Linear algebra and group theory

Teo Banica

Department of Mathematics, University of Cergy-Pontoise, F-95000 Cergy-Pontoise, France. teo.banica@gmail.com

2010 *Mathematics Subject Classification.* 15B10

Key words and phrases. Square matrix, Classical group

Abstract. This is an introduction to linear algebra and group theory. We first review the linear algebra basics, namely the determinant, the diagonalization procedure and more, and with the determinant being constructed as it should, as a signed volume. We discuss then the basic applications of linear algebra to questions in analysis. Then we get into the study of the closed groups of unitary matrices $G\subset U_N$, with some basic algebraic theory, and with a number of probability computations, in the finite group case. In the general case, where $G\subset U_N$ is compact, we explain how the Weingarten integration formula works, and we present some basic $N\to\infty$ applications.

Preface

Linear algebra is the source of many good things in this world. First of all, everything algebra, for sure. But also geometry and analysis, because any smooth function or manifold, taken locally, perturbes a certain linear transformation of $\mathbb{R}^N$. And finally probability too, remember indeed that Gauss integral needed for talking about normal laws, which can only be computed by using polar coordinates and their Jacobian.

The purpose of this book is to talk about linear algebra in a large sense, theory and applications, at a somewhat more advanced level than the beginner one, and by insisting on beautiful things. And with some graduate level mathematics, and quantum physics too, in mind. We will particularly insist on the groups of matrices, which are extremely useful for all sorts of mathematics and physics, and which are perhaps the most beautiful topic one could study, once the basics of linear algebra and matrices understood.

The book itself is accessible with 0 knowledge or almost, with page 1 of chapter 1 talking about vectors in $\mathbb{R}^2$. However, things will escalate quickly, and some prior knowledge of linear algebra, even in some rough form, is recommended. All in all, the book covers what can be taught during a 1-year upper division undergraduate course. Some other forms of applications are possible as well, and more on this later.

The first half of the book is concerned with linear algebra and its applications. Part I is a quick journey through basic linear algebra, from basic definitions and fun with 2×2 matrices, up to the Spectral Theorem in its most general form, for the normal matrices $A \in M_N(\mathbb{C})$. Among the features of our presentation, the determinant will be introduced as it should, as a signed volume of a system of vectors. And also, we will discuss all sorts of useful matrix tricks, which are more advanced, and good to know.

As a continuation of this, Part II deals with various applications of linear algebra, to questions in analysis. After a quick look at differentiation and integration, which in several variables are intimately related to matrix theory, via the Jacobian, Hessian and so on, we will develop some useful probability theory, in relation with the normal and hyperspherical laws, by using spherical coordinates and their Jacobian. We will also discuss some other analytic topics, such as special matrices and spectral theory.

The second half of the book is concerned with matrix groups. As already mentioned, this is perhaps the most beautiful topic one could study, once the basics of linear algebra understood. The subject is however huge, and Part III will be a modest introduction to it. Our philosophy will be that of talking about all sorts of interesting closed subgroups $G \subset U_N$, finite and continuous alike, and by using very basic methods, coming from standard calculus, combinatorics and probability, for their study.

As a conclusion to this, the finite group case will appear to be reasonably understood, while the continuous case, not. Part IV will be dedicated to the study of the closed subgroups $G \subset U_N$, and more specifically the continuous ones, by using heavy machinery, as heavy as it gets. We will discuss here the basics of representation theory, then the existence of the Haar measure, and the Peter-Weyl theory, and then more advanced topics, such as Tannakian duality, Brauer theorems, and Weingarten calculus.

All in all, as already mentioned, we will cover here what can be taught during a 1-year upper division undergraduate course. There are of course some other options too, as for instance using Parts I and III, with minimal input from Parts II and IV, for a 1-semester upper division undergraduate course, on linear algebra and group theory.

At the level of things which are not done in this book, notable topics include the Jordan decomposition, which is the nightmare of everyone involved, teacher or student, and this remains between us, as well as some basic Lie algebra theory, which would have perfectly make sense to include, as a main topic for the last part, but that we preferred to replace by representation theory, and its relation with combinatorics and probability, which are somewhat more elementary, and fitting better with the rest of the book.

Finally, let us mention that this way of presenting things has its origins in some recent research work on the quantum groups, and more specifically on the so-called easy quantum groups. The idea there is that there is no much smoothness and geometry, with the main tools belonging to combinatorics and probability. Thus, as main philosophy, the present book, while dealing with classical topics, is written with a "quantum" touch.

Most of this book is based on lecture notes from various classes at Cergy, and I would like to thank my students. The final part goes into research topics, and I am grateful to Benoît Collins, Steve Curran and Jean-Marc Schlenker, for our joint work on the subject. Many thanks go as well to my cats. There is so much to learn from them, too.

Contents

Part I

Linear algebra

So close, no matter how far
Couldn't be much more from the heart
Forever trusting who we are
And nothing else matters

CHAPTER 1

Real matrices

1a. Linear maps

We are interested in what follows in symmetries, rotations, projections and other such basic transformations, in 2, 3 or even more dimensions. Such transformations appear a bit everywhere, in physics. To be more precise, each physical problem or equation has some "symmetries", and exploiting these symmetries is usually a useful thing.

Let us start with 2 dimensions, and leave 3 and more dimensions for later. The transformations of the plane $\mathbb{R}^2$ that we are interested in are as follows:

Definition 1.1. *A map $f:\mathbb{R}^2\to\mathbb{R}^2$ is called affine when it maps lines to lines,*

$$f(tx+(1-t)y)=tf(x)+(1-t)f(y)$$

for any $x,y\in\mathbb{R}^2$ and any $t\in\mathbb{R}$. If in addition $f(0)=0$, we call f linear.

As a first observation, our "maps lines to lines" interpretation of the equation in the statement assumes that the points are degenerate lines, and this in order for our interpretation to work when $x=y$, or when $f(x)=f(y)$. Also, what we call line is not exactly a set, but rather a dynamic object, think trajectory of a point on that line. We will be back to this later, once we will know more about such maps.

Here are some basic examples of symmetries, all being linear in the above sense:

Proposition 1.2. *The symmetries with respect to Ox and Oy are:*

$$\begin{pmatrix}x\\ y\end{pmatrix}\to\begin{pmatrix}x\\ -y\end{pmatrix}\quad,\quad\begin{pmatrix}x\\ y\end{pmatrix}\to\begin{pmatrix}-x\\ y\end{pmatrix}$$

The symmetries with respect to the $x=y$ and $x=-y$ diagonals are:

$$\begin{pmatrix}x\\ y\end{pmatrix}\to\begin{pmatrix}y\\ x\end{pmatrix}\quad,\quad\begin{pmatrix}x\\ y\end{pmatrix}\to\begin{pmatrix}-y\\ -x\end{pmatrix}$$

All these maps are linear, in the above sense.

Proof. The fact that all these maps are linear is clear, because they map lines to lines, in our sense, and they also map 0 to 0. As for the explicit formulae in the statement, these are clear as well, by drawing pictures for each of the maps involved. $\square$

Here are now some basic examples of rotations, once again all being linear:

PROPOSITION 1.3. *The rotations of angle* $0°$ *and of angle* $90°$ *are:*

$$\begin{pmatrix}x\\y\end{pmatrix}\to\begin{pmatrix}x\\y\end{pmatrix}\quad,\quad\begin{pmatrix}x\\y\end{pmatrix}\to\begin{pmatrix}-y\\x\end{pmatrix}$$

The rotations of angle $180°$ *and of angle* $270°$ *are:*

$$\begin{pmatrix}x\\y\end{pmatrix}\to\begin{pmatrix}-x\\-y\end{pmatrix}\quad,\quad\begin{pmatrix}x\\y\end{pmatrix}\to\begin{pmatrix}y\\-x\end{pmatrix}$$

All these maps are linear, in the above sense.

PROOF. As before, these rotations are all linear, for obvious reasons. As for the formulae in the statement, these are clear as well, by drawing pictures. □

Here are some basic examples of projections, once again all being linear:

PROPOSITION 1.4. *The projections on* Ox *and* Oy *are:*

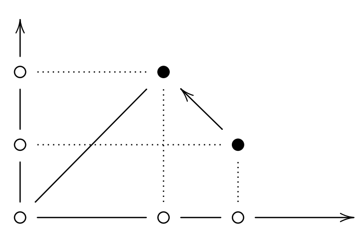
$$\begin{pmatrix}x\\y\end{pmatrix}\to\begin{pmatrix}x\\0\end{pmatrix}\quad,\quad\begin{pmatrix}x\\y\end{pmatrix}\to\begin{pmatrix}0\\y\end{pmatrix}$$

The projections on the $x=y$ *and* $x=-y$ *diagonals are:*

$$\begin{pmatrix}x\\y\end{pmatrix}\to\frac{1}{2}\begin{pmatrix}x+y\\x+y\end{pmatrix}\quad,\quad\begin{pmatrix}x\\y\end{pmatrix}\to\frac{1}{2}\begin{pmatrix}x-y\\y-x\end{pmatrix}$$

All these maps are linear, in the above sense.

PROOF. Again, these projections are all linear, and the formulae are clear as well, by drawing pictures, with only the last 2 formulae needing some explanations. In what regards the projection on the $x=y$ diagonal, the picture here is as follows:



But this gives the result, since the $45°$ triangle shows that this projection leaves invariant $x+y$, so we can only end up with the average $(x+y)/2$, as double coordinate. As for the projection on the $x=-y$ diagonal, the proof here is similar. □

Finally, we have the translations, which are as follows:

PROPOSITION 1.5. *The translations are exactly the maps of the form*

$$\begin{pmatrix}x\\y\end{pmatrix}\to\begin{pmatrix}x+p\\y+q\end{pmatrix}$$

with $p,q\in\mathbb{R}$, and these maps are all affine, in the above sense.

PROOF. A translation $f:\mathbb{R}^2\to\mathbb{R}^2$ is clearly affine, because it maps lines to lines. Also, such a translation is uniquely determined by the following vector:

$$f\begin{pmatrix}0\\0\end{pmatrix}=\begin{pmatrix}p\\q\end{pmatrix}$$

To be more precise, f must be the map which takes a vector $\binom{x}{y}$, and adds this vector $\binom{p}{q}$ to it. But this gives the formula in the statement. $\square$

Summarizing, we have many interesting examples of linear and affine maps. Let us develop now some general theory, for such maps. As a first result, we have:

THEOREM 1.6. *For a map $f:\mathbb{R}^2\to\mathbb{R}^2$, the following are equivalent:*

1. *f is linear in our sense, mapping lines to lines, and 0 to 0.*
2. *f maps sums to sums, $f(x+y)=f(x)+f(y)$, and satisfies $f(\lambda x)=\lambda f(x)$.*

PROOF. This is something which comes from definitions, as follows:

(1) $\implies$ (2) We know that f satisfies the following equation, and $f(0)=0$:

$$f(tx+(1-t)y)=tf(x)+(1-t)f(y)$$

By setting $y=0$, and by using our assumption $f(0)=0$, we obtain, as desired:

$$f(tx)=tf(x)$$

As for the first condition, regarding sums, this can be established as follows:

$$\begin{aligned}f(x+y)&=f\left(2\cdot\frac{x+y}{2}\right)\\&=2f\left(\frac{x+y}{2}\right)\\&=2\cdot\frac{f(x)+f(y)}{2}\\&=f(x)+f(y)\end{aligned}$$

(2) $\implies$ (1) Conversely now, assuming that f satisfies $f(x+y)=f(x)+f(y)$ and $f(\lambda x)=\lambda f(x)$, then f must map lines to lines, as shown by:

$$\begin{aligned}f(tx+(1-t)y)&=f(tx)+f((1-t)y)\\&=tf(x)+(1-t)f(y)\end{aligned}$$

Also, we have $f(0)=f(2\cdot 0)=2f(0)$, which gives $f(0)=0$, as desired. $\square$

The above result is very useful, and in practice, we will often use the condition (2) there, somewhat as a new definition for the linear maps. Let us record this as follows:

DEFINITION 1.7 (upgrade). *A map $f:\mathbb{R}^2\to\mathbb{R}^2$ is called:*

(1) *Linear, when it satisfies $f(x+y)=f(x)+f(y)$ and $f(\lambda x)=\lambda f(x)$.*
(2) *Affine, when it is of the form $f=g+x$, with g linear, and $x\in\mathbb{R}^2$.*

Before getting into the mathematics of linear maps, let us comment a bit more on the "maps lines to lines" feature of such maps. As mentioned after Definition 1.1, this feature requires thinking at lines as being "dynamic" objects, the point being that, when thinking at lines as being sets, this interpretation fails, as shown by the following map:

$$f\begin{pmatrix}x\\ y\end{pmatrix}=\begin{pmatrix}x^3\\ 0\end{pmatrix}$$

However, in relation with all this we have the following useful result:

THEOREM 1.8. *For a continuous injective $f:\mathbb{R}^2\to\mathbb{R}^2$, the following are equivalent:*

(1) *f is affine in our sense, mapping lines to lines.*
(2) *f maps set-theoretical lines to set-theoretical lines.*

PROOF. By composing f with a translation, we can assume that we have $f(0)=0$. With this assumption made, the proof goes as follows:

(1) $\implies$ (2) This is clear from definitions.

(2) $\implies$ (1) Let us first prove that we have $f(x+y)=f(x)+f(y)$. We do this first in the case where our vectors are not proportional, $x\not\sim y$. In this case we have a proper parallelogram $(0,x,y,x+y)$, and since f was assumed to be injective, it must map parallel lines to parallel lines, and so must map our parallelogram into a parallelogram $(0,f(x),f(y),f(x+y))$. But this latter parallelogram shows that we have:

$$f(x+y)=f(x)+f(y)$$

In the remaining case where our vectors are proportional, $x\sim y$, we can pick a sequence $x_n\to x$ satisfying $x_n\not\sim y$ for any n, and we obtain, as desired:

$$\begin{aligned}x_n\to x, x_n\not\sim y,\forall n &\implies f(x_n+y)=f(x_n)+f(y),\forall n\\ &\implies f(x+y)=f(x)+f(y)\end{aligned}$$

Regarding now $f(\lambda x)=\lambda f(x)$, since f maps lines to lines, it must map the line $0-x$ to the line $0-f(x)$, so we have a formula as follows, for any λ,x:

$$f(\lambda x)=\varphi_x(\lambda)f(x)$$

But since f maps parallel lines to parallel lines, by Thales the function $\varphi_x:\mathbb{R}\to\mathbb{R}$ does not depend on x. Thus, we have a formula as follows, for any λ,x:

$$f(\lambda x)=\varphi(\lambda)f(x)$$

We know that we have $\varphi(0)=0$ and $\varphi(1)=1$, and we must prove that we have $\varphi(\lambda)=\lambda$ for any λ. For this purpose, we use a trick. On one hand, we have:

$$f((\lambda+\mu)x)=\varphi(\lambda+\mu)f(x)$$

On the other hand, since f maps sums to sums, we have as well:

$$\begin{aligned}
f((\lambda+\mu)x) &= f(\lambda x)+f(\mu x)\\
&= \varphi(\lambda)f(x)+\varphi(\mu)f(x)\\
&= (\varphi(\lambda)+\varphi(\mu))f(x)
\end{aligned}$$

Thus our rescaling function $\varphi:\mathbb{R}\to\mathbb{R}$ satisfies the following conditions:

$$\varphi(0)=0\quad,\quad\varphi(1)=1\quad,\quad\varphi(\lambda+\mu)=\varphi(\lambda)+\varphi(\mu)$$

But with these conditions in hand, it is clear that we have $\varphi(\lambda)=\lambda$, first for all the inverses of integers, $\lambda=1/n$ with $n\in\mathbb{N}$, then for all rationals, $\lambda\in\mathbb{Q}$, and finally by continuity for all reals, $\lambda\in\mathbb{R}$. Thus, we have proved the following formula:

$$f(\lambda x)=\lambda f(x)$$

But this finishes the proof of (2) $\implies$ (1), and we are done. $\square$

All this is very nice, and there are some further things that can be said, but getting to business, Definition 1.7 is what we need. Indeed, we have the following powerful result, stating that the linear/affine maps $f:\mathbb{R}^2\to\mathbb{R}^2$ are fully described by 4/6 parameters:

THEOREM 1.9. *The linear maps $f:\mathbb{R}^2\to\mathbb{R}^2$ are precisely the maps of type*

$$f\begin{pmatrix}x\\y\end{pmatrix}=\begin{pmatrix}ax+by\\cx+dy\end{pmatrix}$$

and the affine maps $f:\mathbb{R}^2\to\mathbb{R}^2$ are precisely the maps of type

$$f\begin{pmatrix}x\\y\end{pmatrix}=\begin{pmatrix}ax+by\\cx+dy\end{pmatrix}+\begin{pmatrix}p\\q\end{pmatrix}$$

with the conventions from Definition 1.7 for such maps.

PROOF. Assuming that f is linear in the sense of Definition 1.7, we have:

$$\begin{aligned}
f\begin{pmatrix}x\\y\end{pmatrix} &= f\left(\begin{pmatrix}x\\0\end{pmatrix}+\begin{pmatrix}0\\y\end{pmatrix}\right)\\
&= f\begin{pmatrix}x\\0\end{pmatrix}+f\begin{pmatrix}0\\y\end{pmatrix}\\
&= f\left(x\begin{pmatrix}1\\0\end{pmatrix}\right)+f\left(y\begin{pmatrix}0\\1\end{pmatrix}\right)\\
&= xf\begin{pmatrix}1\\0\end{pmatrix}+yf\begin{pmatrix}0\\1\end{pmatrix}
\end{aligned}$$

Thus, we obtain the formula in the statement, with $a,b,c,d\in\mathbb R$ being given by:

$$f\begin{pmatrix}1\\0\end{pmatrix}=\begin{pmatrix}a\\c\end{pmatrix}\quad,\quad f\begin{pmatrix}0\\1\end{pmatrix}=\begin{pmatrix}b\\d\end{pmatrix}$$

In the affine case now, we have as extra piece of data a vector, as follows:

$$f\begin{pmatrix}0\\0\end{pmatrix}=\begin{pmatrix}p\\q\end{pmatrix}$$

Indeed, if $f:\mathbb R^2\to\mathbb R^2$ is affine, then the following map is linear:

$$f-\begin{pmatrix}p\\q\end{pmatrix}:\mathbb R^2\to\mathbb R^2$$

Thus, by using the formula in (1) we obtain the result. $\square$

Moving ahead now, Theorem 1.9 is all that we need for doing some non-trivial mathematics, and so in practice, that will be our "definition" for the linear and affine maps. In order to simplify now all that, which might be a bit complicated to memorize, the idea will be to put our parameters a,b,c,d into a matrix, in the following way:

DEFINITION 1.10. *A matrix $A\in M_2(\mathbb R)$ is an array as follows:*

$$A=\begin{pmatrix}a&b\\c&d\end{pmatrix}$$

These matrices act on the vectors in the following way,

$$\begin{pmatrix}a&b\\c&d\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}=\begin{pmatrix}ax+by\\cx+dy\end{pmatrix}$$

the rule being "multiply the rows of the matrix by the vector".

The above multiplication formula might seem a bit complicated, at a first glance, but it is not. Here is an example for it, quickly worked out:

$$\begin{pmatrix}1&2\\5&6\end{pmatrix}\begin{pmatrix}3\\1\end{pmatrix}=\begin{pmatrix}1\cdot3+2\cdot1\\5\cdot3+6\cdot1\end{pmatrix}=\begin{pmatrix}5\\21\end{pmatrix}$$

As already mentioned, all this comes from our findings from Theorem 1.9. Indeed, with the above multiplication convention for matrices and vectors, we can turn Theorem 1.9 into something much simpler, and better-looking, as follows:

THEOREM 1.11. *The linear maps $f:\mathbb R^2\to\mathbb R^2$ are precisely the maps of type*

$$f(v)=Av$$

and the affine maps $f:\mathbb R^2\to\mathbb R^2$ are precisely the maps of type

$$f(v)=Av+w$$

with A being a 2×2 matrix, and with $v,w\in\mathbb R^2$ being vectors, written vertically.

Proof. With the above conventions, the formulae in Theorem 1.9 read:

$$f\begin{pmatrix}x\\ y\end{pmatrix}=\begin{pmatrix}a&b\\ c&d\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix}$$

$$f\begin{pmatrix}x\\ y\end{pmatrix}=\begin{pmatrix}a&b\\ c&d\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix}+\begin{pmatrix}p\\ q\end{pmatrix}$$

But these are exactly the formulae in the statement, with:

$$A=\begin{pmatrix}a&b\\ c&d\end{pmatrix}\quad,\quad v=\begin{pmatrix}x\\ y\end{pmatrix}\quad,\quad w=\begin{pmatrix}p\\ q\end{pmatrix}$$

Thus, we have proved our theorem. □

Before going further, let us discuss some examples. First, we have:

Proposition 1.12. *The symmetries with respect to Ox and Oy are given by:*

$$\begin{pmatrix}1&0\\ 0&-1\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix}\quad,\quad\begin{pmatrix}-1&0\\ 0&1\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix}$$

The symmetries with respect to the $x=y$ and $x=-y$ diagonals are given by:

$$\begin{pmatrix}0&1\\ 1&0\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix}\quad,\quad\begin{pmatrix}0&-1\\ -1&0\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix}$$

Proof. According to Proposition 1.2, the above transformations map $\binom{x}{y}$ to:

$$\begin{pmatrix}x\\ -y\end{pmatrix}\quad,\quad\begin{pmatrix}-x\\ y\end{pmatrix}\quad,\quad\begin{pmatrix}y\\ x\end{pmatrix}\quad,\quad\begin{pmatrix}-y\\ -x\end{pmatrix}$$

But this gives the formulae in the statement, by guessing in each case the matrix which does the job, in the obvious way. □

Regarding now the basic rotations, we have here:

Proposition 1.13. *The rotations of angle 0° and of angle 90° are given by:*

$$\begin{pmatrix}1&0\\ 0&1\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix}\quad,\quad\begin{pmatrix}0&-1\\ 1&0\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix}$$

The rotations of angle 180° and of angle 270° are given by:

$$\begin{pmatrix}-1&0\\ 0&-1\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix}\quad,\quad\begin{pmatrix}0&1\\ -1&0\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix}$$

Proof. As before, but by using Proposition 1.3, the vector $\binom{x}{y}$ maps to:

$$\begin{pmatrix}x\\ y\end{pmatrix}\quad,\quad\begin{pmatrix}-y\\ x\end{pmatrix}\quad,\quad\begin{pmatrix}-x\\ -y\end{pmatrix}\quad,\quad\begin{pmatrix}y\\ -x\end{pmatrix}$$

But this gives the formulae in the statement, as before by guessing the matrix. □

Finally, regarding the basic projections, we have here:

PROPOSITION 1.14. *The projections on Ox and Oy are given by:*

$$\begin{pmatrix}1&0\\0&0\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}\quad,\quad\begin{pmatrix}0&0\\0&1\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}$$

The projections on the $x=y$ and $x=-y$ diagonals are given by:

$$\frac{1}{2}\begin{pmatrix}1&1\\1&1\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}\quad,\quad\frac{1}{2}\begin{pmatrix}1&-1\\-1&1\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}$$

PROOF. As before, but according now to Proposition 1.4, the vector $\binom{x}{y}$ maps to:

$$\begin{pmatrix}x\\0\end{pmatrix}\quad,\quad\begin{pmatrix}0\\y\end{pmatrix}\quad,\quad\frac{1}{2}\begin{pmatrix}x+y\\x+y\end{pmatrix}\quad,\quad\frac{1}{2}\begin{pmatrix}x-y\\y-x\end{pmatrix}$$

But this gives the formulae in the statement, as before by guessing the matrix. □

In addition to the above transformations, there are many other examples. We have for instance the null transformation, which is given by:

$$\begin{pmatrix}0&0\\0&0\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix}$$

Here is now a more bizarre map, which can still be understood, however, as being the map which "switches the coordinates, then kills the second one":

$$\begin{pmatrix}0&1\\0&0\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}=\begin{pmatrix}y\\0\end{pmatrix}$$

Even more bizarrely now, here is a certain linear map, whose interpretation is more complicated, and is left to you, reader:

$$\begin{pmatrix}1&1\\0&0\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}=\begin{pmatrix}x+y\\0\end{pmatrix}$$

And here is another linear map, which once again, being something geometric, in 2 dimensions, can definitely be understood, at least in theory:

$$\begin{pmatrix}1&1\\0&1\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}=\begin{pmatrix}x+y\\y\end{pmatrix}$$

Let us discuss now the computation of the arbitrary symmetries, rotations and projections. We begin with the rotations, whose formula is a must-know:

THEOREM 1.15. *The rotation of angle $t\in\mathbb{R}$ is given by the matrix*

$$R_t=\begin{pmatrix}\cos t&-\sin t\\\sin t&\cos t\end{pmatrix}$$

depending on $t\in\mathbb{R}$ taken modulo 2π.

Proof. The rotation being linear, it must correspond to a certain matrix:

$$R_t = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

We can guess this matrix, via its action on the basic coordinate vectors $\binom{1}{0}$ and $\binom{0}{1}$. A quick picture shows that we must have:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \cos t \\ \sin t \end{pmatrix}$$

Also, by paying attention to positives and negatives, we must have:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -\sin t \\ \cos t \end{pmatrix}$$

Guessing now the matrix is not complicated, because the first equation gives us the first column, and the second equation gives us the second column:

$$\begin{pmatrix} a \\ c \end{pmatrix} = \begin{pmatrix} \cos t \\ \sin t \end{pmatrix} \quad , \quad \begin{pmatrix} b \\ d \end{pmatrix} = \begin{pmatrix} -\sin t \\ \cos t \end{pmatrix}$$

Thus, we can just put together these two vectors, and we obtain our matrix. $\square$

Regarding now the symmetries, the formula here is as follows:

Theorem 1.16. *The symmetry with respect to the Ox axis rotated by an angle $t/2 \in \mathbb{R}$ is given by the matrix*

$$S_t = \begin{pmatrix} \cos t & \sin t \\ \sin t & -\cos t \end{pmatrix}$$

depending on $t \in \mathbb{R}$ taken modulo 2π.

Proof. As before, we can guess the matrix via its action on the basic coordinate vectors $\binom{1}{0}$ and $\binom{0}{1}$. A quick picture shows that we must have:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \cos t \\ \sin t \end{pmatrix}$$

Also, by paying attention to positives and negatives, we must have:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} \sin t \\ -\cos t \end{pmatrix}$$

Guessing now the matrix is not complicated, because we must have:

$$\begin{pmatrix} a \\ c \end{pmatrix} = \begin{pmatrix} \cos t \\ \sin t \end{pmatrix} \quad , \quad \begin{pmatrix} b \\ d \end{pmatrix} = \begin{pmatrix} \sin t \\ -\cos t \end{pmatrix}$$

Thus, we can just put together these two vectors, and we obtain our matrix. $\square$

Finally, regarding the projections, the formula here is as follows:

THEOREM 1.17. *The projection on the Ox axis rotated by an angle $t/2 \in \mathbb{R}$ is given by the matrix*

$$P_t = \frac{1}{2}\begin{pmatrix}1+\cos t&\sin t\\ \sin t&1-\cos t\end{pmatrix}$$

depending on $t \in \mathbb{R}$ taken modulo 2π.

PROOF. We will need here some trigonometry, and more precisely the formulae for the duplication of the angles. Regarding the sine, the formula here is:

$$\sin(2t) = 2\sin t\cos t$$

Regarding the cosine, we have here 3 equivalent formulae, as follows:

$$\begin{aligned}\cos(2t) &= \cos^2 t - \sin^2 t\\ &= 2\cos^2 t - 1\\ &= 1 - 2\sin^2 t\end{aligned}$$

Getting back now to our problem, some quick pictures, using similarity of triangles, and then the above trigonometry formulae, show that we must have:

$$P_t\begin{pmatrix}1\\0\end{pmatrix} = \cos\frac{t}{2}\begin{pmatrix}\cos\frac{t}{2}\\ \sin\frac{t}{2}\end{pmatrix} = \frac{1}{2}\begin{pmatrix}1+\cos t\\ \sin t\end{pmatrix}$$

$$P_t\begin{pmatrix}0\\1\end{pmatrix} = \sin\frac{t}{2}\begin{pmatrix}\cos\frac{t}{2}\\ \sin\frac{t}{2}\end{pmatrix} = \frac{1}{2}\begin{pmatrix}\sin t\\ 1-\cos t\end{pmatrix}$$

Now by putting together these two vectors, and we obtain our matrix. □

1b. Matrix calculus

In order to formulate now our second theorem, dealing with compositions of maps, let us make the following multiplication convention, between matrices and matrices:

$$\begin{pmatrix}a&b\\c&d\end{pmatrix}\begin{pmatrix}p&q\\r&s\end{pmatrix} = \begin{pmatrix}ap+br&aq+bs\\cp+dr&cq+ds\end{pmatrix}$$

This might look a bit complicated, but as before, in what was concerning multiplying matrices and vectors, the idea is very simple, namely "multiply the rows of the first matrix by the columns of the second matrix". With this convention, we have:

THEOREM 1.18. *If we denote by $f_A : \mathbb{R}^2 \to \mathbb{R}^2$ the linear map associated to a matrix A, given by the formula*

$$f_A(v) = Av$$

then we have the following multiplication formula for such maps:

$$f_A f_B = f_{AB}$$

That is, the composition of linear maps corresponds to the multiplication of matrices.

Proof. We want to prove that we have the following formula, valid for any two matrices $A,B\in M_2(\mathbb{R})$, and any vector $v\in\mathbb{R}^2$:

$$A(Bv)=(AB)v$$

For this purpose, let us write our matrices and vector as follows:

$$A=\begin{pmatrix}a&b\\ c&d\end{pmatrix}\quad,\quad B=\begin{pmatrix}p&q\\ r&s\end{pmatrix}\quad,\quad v=\begin{pmatrix}x\\ y\end{pmatrix}$$

The formula that we want to prove becomes:

$$\begin{pmatrix}a&b\\ c&d\end{pmatrix}\left[\begin{pmatrix}p&q\\ r&s\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix}\right]=\left[\begin{pmatrix}a&b\\ c&d\end{pmatrix}\begin{pmatrix}p&q\\ r&s\end{pmatrix}\right]\begin{pmatrix}x\\ y\end{pmatrix}$$

But this is the same as saying that:

$$\begin{pmatrix}a&b\\ c&d\end{pmatrix}\begin{pmatrix}px+qy\\ rx+sy\end{pmatrix}=\begin{pmatrix}ap+br&aq+bs\\ cp+dr&cq+ds\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix}$$

And this latter formula does hold indeed, because on both sides we get:

$$\begin{pmatrix}apx+aqy+brx+bsy\\ cpx+cqy+drx+dsy\end{pmatrix}$$

Thus, we have proved the result. □

As a verification for the above result, let us compose two rotations. The computation here is as follows, yieding a rotation, as it should, and of the correct angle:

$$\begin{aligned}R_sR_t&=\begin{pmatrix}\cos s&-\sin s\\ \sin s&\cos s\end{pmatrix}\begin{pmatrix}\cos t&-\sin t\\ \sin t&\cos t\end{pmatrix}\\&=\begin{pmatrix}\cos s\cos t-\sin s\sin t&-\cos s\sin t-\sin t\cos s\\ \sin s\cos t+\cos s\sin t&-\sin s\sin t+\cos s\cos t\end{pmatrix}\\&=\begin{pmatrix}\cos(s+t)&-\sin(s+t)\\ \sin(s+t)&\cos(s+t)\end{pmatrix}\\&=R_{s+t}\end{aligned}$$

We are ready now to pass to 3 dimensions. The idea is to select what we learned in 2 dimensions, nice results only, and generalize to 3 dimensions. We obtain:

Theorem 1.19. *Consider a map $f:\mathbb{R}^3\to\mathbb{R}^3$.*

1. *f is linear when it is of the form $f(v)=Av$, with $A\in M_3(\mathbb{R})$.*
2. *f is affine when $f(v)=Av+w$, with $A\in M_3(\mathbb{R})$ and $w\in\mathbb{R}^3$.*
3. *We have the composition formula $f_Af_B=f_{AB}$, similar to the 2D one.*

PROOF. Here (1,2) can be proved exactly as in the 2D case, with the multiplication convention being as usual, "multiply the rows of the matrix by the vector":

$$\begin{pmatrix}a&b&c\\d&e&f\\g&h&i\end{pmatrix}\begin{pmatrix}x\\y\\z\end{pmatrix}=\begin{pmatrix}ax+by+cz\\dx+ey+fz\\gx+hy+iz\end{pmatrix}$$

As for (3), once again the 2D idea applies, with the same product rule, "multiply the rows of the first matrix by the columns of the second matrix":

$$\begin{pmatrix}a&b&c\\d&e&f\\g&h&i\end{pmatrix}\begin{pmatrix}p&q&r\\s&t&u\\v&w&x\end{pmatrix}=\begin{pmatrix}ap+bs+cv&aq+bt+cw&ar+bu+cx\\dp+es+fv&dq+et+fw&dr+eu+fx\\gp+hs+iv&gq+ht+iw&gr+hu+ix\end{pmatrix}$$

Thus, we have proved our theorem. Of course, we are going a bit fast here, and some verifications are missing, but we will discuss all this in detail, in N dimensions. $\square$

We are now ready to discuss 4 and more dimensions. Before doing so, let us point out however that the maps of type $f:\mathbb{R}^3\to\mathbb{R}^2$, or $f:\mathbb{R}\to\mathbb{R}^2$, and so on, are not covered by our results. Since there are many interesting such maps, say obtained by projecting and then rotating, and so on, we will be interested here in the maps $f:\mathbb{R}^N\to\mathbb{R}^M$.

A bit of thinking suggests that such maps should come from the $M\times N$ matrices. Indeed, this is what happens at $M=N=2$ and $M=N=3$, of course. But this happens as well at $N=1$, because a linear map $f:\mathbb{R}\to\mathbb{R}^M$ can only be something of the form $f(\lambda)=\lambda v$, with $v\in\mathbb{R}^M$. But $v\in\mathbb{R}^M$ means that v is a $M\times 1$ matrix. So, let us start with the product rule for such matrices, which is as follows:

DEFINITION 1.20. *We can multiply the $M\times N$ matrices with $N\times K$ matrices,*

$$\begin{pmatrix}a_{11}&\ldots&a_{1N}\\\vdots&&\vdots\\a_{M1}&\ldots&a_{MN}\end{pmatrix}\begin{pmatrix}b_{11}&\ldots&b_{1K}\\\vdots&&\vdots\\b_{N1}&\ldots&b_{NK}\end{pmatrix}$$

the product being the $M\times K$ matrix given by the following formula,

$$\begin{pmatrix}a_{11}b_{11}+\ldots+a_{1N}b_{N1}&\ldots\ldots&a_{11}b_{1K}+\ldots+a_{1N}b_{NK}\\\vdots&&\vdots\\\vdots&&\vdots\\a_{M1}b_{11}+\ldots+a_{MN}b_{N1}&\ldots\ldots&a_{M1}b_{1K}+\ldots+a_{MN}b_{NK}\end{pmatrix}$$

obtained via the usual rule "multiply rows by columns".

Observe that this formula generalizes all the multiplication rules that we have been using so far, between various types of matrices and vectors. Thus, in practice, we can simply forget all the previous multiplication rules, and simply memorize this one.

In case the above formula looks hard to memorize, here is an alternative formulation of it, which is simpler and more powerful, by using the standard algebraic notation for the matrices, $A = (A_{ij})$, that we will heavily use, in what follows:

PROPOSITION 1.21. *The matrix multiplication is given by formula*

$$(AB)_{ij} = \sum_k A_{ik} B_{kj}$$

with A_{ij} standing for the entry of A at row i and column j.

PROOF. This is indeed just a shorthand for the formula in Definition 1.20, by following the rule there, namely "multiply the rows of A by the columns of B". □

As an illustration for the power of the convention in Proposition 1.21, we have:

PROPOSITION 1.22. *We have the following formula, valid for any matrices A, B, C,*

$$(AB)C = A(BC)$$

provided that the sizes of our matrices A, B, C fit.

PROOF. We have the following computation, using indices as above:

$$((AB)C)_{ij} = \sum_k (AB)_{ik} C_{kj} = \sum_{kl} A_{il} B_{lk} C_{kj}$$

On the other hand, we have as well the following computation:

$$(A(BC))_{ij} = \sum_l A_{il} (BC)_{lj} = \sum_{kl} A_{il} B_{lk} C_{kj}$$

Thus we have $(AB)C = A(BC)$, and we have proved our result. □

We can now talk about linear maps between spaces of arbitrary dimension, generalizing what we have been doing so far. The main result here is as follows:

THEOREM 1.23. *Consider a map $f : \mathbb{R}^N \to \mathbb{R}^M$.*

(1) *f is linear when it is of the form $f(v) = Av$, with $A \in M_{M\times N}(\mathbb{R})$.*
(2) *f is affine when $f(v) = Av + w$, with $A \in M_{M\times N}(\mathbb{R})$ and $w \in \mathbb{R}^M$.*
(3) *We have the composition formula $f_A f_B = f_{AB}$, whenever the sizes fit.*

PROOF. We already know that this happens at $M = N = 2$, and at $M = N = 3$ as well. In general, the proof is similar, by doing some elementary computations. □

As a first example here, we have the identity matrix, acting as the identity:

$$\begin{pmatrix}1&&0\\&\ddots&\\0&&1\end{pmatrix}\begin{pmatrix}x_1\\\vdots\\x_N\end{pmatrix}=\begin{pmatrix}x_1\\\vdots\\x_N\end{pmatrix}$$

We have as well the null matrix, acting as the null map:

$$\begin{pmatrix}0&\ldots&0\\ \vdots&&\vdots\\ 0&\ldots&0\end{pmatrix}\begin{pmatrix}x_1\\ \vdots\\ x_N\end{pmatrix}=\begin{pmatrix}0\\ \vdots\\ 0\end{pmatrix}$$

Here is now an important result, providing us with many examples:

Proposition 1.24. *The diagonal matrices act as follows:*

$$\begin{pmatrix}\lambda_1&&0\\ &\ddots&\\ 0&&\lambda_N\end{pmatrix}\begin{pmatrix}x_1\\ \vdots\\ x_N\end{pmatrix}=\begin{pmatrix}\lambda_1x_1\\ \vdots\\ \lambda_Nx_N\end{pmatrix}$$

Proof. This is clear, indeed, from definitions. □

As a more specialized example now, we have:

Proposition 1.25. *The flat matrix, which is as follows,*

$$\mathbb{I}_N=\begin{pmatrix}1&\ldots&1\\ \vdots&&\vdots\\ 1&\ldots&1\end{pmatrix}$$

acts via N times the projection on the all-one vector.

Proof. The flat matrix acts in the following way:

$$\begin{pmatrix}1&\ldots&1\\ \vdots&&\vdots\\ 1&\ldots&1\end{pmatrix}\begin{pmatrix}x_1\\ \vdots\\ x_N\end{pmatrix}=\begin{pmatrix}x_1+\ldots+x_N\\ \vdots\\ x_1+\ldots+x_N\end{pmatrix}$$

Thus, in terms of the matrix $P=\mathbb{I}_N/N$, we have the following formula:

$$P\begin{pmatrix}x_1\\ \vdots\\ x_N\end{pmatrix}=\frac{x_1+\ldots+x_N}{N}\begin{pmatrix}1\\ \vdots\\ 1\end{pmatrix}$$

Since the linear map $f(x)=Px$ satisfies $f^2=f$, and since $Im(f)$ consists of the scalar multiples of the all-one vector $\xi\in\mathbb{R}^N$, we conclude that f is a projection on $\mathbb{R}\xi$. Also, with the standard scalar product convention $<x,y>=\sum x_iy_i$, we have:

$$\begin{aligned}<f(x)-x,\xi> &= <f(x),\xi>-<x,\xi>\\ &= \frac{\sum x_i}{N}\times N-\sum x_i\\ &= 0\end{aligned}$$

Thus, our projection is an orthogonal projection, and we are done. □

1c. Diagonalization

Let us develop now some general theory for the square matrices. We will need the following standard result, regarding the changes of coordinates in $\mathbb{R}^N$:

THEOREM 1.26. *For a system $\{v_1,\ldots,v_N\}\subset\mathbb{R}^N$, the following are equivalent:*

(1) *The vectors v_i form a basis of $\mathbb{R}^N$, in the sense that each vector $x\in\mathbb{R}^N$ can be written in a unique way as a linear combination of these vectors:*

$$x=\sum\lambda_iv_i$$

(2) *The following linear map associated to these vectors is bijective:*

$$f:\mathbb{R}^N\to\mathbb{R}^N\quad,\quad\lambda\to\sum\lambda_iv_i$$

(3) *The matrix formed by these vectors, regarded as usual as column vectors,*

$$P=[v_1,\ldots,v_N]\in M_N(\mathbb{R})$$

is invertible, with respect to the usual multiplication of the matrices.

PROOF. Here the equivalence (1) $\iff$ (2) is clear from definitions, and the equivalence (2) $\iff$ (3) is clear as well, because we have $f(x)=Px$. $\square$

Getting back now to the matrices, as an important definition, we have:

DEFINITION 1.27. *Let $A\in M_N(\mathbb{R})$ be a square matrix. We say that $v\in\mathbb{R}^N$ is an eigenvector of A, with corresponding eigenvalue $\lambda\in\mathbb{R}^N$, when:*

$$Av=\lambda v$$

Also, we say that A is diagonalizable when $\mathbb{R}^N$ has a basis formed by eigenvectors of A.

We will see in a moment examples of eigenvectors and eigenvalues, and of diagonalizable matrices. However, even before seeing the examples, it is quite clear that these are key notions. Indeed, for a matrix $A\in M_N(\mathbb{R})$, being diagonalizable is the best thing that can happen, because in this case, once the basis changed, A becomes diagonal.

To be more precise here, we have the following result:

PROPOSITION 1.28. *Assuming that $A\in M_N(\mathbb{R})$ is diagonalizable, we have the formula*

$$A=\begin{pmatrix}\lambda_1&&\\&\ddots&\\&&\lambda_N\end{pmatrix}$$

with respect to the basis $\{v_1,\ldots,v_N\}$ of $\mathbb{R}^N$ consisting of eigenvectors of A.

PROOF. This is clear from the definition of eigenvalues and eigenvectors, and from the formula of linear maps associated to diagonal matrices, from Proposition 1.24. $\square$

Here is an equivalent form of the above result, which is often used in practice, when we prefer not to change the basis, and stay with the usual basis of $\mathbb{R}^N$:

THEOREM 1.29. *Assuming that $A \in M_N(\mathbb{R})$ is diagonalizable, with*

$$v_1, \ldots, v_N \in \mathbb{R}^N \quad , \quad \lambda_1, \ldots, \lambda_N \in \mathbb{R}$$

as eigenvectors and corresponding eigenvalues, we have the formula

$$A = PDP^{-1}$$

with the matrices $P, D \in M_N(\mathbb{R})$ being given by the formulae

$$P = [v_1, \ldots, v_N] \quad , \quad D = diag(\lambda_1, \ldots, \lambda_N)$$

and respectively called passage matrix, and diagonal form of A.

PROOF. This can be viewed in two possible ways, as follows:

(1) As already mentioned, with respect to the basis $v_1, \ldots, v_N \in \mathbb{R}^N$ formed by the eigenvectors, our matrix A is given by:

$$A = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_N \end{pmatrix}$$

But this corresponds precisely to the formula $A = PDP^{-1}$ from the statement, with P and its inverse appearing there due to our change of basis.

(2) We can equally establish the formula in the statement by a direct computation. Indeed, we have $Pe_i = v_i$, where $\{e_1, \ldots, e_N\}$ is the standard basis of $\mathbb{R}^N$, and so:

$$APe_i = Av_i = \lambda_i v_i$$

On the other hand, once again by using $Pe_i = v_i$, we have as well:

$$PDe_i = P\lambda_i e_i = \lambda_i Pe_i = \lambda_i v_i$$

Thus we have $AP = PD$, and so $A = PDP^{-1}$, as claimed. □

Let us discuss now some basic examples, namely the rotations, symmetries and projections in 2 dimensions. The situation is very simple for the projections, as follows:

PROPOSITION 1.30. *The projection on the Ox axis rotated by an angle $t/2 \in \mathbb{R}$,*

$$P_t = \frac{1}{2}\begin{pmatrix} 1+\cos t & \sin t \\ \sin t & 1-\cos t \end{pmatrix}$$

is diagonalizable, its diagonal form being as follows:

$$P_t \sim \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

PROOF. This is clear, because if we denote by L the line where our projection projects, we can pick any vector $v \in L$, and this will be an eigenvector with eigenvalue 1, and then pick any vector $w \in L^\perp$, and this will be an eigenvector with eigenvalue 0. Thus, even without computations, we are led to the conclusion in the statement. $\square$

The computation for the symmetries is similar, as follows:

PROPOSITION 1.31. *The symmetry with respect to the Ox axis rotated by $t/2 \in \mathbb{R}$,*

$$S_t = \begin{pmatrix}\cos t & \sin t\\ \sin t & -\cos t\end{pmatrix}$$

is diagonalizable, its diagonal form being as follows:

$$S_t \sim \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}$$

PROOF. This is once again clear, because if we denote by L the line with respect to which our symmetry symmetrizes, we can pick any vector $v \in L$, and this will be an eigenvector with eigenvalue 1, and then pick any vector $w \in L^\perp$, and this will be an eigenvector with eigenvalue -1. Thus, we are led to the conclusion in the statement. $\square$

Regarding now the rotations, here the situation is different, as follows:

PROPOSITION 1.32. *The rotation of angle $t \in [0, 2\pi)$, given by the formula*

$$R_t = \begin{pmatrix}\cos t & -\sin t\\ \sin t & \cos t\end{pmatrix}$$

is diagonal at $t = 0, \pi$, and is not diagonalizable at $t \neq 0, \pi$.

PROOF. The first assertion is clear, because at $t = 0, \pi$ the rotations are:

$$R_0 = \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix} \quad , \quad R_\pi = \begin{pmatrix}-1 & 0\\ 0 & -1\end{pmatrix}$$

As for the rotations of angle $t \neq 0, \pi$, these clearly cannot have eigenvectors. $\square$

Finally, here is one more example, which is the most important of them all:

THEOREM 1.33. *The following matrix is not diagonalizable,*

$$J = \begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}$$

because it has only 1 eigenvector.

PROOF. The above matrix, called J en hommage to Jordan, acts as follows:

$$\begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix} = \begin{pmatrix}y\\ 0\end{pmatrix}$$

Thus the eigenvector/eigenvalue equation $Jv = \lambda v$ reads:

$$\begin{pmatrix}y\\0\end{pmatrix} = \begin{pmatrix}\lambda x\\ \lambda y\end{pmatrix}$$

We have then two cases, depending on λ, as follows, which give the result:

(1) For $\lambda \neq 0$ we must have $y = 0$, coming from the second row, and so $x = 0$ as well, coming from the first row, so we have no nontrivial eigenvectors.

(2) As for the case $\lambda = 0$, here we must have $y = 0$, coming from the first row, and so the eigenvectors here are the vectors of the form $\binom{x}{0}$. $\square$

1d. Scalar products

In order to discuss some interesting examples of matrices, and their diagonalization, in arbitrary dimensions, we will need the following standard fact:

PROPOSITION 1.34. *Consider the scalar product on $\mathbb{R}^N$, given by:*

$$< x, y >= \sum_i x_i y_i$$

We have then the following formula, valid for any vectors x, y and any matrix A,

$$< Ax, y >=< x, A^t y >$$

with A^t being the transpose matrix.

PROOF. By linearity, it is enough to prove the above formula on the standard basis vectors $e_1, \ldots, e_N$ of $\mathbb{R}^N$. Thus, we want to prove that for any i, j we have:

$$< Ae_j, e_i >=< e_j, A^t e_i >$$

The scalar product being symmetric, this is the same as proving that:

$$< Ae_j, e_i >=< A^t e_i, e_j >$$

On the other hand, for any matrix M we have the following formula:

$$M_{ij} =< Me_j, e_i >$$

Thus, the formula to be proved simply reads:

$$A_{ij} = (A^t)_{ji}$$

But this precisely the definition of A^t, and we are done. $\square$

With this, we can develop some theory. We first have:

THEOREM 1.35. *The orthogonal projections are the matrices satisfying:*

$$P^2 = P = P^t$$

These projections are diagonalizable, with eigenvalues $0, 1$.

PROOF. It is obvious that a linear map $f(x) = Px$ is a projection precisely when:

$$P^2 = P$$

In order now for this projection to be an orthogonal projection, the condition to be satisfied can be written and then processed as follows:

$$\begin{aligned} < Px - Py, Px - x >= 0 &\iff < x - y, P^tPx - P^tx >= 0 \\ &\iff P^tPx - P^tx = 0 \\ &\iff P^tP - P^t = 0 \end{aligned}$$

Thus we must have $P^t = P^tP$. Now observe that by transposing, we have as well:

$$P = (P^tP)^t = P^t(P^t)^t = P^tP$$

Thus we must have $P = P^t$, as claimed. Finally, regarding the diagonalization assertion, this is clear by taking a basis of $Im(f)$, which consists of 1-eigenvectors, and then completing with 0-eigenvectors, which can be found inside the orthogonal of $Im(f)$. □

Here is now a key computation of such projections:

THEOREM 1.36. *The rank* 1 *projections are given by the formula*

$$P_x = \frac{1}{||x||^2}(x_ix_j)_{ij}$$

where the constant, namely

$$||x|| = \sqrt{\sum_i x_i^2}$$

is the length of the vector.

PROOF. Consider a vector $y \in \mathbb{R}^N$. Its projection on $\mathbb{R}x$ must be a certain multiple of x, and we are led in this way to the following formula:

$$P_xy = \frac{< y, x >}{< x, x >}x = \frac{1}{||x||^2} < y, x > x$$

With this in hand, we can now compute the entries of P_x, as follows:

$$\begin{aligned} (P_x)_{ij} &= < P_xe_j, e_i > \\ &= \frac{1}{||x||^2} < e_j, x >< x, e_i > \\ &= \frac{x_jx_i}{||x||^2} \end{aligned}$$

Thus, we are led to the formula in the statement. □

As an application, we can recover a result that we already know, namely:

Proposition 1.37. *In 2 dimensions, the rank 1 projections, which are the projections on the Ox axis rotated by an angle $t/2 \in [0,\pi)$, are given by the following formula:*

$$P_t = \frac{1}{2}\begin{pmatrix}1+\cos t & \sin t\\ \sin t & 1-\cos t\end{pmatrix}$$

Together with the following two matrices, which are the rank 0 and 2 projections in $\mathbb{R}^2$,

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

these are all the projections in 2 dimensions.

Proof. The first assertion follows from the general formula in Theorem 1.36, by plugging in the following vector, depending on a parameter $s \in [0,\pi)$:

$$x = \begin{pmatrix}\cos s\\ \sin s\end{pmatrix}$$

We obtain in this way the following matrix, which with $t = 2s$ is the one in the statement, via some trigonometry:

$$P_{2s} = \begin{pmatrix}\cos^2 s & \cos s\sin s\\ \cos s\sin s & \sin^2 s\end{pmatrix}$$

As for the second assertion, this is clear from the first one, because outside rank 1 we can only have rank 0 or rank 2, corresponding to the matrices in the statement. □

Here is another interesting application, this time in N dimensions:

Proposition 1.38. *The projection on the all-1 vector $\xi \in \mathbb{R}^N$ is*

$$P_\xi = \frac{1}{N}\begin{pmatrix}1&\ldots&1\\ \vdots&&\vdots\\ 1&\ldots&1\end{pmatrix}$$

with the all-1 matrix on the right being called the flat matrix.

