eigenvalues of $\mathbf{L}$ instead of those of $\mathbf{A}$.

4.6 Trees

A spanning tree of a connected graph G on n vertices is a subset of its edges that connects all vertices without forming any cycle. Necessarily, any spanning tree will contain $n-1$ edge.

A first result we shall see about the Laplacian matrix is actually a quite surprising one. We learned some weeks ago how to count how many spanning trees K_n has (n^{n-2}). Today we shall see that we can actually efficiently count how many spanning trees any graph has.

Let $\tau(G)$ denote the number of spanning trees G has. Recall the notation for edge deletion and contraction: $G\backslash e$ is the graph G with e removed, and G/e is the graph G with e removed and its incident vertices identified.

Lemma 4.36. *For any graph G and edge e, we have*

$$\tau(G) = \tau(G\backslash e) + \tau(G/e).$$

Exercise 4.37. Why?

We can now state the Matrix-Tree Theorem (due to Kirchkoff).

Theorem 4.38. *Let G be a graph, Laplacian $\mathbf{L}$. Let $a \in V(G)$, and $\mathbf{L}[a]$ denote the submatrix of $\mathbf{L}$ obtained upon deleting row and column corresponding to a. Then*

$$\tau(G) = \det \mathbf{L}[a].$$

Proof. This will be a proof by induction on the number of edges. You should check a few base cases on your own. Let us now assume G has m edges, and the result holds for any graph on fewer edges. Let $e \in E(G)$, with $e = \{a, b\}$. In G/e, vertices a and b are identified — let c be the name they receive in this case. If we show that

$$\det \mathbf{L}(G)[a] = \det \mathbf{L}(G\backslash e)[a] + \det \mathbf{L}(G/e)[c],$$

then, by induction and the lemma above, we will be done. So this equality above is now our task. In computing $\det \mathbf{L}(G)[a]$, we will perform row expansion in the row corresponding to b. Note that all terms of this expansion coming from an off-diagonal position will appear exactly the same in $\det \mathbf{L}(G\backslash e)[a]$. The only problem is the diagonal position — it is one unit larger in $\mathbf{L}(G)[a]$ then in $\mathbf{L}(G\backslash e)[a]$. Now the submatrix corresponding to excluding row and column b from $\mathbf{L}(G\backslash e)[a]$ is precisely $\mathbf{L}(G/e)[c]$, that is

$$\det \mathbf{L}(G)[a] = \det \mathbf{L}(G\backslash e)[a] + \det \mathbf{L}(G\backslash e)[a, b] = \det \mathbf{L}(G\backslash e)[a] + \det \mathbf{L}(G/e)[c],$$

as wished. □

Exercise 4.39. Very easily now you can verify that the number of spanning trees on n vertices is n^{n-2}.

Exercise 4.40. Prove that the number of spanning trees of G that contain a given edge $e = ab$ is equal to $\det \mathbf{L}[a, b]$.

As we have all learned before, for any square matrix $\mathbf{M}$,

$$\mathbf{M} \operatorname{adj}(\mathbf{M}) = \det(\mathbf{M})\mathbf{I},$$

where $\operatorname{adj}(\mathbf{M})_{ij} = (-1)^{i+j} \det \mathbf{M}(j, i)$. As we have just seen from above, all diagonal entries of $\operatorname{adj} \mathbf{L}(G)$ are equal to $\tau(G)$.

However for any G, $\det \mathbf{L}(G) = 0$. If we now assume G is connected, we know that there is essentially only one eigenvector to the eigenvalue 0, thus the equality

$$\mathbf{L}(G) \ \operatorname{adj} \mathbf{L}(G) = \mathbf{0}$$

implies that all columns of $\operatorname{adj} \mathbf{L}(G)$ are constant, and therefore all entries of $\operatorname{adj} \mathbf{L}(G)$ are equal to $\tau(G)$. It is immediate to verify all comments above hold if G is disconnected, in which case $\tau(G) = 0$.

Corollary 4.41. *For any graph G, we have*

$$\operatorname{adj} \mathbf{L}(G) = \tau(G)\mathbf{J}.$$

□

Exercise 4.42. Prove that for any graph G with Laplacian eigenvalues $\lambda_1 \leq ... \leq \lambda_n$, it holds that

$$\tau(G) = \frac{1}{n} \prod_{i=2}^{n} \lambda_i.$$

Hint: let $\psi(x)$ be the characteristic polynomial of $\mathbf{L}$. Arrive at the result realizing that

$$\prod_{i=1}^{n} (x - \lambda_i) = \psi(x) = \det(x\mathbf{I} - \mathbf{L}).$$

4.7 Representation, springs and energy

A *representation of a graph* is a map $\rho : V(G) \to \mathbb{R}^m$ (you can think of it as an m-dimensional drawing of G). You can associate ρ to an $n \times m$ matrix $\mathbf{R}$ — each row is the image of the corresponding vertex. A representation ρ is called balanced if $\sum_{a \in V} \rho(a) = \mathbf{0}$, that is, if $\mathbf{1}^T\mathbf{R} = \mathbf{0}$. Upon assuming we can freely translate a representation, we can always assume it is balanced. Moreover, we shall also assume the columns of $\mathbf{R}$ are linearly independent (otherwise we simply look at the representation to the subspace of $\mathbb{R}^m$ and rewrite ρ upon a change of basis so that $\mathbf{R}$ has fewer columns).

Now imagine a physical model, in which the vertices have been placed in $\mathbb{R}^m$. Some of them, $U \subseteq V(G)$, are "nailed", some of them, $V(G) - U$, are free. The edges are *springs*. For now, identical springs, with spring constant 1. By Hooke's law, the force the spring between a and b exerts in a is equal to $\rho(b) - \rho(a)$. Note that a configuration is in equilibrium if and only if the net force at each vertex in $V - U$ is 0. This is equivalent to requiring, for all $a \in V - U$, that

$$\sum_{b \sim a} \rho(b) - \rho(a) = 0 \iff \deg(a)\rho(a) - \sum_{b \sim a} \rho(b) = \mathbf{0}.$$

In other words, $\mathbf{LR}$ must have a rectangle of 0s in the rows corresponding to the vertices in $V - U$. Once the entries of $\mathbf{R}$ corresponding to vertices in U have been determined, finding the remaining entries of $\mathbf{R}$ so that this holds is equivalent to solving m systems of equation whose coefficient matrix is $\mathbf{L}[U]$. All these systems have unique solutions if the graph is connected and $U \neq$, because $\mathbf{L}[U]$ is positive definite.

Exercise 4.43. Let $\mathbf{L}[U]$ denote the submatrix of $\mathbf{L}$ obtained upon removing rows and columns corresponding to the vertices in subset U. Assume the graph is connected, and U non-empty. Prove that all eigenvalues of $\mathbf{L}[U]$ are positive.

Exercise 4.44. Convince yourself that nothing really changes if we assume the spring between a and b to have spring constant ω_{ab}.

Physics also teaches us that vertices will settle in the position the minimizes the potential energy. The potential energy of a spring with constant ω and stretched to a length ℓ is $(1/2)\omega\ell^2$ (we will ignore the fraction). Thus, the potential energy of a configuration is

$$\mathcal{E}(\rho) = \sum_{ab \in E(G)} \omega_{ab} ||\rho(a) - \rho(b)||^2.$$

Let W be a diagonal matrix, indexed by $E(G)$, whose diagonal entry is equal to ω_{ab}. As before, let $\mathbf{N}$ be the incidence matrix of an orientation of G, and $\mathbf{R}$ the matrix of the representation. It is immediate to verify that

$$\mathcal{E}(\rho) = \operatorname{tr} \mathbf{R}^T\mathbf{N}\mathbf{W}\mathbf{N}^T\mathbf{R}.$$

Note that $\mathbf{NWN}^T$ is simply a weighted Laplacian (and if $\mathbf{W}$ has positive diagonal and the graph is connected, then $\mathbf{NWN}^T$ is positive semidefinite, and 0 is a simple eigenvalue with eigenvector $\mathbf{1}$).

A representation is called *orthogonal* if the columns of $\mathbf{R}$ are orthonormal. In this case, $\mathbf{R}^T\mathbf{R} = \mathbf{I}$. Requiring a representation to be orthogonal is also a way of imposing a shape. We do not need to "nail" vertices in this case, as the following theorem shows.

Theorem 4.45. *Let G be a graph, with a weighted Laplacian matrix Ł, with eigenvalues $0 = \lambda_1 < \lambda_2 \leq \lambda_3 \leq ... \leq \lambda_n$. The minimum energy of a balanced orthogonal representation into $\mathbb{R}^m$ is equal to*

$$\sum_{r=2}^{m+1} \lambda_r.$$

Proof. To any orthogonal representation into $\mathbf{R}^k$ whose first column is $\mathbf{1}$ corresponds a balanced orthogonal representation in $\mathbb{R}^{k-1}$ with the same energy, obtained upon ignoring this first column. Thus the minimum energy of a balanced orthogonal representation into $\mathbb{R}^m$ is equal to the minimum energy of an orthogonal representation into $\mathbb{R}^{m+1}$ whose first column is $\mathbf{1}$. Let $\mathbf{R}$ be the matrix of one such representation. Its energy is

$$\operatorname{tr} \mathbf{R}^T\mathbf{L}\mathbf{R},$$

which, by interlacing, is at least $\sum_{r=1}^{m+1} \lambda_r$. Recall that $\lambda_1 = 1$. Moreover, one representation meeting this energy exists: simply write the eigenvectors corresponding $\lambda_1, ..., \lambda_{m+1}$ as the columns of $\mathbf{R}$. □

An immediate consequence is a method to draw graphs in $\mathbb{R}^m$ that is balanced and will somehow look "rigid" and "having a volume" — the so called spring embedding. Just pick the eigenvectors of $\mathbf{L}$ corresponding to $\lambda_2, ..., \lambda_{m+1}$, line them up as columns of a matrix, and map each vertex to the corresponding row. You should try this method to draw graphs in $\mathbb{R}^2$ and $\mathbb{R}^3$ using your favourite software.

4.8 Electrical currents

We define a physical model of a graph in which each edge corresponds to a wire. Let us say each edge has weight ω_{ab}, and this will mean to us that its resistance is $1/\omega_{ab}$ (a small weight corresponds to a big resistance). Ohm's law says that the potential drop across a resistor is equal to the current flowing times the resistance. If the current from a to b is $i(a, b)$ and the potentials in a and b are $v(a)$ and $v(b)$, then

$$v(a) - v(b) = \frac{i(a, b)}{\omega_{ab}}.$$

If $\mathbf{N}$ is the incidence matrix of an orientation of G, $\mathbf{W}$ the diagonal matrix with edge weights, $\mathbf{v}$ the vector with vertex potentials, and $\mathbf{i}$ the vector of edge currents, we now have

$$\mathbf{i} = \mathbf{W}\mathbf{N}^T\mathbf{v}.$$

Let $\mathbf{j}$ be a vector indexed by vertices whose ath entry denotes the net current entering or leaving the network at a. Recall that by Kirchkoff's law, the current entering a node is equal to the current exiting. Thus

$$\mathbf{j} = \mathbf{N}\mathbf{i}.$$

All together, and again making $\mathbf{L} = \mathbf{N}\mathbf{W}\mathbf{N}^T$ the weighted Laplacian matrix, we have

$$\mathbf{j} = \mathbf{L}\mathbf{v}.$$

As a consequence of this fact, it must be that $\mathbf{1}^T\mathbf{j} = 0$.

On the other hand, assume now that we are given a vector indicating currents entering and leaving the network satisfying $\mathbf{1}^T\mathbf{j} = 0$. Is this enough to find the voltages that correspond to the system?

A solution for $\mathbf{v}$ can be found computing the pseudo-inverse $\mathbf{L}^+$ of $\mathbf{L}$. That is, if $\mathbf{L}$ has spectral decomposition

$$\mathbf{L} = \sum_{i=1}^{n} \lambda_i \mathbf{E}_i,$$

with $0 = \lambda_1 < \lambda_2 \leq ... \leq \lambda_n$, then, given $\mathbf{j}$, with $\mathbf{1}^T\mathbf{j} = 0$, a solution for $\mathbf{v}$ can be found as

$$\mathbf{v} = \left(\sum_{i=2}^{n} \frac{1}{\lambda_i} \mathbf{E}_i \right) \mathbf{j}.$$

(Note that if $\mathbf{v}$ is as above, then $\mathbf{v} + \alpha\mathbf{1}$ also satisfies $\mathbf{L}(\mathbf{v} + \alpha\mathbf{1}) = \mathbf{j}$ for any α).

Now assume a and b are neighbours, and imagine one unit of current is pushed into a, and one unit extracted from b (meaning: $\mathbf{j}_a = -\mathbf{j}_b = 1$, 0 elsewhere, or simply $\mathbf{j} = \mathbf{e}_a - \mathbf{e}_b$). We can solve which potential arrangement at all vertices allows for this, and the difference of potential between a and b is defined as their *effective resistance*. In other words

$$R_{\text{eff}}(a, b) = (\mathbf{e}_a - \mathbf{e}_b)^T \mathbf{L}^+ (\mathbf{e}_a - \mathbf{e}_b).$$

Exercise 4.46. Suppose you have two edges, ab and cd. Prove that the difference of potential between c and d when you push one unit of current at a a and remove it at b is the same as the difference of potential between a and b when you push one unit of current at a c and remove it at d.

4.9 Connectivity and interlacing

Again, $\mathbf{L}$ is the Laplacian matrix, and $\lambda_1 \leq \lambda_2 \leq ... \leq \lambda_n$ its eigenvalues.

Over the next few sections, we will learn that λ_2 carries powerful information about a graph. We start with a bound associating λ_2 to the connectivity. Given a graph G, it is k-vertex-connected if it has more than k vertices, and remains connected whenever fewer than k vertices are removed. The *vertex connectivity* of a graph, denoted $\kappa_0(G)$, is the largest k to so that G is k-vertex connected. For all graphs which are not complete, this definition is equivalent to saying that $\kappa_0(G)$ is the smallest size of a subset of vertices whose removal disconnects G.

Computing the vertex connectivity of a graph is not difficult — Menger's theorem says that the size of a minimum cut in a graph is equal to the maximum number of disjoint paths that can be found between any pair of vertices. I invite you to prove this result using linear programming duality.

Nevertheless, an eigenvalue bound can always be useful.

Theorem 4.47. *Suppose $U \subseteq V(G)$. Then*

$$\lambda_2(G) \leq \lambda_2(G\backslash U) + |U|.$$

Proof. Let $\mathbf{v}'$ be a normalized $\lambda_2(G\backslash U)$ eigenvector of $\mathbf{L}(G\backslash U)$. Let $\mathbf{v}$ be the extension of $\mathbf{v}'$ to $\mathbf{R}^{V(G)}$, adding 0s in the remaining entries. By Courant-Fisher-Weyl, we have

$$\begin{aligned}\lambda_2(G) &\leq \sum_{ab\in E(G)} (\mathbf{v}_a - \mathbf{v}_b)^2 \\ &\leq \sum_{a\in U}\sum_{b\sim a} \mathbf{v}_b^2 + \sum_{ab\in E(G\backslash U)} (\mathbf{v}_a - \mathbf{v}_b)^2 \\ &\leq |U| + \lambda_2(G\backslash U).\end{aligned}$$

□

If U is a cut-set, then $G\backslash U$ is disconnected, thus 0 has multiplicity bigger than 1, and therefore

$$\lambda_2(G) \leq \kappa_0(G).$$

This immediately implies that $\lambda_2(G) \leq \delta(G)$.

Exercise 4.48. Prove that for all trees on more than 2 vertices, $\lambda_2 \leq 1$. Prove that equality holds if and only if the tree is a star.

Interlacing "works" for $\mathbf{L}$, but the problem is that the submatrices of $\mathbf{L}$ are not Laplacian matrices of subgraphs. If we would like to related the eigenvalues of $\mathbf{L}(G)$ with those of the Laplacians of subgraphs, we must use different methods. The following exercise can be proved elementarily, just as we did above.

Exercise 4.49. Let G be a graph, and $e \in E(G)$. Prove that

$$\lambda_2(G\backslash e) \leq \lambda_2(G) \leq \lambda_2(G\backslash e) + 2.$$

Show that equality holds in the second bound if and only if G is complete.

4.10 Partitioning and cuts

Even though finding the minimum cut in a graph amount to an easy task, other problems involving cuts or partitions into clusters are significantly harder. We introduce three problems related to edge cuts:

(a) bipartition width: finding the minimum over all $e(U, \overline{U})$ where $U \subseteq V(G)$, and $|U| = \lfloor n/2 \rfloor$.

(b) maxcut: finding the maximum cut, meaning a non-empty propoer subset U of $V(G)$ so that $e(U, \overline{U})$ is maximized.

(c) finding the conductance, meaning, the minimum over all $e(U, \overline{U})/|U|$, with $U \subseteq V(G)$, $0 < |U| \leq n/2$.

These parameters are all NP-hard to compute, but we can find some interesting bounds or approximations using the eigenvalues λ_2 or λ_n. We start with an easy observation.

Lemma 4.50. *For all $U \subseteq V(G)$, we have*

$$\lambda_2 \frac{|U|(n-|U|)}{n} \leq e(U,\overline{U}) \leq \lambda_n \frac{|U|(n-|U|)}{n}.$$

Proof. Both bounds follow immediately from Exercise 4.34. □

This immediately leads to a lower bound to the bipartition width of the graph, caled $\mathrm{bw}(G)$. We have

$$\mathrm{bw}(G) \geq \frac{1}{4} n \lambda_2(G).$$

It also implies an immediate upper bound to the maxcut, labelled $\mathrm{mc}(G)$. We have

$$\mathrm{mc}(G) \leq \frac{1}{4} n \lambda_n(G).$$

Both these bounds can be made stronger by solving semidefinite programs. I won't get into details, but I will hint where in the expression we are allowed to put "new variables".