Proof. As already pointed out in the proof of Proposition 1.25, the matrix in the statement acts in the following way:

$$P_\xi\begin{pmatrix}x_1\\ \vdots\\ x_N\end{pmatrix} = \frac{x_1+\ldots+x_N}{N}\begin{pmatrix}1\\ \vdots\\ 1\end{pmatrix}$$

Thus P_ξ is indeed a projection onto $\mathbb{R}\xi$, and the fact that this projection is indeed the orthogonal one follows either by a direct orthogonality computation, or by using the general formula in Theorem 1.36, by plugging in the all-1 vector ξ. □

Let us discuss now, as a final topic of this chapter, the isometries of $\mathbb{R}^N$. We have here the following general result:

THEOREM 1.39. *The linear maps $f : \mathbb{R}^N \to \mathbb{R}^N$ which are isometries, in the sense that they preserve the distances, are those coming from the matrices satisfying:*

$$U^t = U^{-1}$$

These latter matrices are called orthogonal, and they form a set $O_N \subset M_N(\mathbb{R})$ which is stable under taking compositions, and inverses.

PROOF. We have several things to be proved, the idea being as follows:

(1) We recall that we can pass from scalar products to distances, as follows:

$$||x|| = \sqrt{< x, x >}$$

Conversely, we can compute the scalar products in terms of distances, by using the parallelogram identity, which is as follows:

$$\begin{aligned} ||x+y||^2 - ||x-y||^2 &= ||x||^2 + ||y||^2 + 2 < x, y > -||x||^2 - ||y||^2 + 2 < x, y > \\ &= 4 < x, y > \end{aligned}$$

Now given a matrix $U \in M_N(\mathbb{R})$, we have the following equivalences, with the first one coming from the above identities, and with the other ones being clear:

$$\begin{aligned} ||Ux|| = ||x|| &\iff < Ux, Uy >=< x, y > \\ &\iff < x, U^tUy >=< x, y > \\ &\iff U^tUy = y \\ &\iff U^tU = 1 \\ &\iff U^t = U^{-1} \end{aligned}$$

(2) The second assertion is clear from the definition of the isometries, and can be established as well by using matrices, and the $U^t = U^{-1}$ criterion. □

As a basic illustration here, we have:

THEOREM 1.40. *The rotations and symmetries in the plane, given by*

$$R_t = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix} \quad , \quad S_t = \begin{pmatrix} \cos t & \sin t \\ \sin t & -\cos t \end{pmatrix}$$

are isometries. These are all the isometries in 2 dimensions.

PROOF. We already know that R_t is the rotation of angle t. As for S_t, this is the symmetry with respect to the Ox axis rotated by $t/2 \in \mathbb{R}$. But this gives the result, since the isometries in 2 dimensions are either rotations, or symmetries. □

As a conclusion, the set O_N from Theorem 1.39 is a quite fundamental object, with O_2 already consisting of some interesting 2×2 matrices, namely the matrices R_t, S_t. We will be back to O_N, which is a so-called group, and is actually one of the most important examples of groups, on several occasions, in what follows.

1e. Exercises

The key thing in linear algebra is that of geometrically understanding the linear maps $x \to Ax$ associated to the matrices $A \in M_N(\mathbb{R})$. Here is an exercise on this:

EXERCISE 1.41. *Work out the geometric interpretation of the map $f(x) = Ax$, with*

$$A \in M_2(\pm 1)$$

and then discuss as well the diagonalization of these matrices.

To be more precise, there are $2^4 = 16$ matrices here, some of which were already discussed in the above. As a bonus exercise, you can try as well $A \in M_2(0,1)$, which is 16 more matrices. And for the black belt, try $A \in M_2(-1,0,1)$.

EXERCISE 1.42. *Diagonalize explicitely the third flat matrix, namely*

$$\mathbb{I}_3 = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

and then study as well the general case, that of the matrix $\mathbb{I}_N$.

Here we already know from the above that the diagonal form is $D = (N, 0, \ldots, 0)$, and the problem is that of finding the passage matrix P, as to write the diagonalization formula $\mathbb{I}_N = PDP^{-1}$. The case to start with, as a warm-up for the exercise, is $N = 2$, where $\mathbb{I}_2$ is twice the orthogonal projection on the $x = y$ diagonal, which was already discussed in the above. Then, go with $N = 3$, and then with general $N \in \mathbb{N}$.

EXERCISE 1.43. *Work out the trigonometry formulae*

$$\sin(2t) = 2 \sin t \cos t \quad , \quad \cos(2t) = 2 \cos^2 t - 1$$

by using elementary methods, coming from plane geometry.

There are many ways of solving this exercise, and of course enjoy.

EXERCISE 1.44. *Prove that the isometries in 2 dimensions are either rotations, or symmetries, as to complete the proof of Theorem 1.40.*

As before, there are many ways of dealing with this, all being nice geometry.

EXERCISE 1.45. *Develop a theory of angles between the vectors $x, y \in \mathbb{R}^N$, by using the well-known formula*

$$< x, y >= ||x|| \cdot ||y|| \cdot \cos t$$

that you should by the way fully understand first, in $N = 2$ dimensions.

To be more precise, you must first make sure that the above formula holds indeed at $N = 2$, as a theorem. Then, based on this, you can use this formula at $N \geq 3$ too, but this time as a definition for the angle t between x, y. There are many things that can be done here, and the more complete the theory that you develop, the better.

CHAPTER 2

The determinant

2a. Matrix inversion

We have seen in the previous chapter that most of the interesting maps $f:\mathbb R^N\to\mathbb R^N$ that we know, such as the rotations, symmetries and projections, are linear, and can be written in the following form, with $A\in M_N(\mathbb R)$ being a square matrix:

$$f(v)=Av$$

In this chapter we develop more general theory for such linear maps. We are mostly motivated by the following fundamental result, which has countless concrete applications, and which is actually at the origin of the whole linear algebra theory:

THEOREM 2.1. *Any linear system of equations*

$$\begin{cases}a_{11}x_1+a_{12}x_2+\ldots+a_{1N}x_N&=&v_1\\ a_{21}x_1+a_{22}x_2+\ldots+a_{2N}x_N&=&v_2\\ \quad\vdots\\ a_{N1}x_1+a_{N2}x_2+\ldots+a_{NN}x_N&=&v_N\end{cases}$$

can be written in matrix form, as follows,

$$Ax=v$$

and when A is invertible, its solution is given by $x=A^{-1}v$.

PROOF. With linear algebra conventions, our system reads:

$$\begin{pmatrix}a_{11}&a_{12}&\ldots&a_{1N}\\ a_{21}&a_{22}&\ldots&a_{2N}\\ \vdots&&&\vdots\\ a_{N1}&a_{N2}&\ldots&a_{NN}\end{pmatrix}\begin{pmatrix}x_1\\ x_2\\ \vdots\\ x_N\end{pmatrix}=\begin{pmatrix}v_1\\ v_2\\ \vdots\\ v_N\end{pmatrix}$$

Thus, we are led to the conclusions in the statement. $\square$

In practice, we are led to the question of inverting the matrices $A\in M_N(\mathbb R)$. And this is the same question as inverting the linear maps $f:\mathbb R^N\to\mathbb R^N$, due to:

Theorem 2.2. *A linear map $f:\mathbb{R}^N\to\mathbb{R}^N$, written as*

$$f(v)=Av$$

is invertible precisely when A is invertible, and in this case we have $f^{-1}(v)=A^{-1}v$.

Proof. This is something that we basically know, coming from the fact that, with the notation $f_A(v)=Av$, we have the following formula:

$$f_Af_B=f_{AB}$$

Thus, we are led to the conclusion in the statement. □

In order to study invertibility questions, for matrices or linear maps, let us begin with some examples. In the simplest case, in 2 dimensions, the result is as follows:

Theorem 2.3. *We have the following inversion formula, for the 2×2 matrices:*

$$\begin{pmatrix}a&b\\c&d\end{pmatrix}^{-1}=\frac{1}{ad-bc}\begin{pmatrix}d&-b\\-c&a\end{pmatrix}$$

When $ad-bc=0$, the matrix is not invertible.

Proof. We have two assertions to be proved, the idea being as follows:

(1) As a first observation, when $ad-bc=0$ we must have, for some $\lambda\in\mathbb{R}$:

$$b=\lambda a\quad,\quad d=\lambda c$$

Thus our matrix must be of the following special type:

$$\begin{pmatrix}a&b\\c&d\end{pmatrix}=\begin{pmatrix}a&\lambda a\\a&\lambda c\end{pmatrix}$$

But in this case the columns are proportional, and so the linear map associated to the matrix is not invertible, and so the matrix itself is not invertible either.

(2) When $ad-bc\neq 0$, let us look for an inversion formula of the following type:

$$\begin{pmatrix}a&b\\c&d\end{pmatrix}^{-1}=\frac{1}{ad-bc}\begin{pmatrix}*&*\\ *&*\end{pmatrix}$$

We must therefore solve the following equations:

$$\begin{pmatrix}a&b\\c&d\end{pmatrix}\begin{pmatrix}*&*\\ *&*\end{pmatrix}=\begin{pmatrix}ad-bc&0\\0&ad-bc\end{pmatrix}$$

The obvious solution here is as follows:

$$\begin{pmatrix}a&b\\c&d\end{pmatrix}\begin{pmatrix}d&-b\\-c&a\end{pmatrix}=\begin{pmatrix}ad-bc&0\\0&ad-bc\end{pmatrix}$$

Thus, we are led to the formula in the statement. □

2b. The determinant

In order to deal now with the inversion problem in general, for the arbitrary matrices $A \in M_N(\mathbb{R})$, we will use the same method as the one above, at $N = 2$. Let us write indeed our matrix as follows, with $v_1, \ldots, v_N \in \mathbb{R}^N$ being its column vectors:

$$A = [v_1, \ldots, v_N]$$

We know from the general results from chapter 1 that, in order for A to be invertible, the vectors $v_1, \ldots, v_N$ must be linearly independent. Thus, following the observations (1) from the above proof of Theorem 2.3, we are led into the question of understanding when a family of vectors $v_1, \ldots, v_N \in \mathbb{R}^N$ are linearly independent.

In order to deal with this latter question, let us introduce the following notion:

Definition 2.4. *Associated to any vectors $v_1, \ldots, v_N \in \mathbb{R}^N$ is the volume*

$$\det{}^+(v_1 \ldots v_N) = vol < v_1, \ldots, v_N >$$

of the parallelepiped made by these vectors.

Here the volume is taken in the standard N-dimensional sense. At $N = 1$ this volume is a length, at $N = 2$ this volume is an area, at $N = 3$ this is the usual 3D volume, and so on. In general, the volume of a body $X \subset \mathbb{R}^N$ is by definition the number $vol(X) \in [0, \infty]$ of copies of the unit cube $C \subset \mathbb{R}^N$ which are needed for filling X.

In order to compute this volume we can use various geometric techniques, and we will see soon that, in what regards the case that we are interested in, namely that of the parallelepipeds $P \subset \mathbb{R}^N$, we can basically compute here everything, just by using very basic geometric techniques, essentially based on the Thales theorem.

In relation with our inversion problem, we have the following statement:

Proposition 2.5. *The quantity* $\det^+$ *that we constructed, regarded as a function of the corresponding square matrices, formed by column vectors,*

$$\det{}^+ : M_N(\mathbb{R}) \to \mathbb{R}_+$$

has the property that a matrix $A \in M_N(\mathbb{R})$ is invertible precisely when $\det^+(A) > 0$.

Proof. This follows from Theorem 2.2, and from the general results from chapter 1, which tell us that a matrix $A \in M_N(\mathbb{R})$ is invertible precisely when its column vectors $v_1, \ldots, v_N \in \mathbb{R}^N$ are linearly independent. But this latter condition is equivalent to the fact that we must have the following strict inequality:

$$vol < v_1, \ldots, v_N >> 0$$

Thus, we are led to the conclusion in the statement. ∎

Summarizing, all this leads us into the explicit computation of $\det^+$. As a first observation, in 1 dimension we obtain the absolute value of the real numbers:

$$\det{}^+(a) = |a|$$

In 2 dimensions now, the computation is non-trivial, and we have the following result, making the link with our main result so far, namely Theorem 2.3 above:

THEOREM 2.6. *In 2 dimensions we have the following formula,*

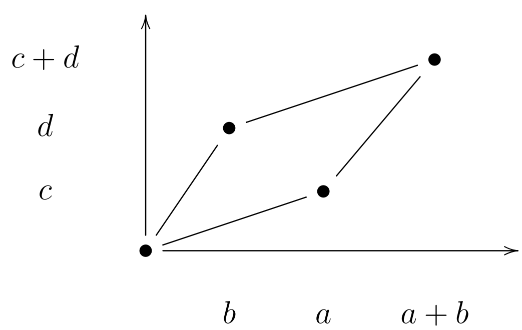
$$\det{}^+\begin{pmatrix} a & b \\ c & d \end{pmatrix} = |ad - bc|$$

with $\det^+ : M_2(\mathbb{R}) \to \mathbb{R}_+$ *being the function constructed above.*

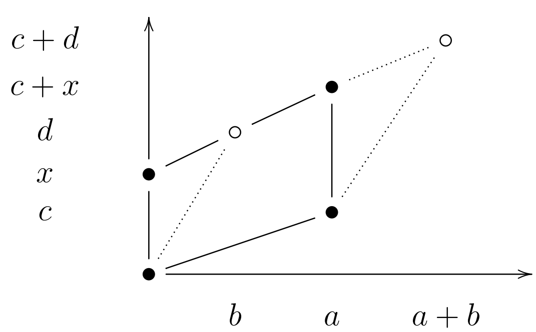
PROOF. We must show that the area of the parallelogram formed by $\binom{a}{c}, \binom{b}{d}$ equals $|ad-bc|$. We can assume $a,b,c,d > 0$ for simplifying, the proof in general being similar. Moreover, by switching if needed the vectors $\binom{a}{c}, \binom{b}{d}$, we can assume that we have:

$$\frac{a}{c} > \frac{b}{d}$$

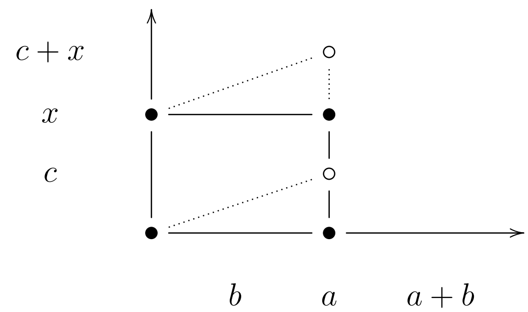
According to these conventions, the picture of our parallelogram is as follows:



Now let us slide the upper side downwards left, until we reach the Oy axis. Our parallelogram, which has not changed its area in this process, becomes:



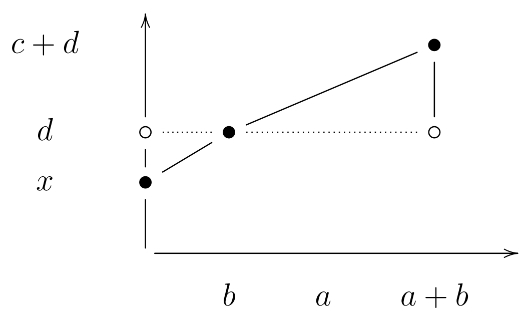
We can further modify this parallelogram, once again by not altering its area, by sliding the right side downwards, until we reach the Ox axis:



Let us compute now the area. Since our two sliding operations have not changed the area of the original parallelogram, this area is given by:

$$A = ax$$

In order to compute the quantity x, observe that in the context of the first move, we have two similar triangles, according to the following picture:



Thus, we are led to the following equation for the number x:

$$\frac{d-x}{b} = \frac{c}{a}$$

By solving this equation, we obtain the following value for x:

$$x = d - \frac{bc}{a}$$

Thus the area of our parallelogram, or rather of the final rectangle obtained from it, which has the same area as the original parallelogram, is given by:

$$ax = ad - bc$$

Thus, we are led to the conclusion in the statement. □

All this is very nice, and we obviously have a beginning of theory here. However, when looking carefully, we can see that our theory has a weakness, because:

(1) In 1 dimension the number a, which is the simplest function of a itself, is certainly a better quantity than the number $|a|$.

(2) In 2 dimensions the number $ad-bc$, which is linear in a,b,c,d, is certainly a better quantity than the number $|ad-bc|$.

So, let us upgrade now our theory, by constructing a better function, which does the same job, namely checking if the vectors are proportional, of the following type:

$$\det : M_N(\mathbb{R}) \to \mathbb{R}$$

$$\det = \pm \det{}^+$$

That is, we would like to have a clever, signed version of $\det^+$, satisfying:

$$\det(a) = a$$

$$\det\begin{pmatrix}a & b\\ c & d\end{pmatrix} = ad-bc$$

In order to do this, we must come up with a way of splitting the systems of vectors $v_1,\ldots,v_N \in \mathbb{R}^N$ into two classes, call them positive and negative. And here, the answer is quite clear, because a bit of thinking leads to the following definition:

Definition 2.7. *A system of vectors $v_1,\ldots,v_N \in \mathbb{R}^N$ is called:*

(1) *Oriented, if one can continuously pass from the standard basis to it.*

(2) *Unoriented, otherwise.*

The associated sign is $+$ in the oriented case, and $-$ in the unoriented case.

As a first example, in 1 dimension the basis consists of the single vector $e=1$, which can be continuously deformed into any vector $a>0$. Thus, the sign is the usual one:

$$sgn(a) = \begin{cases}+ & \text{if } a>0\\ - & \text{if } a<0\end{cases}$$

Thus, in connection with our original question, we are definitely on the good track, because when multiplying $|a|$ by this sign we obtain a itself, as desired:

$$a = sgn(a)|a|$$

In 2 dimensions now, the explicit formula of the sign is as follows:

Proposition 2.8. *We have the following formula, valid for any 2 vectors in $\mathbb{R}^2$,*

$$sgn\left[\begin{pmatrix}a\\ c\end{pmatrix},\begin{pmatrix}b\\ d\end{pmatrix}\right] = sgn(ad-bc)$$

with the sign function on the right being the usual one, in 1 dimension.

PROOF. According to our conventions, the sign of $\binom{a}{c},\binom{b}{d}$ is as follows:

(1) The sign is + when these vectors come in this order with respect to the counter-clockwise rotation in the plane, around 0.

(2) The sign is − otherwise, meaning when these vectors come in this order with respect to the clockwise rotation in the plane, around 0.

If we assume now $a,b,c,d>0$ for simplifying, we are left with comparing the angles having the numbers c/a and d/b as tangents, and we obtain in this way:

$$sgn\left[\begin{pmatrix}a\\c\end{pmatrix},\begin{pmatrix}b\\d\end{pmatrix}\right]=\begin{cases}+&\text{if }\frac{c}{a}<\frac{d}{b}\\-&\text{if }\frac{c}{a}>\frac{d}{b}\end{cases}$$

But this gives the formula in the statement. The proof in general is similar. □

Once again, in connection with our original question, we are on the good track, because when multiplying $|ad-bc|$ by this sign we obtain $ad-bc$ itself, as desired:

$$ad-bc=sgn(ad-bc)|ad-bc|$$

Let us look as well into the case $N=3$. Things here are quite complicated, and we will discuss this later on. However, we have the following basic result:

PROPOSITION 2.9. *Consider the standard basis of $\mathbb{R}^3$, namely:*

$$e_1=\begin{pmatrix}1\\0\\0\end{pmatrix}\quad,\quad e_2=\begin{pmatrix}0\\1\\0\end{pmatrix}\quad,\quad e_3=\begin{pmatrix}0\\0\\1\end{pmatrix}$$

We have then the following sign computations:

(1) $sgn(e_1,e_2,e_3)=+$.
(2) $sgn(e_1,e_3,e_2)=-$.
(3) $sgn(e_2,e_1,e_3)=-$.
(4) $sgn(e_2,e_3,e_1)=+$.
(5) $sgn(e_3,e_1,e_2)=+$.
(6) $sgn(e_3,e_2,e_1)=-$.

PROOF. In each case the problem is whether one can continuously pass from (e_1,e_2,e_3) to the basis in statement, and the computations can be done as follows:

(1) In three of the cases under investigation, namely (2,3,6), one of the vectors is unchanged, and the other two are switched. Thus, we are more or less in 2 dimensions, and since the switch here clearly corresponds to −, the sign in these cases is −.

(2) As for the remaining three cases, namely (1,4,5), here the sign can only be +, since things must be 50-50 between + and −, say by symmetry reasons. And this is indeed the case, because what we have here are rotations of the standard basis. □

As already mentioned, we will be back to this later, with a general formula for the sign in 3 dimensions. This formula is quite complicated, the idea being that of making out of the $3 \times 3 = 9$ entries of our vectors a certain quantity, somewhat in the spirit of the one in Proposition 2.8, and then taking the sign of this quantity.

At the level of the general results now, we have:

PROPOSITION 2.10. *The orientation of a system of vectors changes as follows:*
(1) *If we switch the sign of a vector, the associated sign switches.*
(2) *If we permute two vectors, the associated sign switches as well.*

PROOF. Both these assertions are clear from the definition of the sign, because the two operations in question change the orientation of the system of vectors. $\square$

With the above notion in hand, we can now formulate:

DEFINITION 2.11. *The determinant of $v_1, \ldots, v_N \in \mathbb{R}^N$ is the signed volume*

$$\det(v_1 \ldots v_N) = \pm vol < v_1, \ldots, v_N >$$

of the parallelepiped made by these vectors.

In other words, we are upgrading here Definition 2.4, by adding a sign to the quantity $\det^+$ constructed there, as to potentially reach to good additivity properties:

$$\det(v_1 \ldots v_N) = \pm \det^+(v_1 \ldots v_N)$$

In relation with our original inversion problem for the square matrices, this upgrade does not change what we have so far, and we have the following statement:

THEOREM 2.12. *The quantity* det *that we constructed, regarded as a function of the corresponding square matrices, formed by column vectors,*

$$\det : M_N(\mathbb{R}) \to \mathbb{R}$$

has the property that a matrix $A \in M_N(\mathbb{R})$ is invertible precisely when $\det(A) \neq 0$.

PROOF. We know from Proposition 2.5 that a matrix $A \in M_N(\mathbb{R})$ is invertible precisely when $\det^+(A) = |\det A|$ is strictly positive, and this gives the result. $\square$

In the matrix context, we will often use the symbol $|\,.\,|$ instead of det:

$$|A| = \det A$$

Let us try now to compute the determinant. In 1 dimension we have of course the formula $\det(a) = a$, because the absolute value fits, and so does the sign:

$$\begin{aligned} \det(a) &= sgn(a) \times |a| \\ &= a \end{aligned}$$

In 2 dimensions now, we have the following result:

THEOREM 2.13. *In 2 dimensions we have the following formula,*

$$\begin{vmatrix}a&b\\c&d\end{vmatrix}=ad-bc$$

with $|\,.\,|=\det$ *being the determinant function constructed above.*

PROOF. According to our definition, to the computation in Theorem 2.6, and to sign formula from Proposition 2.8, the determinant of a 2×2 matrix is given by:

$$\begin{aligned}\det\begin{pmatrix}a&b\\c&d\end{pmatrix}&=sgn\left[\begin{pmatrix}a\\c\end{pmatrix},\begin{pmatrix}b\\d\end{pmatrix}\right]\times\det{}^+\begin{pmatrix}a&b\\c&d\end{pmatrix}\\&=sgn\left[\begin{pmatrix}a\\c\end{pmatrix},\begin{pmatrix}b\\d\end{pmatrix}\right]\times|ad-bc|\\&=sgn(ad-bc)\times|ad-bc|\\&=ad-bc\end{aligned}$$

Thus, we have obtained the formula in the statement. □

2c. Basic properties

In order to discuss now arbitrary dimensions, we will need a number of theoretical results. Here is a first series of formulae, coming straight from the definitions:

THEOREM 2.14. *The determinant has the following properties:*

(1) *When multiplying by scalars, the determinant gets multiplied as well:*

$$\det(\lambda_1v_1,\ldots,\lambda_Nv_N)=\lambda_1\ldots\lambda_N\det(v_1,\ldots,v_N)$$

(2) *When permuting two columns, the determinant changes the sign:*

$$\det(\ldots,u,\ldots,v,\ldots)=-\det(\ldots,v,\ldots,u,\ldots)$$

(3) *The determinant* $\det(e_1,\ldots,e_N)$ *of the standard basis of* $\mathbb{R}^N$ *is* 1.

PROOF. All this is clear from definitions, as follows:

(1) This follows from definitions, and from Proposition 2.10 (1).

(2) This follows as well from definitions, and from Proposition 2.10 (2).

(3) This is clear from our definition of the determinant. □

As an application of the above result, we have:

THEOREM 2.15. *The determinant of a diagonal matrix is given by:*

$$\begin{vmatrix}\lambda_1&&\\&\ddots&\\&&\lambda_N\end{vmatrix}=\lambda_1\ldots\lambda_N$$

That is, we obtain the product of diagonal entries, or of eigenvalues.

Proof. The formula in the statement is clear by using the rules (1) and (3) in Theorem 2.14 above, which in matrix terms give:

$$\begin{aligned}\begin{vmatrix}\lambda_1&&\\&\ddots&\\&&\lambda_N\end{vmatrix}&=\lambda_1\ldots\lambda_N\begin{vmatrix}1&&\\&\ddots&\\&&1\end{vmatrix}\\&=\lambda_1\ldots\lambda_N\end{aligned}$$

As for the last assertion, this is rather a remark. □

We will see in a moment that, more generally, the determinant of any diagonalizable matrix is the product of its eigenvalues.

In order to reach to a more advanced theory, let us adopt now the linear map point of view. In this setting, the definition of the determinant reformulates as follows:

Theorem 2.16. *Given a linear map, written as $f(v) = Av$, its "inflation coefficient", obtained as the signed volume of the image of the unit cube, is given by:*

$$I_f = \det A$$

More generally, I_f is the inflation ratio of any parallelepiped in $\mathbb{R}^N$, via the transformation f. In particular f is invertible precisely when $\det A \neq 0$.

Proof. The only non-trivial thing in all this is the fact that the inflation coefficient I_f, as defined above, is independent of the choice of the parallelepiped. But this is a generalization of the Thales theorem, which follows from the Thales theorem itself. □

As a first application of the above linear map viewpoint, we have:

Theorem 2.17. *We have the following formula, valid for any matrices A, B:*

$$\det(AB) = \det A \cdot \det B$$

In particular, we have $\det(AB) = \det(BA)$.

Proof. The decomposition formula in the statement follows by using the associated linear maps, which multiply as follows:

$$f_{AB} = f_A f_B$$

Indeed, when computing the determinant, by using the "inflation coefficient" viewpoint from Theorem 2.16, we obtain the same thing on both sides.

As for the formula $\det(AB) = \det(BA)$, this is clear from the first formula. □

Getting back now to explicit computations, we have the following key result:

THEOREM 2.18. *The determinant of a diagonalizable matrix*

$$A \sim \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_N \end{pmatrix}$$

is the product of its eigenvalues, $\det A = \lambda_1 \ldots \lambda_N$.

PROOF. We know that a diagonalizable matrix can be written in the form $A = PDP^{-1}$, with $D = diag(\lambda_1, \ldots, \lambda_N)$. Now by using Theorem 2.17, we obtain:

$$\begin{aligned} \det A &= \det(PDP^{-1}) \\ &= \det(DP^{-1}P) \\ &= \det D \\ &= \lambda_1 \ldots \lambda_N \end{aligned}$$

Thus, we are led to the formula in the statement. □

Here is another important result, which is very useful for diagonalization:

THEOREM 2.19. *The eigenvalues of a matrix* $A \in M_N(\mathbb{R})$ *are the roots of*

$$P(x) = \det(A - x1_N)$$

called characteristic polynomial of the matrix.

PROOF. We have the following computation, using the fact that a linear map is bijective precisely when the determinant of the associated matrix is nonzero:

$$\begin{aligned} \exists v, Av = \lambda v &\iff \exists v, (A - \lambda 1_N)v = 0 \\ &\iff \det(A - \lambda 1_N) = 0 \end{aligned}$$

Thus, we are led to the conclusion in the statement. □

Here are now some other computations, once again in arbitrary dimensions:

PROPOSITION 2.20. *We have the following results:*

(1) *The determinant of an orthogonal matrix must be* ± 1.
(2) *The determinant of a projection must be* 0 *or* 1.

PROOF. These are elementary results, the idea being as follows:

(1) Here the determinant must be indeed ± 1, because the orthogonal matrices map the unit cube to a copy of the unit cube.

(2) Here the determinant is in general 0, because the projections flatten the unit cube, unless we have the identity, where the determinant is 1. □

In general now, at the theoretical level, we have the following key result:

Theorem 2.21. *The determinant has the additivity property*

$$\det(\ldots,u+v,\ldots)=\det(\ldots,u,\ldots)+\det(\ldots,v,\ldots)$$

valid for any choice of the vectors involved.

Proof. This follows by doing some elementary geometry, in the spirit of the computations in the proof of Theorem 2.6 above, as follows:

(1) We can either use the Thales theorem, and then compute the volumes of all the parallelepipeds involved, by using basic algebraic formulae.

(2) Or we can solve the problem in "puzzle" style, the idea being to cut the big parallelepiped, and then recover the small ones, after some manipulations.

(3) We can do as well something hybrid, consisting in deforming the parallelepipeds involved, without changing their volumes, and then cutting and gluing. □

As a basic application of the above result, we have:

Theorem 2.22. *We have the following results:*

1. *The determinant of a diagonal matrix is the product of diagonal entries.*
2. *The same is true for the upper triangular matrices.*
3. *The same is true for the lower triangular matrices.*

Proof. All this can be deduced by using our various general formulae, as follows:

(1) This is something that we already know, from Theorem 2.15 above.

(2) This follows by using Theorem 2.14 and Theorem 2.21, then (1), as follows:

$$\begin{aligned}
\begin{vmatrix}\lambda_1&&&*\\&\lambda_2&&\\&&\ddots&\\0&&&\lambda_N\end{vmatrix}
&=\begin{vmatrix}\lambda_1&0&&*\\&\lambda_2&&\\&&\ddots&\\0&&&\lambda_N\end{vmatrix}\\
&\ \vdots\\
&\ \vdots\\
&=\begin{vmatrix}\lambda_1&&&0\\&\lambda_2&&\\&&\ddots&\\0&&&\lambda_N\end{vmatrix}\\
&=\lambda_1\ldots\lambda_N
\end{aligned}$$

(3) This follows as well from Theorem 2.14 and Theorem 2.21, then (1), by proceeding this time from right to left, from the last column towards the first column. □

We can see from the above that the rules in Theorem 2.14 and Theorem 2.21 are quite powerful, taken altogether. For future reference, let us record these rules:

THEOREM 2.23. *The determinant has the following properties:*

(1) *When adding two columns, the determinants get added:*

$$\det(\ldots,u+v,\ldots)=\det(\ldots,u,\ldots)+\det(\ldots,v,\ldots)$$

(2) *When multiplying columns by scalars, the determinant gets multiplied:*

$$\det(\lambda_1v_1,\ldots,\lambda_Nv_N)=\lambda_1\ldots\lambda_N\det(v_1,\ldots,v_N)$$

(3) *When permuting two columns, the determinant changes the sign:*

$$\det(\ldots,u,\ldots,v,\ldots)=-\det(\ldots,v,\ldots,u,\ldots)$$

(4) *The determinant* $\det(e_1,\ldots,e_N)$ *of the standard basis of* $\mathbb{R}^N$ *is* 1.

PROOF. This is something that we already know, which follows by putting together the various formulae from Theorem 2.14 and Theorem 2.21 above. □

As an important theoretical result now, which will ultimately lead to an algebraic reformulation of the whole determinant problematics, we have:

THEOREM 2.24. *The determinant of square matrices is the unique map*

$$\det:M_N(\mathbb{R})\to\mathbb{R}$$

satisfying the conditions in Theorem 2.23 above.

PROOF. This can be done in two steps, as follows:

(1) Our first claim is that any map $\det':M_N(\mathbb{R})\to\mathbb{R}$ satisfying the conditions in Theorem 2.23 must coincide with det on the upper triangular matrices. But this is clear from the proof of Theorem 2.22, which only uses the rules in Theorem 2.23.

(2) Our second claim is that we have $\det'=\det$, on all matrices. But this can be proved by putting the matrix in upper triangular form, by using operations on the columns, in the spirit of the manipulations from the proof of Theorem 2.22 above. □

Here is now another important theoretical result:

THEOREM 2.25. *The determinant is subject to the row expansion formula*

$$\begin{vmatrix}a_{11}&\ldots&a_{1N}\\ \vdots&&\vdots\\ a_{N1}&\ldots&a_{NN}\end{vmatrix}=a_{11}\begin{vmatrix}a_{22}&\ldots&a_{2N}\\ \vdots&&\vdots\\ a_{N2}&\ldots&a_{NN}\end{vmatrix}-a_{12}\begin{vmatrix}a_{21}&a_{23}&\ldots&a_{2N}\\ \vdots&\vdots&&\vdots\\ a_{N1}&a_{N3}&\ldots&a_{NN}\end{vmatrix}$$

$$+\ldots\ldots+(-1)^{N+1}a_{1N}\begin{vmatrix}a_{21}&\ldots&a_{2,N-1}\\ \vdots&&\vdots\\ a_{N1}&\ldots&a_{N,N-1}\end{vmatrix}$$

and this method fully computes it, by recurrence.

PROOF. This follows from the fact that the formula in the statement produces a certain function $\det : M_N(\mathbb{R}) \to \mathbb{R}$, which has the 4 properties in Theorem 2.23. □

We can expand as well over the columns, as follows:

THEOREM 2.26. *The determinant is subject to the column expansion formula*

$$\begin{vmatrix}a_{11}&\ldots&a_{1N}\\ \vdots&&\vdots\\ a_{N1}&\ldots&a_{NN}\end{vmatrix} = a_{11}\begin{vmatrix}a_{22}&\ldots&a_{2N}\\ \vdots&&\vdots\\ a_{N2}&\ldots&a_{NN}\end{vmatrix} - a_{21}\begin{vmatrix}a_{12}&\ldots&a_{1N}\\ a_{32}&\ldots&a_{3N}\\ \vdots&&\vdots\\ a_{N2}&\ldots&a_{NN}\end{vmatrix}$$
$$+\ldots\ldots+(-1)^{N+1}a_{N1}\begin{vmatrix}a_{12}&\ldots&a_{1N}\\ \vdots&&\vdots\\ a_{N-1,2}&\ldots&a_{N-1,N}\end{vmatrix}$$

and this method fully computes it, by recurrence.

PROOF. This follows by using the same argument as for the rows. □

We can now complement Theorem 2.23 with a similar result for the rows:

THEOREM 2.27. *The determinant has the following properties:*

(1) *When adding two rows, the determinants get added:*

$$\det\begin{pmatrix}\vdots\\ u+v\\ \vdots\end{pmatrix} = \det\begin{pmatrix}\vdots\\ u\\ \vdots\end{pmatrix} + \det\begin{pmatrix}\vdots\\ v\\ \vdots\end{pmatrix}$$

(2) *When multiplying row by scalars, the determinant gets multiplied:*

$$\det\begin{pmatrix}\lambda_1 v_1\\ \vdots\\ \lambda_N v_N\end{pmatrix} = \lambda_1\ldots\lambda_N \det\begin{pmatrix}v_1\\ \vdots\\ v_N\end{pmatrix}$$

(3) *When permuting two rows, the determinant changes the sign.*

PROOF. This follows indeed by using the using various formulae established above, and is best seen by using the column expansion formula from Theorem 2.26. □

We can see from the above that the determinant is the subject to many interesting formulae, and that some of these formulae, when taken altogether, uniquely determine it. In all this, what is the most luminous is certainly the definition of the determinant as a volume. As for the second most luminous of our statements, this is Theorem 2.24 above, which is something a bit abstract, but both beautiful and useful. So, as a final theoretical statement now, here is an alternative reformulation of Theorem 2.24:

Theorem 2.28. *The determinant of the systems of vectors*

$$\det:\mathbb{R}^N\times\ldots\times\mathbb{R}^N\to\mathbb{R}$$

is multilinear, alternate and unital, and unique with these properties.

Proof. This is a fancy reformulation of Theorem 2.24, with the various properties of det from the statement being those from Theorem 2.23. □

As a conclusion to all this, we have now a full theory for the determinant, and we can freely use all the above results, definitions and theorems alike, and even start forgetting what is actually definition, and what is theorem.

2d. Sarrus and beyond

As a first application of the above methods, we can now prove:

Theorem 2.29. *The determinant of the* 3×3 *matrices is given by*

$$\begin{vmatrix}a&b&c\\d&e&f\\g&h&i\end{vmatrix}=aei+bfg+cdh-ceg-bdi-afh$$

which can be memorized by using Sarrus' triangle method,

$$\begin{aligned}\det\ &=\ \begin{pmatrix}*&&\\&*&\\&&*\end{pmatrix}+\begin{pmatrix}&*&\\&&*\\*&&\end{pmatrix}+\begin{pmatrix}&&*\\*&&\\&*&\end{pmatrix}\\&-\ \begin{pmatrix}&&*\\&*&\\*&&\end{pmatrix}+\begin{pmatrix}&*&\\*&&\\&&*\end{pmatrix}+\begin{pmatrix}*&&\\&&*\\&*&\end{pmatrix}\end{aligned}$$

"triangles parallel to the diagonal, minus triangles parallel to the antidiagonal".

Proof. Here is the computation, using Theorem 2.25 above:

$$\begin{aligned}\begin{vmatrix}a&b&c\\d&e&f\\g&h&i\end{vmatrix}&=a\begin{vmatrix}e&f\\h&i\end{vmatrix}-b\begin{vmatrix}d&f\\g&i\end{vmatrix}+c\begin{vmatrix}d&e\\g&h\end{vmatrix}\\&=a(ei-fh)-b(di-fg)+c(dh-eg)\\&=aei-afh-bdi+bfg+cdh-ceg\\&=aei+bfg+cdh-ceg-bdi-afh\end{aligned}$$

Thus, we obtain the formula in the statement. □

As a first application, let us go back to the inversion problem for the 3×3 matrices, that we left open in the above. We can now solve this problem, as follows:

Theorem 2.30. *The inverses of the 3×3 matrices are given by*

$$\begin{pmatrix}a&b&c\\ d&e&f\\ g&h&i\end{pmatrix}^{-1}=\frac{1}{D}\begin{pmatrix}ei-fh&ch-bi&bf-ce\\ fg-di&ai-cg&cd-af\\ dh-eg&bg-ah&ae-bd\end{pmatrix}$$

with D being the determimant. When $D=0$, the matrix is not invertible.

Proof. We can use here the same method as for the 2×2 matrices. To be more precise, in order for the matrix to be invertible, we must have:

$$D \neq 0$$

The trick now is to look for solutions of the following problem:

$$\begin{pmatrix}a&b&c\\ d&e&f\\ g&h&i\end{pmatrix}\begin{pmatrix}*&*&*\\ *&*&*\\ *&*&*\end{pmatrix}=\begin{pmatrix}D&0&0\\ 0&D&0\\ 0&0&D\end{pmatrix}$$

We know from Theorem 2.29 above that the determinant is given by:

$$D=aei+bfg+cdh-ceg-bdi-afh$$

But this leads, via some obvious choices, to the following solution:

$$\begin{pmatrix}*&*&*\\ *&*&*\\ *&*&*\end{pmatrix}=\begin{pmatrix}ei-fh&ch-bi&bf-ce\\ fg-di&ai-cg&cd-af\\ dh-eg&bg-ah&ae-bd\end{pmatrix}$$

Thus, by rescaling, we obtain the formula in the statement. □

In fact, we can now fully solve the inversion problem, as follows:

Theorem 2.31. *The inverse of a square matrix, having nonzero determinant,*

$$A=\begin{pmatrix}a_{11}&\ldots&a_{1N}\\ \vdots&&\vdots\\ a_{N1}&\ldots&a_{NN}\end{pmatrix}$$

is given by the following formula,

$$A^{-1}=\frac{1}{\det A}\begin{pmatrix}\det A^{(11)}&-\det A^{(21)}&\det A^{(31)}&\ldots\\ -\det A^{(12)}&\det A^{(22)}&-\det A^{(32)}&\ldots\\ \det A^{(13)}&-\det A^{(23)}&\det A^{(33)}&\ldots\\ \vdots&\vdots&\vdots\end{pmatrix}$$

where $A^{(ij)}$ is the matrix A, with the i-th row and j-th column removed.

Proof. This follows indeed by using the row expansion formula from Theorem 2.25, which in terms of the matrix A^{-1} in the statement reads $AA^{-1}=1$. □

In practice, the above result leads to the following algorithm, which is quite easy to memorize, for computing the inverse:

(1) Delete rows and columns, and compute the corresponding determinants.

(2) Transpose, and add checkered signs.

(3) Divide by the determinant.

Observe that this generalizes our previous computations at $N = 2, 3$. As an illustration, consider an arbitrary 2×2 matrix, written as follows:

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

By deleting rows and columns we obtain 1×1 matrices, and so the matrix formed by the determinants $\det(A^{(ij)})$ is as follows:

$$M = \begin{pmatrix} d & c \\ b & a \end{pmatrix}$$

Now by transposing, adding checkered signs and dividing by $\det A$, we obtain:

$$A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

Similarly, at $N = 3$ what we obtain is the Sarrus formula, from Theorem 2.29.

As a new application now, let us record the following result, at $N = 4$:

THEOREM 2.32. *The determinant of the 4×4 matrices is given by*

$$\begin{aligned} & \begin{vmatrix} a_1 & a_2 & a_3 & a_4 \\ b_1 & b_2 & b_3 & b_4 \\ c_1 & c_2 & c_3 & c_4 \\ d_1 & d_2 & d_3 & d_4 \end{vmatrix} \\ = \;& a_1b_2c_3d_4 - a_1b_2c_4d_3 - a_1b_3c_2d_4 + a_1b_3c_4d_2 + a_1b_4c_2d_3 - a_1b_4c_3d_2 \\ - \;& a_2b_1c_3d_4 + a_2b_1c_4d_3 + a_2b_3c_1d_4 - a_2b_3c_4d_1 - a_2b_4c_1d_3 + a_2b_4c_3d_1 \\ + \;& a_3b_1c_2d_4 + a_3b_1c_4d_2 - a_3b_2c_1d_4 + a_3b_2c_4d_1 + a_3b_4c_1d_2 - a_3b_4c_2d_1 \\ - \;& a_4b_1c_2d_3 + a_4b_1c_3d_2 - a_4b_2c_1d_3 - a_4b_2c_3d_1 - a_4b_3c_1d_2 + a_4b_3c_2d_1 \end{aligned}$$

and the formula of the inverse is as follows, involving 16 Sarrus determinants,

$$A^{-1} = \frac{1}{\det A} \begin{pmatrix} \det A^{(11)} & -\det A^{(21)} & \det A^{(31)} & -\det A^{(41)} \\ -\det A^{(12)} & \det A^{(22)} & -\det A^{(32)} & \det A^{(42)} \\ \det A^{(13)} & -\det A^{(23)} & \det A^{(33)} & -\det A^{(43)} \\ -\det A^{(14)} & \det A^{(24)} & -\det A^{(34)} & \det A^{(44)} \end{pmatrix}$$

where $A^{(ij)}$ is the matrix A, with the i-th row and j-th column removed.

PROOF. The formula for the determinant follows by developing over the first row, then by using the Sarrus formula, for each of the 4 smaller determinants which appear:

$$\begin{vmatrix}a_1&a_2&a_3&a_4\\b_1&b_2&b_3&b_4\\c_1&c_2&c_3&c_4\\d_1&d_2&d_3&d_4\end{vmatrix} = a_1\begin{vmatrix}b_2&b_3&b_4\\c_2&c_3&c_4\\d_2&d_3&d_4\end{vmatrix} - a_2\begin{vmatrix}b_1&b_3&b_4\\c_1&c_3&c_4\\d_1&d_3&d_4\end{vmatrix}$$

$$+ a_3\begin{vmatrix}b_1&b_2&b_4\\c_1&c_2&c_4\\d_1&d_2&d_4\end{vmatrix} - a_4\begin{vmatrix}b_1&b_2&b_3\\c_1&c_2&c_3\\d_1&d_2&d_3\end{vmatrix}$$

As for the formula of the inverse, this is something that we already know. $\square$

Let us discuss now the general formula of the determinant, at arbitrary values $N \in \mathbb{N}$ of the matrix size, generalizing those that we have at $N = 2, 3, 4$. We will need:

DEFINITION 2.33. *A permutation of $\{1, \ldots, N\}$ is a bijection, as follows:*

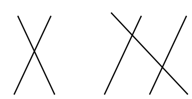
$$\sigma : \{1, \ldots, N\} \to \{1, \ldots, N\}$$

The set of such permutations is denoted S_N.

There are many possible notations for the permutations, the simplest one consisting in writing the numbers $1, \ldots, N$, and below them, their permuted versions:

$$\sigma = \begin{pmatrix}1&2&3&4&5\\2&1&4&5&3\end{pmatrix}$$

Another method, which is better, is by denoting the permutations as diagrams, going from top to bottom:

$\sigma =$ 

Here are some basic properties of the permutations:

THEOREM 2.34. *The permutations have the following properties:*

(1) *There are $N!$ of them.*
(2) *They are stable by composition, and inversion.*

PROOF. In order to construct a permutation $\sigma \in S_N$, we have:

– N choices for the value of $\sigma(N)$.
– $(N-1)$ choices for the value of $\sigma(N-1)$.
– $(N-2)$ choices for the value of $\sigma(N-2)$.

$\vdots$

– and so on, up to 1 choice for the value of $\sigma(1)$.

Thus, we have $N!$ choices, as claimed. As for the second assertion, this is clear. $\square$

We will need the following key result:

THEOREM 2.35. *The permutations have a signature function*

$$\varepsilon : S_N \to \{\pm 1\}$$

which can be defined in the following equivalent ways:

1. *As $(-1)^c$, where c is the number of inversions.*
2. *As $(-1)^t$, where t is the number of transpositions.*
3. *As $(-1)^o$, where o is the number of odd cycles.*
4. *As $(-1)^x$, where x is the number of crossings.*
5. *As the sign of the corresponding permuted basis of $\mathbb{R}^N$.*

PROOF. This is something important, and quite subtle, to be systematically used in what follows. As a first observation, we can see right away a relation with the determinant, coming from (5) above. Thus, we already have some knowledge here, for instance coming from Proposition 2.9, which computes the signature of the permutations $\sigma \in S_3$.

In practice now, we have explain what the numbers c, t, o, x appearing in (1-4) above exactly are, then why they are well-defined modulo 2, then why they are equal to each other, and finally why the constructions (1-4) yield the same sign as (5).

Let us begin with the first two steps, namely precise definition of c, t, o, x, and fact that these numbers are well-defined modulo 2:

(1) The idea here is that given any two numbers $i < j$ among $1, \ldots, N$, the permutation can either keep them in the same order, $\sigma(i) < \sigma(j)$, or invert them:

$$\sigma(j) > \sigma(i)$$

Now by making $i < j$ vary over all pairs of numbers in $1, \ldots, N$, we can count the number of inversions, and call it c. This is an integer, $c \in \mathbb{N}$, which is well-defined.

(2) Here the idea, which is something quite intuitive, is that any permutation appears as a product of switches, also called transpositions:

$$i \leftrightarrow j$$

The decomposition as a product of transpositions is not unique, but the number t of the needed transpositions is unique, when considered modulo 2. This follows for instance from the equivalence of (2) with (1,3,4,5), explained below.

(3) Here the point is that any permutation decomposes, in a unique way, as a product of cycles, which are by definition permutations of the following type:

$$i_1 \to i_2 \to i_3 \to \ldots\ldots \to i_k \to i_1$$

Some of these cycles have even length, and some others have odd length. By counting those having odd length, we obtain a well-defined number $o \in \mathbb{N}$.

(4) Here the method is that of drawing the permutation, as we usually do, and by avoiding triple crossings, and then counting the number of crossings. This number x depends on the way we draw the permutations, but modulo 2, we always get the same number. Indeed, this follows from the fact that we can continuously pass from a drawing to each other, and that when doing so, the number of crossings can only jump by ± 2.

Summarizing, we have 4 different definitions for the signature of the permutations, which all make sense, constructed according to (1-4) above. Regarding now the fact that we always obtain the same number, this can be established as follows:

(1)=(2) This is clear, because any transposition inverts once, modulo 2.

(1)=(3) This is clear as well, because the odd cycles invert once, modulo 2.

(1)=(4) This comes from the fact that the crossings correspond to inversions.

(2)=(3) This follows by decomposing the cycles into transpositions.

(2)=(4) This comes from the fact that the crossings correspond to transpositions.

(3)=(4) This follows by drawing a product of cycles, and counting the crossings.

Finally, in what regards the equivalence of all these constructions with (5), here simplest is to use (2). Indeed, we already know that the sign of a system of vectors switches when interchanging two vectors, and so the equivalence between (2,5) is clear. □

We can now formulate a key result, as follows:

THEOREM 2.36. *We have the following formula for the determinant,*

$$\det A = \sum_{\sigma \in S_N} \varepsilon(\sigma) A_{1\sigma(1)} \ldots A_{N\sigma(N)}$$

with the signature function being the one introduced above.

PROOF. This follows by recurrence over $N \in \mathbb{N}$, as follows:

(1) When developing the determinant over the first column, we obtain a signed sum of N determinants of size $(N-1)\times(N-1)$. But each of these determinants can be computed by developing over the first column too, and so on, and we are led to the conclusion that we have a formula as in the statement, with $\varepsilon(\sigma) \in \{-1, 1\}$ being certain coefficients.

(2) But these latter coefficients $\varepsilon(\sigma) \in \{-1, 1\}$ can only be the signatures of the corresponding permutations $\sigma \in S_N$, with this being something that can be viewed again by recurrence, with either of the definitions (1-5) in Theorem 2.35 for the signature. □

The above result is something quite tricky, and in order to get familiar with it, there is nothing better than doing some computations. As a first, basic example, in 2 dimensions we recover the usual formula of the determinant, the details being as follows:

$$\begin{aligned}\begin{vmatrix}a & b\\ c & d\end{vmatrix} &= \varepsilon(|\,|)\cdot ad+\varepsilon(\times)\cdot cb\\ &= 1\cdot ad+(-1)\cdot cb\\ &= ad-bc\end{aligned}$$

In 3 dimensions now, we recover the Sarrus formula:

$$\begin{vmatrix}a & b & c\\ d & e & f\\ g & h & i\end{vmatrix}=aei+bfg+cdh-ceg-bdi-afh$$

Observe that the triangles in the Sarrus formula correspond to the permutations of $\{1,2,3\}$, and their signs correspond to the signatures of these permutations:

$$\begin{aligned}\det &= \begin{pmatrix}* & & \\ & * & \\ & & *\end{pmatrix}+\begin{pmatrix} & * & \\ & & *\\ * & & \end{pmatrix}+\begin{pmatrix} & & *\\ * & & \\ & * & \end{pmatrix}\\ &- \begin{pmatrix} & & *\\ & * & \\ * & & \end{pmatrix}+\begin{pmatrix} & * & \\ * & & \\ & & *\end{pmatrix}+\begin{pmatrix}* & & \\ & & *\\ & * & \end{pmatrix}\end{aligned}$$

Also, in 4 dimensions, we recover the formula that we already know, as follows:

THEOREM 2.37. *The determinant of the 4×4 matrices is given by*

$$\begin{aligned}&\begin{vmatrix}a_1 & a_2 & a_3 & a_4\\ b_1 & b_2 & b_3 & b_4\\ c_1 & c_2 & c_3 & c_4\\ d_1 & d_2 & d_3 & d_4\end{vmatrix}\\ =\ & a_1b_2c_3d_4-a_1b_2c_4d_3-a_1b_3c_2d_4+a_1b_3c_4d_2+a_1b_4c_2d_3-a_1b_4c_3d_2\\ -\ & a_2b_1c_3d_4+a_2b_1c_4d_3+a_2b_3c_1d_4-a_2b_3c_4d_1-a_2b_4c_1d_3+a_2b_4c_3d_1\\ +\ & a_3b_1c_2d_4+a_3b_1c_4d_2-a_3b_2c_1d_4+a_3b_2c_4d_1+a_3b_4c_1d_2-a_3b_4c_2d_1\\ -\ & a_4b_1c_2d_3+a_4b_1c_3d_2-a_4b_2c_1d_3-a_4b_2c_3d_1-a_4b_3c_1d_2+a_4b_3c_2d_1\end{aligned}$$

with the generic term being of the following form, with $\sigma\in S_4$,

$$\pm a_{\sigma(1)}b_{\sigma(2)}c_{\sigma(3)}d_{\sigma(4)}$$

and with the sign being $\varepsilon(\sigma)$, computable by using Theorem 2.35.