Theorem 4.51. *Let G be a graph, of even order n. Then*

$$bw(G) \geq \frac{1}{4} n \max_{\mathbf{v} \perp \boldsymbol{1}} \min_{\mathbf{u} \perp \boldsymbol{1}} \frac{\langle (\mathbf{L} + \boldsymbol{diag}(\mathbf{c}))\mathbf{u}, \mathbf{u} \rangle}{\langle \mathbf{u}, \mathbf{u} \rangle},$$

Proof. Let S be a set of cardinality $n/2$ with $e(S,\overline{S}) = \mathrm{bw}(G)$, and define $\mathbf{w} \in \mathbb{R}^V$ to be $+1$ in S and -1 in $\overline{S}$. Note that $\mathbf{w} \perp \mathbf{1}$. Also,

$$\langle \mathbf{diag}(\mathbf{v})\mathbf{w}, \mathbf{w} \rangle = 0.$$

Therefore

$$\begin{aligned} \frac{\langle (\mathbf{L} + \mathbf{diag}(\mathbf{v}))\mathbf{w}, \mathbf{w} \rangle}{\langle \mathbf{w}, \mathbf{w} \rangle} = \frac{\langle \mathbf{L}\mathbf{w}, \mathbf{w} \rangle}{\langle \mathbf{w}, \mathbf{w} \rangle} &= \frac{\sum_{ab \in E} (\mathbf{w}_a - \mathbf{w}_b)^2}{\sum_{a \in V} \mathbf{w}_a^2} \\ &= \frac{4e(S,\overline{S})}{n} = \frac{4}{n}\mathrm{bw}(G). \end{aligned}$$

□

Exercise 4.52. Let $\mathbf{Q}$ be a $n \times (n-1)$ matrix with orthonormal columns and $\mathbf{1}$ in its left kernel. Argue why we also have

$$\mathrm{bw}(G) \geq \frac{1}{4} n \max_{\mathbf{v} \perp \mathbf{1}} \lambda_1(\mathbf{Q}^T(\mathbf{L} + \mathbf{diag}(\mathbf{v})\mathbf{Q}).$$

Exercise 4.53. Prove that

$$\mathrm{mc}(G) \leq \frac{1}{4} n \min_{\mathbf{v} \perp \mathbf{1}} \lambda_n((\mathbf{L} + \mathbf{diag}(\mathbf{v})).$$

(Hint: it is similar to the Theorem above).

For the third parameter we defined, the conductance, denoted by $\Phi(G)$ and also called the isoperimetric number, Lemma 4.50 implies that

$$\Phi(G) \geq \lambda_2/2.$$

For this parameter, we can bound it from the other side as well.

Theorem 4.54. *Given a graph G, we have*

$$\Phi(G) < \sqrt{2\Delta(G)\lambda_2(G)}.$$

Proof. We consider a normalized eigenvector $\mathbf{v}$ for λ_2, and we assume without loss of generality that the vertices are ordered, meaning $V(G) = \{1, 2, ..., n\}$, in such a way that $\mathbf{v}_i \geq \mathbf{v}_{i+1}$ for all i. Let V_+ be the vertices with $\mathbf{v}_i > 0$, and assume $\mathbf{v}$ is signed so that $|V_+| \leq n/2$. Also, define $\mathbf{u}$ vector with $\mathbf{u}_i = \mathbf{v}_i$ if $\mathbf{v}_i > 0$, and $\mathbf{u}_j = 0$ otherwise. We finally define E_+ the set of edges incident to one vertex in V_+.

To each i, $1 \leq i \leq |V(G)|$, we consider the cut

$$C_i = \{\{j, k\} \in E(G) : 1 \leq j \leq i < k \leq n\}.$$

Let

$$\alpha = \min_{1 \leq i \leq n} \frac{|C_i|}{\min\{i, n-i\}},$$

whence $\alpha \geq \Phi(G)$. Let $\mathbf{P}$ be the projection onto the subspace spanned by the characteristic

vectors of the vertices in V_+, that is, $\mathbf{Pv} = \mathbf{u}$. We now have

$$\begin{aligned}
\lambda_2 = \frac{\mathbf{v}^T\mathbf{PLv}}{\mathbf{v}^T\mathbf{Pv}} &= \frac{(\mathbf{N}^T\mathbf{Pv})^T(\mathbf{N}^T\mathbf{v})}{\sum_{i\in V_+}\mathbf{v}_i^2} \\
&= \frac{\sum_{ij\in E_+}(\mathbf{u}_i-\mathbf{u}_j)(\mathbf{v}_i-\mathbf{v}_j)}{\sum_{i\in V_+}\mathbf{v}_i^2} \\
&> \frac{\sum_{ij\in E_+}(\mathbf{u}_i-\mathbf{u}_j)^2}{\sum_{i\in V_+}\mathbf{u}_i^2} \\
&= \frac{\sum_{ij\in E_+}(\mathbf{u}_i-\mathbf{u}_j)^2\sum_{ij\in E_+}(\mathbf{u}_i+\mathbf{u}_j)^2}{\sum_{i\in V_+}\mathbf{u}_i^2\sum_{ij\in E_+}(\mathbf{u}_i+\mathbf{u}_j)^2} \\
&\geq \frac{\left(\sum_{i\sim j}|\mathbf{u}_i^2-\mathbf{u}_j^2|\right)^2}{2\Delta\left(\sum_{i\in V_+}\mathbf{u}_i^2\right)^2} \\
&\geq \frac{\left(\sum_i|\mathbf{u}_i^2-\mathbf{u}_{i+1}^2||C_i|\right)^2}{2\Delta\left(\sum_{i\in V_+}\mathbf{u}_i^2\right)^2} \\
&\geq \frac{\left(\sum_i|\mathbf{u}_i^2-\mathbf{u}_{i+1}^2|\alpha i\right)^2}{2\Delta\left(\sum_{i\in V_+}\mathbf{u}_i^2\right)^2} \\
&\geq \frac{\alpha^2}{2\Delta} \\
&\geq \frac{\Phi^2}{2\Delta}
\end{aligned}$$

□

Exercise 4.55. Justify in details each of the steps of the inequality chain above.

Note that not only the result above gives a bound, but it also has an implicit algorithm in its proof. In fact, we were able to efficiently find a set of vertices U so that

$$\Phi \leq \frac{e(U,\overline{U})}{|U|} \leq \sqrt{2\Delta\lambda_2} \leq 2\sqrt{\Delta\Phi}.$$

4.11 Normalized Laplacian

As we now know,

$$\mathbf{L} = \mathbf{D} - \mathbf{A}.$$

We assume G has no isolated vertices. We define $\mathbf{Q}$ as

$$\mathbf{Q} = \mathbf{D}^{-1/2}\mathbf{L}\mathbf{D}^{-1/2} = \mathbf{I} - \mathbf{D}^{-1/2}\mathbf{A}\mathbf{D}^{-1/2}.$$

Note that it is positive semidefinite.

Exercise 4.56. Prove that

$$R_{\mathbb{Q}}(\mathbf{u}) = \frac{\sum_{ab\in E}(\mathbf{v}_a - \mathbf{v}_b)^2}{\sum_{a\in V}\mathbf{v}_a^2 d(a)},$$

where $\mathbf{v} = \mathbf{D}^{-1/2}\mathbf{u}$.

We also prove some basic properties.

Theorem 4.57. *Let G be a graph with no isolated vertices, and denote the eigenvalues of $\mathbf{Q}$ by $\mu_1 \leq ... \leq \mu_n$. Then*

(i) $\sum_j \mu_j = n$.

(ii) For $n \geq 2$, $\mu_2 \leq n/(n-1)$, and equality holds if and only if G is the complete graph. Also, $\mu_n \geq n/(n-1)$.

(iii) For a graph not complete, we have $\mu_2 \leq 1$.

(iv) The multiplicity of 0 as an eigenvalue is the number of connected components of G.

(v) We have $\mu_n \leq 2$, and equality holds if and only if G is bipartite. In this case, for all μ eigenvalue of $\mathbf{Q}$, $2-\mu$ is also eigenvalue.

Exercise 4.58. Prove the properties above.

Exercise 4.59. Let G be a graph, connected, diameter d. Prove that

$$\mu_2 \geq \frac{1}{d\sum_{a\in V} d(a)}.$$

4.12 Random Walks

Let G be a weighted graph (edge weights are given by a function ω). We assume a walker is sitting at a vertex, and then decides to move with certain probability. At each step, this walker hops from one vertex to another with probability proportional to the edge weight of the corresponding edge. This model is equivalent to a Markov chain defined on a finite measurable state space.

We will be specially interested in the expected behaviour of a random walker, rather than on some fixed particular instance of this experiment.

Let $\mathbf{p}_t \in \mathbb{R}^V$ denote the probability distribution of the walker at time t. As such, $\mathbf{p}_t(a) \geq 0$ for all $a \in V$, and

$$\sum_{a\in V}\mathbf{p}_t(a) = 1.$$

Because the probability of arriving at a at $t+1$ steps is only determined by the probability distribution at time t and the edge weights, we have

$$\mathbf{p}_{t+1}(a) = \sum_{b\sim a}\frac{\omega_{ab}}{d_b}\mathbf{p}_t(b).$$

with d_b meaning the sum of the weights of edges incident to b. Equivalently,

$$\mathbf{p}_{t+1} = \mathbf{A}\mathbf{D}^{-1}\mathbf{p}_t,$$

where $\mathbf{A}$ is the weighted adjacency matrix and $\mathbf{D}$ the diagonal matrix of (weighted) degrees. Let $\mathbf{W} = \mathbf{A}\mathbf{D}^{-1}$. It is immediate to verify that $\mathbf{p}_{t+k} = \mathbf{W}^k\mathbf{p}_t$, where $\mathbf{p}_0$ typically stands for the starting distribution. We also see that

$$\mathbf{D}^{-1/2}\mathbf{W}\mathbf{D}^{1/2} = \mathbf{I} - \mathbf{Q},$$

where $\mathbf{Q}$ is the normalized Laplacian. If there is a probability that the walker does not move, say $1/2$, we now have

$$\mathbf{p}_{t+1} = (1/2)\mathbf{I}\mathbf{p}_t + (1/2)\mathbf{A}\mathbf{D}^{-1}\mathbf{p}_t.$$

Let $\mathbf{Z} = (1/2)(\mathbf{I} + \mathbf{A}\mathbf{D}^{-1})$. You can now see that

$$\mathbf{D}^{-1/2}\mathbf{Z}\mathbf{D}^{1/2} = \mathbf{I} - (1/2)\mathbf{Q},$$

Matrices $\mathbf{Z}$ or $\mathbf{W}$ are not symmetric, but they are both similar to a symmetric matrix. This gives that they are diagonalizable with real eigenvectors.

Exercise 4.60. If $\mathbf{v}$ is λ-eigenvector of $\mathbf{Q}$, to which eigenpair of $\mathbf{W}$ or $\mathbf{Z}$ they relate? Later, prove that all eigenvalues of $\mathbf{W}$ are between -1 and 1, and those of $\mathbf{Z}$ lie between 0 and 1.

Exercise 4.61. What do you obtain if you replace $1/2$ by another probability?

A random walk $\mathbf{W}$ converges to a distribution $\mathbf{p}$ if for any given ε and any distribution $\mathbf{q}$, there is an n so that

$$||\mathbf{W}^n\mathbf{q} - \mathbf{p}|| < \varepsilon.$$

Exercise 4.62. Can you show that if $\mathbf{W}$ converges to $\mathbf{p}$, then $\mathbf{p}$ is "stable", meaning, $\mathbf{W}\mathbf{p} = \mathbf{p}$? Later, prove that every graph contains a stable distribution, and if the graph is connected, this is unique (it also does not depend on the probability of staying put).

Exercise 4.63. Show that if there is any probability that a walker stays put (we call these random walks "lazy"), then the random walk will converge to the stable distribution. Describe the graphs for which a non-lazy random walk does not converge.

Example 4.64. Imagine now the following experiment. A deck of n cards c_1,...,c_n is lying on a table. We will shuffle these cards in a very stupid way: at each time step, we select i, j from 1 to n uniformly at random and exchange the positions of cards i and j (that includes choosing $i = j$ and doing nothing). How fast does this procedure produces a good shuffling?

The graph here is the one whose vertex set corresponds to the permutations on n elements. Vertices adjacent if one can be obtained from the other by applying a transposition (this is called the Cayley Graph $\mathrm{Cay}(S_n, T)$).

The weights here are simply determined:

$$\mathbf{p}_{t+1} = (1/n)\mathbf{I}\mathbf{p}_t + [(n-1)/n]\mathbf{A}\mathbf{D}^{-1}\mathbf{p}_t,$$

with $\mathbf{A}_{\sigma\tau} = 1$ if $\sigma\tau^{-1}$ is a transposition, and $= 0$ otherwise.

We thus want to know how fast $\mathbf{p}_t$ becomes the stable distribution.

Exercise 4.65. What is the stable distribution of the example above?

We can now provide a reasonably good estimate.

Theorem 4.66. *Let* $\mathbf{W}$ *be the transition matrix of a random walk, having laziness probability* $1-\rho$*, and with eigenvalues* $1 = \omega_1 \geq \omega_2 \geq ... \geq \omega_n$*. Assume* $\omega = \max\{|\omega_2|, |\omega_n|\} < 1$*, meaning, the random walk converges, say, to* $\mathbf{p}$*. Let* $\mathbf{p}_0 = \mathbf{e}_a$*, meaning, the walk starts at* a*. Then*

$$|\mathbf{p}_t(b) - \mathbf{p}(b)| < \sqrt{\frac{d(b)}{d(a)}}\omega^t.$$

Proof. Let

$$\mathbf{Q} = \sum_{i=1}^{n} \lambda_i \mathbf{F}_i$$

be the spectral decomposition of the normalized Laplacian. Then

$$\mathbf{W} = \mathbf{I} - \rho\mathbf{D}^{1/2}\mathbf{Q}\mathbf{D}^{-1/2} = \sum_{i=1}^{n}(1-\rho\lambda_i)\mathbf{D}^{1/2}\mathbf{F}_i\mathbf{D}^{-1/2} = \sum_{i=1}^{n}\omega_i\mathbf{E}_i.$$

Note that

$$\mathbf{E}_1 = \frac{1}{\sum_{a\in V} d(a)}\mathbf{D}^{1/2}\mathbf{D}^{1/2}\mathbf{1}\mathbf{1}^T\mathbf{D}^{1/2}\mathbf{D}^{-1/2} = \frac{1}{\sum_{a\in V} d(a)}\mathbf{D}\mathbf{1}\mathbf{1}^T.$$

Then

$$\begin{aligned}
|\mathbf{p}_t(b) - \mathbf{p}(b)| &= |\mathbf{e}_b^T\mathbf{W}^t\mathbf{e}_a - \mathbf{e}_b^T\mathbf{E}_1\mathbf{e}_b| \\
&= \left|\sum_{i=2}^{n}\omega_i^t(\mathbf{e}_b^T\mathbf{E}_i\mathbf{e}_a)\right| \\
&\leq \omega^t\sum_{i=2}^{n}\left|\mathbf{e}_b^T\mathbf{E}_i\mathbf{e}_a\right| \\
&\leq \omega^t\sqrt{\frac{d(b)}{d(a)}}\sum_{i=2}^{n}\left|\mathbf{e}_b^T\mathbf{F}_i\mathbf{e}_a\right| \\
&\leq \omega^t\sqrt{\frac{d(b)}{d(a)}}\sum_{i=2}^{n}\sqrt{\mathbf{e}_b^T\mathbf{F}_i\mathbf{e}_b}\sqrt{\mathbf{e}_a^T\mathbf{F}_i\mathbf{e}_a} \\
&\leq \omega^t\sqrt{\frac{d(b)}{d(a)}}\sqrt{\sum_{i=2}^{n}\mathbf{e}_b^T\mathbf{F}_i\mathbf{e}_b}\sqrt{\sum_{i=2}^{n}\mathbf{e}_a^T\mathbf{F}_i\mathbf{e}_a} \\
&\leq \omega^t\sqrt{\frac{d(b)}{d(a)}}\sqrt{1-\mathbf{e}_b^T\mathbf{F}_1\mathbf{e}_b}\sqrt{1-\mathbf{e}_a^T\mathbf{F}_1\mathbf{e}_a} \\
&< \omega^t\sqrt{\frac{d(b)}{d(a)}}.
\end{aligned}$$

□

Exercise 4.67. Assume G is connected and non-bipartite, with an initial probability distribution $\mathbf{q}$. Let $\mathbf{W}$ be the transition matrix of a non-lazy random walk, and, as before, let $\omega = \max\{|\omega_2|, |\omega_n|\}$. Let $\mathbf{p}$ be the stable distribution. Prove that

$$||\mathbf{p}_t - \mathbf{p}|| < \omega^t.$$

We end this section with a nice exercise.

Exercise 4.68. Let $\mathbf{L}$ be the combinatorial Laplacian, with eigenvalues $\lambda_1 \leq ... \leq \lambda_n$, and $\mathbf{Q}$ the normalized version, with eigenvalues $\mu_1 \leq ... \leq \mu_n$. Let Δ and δ be the largest and smallest degrees of the graph. Verify that

$$\frac{\lambda_i}{\Delta} \leq \mu_i \leq \frac{\lambda_i}{\delta}.$$

(Hint: Use the Courant-Fisher-Weyl theorem — and apply the transformation $\mathbf{D}^{1/2}$).

4.13 References

Here is the set of references used to write the past few pages.

W. Haemers's paper "Interlacing Eigenvalues of Graphs" is a standard reference for applications of interlacing to combinatorics.

More interlacing resources are Brouwer and Haemers's textbook "Spectra of Graphs", and Godsil and Royle's "Algebraic Graph Theory", Chapter 9.

For the theorem associating eigenvalues and matchings, the reference is Brouwer and Haemers's paper "Eigenvalues and Perfect Matchings".

The diameter bound is due to Fan Chung "Diameters and Eigenvalues".

The initial material on Laplacian matrix was mostly based on Godsil and Royle's, Chapter 13.

Fan Chung's book "Spectral Graph Theory" is the standard reference on the Normalized Laplacian.

Bojan Mohar has several articles about Laplacian matrices: "Some Applications of Laplace Eigenvalues of Graphs", "The Laplacian Spectrum of Graphs", "Eigenvalues in combinatorial optimization" (with S. Poljak), and others.

Finally, I also acknowledge D. Spielman's course notes (2018), specially for the last section on random walks.

5 Polynomial method

5.1 DeMillo-Lipton-Zipper-Schwartz

Let x_1, ..., x_n be variables (we will typically denote $\mathbf{x} = (x_1, ..., x_n)$). A *monomial* of degree t is a product of these variables whose total degree sums to t. The constant 1 is the only monomial of degree 0. For a fixed field $\mathbb{F}$ (this could be $\mathbb{R}$, $\mathbb{C}$, $\mathbb{Z}_p$ for prime p, or another finite set with sum, commutative product, and all properties you are used to), let $\mathbb{F}[\mathbf{x}] = \mathbb{F}[x_1, ..., x_n]$ denote the ring of all multivariate polynomials whose coefficients lie in $\mathbb{F}$. The degree of $f \in \mathbb{F}[\mathbf{x}]$ is the largest degree of its monomials. We say f is homogeneous if the degree of all of its monomials are equal. An element $\mathbf{x}$ of $\mathbb{F}^n$ is a root of f if $f(\mathbf{x}) = 0$ (we could also say f vanishes on $\mathbf{x}$, or even on an entire subset $S \subseteq \mathbb{F}^n$).