PROOF. We can indeed recover this formula as well as a particular case of Theorem 2.36. To be more precise, the permutations in the statement are listed according to the lexicographic order, and the computation of the corresponding signatures is something elementary, by using the various rules from Theorem 2.35. $\square$

As another application, we have the following key result:

THEOREM 2.38. *We have the formula*

$$\det A = \det A^t$$

valid for any square matrix A.

PROOF. This follows from the formula in Theorem 2.36. Indeed, we have:

$$\begin{aligned}
\det A^t &= \sum_{\sigma\in S_N}\varepsilon(\sigma)(A^t)_{1\sigma(1)}\ldots(A^t)_{N\sigma(N)}\\
&= \sum_{\sigma\in S_N}\varepsilon(\sigma)A_{\sigma(1)1}\ldots A_{\sigma(N)N}\\
&= \sum_{\sigma\in S_N}\varepsilon(\sigma)A_{1\sigma^{-1}(1)}\ldots A_{N\sigma^{-1}(N)}\\
&= \sum_{\sigma\in S_N}\varepsilon(\sigma^{-1})A_{1\sigma^{-1}(1)}\ldots A_{N\sigma^{-1}(N)}\\
&= \sum_{\sigma\in S_N}\varepsilon(\sigma)A_{1\sigma(1)}\ldots A_{N\sigma(N)}\\
&= \det A
\end{aligned}$$

Thus, we are led to the formula in the statement. □

Good news, this is the end of the general theory that we wanted to develop. We have now in our bag all the needed techniques for computing the determinant.

Here is however a nice and important example of a determinant, whose computation uses some interesting new techniques, going beyond what has been said above:

THEOREM 2.39. *We have the Vandermonde determinant formula*

$$\begin{vmatrix}
1 & 1 & 1 & \ldots\ldots & 1\\
x_1 & x_2 & x_3 & \ldots\ldots & x_N\\
x_1^2 & x_2^2 & x_3^2 & \ldots\ldots & x_N\\
\vdots & \vdots & \vdots & & \vdots\\
\vdots & \vdots & \vdots & & \vdots\\
x_1^{N-1} & x_2^{N-1} & x_3^{N-1} & \ldots\ldots & x_N^{N-1}
\end{vmatrix} = \prod_{i>j}(x_i - x_j)$$

valid for any $x_1,\ldots,x_N\in\mathbb{R}$.

PROOF. By expanding over the columns, we see that the determinant in question, say D, is a polynomial in the variables $x_1,\ldots,x_N$, having degree $N-1$ in each variable.

Now observe that when setting $x_i = x_j$, for some indices $i \neq j$, our matrix will have two identical columns, and so its determinant D will vanish:

$$x_i = x_j \implies D = 0$$

But this gives us the key to the computation of D. Indeed, D must be divisible by $x_i - x_j$ for any $i \neq j$, and so we must have a formula of the following type:

$$D = c\prod_{i>j}(x_i - x_j)$$

Moreover, since the product on the right is, exactly as D itself, a polynomial in the variables $x_1, \ldots, x_N$, having degree $N-1$ in each variable, we conclude that the quantity c must be a constant, not depending on any of the variables $x_1, \ldots, x_N$:

$$c \in \mathbb{R}$$

In order to finish the computation, it remains to find the value of this constant c. But this can be done for instance by recurrence, and we obtain:

$$c = 1$$

Thus, we are led to the formula in the statement. $\square$

Getting back now to generalities, and to what we want to do with our linear algebra theory, now that we are experts in the computation of the determinant, we should investigate the next problem, namely the diagonalization one.

And here, we know from Theorem 2.19 that the eigenvalues of a matrix $A \in M_N(\mathbb{R})$ appear as roots of the characteristic polynomial:

$$P(x) = \det(A - x1_N)$$

Thus, with the determinant theory developed above, we can in principle compute these eigenvalues, and solve the diagonalization problem afterwards.

The problem, however, is that certain real matrices can have characteristic polynomials of type $P(x) = x^2 + 1$, and this suggests that these matrices might be not diagonalizable over $\mathbb{R}$, but be diagonalizable over $\mathbb{C}$ instead. And so, before getting into diagonalization problems, we must upgrade our theory, and talk about complex matrices. We will do this in the next chapter, and afterwards, we will go back to the diagonalization problem.

2e. Exercises

There has been a lot of exciting theory in this chapter, with some details sometimes missing, and our first exercises will be about this. First, we have:

EXERCISE 2.40. *Fill in all the geometric details in the basic theory of the determinant, by using the same type of arguments as those in the proof of*

$$\det\begin{pmatrix}a&b\\c&d\end{pmatrix}=ad-bc$$

which was fully proved in the above, namely geometric manipulations, and Thales.

To be more precise here, passed some issues with the sign and orientation, which are all elementary, the above 2×2 determinant formula was subject of Theorem 2.6, coming with a full and honest proof. The problem is that of using the same arguments, namely basic geometry, as to have a full proof of Theorem 2.16 and Theorem 2.21 as well.

EXERCISE 2.41. *Prove with full details, based on the above, that the determinant of the systems of vectors*

$$\det:\mathbb{R}^N\times\ldots\times\mathbb{R}^N\to\mathbb{R}$$

is multilinear, alternate and unital, and unique with these properties. Then try to prove as well this directly, without any reference to geometry.

To be more precise, in what regards the first question, this is something that we already discussed in the above, with only a few details missing, and the problem is that of recovering these details. As for the second question, this is something more tricky, and there are several possible approaches here, all being interesting and enjoyable.

EXERCISE 2.42. *Work out, with full details, the theory of the signature map*

$$\varepsilon:S_N\to\{\pm1\}$$

as outlined in Theorem 2.35 above and its proof.

As before, these are things that we already discussed, with a few details missing.

EXERCISE 2.43. *Prove that for a matrix $H\in M_N(\pm1)$, we have*

$$|\det H|\leq N^{N/2}$$

and then find the maximizers of $|\det H|$, at small values of N.

Here the first question is theoretical, and its proof should not be difficult. As for the second question, which is quite tricky, the higher the $N\in\mathbb{N}$ you get to, the better.

CHAPTER 3

Complex matrices

3a. Complex numbers

We have seen that the study of the real matrices $A \in M_N(\mathbb{R})$ suggests the use of the complex numbers. Indeed, even simple matrices like the 2×2 ones can, at least in a formal sense, have complex eigenvalues. In what follows we discuss the complex matrices $A \in M_N(\mathbb{C})$. We will see that the theory here is much more complete than in the real case. As an application, we will solve in this way problems left open in the real case.

Let us begin with the complex numbers. There is a lot of magic here, and we will carefully explain this material. Their definition is as follows:

Definition 3.1. *The complex numbers are variables of the form*

$$x = a + ib$$

which add in the obvious way, and multiply according to the following rule:

$$i^2 = -1$$

In other words, we consider variables as above, without bothering for the moment with their precise meaning. Now consider two such complex numbers:

$$x = a + ib \quad , \quad y = c + id$$

The formula for the sum is then the obvious one, as follows:

$$x + y = (a + c) + i(b + d)$$

As for the formula of the product, by using the rule $i^2 = -1$, we obtain:

$$\begin{aligned} xy &= (a + ib)(c + id) \\ &= ac + iad + ibc + i^2bd \\ &= ac + iad + ibc - bd \\ &= (ac - bd) + i(ad + bc) \end{aligned}$$

Thus, the complex numbers as introduced above are well-defined. The multiplication formula is of course quite tricky, and hard to memorize, but we will see later some alternative ways, which are more conceptual, for performing the multiplication.

The advantage of using the complex numbers comes from the fact that the equation $x^2 = 1$ has now a solution, $x = i$. In fact, this equation has two solutions, namely:

$$x = \pm i$$

This is of course very good news. More generally, we have the following result:

THEOREM 3.2. *The complex solutions of $ax^2 + bx + c = 0$ with $a, b, c \in \mathbb{R}$ are*

$$x_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

with the square root of negative real numbers being defined as:

$$\sqrt{-m} = \pm i\sqrt{m}$$

PROOF. We can write our equation in the following way:

$$\begin{aligned} ax^2 + bx + c = 0 &\iff x^2 + \frac{b}{a}x + \frac{c}{a} = 0 \\ &\iff \left(x + \frac{b}{2a}\right)^2 - \frac{b^2}{4a^2} + \frac{c}{a} = 0 \\ &\iff \left(x + \frac{b}{2a}\right)^2 = \frac{b^2 - 4ac}{4a^2} \\ &\iff x + \frac{b}{2a} = \pm\frac{\sqrt{b^2 - 4ac}}{2a} \end{aligned}$$

Thus, we are led to the conclusion in the statement. $\square$

We will be back later to this, with generalizations. Getting back now to Definition 3.1 as it is, we can represent the complex numbers in the plane, as follows:

PROPOSITION 3.3. *The complex numbers, written as usual*

$$x = a + ib$$

can be represented in the plane, according to the following identification:

$$x = \begin{pmatrix} a \\ b \end{pmatrix}$$

With this convention, the sum of complex numbers is the usual sum of vectors.

PROOF. Consider indeed two arbitrary complex numbers:

$$x = a + ib \quad , \quad y = c + id$$

Their sum is then by definition the following complex number:

$$x + y = (a + c) + i(b + d)$$

Now let us represent x, y in the plane, as in the statement:

$$x = \begin{pmatrix} a \\ b \end{pmatrix} \quad , \quad y = \begin{pmatrix} c \\ d \end{pmatrix}$$

In this picture, their sum is given by the following formula:

$$x + y = \begin{pmatrix} a + c \\ b + d \end{pmatrix}$$

But this is indeed the vector corresponding to $x + y$, so we are done. $\square$

Observe that in the above picture, the real numbers correspond to the numbers on the Ox axis. As for the purely imaginary numbers, these lie on the Oy axis, with:

$$i = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

All this is very nice, but in order to understand now the multiplication, we must do something more complicated, namely using polar coordinates. Let us start with:

Definition 3.4. *The complex numbers $x = a + ib$ can be written in polar coordinates,*

$$x = r(\cos t + i \sin t)$$

with the connecting formulae being

$$a = r \cos t \quad , \quad b = r \sin t$$

and in the other sense being

$$r = \sqrt{a^2 + b^2} \quad , \quad \tan t = b/a$$

and with r, t being called modulus, and argument.

There is a clear relation here with the vector notation from Proposition 3.3, because r is the length of the vector, and t is the angle made by the vector with the Ox axis. As a basic example here, the number i takes the following form:

$$i = \cos\left(\frac{\pi}{2}\right) + i \sin\left(\frac{\pi}{2}\right)$$

The point now is that in polar coordinates, the multiplication formula for the complex numbers, which was so far something quite opaque, takes a very simple form:

Theorem 3.5. *Two complex numbers written in polar coordinates,*

$$x = r(\cos s + i \sin s) \quad , \quad y = p(\cos t + i \sin t)$$

multiply according to the following formula:

$$xy = rp(\cos(s + t) + i \sin(s + t))$$

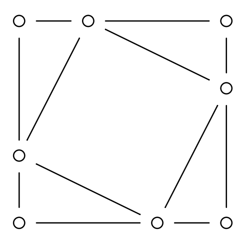
In other words, the moduli multiply, and the arguments sum up.

PROOF. This can be proved by doing some trigonometry, as follows:

(1) Recall first the definition of $\sin, \cos$, as being the sides of a right triangle having angle t. Our first claim is that we have the Pythagoras' theorem, namely:

$$\sin^2 t + \cos^2 t = 1$$

But this comes from the following well-known, and genius picture, with the edges of the outer and inner square being respectively $\sin t + \cos t$ and 1:



Indeed, when computing the area of the outer square, we obtain:

$$(\sin t + \cos t)^2 = 1 + 4 \times \frac{\sin t \cos t}{2}$$

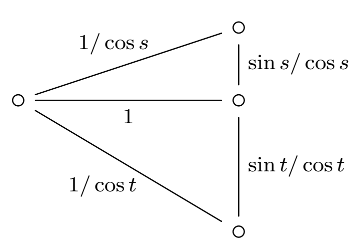
Now when expanding we obtain $\sin^2 t + \cos^2 t = 1$, as claimed.

(2) Next in line, our claim is that we have the following formulae:

$$\sin(s + t) = \cos s \sin t + \sin s \cos t$$

$$\cos(s + t) = \cos s \cos t - \sin s \sin t$$

To be more precise, let us first establish this formula. In order to do so, consider the following picture, consisting of a length 1 line segment, with angles s, t drawn on each side, and with everything being completed, and lengths computed, as indicated:



Now let us compute the area of the big triangle, or rather the double of that area. We can do this in two ways, either directly, with a formula involving $\sin(s + t)$, or by using the two small triangles, involving functions of s, t. We obtain in this way:

$$\frac{1}{\cos s} \cdot \frac{1}{\cos t} \cdot \sin(s + t) = \frac{\sin s}{\cos s} \cdot 1 + \frac{\sin t}{\cos t} \cdot 1$$

But this gives the formula for $\sin(s+t)$ claimed above. Now by using this formula for $\sin(s+t)$ we can deduce as well the formula for $\cos(s+t)$, as follows:

$$\begin{aligned}\cos(s+t) &= \sin\left(\frac{\pi}{2}-s-t\right)\\ &= \sin\left[\left(\frac{\pi}{2}-s\right)+(-t)\right]\\ &= \sin\left(\frac{\pi}{2}-s\right)\cos(-t)+\cos\left(\frac{\pi}{2}-s\right)\sin(-t)\\ &= \cos s\cos t-\sin s\sin t\end{aligned}$$

(3) Now back to complex numbers, we want to prove that $x=r(\cos s+i\sin s)$ and $y=p(\cos t+i\sin t)$ multiply according to the following formula:

$$xy=rp(\cos(s+t)+i\sin(s+t))$$

We can assume that we have $r=p=1$, by dividing everything by these numbers. Now with this assumption made, we have the following computation:

$$\begin{aligned}xy &= (\cos s+i\sin s)(\cos t+i\sin t)\\ &= (\cos s\cos t-\sin s\sin t)+i(\cos s\sin t+\sin s\cos t)\\ &= \cos(s+t)+i\sin(s+t)\end{aligned}$$

Thus, we are led to the conclusion in the statement. □

The above result, which was based on some non-trivial trigonometry, is quite powerful. As a basic application of it, we can now compute powers, as follows:

Theorem 3.6. *The powers of a complex number, written in polar form,*

$$x=r(\cos t+i\sin t)$$

are given by the following formula, valid for any exponent $k\in\mathbb{N}$:

$$x^k=r^k(\cos kt+i\sin kt)$$

Moreover, this formula holds in fact for any $k\in\mathbb{Z}$, and even for any $k\in\mathbb{Q}$.

Proof. Given a complex number x, written in polar form as above, and an exponent $k\in\mathbb{N}$, we have indeed the following computation, with k terms everywhere:

$$\begin{aligned}x^k &= x\ldots x\\ &= r(\cos t+i\sin t)\ldots r(\cos t+i\sin t)\\ &= r\ldots r([\cos(t+\ldots+t)+i\sin(t+\ldots+t))\\ &= r^k(\cos kt+i\sin kt)\end{aligned}$$

Thus, we are done with the case $k\in\mathbb{N}$. Regarding now the generalization to the case $k\in\mathbb{Z}$, it is enough here to do the verification for $k=-1$, where the formula is:

$$x^{-1}=r^{-1}(\cos(-t)+i\sin(-t))$$

But this number x^{-1} is indeed the inverse of x, because:

$$\begin{aligned} xx^{-1} &= r(\cos t + i\sin t)\cdot r^{-1}(\cos(-t)+i\sin(-t))\\ &= \cos(t-t)+i\sin(t-t)\\ &= \cos 0 + i\sin 0\\ &= 1 \end{aligned}$$

Finally, regarding the generalization to the case $k\in\mathbb{Q}$, it is enough to do the verification for exponents of type $k=1/n$, with $n\in\mathbb{N}$. The claim here is that:

$$x^{1/n} = r^{1/n}\left[\cos\left(\frac{t}{n}\right)+i\sin\left(\frac{t}{n}\right)\right]$$

In order to prove this, let us compute the n-th power of this number. We can use the power formula for the exponent $n\in\mathbb{N}$, that we already established, and we obtain:

$$\begin{aligned} (x^{1/n})^n &= (r^{1/n})^n\left[\cos\left(n\cdot\frac{t}{n}\right)+i\sin\left(n\cdot\frac{t}{n}\right)\right]\\ &= r(\cos t + i\sin t)\\ &= x \end{aligned}$$

Thus, we have indeed a n-th root of x, and our proof is now complete. □

We should mention that there is a bit of ambiguity in the above, in the case of the exponents $k\in\mathbb{Q}$, due to the fact that the square roots, and the higher roots as well, can take multiple values, in the complex number setting. We will be back to this.

Let us discuss now the final and most convenient writing of the complex numbers, which is a well-known variation on the polar writing, as follows:

$$x = re^{it}$$

In what follows we will not really need the true power of this formula, which is of analytic nature, due to occurrence of the number e. However, we would like to use the notation $x=re^{it}$, as everyone does, among others because it simplifies the writing. The point indeed with the above formula comes from the following deep result:

THEOREM 3.7. *We have the following formula, valid for any $t\in\mathbb{R}$,*

$$e^{it} = \cos t + i\sin t$$

where $e = 2.7182\ldots$ is the usual constant from analysis.

PROOF. In order to prove such a result, we must first recall what e is, and what e^x is. One way of viewing things is that e^x is the unique function satisfying:

$$(e^x)' = e^x \quad , \quad e^0 = 1$$

Then, we can set $e = e^1$, and then prove that e^x equals indeed e to the power x. This is a bit abstract, but is convenient for our purposes. Indeed, the solution to the above derivative problem is easy to work out, as a series, by recurrence, the answer being:

$$e^x = \sum_k \frac{x^k}{k!}$$

Now let us plug $x = it$ in this formula. We obtain the following formula:

$$\begin{aligned}
e^{it} &= \sum_k \frac{(it)^k}{k!} \\
&= \sum_{k=2l} \frac{(it)^k}{k!} + \sum_{k=2l+1} \frac{(it)^k}{k!} \\
&= \sum_l (-1)^l \frac{t^{2l}}{(2l)!} + i \sum_l (-1)^l \frac{t^{2l+1}}{(2l+1)!}
\end{aligned}$$

Our claim now, which will complete the proof, is that we have:

$$\cos t = \sum_l (-1)^l \frac{t^{2l}}{(2l)!} \quad , \quad \sin t = \sum_l (-1)^l \frac{t^{2l+1}}{(2l+1)!}$$

In order to prove this claim, let us compute the Taylor series of cos and sin. By using the formulae for sums of angles, used in the proof of Theorem 3.5, we have:

$$\sin' = \cos \quad , \quad \cos' = -\sin$$

Thus, we know how to differentiate sin and cos, once, then twice, then as many times as we want to, and with this we can compute the corresponding Taylor series, the answers being those given above. Now by putting everything together, we have:

$$e^{it} = \cos t + i \sin t$$

Thus, we are led to the conclusion in the statement. $\square$

All this was quite brief, but we will be back to it with details in chapters 5-8 below, when doing analysis. For the moment, let us just enjoy all this. We first have:

THEOREM 3.8. *We have the following formula,*

$$e^{\pi i} = -1$$

and we have $E = mc^2$ as well.

PROOF. We have two assertions here, the idea being as follows:

(1) The first formula, $e^{\pi i}=-1$, which is actually the main formula in mathematics, comes from Theorem 3.7, by setting $t=\pi$. Indeed, we obtain:

$$\begin{aligned} e^{\pi i} &= \cos\pi + i\sin\pi \\ &= -1 + i\cdot 0 \\ &= -1 \end{aligned}$$

(2) As for $E=mc^2$, which is the main formula in physics, this is something deep as well. Although we will not really need it here, we recommend learning it too, for symmetry reasons between math and physics, say from Feynman [37], [38], [39]. $\square$

Now back to our $x=re^{it}$ objectives, with the above theory in hand we can indeed use from now on this notation, the complete statement being as follows:

THEOREM 3.9. *The complex numbers $x=a+ib$ can be written in polar coordinates,*

$$x=re^{it}$$

with the connecting formulae being

$$a=r\cos t \quad , \quad b=r\sin t$$

and in the other sense being

$$r=\sqrt{a^2+b^2} \quad , \quad \tan t=b/a$$

and with r,t being called modulus, and argument.

PROOF. This is just a reformulation of Definition 3.4, by using the formula $e^{it}=\cos t+i\sin t$ from Theorem 3.7, and multiplying everything by r. $\square$

We can now go back to the basics, and we have the following result:

THEOREM 3.10. *In polar coordinates, the complex numbers multiply as*

$$re^{is}\cdot pe^{it}=rp\,e^{i(s+t)}$$

with the arguments s,t being taken modulo 2π.

PROOF. This is something that we already know, from Theorem 3.5, reformulated by using the notations from Theorem 3.9. Observe that this follows as well directly, from the fact that we have $e^{a+b}=e^ae^b$, that we know from analysis. $\square$

We can now investigate more complicated operations, as follows:

THEOREM 3.11. *We have the following operations on the complex numbers:*

(1) *Inversion:* $(re^{it})^{-1}=r^{-1}e^{-it}$.
(2) *Square roots:* $\sqrt{re^{it}}=\pm\sqrt{r}e^{it/2}$.
(3) *Powers:* $(re^{it})^a=r^ae^{ita}$.

PROOF. This is something that we already know, from Theorem 3.6, but we can now discuss all this, from a more conceptual viewpoint, the idea being as follows:

(1) We have indeed the following computation, using Theorem 3.10:

$$(re^{it})(r^{-1}e^{-it}) = rr^{-1} \cdot e^{i(t-t)} = 1$$

(2) Once again by using Theorem 3.10, we have:

$$(\pm\sqrt{r}e^{it/2})^2 = (\sqrt{r})^2 e^{i(t/2+t/2)} = re^{it}$$

(3) Given an arbitrary number $a \in \mathbb{R}$, we can define, as stated:

$$(re^{it})^a = r^a e^{ita}$$

Due to Theorem 3.10, this operation $x \to x^a$ is indeed the correct one. □

We can now go back to the degree 2 equations, and we have:

THEOREM 3.12. *The complex solutions of $ax^2 + bx + c = 0$ with $a, b, c \in \mathbb{C}$ are*

$$x_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

with the square root of complex numbers being defined as above.

PROOF. This is clear, the computations being the same as in the real case. To be more precise, our degree 2 equation can be written as follows:

$$\left(x + \frac{b}{2a}\right)^2 = \frac{b^2 - 4ac}{4a^2}$$

Now since we know from Theorem 3.11 (2) that any complex number has a square root, we are led to the conclusion in the statement. □

More generally now, we can prove that any polynomial equation, of arbitrary degree $N \in \mathbb{N}$, has exactly N complex solutions, counted with multiplicities:

THEOREM 3.13. *Any polynomial $P \in \mathbb{C}[X]$ decomposes as*

$$P = c(X - a_1) \ldots (X - a_N)$$

with $c \in \mathbb{C}$ and with $a_1, \ldots, a_N \in \mathbb{C}$.

PROOF. The problem is that of proving that our polynomial has at least one root, because afterwards we can proceed by recurrence. We prove this by contradiction. So, assume that P has no roots, and pick a number $z \in \mathbb{C}$ where $|P|$ attains its minimum:

$$|P(z)| = \min_{x \in \mathbb{C}} |P(x)| > 0$$

Since $Q(t)=P(z+t)-P(z)$ is a polynomial which vanishes at $t=0$, this polynomial must be of the form ct^k + higher terms, with $c\neq 0$, and with $k\geq 1$ being an integer. We obtain from this that, with $t\in\mathbb C$ small, we have the following estimate:

$$P(z+t)\simeq P(z)+ct^k$$

Now let us write $t=rw$, with $r>0$ small, and with $|w|=1$. Our estimate becomes:

$$P(z+rw)\simeq P(z)+cr^kw^k$$

Now recall that we have assumed $P(z)\neq 0$. We can therefore choose $w\in\mathbb T$ such that cw^k points in the opposite direction to that of $P(z)$, and we obtain in this way:

$$\begin{aligned}|P(z+rw)| &\simeq |P(z)+cr^kw^k|\\ &= |P(z)|(1-|c|r^k)\end{aligned}$$

Now by choosing $r>0$ small enough, as for the error in the first estimate to be small, and overcame by the negative quantity $-|c|r^k$, we obtain from this:

$$|P(z+rw)|<|P(z)|$$

But this contradicts our definition of $z\in\mathbb C$, as a point where $|P|$ attains its minimum. Thus P has a root, and by recurrence it has N roots, as stated. □

All this is very nice, and we will see applications in a moment. As a last topic now regarding the complex numbers, we have:

THEOREM 3.14. *The equation $x^N=1$ has N complex solutions, namely*

$$\left\{w^k\Big|k=0,1,\ldots,N-1\right\}\quad,\quad w=e^{2\pi i/N}$$

which are called roots of unity of order N.

PROOF. This follows from Theorem 3.10. Indeed, with $x=re^{it}$ our equation reads:

$$r^Ne^{itN}=1$$

Thus $r=1$, and $t\in[0,2\pi)$ must be a multiple of $2\pi/N$, as stated. □

As an illustration here, the roots of unity of small order, along with some of their basic properties, which are very useful for computations, are as follows:

<u>$N=1$</u>. Here the unique root of unity is 1.

<u>$N=2$</u>. Here we have two roots of unity, namely 1 and -1.

<u>$N=3$</u>. Here we have 1, then $w=e^{2\pi i/3}$, and then $w^2=\bar w=e^{4\pi i/3}$.

<u>$N=4$</u>. Here the roots of unity, read as usual counterclockwise, are $1,i,-1,-i$.

<u>$N=5$</u>. Here, with $w=e^{2\pi i/5}$, the roots of unity are $1,w,w^2,w^3,w^4$.

<u>$N=6$</u>. Here a useful alternative writing is $\{\pm1,\pm w,\pm w^2\}$, with $w=e^{2\pi i/3}$.

The roots of unity are very useful variables, and have many interesting properties. As a first application, we can now solve the ambiguity questions related to the extraction of N-th roots, from Theorem 3.6 and Theorem 3.11, the statement being as follows:

THEOREM 3.15. *Any nonzero complex number, written as*

$$x = re^{it}$$

has exactly N roots of order N, which appear as

$$y = r^{1/N}e^{it/N}$$

multiplied by the N roots of unity of order N.

PROOF. We must solve the equation $z^N = x$, over the complex numbers. Since the number y in the statement clearly satisfies $y^N = x$, our equation is equivalent to:

$$z^N = y^N$$

Now observe that we can write this equation as follows:

$$\left(\frac{z}{y}\right)^N = 1$$

We conclude that the solutions z appear by multiplying y by the solutions of $t^N = 1$, which are the N-th roots of unity, as claimed. □

The roots of unity appear in connection with many other questions, and there are many useful formulae relating them, which are good to know, as for instance:

THEOREM 3.16. *The roots of unity, $\{w^k\}$ with $w = e^{2\pi i/N}$, have the property*

$$\sum_{k=0}^{N-1}(w^k)^s = N\delta_{N|s}$$

for any exponent $s \in \mathbb{N}$, where on the right we have a Kronecker symbol.

PROOF. The numbers in the statement, when written more conveniently as $(w^s)^k$ with $k = 0, \ldots, N-1$, form a certain regular polygon in the plane P_s. Thus, if we denote by C_s the barycenter of this polygon, we have the following formula:

$$\frac{1}{N}\sum_{k=0}^{N-1} w^{ks} = C_s$$

Now observe that in the case $N \nmid s$ our polygon P_s is non-degenerate, circling around the unit circle, and having center $C_s = 0$. As for the case $N|s$, here the polygon is degenerate, lying at 1, and having center $C_s = 1$. Thus, we have the following formula:

$$C_s = \delta_{N|s}$$

Thus, we obtain the formula in the statement. □

Before moving ahead, let us recommend some reading. The mathematics in this book requires a deep commitment to the complex numbers. And here, while there are surely plenty of good logical reasons for using $\mathbb{C}$ instead of $\mathbb{R}$, such as those evoked above, regarding real polynomials which can have complex roots, or real matrices which can have complex eigenvalues, and so on, things remain a bit abstract.

In order to overcome this hurdle, and get to love complex numbers, nothing better than learning some physics. Indeed, physics is all about waves. Water waves of course, but also light and other electromagnetic waves, and also elementary particles, which are treated like waves too. And all these waves naturally live over $\mathbb{C}$. And even more, you will learn in fact that quantum mechanics itself lives over $\mathbb{C}$, and so farewell $\mathbb{R}$.

The standard books for learning physics are those of Feynman [37], [38], [39]. At a more advanced level, equally fun and entertaining, you have the books of Griffiths [41], [42], [43]. There are also plenty of more popular books that you can start with, such as Griffiths again [44], or Huang [49], or Kumar [57]. Up to you here, and in the hope that you will follow my advice. There ain't such thing as mathematics without physics, and the answer to any mathematical question that you might have, technical or philosophical, now and in the future, always, but really always, believe me, lies in physics.

3b. Linear maps

Back now to linear algebra, our first task will be that of extending the results that we know, from the real case, to the complex case. We first have:

THEOREM 3.17. *The linear maps $f:\mathbb{C}^N\to\mathbb{C}^M$ are the maps of the form*

$$f(x)=Ax$$

with A being a rectangular matrix, $A\in M_{M\times N}(\mathbb{C})$.

PROOF. This follows as in the real case. Indeed, $f:\mathbb{C}^N\to\mathbb{C}^M$ must send a vector $x\in\mathbb{C}^N$ to a certain vector $f(x)\in\mathbb{C}^M$, all whose components are linear combinations of the components of x. Thus, we can write, for certain complex numbers $a_{ij}\in\mathbb{C}$:

$$f\begin{pmatrix}x_1\\ \vdots\\ \vdots\\ x_N\end{pmatrix}=\begin{pmatrix}a_{11}x_1+\ldots+a_{1N}x_N\\ \vdots\\ \vdots\\ a_{M1}x_1+\ldots+a_{MN}x_N\end{pmatrix}$$

But the parameters $a_{ij}\in\mathbb{C}$ can be regarded as being the entries of a matrix:

$$A=(a_{ij})\in M_{M\times N}(\mathbb{C})$$

Now with the usual convention for the rectangular matrix multiplication, exactly as in the real case, the above formula is precisely the one in the statement. $\square$

We have as well the following result:

THEOREM 3.18. *A linear map $f:\mathbb{C}^N\to\mathbb{C}^M$, written as*

$$f(v)=Av$$

is invertible precisely when A is invertible, and in this case we have:

$$f^{-1}(v)=A^{-1}v$$

PROOF. As in the real case, with the convention $f_A(v)=Av$, we have the following multiplication formula for such linear maps:

$$f_Af_B(v)=f_{AB}(v)$$

But this shows that $f_Af_B=1$ is equivalent to $AB=1$, as desired. □

With respect to the real case, some subtleties appear at the level of the scalar products, isometries and projections. The basic theory here is as follows:

THEOREM 3.19. *Consider the usual scalar product $<x,y>=\sum_i x_i\bar{y}_i$ on $\mathbb{C}^N$.*

(1) *We have the following formula, where $(A^*)_{ij}=\bar{A}_{ji}$ is the adjoint matrix:*

$$<Ax,y>=<x,A^*y>$$

(2) *A linear map $f:\mathbb{C}^N\to\mathbb{C}^N$, written as $f(x)=Ux$ with $U\in M_N(\mathbb{C})$, is an isometry precisely when U is unitary, in the sense that:*

$$U^*=U^{-1}$$

(3) *A linear map $f:\mathbb{C}^N\to\mathbb{C}^N$, written as $f(x)=Px$ with $P\in M_N(\mathbb{C})$, is a porojection precisely when P is projection, in the sense that:*

$$P=P^2=P^*$$

(4) *The formula for the rank 1 projections is as follows:*

$$P_x=\frac{1}{||x||^2}(x_i\bar{x}_j)_{ij}$$

PROOF. This follows as in the real case, by performing modifications where needed:

(1) By using the standard basis of $\mathbb{C}^N$, we want to prove that for any i,j we have:

$$<Ae_j,e_i>=<e_j,A^*e_i>$$

The scalar product being now antisymmetric, this is the same as proving that:

$$<Ae_j,e_i>=\overline{<A^*e_i,e_j>}$$

On the other hand, for any matrix M we have the following formula:

$$M_{ij}=<Me_j,e_i>$$

Thus, the formula to be proved simply reads:

$$A_{ij} = \overline{(A^*)_{ji}}$$

But this is precisely the definition of A^*, and we are done.

(2) Let first recall that we can pass from scalar products to distances, as follows:

$$||x|| = \sqrt{< x, x >}$$

Conversely, we can compute the scalar products in terms of distances, by using the complex polarization identity, which is as follows:

$$\begin{aligned}
& ||x+y||^2 - ||x-y||^2 + i||x+iy||^2 - i||x-iy||^2 \\
= & ||x||^2 + ||y||^2 - ||x||^2 - ||y||^2 + i||x||^2 + i||y||^2 - i||x||^2 - i||y||^2 \\
& +2Re(< x, y >) + 2Re(< x, y >) + 2iIm(< x, y >) + 2iIm(< x, y >) \\
= & 4 < x, y >
\end{aligned}$$

Now given a matrix $U \in M_N(\mathbb{C})$, we have the following equivalences, with the first one coming from the above identities, and with the other ones being clear:

$$\begin{aligned}
||Ux|| = ||x|| & \iff < Ux, Uy >=< x, y > \\
& \iff < x, U^*Uy >=< x, y > \\
& \iff U^*Uy = y \\
& \iff U^*U = 1 \\
& \iff U^* = U^{-1}
\end{aligned}$$

(3) As in the real case, P is an abstract projection, not necessarily orthogonal, when $P^2 = P$. The point now is that this projection is orthogonal when:

$$\begin{aligned}
< Px - Py, Px - x >= 0 & \iff < x - y, P^*Px - P^*x >= 0 \\
& \iff P^*Px - P^*x = 0 \\
& \iff P^*P - P^* = 0
\end{aligned}$$

Thus we must have $P^* = P^*P$. Now observe that by conjugating, we obtain:

$$\begin{aligned}
P & = (P^*P)^* \\
& = P^*(P^*)^* \\
& = P^*P
\end{aligned}$$

Now by comparing with the original relation, $P^* = P^*P$, we conclude that $P = P^*$. Thus, we have shown that any orthogonal projection must satisfy, as claimed:

$$P^2 = P = P^*$$

Conversely, if this condition is satisfied, $P^2 = P$ shows that P is a projection, and $P = P^*$ shows via the above computation that P is indeed orthogonal.

(4) Once again in analogy with the real case, we have the following formula:

$$\begin{aligned} P_x y &= \frac{< y, x >}{< x, x >} x \\ &= \frac{1}{||x||^2} < y, x > x \end{aligned}$$

With this in hand, we can now compute the entries of P_x, as follows:

$$\begin{aligned} (P_x)_{ij} &= < P_x e_j, e_i > \\ &= \frac{1}{||x||^2} < e_j, x >< x, e_i > \\ &= \frac{\bar{x}_j x_i}{||x||^2} \end{aligned}$$

Thus, we are led to the formula in the statement. □

3c. Diagonalization

In the present complex matrix setting, we can talk as well about eigenvalues and eigenvectors, exactly as in the real case, as follows:

Definition 3.20. *Let $A \in M_N(\mathbb{C})$ be a square matrix. When $Av = \lambda v$ we say that:*

(1) *v is an eigenvector of A.*
(2) *λ is an eigenvalue of A.*

We say that A is diagonalizable when $\mathbb{C}^N$ has a basis of eigenvectors of A.

When A is diagonalizable, in that basis of eigenvectors we can write:

$$A = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_N \end{pmatrix}$$

In general, this means that we have a formula as follows, with D diagonal:

$$A = PDP^{-1}$$

Indeed, we can take P to be the matrix formed by the eigenvectors:

$$P = [v_1 \ldots v_N]$$

As a first interesting result now, regarding the real matrices, we have:

Theorem 3.21. *The eigenvalues of a real matrix $A \in M_N(\mathbb{R})$ are the roots of the characteristic polynomial, given by:*

$$P(x) = \det(A - x1_N)$$

In particular, any such matrix $A \in M_N(\mathbb{R})$ has at least 1 complex eigenvalue.

Proof. The first assertion is something that we already know from chapter 2 above, which follows from the following computation:

$$\begin{aligned}\exists v, Av=\lambda v &\iff \exists v,(A-\lambda 1_N)v=0\\ &\iff \det(A-\lambda 1_N)=0\end{aligned}$$

As for the second assertion, this follows from the first assertion, and from Theorem 3.13, which shows in particular that P has at least 1 complex root. $\square$

It is possible to further build on these results, but this is quite long, and we will rather do this in the next chapter. For the moment, let us just keep in mind the conclusion that a real matrix $A\in M_N(\mathbb{R})$ has substantially more chances of being diagonalizable over the complex numbers, than over the real numbers. As an illustration for this principle, and as a first concrete result, which is of true complex nature, we have:

Theorem 3.22. *The rotation of angle $t\in\mathbb{R}$ in the real plane, namely*

$$R_t=\begin{pmatrix}\cos t&-\sin t\\ \sin t&\cos t\end{pmatrix}$$

can be diagonalized over the complex numbers, as follows:

$$R_t=\frac{1}{2}\begin{pmatrix}1&1\\ i&-i\end{pmatrix}\begin{pmatrix}e^{-it}&0\\ 0&e^{it}\end{pmatrix}\begin{pmatrix}1&-i\\ 1&i\end{pmatrix}$$

Over the real numbers this is impossible, unless $t=0,\pi$.

Proof. The last assertion is something clear, that we already know, coming from the fact that at $t\neq 0,\pi$ our rotation is a "true" rotation, having no eigenvectors in the plane. Regarding the first assertion, the point is that we have the following computation:

$$\begin{aligned}R_t\begin{pmatrix}1\\ i\end{pmatrix}&=\begin{pmatrix}\cos t&-\sin t\\ \sin t&\cos t\end{pmatrix}\begin{pmatrix}1\\ i\end{pmatrix}\\ &=\begin{pmatrix}\cos t-i\sin t\\ i\cos t+\sin t\end{pmatrix}\\ &=e^{-it}\begin{pmatrix}1\\ i\end{pmatrix}\end{aligned}$$

We have as well a second eigenvector, as follows:

$$\begin{aligned}R_t\begin{pmatrix}1\\ -i\end{pmatrix}&=\begin{pmatrix}\cos t&-\sin t\\ \sin t&\cos t\end{pmatrix}\begin{pmatrix}1\\ -i\end{pmatrix}\\ &=\begin{pmatrix}\cos t+i\sin t\\ -i\cos t+\sin t\end{pmatrix}\\ &=e^{it}\begin{pmatrix}1\\ -i\end{pmatrix}\end{aligned}$$

Thus our matrix R_t is diagonalizable over $\mathbb{C}$, with the diagonal form being:

$$R_t \sim \begin{pmatrix} e^{-it} & 0 \\ 0 & e^{it} \end{pmatrix}$$

As for the passage matrix, obtained by putting together the eigenvectors, this is:

$$P = \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix}$$

In order to invert now P, we can use the standard inversion formula for the 2×2 matrices, which is the same as the one in the real case, and which gives:

$$P^{-1} = \frac{1}{-2i} \begin{pmatrix} -i & -1 \\ -i & 1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix}$$

Our diagonalization formula is therefore as follows:

$$R_t = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix} \begin{pmatrix} e^{-it} & 0 \\ 0 & e^{it} \end{pmatrix} \begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix}$$

Thus, we are led to the conclusion in the statement. □

3d. The determinant

Regarding now the determinant, for the complex matrices it is more convenient to follow an abstract approach, and this due to our lack of geometric intuition with the space $\mathbb{C}^N$, at $N \geq 2$, and with the volumes of the bodies there. We have:

Definition 3.23. *The determinant of a complex matrix $A \in M_N(\mathbb{C})$ is given by*

$$\det A = \sum_{\sigma \in S_N} \varepsilon(\sigma) A_{1\sigma(1)} \ldots A_{N\sigma(N)}$$

with $\varepsilon = \pm 1$ being the signature of the permutations.

Generally speaking, the theory of the determinant from the real case extends well. To be more precise, we first have the following result, summarizing the needed properties and formulae of the determinant of the complex matrices:

Theorem 3.24. *The determinant has the following properties:*

(1) *When adding two columns, the determinants get added:*

$$\det(\ldots, u+v, \ldots) = \det(\ldots, u, \ldots) + \det(\ldots, v, \ldots)$$

(2) *When multiplying columns by scalars, the determinant gets multiplied:*

$$\det(\lambda v_1, \ldots, \lambda_N v_N) = \lambda_1 \ldots \lambda_N \det(v_1, \ldots, v_N)$$

(3) *When permuting two columns, the determinant changes the sign:*

$$\det(\ldots, v, \ldots, w, \ldots) = -\det(\ldots, w, \ldots, v, \ldots)$$

Proof. This follows indeed by doing some elementary algebraic computations with permutations, which are similar to those in the real case, but now done backwards, based on the formula of the determinant from Definition 3.23. □

We have as well a similar result for the rows, which completes the list of needed operations which are usually needed in practice, as follows:

Theorem 3.25. *The determinant has the following properties:*

(1) *When adding two rows, the determinants get added:*

$$\det\begin{pmatrix}\vdots\\ u+v\\ \vdots\end{pmatrix}=\det\begin{pmatrix}\vdots\\ u\\ \vdots\end{pmatrix}+\det\begin{pmatrix}\vdots\\ v\\ \vdots\end{pmatrix}$$

(2) *When multiplying rows by scalars, the determinant gets multiplied:*

$$\det\begin{pmatrix}\lambda_1v_1\\ \vdots\\ \lambda_Nv_N\end{pmatrix}=\lambda_1\ldots\lambda_N\det\begin{pmatrix}v_1\\ \vdots\\ v_N\end{pmatrix}$$

(3) *When permuting two rows, the determinant changes the sign.*

Proof. This follows once again by doing some algebraic computations with permutations, based on the formula of the determinant from Definition 3.23. □

Next in line, we have the following result, which is very useful in practice:

Theorem 3.26. *The determinant is subject to the row expansion formula*

$$\begin{aligned}\begin{vmatrix}a_{11}&\ldots&a_{1N}\\ \vdots&&\vdots\\ a_{N1}&\ldots&a_{NN}\end{vmatrix}&=&a_{11}\begin{vmatrix}a_{22}&\ldots&a_{2N}\\ \vdots&&\vdots\\ a_{N2}&\ldots&a_{NN}\end{vmatrix}\\ &-&a_{12}\begin{vmatrix}a_{21}&a_{23}&\ldots&a_{2N}\\ \vdots&\vdots&&\vdots\\ a_{N1}&a_{N3}&\ldots&a_{NN}\end{vmatrix}\\ &\vdots&\\ &\vdots&\\ &+&(-1)^{N+1}a_{1N}\begin{vmatrix}a_{21}&\ldots&a_{2,N-1}\\ \vdots&&\vdots\\ a_{N1}&\ldots&a_{N,N-1}\end{vmatrix}\end{aligned}$$

and this method fully computes it, by recurrence.

Proof. This follows indeed by doing some elementary algebraic computations. □

We can expand as well over the columns, as follows:

THEOREM 3.27. *The determinant is subject to the column expansion formula*

$$\begin{aligned}
\begin{vmatrix}a_{11}&\ldots&a_{1N}\\ \vdots&&\vdots\\ a_{N1}&\ldots&a_{NN}\end{vmatrix}
&= a_{11}\begin{vmatrix}a_{22}&\ldots&a_{2N}\\ \vdots&&\vdots\\ a_{N2}&\ldots&a_{NN}\end{vmatrix}\\
&- a_{21}\begin{vmatrix}a_{12}&\ldots&a_{1N}\\ a_{32}&\ldots&a_{3N}\\ \vdots&&\vdots\\ a_{N2}&\ldots&a_{NN}\end{vmatrix}\\
&\vdots\\
&\vdots\\
&+ (-1)^{N+1}a_{N1}\begin{vmatrix}a_{12}&\ldots&a_{1N}\\ \vdots&&\vdots\\ a_{N-1,2}&\ldots&a_{N-1,N}\end{vmatrix}
\end{aligned}$$

and this method fully computes it, by recurrence.

PROOF. Once again, this follows by doing some algebraic computations. □

Still in analogy with the real case, we have the following result:

THEOREM 3.28. *The determinant of the systems of vectors*

$$\det : \mathbb{C}^N \times \ldots \times \mathbb{C}^N \to \mathbb{C}$$

is multilinear, alternate and unital, and unique with these properties.

PROOF. This is something that we know in the real case, and the proof in the complex case is similar, with the conditions in the statement corresponding to those in Theorem 3.24. It is possible to prove this result as well directly, by doing some abstract algebra. □

Finally, once again at the general level, let us record the following result:

THEOREM 3.29. *We have the following formulae,*

$$\det \bar{A} = \overline{\det A}$$

$$\det A^t = \det A$$

$$\det A^* = \overline{\det A}$$

valid for any square matrix $A \in M_N(\mathbb{C})$.

PROOF. The first formula is clear from Definition 3.23, because when conjugating the entries of A, the determinant will get conjugated:

$$\det\bar{A}=\sum_{\sigma\in S_N}\varepsilon(\sigma)\overline{A_{1\sigma(1)}\ldots A_{N\sigma(N)}}$$

The second formula follows as in the real case, as follows:

$$\begin{aligned}\det A^t&=\sum_{\sigma\in S_N}\varepsilon(\sigma)(A^t)_{1\sigma(1)}\ldots(A^t)_{N\sigma(N)}\\&=\sum_{\sigma\in S_N}\varepsilon(\sigma)A_{\sigma(1)1}\ldots A_{\sigma(N)N}\\&=\sum_{\sigma\in S_N}\varepsilon(\sigma)A_{1\sigma^{-1}(1)}\ldots A_{N\sigma^{-1}(N)}\\&=\sum_{\sigma\in S_N}\varepsilon(\sigma^{-1})A_{1\sigma^{-1}(1)}\ldots A_{N\sigma^{-1}(N)}\\&=\sum_{\sigma\in S_N}\varepsilon(\sigma)A_{1\sigma(1)}\ldots A_{N\sigma(N)}\\&=\det A\end{aligned}$$

As for the third formula, this follows from the first two formulae, by using:

$$\det A^*=\det\bar{A}^t$$

Thus, we are led to the conclusions in the statement. □

Summarizing, the theory from the real case extends well, and we have complex analogues of all results. As in the real case, as a main application of all this, we have:

THEOREM 3.30. *The inverse of a square matrix, having nonzero determinant,*

$$A=\begin{pmatrix}a_{11}&\ldots&a_{1N}\\\vdots&&\vdots\\a_{N1}&\ldots&a_{NN}\end{pmatrix}$$

is given by the following formula,

$$A^{-1}=\frac{1}{\det A}\begin{pmatrix}\det A^{(11)}&-\det A^{(21)}&\det A^{(31)}&\ldots\\-\det A^{(12)}&\det A^{(22)}&-\det A^{(32)}&\ldots\\\det A^{(13)}&-\det A^{(23)}&\det A^{(33)}&\ldots\\\vdots&\vdots&\vdots\end{pmatrix}$$

where $A^{(ij)}$ is the matrix A, with the i-th row and j-th column removed.

PROOF. This follows indeed by using the row expansion formula from Theorem 3.26, which in terms of the matrix A^{-1} in the statement reads $AA^{-1}=1$. □

As a final topic now, regarding the complex matrices, let us discuss some interesting examples of such matrices, which definitely do not exist in the real setting, and which are very useful, even in connection with real matrix questions. Let us start with:

DEFINITION 3.31. *The Fourier matrix is as follows,*

$$F_N = (w^{ij})_{ij}$$

with $w = e^{2\pi i/N}$, and with the convention that the indices are

$$i, j \in \{0, 1, \ldots, N-1\}$$

and are usually taken modulo N.

Here the conventions regarding the indices are standard, and are there for various reasons, as for instance for having the first row and column consisting of 1 entries. Indeed, in standard matrix form, and with the above conventions for the indices, we have:

$$F_N = \begin{pmatrix} 1 & 1 & 1 & \ldots & 1 \\ 1 & w & w^2 & \ldots & w^{N-1} \\ 1 & w^2 & w^4 & \ldots & w^{2(N-1)} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & w^{N-1} & w^{2(N-1)} & \ldots & w^{(N-1)^2} \end{pmatrix}$$

Thus, what we have here is a Vandermonde matrix, in the sense of chapter 2, of very special type. Let us record as well the first few values of these matrices:

PROPOSITION 3.32. *The second Fourier matrix is as follows:*

$$F_2 = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

The third Fourier matrix is as follows, with $w = e^{2\pi i/3}$:

$$F_3 = \begin{pmatrix} 1 & 1 & 1 \\ 1 & w & w^2 \\ 1 & w^2 & w \end{pmatrix}$$

As for the fourth Fourier matrix, this is as follows:

$$F_4 = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & i & -1 & -i \\ 1 & -1 & 1 & -1 \\ 1 & -i & -1 & i \end{pmatrix}$$

PROOF. All these formulae are clear from definitions, with our usual convention for the indices of the Fourier matrices, from Definition 3.31. $\square$

Our claim now is that the Fourier matrix can be used in order to solve a variety of linear algebra questions, a bit in a same way as the Fourier transform can be used in order to solve analysis questions. Before discussing all this, however, let us analyze the Fourier matrix F_N, from a linear algebra perspective. We have the following result:

THEOREM 3.33. *The Fourier matrix F_N has the following properties:*
(1) *It is symmetric, $F_N^t = F_N$.*
(2) *The matrix $F_N/\sqrt{N}$ is unitary.*
(3) *Its inverse is the matrix F_N^*/N.*

PROOF. This is a collection of elementary results, the idea being as follows:

(1) This is clear from definitions.

(2) The row vectors $R_0, \ldots, R_{N-1}$ of the rescaled matrix $F_N/\sqrt{N}$ have all length 1, and by using the barycenter formula in Theorem 3.16, we have, for any $i \neq j$:

$$\begin{aligned} < R_i, R_j > &= \frac{1}{N}\sum_k w^{ik} w^{-jk} \\ &= \frac{1}{N}\sum_k (w^{i-j})^k \\ &= 0 \end{aligned}$$

Thus, $R_0, \ldots, R_{N-1}$ are pairwise orthogonal, and so $F_N/\sqrt{N}$ is unitary, as claimed.