Exercise 5.1. Let V_d be the vector space whose vectors are polynomials in $\mathbb{F}[\mathbf{x}]$ of degree at most d. What is the dimension of V_d?

In one variable case, we know that for every finite subset S of $\mathbb{F}$, there is a polynomial in $\mathbb{F}[x_1]$ of degree $|S|$ that vanishes on S. We can extend this to the multivariate case (it is essentially a dimensionality argument).

Lemma 5.2. *Given $S \subseteq \mathbb{F}^n$, with $|S| < \binom{n+d}{d}$, there is a nonzero polynomial $f \in \mathbb{F}[\mathbf{x}]$ of degree at most d that vanishes on every element of S.*

Proof. Let $S = \{\mathbf{s}_1, ..., \mathbf{s}_m\}$. Consider the evaluation map

$$\begin{aligned} e : V_d &\to \mathbb{F}^S \\ f &\mapsto (f(\mathbf{s}_1), f(\mathbf{s}_2), ..., f(\mathbf{s}_m)) \end{aligned}$$

Because $\dim V_d > \dim \mathbb{F}^S = |S|$, it follows that this map is non-injective. Therefore there are polynomials f_1 and f_2 so that $e(f_1) = e(f_2)$. Hence $f_1 - f_2$ belongs to V_d and vanishes on S. □

We also know that every polynomial of degree d in one variable has at most d roots (division algorithm?). This can also be generalized (for finite fields, of course).

Lemma 5.3. *Let $\mathbb{F}$ be a finite field with q elements (this is actually unique up to isomorphism, usually denoted by $\mathbb{F}_q$.) Every $f \in \mathbb{F}[\mathbf{x}]$ that is not zero of degree d has at most dq^{n-1} roots.*

Proof. We assume $n \geq 2$, $1 \leq d \leq q$. Let us separate f into the monomials of degree d and the rest, ie, $f = g + h$ with g homogeneous of degree d. Let $\mathbf{y}$ be so that $g(\mathbf{y}) \neq 0$. We now partition $\mathbb{F}^n$ into q^{n-1} lines, namely, sets of the form

$$L_{\mathbf{u}} = \{\mathbf{u} + t\mathbf{w} : t \in \mathbb{F}\}.$$

We now observe that $p_{\mathbf{u}}(t) = f(\mathbf{u} + t\mathbf{w})$ is a polynomial in t of degree at most d, and non-identically zero because the coefficient of t^d is $g(\mathbf{w})$. Therefore f vanishes on at most d points in each $L_{\mathbf{u}}$, and because there are q^{n-1} of these and they partition $\mathbb{F}^n$, we have that f vanishes on at most dq^{n-1} points. □

DeMillo and Lipton, then Zippel and later Schwartz, all independently, proved the following result. The field now is no longer necessarily finite.

Lemma 5.4. *For every $S \subseteq \mathbb{F}$, $|S| \geq d$, every non-zero polynomial $f \in \mathbb{F}[\mathbf{x}]$ of degree d has at most $d|S|^{n-1}$ roots in S^n.*

Proof. The result is by induction on the number of variables. It is clearly true for $n = 1$. Now we write

$$f = f_0 + f_1 x_n + f_2 x_n^2 + ... + f_t x_n^t,$$

where each $f_i \in \mathbb{F}[x_1, ..., x_{n-1}]$. Note that $t \leq d$. Since f_t has degree $d - t$, there are, by induction, at most $(d-t)|S|^{n-2}$ points of S^{n-1} where it vanishes. Thus there are at most $(d-t)|S|^{n-1}$ points $(\mathbf{a}, b) \in S^{n-1} \times S$ where $f(\mathbf{a}, b) = 0$ and $f_t(\mathbf{a}) = 0$.

On the other hand, fixing $\mathbf{a} \in S^{n-1}$ with $f_t(\mathbf{a}) \neq 0$, we have the polynomial $f(\mathbf{a}, x_n)$ has degree d in x_n, thus at most t roots. Hence there are at most $t|S|^{n-1}$ points $(\mathbf{a}, b) \in S^{n-1} \times S$ where $f(\mathbf{a}, b) = 0$ and $f_t(\mathbf{a}) \neq 0$.

All together, at most $d|S|^{n-1}$ points in S^n where f vanishes. □

This lemma is specially useful when investigating whether a given polynomial is identically 0 or not. More specifically, if there is a black box that allows one to evaluate the polynomial at given points, but not to see the coefficients of the polynomial, how many queries are enough to give enough certainty that the polynomial is not identically 0? (of course, if any query returns a non-zero answer, then the polynomial cannot be identically 0). Lemma 5.4 can be reformulated in probabilistic terms.

Lemma 5.5. *Let $f \in \mathbb{F}[\mathbf{x}]$, nonzero, of degree d, and $S \subseteq \mathbb{F}$ non-empty subset. Then the probability that $f(\mathbf{s}) = 0$ for some $\mathbf{s} \in S^n$ selected uniformly and independently is less or equal than $d/|S|$.* □

Naturally, selecting S with $|S| = 2d$ and repeating this several times will yield a fairly certain probability that f must be 0.

We will see a nice application of the DeMillo-Lipton-Zippel-Schwartz lemma to decide whether a graph contains a perfect matching or not. But for now, a quite interesting application to another problem.

5.2 The Kakeya problem

First, watch this

`https://www.youtube.com/watch?v=IM-n9c-ARHU&t=2s`

The Kakeya problem is the following: what is the smallest set in the plane in which one can rotate a unit length needle around completely? (the disk of diameter 1 is clearly not the best choice).

For dimensions larger than 2, the question rephrases as "what is the *Hausdorff dimension* of a subset of $\mathbb{R}^n$ that contains a unit line segment in every direction?" This question remains open. However a finite field version has been proposed and answered quite simply by Dvir in 2009.

The question is simple. Let $\mathbb{F}$ be a finite field. A Kakeya set K is a subset of $\mathbb{F}^n$ that contains a line in every possible direction. In other words, to any $\mathbf{w}$ (direction), there is a $\mathbf{v}$ so that $\mathbf{v} + t\mathbf{w} \in K$ for all t. How big must K be?

Lemma 5.6. *Let $\mathbb{F} = \mathbb{F}_q$. Let $f \in \mathbb{F}[\mathbf{x}]$, degree at most $q-1$. If f vanishes on a Kakeya set K, then f is the 0 polynomial.*

Proof. Write $f = f_0 + ... + f_d$, where each f_i is homogeneous of degree i (and thus $d \leq q-1$). Given $\mathbf{w}$, there is a $\mathbf{v}$ so that $\mathbf{v} + t\mathbf{w} \in K$ for all t, and thus

$$f(\mathbf{v} + t\mathbf{w}) = 0.$$

Note that fixing $\mathbf{v}$ and $\mathbf{w}$, f is a polynomial in t of degree at most $q-1$ but with q roots, thus all coefficients of t in $f(\mathbf{v} + t\mathbf{w})$ are 0. In particular, the coefficient of t^d is 0, and that is $f_d(\mathbf{w})$. Choosing another $\mathbf{w}$, we reach the conclusion that f_d is everywhere 0. But f_d, if nonzero, should have at most dq^{n-1} roots, and not q^n. Thus f_d is the zero polynomial. The same will hold for the other f_i. □

Theorem 5.7 (Dvir)**.** *Let K be a Kakeya set in $\mathbb{F}^n$, with $\mathbb{F} = \mathbb{F}_q$. Then*

$$|K| \geq \binom{q+n-1}{n} \geq \frac{q^n}{n!}.$$

Proof. If $|K| < \binom{q+n-1}{n}$, then there is a nonzero polynomial $f \in \mathbb{F}[\mathbf{x}]$ of degree at most $q-1$ that vanishes on K (as per Lemma 5.2), a contradiction to the previous lemma. □

Exercise 5.8. Let $\mathbb{F} = \mathbb{F}_q$. Assume there is a subset $B \subseteq \mathbb{F}^N$, called a Nikodym set, so that, for each point $\mathbf{x} \notin B$, there is a line $L_{\mathbf{x}} = \{\mathbf{x} + t\mathbf{w} : t \in \mathbb{F}\}$ so that $|L_{\mathbf{x}} \cap B| = q-1$. Prove that

$$|B| \geq \binom{n+q-2}{n}.$$

We make a detour now to talk about matchings again. We will see why the DeMillo-Lipton-Zippel-Schwartz lemma is a key piece in designing and efficient algorithm to check the existence of a perfect matching in a graph.

5.3 Pfaffians and determinants

A square $n \times n$ matrix $\mathbf{A}$ with entries taken from $\mathbb{F}[\mathbf{x}]$ is *skew-symmetric* if $\mathbf{A}^T = -\mathbf{A}$. Our goal here is to show that the determinant of $\mathbf{A}$ is always a perfect square. To that effect, we shall define what is the Pfaffian of a matrix.

For this section, assume n is even, equal to $2m$.

First recall the Leibniz expression of a determinant:

$$\det \mathbf{A} = \sum_{\sigma \in S_n} (-1)^{\epsilon(\sigma)} \prod_{i=1}^{n} \mathbf{A}_{i\sigma(i)}$$

(The sum running over all permutations of $\{1, ..., n\}$, and $\epsilon(\sigma)$ being the number of cycles of even length in the decomposition of σ as a product of disjoint cycles.)

Exercise 5.9. Let G be a bipartite graph, classes U and V, and assume each class has n vertices. To each edge $uv \in E(G)$, consider the variable x_{uv}. Define the square $n \times n$ matrix $\mathbf{B}$, whose rows are indexed by vertices in U and columns by vertices in V, with $\mathbf{B}_{u,v} = x_{uv}$ if $uv \in E$, and 0 otherwise. Show that $\det \mathbf{B}$ is not identically equal to 0 if and only if G has a perfect matching.

Let us define another number associated to a matrix $\mathbf{A}$. Each permutation σ of S_{2m} determines the perfect matching of K_{2m} with edges $\{\sigma(2i-1), \sigma(2i)\}$, for $i = 1, ..., m$.

Exercise 5.10. Verify that each matching of K_{2m} can be obtained from precisely $2^m m!$ distinct permutations.

Given $\mathbf{A}$ skew symmetric, we define the weight of a perfect matching $\mathfrak{m}$ of K_{2m} obtained from σ by

$$\operatorname{wt} \mathfrak{m} = (-1)^{\epsilon(\sigma)} \prod_{i=1}^{m} \mathbf{A}_{\sigma(2i-1),\sigma(2i)}.$$

It is not at all obvious at first sight that this definition does not depend on σ. To see this, note that given a permutation σ that determines $\mathfrak{m}$, the permutations $\sigma \circ (2i-1\ 2i)$ and $\sigma \circ (2j-1\ 2i-1) \circ (2j\ 2i)$ both determine the same matching. Moreover, as the matrix is skew-symmetric, the weight of $\mathfrak{m}$ computed according to both permutation remains unchanged. Finally, any other permutation that determines the same matching $\mathfrak{m}$ can be obtained from σ upon applying these compositions with transpositions.

The Pfaffian of a skew-symmetric matrix $\mathbf{A}$ is defined as

$$\operatorname{pf} \mathbf{A} = \sum_{\mathfrak{m}} \operatorname{wt} \mathfrak{m},$$

with the sum being over all perfect matchings of K_n (note that the Pfaffian is equal to 0 if n is odd).

Theorem 5.11 (Cayley)**.** *Let $\mathbf{A}$ be a $n \times n$ skew-symmetric matrix (with entries taken from $\mathbb{F}[\mathbf{x}]$). Then*

$$\det \mathbf{A} = (\operatorname{pf} \mathbf{A})^2.$$

Proof. Let $\mathcal{E}_n \subseteq S_n$ be the set of permutations composed exactly by those permutation whose all cycles have even length. We start by noticing that if $\mathbf{A}$ is a skew-symmetric matrix, then:

$$\det \mathbf{A} = \sum_{\sigma \in \mathcal{E}_n} (-1)^{\epsilon(\sigma)} \prod_{i=1}^{n} \mathbf{A}_{i\sigma(i)}$$

To see that, we pair up the permutations which contain odd cycles by defining to each permutation α with an odd cycle the permutation α' obtained from α by reversing the odd cycle containing the smallest least element. Note that $(\alpha')' = \alpha$, and that $(-1)^{\epsilon(\alpha)} = (-1)^{\epsilon(\alpha')}$, but because the matrix is skew-symmetric, $\prod \mathbf{A}_{i\alpha(i)} = -\prod \mathbf{A}_{i\alpha'(i)}$.

Let $\mathcal{M}(K_n)$ be the set of all perfect matchings of K_n. Consider the function:

$$\Phi : \mathcal{E}_n \to \ \mathcal{M}(K_n) \times \mathcal{M}(K_n)$$

that takes a permutation

$$\alpha = (\alpha_{11}\alpha_{12}...\alpha_{1i_1})(\alpha_{21}\alpha_{22}...\alpha_{2i_2})....(\alpha_{k1}\alpha_{k2}...\alpha_{ki_k})$$

(assume α_{*1} is the smallest element in each cycle) and creates two matchings:

$$\begin{aligned}\mathfrak{m}_1 &= \{\alpha_{11}\alpha_{12}, \alpha_{13}\alpha_{14}, ..., \alpha_{1(i_1-1)}\alpha_{1i_1}\}\cup \\ &\quad \cup \{\alpha_{21}\alpha_{22}, \alpha_{23}\alpha_{24}, ..., \alpha_{2(i_2-1)}\alpha_{2i_2}\}\cup \\ &\quad \cup ... \cup \{\alpha_{k1}\alpha_{k2}, \alpha_{k3}\alpha_{k4}, ..., \alpha_{k(i_k-1)}\alpha_{ki_k}\} \\ \mathfrak{m}_2 &= \{\alpha_{12}\alpha_{13}, \alpha_{14}\alpha_{15}, ..., \alpha_{1i_1}\alpha_{11}\}\cup \\ &\quad \cup \{\alpha_{22}\alpha_{23}, \alpha_{24}\alpha_{25}, ..., \alpha_{2i_2}\alpha_{21}\}\cup \\ &\quad \cup ... \cup \{\alpha_{k2}\alpha_{k3}, \alpha_{k4}\alpha_{k5}, ..., \alpha_{ki_k}\alpha_{k1}\}\end{aligned}$$

Convince yourself that this function is a bijection. Given α with $\Phi(\alpha) = (\mathfrak{m}_1, \mathfrak{m}_2)$, we claim now that we can choose σ_1 and σ_2 so that each define $\mathfrak{m}_1$ and $\mathfrak{m}_2$ respectively, as we previously described, and so that

$$\alpha = \sigma_2 \circ \sigma_1.$$

In fact, simply choose the σ_1 as the permutation given by how α was written above, and define σ_2 as the unique permutation that makes equality hold. It is immediate to check that σ_2 defines $\mathfrak{m}_2$. Therefore the following equality holds:

$$\begin{aligned}(-1)^{\epsilon(\alpha)}\prod_{i=1}^{n}\mathbf{A}_{i\alpha(i)} &= \left((-1)^{\epsilon(\sigma_1)}\prod_{i=1}^{m}\mathbf{A}_{\sigma_1(2i-1),\sigma_1(2i)}\right)\left((-1)^{\epsilon(\sigma_2)}\prod_{i=1}^{m}\mathbf{A}_{\sigma_2(2i-1),\sigma_2(2i)}\right) \\ &= \operatorname{wt}(\mathfrak{m}_1)\operatorname{wt}(\mathfrak{m}_2).\end{aligned}$$

Because Φ is a bijection, the following equality also holds:

$$\sum_{\alpha\in\mathcal{E}_n}(-1)^{\epsilon(\alpha)}\prod_{i=1}^{n}\mathbf{A}_{i\alpha(i)} = \sum_{(\mathfrak{m}_1,\mathfrak{m}_2)}\operatorname{wt}(\mathfrak{m}_1)\operatorname{wt}(\mathfrak{m}_2) = \left(\sum_{\mathfrak{m}\in\mathcal{M}}\operatorname{wt}(\mathfrak{m})\right)^2.$$

But the left hand side is the determinant of $\mathbf{A}$ because of the first remarks in the proof, and the right hand side is precisely $\operatorname{pf}(\mathbf{A})^2$. □

5.4 Tutte matrix, and perfect matchings

Let $G = (V, E)$ be a simple undirected graph. We write $V = \{v_1, ..., v_n\}$. Let $G' = (V, A)$ be an orientation of the edges of G. To each edge $e \in A$, we assign a variable x_e. The *Tutte Matrix of the graph* G $\mathbf{T} = \mathbf{T}_G(\mathbf{x})$ is defined as:

$$\mathbf{T}_{ij} = \begin{cases} x_e, & \text{if } e = (v_i, v_j); \\ -x_e, & \text{if } e = (v_j, v_i); \\ 0, & \text{otherwise.}\end{cases}$$

Observe that this is a skew-symmetric matrix whose entries lie on $\mathbb{F}[\mathbf{x}]$.

Theorem 5.12 (Tutte - 1947)**.** *A graph G has a perfect matching if and only if* $\det \mathbf{T}_G(\boldsymbol{x})$ *is non-identically zero.*

Proof. Both direction follow easily from Cayley's Theorem: if G has no perfect matching, than there is no non-zero term on the expression for the Pfaffian of $\mathbf{T}_G(\mathbf{x})$. On the other hand, if G has some perfect matchings, than every non-zero term on the expansion of the Pfaffian correspond exactly to one of them. No two of these terms can use the same set of variables, otherwise the matchings would be the same, and as variables are algebraically independent, there can be no non-trivial algebraic relation of them giving zero. So the Pfaffian is non-zero, hence so is the determinant. □

The practical issue with the result above is that a symbolic computation of a determinant means writing an exponentially large expression, hence it is inefficient.

The computation of determinants, however, is efficient if the entries are elements of a field. As pointed by Lovász in 1979, if $\det \mathbf{T}$ is not identically 0, then the subset of $[0,1]^m$ whose attribution to variables makes $\det \mathbf{T} = 0$ has measure 0, and therefore one could simply evaluate $\det \mathbf{T}$ at some randomly chosen real numbers to the variables. This of course cannot be implemented, thus we can finally use the DeMillo-Lipton-Zippel-Schwartz. From Lemma 5.5, we have:

Theorem 5.13. *If* $\det \mathbf{T}$ *has degree at most d, and if we randomly attribute values from a finite set S of* $\mathbb{R}$ *to the variables used in defining* $\mathbf{T}$*, its rank is preserved with probability at least* $1 - d/|S|$.

Thus we have the following randomized algorithm to find whether a graph has a perfect matching, and in which case, to actually find it (we present it with $\mathbb{F}$ being chosen to be finite).