(3) This follows from (1) and (2), because for a symmetric matrix, the adjoint is the conjugate, and in the unitary case, this is the inverse. □

Now back to our motivations, we were saying before that the Fourier matrix is to linear algebra what the Fourier transform is to analysis, namely advanced technology. In order to discuss now an illustrating application of the theory developed above, let us go back to our favorite example of a $N \times N$ matrix, namely the flat matrix:

$$\mathbb{I}_N = \begin{pmatrix} 1 & \ldots & 1 \\ \vdots & & \vdots \\ 1 & \ldots & 1 \end{pmatrix}$$

This is a real matrix, and we know that we have $\mathbb{I}_N = NP_N$, with P_N being the projection on the all-1 vector $\xi = (1)_i \in \mathbb{R}^N$. Thus, $\mathbb{I}_N$ diagonalizes over $\mathbb{R}$:

$$\mathbb{I}_N \sim \begin{pmatrix} N & & & \\ & 0 & & \\ & & \ddots & \\ & & & 0 \end{pmatrix}$$

The problem, however, is that when looking for 0-eigenvectors, in order to have an explicit diagonalization formula, we must solve the following equation:

$$x_1 + \ldots + x_N = 0$$

And this is not an easy task, if our objective is that of finding a nice and explicit basis for the space of solutions. To be more precise, if we want linearly independent vectors $v_1, \ldots, v_{N-1} \in \mathbb{R}^N$, each with components summing up to 0, and which are given by simple formulae, of type $(v_i)_j =$ explicit function of i, j, we are in trouble.

Fortunately, complex numbers and Fourier analysis are there, and we have:

Theorem 3.34. *The flat matrix, namely*

$$\mathbb{I}_N = \begin{pmatrix} 1 & \ldots & 1 \\ \vdots & & \vdots \\ 1 & \ldots & 1 \end{pmatrix}$$

has the following explicit diagonalization, over the complex numbers,

$$\mathbb{I}_N = \frac{1}{N} F_N Q F_N^*$$

with $F_N = (w^{ij})_{ij}$ being the Fourier matrix, and with $Q = diag(N, 0, \ldots, 0)$.

Proof. Indeed, the 0-eigenvector problem discussed above can be solved explicitely over the complex numbers, with the solution $(v_i)_j = w^{ij}$, with $w = e^{2\pi i/N}$. $\square$

There are many other uses of F_N, along the same lines. We will be back to all this in chapter 7 below, with a complete discussion of the Fourier matrices, and of their generalizations, called complex Hadamard matrices.

3e. Exercises

As a first exercise, in relation with the complex numbers, we have:

Exercise 3.35. *Try to use a complex numbers type idea in order to deal with the vectors of $\mathbb{R}^3$, then $\mathbb{R}^4$, and report on what you found.*

This is something quite tricky, and a piece of hint, do not worry if you find nothing interesting at $N = 3$. However, the $N = 4$ case is definitely worth some study.

Exercise 3.36. *Can you use complex numbers in order to explicitely find the roots of arbitrary degree 3 polynomials, a bit in the same way as in degree 2?*

As an indication here, better ignore what comes out of an internet search on this question. Or rather read that stuff, then try applying it to some concrete degree 3 polynomial, see what you get, if you are satisfied with the answer or not.

Exercise 3.37. *Write down a complete proof for the formula*

$$e^{\pi i} = -1$$

using any method of your choice.

This is something that we talked about in the above, with the proof however still missing a few details, regarding the basic properties of the function e^x. Thus, you can either try to recover all these details, or go with some other idea, of your choice.

Exercise 3.38. *Find a geometric interpretation of the formula*

$$\begin{pmatrix}\cos t&-\sin t\\ \sin t&\cos t\end{pmatrix}=\frac{1}{2}\begin{pmatrix}1&1\\ i&-i\end{pmatrix}\begin{pmatrix}e^{-it}&0\\ 0&e^{it}\end{pmatrix}\begin{pmatrix}1&-i\\ 1&i\end{pmatrix}$$

which diagonalizes the rotation of angle $t \in \mathbb{R}$ in the real plane.

This is something quite tricky, and of course, enjoy.

Exercise 3.39. *Develop a complete theory of diagonalization for the 2×2 matrices, notably by deciding when exactly such a matrix is diagonalizable.*

This is quite non-trivial, but all the needed ingredients are in the above.

Exercise 3.40. *Work out all the details of the diagonalization formula*

$$\mathbb{I}_N = \frac{1}{N} F_N Q F_N^*$$

with $Q = diag(N, 0, \ldots, 0)$, and then try formulating a generalization of this.

Here the first question is standard, amounting in completing the proof that was given in the above. As for the second question, this is something more tricky.

CHAPTER 4

Diagonalization

4a. Linear spaces

In this chapter we discuss the diagonalization question, with a number of advanced results, for the complex matrices $A\in M_N(\mathbb C)$. Our techniques will apply in particular to the real case, $A\in M_N(\mathbb R)$, and we will obtain a number of non-trivial results regarding the diagonalization of such matrices, over the complex numbers.

In order to obtain our results, the idea will be that of looking at the eigenvectors $v\in\mathbb C^N$ corresponding to a given eigenvalue $\lambda\in\mathbb C$, taken altogether:

Definition 4.1. *Given a matrix $A\in M_N(\mathbb C)$, its eigenspaces are given by:*

$$E_\lambda=\left\{v\in\mathbb C^N\Big|Av=\lambda v\right\}$$

That is, E_λ collects all the eigenvectors having a given eigenvalue $\lambda\in\mathbb C$.

As a first observation, these eingenspaces are stable under taking sums, and also under multiplying by scalars, and this due to the following trivial facts:

$$Av=\lambda v,Aw=\lambda w\implies A(v+w)=\lambda(v+w)$$

$$Av=\lambda v,\mu\in\mathbb C\implies A(\mu v)=\lambda(\mu v)$$

Abstractly speaking, these two properties tell us that the eigenspaces E_λ are linear subspaces of the vector space $\mathbb C^N$, in the following sense:

Definition 4.2. *A subset $V\subset\mathbb C^N$ is called a linear space, or vector space, when:*

$$v,w\in V\implies v+w\in V$$

$$\lambda\in\mathbb C,v\in V\implies\lambda v\in V$$

That is, V must be stable under taking sums, and multiplying by scalars.

Before getting into the study of the eigenspaces of a given matrix $A\in M_N(\mathbb C)$, let us develop some general theory for the arbitrary linear spaces $V\subset\mathbb C^N$.

As a first objective, we would like to talk about bases of such subspaces $V\subset\mathbb C^N$. We already know, from chapter 1 above, how to do this for $V=\mathbb C^N$ itself. In order to review and improve the material there, let us start with the following key result:

Proposition 4.3. *For a matrix $A\in M_N(\mathbb C)$, the following are equivalent:*

(1) *A is right invertible: $\exists B\in M_N(\mathbb C)$, $AB=1$.*
(2) *A is left invertible: $\exists B\in M_N(\mathbb C)$, $BA=1$.*
(3) *A is invertible: $\exists B\in M_N(\mathbb C)$, $AB=BA=1$.*
(4) *A has nonzero determinant: $\det A\neq 0$.*

Proof. This follows indeed from the theory of the determinant, as developed in chapter 3 above, with all the conditions being equivalent to (4). □

In terms of linear maps, we are led to the following statement:

Proposition 4.4. *For a linear map $f:\mathbb C^N\to\mathbb C^N$, the following are equivalent:*

(1) *f is injective.*
(2) *f is surjective.*
(3) *f is bijective.*
(4) *$f(x)=Ax$, with A invertible.*

Proof. This follows indeed from Proposition 4.3, by using the composition formula $f_{AB}=f_Af_B$ for the linear maps $\mathbb C^N\to\mathbb C^N$, written in the form $f_A(x)=Ax$. □

Getting now to the point where we wanted to get, we have:

Theorem 4.5. *For a system of vectors $v_1,\ldots,v_N\in\mathbb C^N$, the following are equivalent:*

(1) *The vectors v_i are linearly independent, in the sense that:*
$$\sum\lambda_iv_i=0\implies\lambda_i=0$$
(2) *The vectors v_i span $\mathbb C^N$, in the sense that any $x\in\mathbb C^N$ can be written as:*
$$x=\sum\lambda_iv_i$$
(3) *The vectors v_i form a basis of $\mathbb C^N$, in the sense that any vector $x\in\mathbb C^N$ can be written in a unique way as:*
$$x=\sum\lambda_iv_i$$
(4) *The matrix formed by these vectors, regarded as usual as column vectors,*
$$P=[v_1,\ldots,v_N]\in M_N(\mathbb C)$$
is invertible, with respect to the usual multiplication of the matrices.

Proof. Given a family of vectors $\{v_1,\ldots,v_N\}\subset\mathbb C^N$ as in the statement, consider the following linear map, associated to these vectors:
$$f:\mathbb C^N\to\mathbb C^N\quad,\quad\lambda\to\sum\lambda_iv_i$$

The conditions (1,2,3) in the statement tell us that this map must be respectively injective, surjective, bijective. As for the condition (4) in the statement, this is related as well to this map, because the matrix of this map is the matrix P there:

$$f(\lambda)=P\lambda$$

With these observations in hand, the result follows from Proposition 4.4. $\square$

More generally now, we have the following result:

THEOREM 4.6. *For a system of vectors $\{v_1,\ldots,v_M\}\subset V$ belonging to a linear subspace $V\subset\mathbb C^N$, the following conditions are equivalent:*

(1) *The vectors v_i are linearly independent, in the sense that:*

$$\sum\lambda_iv_i=0\implies\lambda_i=0$$

(2) *The vectors v_i span V, in the sense that any $x\in V$ can be written as:*

$$x=\sum\lambda_iv_i$$

(3) *The vectors v_i form a basis of V, in the sense that any vector $x\in V$ can be written in a unique way as:*

$$x=\sum\lambda_iv_i$$

If these conditions are satisfied, we say that V has dimension M, and this dimension is independent on the choice of the basis.

PROOF. This is something that we know for $V=\mathbb C^N$ itself, coming from Theorem 4.5 above, with the dimension discussion at the end being something elementary. As for the extension to the general case, involving an arbitrary linear space $V\subset\mathbb C^N$, this is something which is routine, by using recurrence arguments, where needed. $\square$

4b. Diagonalization

Let us begin with a reminder of the basic diagonalization theory, that we already know. The basic theory that we have so far can be summarized as follows:

THEOREM 4.7. *Assuming that a matrix $A\in M_N(\mathbb C)$ is diagonalizable, in the sense that $\mathbb C^N$ has a basis formed by eigenvectors of A, we have*

$$A=PDP^{-1}$$

where $P=[v_1\ldots v_N]$ is the square matrix formed by the eigenvectors of A, and $D=diag(\lambda_1,\ldots,\lambda_N)$ is the diagonal matrix formed by the corresponding eigenvalues.

Proof. This is something that we already know, coming by changing the basis. We can prove this by direct computation as well, because we have $Pe_i = v_i$, and so the matrices A and PDP^{-1} follow to act in the same way on the basis vectors v_i:

$$\begin{aligned} PDP^{-1}v_i &= PDe_i \\ &= P\lambda_i e_i \\ &= \lambda_i Pe_i \\ &= \lambda_i v_i \end{aligned}$$

Thus, the matrices A and PDP^{-1} coincide, as stated. □

In general, in order to study the diagonalization problem, the idea is that the eigenvectors can be grouped into linear spaces, called eigenspaces, as follows:

Theorem 4.8. *Let $A \in M_N(\mathbb{C})$, and for any eigenvalue $\lambda \in \mathbb{C}$ define the corresponding eigenspace as being the vector space formed by the corresponding eigenvectors:*

$$E_\lambda = \left\{ v \in \mathbb{C}^N \Big| Av = \lambda v \right\}$$

These eigenspaces E_λ are then in a direct sum position, in the sense that given vectors $v_1 \in E_{\lambda_1}, \ldots, v_k \in E_{\lambda_k}$ corresponding to different eigenvalues $\lambda_1, \ldots, \lambda_k$, we have:

$$\sum_i c_i v_i = 0 \implies c_i = 0$$

In particular, we have the following dimension inequality, with the sum being over all the eigenvalues $\lambda \in \mathbb{C}$ of our matrix A,

$$\sum_\lambda \dim(E_\lambda) \leq N$$

and our matrix is diagonalizable precisely when we have equality.

Proof. We prove the first assertion by recurrence on $k \in \mathbb{N}$. Assume by contradiction that we have a formula as follows, with the scalars $c_1, \ldots, c_k$ being not all zero:

$$c_1 v_1 + \ldots + c_k v_k = 0$$

By dividing by one of these scalars, we can assume that our formula is:

$$v_k = c_1 v_1 + \ldots + c_{k-1} v_{k-1}$$

Now let us apply A to this vector. On the left we obtain:

$$Av_k = \lambda_k c_1 v_1 + \ldots + \lambda_k c_{k-1} v_{k-1}$$

On the right we obtain something different, as follows:

$$\begin{aligned} A(c_1 v_1 + \ldots + c_{k-1} v_{k-1}) &= c_1 A v_1 + \ldots + c_{k-1} A v_{k-1} \\ &= c_1 \lambda_1 v_1 + \ldots + c_{k-1} \lambda_{k-1} v_{k-1} \end{aligned}$$

We conclude from this that the following equality must hold:

$$\lambda_k c_1 v_1 + \ldots + \lambda_k c_{k-1} v_{k-1} = c_1 \lambda_1 v_1 + \ldots + c_{k-1} \lambda_{k-1} v_{k-1}$$

On the other hand, we know by recurrence that the vectors $v_1, \ldots, v_{k-1}$ must be linearly independent. Thus, the coefficients must be equal, at right and at left:

$$\lambda_k c_1 = c_1 \lambda_1$$

$$\vdots$$

$$\lambda_k c_{k-1} = c_{k-1} \lambda_{k-1}$$

Now since at least one c_i must be nonzero, from $\lambda_k c_i = c_i \lambda_i$ we obtain $\lambda_k = \lambda_i$, which is a contradiction. Thus our proof by recurrence of the first assertion is complete.

As for the second assertion, this follows from the first one. $\square$

In order to reach now to more advanced results, we can use the characteristic polynomial. Here is a result summarizing and improving our knowledge of the subject:

THEOREM 4.9. *Given a matrix $A \in M_N(\mathbb{C})$, consider its characteristic polynomial:*

$$P(x) = \det(A - x1_N)$$

The eigenvalues of A are then the roots of P. Also, we have the inequality

$$\dim(E_\lambda) \leq m_\lambda$$

where m_λ is the multiplicity of λ, as root of P.

PROOF. The first assertion follows from the following computation, using the fact that a linear map is bijective when the determinant of the associated matrix is nonzero:

$$\begin{aligned} \exists v, Av = \lambda v &\iff \exists v, (A - \lambda 1_N)v = 0 \\ &\iff \det(A - \lambda 1_N) = 0 \end{aligned}$$

Regarding now the second assertion, given an eigenvalue λ of our matrix A, consider the dimension of the corresponding eigenspace:

$$d_\lambda = \dim(E_\lambda)$$

By changing the basis of $\mathbb{C}^N$, as for the eigenspace E_λ to be spanned by the first d_λ basis elements, our matrix becomes as follows, with B being a certain smaller matrix:

$$A \sim \begin{pmatrix} \lambda 1_{d_\lambda} & 0 \\ 0 & B \end{pmatrix}$$

We conclude that the characteristic polynomial of A is of the following form:

$$\begin{aligned} P_A &= P_{\lambda 1_{d_\lambda}} P_B \\ &= (\lambda - x)^{d_\lambda} P_B \end{aligned}$$

Thus we have $m_\lambda \geq d_\lambda$, which leads to the conclusion in the statement. $\square$

We can put together Theorem 4.8 and Theorem 4.9, and by using as well the fact that any complex polynomial of degree N has exactly N complex roots, when counted with multiplicities, that we know from chapter 3 above, we obtain the folowing result:

THEOREM 4.10. *Given a matrix $A \in M_N(\mathbb{C})$, consider its characteristic polynomial*

$$P(X) = \det(A - X1_N)$$

then factorize this polynomial, by computing the complex roots, with multiplicities,

$$P(X) = (-1)^N (X - \lambda_1)^{n_1} \ldots (X - \lambda_k)^{n_k}$$

and finally compute the corresponding eigenspaces, for each eigenvalue found:

$$E_i = \left\{ v \in \mathbb{C}^N \Big| Av = \lambda_i v \right\}$$

The dimensions of these eigenspaces satisfy then the following inequalities,

$$\dim(E_i) \leq n_i$$

and A is diagonalizable precisely when we have equality for any i.

PROOF. This follows indeed from Theorem 4.8 and Theorem 4.9, and from the above-mentioned result regarding the complex roots of complex polynomials. □

This was for the main result of linear algebra. There are countless applications of this, and generally speaking, advanced linear algebra consists in building on Theorem 4.10. Let us record as well a useful algorithmic version of the above result:

THEOREM 4.11. *The square matrices $A \in M_N(\mathbb{C})$ can be diagonalized as follows:*

(1) *Compute the characteristic polynomial.*
(2) *Factorize the characteristic polynomial.*
(3) *Compute the eigenvectors, for each eigenvalue found.*
(4) *If there are no N eigenvectors, A is not diagonalizable.*
(5) *Otherwise, A is diagonalizable, $A = PDP^{-1}$.*

PROOF. This is an informal reformulation of Theorem 4.10 above, with (4) referring to the total number of linearly independent eigenvectors found in (3), and with $A = PDP^{-1}$ in (5) being the usual diagonalization formula, with P, D being as before. □

As a remark here, in step (3) it is always better to start with the eigenvalues having big multiplicity. Indeed, a multiplicity 1 eigenvalue, for instance, can never lead to the end of the computation, via (4), simply because the eigenvectors always exist.

4c. Matrix tricks

At the level of basic examples of diagonalizable matrices, we first have:

THEOREM 4.12. *For a matrix $A \in M_N(\mathbb{C})$ the following conditions are equivalent,*

1. *The eigenvalues are different, $\lambda_i \neq \lambda_j$,*
2. *The characteristic polynomial P has simple roots,*
3. *The characteristic polynomial satisfies $(P, P') = 1$,*

and in this case, the matrix is diagonalizable.

PROOF. The equivalences in the statement are clear, the idea being as follows:

(1) $\iff$ (2) This follows from Theorem 4.10.

(2) $\iff$ (3) This is standard, the double roots of P being roots of P'.

Regarding now the last assertion, this follows from Theorem 4.10 as well. Indeed, the criterion there for diagonalization involving the sum of dimensions of eigenspaces is trivially satisfied when the eigenvalues are different, because:

$$\sum_\lambda \dim(E_\lambda) = \sum_\lambda 1 = N$$

Thus, we are led to the conclusion in the statement. $\square$

As an important comment, the assumptions of Theorem 4.12 can be effectively verified in practice, without the need for factorizing polynomials, the idea here being that of using the condition (3) there. In order to discuss this, let us start with:

THEOREM 4.13. *Given two polynomials $P, Q \in \mathbb{C}[X]$, written as follows,*

$$P = c(X - a_1) \ldots (X - a_k)$$

$$Q = d(X - b_1) \ldots (X - b_l)$$

the following quantity, which is called resultant of P, Q,

$$R(P, Q) = c^l d^k \prod_{ij} (a_i - b_j)$$

is a polynomial in the coefficients of P, Q, with integer coefficients, and we have

$$R(P, Q) = 0$$

precisely when P, Q have a common root.

PROOF. Given $P, Q \in \mathbb{C}[X]$, we can certainly construct the quantity $R(P, Q)$ in the statement, and we have then $R(P, Q) = 0$ precisely when P, Q have a common root.

The whole point is that of proving that $R(P, Q)$ is a polynomial in the coefficients of P, Q, with integer coefficients. But this can be checked as follows:

(1) We can expand the formula of $R(P,Q)$, and in what regards $a_1,\ldots,a_k$, which are the roots of P, we obtain in this way certain symmetric functions in these variables, which will be therefore polynomials in the coefficients of P, with integer coefficients.

(2) We can then look what happens with respect to the remaining variables $b_1,\ldots,b_l$, which are the roots of Q. Once again what we have here are certain symmetric functions, and so polynomials in the coefficients of Q, with integer coefficients.

(3) Thus, we are led to the conclusion in the statement, that $R(P,Q)$ is a polynomial in the coefficients of P,Q, with integer coefficients, and with the remark that the c^ld^k factor is there for these latter coefficients to be indeed integers, instead of rationals. $\square$

All this might seem a bit complicated, so as an illustration, let us work out an example. Consider the case of a polynomial of degree 2, and a polynomial of degree 1:

$$P = ax^2 + bx + c$$

$$Q = dx + e$$

In order to compute the resultant, let us factorize our polynomials:

$$P = a(x-p)(x-q)$$

$$Q = d(x-r)$$

The resultant can be then computed as follows, by using the method above:

$$\begin{aligned}
R(P,Q) &= ad^2(p-r)(q-r)\\
&= ad^2(pq-(p+q)r+r^2)\\
&= cd^2+bd^2r+ad^2r^2\\
&= cd^2-bde+ae^2
\end{aligned}$$

Finally, observe that $R(P,Q)=0$ corresponds indeed to the fact that P,Q have a common root. Indeed, the root of Q is $r=-e/d$, and we have:

$$\begin{aligned}
P(r) &= \frac{ae^2}{d^2}-\frac{be}{d}+c\\
&= \frac{R(P,Q)}{d^2}
\end{aligned}$$

Thus $P(r)=0$ precisely when $R(P,Q)=0$, as predicted by Theorem 4.13.

Regarding now the explicit formula of the resultant $R(P,Q)$, this is something quite complicated, and there are several methods for dealing with this problem. There is a slight similarity between Theorem 4.13 and the Vandermonde determinants discussed in chapter 2 above, and we have in fact the following formula for $R(P,Q)$:

Theorem 4.14. *The resultant of two polynomials, written as*

$$P = p_k X^k + \ldots + p_1 X + p_0$$

$$Q = q_l X^l + \ldots + q_1 X + q_0$$

appears as the determinant of an associated matrix, as follows,

$$R(P,Q) = \begin{vmatrix} p_k & & & q_l & & \\ \vdots & \ddots & & \vdots & \ddots & \\ p_0 & & p_k & q_0 & & q_k \\ & \ddots & \vdots & & \ddots & \vdots \\ & & p_0 & & & q_0 \end{vmatrix}$$

with the matrix having size $k+l$, and having 0 coefficients at the blank spaces.

Proof. This follows by doing some linear algebra computations, mixed with algebra, in the spirit of those from the proof of the Vandermonde determinant theorem. □

In what follows we will not really need the above formula, so let us just check now that this formula works indeed. Consider our favorite polynomials, as before:

$$P = ax^2 + bx + c$$

$$Q = dx + e$$

According to the above result, the resultant should be then, as it should:

$$\begin{aligned} R(P,Q) &= \begin{vmatrix} a & d & 0 \\ b & e & d \\ c & 0 & e \end{vmatrix} \\ &= ae^2 - bde + cd^2 \end{aligned}$$

Now back to our diagonalization questions, we want to compute $R(P,P')$, where P is the characteristic polynomial. So, we need one more piece of theory, as follows:

Theorem 4.15. *Given a polynomial $P \in \mathbb{C}[X]$, written as*

$$P(X) = cX^N + dX^{N-1} + \ldots$$

its discriminant, defined as being the following quantity,

$$\Delta(P) = \frac{(-1)^{\binom{N}{2}}}{c} R(P,P')$$

is a polynomial in the coefficients of P, with integer coefficients, and

$$\Delta(P) = 0$$

happens precisely when P has a double root.

PROOF. This follows from Theorem 4.13, applied with $P = Q$, with the division by c being indeed possible, under $\mathbb{Z}$, and with the sign being there for various reasons, including the compatibility with some well-known formulae, at small values of $N \in \mathbb{N}$. $\square$

Observe that, by using Theorem 4.14, we have an explicit formula for the discriminant, as the determinant of a certain associated matrix. There is a lot of theory here, and in order to terminate the discussion, let us see what happens in degree 2. Here we have:

$$P = aX^2 + bX + c$$

$$P' = 2aX + b$$

Thus, the resultant is given by the following formula:

$$\begin{aligned} R(P, P') &= ab^2 - b(2a)b + c(2a)^2 \\ &= 4a^2c - ab^2 \\ &= -a(b^2 - 4ac) \end{aligned}$$

With the normalizations in Theorem 4.15 made, we obtain, as we should:

$$\Delta(P) = b^2 - 4ac$$

Now back to our questions, we can upgrade Theorem 4.12, as follows:

THEOREM 4.16. *For a matrix $A \in M_N(\mathbb{C})$ the following conditions are equivalent,*

1. *The eigenvalues are different, $\lambda_i \neq \lambda_j$,*
2. *The characteristic polynomial P has simple roots,*
3. *The discriminant of P is nonzero, $\Delta(P) \neq 0$,*

and in this case, the matrix is diagonalizable.

PROOF. This is indeed an upgrade of Theorem 4.12 above, by replacing the condition (3) there with the condition $\Delta(P) \neq 0$, which is something much better, because it is something algorithmic, ultimately requiring only computations of determinants. $\square$

As already mentioned before, one can prove that the matrices having distinct eigenvalues are "generic", and so the above result basically captures the whole situation. We have in fact the following collection of density results, all being very useful:

THEOREM 4.17. *The following happen, inside $M_N(\mathbb{C})$:*

1. *The invertible matrices are dense.*
2. *The matrices having distinct eigenvalues are dense.*
3. *The diagonalizable matrices are dense.*

PROOF. These are quite advanced linear algebra results, which can be proved as follows, with the technology that we have so far:

(1) This is clear, intuitively speaking, because the invertible matrices are given by the condition $\det A \neq 0$. Thus, the set formed by these matrices appears as the complement of the surface $\det A = 0$, and so must be dense inside $M_N(\mathbb{C})$, as claimed.

(2) Here we can use a similar argument, this time by saying that the set formed by the matrices having distinct eigenvalues appears as the complement of the surface given by $\Delta(P_A) = 0$, and so must be dense inside $M_N(\mathbb{C})$, as claimed.

(3) This follows from (2), via the fact that the matrices having distinct eigenvalues are diagonalizable, that we know from Theorem 4.16 above. There are of course some other proofs as well, for instance by putting the matrix in Jordan form. $\square$

As an application of the above results, and of our methods in general, we can now establish a number of useful and interesting linear algebra results, as follows:

THEOREM 4.18. *The following happen:*
1. *We have $P_{AB} = P_{BA}$, for any two matrices $A, B \in M_N(\mathbb{C})$.*
2. *AB, BA have the same eigenvalues, with the same multiplicities.*
3. *If A has eigenvalues $\lambda_1, \ldots, \lambda_N$, then $f(A)$ has eigenvalues $f(\lambda_1), \ldots, f(\lambda_N)$.*

PROOF. These results can be deduced by using Theorem 4.17, as follows:

(1) It follows from definitions that the characteristic polynomial of a matrix is invariant under conjugation, in the sense that we have:

$$P_C = P_{ACA^{-1}}$$

Now observe that, when assuming that A is invertible, we have:

$$AB = A(BA)A^{-1}$$

Thus, we have the result when A is invertible. By using now Theorem 4.17 (1) above, we conclude that this formula holds for any matrix A, by continuity.

(2) This is a reformulation of (1) above, via the fact that P encodes the eigenvalues, with multiplicities, which is hard to prove with bare hands.

(3) This is something more informal, the idea being that this is clear for the diagonal matrices D, then for the diagonalizable matrices PDP^{-1}, and finally for all the matrices, by using Theorem 4.17 (3), provided that f has suitable regularity properties. $\square$

We will be back to the above results, with more details, at the end of the present chapter, and then in chapter 7 below, when doing spectral theory.

Getting away now from all this advanced mathematics, let us go back to the main problem raised by the diagonalization procedure, namely the computation of the roots of characteristic polynomials. As a first trivial observation, in degree 2 we have:

PROPOSITION 4.19. *The roots of a degree* 2 *polynomial of the form*

$$P = X^2 - aX + b$$

are precisely the numbers r, s *satisfying* $r + s = a$, $rs = b$.

PROOF. This is trivial, coming from $P = (X - r)(X - s)$. In practice, this is very useful, and this is how calculus teachers are often faster than their students. $\square$

In the matrix setting now, the result is as follows:

PROPOSITION 4.20. *The eigenvalues of a square matrix*

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

are precisely the numbers r, s *satisfying* $r + s = a + d$, $rs = ad - bc$.

PROOF. Once again, this is something trivial, coming from Proposition 4.19. As before, this is a well-known calculus teacher trick. $\square$

Finally, we have the following result, which is useful as well:

PROPOSITION 4.21. *Assume that we have a polynomial as follows, with integer coefficients, and with the leading term being* 1*:*

$$P = X^N + a_{N-1}X^{N-1} + \ldots + a_1X + a_0$$

The integer roots of P *must then divide the last coefficient* a_0.

PROOF. This is clear, because any integer root $c \in \mathbb{Z}$ of our polynomial must satisfy:

$$c^N + a_{N-1}c^{N-1} + \ldots + a_1c + a_0 = 0$$

But modulo c, this equation simply reads $a_0 = 0$, as desired. $\square$

Getting now into more conceptual mathematics, the two formulae from Proposition 4.20 above are quite interesting, and can be generalized as follows:

THEOREM 4.22. *The complex eigenvalues of a matrix* $A \in M_N(\mathbb{C})$*, counted with multiplicities, have the following properties:*

(1) *Their sum is the trace.*
(2) *Their product is the determinant.*

PROOF. Consider indeed the characteristic polynomial P of the matrix:

$$\begin{aligned} P(X) &= \det(A - X1_N) \\ &= (-1)^N X^N + (-1)^{N-1} Tr(A) X^{N-1} + \ldots + \det(A) \end{aligned}$$

We can factorize this polynomial, by using its N complex roots, and we obtain:

$$\begin{aligned} P(X) &= (-1)^N(X-\lambda_1)\ldots(X-\lambda_N) \\ &= (-1)^N X^N + (-1)^{N-1}\left(\sum_i \lambda_i\right) X^{N-1} + \ldots + \prod_i \lambda_i \end{aligned}$$

Thus, we are led to the conclusion in the statement. $\square$

Regarding now the intermediate terms, we have here:

THEOREM 4.23. *Assume that $A \in M_N(\mathbb{C})$ has eigenvalues $\lambda_1, \ldots, \lambda_N \in \mathbb{C}$, counted with multiplicities. The basic symmetric functions of these eigenvalues, namely*

$$c_k = \sum_{i_1 < \ldots < i_k} \lambda_{i_1} \ldots \lambda_{i_k}$$

are then given by the fact that the characteristic polynomial of the matrix is:

$$P(X) = (-1)^N \sum_{k=0}^N (-1)^k c_k X^k$$

Moreover, all symmetric functions of the eigenvalues, such as the sums of powers

$$d_s = \lambda_1^s + \ldots + \lambda_N^s$$

appear as polynomials in these characteristic polynomial coefficients c_k.

PROOF. These results can be proved by doing some algebra, as follows:

(1) Consider indeed the characteristic polynomial P of the matrix, factorized by using its N complex roots, taken with multiplicities. By expanding, we obtain:

$$\begin{aligned} P(X) &= (-1)^N(X-\lambda_1)\ldots(X-\lambda_N) \\ &= (-1)^N X^N + (-1)^{N-1}\left(\sum_i \lambda_i\right) X^{N-1} + \ldots + \prod_i \lambda_i \\ &= (-1)^N X^N + (-1)^{N-1} c_1 X^{N-1} + \ldots + (-1)^0 c_N \\ &= (-1)^N \left(X^N - c_1 X^{N-1} + \ldots + (-1)^N c_N\right) \end{aligned}$$

With the convention $c_0 = 1$, we are led to the conclusion in the statement.

(2) This is something standard, coming by doing some abstract algebra. Working out the formulae for the sums of powers $d_s = \sum_i \lambda_i^s$, at small values of the exponent $s \in \mathbb{N}$, is an excellent exercise, which shows how to proceed in general, by recurrence. $\square$

4d. Spectral theorems

Let us go back now to the diagonalization question. Here is a key result:

THEOREM 4.24. *Any matrix $A\in M_N(\mathbb C)$ which is self-adjoint, $A=A^*$, is diagonalizable, with the diagonalization being of the following type,*

$$A=UDU^*$$

with $U\in U_N$, and with $D\in M_N(\mathbb R)$ diagonal. The converse holds too.

PROOF. As a first remark, the converse trivially holds, because if we take a matrix of the form $A=UDU^*$, with U unitary and D diagonal and real, then we have:

$$\begin{aligned}A^* &= (UDU^*)^*\\ &= UD^*U^*\\ &= UDU^*\\ &= A\end{aligned}$$

In the other sense now, assume that A is self-adjoint, $A=A^*$. Our first claim is that the eigenvalues are real. Indeed, assuming $Av=\lambda v$, we have:

$$\begin{aligned}\lambda<v,v> &= <\lambda v,v>\\ &= <Av,v>\\ &= <v,Av>\\ &= <v,\lambda v>\\ &= \bar\lambda<v,v>\end{aligned}$$

Thus we obtain $\lambda\in\mathbb R$, as claimed. Our next claim now is that the eigenspaces corresponding to different eigenvalues are pairwise orthogonal. Assume indeed that:

$$Av=\lambda v\quad,\quad Aw=\mu w$$

We have then the following computation, using $\lambda,\mu\in\mathbb R$:

$$\begin{aligned}\lambda<v,w> &= <\lambda v,w>\\ &= <Av,w>\\ &= <v,Aw>\\ &= <v,\mu w>\\ &= \mu<v,w>\end{aligned}$$

Thus $\lambda\neq\mu$ implies $v\perp w$, as claimed. In order now to finish, it remains to prove that the eigenspaces span $\mathbb C^N$. For this purpose, we will use a recurrence method. Let us pick an eigenvector, $Av=\lambda v$. Assuming $v\perp w$, we have:

$$\begin{aligned}<Aw,v> &= <w,Av>\\ &= <w,\lambda v>\\ &= \lambda<w,v>\\ &= 0\end{aligned}$$

Thus, if v is an eigenvector, then the vector space $v^\perp$ is invariant under A. In order to do the recurrence, it still remains to prove that the restriction of A to the vector space $v^\perp$ is self-adjoint. But this comes from a general property of the self-adjoint matrices, that we will explain now. Our claim is that an arbitary square matrix A is self-adjoint precisely when the following happens, for any vector v:

$$<Av,v>\in\mathbb{R}$$

Indeed, the fact that the above scalar product is real is equivalent to:

$$<(A-A^*)v,v>=0$$

But this is equivalent to $A=A^*$, by using the complex polarization identity. Now back to our questions, it is clear from our self-adjointness criterion above that the restriction of A to any invariant subspace, and in particular to the subspace $v^\perp$, is self-adjoint. Thus, we can proceed by recurrence, and we obtain the result. □

Let us record as well the real version of the above result:

Theorem 4.25. *Any matrix $A\in M_N(\mathbb{R})$ which is symmetric, in the sense that*

$$A=A^t$$

is diagonalizable, with the diagonalization being of the following type,

$$A=UDU^t$$

with $U\in O_N$, and with $D\in M_N(\mathbb{R})$ diagonal. The converse holds too.

Proof. As before, the converse trivially holds, because if we take a matrix of the form $A=UDU^t$, with U orthogonal and D diagonal and real, then we have $A^t=A$. In the other sense now, this follows from Theorem 4.24, and its proof. □

As basic examples of self-adjoint matrices, we have the orthogonal projections. The diagonalization result regarding them is as follows:

Proposition 4.26. *The matrices $P\in M_N(\mathbb{C})$ which are projections, $P^2=P=P^*$, are precisely those which diagonalize as follows,*

$$P=UDU^*$$

with $U\in U_N$, and with $D\in M_N(0,1)$ being diagonal.

PROOF. We know that the equation for the projections is as follows:

$$P^2 = P = P^*$$

Thus the eigenvalues are real, and then, by using $P^2 = P$, we obtain:

$$\begin{aligned}
\lambda < v, v > &= < \lambda v, v > \\
&= < Pv, v > \\
&= < P^2 v, v > \\
&= < Pv, Pv > \\
&= < \lambda v, \lambda v > \\
&= \lambda^2 < v, v >
\end{aligned}$$

We therefore obtain $\lambda \in \{0, 1\}$, as claimed, and as a final conclusion here, the diagonalization of the self-adjoint matrices is as follows, with $e_i \in \{0, 1\}$:

$$P \sim \begin{pmatrix} e_1 & & \\ & \ddots & \\ & & e_N \end{pmatrix}$$

To be more precise, the number of 1 values is the dimension of the image of P, and the number of 0 values is the dimension of space of vectors sent to 0 by P. $\square$

In the real case, the result regarding the projections is as follows:

PROPOSITION 4.27. *The matrices $P \in M_N(\mathbb{R})$ which are projections,*

$$P^2 = P = P^t$$

are precisely those which diagonalize as follows,

$$P = UDU^t$$

with $U \in O_N$, and with $D \in M_N(0, 1)$ being diagonal.

PROOF. This follows indeed from Proposition 4.26, and its proof. $\square$

An important class of self-adjoint matrices, that we will discuss now, which includes all the projections, are the positive matrices. The general theory here is as follows:

THEOREM 4.28. *For a matrix $A \in M_N(\mathbb{C})$ the following conditions are equivalent, and if they are satisfied, we say that A is positive:*

(1) *$A = B^2$, with $B = B^*$.*
(2) *$A = CC^*$, for some $C \in M_N(\mathbb{C})$.*
(3) *$< Ax, x > \geq 0$, for any vector $x \in \mathbb{C}^N$.*
(4) *$A = A^*$, and the eigenvalues are positive, $\lambda_i \geq 0$.*
(5) *$A = UDU^*$, with $U \in U_N$ and with $D \in M_N(\mathbb{R}_+)$ diagonal.*

PROOF. The idea is that the equivalences in the statement basically follow from some elementary computations, with only Theorem 4.24 needed, at some point:

(1) $\implies$ (2) This is clear, because we can take $C = B$.

(2) $\implies$ (3) This follows from the following computation:

$$\begin{aligned} < Ax, x > &= < CC^*x, x > \\ &= < C^*x, C^*x > \\ &\geq 0 \end{aligned}$$

(3) $\implies$ (4) By using the fact that $< Ax, x >$ is real, we have:

$$\begin{aligned} < Ax, x > &= < x, A^*x > \\ &= < A^*x, x > \end{aligned}$$

Thus we have $A = A^*$, and the remaining assertion, regarding the eigenvalues, follows from the following computation, assuming $Ax = \lambda x$:

$$\begin{aligned} < Ax, x > &= < \lambda x, x > \\ &= \lambda < x, x > \\ &\geq 0 \end{aligned}$$

(4) $\implies$ (5) This follows by using Theorem 4.24 above.

(5) $\implies$ (1) Assuming $A = UDU^*$ as in the statement, we can set $B = U\sqrt{D}U^*$. Then this matrix B is self-adjoint, and its square is given by:

$$\begin{aligned} B^2 &= U\sqrt{D}U^* \cdot U\sqrt{D}U^* \\ &= UDU^* \\ &= A \end{aligned}$$

Thus, we are led to the conclusion in the statement. $\square$

Let us record as well the following technical version of the above result:

THEOREM 4.29. *For a matrix $A \in M_N(\mathbb{C})$ the following conditions are equivalent, and if they are satisfied, we say that A is strictly positive:*

1. *$A = B^2$, with $B = B^*$, invertible.*
2. *$A = CC^*$, for some $C \in M_N(\mathbb{C})$ invertible.*
3. *$< Ax, x >> 0$, for any nonzero vector $x \in \mathbb{C}^N$.*
4. *$A = A^*$, and the eigenvalues are strictly positive, $\lambda_i > 0$.*
5. *$A = UDU^*$, with $U \in U_N$ and with $D \in M_N(\mathbb{R}_+^*)$ diagonal.*

PROOF. This follows either from Theorem 4.28, by adding the various extra assumptions in the statement, or from the proof of Theorem 4.28, by modifying where needed. $\square$

The positive matrices are quite important, for a number of reasons. On one hand, these are the matrices $A\in M_N(\mathbb C)$ having a square root $\sqrt{A}\in M_N(\mathbb C)$, as shown by our positivity condition (1). On the other hand, any matrix $A\in M_N(\mathbb C)$ produces the positive matrix $A^*A\in M_N(\mathbb C)$, as shown by our positivity condition (2). We can combine these two observations, and we are led to the following construction, for any $A\in M_N(\mathbb C)$:

$$A\to\sqrt{A^*A}$$

This is something quite interesting, because at $N=1$ what we have here is the construction of the absolute value of the complex numbers, $|z|=\sqrt{z\bar{z}}$. This suggests using the notation $|A|=\sqrt{A^*A}$, and then looking for a decomposition result of type:

$$A=U|A|$$

We will be back to this type of decomposition later, called polar decomposition, at the end of the present chapter, after developing some more general theory.

Let us discuss now the case of the unitary matrices. We have here:

THEOREM 4.30. *Any matrix $U\in M_N(\mathbb C)$ which is unitary, $U^*=U^{-1}$, is diagonalizable, with the eigenvalues on $\mathbb T$. More precisely we have*

$$U=VDV^*$$

with $V\in U_N$, and with $D\in M_N(\mathbb T)$ diagonal. The converse holds too.

PROOF. As a first remark, the converse trivially holds, because given a matrix of type $U=VDV^*$, with $V\in U_N$, and with $D\in M_N(\mathbb T)$ being diagonal, we have:

$$\begin{aligned}U^*&=(VDV^*)^*\\&=VD^*V^*\\&=VD^{-1}V^{-1}\\&=(V^*)^{-1}D^{-1}V^{-1}\\&=(VDV^*)^{-1}\\&=U^{-1}\end{aligned}$$

Let us prove now the first assertion, stating that the eigenvalues of a unitary matrix $U\in U_N$ belong to $\mathbb T$. Indeed, assuming $Uv=\lambda v$, we have:

$$\begin{aligned}<v,v>&=<U^*Uv,v>\\&=<Uv,Uv>\\&=<\lambda v,\lambda v>\\&=|\lambda|^2<v,v>\end{aligned}$$

Thus we obtain $\lambda \in \mathbb{T}$, as claimed. Our next claim now is that the eigenspaces corresponding to different eigenvalues are pairwise orthogonal. Assume indeed that:

$$Uv = \lambda v \quad , \quad Uw = \mu w$$

We have then the following computation, using $U^* = U^{-1}$ and $\lambda, \mu \in \mathbb{T}$:

$$\begin{aligned} \lambda < v, w > &= < \lambda v, w > \\ &= < Uv, w > \\ &= < v, U^* w > \\ &= < v, U^{-1} w > \\ &= < v, \mu^{-1} w > \\ &= \mu < v, w > \end{aligned}$$

Thus $\lambda \neq \mu$ implies $v \perp w$, as claimed. In order now to finish, it remains to prove that the eigenspaces span $\mathbb{C}^N$. For this purpose, we will use a recurrence method. Let us pick an eigenvector, $Uv = \lambda v$. Assuming $v \perp w$, we have:

$$\begin{aligned} < Uw, v > &= < w, U^* v > \\ &= < w, U^{-1} v > \\ &= < w, \lambda^{-1} v > \\ &= \lambda < w, v > \\ &= 0 \end{aligned}$$

Thus, if v is an eigenvector, then the vector space $v^\perp$ is invariant under U. Now since U is an isometry, so is its restriction to this space $v^\perp$. Thus this restriction is a unitary, and so we can proceed by recurrence, and we obtain the result. $\square$

Let us record as well the real version of the above result, in a weak form:

Theorem 4.31. *Any matrix $U \in M_N(\mathbb{R})$ which is orthogonal, $U^t = U^{-1}$, is diagonalizable, with the eigenvalues on $\mathbb{T}$. More precisely we have*

$$U = VDV^*$$

with $V \in U_N$, and with $D \in M_N(\mathbb{T})$ being diagonal.

Proof. This follows indeed from Theorem 4.30 above. $\square$

Observe that the above result does not provide us with a complete characterization of the matrices $U \in M_N(\mathbb{R})$ which are orthogonal. To be more precise, the question left is that of understanding when the matrices of type $U = VDV^*$, with $V \in U_N$, and with $D \in M_N(\mathbb{T})$ being diagonal, are real, and this is something non-trivial.

Back to generalities, the self-adjoint matrices and the unitary matrices are particular cases of the general notion of a “normal matrix”, and we have here:

THEOREM 4.32. *Any matrix $A \in M_N(\mathbb{C})$ which is normal, $AA^* = A^*A$, is diagonalizable, with the diagonalization being of the following type,*

$$A = UDU^*$$

with $U \in U_N$, and with $D \in M_N(\mathbb{C})$ diagonal. The converse holds too.

PROOF. As a first remark, the converse trivially holds, because if we take a matrix of the form $A = UDU^*$, with U unitary and D diagonal, then we have:

$$\begin{aligned} AA^* &= UDU^* \cdot UD^*U^* \\ &= UDD^*U^* \\ &= UD^*DU^* \\ &= UD^*U^* \cdot UDU^* \\ &= A^*A \end{aligned}$$

In the other sense now, this is something more technical. Our first claim is that a matrix A is normal precisely when the following happens, for any vector v:

$$||Av|| = ||A^*v||$$

Indeed, the above equality can be written as follows:

$$< AA^*v, v >=< A^*Av, v >$$

But this is equivalent to $AA^* = A^*A$, by using the polarization identity. Our claim now is that A, A^* have the same eigenvectors, with conjugate eigenvalues:

$$Av = \lambda v \implies A^*v = \bar{\lambda}v$$

Indeed, this follows from the following computation, and from the trivial fact that if A is normal, then so is any matrix of type $A - \lambda 1_N$:

$$\begin{aligned} ||(A^* - \bar{\lambda}1_N)v|| &= ||(A - \lambda 1_N)^*v|| \\ &= ||(A - \lambda 1_N)v|| \\ &= 0 \end{aligned}$$

Let us prove now, by using this, that the eigenspaces of A are pairwise orthogonal. Assuming $Av = \lambda v$ and $Aw = \mu w$ with $\lambda \neq \mu$, we have:

$$\begin{aligned} \lambda < v, w > &= < \lambda v, w > \\ &= < Av, w > \\ &= < v, A^*w > \\ &= < v, \bar{\mu}w > \\ &= \mu < v, w > \end{aligned}$$

Thus $\lambda \neq \mu$ implies $v \perp w$, as claimed. In order to finish now the proof, it remains to prove that the eigenspaces of A span the whole $\mathbb{C}^N$. This is something that we have already seen for the self-adjoint matrices, and for the unitaries, and we will use here these results, in order to deal with the general normal case. As a first observation, given an arbitrary matrix A, the matrix AA^* is self-adjoint:

$$(AA^*)^* = AA^*$$

Thus, we can diagonalize this matrix AA^*, as follows, with the passage matrix being a unitary, $V \in U_N$, and with the diagonal form being real, $E \in M_N(\mathbb{R})$:

$$AA^* = VEV^*$$

Now observe that, for matrices of type $A = UDU^*$, which are those that we supposed to deal with, we have $V = U, E = D\bar{D}$. In particular, A and AA^* have the same eigenspaces. So, this will be our idea, proving that the eigenspaces of AA^* are eigenspaces of A. In order to do so, let us pick two eigenvectors v, w of the matrix AA^*, corresponding to different eigenvalues, $\lambda \neq \mu$. The eigenvalue equations are then as follows:

$$AA^*v = \lambda v \quad , \quad AA^*w = \mu w$$

We have the following computation, using the normality condition $AA^* = A^*A$, and the fact that the eigenvalues of AA^*, and in particular μ, are real:

$$\begin{aligned}
\lambda < Av, w > &= < \lambda Av, w > \\
&= < A\lambda v, w > \\
&= < AAA^*v, w > \\
&= < AA^*Av, w > \\
&= < Av, AA^*w > \\
&= < Av, \mu w > \\
&= \mu < Av, w >
\end{aligned}$$

We conclude that we have $< Av, w >= 0$. But this reformulates as follows:

$$\lambda \neq \mu \implies A(E_\lambda) \perp E_\mu$$

Now since the eigenspaces of AA^* are pairwise orthogonal, and span the whole $\mathbb{C}^N$, we deduce from this that these eigenspaces are invariant under A:

$$A(E_\lambda) \subset E_\lambda$$

But with this result in hand, we can finish. Indeed, we can decompose the problem, and the matrix A itself, following these eigenspaces of AA^*, which in practice amounts in saying that we can assume that we only have 1 eigenspace. By rescaling, this is the same as assuming that we have $AA^* = 1$, and so we are now into the unitary case, that we know how to solve, as explained in Theorem 4.30 above. $\square$

Let us discuss now, as a final topic, one more important result, namely the polar decomposition. The idea will be that of writing a formula as follows:

$$A = U|A|$$

To be more precise, U will be here a partial isometry, a generalization of the notion of isometry, and the matrix $|A|$ will be a kind of absolute value of A.

In order to discuss this, let us first discuss the absolute values. We have here:

THEOREM 4.33. *Given a matrix $A \in M_N(\mathbb{C})$, we can construct a matrix $|A|$ as follows, by using the fact that A^*A is diagonalizable, with positive eigenvalues:*

$$|A| = \sqrt{A^*A}$$

This matrix $|A|$ is then positive, and its square is $|A|^2 = A$. In the case $N = 1$, we obtain in this way the usual absolute value of the complex numbers.

PROOF. Consider indeed the matrix A^*A, which is normal. According to Theorem 4.32, we can diagonalize this matrix as follows, with $U \in U_N$, and with D diagonal:

$$A = UDU^*$$

Since we have $A^*A \geq 0$, it follows that we have $D \geq 0$, which means that the entries of D are real, and positive. Thus we can extract the square root $\sqrt{D}$, and then set:

$$\sqrt{A^*A} = U\sqrt{D}U^*$$

Now if we call this latter matrix $|A|$, we are led to the conclusions in the statement, namely $|A| \geq 0$, and $|A|^2 = A$. Finally, the last assertion is clear from definitions. $\square$

We can now formulate a first polar decomposition result, as follows:

THEOREM 4.34. *Any invertible matrix $A \in M_N(\mathbb{C})$ decomposes as*

$$A = U|A|$$

with $U \in U_N$, and with $|A| = \sqrt{A^*A}$ as above.

PROOF. This is routine, and follows by comparing the actions of $A, |A|$ on the vectors $v \in \mathbb{C}^N$, and deducing from this the existence of a unitary $U \in U_N$ as above. $\square$

Observe that at $N = 1$ we obtain in this way the usual polar decomposition of the nonzero complex numbers. There are of course many other examples.

More generally now, we have the following result:

THEOREM 4.35. *Any square matrix $A \in M_N(\mathbb{C})$ decomposes as*

$$A = U|A|$$

with U being a partial isometry, and with $|A| = \sqrt{A^*A}$ as above.

PROOF. Once again, this follows by comparing the actions of $A, |A|$ on the vectors $v \in \mathbb{C}^N$, and deducing from this the existence of a partial isometry U as above. Alternatively, we can get this from Theorem 4.34, applied on the complement of the 0-eigenvectors. $\square$

We will be back to this in chapter 8 below, when doing spectral theory.

Good news, we are done with linear algebra. We have learned many things in the past 100 pages, and our knowledge of the subject is quite decent, and we will stop here. In the remainder of the present book we will be rather looking into applications.

This being said, there are many possible continuations of what we learned. As a first piece of advice, the more linear algebra that you know, the better your mathematics will be, so try reading some more. A good linear algebra book, written by an analyst, is the one by Lax [**63**]. Another good linear algebra book, or rather book about algebra at large, written this time by an algebraist, is the one by Lang [**62**]. A more advanced book, which is more than enough for most daily tasks, is the one by Horn and Johnson [**47**]. So, keep an eye on these books, and have them ready for a quick read, when needed.

4e. Exercises

Things have been quite dense in this chapter, which was our last one on basic linear algebra, and the main purpose of the exercises below will be that of filling some gaps in the above material, and also discussing some important things which must be known as well. As a first exercise, in relation with linear spaces, bases and dimensions, we have:

EXERCISE 4.36. *Clarify the theory of bases and dimensions for the linear subspaces $V \subset \mathbb{R}^N$, notably by establishing the formula*

$$\dim(\ker f) + \dim(Imf) = N$$

valid for any linear map $f : \mathbb{C}^N \to \mathbb{C}^N$, and then extend this into a theory of abstract linear spaces V, which are not necessarily subspaces of $\mathbb{C}^N$.