(1) Make $G' = (V', E')$ to be equal to $G = (V, E)$, $|V| = n$, $|E| = m$. Let $q \geq 2n$ be a prime power and make $U = \emptyset$. Let λ be a desirable failure probability.

(2) Make $\mathbf{x} = \mathbf{r}$, where $\mathbf{r} \in \mathbb{F}_q{}^m$ is uniformly and randomly chosen. Compute the determinant of $\mathbf{T}_{G'}((\mathbf{x}))$. If it is non-zero, step to (4). If not, repeat (2) at most $\log_2 \lambda$ times.

(3) If $\det \mathbf{T}_{G'}(\mathbf{x}) = \mathbf{0}$ in all trials and if $E = E'$, stop. Return *no perfect matching exists.*

(4) If $\det \mathbf{T}_{G'}(\mathbf{x}) \neq \mathbf{0}$ for some trial, than choose a random edge $e \in E' \backslash U$ and make $G' = G \backslash e$. Return to (2).

(5) If $\det \mathbf{T}_{G'}(\mathbf{x}) = \mathbf{0}$ in all trials and if $E \neq E'$, than put the last chosen e in U and back again in E'.

> If U becomes equal to E', return U *is a perfect matching.*
>
> Else, choose random edge $f \in E' \backslash U$ and make $G' = G \backslash f$. Return to (2).

Observe that the test in (2) will be run for at most m edges, and each one of them has a probability of failing of λ. Hence the total probability of failing is at most $m\lambda$, which can be made arbitrarily small. The time for each determinant computation is $O(n^w)$, and this shall be performed at most $O(m)$ times, hence the algorithms takes at most $O(mn^w)$ steps. It can be improved to run in $O(n^w)$ steps (see references).

5.5 Combinatorial Nullstenllensatz

Recall the DeMillo-Lipton-Zippel-Schwartz lemma.

Lemma. *For every $S \subseteq \mathbb{F}$, $|S| \geq d$, every non-zero polynomial $f \in \mathbb{F}[\mathbf{x}]$ of degree d has at most $d|S|^{n-1}$ roots in S^n.*

Now we introduce a somewhat granulated version, which will be useful to the next result to come.

Lemma 5.14. *Let $f \in \mathbb{F}[\mathbf{x}]$, and suppose f has degree at most d_i in variable x_i. Let $S_i \subseteq \mathbb{F}$, and assume $|S_i| \geq d_i + 1$. If f is not the zero polynomial, then there is $\mathbf{z} \in S_1 \times ... \times S_n$ so that $f(\mathbf{z}) \neq 0$.*

Proof. The proof is by induction, and basically the same as we had before. For $n = 1$, the result follows from the division algorithm. Next, we write

$$f = \sum_{i=0}^{d_n} f_i x_n^i,$$

where each f_i is a polynomial in $\mathbb{F}[x_1, ..., x_{n-1}]$. As f is non-zero, some f_j is non-zero, and thus, by induction, there is $\mathbf{z}' \in S_1 \times ... \times S_{n-1}$ so that $f_j(\mathbf{z}') \neq 0$. We now consider the one variable polynomial

$$f(z_1, ..., z_{n-1}; x_n) = \sum_{i=0}^{d_n} f_i(z_1, ..., z_{n-1}) x_n^i,$$

which is non-zero and has degree at most d_n, thus implies that at least one element in S_n, say z_n, is not its root. Hence $f(z_1, ..., z_n) \neq 0$. □

Hilbert's Nullstellensatz says that if $\mathbb{F}$ is an algebraically closed field (meaning, all polynomials with coefficients in $\mathbb{F}$ have all their roots in $\mathbb{F}$), and if f and $g_1, ..., g_m$ are polynomials in $\mathbb{F}[\mathbf{x}]$ so that f vanishes in all common roots of $g_1, ..., g_m$, then there is an integer k and polynomials $h_1, ..., h_m$ so that

$$f^k = \sum_{i=1}^{m} h_i g_i.$$

For several applications in combinatorics, it is enough to consider a version with a more hypothesis where a stronger conclusion holds. The result is due to Noga Alon, in 1999.

Theorem 5.15. *Let $\mathbb{F}$ be an arbitrary field, $f \in \mathbb{F}[x_1, ..., x_n]$. Assume $x_1^{t_1} \cdot ... \cdot x_n^{t_n}$ is the monomial of largest degree (and non-zero coefficient) in f. Assume $S_1, ..., S_n \subseteq \mathbb{F}$.*

(i) Special case of Hilbert's Nullstellensatz: Define the univariate polynomials

$$g_i = \prod_{s \in S_i} (x_i - s).$$

If $f(\mathbf{s}) = 0$ for all $\mathbf{s} \in S_1 \times ... \times S_n$, then there are polynomials $h_1, ..., h_n \in \mathbb{F}[\mathbf{x}]$, with $\deg h_i + \deg g_i \leq \deg f$, *so that*

$$f = \sum_{i=1}^{n} h_i g_i.$$

Moreover, if the coefficients of f and g_is are in a subring of $\mathbb{F}$, then so are those of the h_is.

(ii) Combinatorial Nullstellensatz: If $|S_i| \geq t_i + 1$, then there is $\mathbf{z} \in S_1 \times ... \times S_n$ so that $f(\mathbf{z}) \neq 0$.

Proof. First we prove (a). By definition of the g_is, for each $\mathbf{s} \in S_1 \times ... \times S_n$, $g_i(s_i) = 0$, thus

$$s_i^{|S_i|} = \sum_{j=0}^{|S_i|-1} g_{ij} s_i^j.$$

Let $\overline{f}$ be obtained from f upon replacing each occurrence of $x_i^{\ell_i}$ with $\ell_i \geq |S_i|$ by the linear combination of smaller powers of x_i given by the g_{ij}s, as above. In particular, $\overline{f}$ is obtained from f upon subtracting terms of the form $h_i g_i$ where $h_i \in \mathbb{F}[\mathbf{x}]$, with its degree at most $\deg f - \deg g_i$.

Note that $\overline{f}$ has degree at most $|S_i|-1$ in each x_i, satisfies $\overline{f}(\mathbf{s}) = 0$ for all $\mathbf{s} \in S_1 \times ... \times S_n$, and therefore $\overline{f} \equiv 0$, by the previous lemma. Therefore $f = \sum h_i g_i$.

To see (b) now, assume to the contrary that f vanishes in the entirety of $S_1 \times ... \times S_n$. We define g_is as in (a), which implies existence of h_is of degree $\leq \deg f - \deg g_i$ with $f = \sum h_i g_i$. The monomial $x_1^{t_1} \cdot ... \cdot x_n^{t_n}$ has non-zero coefficient in f and has largest total degree, but any term of largest total degree in each $h_i g_i$ contains $x_i^{|S_i|}$, and $|S_i| > t_i$. A contradiction. □

5.6 Combinatorial number theory

The following result has been applied successfully several times in Additive Number Theory. In this subsection, we show some of its applications. If A and B are subsets of $\mathbb{F}$, then we define

$$A + B = \{a + b : a \in A,\ b \in B\}.$$

Exercise 5.16. Show that if $|A| + |B| > p$, then $A + B = \mathbb{Z}_p$.

Theorem 5.17 (Cauchy-Davenport)**.** *If p is a prime, $A, B \subseteq \mathbb{Z}_p$, then*

$$|A + B| \geq \min\{p, |A| + |B| - 1\}.$$

Proof. Due to the exercise, assume $|A| + |B| \leq p$, and, to find a contradiction, assume $|A + B| \leq |A| + |B| - 2$. Let C contain $A + B$, with $|C| = |A| + |B| - 2$. Define the two variable polynomial $f(x, y) = \prod_{t \in C}(x + y - t)$. Note that f vanishes in $A \times B$. Note that the coefficient of

$$x^{|A|-1} y^{|B|-1}$$

in f is the binomial coefficient $\binom{|A|+|B|-2}{|A|-1}$, which is non-zero in $\mathbb{Z}_p$ because $|A| + |B| - 2 < p$. The Combinatorial Nullstellensatz implies the existence of $(a, b) \in A \times B$ with $f(a, b) \neq 0$, a contradiction. □

Assume now we have a sequence of n integers, $a_1, ..., a_n$. We can show (easily) that it contains a consecutive subsequence whose sum is divisible by n. In fact, consider the remainders of a_1, $a_1 + a_2$, ... , $a_1 + ... + a_n$ when divided by n. If any is equal to 0 we are done. Otherwise, two of them are equal, and we can simply subtract one from the other to find the consecutive subsequence whose sum is divisible by n.

To make things more fun, we can now ask the following: given n, what is the smallest N so that any sequence of N integers contains a subsequence of n numbers (not necessarily consecutive) whose sum is divisible by n?

We can very easily build a sequence of length $2n-2$ that does not contain a subsequence of length n whose sum is divisible by n.

Theorem 5.18 (Erdős-Ginzburb-Ziv)**.** *Any sequence of $2n-1$ integers contains a subsequence of size n whose sum is divisible by n.*

Proof. First assume $n = p$, a prime. Let $a_1 \leq ... \leq a_{2p-1}$ be integers. If $a_i = a_{i+p-1}$ for some i, then the result follows trivially. Hence, define $A_i = \{a_i, a_{i+p-1}\}$, for $i = 1, ..., p-1$. Upon repeatedly applying Cauchy-Davenport Theorem, we have

$$|A_1 + ... + A_{p-1}| \geq \min\{p, |A_2 + ... + A_{p-1}| + 1\} \geq ... \geq p.$$

Thus every number in $\mathbb{Z}_p$ is a sum of precisely $p-1$ of the first $2p-2$ elements, and $-a_{2p-1}$ is one of them. Hence we have the result.