Here the first question is something quite standard, by using the material that we already have. As for the second question, things are a bit more tricky here, because once the abstract linear spaces V being defined, the only available tool is recurrence.

EXERCISE 4.37. *Work out what happens to the main diagonalization theorem for the matrices $A \in M_N(\mathbb{C})$, in the cases $A \in M_2(\mathbb{C})$, $A \in M_N(\mathbb{R})$, and $A \in M_2(\mathbb{R})$.*

As before, this is a rather theoretical exercise, the point being that of carefully reviewing all the material above, in the 3 particular cases which are indicated.

Exercise 4.38. *Establish with full details the resultant formula*

$$R(P,Q)=\begin{vmatrix}p_k&&&q_l&&\\ \vdots&\ddots&&\vdots&\ddots&\\ p_0&&p_k&q_0&&q_k\\ &\ddots&\vdots&&\ddots&\vdots\\ &&p_0&&&q_0\end{vmatrix}$$

and discuss as well what happens in the context of the discriminant.

As before, there are many things that can be done or learned here.

Exercise 4.39. *Clarify which functions can be applied to which matrices, as to have results stating that the eigenvalues of $f(A)$ are $f(\lambda_1),\ldots,f(\lambda_N)$.*

This exercise is actually quite difficult. We will be back to this.

Exercise 4.40. *Work out specialized spectral theorems for the orthogonal matrices $U\in O_N$, going beyond what has been said in the above.*

To be more precise here, we have proved many spectral theorems in the above, but the case $U\in O_N$, where our statement here was something quite weak, coming without a converse, is obviously still in need of some discussion.

Exercise 4.41. *Prove that any matrix can be put in Jordan form,*

$$A\sim\begin{pmatrix}J_1&&\\ &\ddots&\\ &&J_k\end{pmatrix}$$

with each of the blocks which appear, called Jordan blocks, being as follows,

$$J_i=\begin{pmatrix}\lambda_i&1&&&\\ &\lambda_i&1&&\\ &&\ddots&\ddots&\\ &&&\lambda_i&1\\ &&&&\lambda_i\end{pmatrix}$$

with the size being the multiplicity of λ_i.

This is something useful, somewhat the "nuclear option" in linear algebra.

Part II

Matrix analysis

Everything dies, baby, that's a fact
But maybe everything that dies some day comes back
Put your makeup on, fix your hair up pretty
And meet me tonight in Atlantic City

CHAPTER 5

The Jacobian

5a. Basic calculus

We discuss in what follows some applications of the theory that we developed above, for the most to questions in analysis. The idea will be that the functions of several variables $f:\mathbb{R}^N\to\mathbb{R}^M$ can be locally approximated by linear maps, in the same way as the functions $f:\mathbb{R}\to\mathbb{R}$ can be locally approximated by using derivatives:

$$f(x+t)\simeq f(x)+f'(x)t\quad,\quad f'(x)\in M_{M\times N}(\mathbb{R})$$

There are many things that can be said here, and we will be brief. As a plan for this chapter and the next one, we would like to quickly review the one-variable calculus, then develop the basics of multivariable calculus, and then get introduced to the Gaussian laws, and to probability theory in general. The instructions being as follows:

(1) In case all this is totally new to you, it is better to stop at this point with reading the present book, and quickly read some calculus. There are plenty of good books here, a standard choice being for instance the books of Lax-Terrell [**65**], [**66**].

(2) In case you know a bit about all this, stay with us. But have the books of Rudin [**73**], [**74**] nearby, for things not explained in what follows. And have a probability book nearby too, such as Feller [**35**] or Durrett [**33**], for some extra help with probability.

(3) Finally, in the case you know well analysis, have of course some fun in quickly reading the material below. But, in parallel to this, upgrade of course, always aim higher, by learning some differential geometry, say from the books of do Carmo [**29**], [**30**].

Getting started now, let us first discuss the simplest case, $f:\mathbb{R}\to\mathbb{R}$. Here we have the following result, which is the starting point for everything in analysis:

THEOREM 5.1. *Any function of one variable $f:\mathbb{R}\to\mathbb{R}$ is locally affine,*

$$f(x+t)\simeq f(x)+f'(x)t$$

with $f'(x)\in\mathbb{R}$ being the derivative of f at the point x, given by

$$f'(x)=\lim_{t\to 0}\frac{f(x+t)-f(x)}{t}$$

provided that this latter limit converges indeed.

PROOF. Assume indeed that the limit in the statement converges. By multiplying by t, we obtain that we have, once again in the $t \to 0$ limit:

$$f(x+t) - f(x) \simeq f'(x)t$$

Thus, we are led to the conclusion in the statement. $\square$

As an illustration, the derivatives of the power functions are as follows:

PROPOSITION 5.2. *We have the differentiation formula*

$$(x^p)' = px^{p-1}$$

valid for any exponent $p \in \mathbb{R}$.

PROOF. We can do this in three steps, as follows:

(1) In the case $p \in \mathbb{N}$ we can use the binomial formula, which gives, as desired:

$$\begin{aligned}
(x+t)^p &= \sum_{k=0}^{n} \binom{p}{k} x^{p-k} t^k \\
&= x^p + px^{p-1}t + \ldots + t^p \\
&\simeq x^p + px^{p-1}t
\end{aligned}$$

(2) Let us discuss now the general case $p \in \mathbb{Q}$. We write $p = m/n$, with $m \in \mathbb{N}$ and $n \in \mathbb{Z}$. In order to do the computation, we use the following formula:

$$a^n - b^n = (a-b)(a^{n-1} + a^{n-2}b + \ldots + b^{n-1})$$

To be more precise, we will use this formula, written as follows:

$$a - b = \frac{a^n - b^n}{a^{n-1} + a^{n-2}b + \ldots + b^{n-1}}$$

We set in this formula $a = (x+t)^{m/n}$ and $b = x^{m/n}$. We obtain in this way, by using the binomial formula for approximating on top, as desired:

$$\begin{aligned}
(x+t)^{m/n} - x^{m/n} &= \frac{(x+t)^m - x^m}{(x+t)^{m(n-1)/n} + \ldots + x^{m(n-1)/n}} \\
&\simeq \frac{(x+t)^m - x^m}{nx^{m(n-1)/n}} \\
&\simeq \frac{mx^{m-1}t}{nx^{m(n-1)/n}} \\
&= \frac{m}{n} \cdot x^{m-1-m+m/n} \cdot t \\
&= \frac{m}{n} \cdot x^{m/n-1} \cdot t
\end{aligned}$$

(3) In the general case now, where $p \in \mathbb{R}$ is real, the same formula holds, namely $(x^p)' = px^{p-1}$, by using what we found above, and a continuity argument. $\square$

There are many other computations that can be done, and we will be back to this later. Let us record however, at the general level, the following key result:

THEOREM 5.3. *The derivatives are subject to the following rules:*

(1) *Leibnitz rule:* $(fg)' = f'g + fg'$.
(2) *Chain rule:* $(f \circ g)' = f'(g)g'$.

PROOF. Both formulae follow from the definition of the derivative, as follows:

(1) Regarding products, we have the following computation:

$$\begin{aligned}(fg)(x+t) &= f(x+t)g(x+t)\\ &\simeq (f(x)+f'(x)t)(g(x)+g'(x)t)\\ &\simeq f(x)g(x)+(f'(x)g(x)+f(x)g'(x))t\end{aligned}$$

(2) Regarding compositions, we have the following computation:

$$\begin{aligned}(f\circ g)(x+t) &= f(g(x+t))\\ &\simeq f(g(x)+g'(x)t)\\ &\simeq f(g(x))+f'(g(x))g'(x)t\end{aligned}$$

Thus, we are led to the conclusions in the statement. □

There are many applications of the derivative, and we have for instance:

PROPOSITION 5.4. *The local minima and maxima of a differentiable function $f : \mathbb{R} \to \mathbb{R}$ appear at the points $x \in \mathbb{R}$ where:*

$$f'(x) = 0$$

However, the converse of this fact is not true in general.

PROOF. The first assertion is clear from the formula in Theorem 5.1, namely:

$$f(x+t) \simeq f(x) + f'(x)t$$

As for the converse, the simplest counterexample is $f(x) = x^3$, at $x = 0$. □

At a more advanced level now, we have the following result:

THEOREM 5.5. *Any function of one variable $f : \mathbb{R} \to \mathbb{R}$ is locally quadratic,*

$$f(x+t) \simeq f(x) + f'(x)t + \frac{f''(x)}{2}\,t^2$$

where $f''(x)$ is the derivative of the function $f' : \mathbb{R} \to \mathbb{R}$ at the point x.

PROOF. This is something quite intuitive, when thinking geometrically. In practice, we can use L'Hôpital's rule, stating that the $0/0$ type limits can be computed as:

$$\frac{f(x)}{g(x)} \simeq \frac{f'(x)}{g'(x)}$$

Observe that this formula holds indeed, as an application of Theorem 5.1. Now by using this, if we denote by $\varphi(t) \simeq P(t)$ the formula to be proved, we have:

$$\begin{aligned}\frac{\varphi(t)-P(t)}{t^2} &\simeq \frac{\varphi'(t)-P'(t)}{2t}\\ &\simeq \frac{\varphi''(t)-P''(t)}{2}\\ &= \frac{f''(x)-f''(x)}{2}\\ &= 0\end{aligned}$$

Thus, we are led to the conclusion in the statement. □

The above result substantially improves Theorem 5.1, and there are many applications of this. We can improve for instance Proposition 5.4, as follows:

PROPOSITION 5.6. *The local minima and maxima of a twice differentiable function $f : \mathbb{R} \to \mathbb{R}$ appear at the points $x \in \mathbb{R}$ where*

$$f'(x) = 0$$

with the local minima corresponding to the case $f'(x) \geq 0$, and with the local maxima corresponding to the case $f''(x) \leq 0$.

PROOF. The first assertion is something that we already know. As for the second assertion, we can use the formula in Theorem 5.5, which in the case $f'(x) = 0$ reads:

$$f(x+t) \simeq f(x) + \frac{f''(x)}{2}\, t^2$$

Indeed, assuming $f''(x) \neq 0$, it is clear that the condition $f''(x) > 0$ will produce a local minimum, and that the condition $f''(x) < 0$ will produce a local maximum. □

We can further develop the above method, at order 3, at order 4, and so on, the ultimate result on the subject, called Taylor formula, being as follows:

THEOREM 5.7. *Any function $f : \mathbb{R} \to \mathbb{R}$ can be locally approximated as*

$$f(x+t) = \sum_{k=0}^{\infty} \frac{f^{(k)}(x)}{k!}\, t^k$$

where $f^{(k)}(x)$ are the higher derivatives of f at the point x.

PROOF. We use the same method as in the proof of Theorem 5.5. If we denote by $\varphi(t) \simeq P(t)$ the approximation to be proved, we have:

$$\begin{aligned}
\frac{\varphi(t)-P(t)}{t^n} &\simeq \frac{\varphi'(t)-P'(t)}{nt^{n-1}} \\
&\simeq \frac{\varphi''(t)-P''(t)}{n(n-1)t^{n-2}} \\
&\vdots \\
&\simeq \frac{\varphi^{(n)}(t)-P^{(n)}(t)}{n!} \\
&= 0
\end{aligned}$$

Thus, we are led to the conclusion in the statement. □

As a basic application of the Taylor series, we have:

THEOREM 5.8. *We have the following formulae,*

$$\sin x = \sum_{l=0}^\infty (-1)^l \frac{x^{2l+1}}{(2l+1)!} \quad , \quad \cos x = \sum_{l=0}^\infty (-1)^l \frac{x^{2l}}{(2l)!}$$

as well as the following formulae,

$$e^x = \sum_{k=0}^\infty \frac{x^k}{k!} \quad , \quad \log(1+x) = \sum_{k=0}^\infty (-1)^{k+1} \frac{x^k}{k}$$

as Taylor series, and in general as well, with $|x|<1$ needed for log.

PROOF. There are several statements here, the proofs being as follows:

(1) Regarding sin and cos, we can use here the following well-known formulae:

$$\sin(x+y) = \sin x \cos y + \cos x \sin y$$

$$\cos(x+y) = \cos x \cos y - \sin x \sin y$$

With these formulae in hand we can appproximate both sin and cos, and we get:

$$(\sin x)' = \cos x \quad , \quad (\cos x)' = -\sin x$$

Thus, we can differentiate sin and cos as many times as we want to, and so we can compute the corresponding Taylor series, and we obtain the formulae in the statement.

(2) Regarding exp and log, here the needed formulae, which lead to the formulae in the statement for the corresponding Taylor series, are as follows:

$$(e^x)' = e^x \quad , \quad (\log x)' = x^{-1} \quad , \quad (x^p)' = px^{p-1}$$

(3) Finally, the fact that the formulae in the statement extend beyond the small x setting, coming from Taylor series, is something standard too. □

As another basic application of the Taylor formula, we have:

THEOREM 5.9. *We have the following generalized binomial formula, with $p \in \mathbb{R}$,*

$$(x+t)^p = \sum_{k=0}^{\infty} \binom{p}{k} x^{p-k} t^k$$

with the generalized binomial coefficients being given by the formula

$$\binom{p}{k} = \frac{p(p-1)\ldots(p-k+1)}{k!}$$

valid for any $|t| < |x|$. With $p \in \mathbb{N}$, we recover the usual binomial formula.

PROOF. Consider indeed the following function:

$$f(x) = x^p$$

The derivatives at $x = 1$ are then given by the following formula:

$$f^{(k)}(1) = p(p-1)\ldots(p-k+1)$$

Thus, the Taylor approximation at $x = 1$ is as follows:

$$f(1+t) = \sum_{k=0}^{\infty} \frac{p(p-1)\ldots(p-k+1)}{k!}\, t^k$$

But this is exactly our generalized binomial formula, so we are done with the case where t is small. With a bit more care, we obtain that this holds for any $|t| < 1$. $\square$

We can see from the above the power of the Taylor formula. As an application now of our generalized binomial formula, we can extract square roots, as follows:

PROPOSITION 5.10. *We have the following formula,*

$$\sqrt{1+t} = 1 - 2\sum_{k=1}^{\infty} C_{k-1} \left(\frac{-t}{4}\right)^k$$

with $C_k = \frac{1}{k+1}\binom{2k}{k}$ being the Catalan numbers. Also, we have

$$\frac{1}{\sqrt{1+t}} = \sum_{k=0}^{\infty} D_k \left(\frac{-t}{4}\right)^k$$

with $D_k = \binom{2k}{k}$ being the central binomial coefficients.

PROOF. The above formulae both follow from Theorem 5.9, as follows:

(1) At $p = 1/2$, the generalized binomial coefficients are:

$$\begin{aligned}
\binom{1/2}{k} &= \frac{1/2(-1/2)\ldots(3/2-k)}{k!}\\
&= (-1)^{k-1}\frac{1\cdot 3\cdot 5\ldots(2k-3)}{2^k k!}\\
&= (-1)^{k-1}\frac{(2k-2)!}{2^{k-1}(k-1)!2^k k!}\\
&= \frac{(-1)^{k-1}}{2^{2k-1}}\cdot\frac{1}{k}\binom{2k-2}{k-1}\\
&= -2\left(\frac{-1}{4}\right)^k C_{k-1}
\end{aligned}$$

(2) At $p = -1/2$, the generalized binomial coefficients are:

$$\begin{aligned}
\binom{-1/2}{k} &= \frac{-1/2(-3/2)\ldots(1/2-k)}{k!}\\
&= (-1)^{k}\frac{1\cdot 3\cdot 5\ldots(2k-1)}{2^k k!}\\
&= (-1)^{k}\frac{(2k)!}{2^k k!2^k k!}\\
&= \frac{(-1)^{k}}{4^k}\binom{2k}{k}\\
&= \left(\frac{-1}{4}\right)^k D_k
\end{aligned}$$

Thus, we obtain the formulae in the statement. □

Let us discuss as well the basics of integration theory. We will be very brief here, by insisting on the main concepts, and skipping technicalities. We first have:

THEOREM 5.11. *We have the Riemann integration formula,*

$$\int_a^b f(x)dx = (b-a)\times\lim_{N\to\infty}\frac{1}{N}\sum_{k=1}^N f\left(a+\frac{b-a}{N}\cdot k\right)$$

which can serve as a formal definition for the integral.

PROOF. Assume indeed that we are given a continuous function $f:[a,b]\to\mathbb{R}$, and let us try to compute the signed area below its graph, called integral and denoted $\int_a^b f(x)dx$. Obviously, this signed area equals $b-a$ times the average of the function on $[a,b]$, and we are led to the following formula, with $x_1, \ldots, x_N \in [a,b]$ being randomly chosen:

$$\int_a^b f(x)dx = (b-a) \times \lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^N f(x_i)$$

This is the so-called Monte Carlo integration formula, which is extremely useful in practice, and is used by scientists, engineers and computers. However, for theoretical purposes, it is better assume that $x_1, \ldots, x_N \in [a,b]$ are uniformly distributed. With this choice, which works of course too, the formula that we obtain is as follows:

$$\int_a^b f(x)dx = (b-a) \times \lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^N f\left(a + \frac{b-a}{N} \cdot k\right)$$

Observe that this latter formula can be alternatively written as follows, which makes it clear that the formula holds indeed, as an approximation of an area by rectangles:

$$\int_a^b f(x)dx = \lim_{N\to\infty} \sum_{k=1}^N \frac{b-a}{N} \times f\left(a + \frac{b-a}{N} \cdot k\right)$$

In any case, we have obtained the formula in the statement, and we are done. □

All this was of course extremely brief, and for more on all this, including further functions that can be integrated, we refer to Rudin [73], [74] or Lax-Terrell [65], [66]. As a useful piece of advice, however, always keep in mind the Monte Carlo formula, briefly evoked above, because that is the real thing, in connection with anything integration.

The derivatives and integrals are related in several subtle ways, and we have:

Theorem 5.12. *We have the following formulae, called fundamental theorem of calculus, integration by parts formula, and change of variable formula,*

$$\int_a^b F'(x)dx = \Big[F\Big]_a^b$$

$$\int_a^b (f'g + fg')(x)dx = \Big[fg\Big]_a^b$$

$$\int_a^b f(x)dx = \int_{\varphi^{-1}(a)}^{\varphi^{-1}(b)} f(\varphi(t))\varphi'(t)dt$$

with the convention $[F]_a^b = F(b) - F(a)$, for the first two formulae.

Proof. Again, this is standard, the idea being that the first formula is clear from the area interpretation of the integral, and that the second and third formulae follow from it, by integrating respectively the Leibnitz rule and the chain rule from Theorem 5.3. □

So long for one-variable calculus. For more on all this, we refer to any basic analysis book, good choices here being the books of Lax-Terrell [65], [66], or Rudin [73], [74].

5b. Several variables

Let us discuss now what happens in several variables. At order 1, the situation is quite similar to the one in 1 variable, but this time involving matrices, as follows:

THEOREM 5.13. *Any function $f:\mathbb{R}^N\to\mathbb{R}^M$ can be locally approximated as*

$$f(x+t)\simeq f(x)+f'(x)t$$

with $f'(x)$ being by definition the matrix of partial derivatives at x,

$$f'(x)=\left(\frac{df_i}{dx_j}(x)\right)_{ij}\in M_{M\times N}(\mathbb{R})$$

acting on the vectors $t\in\mathbb{R}^N$ by usual multiplication.

PROOF. As a first observation, the formula in the statement makes sense indeed, as an equality, or rather approximation, of vectors in $\mathbb{R}^M$, as follows:

$$f\begin{pmatrix}x_1+t_1\\ \vdots\\ x_N+t_N\end{pmatrix}\simeq f\begin{pmatrix}x_1\\ \vdots\\ x_N\end{pmatrix}+\begin{pmatrix}\frac{df_1}{dx_1}(x)&\ldots&\frac{df_1}{dx_N}(x)\\ \vdots&&\vdots\\ \frac{df_M}{dx_1}(x)&\ldots&\frac{df_M}{dx_N}(x)\end{pmatrix}\begin{pmatrix}t_1\\ \vdots\\ t_N\end{pmatrix}$$

In order to prove now this formula, which does make sense, the idea is as follows:

(1) First of all, at $N=M=1$ what we have is a usual 1-variable function $f:\mathbb{R}\to\mathbb{R}$, and the formula in the statement is something that we know well, namely:

$$f(x+t)\simeq f(x)+f'(x)t$$

(2) Let us discuss now the case $N=2,M=1$. Here what we have is a function $f:\mathbb{R}^2\to\mathbb{R}$, and by using twice the basic approximation result from (1), we obtain:

$$\begin{aligned}f\begin{pmatrix}x_1+t_1\\ x_2+t_2\end{pmatrix}&\simeq f\begin{pmatrix}x_1+t_1\\ x_2\end{pmatrix}+\frac{df}{dx_2}(x)t_2\\ &\simeq f\begin{pmatrix}x_1\\ x_2\end{pmatrix}+\frac{df}{dx_1}(x)t_1+\frac{df}{dx_2}(x)t_2\\ &= f\begin{pmatrix}x_1\\ x_2\end{pmatrix}+\begin{pmatrix}\frac{df}{dx_1}(x)&\frac{df}{dx_2}(x)\end{pmatrix}\begin{pmatrix}t_1\\ t_2\end{pmatrix}\end{aligned}$$

(3) More generally, we can deal in this way with the general case $M = 1$, with the formula here, obtained via a straightforward recurrence, being as follows:

$$\begin{aligned} f\begin{pmatrix} x_1+t_1 \\ \vdots \\ x_N+t_N \end{pmatrix} &\simeq f\begin{pmatrix} x_1 \\ \vdots \\ x_N \end{pmatrix} + \frac{df}{dx_1}(x)t_1 + \ldots + \frac{df}{dx_N}(x)t_N \\ &= f\begin{pmatrix} x_1 \\ \vdots \\ x_N \end{pmatrix} + \begin{pmatrix} \frac{df}{dx_1}(x) & \ldots & \frac{df}{dx_N}(x) \end{pmatrix} \begin{pmatrix} t_1 \\ \vdots \\ t_N \end{pmatrix} \end{aligned}$$

(4) But this gives the result in the case where both $N, M \in \mathbb{N}$ are arbitrary too. Indeed, consider a function $f : \mathbb{R}^N \to \mathbb{R}^M$, and let us write it as follows:

$$f = \begin{pmatrix} f_1 \\ \vdots \\ f_M \end{pmatrix}$$

We can apply (3) to each of the components $f_i : \mathbb{R}^N \to \mathbb{R}$, and we get:

$$f_i\begin{pmatrix} x_1+t_1 \\ \vdots \\ x_N+t_N \end{pmatrix} \simeq f_i\begin{pmatrix} x_1 \\ \vdots \\ x_N \end{pmatrix} + \begin{pmatrix} \frac{df_i}{dx_1}(x) & \ldots & \frac{df_i}{dx_N}(x) \end{pmatrix} \begin{pmatrix} t_1 \\ \vdots \\ t_N \end{pmatrix}$$

But this collection of M formulae tells us precisely that the following happens, as an equality, or rather approximation, of vectors in $\mathbb{R}^M$:

$$f\begin{pmatrix} x_1+t_1 \\ \vdots \\ x_N+t_N \end{pmatrix} \simeq f\begin{pmatrix} x_1 \\ \vdots \\ x_N \end{pmatrix} + \begin{pmatrix} \frac{df_1}{dx_1}(x) & \ldots & \frac{df_1}{dx_N}(x) \\ \vdots & & \vdots \\ \frac{df_M}{dx_1}(x) & \ldots & \frac{df_M}{dx_N}(x) \end{pmatrix} \begin{pmatrix} t_1 \\ \vdots \\ t_N \end{pmatrix}$$

Thus, we are led to the conclusion in the statement. $\square$

Generally speaking, Theorem 5.13 is what we need to know for upgrading from calculus to multivariable calculus. As a standard result here, we have:

THEOREM 5.14. *We have the chain derivative formula*

$$(f \circ g)'(x) = f'(g(x)) \cdot g'(x)$$

as an equality of matrices.

PROOF. Consider indeed a composition of functions, as follows:

$$f : \mathbb{R}^N \to \mathbb{R}^M \quad , \quad g : \mathbb{R}^K \to \mathbb{R}^N \quad , \quad f \circ g : \mathbb{R}^K \to \mathbb{R}^M$$

According to Theorem 5.13, the derivatives of these functions are certain linear maps, corresponding to certain rectangular matrices, as follows:

$$f'(g(x)) \in M_{M\times N}(\mathbb{R}) \quad , \quad g'(x) \in M_{N\times K}(\mathbb{R}) \qquad (f\circ g)'(x) \in M_{M\times K}(\mathbb{R})$$

Thus, our formula makes sense indeed. As for proof, this comes from:

$$\begin{aligned}(f\circ g)(x+t) &= f(g(x+t))\\ &\simeq f(g(x)+g'(x)t)\\ &\simeq f(g(x))+f'(g(x))g'(x)t\end{aligned}$$

Thus, we are led to the conclusion in the statement. □

Regarding now the higher derivatives, the situation here is more complicated. Let us record, however, the following fundamental result, happening at order 2, and which does the job, the job in analysis being usually that of finding the minima or maxima:

THEOREM 5.15. *Given a function $f:\mathbb{R}^N\to\mathbb{R}$, construct its Hessian, as being:*

$$H(x)=\left(\frac{d^2f}{dx_idx_j}(x)\right)_{ij}$$

We have then the following order 2 approximation of f around a given $x\in\mathbb{R}^N$,

$$f(x+t)\simeq f(x)+f'(x)t+\frac{<H(x)t,t>}{2}$$

relating the positivity properties of H to the local minima and maxima of f.

PROOF. This is quite standard, by using the same method as in the 1D case, namely building on Theorem 5.13, and approximating the function at order 2. □

Getting now to integration, as a key result here, we have:

THEOREM 5.16. *Given a transformation $\varphi=(\varphi_1,\ldots,\varphi_N)$, we have*

$$\int_E f(x)dx=\int_{\varphi^{-1}(E)}f(\varphi(t))|J_\varphi(t)|dt$$

with the J_φ quantity, called Jacobian, being given by

$$J_\varphi(t)=\det\left[\left(\frac{d\varphi_i}{dx_j}(t)\right)_{ij}\right]$$

and with this generalizing the 1-variable formula that we know well.

PROOF. This is something quite tricky, the idea being as follows:

(1) Observe first that this generalizes indeed the change of variable formula in 1 dimension, from Theorem 5.12, the point here being that the absolute value on the derivative appears as to compensate for the lack of explicit bounds for the integral.

(2) In general now, we can first argue that, the formula in the statement being linear in f, we can assume $f = 1$. Thus we want to prove $vol(E) = \int_{\varphi^{-1}(E)} |J_\varphi(t)|dt$, and with $D = \varphi^{-1}(E)$, this amounts in proving $vol(\varphi(D)) = \int_D |J_\varphi(t)|dt$.

(3) Now since this latter formula is additive with respect to D, it is enough to prove that $vol(\varphi(D)) = \int_D J_\varphi(t)dt$, for small cubes D, and assuming $J_\varphi > 0$. But this follows by using the definition of the determinant as a volume, as in chapter 2.

(4) The details and computations however are quite non-trivial, and can be found for instance in Rudin [73]. So, please read Rudin. With this, reading the complete proof of the present theorem from Rudin, being part of the standard math experience. $\square$

5c. Volumes of spheres

We discuss here some more advanced questions, related to the computation of volumes of the spheres, and to the integration over spheres. Before anything, do you know what π is? I bet not, or at least my students usually don't. So, let me teach you:

THEOREM 5.17. *Assuming that the length of the unit circle is*

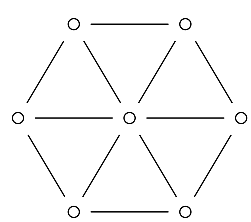
$$L = 2\pi$$

it follows that the area of the unit disk is

$$A = \pi$$

and so the two possible definitions of π are indeed equivalent.

PROOF. This follows by drawing polygons, and taking the $N \to \infty$ limit. To be more precise, let us cut the disk as a pizza, into N slices, and leave aside the rounded parts:



The area to be eaten can be then computed as follows, where H is the height of the slices, S is the length of their sides, and $P = NS$ is the total length of the sides:

$$A = N \times \frac{HS}{2} = \frac{HP}{2} \simeq \frac{1 \times L}{2} = \pi$$

Thus, we are led to the conclusion in the statement. $\square$

In N dimensions now, things are more complicated, and we will need spherical coordinates, in order to deal with such questions. Let us start with:

THEOREM 5.18. *We have polar coordinates in 2 dimensions,*

$$\begin{cases} x &= r\cos t \\ y &= r\sin t \end{cases}$$

the corresponding Jacobian being $J = r$.

PROOF. This is something elementary, the Jacobian being given by:

$$\begin{aligned} J &= \begin{vmatrix} \cos t & -r\sin t \\ \sin t & r\cos t \end{vmatrix} \\ &= r\cos^2 t + r\sin^2 t \\ &= r \end{aligned}$$

Thus, we have indeed the formula in the statement. □

In 3 dimensions now the formula is similar, as follows:

THEOREM 5.19. *We have spherical coordinates in 3 dimensions,*

$$\begin{cases} x &= r\cos s \\ y &= r\sin s\cos t \\ z &= r\sin s\sin t \end{cases}$$

the corresponding Jacobian being $J(r,s,t) = r^2\sin s$.

PROOF. The fact that we have indeed spherical coordinates is clear. Regarding now the Jacobian, this is given by the following formula:

$$\begin{aligned} J(r,s,t) &= \begin{vmatrix} \cos s & -r\sin s & 0 \\ \sin s\cos t & r\cos s\cos t & -r\sin s\sin t \\ \sin s\sin t & r\cos s\sin t & r\sin s\cos t \end{vmatrix} \\ &= r^2\sin s\sin t \begin{vmatrix} \cos s & -r\sin s \\ \sin s\sin t & r\cos s\sin t \end{vmatrix} + r\sin s\cos t \begin{vmatrix} \cos s & -r\sin s \\ \sin s\cos t & r\cos s\cos t \end{vmatrix} \\ &= r\sin s\sin^2 t \begin{vmatrix} \cos s & -r\sin s \\ \sin s & r\cos s \end{vmatrix} + r\sin s\cos^2 t \begin{vmatrix} \cos s & -r\sin s \\ \sin s & r\cos s \end{vmatrix} \\ &= r\sin s(\sin^2 t + \cos^2 t) \begin{vmatrix} \cos s & -r\sin s \\ \sin s & r\cos s \end{vmatrix} \\ &= r\sin s \times 1 \times r \\ &= r^2\sin s \end{aligned}$$

Thus, we have indeed the formula in the statement. □

Let us work out now the spherical coordinate formula in N dimensions. The result here, which generalizes those at $N = 2, 3$, is as follows:

THEOREM 5.20. *We have spherical coordinates in N dimensions,*

$$\begin{cases} x_1 &= r\cos t_1 \\ x_2 &= r\sin t_1\cos t_2 \\ \vdots \\ x_{N-1} &= r\sin t_1\sin t_2\ldots\sin t_{N-2}\cos t_{N-1} \\ x_N &= r\sin t_1\sin t_2\ldots\sin t_{N-2}\sin t_{N-1} \end{cases}$$

the corresponding Jacobian being given by the following formula:

$$J(r,t)=r^{N-1}\sin^{N-2}t_1\sin^{N-3}t_2\,\ldots\,\sin^2t_{N-3}\sin t_{N-2}$$

PROOF. As before, the fact that we have spherical coordinates is clear. Regarding now the Jacobian, also as before, by developing over the last column, we have:

$$\begin{aligned} J_N &= r\sin t_1\ldots\sin t_{N-2}\sin t_{N-1}\times\sin t_{N-1}J_{N-1}\\ &+ r\sin t_1\ldots\sin t_{N-2}\cos t_{N-1}\times\cos t_{N-1}J_{N-1}\\ &= r\sin t_1\ldots\sin t_{N-2}(\sin^2t_{N-1}+\cos^2t_{N-1})J_{N-1}\\ &= r\sin t_1\ldots\sin t_{N-2}J_{N-1} \end{aligned}$$

Thus, we obtain the formula in the statement, by recurrence. □

As a comment here, the above convention for spherical coordinates, which is particularly beautiful, is one among many. Physicists for instance like to write things a bit upside down, and the same is actually true for physicists' notation for scalar products, which is $<x,y>=\sum_i\bar{x}_iy_i$, again upside down, and for many other things. If this bothers you, I can only recommend my physics book [7], written using mathematicians' notations.

By the way, talking physics, again and I insist, you should learn some, normally from Feynman [37], [38], [39], or Griffiths [41], [42], [43], or Weinberg [87], [88], [89]. The point indeed is that using spherical coordinates, while being something usually labelled as "unconceptual", and avoided by mathematicians, is the ABC of physicists, who use it all the time. Want to do some basic electrodynamics computations? Spherical coordinates. Want to solve the hydrogen atom? Spherical coordinates, again. And so on. So, nothing better than learning some physics, in order to get to know, and love, these spherical coordinates. That we will do love in this book, from the bottom of our hearts.

Back to work now, let us compute the volumes of spheres. For this purpose, we must understand how the products of coordinates integrate over spheres. Let us start with the case $N=2$. Here the sphere is the unit circle $\mathbb{T}$, and with $z=e^{it}$ the coordinates are $\cos t,\sin t$. We can first integrate arbitrary powers of these coordinates, as follows:

Proposition 5.21. *We have the following formulae,*

$$\int_0^{\pi/2}\cos^p t\,dt=\int_0^{\pi/2}\sin^p t\,dt=\left(\frac{\pi}{2}\right)^{\varepsilon(p)}\frac{p!!}{(p+1)!!}$$

where $\varepsilon(p)=1$ if p is even, and $\varepsilon(p)=0$ if p is odd, and where

$$m!!=(m-1)(m-3)(m-5)\ldots$$

with the product ending at 2 if m is odd, and ending at 1 if m is even.

Proof. Let us first compute the integral on the left I_p. We have:

$$\begin{aligned}(\cos^p t\sin t)' &= p\cos^{p-1}t(-\sin t)\sin t+\cos^p t\cos t\\ &= p\cos^{p+1}t-p\cos^{p-1}t+\cos^{p+1}t\\ &= (p+1)\cos^{p+1}t-p\cos^{p-1}t\end{aligned}$$

By integrating between 0 and $\pi/2$, we obtain the following formula:

$$(p+1)I_{p+1}=pI_{p-1}$$

Thus we can compute I_p by recurrence, and we obtain:

$$\begin{aligned}I_p &= \frac{p-1}{p}\,I_{p-2}\\ &= \frac{p-1}{p}\cdot\frac{p-3}{p-2}\,I_{p-4}\\ &= \frac{p-1}{p}\cdot\frac{p-3}{p-2}\cdot\frac{p-5}{p-4}\,I_{p-6}\\ &\ \ \vdots\\ &= \frac{p!!}{(p+1)!!}\,I_{1-\varepsilon(p)}\end{aligned}$$

On the other hand, at $p=0$ we have the following formula:

$$I_0=\int_0^{\pi/2}1\,dt=\frac{\pi}{2}$$

Also, at $p=1$ we have the following formula:

$$I_1=\int_0^{\pi/2}\cos t\,dt=1$$

Thus, we obtain the result, by recurrence. As for the second formula, regarding $\sin t$, this follows from the first formula, with the change of variables $t=\frac{\pi}{2}-s$. ☐

We can now compute the volume of the sphere, as follows:

THEOREM 5.22. *The volume of the unit sphere in $\mathbb{R}^N$ is given by*
$$V=\left(\frac{\pi}{2}\right)^{[N/2]}\frac{2^N}{(N+1)!!}$$
with the convention
$$N!!=(N-1)(N-3)(N-5)\ldots$$
with the product ending at 2 if N is odd, and ending at 1 if N is even.

PROOF. If we denote by B^+ the positive part of the unit sphere, we have:
$$\begin{aligned}
V^+ &= \int_{B^+}1\\
&= \int_0^1\int_0^{\pi/2}\ldots\int_0^{\pi/2}r^{N-1}\sin^{N-2}t_1\ldots\sin t_{N-2}\,drdt_1\ldots dt_{N-1}\\
&= \int_0^1 r^{N-1}\,dr\int_0^{\pi/2}\sin^{N-2}t_1\,dt_1\ldots\int_0^{\pi/2}\sin t_{N-2}dt_{N-2}\int_0^{\pi/2}1dt_{N-1}\\
&= \frac{1}{N}\times\left(\frac{\pi}{2}\right)^{[N/2]}\times\frac{(N-2)!!}{(N-1)!!}\cdot\frac{(N-3)!!}{(N-2)!!}\ldots\frac{2!!}{3!!}\cdot\frac{1!!}{2!!}\cdot 1\\
&= \frac{1}{N}\times\left(\frac{\pi}{2}\right)^{[N/2]}\times\frac{1}{(N-1)!!}\\
&= \left(\frac{\pi}{2}\right)^{[N/2]}\frac{1}{(N+1)!!}
\end{aligned}$$
Thus, we are led to the formula in the statement. □

As main particular cases of the above formula, we have:

PROPOSITION 5.23. *The volumes of the low-dimensional spheres are as follows:*
(1) *At $N=1$, the length of the unit interval is $V=2$.*
(2) *At $N=2$, the area of the unit disk is $V=\pi$.*
(3) *At $N=3$, the volume of the unit sphere is $V=\frac{4\pi}{3}$*
(4) *At $N=4$, the volume of the corresponding unit sphere is $V=\frac{\pi^2}{2}$.*

PROOF. These are all particular cases of the formula in Theorem 5.22. □

5d. Basic estimates

In order to obtain estimates for the volumes, in the large N limit, we can use:

THEOREM 5.24. *We have the Stirling formula*
$$N!\simeq\left(\frac{N}{e}\right)^N\sqrt{2\pi N}$$
valid in the $N\to\infty$ limit.

PROOF. This is something quite tricky, the idea being as follows:

(1) We have the following basic approximation, by using a Riemann sum:

$$\begin{aligned}\log(N!) &= \sum_{k=1}^N \log k\\ &\approx \int_1^N \log x\,dx\\ &= \Big[x\log x - x\Big]_0^N\\ &= N\log N - N + 1\end{aligned}$$

(2) By exponentiating we get $N! \approx (N/e)^N e$, which is not bad, but not enough. So, we have to fine-tune our method. By using trapezoids instead of rectangles, we get:

$$\int_1^N \log x\,dx \approx \log(N!) - \frac{\log 1 + \log N}{2} = \log(N!) - \frac{\log N}{2}$$

Thus, we have $\log(N!) \approx N\log N - N + \frac{\log N}{2} + 1$, which by exponentiating gives:

$$N! \approx \left(\frac{N}{e}\right)^N \sqrt{N}\cdot e$$

(3) This is better than before, but still not enough. So, we have to further fine-tune our method, and by using this time some heavy analysis, namely the Taylor formula, we can estimate the error, with $\approx$ becoming $\simeq$, and with the e factor becoming $\sqrt{2\pi}$. □

With the above formula in hand, we have many useful applications, such as:

PROPOSITION 5.25. *We have the following estimate for binomial coefficients,*

$$\binom{N}{K} \simeq \left(\frac{1}{t^t(1-t)^{1-t}}\right)^N \frac{1}{\sqrt{2\pi t(1-t)N}}$$

in the $K \simeq tN \to \infty$ limit, with $t \in (0,1]$. In particular we have

$$\binom{2N}{N} \simeq \frac{4^N}{\sqrt{\pi N}}$$

in the $N \to \infty$ limit, for the central binomial coefficients.

PROOF. All this is very standard, by using the Stirling formula etablished above, for the various factorials which appear, the idea being as follows:

(1) This follows from the definition of the binomial coefficients, namely:

$$\begin{aligned}
\binom{N}{K} &= \frac{N!}{K!(N-K)!} \\
&\simeq \left(\frac{N}{e}\right)^N \sqrt{2\pi N}\left(\frac{e}{K}\right)^K \frac{1}{\sqrt{2\pi K}}\left(\frac{e}{N-K}\right)^{N-K} \frac{1}{\sqrt{2\pi(N-K)}} \\
&= \frac{N^N}{K^K(N-K)^{N-K}}\sqrt{\frac{N}{2\pi K(N-K)}} \\
&\simeq \frac{N^N}{(tN)^{tN}((1-t)N)^{(1-t)N}}\sqrt{\frac{N}{2\pi tN(1-t)N}} \\
&= \left(\frac{1}{t^t(1-t)^{1-t}}\right)^N \frac{1}{\sqrt{2\pi t(1-t)N}}
\end{aligned}$$

Thus, we are led to the conclusion in the statement.

(2) This estimate follows from a similar computation, as follows:

$$\binom{2N}{N} \simeq \left(\frac{2N}{e}\right)^{2N}\sqrt{4\pi N}\left(\frac{e}{N}\right)^{2N}\frac{1}{2\pi N} = \frac{4^N}{\sqrt{\pi N}}$$

Alternatively, we can take $t = 1/2$ in (1), then rescale. Indeed, we have:

$$\binom{N}{[N/2]} \simeq \left(\frac{1}{(\frac{1}{2})^{1/2}(\frac{1}{2})^{1/2}}\right)^N \frac{1}{\sqrt{2\pi\cdot\frac{1}{2}\cdot\frac{1}{2}\cdot N}} = 2^N\sqrt{\frac{2}{\pi N}}$$

Thus with the change $N \to 2N$ we obtain the formula in the statement. $\square$

Summarizing, we have so far complete estimate for the factorials. Regarding now the double factorials, that we will need as well, the result here is as follows:

PROPOSITION 5.26. *We have the following estimate for the double factorials,*

$$N!! \simeq \left(\frac{N}{e}\right)^{N/2} C$$

with $C = \sqrt{2}$ for N even, and $C = \sqrt{\pi}$ for N odd. Alternatively, we have

$$(N+1)!! \simeq \left(\frac{N}{e}\right)^{N/2} D$$

with $D = \sqrt{\pi N}$ for N even, and $D = \sqrt{2N}$ for N odd.

PROOF. Once again this is standard, the idea being as follows:

(1) When $N=2K$ is even, we have the following computation:

$$\begin{aligned}
N!! &= \frac{(2K)!}{2^K K!}\\
&\simeq \frac{1}{2^K}\left(\frac{2K}{e}\right)^{2K}\sqrt{4\pi K}\left(\frac{e}{K}\right)^K\frac{1}{\sqrt{2\pi K}}\\
&= \left(\frac{2K}{e}\right)^K\sqrt{2}\\
&= \left(\frac{N}{e}\right)^{N/2}\sqrt{2}
\end{aligned}$$

(2) When $N=2K+1$ is odd, we have the following computation:

$$\begin{aligned}
N!! &= 2^K K!\\
&\simeq \left(\frac{2K}{e}\right)^K\sqrt{2\pi K}\\
&= \left(\frac{2K+1}{e}\right)^{K+1/2}\sqrt{\frac{e}{2K+1}}\left(\frac{2K}{2K+1}\right)^K\sqrt{2\pi K}\\
&\simeq \left(\frac{N}{e}\right)^{N/2}\sqrt{\frac{e}{2K}}\cdot\frac{1}{\sqrt{e}}\cdot\sqrt{2\pi K}\\
&= \left(\frac{N}{e}\right)^{N/2}\sqrt{\pi}
\end{aligned}$$

(3) Back to the case where $N=2K$ is even, by using (2) we obtain:

$$\begin{aligned}
(N+1)!! &\simeq \left(\frac{N+1}{e}\right)^{(N+1)/2}\sqrt{\pi}\\
&= \left(\frac{N+1}{e}\right)^{N/2}\sqrt{\frac{N+1}{e}}\cdot\sqrt{\pi}\\
&= \left(\frac{N}{e}\right)^{N/2}\left(\frac{N+1}{N}\right)^{N/2}\sqrt{\frac{N+1}{e}}\cdot\sqrt{\pi}\\
&\simeq \left(\frac{N}{e}\right)^{N/2}\sqrt{e}\cdot\sqrt{\frac{N}{e}}\cdot\sqrt{\pi}\\
&= \left(\frac{N}{e}\right)^{N/2}\sqrt{\pi N}
\end{aligned}$$

(4) Finally, back to the case where $N=2K+1$ is odd, by using (1) we obtain:

$$\begin{aligned}(N+1)!! &\simeq \left(\frac{N+1}{e}\right)^{(N+1)/2}\sqrt{2}\\ &= \left(\frac{N+1}{e}\right)^{N/2}\sqrt{\frac{N+1}{e}}\cdot\sqrt{2}\\ &= \left(\frac{N}{e}\right)^{N/2}\left(\frac{N+1}{N}\right)^{N/2}\sqrt{\frac{N+1}{e}}\cdot\sqrt{2}\\ &\simeq \left(\frac{N}{e}\right)^{N/2}\sqrt{e}\cdot\sqrt{\frac{N}{e}}\cdot\sqrt{2}\\ &= \left(\frac{N}{e}\right)^{N/2}\sqrt{2N}\end{aligned}$$

Thus, we have proved the estimates in the statement. □

We can now estimate the volumes of the spheres, as follows:

THEOREM 5.27. *The volume of the unit sphere in $\mathbb{R}^N$ is given by*

$$V\simeq\left(\frac{2\pi e}{N}\right)^{N/2}\frac{1}{\sqrt{\pi N}}$$

in the $N\to\infty$ limit.

PROOF. We use Theorem 5.22. When N is even, the estimate goes as follows:

$$\begin{aligned}V &\simeq \left(\frac{\pi}{2}\right)^{N/2}2^N\left(\frac{e}{N}\right)^{N/2}\frac{1}{\sqrt{\pi N}}\\ &= \left(\frac{2\pi e}{N}\right)^{N/2}\frac{1}{\sqrt{\pi N}}\end{aligned}$$

In the case where N is odd, the estimate goes as follows:

$$\begin{aligned}V &\simeq \left(\frac{\pi}{2}\right)^{(N-1)/2}2^N\left(\frac{e}{N}\right)^{N/2}\frac{1}{\sqrt{2N}}\\ &= \sqrt{\frac{2}{\pi}}\left(\frac{2\pi e}{N}\right)^{N/2}\frac{1}{\sqrt{2N}}\\ &= \left(\frac{2\pi e}{N}\right)^{N/2}\frac{1}{\sqrt{\pi N}}\end{aligned}$$

Thus, we are led to the uniform formula in the statement. □

Getting back now to our main result so far, Theorem 5.22, we can compute in the same way the area of the sphere, the result being as follows:

THEOREM 5.28. *The area of the unit sphere in $\mathbb{R}^N$ is given by*

$$A = \left(\frac{\pi}{2}\right)^{[N/2]} \frac{2^N}{(N-1)!!}$$

with the our usual convention for double factorials, namely:

$$N!! = (N-1)(N-3)(N-5)\ldots$$

In particular, at $N = 2, 3, 4$ we obtain respectively $A = 2\pi, 4\pi, 2\pi^2$.

PROOF. Regarding the first assertion, we can use the slicing argument from the proof of Theorem 5.17, which shows that the area and volume of the sphere in $\mathbb{R}^N$ are related by the following formula, which together with Theorem 5.22 gives the result:

$$A = N \cdot V$$

As for the last assertion, this can be either worked out directly, or deduced from the results for volumes that we have so far, by multiplying by N. $\square$

5e. Exercises

There has been a lot of material in this chapter. In what regards the functions of one variable, and more specifically the second derivative, the standard exercise here is:

EXERCISE 5.29. *Given a convex function $f : \mathbb{R} \to \mathbb{R}$, prove that we have the following Jensen inequality, for any $x_1, \ldots, x_N \in \mathbb{R}$, and any $\lambda_1, \ldots, \lambda_N > 0$ summing up to 1,*

$$f(\lambda_1 x_1 + \ldots + \lambda_N x_N) \leq \lambda_1 f(x_1) + \ldots + \lambda_N x_N$$

with equality when $x_1 = \ldots = x_N$. In particular, by taking the weights λ_i to be all equal, we obtain the following Jensen inequality, valid for any $x_1, \ldots, x_N \in \mathbb{R}$,

$$f\left(\frac{x_1 + \ldots + x_N}{N}\right) \leq \frac{f(x_1) + \ldots + f(x_N)}{N}$$

and once again with equality when $x_1 = \ldots = x_N$. Prove also that a similar statement holds for the concave functions, with all the inequalities being reversed.

This is something very classical, enjoy. For a bonus point, try the functions of several variables as well, and comment on the condition $f'' \geq 0$ in this case.

EXERCISE 5.30. *Prove that for $p \in (1, \infty)$ we have the following Hölder inequality*

$$\left|\frac{x_1 + \ldots + x_N}{N}\right|^p \leq \frac{|x_1|^p + \ldots + |x_N|^p}{N}$$

and that for $p \in (0, 1)$ we have the following reverse Hölder inequality

$$\left|\frac{x_1 + \ldots + x_N}{N}\right|^p \geq \frac{|x_1|^p + \ldots + |x_N|^p}{N}$$

with in both cases equality precisely when $|x_1| = \ldots = |x_N|$.

As a bonus exercise here, try as well, directly, the case $p = 2$.

EXERCISE 5.31. *Further meditate about the definition of π, notably by improving, if needed, our result about the length and area of the unit circle,*

$$L = 2\pi \implies A = \pi$$

into a result as follows, regarding the length and area of a circle of radius R:

$$L = 2\pi R \implies A = \pi R^2$$

Also, reformulate the proof given in the above as to avoid trashing the rounded parts of the pizza, which gastronomically speaking, is an heresy.

Philosophically speaking, the question amounts in understanding what 1 is, geometrically speaking. There might be a discussion about pineapple toppings as well.

EXERCISE 5.32. *Push to the limits the meditations about π, by reformulating things as follows, in case you find areas to be more familiar than lengths,*

$$A = \pi \implies L = 2\pi$$

or even as follows, with V being the volume of unit cylinder, in case you find volumes to be more familiar than areas and lengths:

$$V = \pi \implies L = 2\pi$$

Also, in case you end up with the conclusion that you actually disagree with all the above, work out your own theory of π, say by starting with your favorite $\emptyset$ set.

The point with the first question is that, in real life, as opposed to human-made mathematics as we know it, the volumes are far easier to compute than the areas and lengths, simply by immersing your 3D body into a square glass of water. Also, in the unfortunate case where you ended up disagreeing with all the above, it is probably because you have pushed your meditations too far. Have a beer, with that pizza, and relax.

EXERCISE 5.33. *Develop the theory of the gamma function, defined as*

$$\Gamma(z) = \int_0^\infty x^{z-1} e^{-x} dx$$

notably by establishing the following formula, for any $N \in \mathbb{N}$,

$$\Gamma(N) = (N-1)!$$

and then comment on the formulae for the volumes and areas of spheres.

To be more precise, the first question is that of establishing the well-known formula $\Gamma(z+1) = z\Gamma(z)$. The next step is that of computing $\Gamma(z)$ for $z \in \mathbb{N}/2$, with the above formula in the case $z \in \mathbb{N}$. And then, the problem is that of deciding if all this can be useful in connection with the formulae for the volumes and areas of spheres. With the comment that the answer to this latter question is of course something quite subjective.

CHAPTER 6

Normal laws

6a. Gauss integral

In this chapter we discuss the basics of probability theory, as an application of the methods developed in chapter 5. You might have heard or noticed that most measurements in real life lead to a bell-shaped curve, depending on a parameter $h > 0$, measuring the height of the bell, that we will assume here to appear above the point $x = 0$.

Some further experiments and study show that the bell-shaped curve must appear in fact as a variation of the function e^{-x^2}. Thus, we can expect the formula of this curve to be as follows, with $\lambda > 0$ being related to the height $h > 0$:

$$f_\lambda(x) = C_\lambda \cdot e^{-\lambda x^2}$$

But, the problem is, what is the constant C_λ? Since the total probability for our event to appear is 1, this constant must be the inverse of the integral of $e^{-\lambda x^2}$:

$$C_\lambda = \left(\int_{\mathbb{R}} e^{-\lambda x^2} dx\right)^{-1}$$

Thus, we are led into computing the above integral. By doing a change of variables we can assume $\lambda = 1$, and so we are led into the following question:

$$\int_{\mathbb{R}} e^{-x^2} dx = ?$$

And suprise here, all the methods that we know from basic 1-variable calculus fail, in relation with this question. The function e^{-x^2} has no computable primitive, and all tricks that we know, namely change of variables, partial integration, and more, all fail.

Fortunately multivariable calculus comes to the rescue, and we have:

THEOREM 6.1. *We have the following formula,*

$$\int_{\mathbb{R}} e^{-x^2} dx = \sqrt{\pi}$$

called Gauss integral formula.

PROOF. Let I be the above integral. By using polar coordinates, we obtain:

$$\begin{aligned} I^2 &= \int_{\mathbb{R}}\int_{\mathbb{R}} e^{-x^2-y^2}dxdy \\ &= \int_0^{2\pi}\int_0^\infty e^{-r^2}rdrdt \\ &= 2\pi\int_0^\infty\left(-\frac{e^{-r^2}}{2}\right)'dr \\ &= 2\pi\left[0-\left(-\frac{1}{2}\right)\right] \\ &= \pi \end{aligned}$$

Thus, we are led to the formula in the statement. $\square$

As a main application of the Gauss formula, we can now formulate:

DEFINITION 6.2. *The normal law of parameter* 1 *is the following measure:*

$$g_1 = \frac{1}{\sqrt{2\pi}}e^{-x^2/2}dx$$

More generally, the normal law of parameter $t > 0$ *is the following measure:*

$$g_t = \frac{1}{\sqrt{2\pi t}}e^{-x^2/2t}dx$$

These are also called Gaussian distributions, with "g" standing for Gauss.

We should mention that these laws are traditionally denoted $\mathcal{N}(0,1)$ and $\mathcal{N}(0,t)$, but since we will be doing here in this book all kinds of probability, namely real and complex, discrete and continuous, and so on, we will have to deal with lots of interesting probability laws, and we will be using simplified notations for them. Let us mention as well that the normal laws traditionally have 2 parameters, the mean and the variance. Here we do not need the mean, all our theory in this book using centered laws.

Observe that the above laws have indeed mass 1, as they should. This follows indeed from the Gauss formula, which gives, with $x = \sqrt{2t}\,y$:

$$\begin{aligned} \int_{\mathbb{R}} e^{-x^2/2t}dx &= \int_{\mathbb{R}} e^{-y^2}\sqrt{2t}\,dy \\ &= \sqrt{2t}\int_{\mathbb{R}} e^{-y^2}dy \\ &= \sqrt{2t}\times\sqrt{\pi} \\ &= \sqrt{2\pi t} \end{aligned}$$

Generally speaking, the normal laws appear as bit everywhere, in real life. The reasons behind this phenomenon come from the Central Limit Theorem (CLT), that we will explain in a moment. At the level of basic facts now, we first have:

PROPOSITION 6.3. *We have the variance formula*

$$V(g_t) = t$$

valid for any $t > 0$.

PROOF. The first moment is 0, because our normal law g_t is centered. As for the second moment, this can be computed as follows:

$$\begin{aligned} M_2 &= \frac{1}{\sqrt{2\pi t}} \int_{\mathbb{R}} x^2 e^{-x^2/2t} dx \\ &= \frac{1}{\sqrt{2\pi t}} \int_{\mathbb{R}} (tx) \left(-e^{-x^2/2t}\right)' dx \\ &= \frac{1}{\sqrt{2\pi t}} \int_{\mathbb{R}} t e^{-x^2/2t} dx \\ &= t \end{aligned}$$

We conclude from this that the variance is $V = M_2 = t$. □

More generally now, we have the following result:

THEOREM 6.4. *The moments of the normal law are the numbers*

$$M_k(g_t) = t^{k/2} \times k!!$$

where the double factorials are by definition given by

$$k!! = 1 \cdot 3 \cdot 5 \dots (k-1)$$

with the convention $k!! = 0$ *when* k *is odd.*

PROOF. We have the following computation:

$$\begin{aligned} M_k &= \frac{1}{\sqrt{2\pi t}} \int_{\mathbb{R}} x^k e^{-x^2/2t} dx \\ &= \frac{1}{\sqrt{2\pi t}} \int_{\mathbb{R}} (tx^{k-1}) \left(-e^{-x^2/2t}\right)' dx \\ &= \frac{1}{\sqrt{2\pi t}} \int_{\mathbb{R}} t(k-1) x^{k-2} e^{-x^2/2t} dx \\ &= t(k-1) \times \frac{1}{\sqrt{2\pi t}} \int_{\mathbb{R}} x^{k-2} e^{-x^2/2t} dx \\ &= t(k-1) M_{k-2} \end{aligned}$$

On the other hand, we have $M_0 = 1$, $M_1 = 0$. Thus by recurrence, the even moments vanish, and the odd moments are given by the formula in the statement. □

We can improve the above result, as follows:

THEOREM 6.5. *The moments of the normal law are the numbers*

$$M_k(g_t) = t^{k/2}|P_2(k)|$$

where $P_2(k)$ is the set of pairings of $\{1,\ldots,k\}$.

PROOF. Let us count the pairings of $\{1,\ldots,k\}$. In order to have such a pairing, we must pair 1 with one of $2,\ldots,k$, and then use a pairing of the remaining $k-2$ points. Thus, we have the following recurrence formula:

$$|P_2(k)| = (k-1)|P_2(k-2)|$$

Thus $|P_2(k)| = k!!$, and we are led to the conclusion in the statement. □

We are done done yet, and here is one more improvement:

THEOREM 6.6. *The moments of the normal law are the numbers*

$$M_k(g_t) = \sum_{\pi\in P_2(k)} t^{|\pi|}$$

where $P_2(k)$ is the set of pairings of $\{1,\ldots,k\}$, and $|.|$ is the number of blocks.

PROOF. This follows indeed from Theorem 6.5, because the number of blocks of a pairing of $\{1,\ldots,k\}$ is trivially $k/2$, independently of the pairing. □

We will see later on that many other interesting probability distributions are subject to similar formulae regarding their moments, involving partitions, and a lot of interesting combinatorics. Discussing this will be in fact a main theme of the present book.

6b. Central limits

The fundamental result in probability is the Central Limit Theorem (CLT), and our next task will be that of explaining this. Let us start with:

DEFINITION 6.7. *Let X be a probability space, that is to say, a space with a probability measure, and with the corresponding integration denoted $\mathbb{E}$, and called expectation.*

(1) *The random variables are the real functions as follows:*

$$f \in L^\infty(X)$$

(2) *The moments of such a variable are the following numbers:*

$$M_k(f) = \mathbb{E}(f^k)$$

(3) *The law of such a variable is the measure given by:*

$$M_k(f) = \int_{\mathbb{R}} x^k d\mu_f(x)$$

Here the fact that μ_f exists indeed is not trivial. By linearity, we would like to have a real probability measure making hold the following formula, for any $P \in \mathbb{R}[X]$:

$$\mathbb{E}(P(f)) = \int_{\mathbb{R}} P(x) d\mu_f(x)$$

By using a continuity argument, it is enough to have this formula for the characteristic functions χ_I of the arbitrary measurable sets of real numbers $I \subset \mathbb{R}$:

$$\mathbb{E}(\chi_I(f)) = \int_{\mathbb{R}} \chi_I(x) d\mu_f(x)$$

Thus, we would like to have a measure μ_f such that:

$$\mathbb{P}(f \in I) = \mu_f(I)$$

But this latter formula can serve as a definition for μ_f, with the axioms of real probability measures being trivially satisfied, and so we are done.

Next in line, we need to talk about independence. Once again with the idea of doing things a bit abstractly, the definition here is as follows:

Definition 6.8. *Two variables $f, g \in L^\infty(X)$ are called independent when*

$$\mathbb{E}(f^k g^l) = \mathbb{E}(f^k) \cdot \mathbb{E}(g^l)$$

happens, for any $k, l \in \mathbb{N}$.

Once again, this definition hides some non-trivial things. Indeed, by linearity, we would like to have a formula as follows, valid for any polynomials $P, Q \in \mathbb{R}[X]$:

$$\mathbb{E}(P(f)Q(g)) = \mathbb{E}(P(f)) \cdot \mathbb{E}(Q(g))$$

By continuity, it is enough to have this formula for characteristic functions χ_I, χ_J of the arbitrary measurable sets of real numbers $I, J \subset \mathbb{R}$:

$$\mathbb{E}(\chi_I(f)\chi_J(g)) = \mathbb{E}(\chi_I(f)) \cdot \mathbb{E}(\chi_J(g))$$

Thus, we are led to the usual definition of independence, namely:

$$\mathbb{P}(f \in I, g \in J) = \mathbb{P}(f \in I) \cdot \mathbb{P}(g \in J)$$

All this might seem a bit abstract, but in practice, the idea is of course that f, g must be independent, in an intuitive, real-life sense.

Here is now our first result, regarding this notion of independence:

Proposition 6.9. *Assuming that $f, g \in L^\infty(X)$ are independent, we have*

$$\mu_{f+g} = \mu_f * \mu_g$$

where $*$ is the convolution of real probability measures.

PROOF. We have the following computation, using the independence of f,g:

$$\begin{aligned}M_k(f+g) &= \mathbb{E}((f+g)^k)\\&= \sum_l\binom{k}{l}\mathbb{E}(f^lg^{k-l})\\&= \sum_l\binom{k}{l}M_l(f)M_{k-l}(g)\end{aligned}$$

On the other hand, by using the Fubini theorem, we have as well:

$$\begin{aligned}\int_{\mathbb{R}}x^kd(\mu_f*\mu_g)(x) &= \int_{\mathbb{R}\times\mathbb{R}}(x+y)^kd\mu_f(x)d\mu_g(y)\\&= \sum_l\binom{k}{l}\int_{\mathbb{R}}x^kd\mu_f(x)\int_{\mathbb{R}}y^ld\mu_g(y)\\&= \sum_l\binom{k}{l}M_l(f)M_{k-l}(g)\end{aligned}$$

Thus the measures μ_{f+g} and $\mu_f*\mu_g$ have the same moments:

$$M_k(\mu_{f+g})=M_k(\mu_f*\mu_g)$$

We conclude that these two measures coincide, as stated. □

Here is now our second result, which is something more advanced, providing us with some efficient tools for the study of the independence:

THEOREM 6.10. *Assuming that $f,g\in L^\infty(X)$ are independent, we have*

$$F_{f+g}=F_fF_g$$

where $F_f(x)=\mathbb{E}(e^{ixf})$ is the Fourier transform.

PROOF. We have the following computation, using Proposition 6.9 and Fubini:

$$\begin{aligned}F_{f+g}(x) &= \int_{\mathbb{R}}e^{ixy}d\mu_{f+g}(y)\\&= \int_{\mathbb{R}}e^{ixy}d(\mu_f*\mu_g)(y)\\&= \int_{\mathbb{R}\times\mathbb{R}}e^{ix(y+z)}d\mu_f(y)d\mu_g(z)\\&= \int_{\mathbb{R}}e^{ixy}d\mu_f(y)\int_{\mathbb{R}}e^{ixz}d\mu_g(z)\\&= F_f(x)F_g(x)\end{aligned}$$

Thus, we are led to the conclusion in the statement. □

Getting now to the normal laws, we have the following key result:

THEOREM 6.11. *We have the following formula, valid for any $t > 0$:*

$$F_{g_t}(x) = e^{-tx^2/2}$$

In particular, the normal laws satisfy $g_s * g_t = g_{s+t}$, for any $s, t > 0$.

PROOF. The Fourier transform formula can be established as follows:

$$\begin{aligned}
F_{g_t}(x) &= \frac{1}{\sqrt{2\pi t}} \int_{\mathbb{R}} e^{-y^2/2t+ixy} dy \\
&= \frac{1}{\sqrt{2\pi t}} \int_{\mathbb{R}} e^{-(y/\sqrt{2t}-\sqrt{t/2}ix)^2-tx^2/2} dy \\
&= \frac{1}{\sqrt{2\pi t}} \int_{\mathbb{R}} e^{-z^2-tx^2/2} \sqrt{2t} dz \\
&= \frac{1}{\sqrt{\pi}} e^{-tx^2/2} \int_{\mathbb{R}} e^{-z^2} dz \\
&= \frac{1}{\sqrt{\pi}} e^{-tx^2/2} \cdot \sqrt{\pi} \\
&= e^{-tx^2/2}
\end{aligned}$$

As for the last assertion, this follows from the fact that $\log F_{g_t}$ is linear in t. □

We are now ready to state and prove the CLT, as follows:

THEOREM 6.12 (CLT). *Given random variables $f_1, f_2, f_3, \ldots \in L^\infty(X)$ which are i.i.d., centered, and with variance $t > 0$, we have, with $n \to \infty$, in moments,*

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n f_i \sim g_t$$

where g_t is the Gaussian law of parameter t, having as density $\frac{1}{\sqrt{2\pi t}} e^{-y^2/2t} dy$.

PROOF. We use the Fourier transform, which is by definition given by:

$$F_f(x) = \mathbb{E}(e^{ixf})$$

In terms of moments, we have the following formula:

$$\begin{aligned} F_f(x) &= \mathbb{E}\left(\sum_{k=0}^\infty \frac{(ixf)^k}{k!}\right)\\ &= \sum_{k=0}^\infty \frac{(ix)^k\mathbb{E}(f^k)}{k!}\\ &= \sum_{k=0}^\infty \frac{i^kM_k(f)}{k!}\,x^k \end{aligned}$$

Thus, the Fourier transform of the variable in the statement is:

$$\begin{aligned} F(x) &= \left[F_f\left(\frac{x}{\sqrt{n}}\right)\right]^n\\ &= \left[1-\frac{tx^2}{2n}+O(n^{-2})\right]^n\\ &\simeq \left[1-\frac{tx^2}{2n}\right]^n\\ &\simeq e^{-tx^2/2} \end{aligned}$$

But this latter function being the Fourier transform of g_t, we obtain the result. □

This was for our basic presentation of probability theory, which was of course rather quick. For more on all this, we refer to any standard probability book, such as Feller [35], or Durrett [33]. But we will be back to this, on several occasions.

Needless to say, you can also learn some good probability from physicists, or other scientists. In fact, probability theory was accepted only recently, in the late 20th century, as a respectable branch of mathematics, and if there are some scientists who have taken probability seriously, and this since ever, these are the physicists.

So, in what regards now physics books, there is a bit of probability a bit everywhere, but especially in thermodynamics and statistical mechanics. And here, for getting introduced to the subject, you can go with Feynman [37], or with the classical book of Fermi [36], or with the more recent book of Schroeder [78]. And for more, a standard reference for learning statistical mechanics is the book of Huang [48].

6c. Spherical integrals

Let us discuss now the computation of the arbitrary integrals over the sphere. We will need a technical result extending the formulae from chapter 5, as follows:

THEOREM 6.13. *We have the following formula,*

$$\int_0^{\pi/2}\cos^p t\sin^q t\,dt=\left(\frac{\pi}{2}\right)^{\varepsilon(p)\varepsilon(q)}\frac{p!!q!!}{(p+q+1)!!}$$

where $\varepsilon(p)=1$ if p is even, and $\varepsilon(p)=0$ if p is odd, and where

$$m!!=(m-1)(m-3)(m-5)\ldots$$

with the product ending at 2 if m is odd, and ending at 1 if m is even.

PROOF. Let I_{pq} be the integral in the statement. In order to do the partial integration, observe that we have:

$$\begin{aligned}(\cos^p t\sin^q t)' &= p\cos^{p-1}t(-\sin t)\sin^q t\\ &+ \cos^p t\cdot q\sin^{q-1}t\cos t\\ &= -p\cos^{p-1}t\sin^{q+1}t+q\cos^{p+1}t\sin^{q-1}t\end{aligned}$$

By integrating between 0 and $\pi/2$, we obtain, for $p,q>0$:

$$pI_{p-1,q+1}=qI_{p+1,q-1}$$

Thus, we can compute I_{pq} by recurrence. When q is even we have:

$$\begin{aligned}I_{pq} &= \frac{q-1}{p+1}\,I_{p+2,q-2}\\ &= \frac{q-1}{p+1}\cdot\frac{q-3}{p+3}\,I_{p+4,q-4}\\ &= \frac{q-1}{p+1}\cdot\frac{q-3}{p+3}\cdot\frac{q-5}{p+5}\,I_{p+6,q-6}\\ &= \vdots\\ &= \frac{p!!q!!}{(p+q)!!}\,I_{p+q}\end{aligned}$$

But the last term comes from the formulae in chapter 5, and we obtain the result:

$$\begin{aligned}I_{pq} &= \frac{p!!q!!}{(p+q)!!}\,I_{p+q}\\ &= \frac{p!!q!!}{(p+q)!!}\left(\frac{\pi}{2}\right)^{\varepsilon(p+q)}\frac{(p+q)!!}{(p+q+1)!!}\\ &= \left(\frac{\pi}{2}\right)^{\varepsilon(p)\varepsilon(q)}\frac{p!!q!!}{(p+q+1)!!}\end{aligned}$$

Observe that this gives the result for p even as well, by symmetry. Indeed, we have $I_{pq}=I_{qp}$, by using the following change of variables:

$$t=\frac{\pi}{2}-s$$

In the remaining case now, where both p, q are odd, we can use once again the formula $pI_{p-1,q+1} = qI_{p+1,q-1}$ established above, and the recurrence goes as follows:

$$\begin{aligned}
I_{pq} &= \frac{q-1}{p+1}\, I_{p+2,q-2}\\
&= \frac{q-1}{p+1}\cdot\frac{q-3}{p+3}\, I_{p+4,q-4}\\
&= \frac{q-1}{p+1}\cdot\frac{q-3}{p+3}\cdot\frac{q-5}{p+5}\, I_{p+6,q-6}\\
&= \vdots\\
&= \frac{p!!q!!}{(p+q-1)!!}\, I_{p+q-1,1}
\end{aligned}$$

In order to compute the last term, observe that we have:

$$\begin{aligned}
I_{p1} &= \int_0^{\pi/2} \cos^p t \sin t\, dt\\
&= -\frac{1}{p+1}\int_0^{\pi/2} (\cos^{p+1} t)'\, dt\\
&= \frac{1}{p+1}
\end{aligned}$$

Thus, we can finish our computation in the case p, q odd, as follows:

$$\begin{aligned}
I_{pq} &= \frac{p!!q!!}{(p+q-1)!!}\, I_{p+q-1,1}\\
&= \frac{p!!q!!}{(p+q-1)!!}\cdot\frac{1}{p+q}\\
&= \frac{p!!q!!}{(p+q+1)!!}
\end{aligned}$$

Thus, we obtain the formula in the statement, the exponent of $\pi/2$ appearing there being $\varepsilon(p)\varepsilon(q) = 0 \cdot 0 = 0$ in the present case, and this finishes the proof. □

We can now integrate over the spheres, as follows:

THEOREM 6.14. *The polynomial integrals over the unit sphere $S^{N-1}_{\mathbb{R}} \subset \mathbb{R}^N$, with respect to the normalized, mass 1 measure, are given by the following formula,*

$$\int_{S^{N-1}_{\mathbb{R}}} x_1^{k_1} \ldots x_N^{k_N}\, dx = \frac{(N-1)!!k_1!!\ldots k_N!!}{(N+\Sigma k_i - 1)!!}$$

valid when all exponents k_i are even. If an exponent is odd, the integral vanishes.

PROOF. Assume first that one of the exponents k_i is odd. We can make then the following change of variables, which shows that the integral in the statement vanishes:

$$x_i \to -x_i$$

Assume now that all the exponents k_i are even. As a first observation, the result holds indeed at $N = 2$, due to the formula from Theorem 6.13, which reads:

$$\begin{aligned}\int_0^{\pi/2} \cos^p t \sin^q t \, dt &= \left(\frac{\pi}{2}\right)^{\varepsilon(p)\varepsilon(q)} \frac{p!!q!!}{(p+q+1)!!} \\ &= \frac{p!!q!!}{(p+q+1)!!}\end{aligned}$$

Indeed, this formula computes the integral in the statement over the first quadrant. But since the exponents $p, q \in \mathbb{N}$ are assumed to be even, the integrals over the other quadrants are given by the same formula, so when averaging we obtain the result.

In the general case now, where the dimension $N \in \mathbb{N}$ is arbitrary, the integral in the statement can be written in spherical coordinates, as follows:

$$I = \frac{2^N}{A} \int_0^{\pi/2} \ldots \int_0^{\pi/2} x_1^{k_1} \ldots x_N^{k_N} J \, dt_1 \ldots dt_{N-1}$$

Here A is the area of the sphere, J is the Jacobian, and the 2^N factor comes from the restriction to the $1/2^N$ part of the sphere where all the coordinates are positive. According to our computations in chapter 5, the normalization constant in front of the integral is:

$$\frac{2^N}{A} = \left(\frac{2}{\pi}\right)^{[N/2]} (N-1)!!$$

As for the unnormalized integral, this is given by:

$$\begin{aligned}I' = \int_0^{\pi/2} \ldots \int_0^{\pi/2} \quad & (\cos t_1)^{k_1} (\sin t_1 \cos t_2)^{k_2} \\ & \vdots \\ & (\sin t_1 \sin t_2 \ldots \sin t_{N-2} \cos t_{N-1})^{k_{N-1}} \\ & (\sin t_1 \sin t_2 \ldots \sin t_{N-2} \sin t_{N-1})^{k_N} \\ & \sin^{N-2} t_1 \sin^{N-3} t_2 \ldots \sin^2 t_{N-3} \sin t_{N-2} \\ & dt_1 \ldots dt_{N-1}\end{aligned}$$

By rearranging the terms, we obtain:

$$\begin{aligned} I' &= \int_0^{\pi/2} \cos^{k_1} t_1 \sin^{k_2+\ldots+k_N+N-2} t_1\, dt_1 \\ &\quad \int_0^{\pi/2} \cos^{k_2} t_2 \sin^{k_3+\ldots+k_N+N-3} t_2\, dt_2 \\ &\quad \vdots \\ &\quad \int_0^{\pi/2} \cos^{k_{N-2}} t_{N-2} \sin^{k_{N-1}+k_N+1} t_{N-2}\, dt_{N-2} \\ &\quad \int_0^{\pi/2} \cos^{k_{N-1}} t_{N-1} \sin^{k_N} t_{N-1}\, dt_{N-1} \end{aligned}$$

Now by using the above-mentioned formula at $N = 2$, this gives:

$$\begin{aligned} I' &= \frac{k_1!!(k_2+\ldots+k_N+N-2)!!}{(k_1+\ldots+k_N+N-1)!!}\left(\frac{\pi}{2}\right)^{\varepsilon(N-2)} \\ &\quad \frac{k_2!!(k_3+\ldots+k_N+N-3)!!}{(k_2+\ldots+k_N+N-2)!!}\left(\frac{\pi}{2}\right)^{\varepsilon(N-3)} \\ &\quad \vdots \\ &\quad \frac{k_{N-2}!!(k_{N-1}+k_N+1)!!}{(k_{N-2}+k_{N-1}+l_N+2)!!}\left(\frac{\pi}{2}\right)^{\varepsilon(1)} \\ &\quad \frac{k_{N-1}!!k_N!!}{(k_{N-1}+k_N+1)!!}\left(\frac{\pi}{2}\right)^{\varepsilon(0)} \end{aligned}$$

Now let F be the part involving the double factorials, and P be the part involving the powers of $\pi/2$, so that $I' = F \cdot P$. Regarding F, by cancelling terms we have:

$$F = \frac{k_1!!\ldots k_N!!}{(\Sigma k_i + N - 1)!!}$$

As in what regards P, by summing the exponents, we obtain $P = \left(\frac{\pi}{2}\right)^{[N/2]}$. We can now put everything together, and we obtain:

$$\begin{aligned} I &= \frac{2^N}{A} \times F \times P \\ &= \left(\frac{2}{\pi}\right)^{[N/2]} (N-1)!! \times \frac{k_1!!\ldots k_N!!}{(\Sigma k_i + N - 1)!!} \times \left(\frac{\pi}{2}\right)^{[N/2]} \\ &= \frac{(N-1)!!k_1!!\ldots k_N!!}{(\Sigma k_i + N - 1)!!} \end{aligned}$$

Thus, we are led to the conclusion in the statement. □

We have the following useful generalization of the above formula:

THEOREM 6.15. *We have the following integration formula over the sphere $S^{N-1}_\mathbb{R} \subset \mathbb{R}^N$, with respect to the normalized measure, valid for any exponents $k_i \in \mathbb{N}$,*

$$\int_{S^{N-1}_\mathbb{R}} |x_1^{k_1} \ldots x_N^{k_N}|\,dx = \left(\frac{2}{\pi}\right)^{\Sigma(k_1,\ldots,k_N)} \frac{(N-1)!!k_1!!\ldots k_N!!}{(N+\Sigma k_i-1)!!}$$

with $\Sigma = [odds/2]$ if N is odd and $\Sigma = [(odds+1)/2]$ if N is even, where "odds" denotes the number of odd numbers in the sequence $k_1, \ldots, k_N$.

PROOF. As before, the formula holds at $N = 2$, due to Theorem 6.13. In general, the integral in the statement can be written in spherical coordinates, as follows:

$$I = \frac{2^N}{A}\int_0^{\pi/2} \ldots \int_0^{\pi/2} x_1^{k_1} \ldots x_N^{k_N}\,J\,dt_1 \ldots dt_{N-1}$$

Here A is the area of the sphere, J is the Jacobian, and the 2^N factor comes from the restriction to the $1/2^N$ part of the sphere where all the coordinates are positive. The normalization constant in front of the integral is, as before:

$$\frac{2^N}{A} = \left(\frac{2}{\pi}\right)^{[N/2]}(N-1)!!$$

As for the unnormalized integral, this can be written as before, as follows:

$$\begin{aligned}
I' \quad = \quad & \int_0^{\pi/2} \cos^{k_1} t_1 \sin^{k_2+\ldots+k_N+N-2} t_1\,dt_1\\
& \int_0^{\pi/2} \cos^{k_2} t_2 \sin^{k_3+\ldots+k_N+N-3} t_2\,dt_2\\
& \vdots\\
& \int_0^{\pi/2} \cos^{k_{N-2}} t_{N-2} \sin^{k_{N-1}+k_N+1} t_{N-2}\,dt_{N-2}\\
& \int_0^{\pi/2} \cos^{k_{N-1}} t_{N-1} \sin^{k_N} t_{N-1}\,dt_{N-1}
\end{aligned}$$

Now by using the formula at $N=2$, we get:

$$\begin{aligned}
I' &= \frac{\pi}{2}\cdot\frac{k_1!!(k_2+\ldots+k_N+N-2)!!}{(k_1+\ldots+k_N+N-1)!!}\left(\frac{2}{\pi}\right)^{\delta(k_1,k_2+\ldots+k_N+N-2)}\\
&\quad \frac{\pi}{2}\cdot\frac{k_2!!(k_3+\ldots+k_N+N-3)!!}{(k_2+\ldots+k_N+N-2)!!}\left(\frac{2}{\pi}\right)^{\delta(k_2,k_3+\ldots+k_N+N-3)}\\
&\quad \vdots\\
&\quad \frac{\pi}{2}\cdot\frac{k_{N-2}!!(k_{N-1}+k_N+1)!!}{(k_{N-2}+k_{N-1}+k_N+2)!!}\left(\frac{2}{\pi}\right)^{\delta(k_{N-2},k_{N-1}+k_N+1)}\\
&\quad \frac{\pi}{2}\cdot\frac{k_{N-1}!!k_N!!}{(k_{N-1}+k_N+1)!!}\left(\frac{2}{\pi}\right)^{\delta(k_{N-1},k_N)}
\end{aligned}$$

In order to compute this quantity, let us denote by F the part involving the double factorials, and by P the part involving the powers of $\pi/2$, so that we have:

$$I' = F\cdot P$$

Regarding F, there are many cancellations there, and we end up with:

$$F = \frac{k_1!!\ldots k_N!!}{(\Sigma k_i+N-1)!!}$$

As in what regards P, the δ exponents on the right sum up to the following number:

$$\Delta(k_1,\ldots,k_N) = \sum_{i=1}^{N-1}\delta(k_i,k_{i+1}+\ldots+k_N+N-i-1)$$

In other words, with this notation, the above formula reads:

$$\begin{aligned}
I' &= \left(\frac{\pi}{2}\right)^{N-1}\frac{k_1!!k_2!!\ldots k_N!!}{(k_1+\ldots+k_N+N-1)!!}\left(\frac{2}{\pi}\right)^{\Delta(k_1,\ldots,k_N)}\\
&= \left(\frac{2}{\pi}\right)^{\Delta(k_1,\ldots,k_N)-N+1}\frac{k_1!!k_2!!\ldots k_N!!}{(k_1+\ldots+k_N+N-1)!!}\\
&= \left(\frac{2}{\pi}\right)^{\Sigma(k_1,\ldots,k_N)-[N/2]}\frac{k_1!!k_2!!\ldots k_N!!}{(k_1+\ldots+k_N+N-1)!!}
\end{aligned}$$

Here the formula relating Δ to Σ follows from a number of simple observations, the first of which being the act that, due to obvious parity reasons, the sequence of δ numbers appearing in the definition of Δ cannot contain two consecutive zeroes. Together with $I=(2^N/V)I'$, this gives the formula in the statement. □

Summarizing, we have complete results for the integration over the spheres, with the answers involving various multinomial type coefficients, defined in terms of factorials, or of double factorials. All these formulae are of course very useful, in practice.

As a basic application of all this, we have the following result:

THEOREM 6.16. *The moments of the hyperspherical variables are*

$$\int_{S^{N-1}_\mathbb{R}} x_i^k dx = \frac{(N-1)!!k!!}{(N+k-1)!!}$$

and the normalized hyperspherical variables

$$y_i = \frac{x_i}{\sqrt{N}}$$

become normal and independent with $N \to \infty$.

PROOF. We have two things to be proved, the idea being as follows:

(1) The formula in the statement follows from the general integration formula over the sphere, established above. Indeed, this formula gives:

$$\int_{S^{N-1}_\mathbb{R}} x_i^k dx = \frac{(N-1)!!k!!}{(N+k-1)!!}$$

Now observe that with $N \to \infty$ we have the following estimate:

$$\begin{aligned}\int_{S^{N-1}_\mathbb{R}} x_i^k dx &= \frac{(N-1)!!}{(N+k-1)!!} \times k!! \\ &\simeq N^{k/2} k!! \\ &= N^{k/2} M_k(g_1)\end{aligned}$$

Thus, the variables $y_i = \frac{x_i}{\sqrt{N}}$ become normal with $N \to \infty$.

(2) As for the asymptotic independence result, this is standard as well, once again by using Theorem 6.14, for computing mixed moments, and taking the $N \to \infty$ limit. $\square$

As a comment here, all this might seem quite specialized. However, we will see later on that all this is related to linear algebra, and more specifically to the fine study of the group O_N formed by the orthogonal matrices. But more on this later.

6d. Complex spheres

Let us discuss now the complex analogues of all the above. We must first introduce the complex analogues of the normal laws, and this can be done as follows:

Definition 6.17. *The complex Gaussian law of parameter $t>0$ is*

$$G_t = law\left(\frac{1}{\sqrt{2}}(a+ib)\right)$$

where a,b are independent, each following the law g_t.

The combinatorics of these laws is a bit more complicated than in the real case, and we will be back to this in a moment. But to start with, we have:

Theorem 6.18. *The complex Gaussian laws have the property*

$$G_s * G_t = G_{s+t}$$

for any $s,t>0$, and so they form a convolution semigroup.

Proof. This follows indeed from the real result, for the usual Gaussian laws, established in above, by taking real and imaginary parts. □

We have as well the following complex analogue of the CLT:

Theorem 6.19 (CCLT). *Given complex random variables $f_1,f_2,f_3,\ldots\in L^\infty(X)$, which are i.i.d., centered, and with variance $t>0$, we have, with $n\to\infty$, in moments,*

$$\frac{1}{\sqrt{n}}\sum_{i=1}^n f_i \sim G_t$$

where G_t is the complex Gaussian law of parameter t.

Proof. This follows indeed from the real CLT, established above, simply by taking the real and imaginary parts of all the variables involved. □

Regarding now the moments, things are a bit more complicated than before, because our variables are now complex instead of real. In order to deal with this issue, we will use "colored moments", which are the expectations of the "colored powers", with these latter powers being defined by the following formulae, and multiplicativity:

$$f^\emptyset = 1 \quad , \quad f^\circ = f \quad , \quad f^\bullet = \bar{f}$$

With these conventions made, the result is as follows:

Theorem 6.20. *The moments of the complex normal law are the numbers*

$$M_k(G_t) = \sum_{\pi\in\mathcal{P}_2(k)} t^{|\pi|}$$

where $\mathcal{P}_2(k)$ are the matching pairings of $\{1,\ldots,k\}$, and $|.|$ is the number of blocks.

PROOF. This can be done in several steps, as follows:

(1) We recall from the above that the moments of the real Gaussian law g_1, with respect to integer exponents $k \in \mathbb{N}$, are the following numbers:

$$m_k = |P_2(k)|$$

(2) We will show here that in what concerns the complex Gaussian law G_1, a similar result holds. Numerically, we will prove that we have the following formula, where a colored integer $k = \circ \bullet \bullet \circ \ldots$ is called uniform when it contains the same number of $\circ$ and $\bullet$, and where $|k| \in \mathbb{N}$ is the length of such a colored integer:

$$M_k = \begin{cases} (|k|/2)! & (k \text{ uniform}) \\ 0 & (k \text{ not uniform}) \end{cases}$$

Now since the matching partitions $\pi \in \mathcal{P}_2(k)$ are counted by exactly the same numbers, and this for trivial reasons, we will obtain the formula in the statement, namely:

$$M_k = |\mathcal{P}_2(k)|$$

(3) This was for the plan. In practice now, we must compute the moments, with respect to colored integer exponents $k = \circ \bullet \bullet \circ \ldots$, of the variable in the statement:

$$c = \frac{1}{\sqrt{2}}(a + ib)$$

As a first observation, in the case where such an exponent $k = \circ\bullet\bullet\circ\ldots$ is not uniform in $\circ, \bullet$, a rotation argument shows that the corresponding moment of c vanishes. To be more precise, the variable $c' = wc$ can be shown to be complex Gaussian too, for any $w \in \mathbb{C}$, and from $M_k(c) = M_k(c')$ we obtain $M_k(c) = 0$, in this case.

(4) In the uniform case now, where $k = \circ\bullet\bullet\circ\ldots$ consists of p copies of $\circ$ and p copies of $\bullet$, the corresponding moment can be computed as follows:

$$\begin{aligned}
M_k &= \int (c\bar{c})^p \\
&= \frac{1}{2^p}\int (a^2+b^2)^p \\
&= \frac{1}{2^p}\sum_s \binom{p}{s}\int a^{2s}\int b^{2p-2s} \\
&= \frac{1}{2^p}\sum_s \binom{p}{s}(2s)!!(2p-2s)!! \\
&= \frac{1}{2^p}\sum_s \frac{p!}{s!(p-s)!}\cdot\frac{(2s)!}{2^s s!}\cdot\frac{(2p-2s)!}{2^{p-s}(p-s)!} \\
&= \frac{p!}{4^p}\sum_s \binom{2s}{s}\binom{2p-2s}{p-s}
\end{aligned}$$

(5) In order to finish now the computation, let us recall that we have the following formula, coming from the generalized binomial formula, or from the Taylor formula:

$$\frac{1}{\sqrt{1+t}} = \sum_{k=0}^{\infty}\binom{2k}{k}\left(\frac{-t}{4}\right)^k$$

By taking the square of this series, we obtain the following formula:

$$\begin{aligned}
\frac{1}{1+t} &= \sum_{ks}\binom{2k}{k}\binom{2s}{s}\left(\frac{-t}{4}\right)^{k+s} \\
&= \sum_p \left(\frac{-t}{4}\right)^p \sum_s \binom{2s}{s}\binom{2p-2s}{p-s}
\end{aligned}$$

Now by looking at the coefficient of t^p on both sides, we conclude that the sum on the right equals 4^p. Thus, we can finish the moment computation in (4), as follows:

$$M_p = \frac{p!}{4^p}\times 4^p = p!$$

(6) As a conclusion, if we denote by $|k|$ the length of a colored integer $k = \circ\bullet\bullet\circ\ldots$, the moments of the variable c in the statement are given by:

$$M_k = \begin{cases} (|k|/2)! & (k \text{ uniform}) \\ 0 & (k \text{ not uniform}) \end{cases}$$

On the other hand, the numbers $|\mathcal{P}_2(k)|$ in the statement are given by exactly the same formula. Indeed, in order to have matching pairings of k, our exponent $k = \circ\bullet\bullet\circ\ldots$ must be uniform, consisting of p copies of $\circ$ and p copies of $\bullet$, with:

$$p = \frac{|k|}{2}$$

But then the matching pairings of k correspond to the permutations of the $\bullet$ symbols, as to be matched with $\circ$ symbols, and so we have $p!$ such matching pairings. Thus, we have exactly the same formula as for the moments of c, and this finishes the proof. $\square$

There are of course many other possible proofs for the above result, which are all instructive, and some further theory as well, that can be developed for the complex normal variables, which is very interesting too. We refer here to Feller [**35**], or Durrett [**33**]. We will be back to this, on several occasions, in what follows.

In practice, we also need to know how to compute joint moments of independent normal variables. We have here the following result, to be heavily used later on:

THEOREM 6.21 (Wick formula). *Given independent variables f_i, each following the complex normal law G_t, with $t > 0$ being a fixed parameter, we have the formula*

$$\mathbb{E}\left(f_{i_1}^{k_1}\ldots f_{i_s}^{k_s}\right) = t^{s/2}\#\left\{\pi\in\mathcal{P}_2(k)\Big|\pi\leq\ker(i)\right\}$$

where $k = k_1\ldots k_s$ and $i = i_1\ldots i_s$, for the joint moments of these variables.

PROOF. This is something well-known, and the basis for all possible computations with complex normal variables, which can be proved in two steps, as follows:

(1) Let us first discuss the case where we have a single variable f, which amounts in taking $f_i = f$ for any i in the formula in the statement. What we have to compute here are the moments of f, with respect to colored integer exponents $k = \circ\bullet\bullet\circ\ldots$, and the formula in the statement tells us that these moments must be:

$$\mathbb{E}(f^k) = t^{|k|/2}|\mathcal{P}_2(k)|$$

But this is the formula in Theorem 6.20, so we are done with this case.

(2) In general now, when expanding the product $f_{i_1}^{k_1}\ldots f_{i_s}^{k_s}$ and rearranging the terms, we are left with doing a number of computations as in (1), and then making the product of the expectations that we found. But this amounts in counting the partitions in the statement, with the condition $\pi\leq\ker(i)$ there standing for the fact that we are doing the various type (1) computations independently, and then making the product. $\square$

The above statement is one of the possible formulations of the Wick formula, and there are in fact many more formulations, which are all useful. Here is an alternative such formulation, which is quite popular, and that we will often use in what follows:

THEOREM 6.22 (Wick formula 2). *Given independent variables f_i, each following the complex normal law G_t, with $t>0$ being a fixed parameter, we have the formula*

$$\mathbb{E}\left(f_{i_1}\ldots f_{i_k}f_{j_1}^*\ldots f_{j_k}^*\right)=t^k\#\left\{\pi\in S_k\Big|i_{\pi(r)}=j_r,\forall r\right\}$$

for the non-vanishing joint moments of these variables.

PROOF. This follows from the usual Wick formula, from Theorem 6.21. With some changes in the indices and notations, the formula there reads:

$$\mathbb{E}\left(f_{I_1}^{K_1}\ldots f_{I_s}^{K_s}\right)=t^{s/2}\#\left\{\sigma\in\mathcal{P}_2(K)\Big|\sigma\leq\ker(I)\right\}$$

Now observe that we have $\mathcal{P}_2(K)=\emptyset$, unless the colored integer $K=K_1\ldots K_s$ is uniform, in the sense that it contains the same number of $\circ$ and $\bullet$ symbols. Up to permutations, the non-trivial case, where the moment is non-vanishing, is the case where the colored integer $K=K_1\ldots K_s$ is of the following special form:

$$K=\underbrace{\circ\circ\ldots\circ}_k\ \underbrace{\bullet\bullet\ldots\bullet}_k$$

So, let us focus on this case, which is the non-trivial one. Here we have $s=2k$, and we can write the multi-index $I=I_1\ldots I_s$ in the following way:

$$I=i_1\ldots i_k\ j_1\ldots j_k$$

With these changes made, the above usual Wick formula reads:

$$\mathbb{E}\left(f_{i_1}\ldots f_{i_k}f_{j_1}^*\ldots f_{j_k}^*\right)=t^k\#\left\{\sigma\in\mathcal{P}_2(K)\Big|\sigma\leq\ker(ij)\right\}$$

The point now is that the matching pairings $\sigma\in\mathcal{P}_2(K)$, with $K=\circ\ldots\circ\bullet\ldots\bullet$, of length $2k$, as above, correspond to the permutations $\pi\in S_k$, in the obvious way. With this identification made, the above modified usual Wick formula becomes:

$$\mathbb{E}\left(f_{i_1}\ldots f_{i_k}f_{j_1}^*\ldots f_{j_k}^*\right)=t^k\#\left\{\pi\in S_k\Big|i_{\pi(r)}=j_r,\forall r\right\}$$

Thus, we have reached to the formula in the statement, and we are done. □

Finally, here is one more formulation of the Wick formula, which is useful as well:

THEOREM 6.23 (Wick formula 3). *Given independent variables f_i, each following the complex normal law G_t, with $t>0$ being a fixed parameter, we have the formula*

$$\mathbb{E}\left(f_{i_1}f_{j_1}^*\ldots f_{i_k}f_{j_k}^*\right)=t^k\#\left\{\pi\in S_k\Big|i_{\pi(r)}=j_r,\forall r\right\}$$

for the non-vanishing joint moments of these variables.

PROOF. This follows from our second Wick formula, from Theorem 6.22, simply by permuting the terms, as to have an alternating sequence of plain and conjugate variables. Alternatively, we can start with Theorem 6.21, and then perform the same manipulations as in the proof of Theorem 6.22, but with the exponent being this time as follows:

$$K = \underbrace{\circ\bullet\circ\bullet\ldots\ldots\circ\bullet}_{2k}$$

Thus, we are led to the conclusion in the statement. □

In relation now with the spheres, we first have the following variation of the integration formula in Theorem 6.15, dealing this time with integrals over the complex sphere:

THEOREM 6.24. *We have the following integration formula over the complex sphere $S^{N-1}_\mathbb{C} \subset \mathbb{R}^N$, with respect to the normalized measure,*

$$\int_{S^{N-1}_\mathbb{C}} |z_1|^{2l_1}\ldots|z_N|^{2l_N}\,dz = 4^{\sum l_i}\frac{(2N-1)!l_1!\ldots l_n!}{(2N+\sum l_i-1)!}$$

valid for any exponents $l_i \in \mathbb{N}$. As for the other polynomial integrals in $z_1,\ldots,z_N$ and their conjugates $\bar{z}_1,\ldots,\bar{z}_N$, these all vanish.

PROOF. Consider an arbitrary polynomial integral over $S^{N-1}_\mathbb{C}$, written as follows:

$$I = \int_{S^{N-1}_\mathbb{C}} z_{i_1}\bar{z}_{i_2}\ldots z_{i_{2l-1}}\bar{z}_{i_{2l}}\,dz$$

(1) By using transformations of type $p \to \lambda p$ with $|\lambda| = 1$, we see that I vanishes, unless each z_a appears as many times as $\bar{z}_a$ does, and this gives the last assertion.

(2) Assume now that we are in the non-vanishing case. Then the l_a copies of z_a and the l_a copies of $\bar{z}_a$ produce by multiplication a factor $|z_a|^{2l_a}$, so we have:

$$I = \int_{S^{N-1}_\mathbb{C}} |z_1|^{2l_1}\ldots|z_N|^{2l_N}\,dz$$

Now by using the standard identification $S^{N-1}_\mathbb{C} \simeq S^{2N-1}_\mathbb{R}$, we obtain:

$$\begin{aligned} I &= \int_{S^{2N-1}_\mathbb{R}} (x_1^2+y_1^2)^{l_1}\ldots(x_N^2+y_N^2)^{l_N}\,d(x,y)\\ &= \sum_{r_1\ldots r_N}\binom{l_1}{r_1}\ldots\binom{l_N}{r_N}\int_{S^{2N-1}_\mathbb{R}} x_1^{2l_1-2r_1}y_1^{2r_1}\ldots x_N^{2l_N-2r_N}y_N^{2r_N}\,d(x,y) \end{aligned}$$

(3) By using the formula in Theorem 6.14, we obtain:

$$\begin{aligned}
I &= \sum_{r_1\ldots r_N}\binom{l_1}{r_1}\ldots\binom{l_N}{r_N}\frac{(2N-1)!!(2r_1)!!\ldots(2r_N)!!(2l_1-2r_1)!!\ldots(2l_N-2r_N)!!}{(2N+2\sum l_i-1)!!}\\
&= \sum_{r_1\ldots r_N}\binom{l_1}{r_1}\ldots\binom{l_N}{r_N}\frac{(2N-1)!(2r_1)!\ldots(2r_N)!(2l_1-2r_1)!\ldots(2l_N-2r_N)!}{(2N+\sum l_i-1)!r_1!\ldots r_N!(l_1-r_1)!\ldots(l_N-r_N)!}
\end{aligned}$$

(4) We can rewrite the sum on the right in the following way:

$$\begin{aligned}
I &= \sum_{r_1\ldots r_N}\frac{l_1!\ldots l_N!(2N-1)!(2r_1)!\ldots(2r_N)!(2l_1-2r_1)!\ldots(2l_N-2r_N)!}{(2N+\sum l_i-1)!(r_1!\ldots r_N!(l_1-r_1)!\ldots(l_N-r_N)!)^2}\\
&= \sum_{r_1}\binom{2r_1}{r_1}\binom{2l_1-2r_1}{l_1-r_1}\ldots\sum_{r_N}\binom{2r_N}{r_N}\binom{2l_N-2r_N}{l_N-r_N}\frac{(2N-1)!l_1!\ldots l_N!}{(2N+\sum l_i-1)!}\\
&= 4^{l_1}\times\ldots\times 4^{l_N}\times\frac{(2N-1)!l_1!\ldots l_N!}{(2N+\sum l_i-1)!}
\end{aligned}$$

Thus, we obtain the formula in the statement. □

Regarding now the hyperspherical variables, investigated in the above in the real case, we have similar results for the complex spheres, as follows:

THEOREM 6.25. *The rescaled coordinates on the complex sphere $S^{N-1}_\mathbb{C}$,*

$$w_i = \frac{z_i}{\sqrt{N}}$$

become complex Gaussian and independent with $N\to\infty$.

PROOF. We have two assertions to be proved, the idea being as follows:

(1) The assertion about the laws follows as in the real case, by using this time Theorem 6.24 as a main technical ingredient.

(2) As for the independence result, this follows as well as in the real case, by using this time the Wick formula as a main technical ingredient. □

As a conclusion to all this, we have now a good level in linear algebra, and probability. And this can only open up a whole new set of perspectives, on what further books can be read. Here is a guide, to what can be learned, as a continuation:

(1) Before anything, talking analysis, you should imperatively learn some Fourier analysis and differential equations, topics not covered in this book. With good places here being the books of Arnold [1], [2], [3], [4], Evans [34] and Rudin [74], [75].

(2) You can also embark on the reading of some tough, ultimate physics, say from Landau-Lifshitz [58], [59], [60], [61]. Normally these are books which are quite strong on mathematics, but with your knowledge from here, you will certainly survive.

(3) You can also start reading whatever fancy physics that you might like, such as astrophysics and cosmology from Ryden [76], [77] or Weinberg [89], [90], quantum information from Bengtsson-Życzkowski [14] or Nielsen-Chuang [71], and so on.

(4) And for algebra and probability, stay with us. The story is far from being over with what we learned, and dozens of further interesting things to follow. We still have 250 more pages, and there will be algebra and probability in them, that is promised.

6e. Exercises

We have learned many interesting things in this chapter, and there are many exercises about this. First, in connection with the CLT, we have:

EXERCISE 6.26. *Work out the precise convergence conclusions in the CLT,*

$$\frac{1}{\sqrt{n}}\sum_{i=1}^n f_i \sim g_t$$

going beyond the convergence in moments, which was established in the above.

This is a bit vague, but at this stage, learning more theory would be a good thing. Of course, in case all this looks a bit complicated, don't hesitate to look it up.

EXERCISE 6.27. *Find an alternative proof for the moment formula*

$$M_k(G_t)=\sum_{\pi\in\mathcal{P}_2(k)}t^{|\pi|}$$

using a method of your choice.

Again, this is a bit vague, and many things that you can try. As before, in case you lack a new idea here, don't hesitate to look it up, and report on what you learned.

EXERCISE 6.28. *Find a probability measure μ whose moments are given by*

$$M_k(\mu)=|P(k)|$$

then more generally find a measure μ_t whose moments are given by

$$M_k(\mu_t)=\sum_{\pi\in P(k)}t^{|\pi|}$$

where $P(k)$ stands as usual for all partitions of $\{1,\ldots,k\}$.

This is something quite tricky, and in case you do not find, do not worry. We will be back to such questions, which are quite fundamental, later in this book.

EXERCISE 6.29. *Find a probability measure ν whose moments are given by*

$$M_k(\nu) = |NC_2(k)|$$

then find as well a probability measure η whose moments are given by

$$M_k(\eta) = |NC(k)|$$

where NC stands for "noncrossing". Then try as well the parametric case.

As before with the previous exercise, this is something quite fundamental, and we will be back to this, later in this book. And to end this series of exercises, for a bonus point, solve as well the question left, regarding this time the noncrossing matching pairings. With this latter question being actually quite difficult, but worth studying.

EXERCISE 6.30. *Prove, with full details, that the rescaled coordinates*

$$y_i = \frac{x_i}{\sqrt{N}}$$

become independent with $N \to \infty$, for both the real and complex spheres.

To be more precise here, we have proved in the above that the rescaled coordinates become respectively real and complex Gaussian, in the $N \to \infty$ limit, for the real and complex spheres. The problem now is that of using the same method, namely a straightforward application of our spherical integral formulae, in order to compute and then estimate the joint moments, as to prove the independence result as well.

EXERCISE 6.31. *Compute the density of the hyperspherical law at $N = 4$, that is, the law of one of the coordinates over the unit sphere $S^3_{\mathbb{R}} \subset \mathbb{R}^4$.*

If you find something very interesting, as an answer here, do not be surprised. After all, $S^3_{\mathbb{R}}$ is the sphere of space-time, having its own magic. We will be back to this.

CHAPTER 7

Special matrices

7a. Fourier matrices

In this chapter we go back to basic linear algebra, and we discuss a number of more specialized questions. We will be interested in various classes of "special matrices", and in the tools for dealing with them. Our motivations are as follows:

(1) There are many interesting groups of matrices $G \subset U_N$, that we will study later in this book, and the individual elements $U \in G$, which are defined by some particular conditions, can be usually thought of as being "special matrices". Thus, our study of the special matrices will be a good introduction to the groups of matrices $G \subset U_N$.

(2) Less abstractly, we will see that the study of the special matrices leads us into Fourier analysis, and more specifically into "discrete Fourier analysis". Although nothing beats classical Fourier analysis, that can be learned for instance from Rudin [**74**], [**75**], having something here on discrete Fourier analysis will be certainly a good thing.