If n is not a prime, we write $n = pm$. We will apply induction on the number of prime factors of n. Let $a_1 \leq ... \leq a_{2n-1}$, and, using the case for primes above, there are pairwise disjoint subsets $I_1, ..., I_\ell$ of $\{a_1, ..., a_{2n-1}\}$, each of size p, so that

$$\sum_{j \in I_i} a_j \equiv 0 \pmod{p}.$$

Note that we can assume $\ell \geq 2m-1$. So we define

$$b_i = \frac{1}{p} \sum_{j \in I_i} a_j.$$

By induction now, there is a subsequence of $b_1, ..., b_{2m-1}$ with m elements whose sum is divisible by m. Each of these correspond to an I_i. Taking their union gives a set of n elements whose sum is divisible by n. □

Another way of obtaining the theorem above is by applying the result below (I leave this as an exercise). Before showing the result, we recall a famous elementary result due to Fermat

Lemma 5.19 (Fermat's Little Theorem)**.** *Let p be a prime, and $a \not\equiv 0 \pmod{p}$. Then $a^{p-1} \equiv 1 \pmod{p}$.*

Proof. Note that $\{a, 2a, 3a, ..., (p-1)a\} \equiv \{1, 2, 3, ..., (p-1)\} \pmod{p}$. Thus

$$a^{p-1}(p-1)! \equiv (p-1)! \pmod{p},$$

thus

$$a^{p-1} \equiv 1 \pmod{p}.$$

□

Theorem 5.20 (Chevally-Warning)**.** *Let p be a prime, and define polynomials $f_1, ..., f_m$ in $\mathbb{Z}_p[x_1, ..., x_n]$. If $n > \sum \deg f_i$, then these polynomials cannot have a unique common zero.*

Proof. Suppose this is false, and let $\mathbf{c} = (c_1, ..., c_n)$ be their unique zero. Define the polynomial

$$f = \prod_{i=1}^{m}(1 - f_i^{p-1}) - \delta \prod_{j=1}^{n} \prod_{c\in\mathbb{Z}_p,\ c\neq c_j} (x_j - c).$$

The element δ is chosen so that $f(c_1, ..., c_n) = 0$. Note that $\delta \neq 0$.

The polynomial f is special because f vanishes entirely in $\mathbb{Z}_p^n$. In fact, $\mathbf{c}$ is a root, but if $\mathbf{s} \neq \mathbf{c}$, then there f_j so that $f_j(ß) \neq 0$, thus $(1 - f_j^{p-1}(ß)) = 0$, and, also, for some i where $s_i \neq c_i$, $\prod_{c\in\mathbb{Z}_p,\ c\neq c_i}(s_i - c) = 0$.

The monomial $(x_1^{p-1}x_2^{p-1} \cdot ... \cdot x_n^{p-1})$ has largest degree. Thus, by the Combinatorial Nullstellensatz, there is $ß \in \mathbb{Z}_p^n$ with $f(ß) \neq 0$, a contradiction. □

Exercise 5.21. Chevalley-Warning actually has a stronger statement (and a more elementary proof). In fact, it holds that the number of common zeros, say N, of those polynomials is divisible by p. To see this, define

$$N \equiv \sum_{\mathbf{y}\in\mathbb{F}_p^n} \prod_{j=1}^{m}(1 - f_m(\mathbf{y})^{p-1}).$$

(Why is this true?).

Following, it is enough to expand the product, and note that $\sum_{y\in\mathbb{F}_p} y^r \equiv 0 \pmod p$ if $1 \leq r \leq p-2$. (Why?).

(This result easily generalizes if $|\mathbb{F}|$ is a power of p.)

5.7 Applications to graph theory

We finally arrive to some application to combinatorics and graph theory. The first result displays a clever application, which resolved a well-known conjecture.

Theorem 5.22. *For any prime p, any graph G with average degree bigger than $2p-2$ and maximum degree at most $2p-1$ contains a p-regular graph subgraph.*

Proof. Let N be the incidence matrix, and create a variable to each edge e of G, say x_e. Then, define

$$f = \prod_{v\in V(G)} (1 - \mathbf{e}_v^T N\mathbf{x})^{p-1} - \prod_{e\in E(G)} (1 - x_e).$$

The degree of f is $|E|$. In fact, the degree of the first term is bounded below this number, and the coefficient to the monomial of total degree $|E|$ is $(-1)^{|E|+1}$. Thus, by the Combinatorial Nullstellensatz, there are values $ß \in \{0,1\}^n$ so that $f(ß) = 0$. Note that this cannot be the 0 vector, and, thus, it must be that $\mathbf{e}_v^T Nß$ is zero modulo p for all v. Thus, the edges for which ß is equal to 1 define a subgraph whose degree of all vertices are divisible by p, and, thus are equal to p. □

Exercise 5.23. Let p be a prime, and G a graph on $|V| > d(p-1)$ vertices. Then there is a nonempty subset U of vertices of G so that the number of cliques of d vertices that intersect U is 0 modulo p.

Prove this fact, examining the polynomial

$$f = \prod_{v \in V}(1 - x_v) - 1 + \left(\sum_{\emptyset \neq I \subseteq V} (-1)^{|I|+1} K(I) \prod_{v \in I} x_v \right)^{p-1},$$

where $K(I)$ counts the number of cliques on d vertices of G that contain I as a subset.

Our last application is to the topic of graph colourings.

Let G be a graph and L a function that assigns to each vertex of G a list of positive integers. A *list colouring* of G with respect to L is a proper colouring of G whose colour of each vertex lies in its assigned list. If L is given and G has a list colouring with respect to L, we say G is L-colourable. Note that if $L(v) = \{1, ..., k\}$ for all v, then G is L-colourable if and only if G is k-chromatic.

Exercise 5.24. Contrary to your intuition, it is possible to find a graph which is k-chromatic but so that one can define lists to each vertex of size at least k and yet find no list colouring with respect to them. Find a bipartite graph on 6 vertices satisfying this property.

Given an ordering of the vertices and having $\mathbf{d}$ to be a sequence of n nonnegative integers, a graph is said to be $\mathbf{d}$-list-colourable if every list assignment to the vertices so that vertex v_i receives a list of size d_i allows for a corresponding list colouring.

Our goal below is to provide a strong combinatorial sufficient condition for G to admit a list colouring with relatively small lists.

Given a graph G, we define variables x_v for each $v \in V(G)$. The adjacency polynomial of G is defined as

$$a(G; \mathbf{x}) = \prod_{\substack{a<b \\ ab \in E(G)}} (x_a - x_b)$$

Each monomial in $a(G; \mathbf{x})$ corresponds to a choice of a term in each $(x_a - x_b)$, and therefore corresponds to an orientation of G. Given an ordering of the vertices, each orientation D therefore determines a sign (the sign of the corresponding monomial), which we define as the sign of the orientation, denoted by $\sigma(D)$.

Let $\mathbf{d} = (d_1, ..., d_n)$ be a sequence of nonnegative integers summing to m. The weight of $\mathbf{d}$ is defined as

$$\omega(\mathbf{d}) = \sum \sigma(D),$$

where the sum is running over all orientations of G whose outdegree sequence is $\mathbf{d}$. If we set $\mathbf{x}^{\mathbf{d}}$ to mean $\prod_{i=1}^{n} x_i^{d_i}$, then the adjacency polynomial is given by

$$a(G; \mathbf{x}) = \sum_{\mathbf{d}} \omega(\mathbf{d}) \mathbf{x}^{\mathbf{d}}.$$

Lemma 5.25. *Let G be a graph, D an orientation with outdegree sequence $\mathbf{d}$. If D' is another orientation with the same outdegree sequence, then $\sigma(D) = \sigma(D')$ if and only if the number of arcs in D which are not in D' is even. Moreover, if D has no directed odd cycles, then all orientations of G with outdegree sequence $\mathbf{d}$ have the same sign.*

Proof. Exercise (purely combinatorial). □

Theorem 5.26. *Let G be a graph, and D an orientation of G without odd cycles. Then G is $(\mathbf{d}+\mathbf{1})$-list colourable, where $\mathbf{d}$ is the outdegree sequence of D.*

Proof. From the lemma, it holds that $\omega(\mathbf{d}) \neq 0$. Thus we can immediately apply the Combinatorial Nullstellensatz to the polynomial $a(G;\mathbf{x})$ (note here that an evaluation of $a(G;\mathbf{x})$ is non-zero if and only if variables corresponding to neighbouring vertices take distinct values). □

Note that the same conclusion can be obtained if we assume G has an odd number of orientations D with outdegree sequence D.

5.8 References

Here is the set of references used to write the past few pages.

N. Alon's paper "Combinatorial Nullstellensatz" is certainly the best source for most results in the end of the past section.

Bondy and Murty's book also contains an interesting section on the topic.

The chapter on the polynomial method on Jukna's book "Extremal Combinatorics" served as a guide to the first two subsection and the last.

N. Harvey's paper "Algebraic Algorithms for Matching and Matroid Problems" contains a good account of algebraic algorithms to find perfect matchings.

Chapter 7 in Godsil's "Algebraic Combinatorics" contains the proof of Cayley's theorem I based mine upon.

The main reference on matchings in Lovász and Plummer's book "Matching Theory".