(3) Our study will come as well as a natural continuation of the linear algebra theory developed in chapters 1-4. We have seen there the foundations of the theory, basically reducing everything to the study of the matrices $A \in M_N(\mathbb{C})$, and now if we want to learn more, we will certainly have to assume that our matrices A are of "special type".

(4) Finally, and totally away now from any abstraction, the special matrices that we will study here are related to all sorts of interesting physics and applications, such as quantum groups, quantum information, random matrices, coding theory, and many more. So, we will learn here interesting and useful theory, that can be applied afterwards.

Getting started now, and for having a taste of what we want to do, in this chapter, as a first and central example of a special matrix, we have the flat matrix:

Definition 7.1. *The flat matrix $\mathbb{I}_N$ is the all-one $N \times N$ matrix:*

$$\mathbb{I}_N = \begin{pmatrix} 1 & \dots & 1 \\ \vdots & & \vdots \\ 1 & \dots & 1 \end{pmatrix}$$

Equivalently, $\mathbb{I}_N/N$ is the orthogonal projection on the all-one vector $\xi \in \mathbb{C}^N$.

Observe that $\mathbb{I}_N$ has a lot of interesting properties, such as being circulant, and bistochastic. The idea will be that many techniques that can be applied to $\mathbb{I}_N$, with quite trivial results, apply to such special classes of matrices, with non-trivial consequences.

A first interesting question regarding $\mathbb{I}_N$ concerns its diagonalization. Since $\mathbb{I}_N$ is a multiple of a rank 1 projection, we have right away the following result:

Proposition 7.2. *The flat matrix $\mathbb{I}_N$ diagonalizes as follows,*

$$\begin{pmatrix}1&\ldots&1\\ \vdots&&\vdots\\ 1&\ldots&1\end{pmatrix}=P\begin{pmatrix}N&&&\\ &0&&\\ &&\ddots&\\ &&&0\end{pmatrix}P^{-1}$$

where $P\in M_N(\mathbb{C})$ can be any matrix formed by the all one-vector ξ, followed by $N-1$ linearly independent solutions $x\in\mathbb{C}^N$ of the following equation:

$$x_1+\ldots+x_N=0$$

Proof. This follows indeed from our linear algebra knowledge from chapters 1-4, by using the fact that $\mathbb{I}_N/N$ is the orthogonal projection onto $\mathbb{C}\xi$. □

In practice now, the problem which is left is that of finding an explicit matrix $P\in M_N(\mathbb{C})$, as above. To be more precise, there are plently of solutions here, some of them being even real, $P\in M_N(\mathbb{R})$, and the problem is that of finding a "nice" such solution, say having the property that P_{ij} appears as an explicit function of i,j.

Long story short, we are led to the question of solving, in a somewhat canonical and elegant way, the following equation, over the real or the complex numbers:

$$x_1+\ldots+x_N=0$$

And this is more tricky than it seems. To be more precise, there is no hope of doing this over the real numbers. As in what regards the complex numbers, there is a ray of light here coming from the roots of unity, and more specifically, from:

Proposition 7.3. *We have the formula*

$$\frac{1}{N}\sum_{k=0}^{N-1}z^k=\begin{cases}0&\text{if } z\neq 1\\ 1&\text{if } z=1\end{cases}$$

valid for any N-th root of unity z.

Proof. This is something that we know from chapter 3, coming from the fact that the average in the statement computes the barycenter of the polygon formed by the numbers $1,z,z^2,\ldots,z^{N-1}$, in the complex plane. Indeed, since this polygon is regular and centered at 0, the barycenter in question is obviously 0, and this unless we are in the case $z=1$, where the polygon degenerates, and the barycenter is obviously 1. □

Summarizing, we have a beginning of an answer to our question, with the idea being that of using a matrix $P \in M_N(\mathbb{C})$ formed by the N-th roots of unity. To be more precise, this leads us into the Fourier matrix, which is as follows:

Definition 7.4. *The Fourier matrix F_N is the following matrix, with $w = e^{2\pi i/N}$:*

$$F_N = \begin{pmatrix} 1 & 1 & 1 & \ldots & 1 \\ 1 & w & w^2 & \ldots & w^{N-1} \\ 1 & w^2 & w^4 & \ldots & w^{2(N-1)} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & w^{N-1} & w^{2(N-1)} & \ldots & w^{(N-1)^2} \end{pmatrix}$$

That is, $F_N = (w^{ij})_{ij}$, with indices $i, j \in \{0, 1, \ldots, N-1\}$, taken modulo N.

Before getting further, observe that this matrix F_N is "special" as well, but in a different sense, its main properties being the fact that it is a Vandermonde matrix, and a rescaled unitary. We will axiomatize later on the matrices of this type.

Getting back now to the diagonalization problem for the flat matrix $\mathbb{I}_N$, this can be solved by using the Fourier matrix F_N, in the following elegant way:

Theorem 7.5. *The flat matrix $\mathbb{I}_N$ diagonalizes as follows,*

$$\begin{pmatrix} 1 & \ldots & 1 \\ \vdots & & \vdots \\ 1 & \ldots & 1 \end{pmatrix} = \frac{1}{N} F_N \begin{pmatrix} N & & & \\ & 0 & & \\ & & \ddots & \\ & & & 0 \end{pmatrix} F_N^*$$

with $F_N = (w^{ij})_{ij}$ being the Fourier matrix.

Proof. According to Proposition 7.2, we are left with finding the 0-eigenvectors of $\mathbb{I}_N$, which amounts in solving the following equation:

$$x_0 + \ldots + x_{N-1} = 0$$

For this purpose, we use the root of unity $w = e^{2\pi i/N}$, and more specifically, the following standard formula, coming from Proposition 7.3:

$$\sum_{i=0}^{N-1} w^{ij} = N\delta_{j0}$$

This formula shows that for $j = 1, \ldots, N-1$, the vector $v_j = (w^{ij})_i$ is a 0-eigenvector. Moreover, these vectors are pairwise orthogonal, because we have:

$$\begin{aligned} < v_j, v_k > &= \sum_i w^{ij-ik} \\ &= N\delta_{jk} \end{aligned}$$

Thus, we have our basis $\{v_1,\ldots,v_{N-1}\}$ of 0-eigenvectors, and since the N-eigenvector is $\xi=v_0$, the passage matrix P that we are looking is given by:

$$P=\begin{bmatrix}v_0 & v_1 & \ldots & v_{N-1}\end{bmatrix}$$

But this is precisely the Fourier matrix, $P=F_N$. In order to finish now, observe that the above computation of $<v_i,v_j>$ shows that $F_N/\sqrt{N}$ is unitary, and so:

$$F_N^{-1}=\frac{1}{N}\,F_N^*$$

Thus, we are led to the diagonalization formula in the statement. □

Generally speaking, the above result will be the template for what we will be doing here. On one hand we will have special matrices to be studied, of $\mathbb{I}_N$ type, and on the other hand we will have special matrices that can be used as tools, of F_N type.

Let us begin with a discussion of the "tools". The Fourier matrix F_N certainly has many interesting properties, but what really stands out is the fact that its entries are on the unit circle, and its columns are pairwise orthogonal. Which leads us into:

Definition 7.6. *A complex Hadamard matrix is a square matrix*

$$H\in M_N(\mathbb{T})$$

where $\mathbb{T}$ is the unit circle, satisfying the following equivalent conditions:

(1) *The rows are pairwise orthogonal.*
(2) *The columns are pairwise orthogonal.*
(3) *The rescaled matrix $H/\sqrt{N}$ is unitary.*
(4) *The rescaled matrix $H^t/\sqrt{N}$ is unitary.*

Here the fact that the above conditions are equivalent comes from basic linear algebra, and more specifically from the fact that a matrix $U\in M_N(\mathbb{C})$ is a unitary precisely when the rows, or the columns, have norm 1, and are pairwise orthogonal.

We already know, from the proof of Theorem 7.5, that the Fourier matrix F_N is a complex Hadamard matrix. There are many other examples of complex Hadamard matrices, and the basic theory of such matrices can be summarized as follows:

Theorem 7.7. *The class of the $N\times N$ complex Hadamard matrices is as follows:*

(1) *It contains the Fourier matrix F_N.*
(2) *It is stable under taking tensor products.*
(3) *It is stable under taking transposes, conjugates and adjoints.*
(4) *It is stable under permuting rows, or permuting columns.*
(5) *It is stable under multiplying rows or columns by numbers in $\mathbb{T}$.*

PROOF. All this is elementary, the idea being as follows:

(1) This is something that we already know, from the proof of Theorem 7.5, with the orthogonality coming from the following standard formula:

$$\sum_i w^{ij} = N\delta_{j0}$$

(2) Assume that $H \in M_M(\mathbb{T})$ and $K \in M_N(\mathbb{T})$ are Hadamard matrices, and consider their tensor product, which in double index notation is as follows:

$$(H \otimes K)_{ia,jb} = H_{ij}K_{ab}$$

We have then $H \otimes K \in M_{MN}(\mathbb{T})$, and the rows R_{ia} of this matrix are pairwise orthogonal, as shown by the following computation:

$$\begin{aligned} < R_{ia}, R_{kc} > &= \sum_{jb} H_{ij}K_{ab} \cdot \bar{H}_{kj}\bar{K}_{cb} \\ &= \sum_j H_{ij}\bar{H}_{kj} \sum_b K_{ab}\bar{K}_{cb} \\ &= MN\delta_{ik}\delta_{ac} \end{aligned}$$

(3) We know that the set formed by the $N \times N$ complex Hadamard matrices appears as follows, with the intersection being taken inside $M_N(\mathbb{C})$:

$$X_N = M_N(\mathbb{T}) \cap \sqrt{N}U_N$$

The set $M_N(\mathbb{T})$ is stable under the operations in the statement. As for the set $\sqrt{N}U_N$, here we can use the well-known fact that if a matrix is unitary, $U \in U_N$, then so is its complex conjugate $\bar{U} = (\bar{U}_{ij})$, the inversion formulae being as follows:

$$U^* = U^{-1} \quad , \quad U^t = \bar{U}^{-1}$$

Thus the unitary group U_N is stable under the following operations:

$$U \to U^t \quad , \quad U \to \bar{U} \quad , \quad U \to U^*$$

It follows that the above set X_N is stable as well under these operations, as desired.

(4) This is clear from definitions, because permuting rows or permuting columns leaves invariant both the sets $M_N(\mathbb{T})$ and $\sqrt{N}U_N$.

(5) This is once again clear from definitions, because multiplying rows or columns by numbers in $\mathbb{T}$ leaves invariant both the sets $M_N(\mathbb{T})$ and $\sqrt{N}U_N$. $\square$

In the above result (1,2) are really important, and (3,4,5) are rather technical remarks. As a consequence, coming from (1,2), let us formulate:

THEOREM 7.8. *The following matrices, called generalized Fourier matrices,*

$$F_{N_1,\ldots,N_k}=F_{N_1}\otimes\ldots\otimes F_{N_k}$$

are Hadamard, for any choice of $N_1,\ldots,N_k$. In particular the following matrices,

$$W_N=\begin{pmatrix}1&1\\1&-1\end{pmatrix}^{\otimes k}$$

having size $N=2^k$, and called Walsh matrices, are all Hadamard.

PROOF. This is a consequence of Theorem 7.7, as follows:

(1) The first assertion follows from (1,2) in Theorem 7.7.

(2) The second assertion is a consequence of (1), by taking $N_1=\ldots=N_k=2$. Indeed, the generalized Fourier matrix that we obtain is:

$$F_{2,\ldots,2}=F_2^{\otimes k}=\begin{pmatrix}1&1\\1&-1\end{pmatrix}^{\otimes k}$$

Thus, we are led to the conclusion in the statement. □

As an illustration for the above result, the second Walsh matrix, which is an Hadamard matrix having real entries, as is the case with all the Walsh matrices, is as follows:

$$W_4=\begin{pmatrix}1&1&1&1\\1&-1&1&-1\\1&1&-1&-1\\1&-1&-1&1\end{pmatrix}$$

The generalized Fourier matrices, and in particular the Walsh matrices, have many applications, to questions in coding, radio transmissions, quantum physics, and many more. We refer here for instance to the book of Bengtsson-Życzkowski [**14**], and to the papers of Björck [**15**], Haagerup [**45**], Idel-Wolf [**51**], Jones [**56**], Sylvester [**82**]. We will be back to these matrices later, on several occasions.

7b. Circulant matrices

Let us go back now to the general linear algebra considerations from the beginning of this chapter. We have seen that F_N diagonalizes in an elegant way the flat matrix $\mathbb{I}_N$, and the idea in what follows will be that of F_N, or other real or complex Hadamard matrices, can be used in order to deal with other matrices, of $\mathbb{I}_N$ type.

A first feature of the flat matrix $\mathbb{I}_N$ is that it is circulant, in the following sense:

DEFINITION 7.9. *A real or complex matrix M is called circulant if*

$$M_{ij}=\xi_{j-i}$$

for a certain vector ξ, with the indices taken modulo N.

The circulant matrices are beautiful mathematical objects, which appear of course in many serious problems as well. As an example, at $N = 4$, we must have:

$$M = \begin{pmatrix} a & b & c & d \\ d & a & b & c \\ c & d & a & b \\ b & c & d & a \end{pmatrix}$$

The point now is that, while certainly gently looking, these matrices can be quite diabolic, when it comes to diagonalization, and other problems. For instance, when M is real, the computations with M are usually very complicated over the real numbers. Fortunately the complex numbers and the Fourier matrices are there, and we have:

THEOREM 7.10. *For a matrix $M \in M_N(\mathbb{C})$, the following are equivalent:*

(1) *M is circulant, in the sense that we have, for a certain vector $\xi \in \mathbb{C}^N$:*

$$M_{ij} = \xi_{j-i}$$

(2) *M is Fourier-diagonal, in the sense that, for a certain diagonal matrix Q:*

$$M = F_N Q F_N^*$$

In addition, if these conditions hold, then ξ, Q are related by the formula

$$\xi = F_N^* q$$

where $q \in \mathbb{C}^N$ is the column vector formed by the diagonal entries of Q.

PROOF. This follows from some basic computations with roots of unity, as follows:

(1) $\implies$ (2) Assuming $M_{ij} = \xi_{j-i}$, the matrix $Q = F_N^* M F_N$ is indeed diagonal, as shown by the following computation:

$$\begin{aligned} Q_{ij} &= \sum_{kl} w^{-ik} M_{kl} w^{lj} \\ &= \sum_{kl} w^{jl-ik} \xi_{l-k} \\ &= \sum_{kr} w^{j(k+r)-ik} \xi_r \\ &= \sum_r w^{jr} \xi_r \sum_k w^{(j-i)k} \\ &= N\delta_{ij} \sum_r w^{jr} \xi_r \end{aligned}$$

(2) $\implies$ (1) Assuming $Q = diag(q_1, \ldots, q_N)$, the matrix $M = F_N Q F_N^*$ is indeed circulant, as shown by the following computation:

$$\begin{aligned} M_{ij} &= \sum_k w^{ik} Q_{kk} w^{-jk} \\ &= \sum_k w^{(i-j)k} q_k \end{aligned}$$

To be more precise, in this formula the last term depends only on $j - i$, and so shows that we have $M_{ij} = \xi_{j-i}$, with ξ being the following vector:

$$\begin{aligned} \xi_i &= \sum_k w^{-ik} q_k \\ &= (F_N^* q)_i \end{aligned}$$

Thus, we are led to the conclusions in the statement. $\square$

As a basic illustration for the above result, for the circulant matrix $M = \mathbb{I}_N$ we recover in this way the diagonalization result from Theorem 7.5, namely:

$$\begin{pmatrix} 1 & \ldots & 1 \\ \vdots & & \vdots \\ 1 & \ldots & 1 \end{pmatrix} = \frac{1}{N} F_N \begin{pmatrix} N &&& \\ & 0 && \\ && \ddots & \\ &&& 0 \end{pmatrix} F_N^*$$

The above result is something quite powerful, and very useful, and suggests doing everything in Fourier, when dealing with circulant matrices. And we can use here:

THEOREM 7.11. *The various basic sets of $N \times N$ circulant matrices are as follows, with the convention that associated to any $q \in \mathbb{C}^N$ is the matrix $Q = diag(q_1, \ldots, q_N)$:*

(1) *The set of all circulant matrices is:*

$$M_N(\mathbb{C})^{circ} = \left\{ F_N Q F_N^* \Big| q \in \mathbb{C}^N \right\}$$

(2) *The set of all circulant unitary matrices is:*

$$U_N^{circ} = \left\{ \frac{1}{N} F_N Q F_N^* \Big| q \in \mathbb{T}^N \right\}$$

(3) *The set of all circulant orthogonal matrices is:*

$$O_N^{circ} = \left\{ \frac{1}{N} F_N Q F_N^* \Big| q \in \mathbb{T}^N, \bar{q}_i = q_{-i}, \forall i \right\}$$

In addition, in this picture, the first row vector of $F_N Q F_N^*$ is given by $\xi = F_N^* q$.

PROOF. All this follows from Theorem 7.10, as follows:

(1) This assertion, along with the last one, is Theorem 7.10 itself.

(2) This is clear from (1), and from the fact that the rescaled matrix $F_N/\sqrt{N}$ is unitary, because the eigenvalues of a unitary matrix must be on the unit circle $\mathbb{T}$.

(3) This follows from (2), because the matrix is real when $\xi_i = \bar{\xi}_i$, and in Fourier transform, $\xi = F_N^* q$, this corresponds to the condition $\bar{q}_i = q_{-i}$. □

There are many other things that can be said about the circulant matrices, along these lines. Importantly, all this can be generalized to the setting of the matrices which are $(N_1, \ldots, N_k)$ patterned, in a certain technical sense, and the matrix which does the job here is the corresponding generalized Fourier matrix, namely:

$$F_{N_1,\ldots,N_k} = F_{N_1} \otimes \ldots \otimes F_{N_k}$$

We will be back to this later, towards the end of the present book.

As a last topic regarding the circulant matrices, which is somehow one level above the above considerations, let us discuss the circulant Hadamard matrices. It is convenient to use the following notion, coming from the various results in Theorem 7.7:

DEFINITION 7.12. *Two complex Hadamard matrices are called equivalent, and we write $H \sim K$, when it is possible to pass from H to K via the following operations:*
1. *Permuting the rows, or permuting the columns.*
2. *Multiplying the rows or columns by numbers in $\mathbb{T}$.*

Here we have not taken into account all the results in Theorem 7.7 when formulating the above definition, because the operations $H \to H^t, \bar{H}, H^*$ are far more subtle than those in (1,2) above, and can complicate things, if included in the equivalence.

The point now is that, up to equivalence, we can put the Fourier matrix F_N in circulant form. At small values of N, this can be done as follows:

PROPOSITION 7.13. *The following are circulant and symmetric Hadamard matrices,*

$$F_2' = \begin{pmatrix} i & 1 \\ 1 & i \end{pmatrix} \quad , \quad F_3' = \begin{pmatrix} w & 1 & 1 \\ 1 & w & 1 \\ 1 & 1 & w \end{pmatrix}$$

$$F_4'' = \begin{pmatrix} -1 & \nu & 1 & \nu \\ \nu & -1 & \nu & 1 \\ 1 & \nu & -1 & \nu \\ \nu & 1 & \nu & -1 \end{pmatrix}$$

where $w = e^{2\pi i/3}, \nu = e^{\pi i/4}$, equivalent to the Fourier matrices F_2, F_3, F_4.

PROOF. The orthogonality between rows being clear, we have here complex Hadamard matrices. The fact that we have an equivalence $F_2 \sim F_2'$ follows from:

$$\begin{pmatrix}1&1\\1&-1\end{pmatrix}\sim\begin{pmatrix}i&i\\1&-1\end{pmatrix}\sim\begin{pmatrix}i&1\\1&i\end{pmatrix}$$

At $N=3$ now, the equivalence $F_3\sim F_3'$ can be constructed as follows:

$$\begin{pmatrix}1&1&1\\1&w&w^2\\1&w^2&w\end{pmatrix}\sim\begin{pmatrix}1&1&w\\1&w&1\\w&1&1\end{pmatrix}\sim\begin{pmatrix}w&1&1\\1&w&1\\1&1&w\end{pmatrix}$$

As for the case $N=4$, here the equivalence $F_4\sim F_4''$ can be constructed as follows, where we use the logarithmic notation $[k]_s=e^{2\pi ki/s}$, with respect to $s=8$:

$$\begin{bmatrix}0&0&0&0\\0&2&4&6\\0&4&0&4\\0&6&4&2\end{bmatrix}_8\sim\begin{bmatrix}0&1&4&1\\1&4&1&0\\4&1&0&1\\1&0&1&4\end{bmatrix}_8\sim\begin{bmatrix}4&1&0&1\\1&4&1&0\\0&1&4&1\\1&0&1&4\end{bmatrix}_8$$

Thus, the Fourier matrices F_2,F_3,F_4 can be put indeed in circulant form. □

We will explain later the reasons for denoting the above matrix F_4'', instead of F_4', the idea being that F_4' will be a matrix belonging to a certain series.

In order to discuss now the general case, we will use a technical method for dealing with the circulant matrices, namely Björck's cyclic root formalism [15], as follows:

THEOREM 7.14. *Assume that a matrix $H\in M_N(\mathbb{T})$ is circulant, $H_{ij}=\gamma_{j-i}$. Then H is is a complex Hadamard matrix if and only if the vector*

$$z=(z_0,z_1,\ldots,z_{N-1})$$

given by $z_i=\gamma_i/\gamma_{i-1}$ satisfies the following equations:

$$\begin{aligned}z_0+z_1+\ldots+z_{N-1}&=0\\z_0z_1+z_1z_2+\ldots+z_{N-1}z_0&=0\\&\ldots\\z_0z_1\ldots z_{N-2}+\ldots+z_{N-1}z_0\ldots z_{N-3}&=0\\z_0z_1\ldots z_{N-1}&=1\end{aligned}$$

If so is the case, we say that $z=(z_0,\ldots,z_{N-1})$ is a cyclic N-root.

PROOF. This follows from a direct computation, the idea being that, with $H_{ij}=\gamma_{j-i}$ as above, the orthogonality conditions between the rows are best written in terms of the variables $z_i=\gamma_i/\gamma_{i-1}$, and correspond to the equations in the statement. □

Now back to the Fourier matrices, we have the following result:

THEOREM 7.15. *Given $N \in \mathbb{N}$, construct the following complex numbers:*

$$\nu = e^{\pi i/N} \quad , \quad q = \nu^{N-1} \quad , \quad w = \nu^2$$

We have then a cyclic N-root as follows,

$$(q, qw, qw^2, \ldots, qw^{N-1})$$

and the corresponding complex Hadamard matrix F_N' is circulant and symmetric, and equivalent to the Fourier matrix F_N.

PROOF. Given two numbers $q, w \in \mathbb{T}$, let us find out when $(q, qw, qw^2, \ldots, qw^{N-1})$ is a cyclic root. We have two conditions to be verified, as follows:

(1) In order for the $= 0$ equations in Theorem 7.14 to be satisfied, the value of q is irrelevant, and w must be a primitive N-root of unity.

(2) As for the $= 1$ equation in Theorem 7.14, this states that we must have:

$$q^N w^{\frac{N(N-1)}{2}} = 1$$

Thus, we must have $q^N = (-1)^{N-1}$, so with the values of $q, w \in \mathbb{T}$ in the statement, we have a cyclic N-root. Now construct $H_{ij} = \gamma_{j-i}$ as in Theorem 7.14. We have:

$$\begin{aligned}
\gamma_k = \gamma_{-k} &\iff q^{k+1} w^{\frac{k(k+1)}{2}} = q^{-k+1} w^{\frac{k(k-1)}{2}} \\
&\iff q^{2k} w^k = 1 \\
&\iff q^2 = w^{-1}
\end{aligned}$$

But this latter condition holds indeed, because we have:

$$q^2 = \nu^{2N-2} = \nu^{-2} = w^{-1}$$

We conclude that our circulant matrix H is symmetric as well, as claimed. It remains to construct an equivalence $H \sim F_N$. In order to do this, observe that, due to our conventions $q = \nu^{N-1}, w = \nu^2$, the first row vector of H is given by:

$$\begin{aligned}
\gamma_k &= q^{k+1} w^{\frac{k(k+1)}{2}} \\
&= \nu^{(N-1)(k+1)} \nu^{k(k+1)} \\
&= \nu^{(N+k-1)(k+1)}
\end{aligned}$$

Thus, the entries of H are given by the following formula:

$$\begin{aligned}
H_{-i,j} &= H_{0,i+j} \\
&= \nu^{(N+i+j-1)(i+j+1)} \\
&= \nu^{i^2+j^2+2ij+Ni+Nj+N-1} \\
&= \nu^{N-1} \cdot \nu^{i^2+Ni} \cdot \nu^{j^2+Nj} \cdot \nu^{2ij}
\end{aligned}$$

With this formula in hand, we can now finish. Indeed, the matrix $H=(H_{ij})$ is equivalent to the following matrix:

$$H'=(H_{-i,j})$$

Now regarding this latter matrix H', observe that in the above formula, the factors ν^{N-1}, ν^{i^2+Ni}, ν^{j^2+Nj} correspond respectively to a global multiplication by a scalar, and to row and column multiplications by scalars. Thus H' is equivalent to the matrix H'' obtained from it by deleting these factors. But this latter matrix, given by $H''_{ij}=\nu^{2ij}$ with $\nu=e^{\pi i/N}$, is precisely the Fourier matrix F_N, and we are done. $\square$

As an illustration, at $N=2,3$ we obtain the old matrices F_2',F_3'. As for the case $N=4$, here we obtain the following matrix, with $\nu=e^{\pi i/4}$:

$$F_4'=\begin{pmatrix}\nu^3&1&\nu^7&1\\ 1&\nu^3&1&\nu^7\\ \nu^7&1&\nu^3&1\\ 1&\nu^7&1&\nu^3\end{pmatrix}$$

This matrix is equivalent to the matrix F_4'' from Proposition 7.13, with the equivalence $F_4'\sim F_4''$ being obtained by multiplying everything by the number $\nu=e^{\pi i/4}$.

There are many other things that can be said about the circulant Hadamard matrices, and about the Fourier matrices, and we refer here to Björck [15] and Haagerup [45].

7c. Bistochastic matrices

Getting back now to the main idea behind what we are doing, namely building on the relation between $\mathbb{I}_N$ and F_N, let us study now the class of bistochastic matrices:

DEFINITION 7.16. *A square matrix $M\in M_N(\mathbb{C})$ is called bistochastic if each row and each column sum up to the same number:*

$$\begin{matrix}M_{11}&\ldots&M_{1N}&\to&\lambda\\ \vdots&&\vdots\\ M_{N1}&\ldots&M_{NN}&\to&\lambda\\ \downarrow&&\downarrow\\ \lambda&&\lambda\end{matrix}$$

If this happens only for the rows, or only for the columns, the matrix is called row-stochastic, respectively column-stochastic.

As the name indicates, these matrices are useful in statistics, with the case of the matrices having entries in $[0,1]$, which sum up to $\lambda=1$, being the important one.

As a basic example of a bistochastic matrix, we have of course the flat matrix $\mathbb{I}_N$. In fact, the various above notions of stochasticity are closely related to $\mathbb{I}_N$, or rather to the all-one vector ξ that the matrix $\mathbb{I}_N/N$ projects on, in the following way:

PROPOSITION 7.17. *Let $M \in M_N(\mathbb{C})$ be a square matrix.*

1. *M is row stochastic, with sums λ, when $M\xi = \lambda\xi$.*
2. *M is column stochastic, with sums λ, when $M^t\xi = \lambda\xi$.*
3. *M is bistochastic, with sums λ, when $M\xi = M^t\xi = \lambda\xi$.*

PROOF. All these assertions are clear from definitions, because when multiplying a matrix by ξ, we obtain the vector formed by the row sums. $\square$

As an observation here, we can reformulate if we want the above statement in a purely matrix-theoretic form, by using the flat matrix $\mathbb{I}_N$, as follows:

PROPOSITION 7.18. *Let $M \in M_N(\mathbb{C})$ be a square matrix.*

1. *M is row stochastic, with sums λ, when $M\mathbb{I}_N = \lambda\mathbb{I}_N$.*
2. *M is column stochastic, with sums λ, when $\mathbb{I}_N M = \lambda\mathbb{I}_N$.*
3. *M is bistochastic, with sums λ, when $M\mathbb{I}_N = \mathbb{I}_N M = \lambda\mathbb{I}_N$.*

PROOF. This follows from Proposition 7.17, and from the fact that both the rows and the columns of the flat matrix $\mathbb{I}_N$ are copies of the all-one vector ξ. $\square$

In what follows we will be mainly interested in the unitary bistochastic matrices, which are quite interesting objects. These do not exactly cover the flat matrix $\mathbb{I}_N$, but cover instead the following related matrix, which appears in many linear algebra questions:

$$K_N = \frac{1}{N}\begin{pmatrix} 2-N & & 2 \\ & \ddots & \\ 2 & & 2-N \end{pmatrix}$$

As a first result, regarding such matrices, we have the following statement:

THEOREM 7.19. *For a unitary matrix $U \in U_N$, the following conditions are equivalent:*

1. *H is bistochastic, with sums λ.*
2. *H is row stochastic, with sums λ, and $|\lambda| = 1$.*
3. *H is column stochastic, with sums λ, and $|\lambda| = 1$.*

PROOF. By using a symmetry argument we just need to prove $(1) \iff (2)$, and both the implications are elementary, as follows:

(1) $\implies$ (2) If we denote by $U_1,\ldots,U_N\in\mathbb{C}^N$ the rows of U, we have indeed:

$$\begin{aligned}
1 &= <U_1,U_1>\\
&= \sum_i <U_1,U_i>\\
&= \sum_i\sum_j U_{1j}\bar{U}_{ij}\\
&= \sum_j U_{1j}\sum_i \bar{U}_{ij}\\
&= \sum_j U_{1j}\cdot\bar{\lambda}\\
&= \lambda\cdot\bar{\lambda}\\
&= |\lambda|^2
\end{aligned}$$

(2) $\implies$ (1) Consider the all-one vector $\xi=(1)_i\in\mathbb{C}^N$. The fact that U is row-stochastic with sums λ reads:

$$\begin{aligned}
\sum_j U_{ij}=\lambda,\forall i &\iff \sum_j U_{ij}\xi_j=\lambda\xi_i,\forall i\\
&\iff U\xi=\lambda\xi
\end{aligned}$$

Also, the fact that U is column-stochastic with sums λ reads:

$$\begin{aligned}
\sum_i U_{ij}=\lambda,\forall j &\iff \sum_j U_{ij}\xi_i=\lambda\xi_j,\forall j\\
&\iff U^t\xi=\lambda\xi
\end{aligned}$$

We must prove that the first condition implies the second one, provided that the row sum λ satisfies $|\lambda|=1$. But this follows from the following computation:

$$\begin{aligned}
U\xi=\lambda\xi &\implies U^*U\xi=\lambda U^*\xi\\
&\implies \xi=\lambda U^*\xi\\
&\implies \xi=\bar{\lambda}U^t\xi\\
&\implies U^t\xi=\lambda\xi
\end{aligned}$$

Thus, we have proved both the implications, and we are done. □

The unitary bistochastic matrices are stable under a number of operations, and in particular under taking products, and we have the following result:

THEOREM 7.20. *The real and complex bistochastic groups, which are the sets*

$$B_N\subset O_N\quad,\quad C_N\subset U_N$$

consisting of matrices which are bistochastic, are isomorphic to O_{N-1}, U_{N-1}.

PROOF. Let us pick a unitary matrix $F \in U_N$ satisfying the following condition, where $e_0, \ldots, e_{N-1}$ is the standard basis of $\mathbb{C}^N$, and where ξ is the all-one vector:

$$Fe_0 = \frac{1}{\sqrt{N}}\xi$$

Observe that such matrices $F \in U_N$ exist indeed, the basic example being the normalized Fourier matrix $F_N/\sqrt{N}$. We have then, by using the above property of F:

$$\begin{aligned} u\xi = \xi &\iff uFe_0 = Fe_0 \\ &\iff F^*uFe_0 = e_0 \\ &\iff F^*uF = diag(1, w) \end{aligned}$$

Thus we have isomorphisms as in the statement, given by:

$$w_{ij} \to (F^*uF)_{ij}$$

But this gives both the assertions. □

We will be back to the real and complex bistochastic groups $B_N \subset O_N$ and $C_N \subset U_N$ later on in this book, when systematically doing group theory.

In relation now with Hadamard matrices, as a first remark, the first Walsh matrix W_2 looks better in complex bistochastic form, modulo the standard equivalence relation:

$$\begin{pmatrix}1&1\\1&-1\end{pmatrix} \sim \begin{pmatrix}i&i\\1&-1\end{pmatrix} \sim \begin{pmatrix}i&1\\1&i\end{pmatrix}$$

The second Walsh matrix $W_4 = W_2 \otimes W_2$ can be put as well in complex bistochastic form, and also looks better in this form:

$$\begin{aligned} \begin{pmatrix}1&1&1&1\\1&-1&1&-1\\1&1&-1&-1\\1&-1&-1&1\end{pmatrix} &\sim \begin{pmatrix}1&1&1&1\\1&1&-1&-1\\1&-1&1&-1\\1&-1&-1&1\end{pmatrix} \\ &\sim \begin{pmatrix}1&-1&-1&-1\\1&-1&1&1\\1&1&-1&1\\1&1&1&-1\end{pmatrix} \\ &\sim \begin{pmatrix}-1&1&1&1\\1&-1&1&1\\1&1&-1&1\\1&1&1&-1\end{pmatrix} \end{aligned}$$

In fact, by using the above formulae, we are led to the following statement:

PROPOSITION 7.21. *All the Walsh matrices, $W_N = W_2^{\otimes n}$ with $N = 2^n$, can be put in bistochastic form, up to the standard equivalence relation, as follows:*

(1) *The matrices W_N with $N = 4^n$ admit a real bistochastic form, namely:*

$$W_N \sim \begin{pmatrix} -1 & 1 & 1 & 1 \\ 1 & -1 & 1 & 1 \\ 1 & 1 & -1 & 1 \\ 1 & 1 & 1 & -1 \end{pmatrix}^{\otimes n}$$

(2) *The matrices W_N with $N = 2 \times 4^n$ admit a complex bistochastic form, namely:*

$$W_N \sim \begin{pmatrix} i & 1 \\ 1 & i \end{pmatrix} \otimes \begin{pmatrix} -1 & 1 & 1 & 1 \\ 1 & -1 & 1 & 1 \\ 1 & 1 & -1 & 1 \\ 1 & 1 & 1 & -1 \end{pmatrix}^{\otimes n}$$

PROOF. This follows indeed from the above discussion. □

Regarding now the question of putting the general Hadamard matrices, real or complex, in complex bistochastic form, things here are tricky. We first have:

THEOREM 7.22. *The class of the bistochastic complex Hadamard matrices has the following properties:*

(1) *It contains the circulant symmetric forms F_N' of the Fourier matrices F_N.*
(2) *It is stable under permuting rows and columns.*
(3) *It is stable under taking tensor products.*

In particular, any generalized Fourier matrix $F_{N_1,\ldots,N_k} = F_{N_1} \otimes \ldots \otimes F_{N_k}$ can be put in bistochastic and symmetric form, up to the equivalence relation.

PROOF. We have several things to be proved, the idea being as follows:

(1) We know from the above that any Fourier matrix F_N has a circulant and symmetric form F_N'. But since circulant implies bistochastic, this gives the result.

(2) The claim regarding permuting rows and columns is clear.

(3) Assuming that H, K are bistochastic, with sums λ, μ, we have:

$$\begin{aligned} \sum_{ia} (H \otimes K)_{ia,jb} &= \sum_{ia} H_{ij} K_{ab} \\ &= \sum_i H_{ij} \sum_a K_{ab} \\ &= \lambda\mu \end{aligned}$$

We have as well the following computation:

$$\begin{aligned}\sum_{jb}(H\otimes K)_{ia,jb} &= \sum_{jb}H_{ij}K_{ab}\\ &= \sum_j H_{ij}\sum_b K_{ab}\\ &= \lambda\mu\end{aligned}$$

Thus, the matrix $H\otimes K$ is bistochastic as well.

(4) As for the last assertion, this follows from (1,2,3). □

In general now, putting an arbitrary complex Hadamard matrix in bistochastic form can be theoretically done, according to a general theorem of Idel-Wolf [51]. The proof of this latter theorem is however based on a quite advanced, and non-explicit argument, coming from symplectic geometry, and there are many interesting open questions here.

7d. Hadamard conjecture

As a final topic for this chapter, let us discuss the real Hadamard matrices. The definition here, going back to 19th century work of Sylvester [82], is as follows:

DEFINITION 7.23. *A real Hadamard matrix is a square binary matrix,*

$$H\in M_N(\pm1)$$

whose rows are pairwise orthogonal, with respect to the scalar product on $\mathbb R^N$.

Observe that we do not really need real numbers in order to talk about the Hadamard matrices, because the orthogonality condition tells us that, when comparing two rows, the number of matchings should equal the number of mismatchings.

As a first result regarding such matrices, we have:

PROPOSITION 7.24. *For a square matrix $H\in M_N(\pm1)$, the following are equivalent:*

1. *The rows of H are pairwise orthogonal, and so H is Hadamard.*
2. *The columns of H are pairwise orthogonal, and so H^t is Hadamard.*
3. *The rescaled matrix $U=H/\sqrt N$ is orthogonal, $U\in O_N$.*

PROOF. This is something that we already know for the complex Hadamard matrices, with the orthogonal group O_N being replaced by the unitary group U_N. In the real case the proof is similar, with everything coming from definitions, and linear algebra. □

As an abstract consequence of the above result, let us record:

Theorem 7.25. *The set of the $N \times N$ Hadamard matrices is*

$$Y_N = M_N(\pm 1) \cap \sqrt{N} O_N$$

where O_N is the orthogonal group, the intersection being taken inside $M_N(\mathbb{R})$.

Proof. This follows from Proposition 7.24, which tells us that an arbitrary matrix $H \in M_N(\pm 1)$ belongs to Y_N if and only if it belongs to $\sqrt{N} O_N$. □

As a conclusion here, the set Y_N that we are interested in appears as a kind of set of "special rational points" of the real algebraic manifold $\sqrt{N} O_N$. Thus, we are doing some kind of algebraic geometry here, of precise type to be determined.

As before in the complex matrix case, it is convenient to introduce:

Definition 7.26. *Two real Hadamard matrices are called equivalent, and we write $H \sim K$, when it is possible to pass from H to K via the following operations:*

(1) *Permuting the rows, or the columns.*
(2) *Multiplying the rows or columns by -1.*

Observe that we do not include the transposition operation $H \to H^t$ in our list of allowed operations. This is because Proposition 7.24, while looking quite elementary, rests however on a deep linear algebra fact, namely that the transpose of an orthogonal matrix is orthogonal as well, and this can produce complications later on.

Let us do now some classification work. Here is the result at $N = 4$:

Proposition 7.27. *There is only one Hadamard matrix at $N = 4$, namely*

$$W_4 = W_2 \otimes W_2$$

up to the standard equivalence relation for such matrices.

Proof. Consider an Hadamard matrix $H \in M_4(\pm 1)$, assumed to be dephased:

$$H = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & a & b & c \\ 1 & d & e & f \\ 1 & g & h & i \end{pmatrix}$$

By orthogonality of the first 2 rows we must have $\{a, b, c\} = \{-1, -1, 1\}$, and so by permuting the last 3 columns, we can further assume that our matrix is as follows:

$$H = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & m & n & o \\ 1 & p & q & r \end{pmatrix}$$

By orthogonality of the first 2 columns we must have $\{m,p\}=\{-1,1\}$, and so by permuting the last 2 rows, we can further assume that our matrix is as follows:

$$H=\begin{pmatrix}1&1&1&1\\1&-1&1&-1\\1&1&x&y\\1&-1&z&t\end{pmatrix}$$

But this gives the result, because from the orthogonality of the rows we obtain $x=y=-1$. Indeed, with these values of x,y, from the orthogonality of the columns we obtain $z=-1,t=1$. Thus, up to equivalence we have $H=W_4$, as claimed. $\square$

The case $N=5$ is excluded, because the orthogonality condition forces $N\in 2\mathbb{N}$. The point now is that the case $N=6$ is excluded as well, because we have:

Proposition 7.28. *The size of an Hadamard matrix must be*

$$N\in\{2\}\cup 4\mathbb{N}$$

with this coming from the orthogonality condition between the first 3 *rows.*

Proof. By permuting the rows and columns or by multiplying them by -1, as to rearrange the first 3 rows, we can always assume that our matrix looks as follows:

$$H=\begin{pmatrix}1\ldots\ldots1&1\ldots\ldots1&1\ldots\ldots1&1\ldots\ldots1\\1\ldots\ldots1&1\ldots\ldots1&-1\ldots-1&-1\ldots-1\\1\ldots\ldots1&-1\ldots-1&1\ldots\ldots1&-1\ldots-1\\\underbrace{\ldots\ldots\ldots}_{x}&\underbrace{\ldots\ldots\ldots}_{y}&\underbrace{\ldots\ldots\ldots}_{z}&\underbrace{\ldots\ldots\ldots}_{t}\end{pmatrix}$$

Now if we denote by x,y,z,t the sizes of the 4 block columns, as indicated, the orthogonality conditions between the first 3 rows give the following system of equations:

$$(1\perp 2)\quad:\quad x+y=z+t$$

$$(1\perp 3)\quad:\quad x+z=y+t$$

$$(2\perp 3)\quad:\quad x+t=y+z$$

The numbers x,y,z,t being such that the average of any two equals the average of the other two, and so equals the global average, the solution of our system is:

$$x=y=z=t$$

We therefore conclude that the size of our Hadamard matrix, which is the number $N=x+y+z+t$, must be a multiple of 4, as claimed. $\square$

The above result, and various other findings, suggest the following conjecture:

CONJECTURE 7.29 (Hadamard Conjecture (HC)). *There is at least one Hadamard matrix*

$$H \in M_N(\pm 1)$$

for any integer $N \in 4\mathbb{N}$.

This conjecture, going back to the 19th century, is one of the most beautiful statements in combinatorics, linear algebra, and mathematics in general. Quite remarkably, the numeric verification so far goes up to the number of the beast:

$$\mathfrak{N} = 666$$

Our purpose now will be that of gathering some evidence for this conjecture. At $N = 4, 8$ we have the Walsh matrices W_4, W_8. Thus, the next existence problem comes at $N = 12$. And here, we can use the following key construction, due to Paley:

THEOREM 7.30. *Let $q = p^r$ be an odd prime power, define*

$$\chi : \mathbb{F}_q \to \{-1, 0, 1\}$$

by $\chi(0) = 0$, $\chi(a) = 1$ if $a = b^2$ for some $b \neq 0$, and $\chi(a) = -1$ otherwise, and finally set

$$Q_{ab} = \chi(a - b)$$

We have then constructions of Hadamard matrices, as follows:

(1) *Paley 1: if $q = 3(4)$ we have a matrix of size $N = q + 1$, as follows:*

$$P_N^1 = 1 + \begin{pmatrix} 0 & 1 & \ldots & 1 \\ -1 & & & \\ \vdots & & Q & \\ -1 & & & \end{pmatrix}$$

(2) *Paley 2: if $q = 1(4)$ we have a matrix of size $N = 2q + 2$, as follows:*

$$P_N^2 = \begin{pmatrix} 0 & 1 & \ldots & 1 \\ 1 & & & \\ \vdots & & Q & \\ 1 & & & \end{pmatrix} \quad : \quad 0 \to \begin{pmatrix} 1 & -1 \\ -1 & -1 \end{pmatrix} \quad , \quad \pm 1 \to \pm \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

These matrices are skew-symmetric ($H + H^t = 2$), respectively symmetric ($H = H^t$).

PROOF. In order to simplify the presentation, we will denote by 1 all the identity matrices, of any size, and by $\mathbb{I}$ all the rectangular all-one matrices, of any size as well. It is elementary to check that the matrix $Q_{ab} = \chi(a - b)$ has the following properties:

$$QQ^t = q1 - \mathbb{I}$$

$$Q\mathbb{I} = \mathbb{I}Q = 0$$

In addition, we have the following formulae, which are elementary as well, coming from the fact that -1 is a square in $\mathbb{F}_q$ precisely when $q = 1(4)$:

$$q = 1(4) \implies Q = Q^t$$

$$q = 3(4) \implies Q = -Q^t$$

With these observations in hand, the proof goes as follows:

(1) With our conventions for the symbols 1 and $\mathbb{I}$, the matrix in the statement is:

$$P_N^1 = \begin{pmatrix} 1 & \mathbb{I} \\ -\mathbb{I} & 1+Q \end{pmatrix}$$

With this formula in hand, the Hadamard matrix condition follows from:

$$\begin{aligned} P_N^1 (P_N^1)^t &= \begin{pmatrix} 1 & \mathbb{I} \\ -\mathbb{I} & 1+Q \end{pmatrix} \begin{pmatrix} 1 & -\mathbb{I} \\ \mathbb{I} & 1-Q \end{pmatrix} \\ &= \begin{pmatrix} N & 0 \\ 0 & \mathbb{I} + 1 - Q^2 \end{pmatrix} \\ &= \begin{pmatrix} N & 0 \\ 0 & N \end{pmatrix} \end{aligned}$$

(2) If we denote by G, F the matrices in the statement, which replace respectively the $0, 1$ entries, then we have the following formula for our matrix:

$$P_N^2 = \begin{pmatrix} 0 & \mathbb{I} \\ \mathbb{I} & Q \end{pmatrix} \otimes F + 1 \otimes G$$

With this formula in hand, the Hadamard matrix condition follows from:

$$\begin{aligned} (P_N^2)^2 &= \begin{pmatrix} 0 & \mathbb{I} \\ \mathbb{I} & Q \end{pmatrix}^2 \otimes F^2 + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \otimes G^2 + \begin{pmatrix} 0 & \mathbb{I} \\ \mathbb{I} & Q \end{pmatrix} \otimes (FG + GF) \\ &= \begin{pmatrix} q & 0 \\ 0 & q \end{pmatrix} \otimes 2 + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \otimes 2 + \begin{pmatrix} 0 & \mathbb{I} \\ \mathbb{I} & Q \end{pmatrix} \otimes 0 \\ &= \begin{pmatrix} N & 0 \\ 0 & N \end{pmatrix} \end{aligned}$$

Finally, the last assertion is clear, from the above formulae relating Q, Q^t. □

The above constructions allow us to get well beyond the Walsh matrix level:

THEOREM 7.31. *The HC is verified at least up to $N = 88$, as follows:*

(1) *At $N = 4, 8, 16, 32, 64$ we have Walsh matrices.*
(2) *At $N = 12, 20, 24, 28, 44, 48, 60, 68, 72, 80, 84, 88$ we have Paley 1 matrices.*
(3) *At $N = 36, 52, 76$ we have Paley 2 matrices.*
(4) *At $N = 40, 56$ we have Paley 1 matrices tensored with W_2.*

PROOF. First of all, the numbers in (1-4) are indeed all the multiples of 4, up to 88. As for the various assertions, the proof here goes as follows:

(1) This is clear from the definition of the Walsh matrices.

(2) Since $N-1$ takes the values $q = 11, 19, 23, 27, 43, 47, 59, 67, 71, 79, 83, 87$, all prime powers, we can indeed apply the Paley 1 construction, in all these cases.

(3) Since $N = 4(8)$ here, and $N/2-1$ takes the values $q = 17, 25, 37$, all prime powers, we can indeed apply the Paley 2 construction, in these cases.

(4) At $N = 40$ we have indeed $P_{20}^1 \otimes W_2$, and at $N = 56$ we have $P_{28}^1 \otimes W_2$. □

As a continuation of all this, at $N = 92$ we have $92 - 1 = 7 \times 13$, so the Paley 1 construction does not work, and $92/2 = 46$, so the Paley 2 construction, or tensoring with W_2, does not work either. However, the problem can be solved by using a computer.

Things get even worse at higher values of N, where more and more complicated constructions are needed. The whole subject is quite technical, and, as already mentioned, human knowledge here stops so far at the number of the beast, namely:

$$\mathfrak{N} = 666$$

For a systematic discussion of the Hadamard Conjecture, among others heavily improving Theorem 7.31, we refer to Haagerup [45] and related papers.

Another well-known open question concerns the circulant case. Given a binary vector $\gamma \in (\pm 1)^N$, one can ask whether the matrix $H \in M_N(\pm 1)$ defined by $H_{ij} = \gamma_{j-i}$ is Hadamard or not. Here is a solution to the problem:

$$K_4 = \begin{pmatrix} -1 & 1 & 1 & 1 \\ 1 & -1 & 1 & 1 \\ 1 & 1 & -1 & 1 \\ 1 & 1 & 1 & -1 \end{pmatrix}$$

More generally, any vector $\gamma \in (\pm 1)^4$ satisfying $\sum \gamma_i = \pm 1$ is a solution to the problem. The following conjecture, from the 50s, states that there are no other solutions:

CONJECTURE 7.32 (Circulant Hadamard Conjecture (CHC)). *The only Hadamard matrices which are circulant are*

$$K_4 = \begin{pmatrix} -1 & 1 & 1 & 1 \\ 1 & -1 & 1 & 1 \\ 1 & 1 & -1 & 1 \\ 1 & 1 & 1 & -1 \end{pmatrix}$$

and its conjugates, regardless of the value of $N \in \mathbb{N}$.

The fact that such a simple-looking problem is still open might seem quite surprising. Indeed, if we denote by $S \subset \{1, \ldots, N\}$ the set of positions of the -1 entries of γ, the Hadamard matrix condition is simply, for any $k \neq 0$, taken modulo N:

$$|S \cap (S+k)| = |S| - N/4$$

Thus, the above conjecture simply states that at $N \neq 4$, such a set S cannot exist. This is a well-known problem in combinatorics, raised by Ryser a long time ago.

Summarizing, we have many interesting questions in the real case. The situation is quite different from the one in complex case, where at any $N \in \mathbb{N}$ we have the Fourier matrix F_N, which makes the HC problematics dissapear. Since F_N can be put in circulant form, the CHC dissapears as well. There are however many interesting questions in the complex case, for the most in relation with questions in quantum physics.

7e. Exercises

There have been many questions, ranging from elementary to quite difficult, left open in this chapter, and we will leave things like that. Our exercises here will focus on something new, or rather not really discussed in the above, namely the systematic development of the theory of complex Hadamard matrices. First, we have:

EXERCISE 7.33. *Prove that the Fourier matrices F_2, F_3, which are given by*

$$F_2 = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \quad , \quad F_3 = \begin{pmatrix} 1 & 1 & 1 \\ 1 & w & w^2 \\ 1 & w^2 & w \end{pmatrix}$$

with $w = e^{2\pi i/3}$ are the only Hadamard matrices at $N = 2, 3$, up to equivalence.

Here the question about F_2 is quite trivial, but the question about F_3 is interesting, requiring studying the equation $a + b + c = 0$ in the plane, with $a, b, c \in \mathbb{T}$. Enjoy.

EXERCISE 7.34. *If $H \in M_M(\mathbb{T})$ and $K \in M_N(\mathbb{T})$ are complex Hadamard matrices, prove that so is the matrix*

$$H \otimes_Q K \in M_{MN}(\mathbb{T})$$

given by the following formula, with $Q \in M_{M \times N}(\mathbb{T})$,

$$(H \otimes_Q K)_{ia,jb} = Q_{ib} H_{ij} K_{ab}$$

called Diţă deformation of $H \otimes K$, with parameter Q.

Normally this is just a quick, standard verification. More difficult, however, is the question of explicitely writing down the matrices that can be constructed in this way, because this requires things like struggling with double indices.

EXERCISE 7.35. *Prove that the only complex Hadamard matrices at $N = 4$ are, up to the standard equivalence relation, the matrices*

$$F_4^q = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & q & -1 & -q \\ 1 & -q & -1 & q \end{pmatrix}$$

with $q \in \mathbb{T}$, which appear as Diţă deformations of $W_4 = F_2 \otimes F_2$.

Here the first question is quite standard, in the spirit of the computations at $N = 3$, mentioned before. As for the second question, good luck here with the double indices.

EXERCISE 7.36. *Given an Hadamard matrix $H \in M_5(\mathbb{T})$, chosen dephased,*

$$H = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & a & x & * & * \\ 1 & y & b & * & * \\ 1 & * & * & * & * \\ 1 & * & * & * & * \end{pmatrix}$$

prove that the numbers a, b, x, y must satisfy the following equation:

$$(x - y)(x - ab)(y - ab) = 0$$

This is something quite tricky, called Haagerup lemma, and in case you're stuck with this, you can of course take a look at Haagerup's paper [**45**].

EXERCISE 7.37. *Prove that the only Hadamard matrix at $N = 5$ is the Fourier matrix,*

$$F_5 = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & w & w^2 & w^3 & w^4 \\ 1 & w^2 & w^4 & w & w^3 \\ 1 & w^3 & w & w^4 & w^2 \\ 1 & w^4 & w^3 & w^2 & w \end{pmatrix}$$

with $w = e^{2\pi i/5}$, up to the standard equivalence relation for such matrices.

As a hint here, try first to apply the Haagerup lemma, all across the matrix. Finally, as a bonus exercise, take a look at the case $N = 6$, and report on what you found.

CHAPTER 8

Infinite dimensions

8a. Hilbert spaces

We have seen so far the basics of linear algebra, concerning the relation between linear maps and matrices, the determinant, the diagonalization procedure, along with some applications. In this chapter we discuss what happens in infinite dimensions. Among our motivations is the fact that the spaces of infinite dimensions are of great importance in various branches of theoretical physics, such as quantum mechanics.

To be more precise, among the main discoveries of the 1920s, due to Heisenberg, Schrödinger and others was the fact that small particles like electrons cannot really be described by their position vectors $v\in\mathbb R^3$, and instead we must use their so-called wave functions $\psi:\mathbb R^3\to\mathbb C$. Thus, the natural space for quantum mechanics, or at least for the quantum mechanics of the 1920s, is not our usual $V=\mathbb R^3$, but rather the infinite dimensional space $H=L^2(\mathbb R^3)$ of such wave functions ψ. And more recent versions of quantum mechanics are built on the same idea, infinite dimensional spaces.

For an introduction to all this, have a look at some good old books on quantum mechanics, such as Dirac [**28**], von Neumann [**86**] or Weyl [**92**], and also at Kumar [**57**] for the story. And for more recent aspects, including all sorts of particles, and the whole mess coming with them, both mathematical and physical, go with Griffiths [**43**].

Getting started now, we should look at linear algebra over infinite dimensional spaces. However, all this is not very interesting, due to a number of technical reasons, the idea being that the infinite dimensionality prevents us from doing many basic things, to the point that we cannot even have things started. So, the idea will be that of using infinite dimensional vector spaces with some extra structure, as follows:

Definition 8.1. *A scalar product on a complex vector space H is an operation*

$$H\times H\to\mathbb C$$

denoted $(x,y)\to<x,y>$, satisfying the following conditions:

(1) *$<x,y>$ is linear in x, and antilinear in y.*
(2) *$\overline{<x,y>}=<y,x>$, for any x,y.*
(3) *$<x,x>>0$, for any $x\neq0$.*

As a basic example here, we have the finite dimensional vector space $H = \mathbb{C}^N$, with its usual scalar product, which is as follows:

$$< x, y >= \sum_i x_i \bar{y}_i$$

There are many other examples, and notably various spaces of L^2 functions, which naturally appear in problems coming from physics. We will discuss them later on.

In order to study the scalar products, let us formulate the following definition:

DEFINITION 8.2. *The norm of a vector $x \in H$ is the following quantity:*

$$||x|| = \sqrt{< x, x >}$$

We also call this number length of x, or distance from x to the origin.

In analogy with what happens in finite dimensions, we have two important results regarding the norms. First is the Cauchy-Schwarz inequality, as follows:

THEOREM 8.3. *We have the Cauchy-Schwarz inequality*

$$| < x, y > | \leq ||x|| \cdot ||y||$$

and the equality case holds precisely when x, y are proportional.

PROOF. Consider the following quantity, depending on a real variable $t \in \mathbb{R}$, and on a variable on the unit circle, $w \in \mathbb{T}$:

$$f(t) = ||twx + y||^2$$

By developing f, we see that this is a degree 2 polynomial in t:

$$\begin{aligned} f(t) &= < twx + y, twx + y > \\ &= t^2 < x, x > + tw < x, y > + t\bar{w} < y, x > + < y, y > \\ &= t^2||x||^2 + 2tRe(w < x, y >) + ||y||^2 \end{aligned}$$

Since f is obviously positive, its discriminant must be negative:

$$4Re(w < x, y >)^2 - 4||x||^2 \cdot ||y||^2 \leq 0$$

But this is equivalent to the following condition:

$$|Re(w < x, y >)| \leq ||x|| \cdot ||y||$$

Now the point is that we can arrange for the number $w \in \mathbb{T}$ to be such that the quantity $w < x, y >$ is real. Thus, we obtain the following inequality:

$$| < x, y > | \leq ||x|| \cdot ||y||$$

Finally, the study of the equality case is straightforward, by using the fact that the discriminant of f vanishes precisely when we have a root. But this leads to the conclusion in the statement, namely that the vectors x, y must be proportional. $\square$

As a second main result now, we have the Minkowski inequality:

THEOREM 8.4. *We have the Minkowski inequality*

$$||x+y|| \leq ||x|| + ||y||$$

and the equality case holds precisely when x, y are proportional.

PROOF. This follows indeed from the Cauchy-Schwarz inequality, as follows:

$$\begin{aligned}
& ||x+y|| \leq ||x|| + ||y|| \\
\iff & ||x+y||^2 \leq (||x|| + ||y||)^2 \\
\iff & ||x||^2 + ||y||^2 + 2Re < x, y > \leq ||x||^2 + ||y||^2 + 2||x|| \cdot ||y|| \\
\iff & Re < x, y > \leq ||x|| \cdot ||y||
\end{aligned}$$

As for the equality case, this is clear from Cauchy-Schwarz as well. □

As a consequence of this, we have the following result:

THEOREM 8.5. *The following function is a distance on H,*

$$d(x,y) = ||x-y||$$

in the usual sense, that of the abstract metric spaces.

PROOF. This follows indeed from the Minkowski inequality, which corresponds to the triangle inequality, the other two axioms for a distance being trivially satisfied. □

The above result is quite important, because it shows that we can do geometry in our present setting, a bit as in the finite dimensional case.

Finally, in connection with doing geometry, we have the following key technical result, which shows that everything can be recovered in terms of distances:

PROPOSITION 8.6. *The scalar products can be recovered from distances, via the formula*

$$\begin{aligned}
4 < x, y > &= ||x+y||^2 - ||x-y||^2 \\
&+ i||x+iy||^2 - i||x-iy||^2
\end{aligned}$$

called complex polarization identity.

PROOF. This is something that we have already met in finite dimensions. In arbitrary dimensions the proof is similar, as follows:

$$\begin{aligned}
& ||x+y||^2 - ||x-y||^2 + i||x+iy||^2 - i||x-iy||^2 \\
= & ||x||^2 + ||y||^2 - ||x||^2 - ||y||^2 + i||x||^2 + i||y||^2 - i||x||^2 - i||y||^2 \\
& +2Re(< x, y >) + 2Re(< x, y >) + 2iIm(< x, y >) + 2iIm(< x, y >) \\
= & 4 < x, y >
\end{aligned}$$

Thus, we are led to the conclusion in the statement. □

Let us discuss now some more advanced aspects. In order to do analysis on our spaces, we need the Cauchy sequences that we construct to converge. This is something which is automatic in finite dimensions, but in arbitrary dimensions, this can fail.

Thus, we must add an extra axiom, stating that H is complete with respect to the norm. It is convenient here to formulate a detailed new definition, as follows, which will be the starting point for our various considerations to follow:

DEFINITION 8.7. *A Hilbert space is a complex vector space H given with a scalar product $< x, y >$, satisfying the following conditions:*

(1) *$< x, y >$ is linear in x, and antilinear in y.*
(2) *$\overline{< x, y >} =< y, x >$, for any x, y.*
(3) *$< x, x >> 0$, for any $x \neq 0$.*
(4) *H is complete with respect to the norm $||x|| = \sqrt{< x, x >}$.*

In other words, we have taken here Definition 8.1, and added the condition that H must be complete with respect to the norm $||x|| = \sqrt{< x, x >}$, that we know indeed to be a norm, according to the Minkowski inequality proved above.

As a basic example, we have the space $H = \mathbb{C}^N$, with its usual scalar product:

$$< x, y >= \sum_i x_i \bar{y}_i$$

More generally now, we have the following construction of Hilbert spaces:

PROPOSITION 8.8. *The sequences of numbers $x = (x_i)$ which are square-summable,*

$$\sum_i |x_i|^2 < \infty$$

form a Hilbert space, denoted $l^2(\mathbb{N})$, with the following scalar product:

$$< x, y >= \sum_i x_i \bar{y}_i$$

In fact, given any index set I, we can construct a Hilbert space $l^2(I)$, in this way.

PROOF. The fact that we have indeed a complex vector space with a scalar product is elementary, and the fact that this space is indeed complete is very standard too. $\square$

On the other hand, we can talk as well about spaces of functions, as follows:

PROPOSITION 8.9. *Given an interval $X \subset \mathbb{R}$, the quantity*

$$< f, g >= \int_X f(x)\overline{g(x)}dx$$

is a scalar product, making $H = L^2(X)$ a Hilbert space.

PROOF. Once again this is routine, coming this time from basic measure theory, with $H = L^2(X)$ being the space of square-integrable functions $f : X \to \mathbb{C}$, with the convention that two such functions are identified when they coincide almost everywhere. $\square$

We can unify the above two constructions, as follows:

THEOREM 8.10. *Given a measured space X, the quantity*

$$< f, g >= \int_X f(x)\overline{g(x)}dx$$

is a scalar product, making $H = L^2(X)$ a Hilbert space.

PROOF. Here the first assertion is clear, and the fact that the Cauchy sequences converge is clear as well, by taking the pointwise limit, and using a standard argument. $\square$

Observe that with $X = \{1, \ldots, N\}$ we obtain the space $H = \mathbb{C}^N$. Also, with $X = \mathbb{N}$, with the counting measure, we obtain the space $H = l^2(\mathbb{N})$. In fact, with an arbitrary set I, once again with the counting mesure, we obtain the space $H = l^2(I)$. Thus, the construction in Theorem 8.10 unifies all the Hilbert space constructions that we have.

Quite remarkably, the converse of this holds, in the sense that any Hilbert space must be of the form $L^2(X)$. This follows indeed from the following key result, which tells us that, in addition to this, we can always assume that $X = I$ is a discrete space:

THEOREM 8.11. *Let H be a Hilbert space.*

(1) *Any algebraic basis of this space $\{f_i\}_{i\in I}$ can be turned into an orthonormal basis $\{e_i\}_{i\in I}$, by using the Gram-Schmidt procedure.*
(2) *Thus, H has an orthonormal basis, and so we have $H \simeq l^2(I)$, with I being the indexing set for this orthonormal basis.*

PROOF. This is standard, by recurrence in finite dimensions, using Gram-Schmidt, as stated, and by recurrence as well in infinite, countable dimensions. As for the case of infinite, uncountable dimensions, here the result holds as well, with the proof using transfinite recurrence arguments, borrowed from logic. $\square$

We have the following definition, based on the above:

DEFINITION 8.12. *A Hilbert space H is called separable when the following equivalent conditions are satisfied:*

(1) *H has a countable algebraic basis $\{f_i\}_{i\in\mathbb{N}}$.*
(2) *H has a countable orthonormal basis $\{e_i\}_{i\in\mathbb{N}}$.*
(3) *We have $H \simeq l^2(\mathbb{N})$, isomorphism of Hilbert spaces.*

In what follows we will be mainly interested in the separable Hilbert spaces, where most of the questions coming from physics take place. In view of the above, the following philosophical question appears: why not simply talking about $l^2(\mathbb{N})$?

In answer to this, we cannot really do so, because many of the separable spaces that we are interested in appear as spaces of functions, and such spaces do not necessarily have a very simple or explicit orthonormal basis, as shown by the following result:

PROPOSITION 8.13. *The Hilbert space $H = L^2[0,1]$ is separable, having as orthonormal basis the orthonormalized version of the algebraic basis*

$$f_n = x^n$$

with $n \in \mathbb{N}$, coming from the Weierstrass density theorem.

PROOF. The fact that the space $H = L^2[0,1]$ is indeed separable is clear from the Weierstrass theorem, which provides us with the algebraic basis $f_n = x^n$, which can be orthogonalized by using the Gram-Schmidt procedure, as explained in Theorem 8.11. Working out the details here is actually an excellent exercise. □

As a conclusion to all this, we are interested in 1 space, namely the unique separable Hilbert space H, but due to various technical reasons, it is often better to forget that we have $H = l^2(\mathbb{N})$, and say instead that we have $H = L^2(X)$, with X being a separable measured space, or simply say that H is an abstract separable Hilbert space.

8b. Linear operators

Let us get now into the study of linear operators $T : H \to H$, which will eventually lead us into the correct infinite dimensional version of linear algebra. We first have:

THEOREM 8.14. *Let H be an arbitrary Hilbert space, with orthonormal basis $\{e_i\}_{i\in I}$. The algebra of all linear operators from H to itself,*

$$\mathcal{L}(H) = \left\{T : H \to H \text{ linear}\right\}$$

embeds then into the space of the $I \times I$ complex matrices,

$$M_I(\mathbb{C}) = \left\{(M_{ij})_{i,j\in I}\Big|M_{ij} \in \mathbb{C}\right\}$$

with an operator T corresponding to the following matrix:

$$M_{ij} =< Te_j, e_i >$$

In the case $H = \mathbb{C}^N$ we obtain in this way the usual isomorphism $\mathcal{L}(H) \simeq M_N(\mathbb{C})$. In the separable case we obtain in this way a proper embedding $\mathcal{L}(H) \subset M_\infty(\mathbb{C})$.

PROOF. We have three assertions to be proved, the idea being as follows:

(1) The correspondence $T \to M$ constructed in the statement is indeed linear, and its kernel is $\{0\}$, so we have indeed an embedding as follows, as claimed:

$$\mathcal{L}(H) \subset M_I(\mathbb{C})$$

(2) In finite dimensions we obtain an isomorphism, because any matrix $M \in M_N(\mathbb{C})$ determines an operator $T : \mathbb{C}^N \to \mathbb{C}^N$, according to the formula $< Te_j, e_i >= M_{ij}$.

(3) In infinite dimensions, however, we do not have an isomorphism. For instance on $H = l^2(\mathbb{N})$ the following matrix does not define an operator:

$$M = \begin{pmatrix} 1 & 1 & \ldots \\ 1 & 1 & \ldots \\ \vdots & \vdots \end{pmatrix}$$

Indeed, $T(e_1)$ should be the all-1 vector, but this vector is not square-summable. $\square$

In connection with our previous comments, the above result is something quite theoretical, because for basic Hilbert spaces like $L^2[0,1]$, which do not have a simple orthonormal basis, the embedding $\mathcal{L}(H) \subset M_\infty(\mathbb{C})$ that we obtain is not something very useful.

In short, while the operators $T : H \to H$ are basically some infinite matrices, it is better to think of these operators as being objects on their own.

In what follows we will be interested in the operators $T : H \to H$ which are bounded. Regarding such operators, we have the following result:

THEOREM 8.15. *Given a Hilbert space H, the linear operators $T : H \to H$ which are bounded, in the sense that we have*

$$||T|| = \sup_{||x|| \leq 1} ||Tx|| < \infty$$

form a complex algebra with unit $B(H)$, having the property

$$||ST|| \leq ||S|| \cdot ||T||$$

and which is complete with respect to the norm.

PROOF. The fact that we have indeed an algebra, satisfying the product condition in the statement, follows from the following estimates, which are all elementary:

$$||S + T|| \leq ||S|| + ||T||$$

$$||\lambda T|| = |\lambda| \cdot ||T||$$

$$||ST|| \leq ||S|| \cdot ||T||$$

Regarding now the last assertion, if $\{T_n\} \subset B(H)$ is Cauchy then $\{T_n x\}$ is Cauchy for any $x \in H$, so we can define the limit $T = \lim_{n\to\infty} T_n$ by setting:

$$Tx = \lim_{n\to\infty} T_n x$$

Let us first check that the application $x \to Tx$ is linear. We have:

$$\begin{aligned}
T(x+y) &= \lim_{n\to\infty} T_n(x+y)\\
&= \lim_{n\to\infty} T_n(x)+T_n(y)\\
&= \lim_{n\to\infty} T_n(x)+\lim_{n\to\infty} T_n(y)\\
&= T(x)+T(y)
\end{aligned}$$

Similarly, we have as well the following computation:

$$\begin{aligned}
T(\lambda x) &= \lim_{n\to\infty} T_n(\lambda x)\\
&= \lambda \lim_{n\to\infty} T_n(x)\\
&= \lambda T(x)
\end{aligned}$$

Thus we have $T \in \mathcal{L}(H)$. It remains now to prove that we have $T \in B(H)$, and that we have $T_n \to T$ in norm. For this purpose, observe that we have:

$$\begin{aligned}
& ||T_n - T_m|| \leq \varepsilon\ ,\ \forall n,m \geq N\\
\implies\quad & ||T_nx - T_mx|| \leq \varepsilon\ ,\ \forall ||x|| = 1\ ,\ \forall n,m \geq N\\
\implies\quad & ||T_nx - Tx|| \leq \varepsilon\ ,\ \forall ||x|| = 1\ ,\ \forall n \geq N\\
\implies\quad & ||T_Nx - Tx|| \leq \varepsilon\ ,\ \forall ||x|| = 1\\
\implies\quad & ||T_N - T|| \leq \varepsilon
\end{aligned}$$

As a first consequence, we obtain $T \in B(H)$, because we have:

$$\begin{aligned}
||T|| &= ||T_N + (T - T_N)||\\
&\leq ||T_N|| + ||T - T_N||\\
&\leq ||T_N|| + \varepsilon\\
&< \infty
\end{aligned}$$

As a second consequence, we obtain $T_N \to T$ in norm, and we are done. $\square$

In relation with the construction from Theorem 8.14, we have:

PROPOSITION 8.16. *We have embeddings as follows,*

$$B(H) \subset \mathcal{L}(H) \subset M_I(\mathbb{C})$$

which are both proper, in the infinite dimensional case.

PROOF. According to Theorem 8.14, the algebra $B(H)$ consists of the $I \times I$ complex matrices which define indeed linear maps $T : H \to H$, and which satisfy as well a second boundedness condition, coming from the boundedness of the norm of T:

$$||T|| < \infty$$

In finite dimensions we have equalities everywhere, but in general this is not true, the standard example of a matrix not producing an operator being as follows:

$$M = \begin{pmatrix} 1 & 1 & \ldots \\ 1 & 1 & \ldots \\ \vdots & \vdots \end{pmatrix}$$

As for the examples of linear operators which are not bounded, these are more complicated, coming from logic, and we will not need them in what follows. □

As already mentioned after Theorem 8.14, all this is something quite theoretical, because for basic function spaces like $L^2[0,1]$, which do not have a simple orthonormal basis, the embedding $B(H) \subset M_I(\mathbb{C})$ that we obtain is not very useful.

As a conclusion, while the bounded operators $T : H \to H$ are basically some infinite matrices, it is better to think of these operators as being objects on their own.

8c. Spectral theory

We will be interested in what follows in $B(H)$ and its closed subalgebras $A \subset B(H)$. It is convenient to formulate the following definition:

Definition 8.17. *A Banach algebra is a complex algebra with unit A, having a vector space norm $||.||$ satisfying*

$$||ab|| \leq ||a|| \cdot ||b||$$

and which makes it a Banach space, in the sense that the Cauchy sequences converge.

As said above, the basic examples of Banach algebras, or at least the basic examples that we will be interested in here, are the operator algebra $B(H)$, and its norm closed subalgebras $A \subset B(H)$, such as the algebras $A =< T >$ generated by a single operator $T \in B(H)$. There are many other examples, but more on this later.

Generally speaking, the elements $a \in A$ of a Banach algebra can be thought of as being bounded operators on some Hilbert space, which is not present. With this idea in mind, we can emulate spectral theory in our setting, the starting point being:

Definition 8.18. *The spectrum of an element $a \in A$ is the set*

$$\sigma(a) = \left\{ \lambda \in \mathbb{C} \Big| a - \lambda \notin A^{-1} \right\}$$

where $A^{-1} \subset A$ is the set of invertible elements.

As a basic example, the spectrum of a usual matrix $M \in M_N(\mathbb{C})$ is the collection of its eigenvalues, taken of course without multiplicities. In the case of the trivial algebra $A = \mathbb{C}$, appearing at $N = 1$, the spectrum of an element is the element itself.

As a first, basic result regarding spectra, we have:

PROPOSITION 8.19. *We have the following formula, valid for any $a,b\in A$:*

$$\sigma(ab)\cup\{0\}=\sigma(ba)\cup\{0\}$$

Also, there are examples where $\sigma(ab)\neq\sigma(ba)$.

PROOF. We first prove that we have the following implication:

$$1\notin\sigma(ab)\implies 1\notin\sigma(ba)$$

Assume indeed that $1-ab$ is invertible, with inverse denoted c:

$$c=(1-ab)^{-1}$$

We have then the following formulae, relating our variables a,b,c:

$$abc=cab=c-1$$

By using these formulae, we obtain:

$$\begin{aligned}(1+bca)(1-ba)&=1+bca-ba-bcaba\\&=1+bca-ba-bca+ba\\&=1\end{aligned}$$

A similar computation shows that we have:

$$(1-ba)(1+bca)=1$$

Thus $1-ba$ is invertible, with inverse $1+bca$, which proves our claim. Now by multiplying by scalars, we deduce from this that for any $\lambda\in\mathbb{C}-\{0\}$ we have:

$$\lambda\notin\sigma(ab)\implies\lambda\notin\sigma(ba)$$

But this leads to the conclusion in the statement, namely:

$$\sigma(ab)\cup\{0\}=\sigma(ba)\cup\{0\}$$

Regarding now the last claim, we know from linear algebra that $\sigma(ab)=\sigma(ba)$ holds for the usual matrices, for instace because of the above, and because ab is invertible if any only if ba is. However, this latter fact fails for general operators on Hilbert spaces. Indeed, we can take our operator a to be the shift on the space $l^2(\mathbb{N})$, given by:

$$S(e_i)=e_{i+1}$$

As for b, we can take the adjoint of S, which is the following operator:

$$S^*(e_i)=\begin{cases}e_{i-1}&\text{if } i>0\\0&\text{if } i=0\end{cases}$$

Let us compose now these two operators. In one sense, we have:

$$S^*S=1\implies 0\notin\sigma(SS^*)$$

In the other sense, however, the situation is different, as follows:

$$SS^* = Proj(e_0^\perp) \implies 0 \in \sigma(SS^*)$$

Thus, the spectra do not match on 0, and we have our counterexample, as desired. $\square$

Let us discuss now a second basic result about spectra, which is very useful. Given an arbitrary Banach algebra element $a \in A$, and a rational function $f = P/Q$ having poles outside $\sigma(a)$, we can construct the following element:

$$f(a) = P(a)Q(a)^{-1}$$

For simplicity, and due to the fact that the elements $P(a), Q(a)$ commute, so that the order is irrelevant, we write this element as a usual fraction, as follows:

$$f(a) = \frac{P(a)}{Q(a)}$$

With this convention, we have the following result:

THEOREM 8.20. *We have the "rational functional calculus" formula*

$$\sigma(f(a)) = f(\sigma(a))$$

valid for any rational function $f \in \mathbb{C}(X)$ having poles outside $\sigma(a)$.

PROOF. In order to prove this result, we can proceed in two steps, as follows:

(1) Assume first that we are in the polynomial function case, $f \in \mathbb{C}[X]$. We pick a scalar $\lambda \in \mathbb{C}$, and we write:

$$f(X) - \lambda = c(X - r_1) \ldots (X - r_n)$$

We have then, as desired:

$$\begin{aligned}
\lambda \notin \sigma(f(a)) &\iff f(a) - \lambda \in A^{-1} \\
&\iff c(a - r_1) \ldots (a - r_n) \in A^{-1} \\
&\iff a - r_1, \ldots, a - r_n \in A^{-1} \\
&\iff r_1, \ldots, r_n \notin \sigma(a) \\
&\iff \lambda \notin f(\sigma(a))
\end{aligned}$$

(2) Assume now that we are in the general rational function case, $f \in \mathbb{C}(X)$. We pick a scalar $\lambda \in \mathbb{C}$, we write $f = P/Q$, and we set:

$$F = P - \lambda Q$$

By using now (1), for this polynomial, we obtain:

$$\begin{aligned}\lambda\in\sigma(f(a)) &\iff F(a)\notin A^{-1}\\ &\iff 0\in\sigma(F(a))\\ &\iff 0\in F(\sigma(a))\\ &\iff \exists\mu\in\sigma(a),F(\mu)=0\\ &\iff \lambda\in f(\sigma(a))\end{aligned}$$

Thus, we have obtained the formula in the statement. □

Summarizing, we have so far a beginning of theory. Let us prove now something that we do not know yet, namely that the spectra are non-empty:

$$\sigma(a)\neq\emptyset$$

This is something that we know well for the usual matrices. However, a bit of thinking tells us that, even for the usual matrices, this is something rather advanced.

In the present Banach algebra setting, this is definitely something non-trivial. In order to establish this result, we will need a number of analytic preliminaries, as follows:

PROPOSITION 8.21. *Let A be a Banach algebra.*

(1) $||a||<1\implies(1-a)^{-1}=1+a+a^2+\ldots$
(2) *The set A^{-1} is open.*
(3) *The map $a\to a^{-1}$ is differentiable.*

PROOF. All these assertions are elementary, as follows:

(1) This follows as in the scalar case, the computation being as follows, provided that everything converges under the norm, which amounts in saying that $||a||<1$:

$$\begin{aligned}&(1-a)(1+a+a^2+\ldots)\\ =&\ 1-a+a-a^2+a^2-a^3+\ldots\\ =&\ 1\end{aligned}$$

(2) Assuming $a\in A^{-1}$, let us pick $b\in A$ such that:

$$||a-b||<\frac{1}{||a^{-1}||}$$

We have then the following estimate:

$$\begin{aligned}||1-a^{-1}b|| &= ||a^{-1}(a-b)||\\ &\leq ||a^{-1}||\cdot||a-b||\\ &< 1\end{aligned}$$

Thus by (1) we obtain $a^{-1}b\in A^{-1}$, and so $b\in A^{-1}$, as desired.

(3) This follows as in the scalar case, where the derivative of $f(t) = t^{-1}$ is:

$$f'(t) = -t^{-2}$$

To be more precise, in the present Banach algebra setting the derivative is no longer a number, but rather a linear transformation. But this linear transformation can be found by developing the function $f(a) = a^{-1}$ at order 1, as follows:

$$\begin{aligned}
(a+h)^{-1} &= ((1+ha^{-1})a)^{-1}\\
&= a^{-1}(1+ha^{-1})^{-1}\\
&= a^{-1}(1-ha^{-1}+(ha^{-1})^2-\ldots)\\
&\simeq a^{-1}(1-ha^{-1})\\
&= a^{-1}-a^{-1}ha^{-1}
\end{aligned}$$

We conclude that the derivative that we are looking for is:

$$f'(a)h = -a^{-1}ha^{-1}$$

Thus, we are led to the conclusion in the statement. □

We can now formulate a key theorem about the Banach algebras, as follows:

THEOREM 8.22. *The spectrum of a Banach algebra element $\sigma(a) \subset \mathbb{C}$ is:*

(1) *Compact.*
(2) *Contained in the disc $D_0(||a||)$.*
(3) *Non-empty.*

PROOF. This can be proved by using the above results, as follows:

(1) In view of (2) below, it is enough to prove that $\sigma(a)$ is closed. But this follows from the following computation, with $|\varepsilon|$ being small:

$$\begin{aligned}
\lambda \notin \sigma(a) &\implies a-\lambda \in A^{-1}\\
&\implies a-\lambda-\varepsilon \in A^{-1}\\
&\implies \lambda+\varepsilon \notin \sigma(a)
\end{aligned}$$

(2) This follows from the following computation:

$$\begin{aligned}
\lambda > ||a|| &\implies \left|\left|\frac{a}{\lambda}\right|\right| < 1\\
&\implies 1-\frac{a}{\lambda} \in A^{-1}\\
&\implies \lambda - a \in A^{-1}\\
&\implies \lambda \notin \sigma(a)
\end{aligned}$$

(3) Assume by contradiction $\sigma(a) = \emptyset$. Given a linear form $f \in A^*$, consider the following map, which is well-defined, due to our assumption $\sigma(a) = \emptyset$:

$$\varphi : \mathbb{C} \to \mathbb{C} \quad , \quad \lambda \to f((a-\lambda)^{-1})$$

By using Proposition 8.21 this map is differentiable, and so is a power series:

$$\varphi(\lambda) = \sum_{k=0}^{\infty} c_k \lambda^k$$

On the other hand, we have the following estimate:

$$\begin{aligned} \lambda \to \infty \quad &\implies \quad a - \lambda \to \infty \\ &\implies \quad (a-\lambda)^{-1} \to 0 \\ &\implies \quad \varphi(\lambda) \to 0 \end{aligned}$$

Thus $\varphi = 0$, and so $(a-\lambda)^{-1} = 0$. But this is a contradiction, as desired. □

This was for the basic spectral theory in Banach algebras, which notably applies to the case $A = B(H)$. It is possible to go beyond the above, for instance with a holomorphic function extension of the rational functional calculus formula $\sigma(f(a)) = f(\sigma(a))$ from Theorem 8.20. Also, in the case of the algebras of operators, more can be said.

8d. Operator algebras

Let us get back now to the operator algebra $B(H)$, from Theorem 8.15. The point is that this Banach algebra is of a very special type, due to the following fact:

THEOREM 8.23. *The Banach algebra $B(H)$ has an involution $T \to T^*$, given by*

$$< Tx, y > = < x, T^* y >$$

which is antilinear, antimultiplicative, and is an isometry, in the sense that:

$$||T|| = ||T^*||$$

Moreover, the norm the involution are related as well by the following formula:

$$||TT^*|| = ||T||^2$$

PROOF. We have several things to be proved, the idea being as follows:

(1) As a preliminary fact, that we will need in what follows, our claim is that any linear form $\varphi : H \to \mathbb{C}$ must be of the following type, for a certain vector $z \in H$:

$$\varphi(x) = < x, z >$$

Indeed, this is something clear for any Hilbert space of type $H = l^2(I)$. But, by using a basis, any Hilbert space is of this form, and so we have proved our claim.

(2) The existence of the adjoint operator T^*, given by the formula in the statement, comes from the fact that the function $\varphi(x) =< Tx, y >$ being a linear map $H \to \mathbb{C}$, we must have a formula as follows, for a certain vector $T^*y \in H$:

$$\varphi(x) =< x, T^*y >$$

Moreover, since this vector is unique, T^* is unique too, and we have as well:

$$(S+T)^* = S^* + T^*$$
$$(\lambda T)^* = \bar{\lambda}T^*$$
$$(ST)^* = T^*S^*$$
$$(T^*)^* = T$$

Observe also that we have indeed $T^* \in B(H)$, because:

$$\begin{aligned} ||T|| &= \sup_{||x||=1} \sup_{||y||=1} < Tx, y > \\ &= \sup_{||y||=1} \sup_{||x||=1} < x, T^*y > \\ &= ||T^*|| \end{aligned}$$

Summarizing, we have proved everything, up to the last assertion.

(3) Regarding now the last assertion, observe that we have:

$$||TT^*|| \leq ||T|| \cdot ||T^*|| = ||T||^2$$

On the other hand, we have as well the following estimate:

$$\begin{aligned} ||T||^2 &= \sup_{||x||=1} | < Tx, Tx > | \\ &= \sup_{||x||=1} | < x, T^*Tx > | \\ &\leq ||T^*T|| \end{aligned}$$

By replacing $T \to T^*$ we obtain from this that we have as well:

$$||T||^2 \leq ||TT^*||$$

Thus, we have obtained the needed inequality, and we are done. $\square$

As a first observation, in the context of the construction $T \to M$ from Theorem 8.14, the adjoint operation $T \to T^*$ takes a very simple form, namely:

$$(M^*)_{ij} = \overline{M}_{ji}$$

However, as already explained before, while the operators $T : H \to H$ are basically some infinite matrices, it is better to think of them as being objects on their own.

The above result suggests the following definition:

DEFINITION 8.24. *A C^*-algebra is a complex algebra with unit A, having:*

(1) *A norm $a \to ||a||$, making it a Banach algebra.*
(2) *An involution $a \to a^*$, which satisfies $||aa^*|| = ||a||^2$, for any $a \in A$.*

At the level of the basic examples, we know from Theorem 8.23 that the full operator algebra $B(H)$ is a C^*-algebra, in the above sense. More generally, any closed $*$-subalgebra $A \subset B(H)$ is a C^*-algebra. We will see later on that any C^*-algebra appears in fact in this way, as a closed $*$-subalgebra $A \subset B(H)$, for a certain Hilbert space H.

For the moment, we are interested in developing the theory of C^*-algebras, without reference to operators, or Hilbert spaces. As a first observation, we have:

PROPOSITION 8.25. *If X is an abstract compact space, the algebra $C(X)$ of continuous functions $f : X \to \mathbb{C}$ is a C^*-algebra, with structure as follows:*

(1) *The norm is the usual sup norm, given by:*
$$||f|| = \sup_{x \in X} |f(x)|$$
(2) *The involution is the usual involution, given by:*
$$f^*(x) = \overline{f(x)}$$

This algebra is commutative, in the sense that $fg = gf$, for any f, g.

PROOF. Almost everything here is trivial. Observe that we have indeed:
$$\begin{aligned} ||ff^*|| &= \sup_{x \in X} |f(x)\overline{f(x)}| \\ &= \sup_{x \in X} |f(x)|^2 \\ &= ||f||^2 \end{aligned}$$
Thus, the axioms are satisfied, and finally $fg = gf$ is clear. □

Our claim now is that any commutative C^*-algebra appears as above. This is something non-trivial, which requires a number of preliminaries. We will need:

DEFINITION 8.26. *Given an element $a \in A$, its spectral radius*
$$\rho(a) \in (0, ||a||)$$
is the radius of the smallest disk centered at 0 containing $\sigma(a)$.

Here we have included a number of results that we already know, from Theorem 8.22, namely the fact that the spectrum is nonzero, and contained in the disk $D_0(||a||)$.

We have the following key result, extending our spectral theory knowledge, from the general Banach algebra setting, to the present C^*-algebra setting:

THEOREM 8.27. *Let A be a C^*-algebra.*

1. *The spectrum of a unitary element ($a^* = a^{-1}$) is on the unit circle.*
2. *The spectrum of a self-adjoint element ($a = a^*$) consists of real numbers.*
3. *The spectral radius of a normal element ($aa^* = a^*a$) is equal to its norm.*

PROOF. We use the various results established above, and notably the rational calculus formula from Theorem 8.20, and the various results from Theorem 8.22:

(1) Assuming $a^* = a^{-1}$, we have the following norm computations:

$$||a|| = \sqrt{||aa^*||} = \sqrt{1} = 1$$

$$||a^{-1}|| = ||a^*|| = ||a|| = 1$$

Now if we denote by D the unit disk, we obtain from this:

$$||a|| = 1 \implies \sigma(a) \subset D$$

$$||a^{-1}|| = 1 \implies \sigma(a^{-1}) \subset D$$

On the other hand, by using the rational function $f(z) = z^{-1}$, we have:

$$\sigma(a^{-1}) \subset D \implies \sigma(a) \subset D^{-1}$$

Now by putting everything together we obtain, as desired:

$$\sigma(a) \subset D \cap D^{-1} = \mathbb{T}$$

(2) This follows by using the result (1), just established above, and Theorem 8.20, with the following rational function, depending on $t \in \mathbb{R}$:

$$f(z) = \frac{z + it}{z - it}$$

Indeed, for $t >> 0$ the element $f(a)$ is well-defined, and we have:

$$\begin{aligned}
\left(\frac{a + it}{a - it}\right)^* &= \frac{(a + it)^*}{(a - it)^*} \\
&= \frac{a - it}{a + it} \\
&= \left(\frac{a + it}{a - it}\right)^{-1}
\end{aligned}$$

Thus the element $f(a)$ is a unitary, and by using (1) its spectrum is contained in $\mathbb{T}$. We conclude from this that we have:

$$f(\sigma(a)) = \sigma(f(a)) \subset \mathbb{T}$$

But this shows that we have, as desired:

$$\sigma(a) \subset f^{-1}(\mathbb{T}) = \mathbb{R}$$

(3) We already know that we have the inequality in one sense, $\rho(a) \leq ||a||$, and this for any $a \in A$. For the reverse inequality, when a is normal, we fix a number as follows:

$$\rho > \rho(a)$$

We have then the following computation, with the convention that the integration over the circle $|z| = \rho$ is normalized, as for the integral of the 1 function to be 1:

$$\begin{aligned}
\int_{|z|=\rho} \frac{z^n}{z-a}\, dz &= \int_{|z|=\rho} \sum_{k=0}^{\infty} z^{n-k-1} a^k\, dz \\
&= \sum_{k=0}^{\infty} \left(\int_{|z|=\rho} z^{n-k-1} dz \right) a^k \\
&= \sum_{k=0}^{\infty} \delta_{n,k+1} a^k \\
&= a^{n-1}
\end{aligned}$$

Here we have used the following formula, with $m \in \mathbb{Z}$, whose proof is elementary:

$$\int_{|z|=\rho} z^m\, dz = \delta_{m0}$$

By applying now the norm and taking n-th roots we obtain from the above formula, modulo some elementary manipulations, the following estimate:

$$\rho \geq \lim_{n\to\infty} ||a^n||^{1/n}$$

Now recall that ρ was by definiton an arbitrary number satisfying $\rho > \rho(a)$. Thus, we have obtained the following estimate, valid for any $a \in A$:

$$\rho(a) \geq \lim_{n\to\infty} ||a^n||^{1/n}$$

In order to finish, we must prove that when a is normal, this estimate implies the missing estimate, namely $\rho(a) \geq ||a||$. We can proceed in two steps, as follows:

Step 1. In the case $a = a^*$ we have $||a^n|| = ||a||^n$ for any exponent of the form $n = 2^k$, by using the C^*-algebra condition $||aa^*|| = ||a||^2$, and by taking n-th roots we get:

$$\rho(a) \geq ||a||$$

Thus, we are done with the self-adjoint case, with the result $\rho(a) = ||a||$.

Step 2. In the general normal case $aa^* = a^*a$ we have $a^n(a^n)^* = (aa^*)^n$, and by using this, along with the result from Step 1, applied to aa^*, we obtain:

$$\begin{aligned}\rho(a) &\geq \lim_{n\to\infty} \|a^n\|^{1/n}\\ &= \sqrt{\lim_{n\to\infty} \|a^n(a^n)^*\|^{1/n}}\\ &= \sqrt{\lim_{n\to\infty} \|(aa^*)^n\|^{1/n}}\\ &= \sqrt{\rho(aa^*)}\\ &= \sqrt{\|a\|^2}\\ &= \|a\|\end{aligned}$$

Thus, we are led to the conclusion in the statement. □

As a first comment, the spectral radius formula $\rho(a) = ||a||$ does not hold in general, the simplest counterexample being the following non-normal matrix:

$$M = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

As another comment, we can combine the formula $\rho(a) = ||a||$ for normal elements with the formula $||aa^*|| = ||a||^2$, and we are led to the following statement:

PROPOSITION 8.28. *In a C^*-algebra, the norm is given by*

$$||a|| = \sqrt{\sup\left\{\lambda \in \mathbb{C}\Big| aa^* - \lambda \notin A^{-1}\right\}}$$

and so is an algebraic quantity.

PROOF. We have the following computation, using the condition $||aa^*|| = ||a||^2$, then the spectral radius formula for aa^*, and finally the definition of the spectral radius:

$$\begin{aligned}||a|| &= \sqrt{||aa^*||}\\ &= \sqrt{\rho(aa^*)}\\ &= \sqrt{\sup\left\{\lambda \in \mathbb{C}\Big| \lambda \in \sigma(aa^*)\right\}}\\ &= \sqrt{\sup\left\{\lambda \in \mathbb{C}\Big| aa^* - \lambda \notin A^{-1}\right\}}\end{aligned}$$

Thus, we are led to the conclusion in the statement. □

The above result is quite interesting, because it raises the possibility of axiomatizing the C^*-algebras as being the Banach $*$-algebras having the property that the formula in Proposition 8.28 defines a norm, which must satisfy the usual C^*-algebra conditions. However, this is something rather philosophical, and we will not follow this path.

We are now in position of proving a key result, namely:

THEOREM 8.29 (Gelfand). *Any commutative C^*-algebra is the form*

$$A = C(X)$$

with the compact space X, called spectrum of A, and denoted

$$X = Spec(A)$$

appearing as the space of Banach algebra characters $\chi : A \to \mathbb{C}$.

PROOF. This can be deduced from our spectral theory results, as follows:

(1) Given a commutative C^*-algebra A, we can define indeed X to be the set of characters $\chi : A \to \mathbb{C}$, with the topology making continuous all the evaluation maps:

$$ev_a : \chi \to \chi(a)$$

Then X is a compact space, and $a \to ev_a$ is a morphism of algebras:

$$ev : A \to C(X)$$

(2) We first prove that ev is involutive. We use the following formula:

$$a = \frac{a + a^*}{2} - i \cdot \frac{i(a - a^*)}{2}$$

Thus it is enough to prove the following equality, for self-adjoint elements a:

$$ev_{a^*} = ev_a^*$$

But this is the same as proving that $a = a^*$ implies that ev_a is a real function, which is in turn true, because $ev_a(\chi) = \chi(a)$ is an element of $\sigma(a)$, contained in $\mathbb{R}$.

(3) Since A is commutative, each element is normal, so ev is isometric:

$$||ev_a|| = \rho(a) = ||a||$$

(4) It remains to prove that ev is surjective. But this follows from the Stone-Weierstrass theorem, because $ev(A)$ is a closed subalgebra of $C(X)$, which separates the points. □

As a first consequence of the Gelfand theorem, we can extend the rational calculus formula from Theorem 8.20, in the case of the normal elements, as follows:

THEOREM 8.30. *We have the "continuous functional calculus" formula*

$$\sigma(f(a)) = f(\sigma(a))$$

valid for any normal element $a \in A$, and any continuous function $f \in C(\sigma(a))$.

PROOF. Since our element a is normal, the C^*-algebra $< a >$ that is generates is commutative, and the Gelfand theorem gives an identification as follows:

$$< a >= C(X)$$

In order to compute X, observe that the map $X \to \sigma(a)$ given by evaluation at a is bijective. Thus, we have an identification of compact spaces, as follows:

$$X = \sigma(a)$$

As a conclusion, the Gelfand theorem provides us with an identification as follows:

$$< a >= C(\sigma(a))$$

Now given $f \in C(\sigma(a))$, we can define indeed an element $f(a) \in A$, with $f \to f(a)$ being a morphism of C^*-algebras, and we have $\sigma(f(a)) = f(\sigma(a))$, as claimed. □

The above result adds to a series of similar statements, namely Theorem 8.20, dealing with rational calculus, and the known holomorphic calculus in Banach algebras, briefly mentioned after Theorem 8.22. The story is not over here, because in certain special C^*-algebras, such as the matrix algebras $M_N(\mathbb{C})$, or more generally the so-called von Neumann algebras, we can apply if we want arbitrary measurable functions to the normal elements, and we have $\sigma(f(a)) = f(\sigma(a))$ as well. We will not get into this.

As another important remark, the above result, or rather the formula $< a >= C(\sigma(a))$ from its proof, when applied to the normal operators $T \in B(H)$, is more of less the spectral theorem for such operators. Once again, we will not get into this.

As a last topic, let us discuss now the GNS representation theorem, providing us with embeddings $A \subset B(H)$. We will need some more spectral theory, as follows:

PROPOSITION 8.31. *For a normal element $a \in A$, the following are equivalent:*

(1) *a is positive, in the sense that $\sigma(a) \subset [0, \infty)$.*
(2) *$a = b^2$, for some $b \in A$ satisfying $b = b^*$.*
(3) *$a = cc^*$, for some $c \in A$.*

PROOF. This is something very standard, as follows:

(1) $\implies$ (2) Since a is normal, we can use Theorem 8.30, and set $b = \sqrt{a}$.

(2) $\implies$ (3) This is trivial, because we can set $c = b$.

(3) $\implies$ (1) We proceed by contradiction. By multiplying c by a suitable element of $< cc^* >$, we are led to the existence of an element $d \neq 0$ satisfying $-dd^* \geq 0$. By writing now $d = x + iy$ with $x = x^*, y = y^*$ we have:

$$dd^* + d^*d = 2(x^2 + y^2) \geq 0$$

Thus $d^*d \geq 0$. But this contradicts the elementary fact that $\sigma(dd^*), \sigma(d^*d)$ must coincide outside $\{0\}$, that we know from Proposition 8.19. □

Here is now the GNS representation theorem for the C^*-algebras, due to Gelfand, Naimark and Segal, along with the idea of the proof:

THEOREM 8.32 (GNS theorem). *Let A be a C^*-algebra.*

(1) *A appears as a closed $*$-subalgebra $A\subset B(H)$, for some Hilbert space H.*
(2) *When A is separable (usually the case), H can be chosen to be separable.*
(3) *When A is finite dimensional, H can be chosen to be finite dimensional.*

PROOF. Let us first discuss the commutative case, $A=C(X)$. Our claim here is that if we pick a probability measure on X, we have an embedding as follows:

$$C(X)\subset B(L^2(X))$$

$$f\to(g\to fg)$$

Indeed, given a function $f\in C(X)$, consider the operator $T_f(g)=fg$, acting on $H=L^2(X)$. Observe that T_f is indeed well-defined, and bounded as well, because:

$$\begin{aligned}||fg||_2&=\sqrt{\int_X|f(x)|^2|g(x)|^2dx}\\&\leq||f||_\infty||g||_2\end{aligned}$$

The application $f\to T_f$ being linear, involutive, continuous, and injective as well, we obtain in this way a C^*-algebra embedding $C(X)\subset B(H)$, as claimed.

In general, we can use a similar idea, with the algebraic aspects being fine, and with the positivity issues being taken care of by Proposition 8.31.

Indeed, assuming that a linear form $\varphi:A\to\mathbb C$ has some suitable positivity properties, making it analogous to the integration functionals $\int_X:A\to\mathbb C$ from the commutative case, we can define a scalar product on A, by the following formula:

$$<a,b>=\varphi(ab^*)$$

By completing we obtain a Hilbert space H, and we have an embedding as follows:

$$A\subset B(H)$$

$$a\to(b\to ab)$$

Thus we obtain the assertion (1), and a careful examination of the construction $A\to H$, outlined above, shows that the assertions (2,3) are in fact proved as well. $\square$

There are of course many other things that can be said about the bounded operators and the operator algebras, but for our purposes here, the above material, and especially the Gelfand theorem, will be basically all that we will need, in what follows. For more on all this, we refer as usual to our favorite analysis authors, namely Rudin [**74**] and Lax [**64**]. And for even more, this time in relation with physics, go with Connes [**24**].

8e. Exercises

The present chapter was an introduction to linear algebra in infinite dimensions, and most of our exercises here will be about straightforward continuations of all this. As a first exercise, dealing with the general theory of Hilbert spaces, we have:

EXERCISE 8.33. *Find an explicit orthonormal basis of the separable Hilbert space*

$$H = L^2[0,1]$$

by applying the Gram-Schmidt procedure to the polynomials $f_n = x^n$, with $n \in \mathbb{N}$.

This is something both fundamental and scary, and the answer can be found by doing an internet search with the keyword "orthogonal polynomials", then carefully reading what comes out of that, and adapting it if needed to the $H = L^2[0,1]$ situation.

EXERCISE 8.34. *Develop a theory of projections, isometries and symmetries inside $B(H)$, notably by examining the validity of the formula*

$$\lim_{n\to\infty}(PQ)^n = P \wedge Q$$

when talking about projections, and also by taking into account the fact that

$$UU^* = 1 \iff U^*U = 1$$

does not necessarily hold in infinite dimensions, when talking about isometries.

To be more precise, there are countless things to be done here. Regarding the projections, their definition should be $P^2 = P = P^*$, but there are a few things to be clarified in relation with the fact that an arbitrary subspace $K \subset H$ is not necessarily closed. Once all this understood, the limiting formula in the statement is a good object of study. As in what regards the isometries, and related operators such as the symmetries, here things are tricky as well, and the hint here would be to understand first when the operators UU^*, U^*U are projections, and what is the precise relation between these projections.

EXERCISE 8.35. *Prove that for the usual matrices $A, B \in M_N(\mathbb{C})$ we have*

$$\sigma^+(AB) = \sigma^+(BA)$$

where σ^+ denotes the set of eigenvalues, taken with multiplicities.

As a remark, we have seen in the above that $\sigma(AB) = \sigma(BA)$ holds outside $\{0\}$, and the equality on $\{0\}$ holds as well, because AB is invertible if and only if BA is invertible. However, in what regards the eigenvalues taken with multiplicities, things are more tricky here, and the answer should be somewhere inside your linear algebra knowledge.

EXERCISE 8.36. *Prove that an operator $T \in B(H)$ satisfies the condition*

$$< Tx, x > \geq 0$$

for any $x \in H$ precisely when it is positive in our sense, $\sigma(T) \subset [0,\infty)$.

In one sense this is normally something quite clear, and in the other sense this needs some tricks with vectors and scalar products, such as the polarization identity. Working out first the case of the usual matrices, $M \in M_N(\mathbb{C})$, with not much advanced linear algebra involved, is actually a good preliminary exercise.

EXERCISE 8.37. *Clarify, with examples and counterexamples, the relation between the eigenvalues of an operator $T \in B(H)$, and its spectrum $\sigma(T) \subset \mathbb{C}$.*

Here, as usual, the counterexamples could only come from the shift operator S, on the space $H = l^2(\mathbb{N})$. As a bonus exercise here, try computing the spectrum of S.

EXERCISE 8.38. *Assuming that an operator $T \in B(H)$ is normal, $TT^* = T^*T$, apply the Gelfand theorem to the C^*-algebra that it generates*

$$< T > \subset B(H)$$

in order to deduce a diagonalization theorem for T.

This is actually something quite heavy, so better start with an internet search.

EXERCISE 8.39. *Develop a theory of noncommutative geometry, by formally writing any C^*-algebra, not necessarily commutative, as*

$$A = C(X)$$

with X being a "compact quantum space", and report on what you found.

This is of course a very broad question, and almost anything would be welcome here, depending on your geometric taste and knowledge, plus of course on the ability to generalize what you know, to the case of the quantum spaces as defined above. We will be actually back to such questions and to C^*-algebras later on, when doing group theory.

Part III

Matrix groups

Castles out of fairy tales
Timbers shivered where once there sailed
The lovesick men who caught her eye
And no one knew but Lorelei

CHAPTER 9

Group theory

9a. Basic examples

We have seen so far the basics of linear algebra, with the conclusion that the theory is very useful, and quickly becomes non-trivial. We have seen as well some abstract applications, to questions in analysis and combinatorics, and with some results on the infinite dimensional case as well. All this is of course very useful in physics.

In this second half of this book we discuss a related topic, which is of key interest, namely the matrix groups. The theory here is once again very useful in connection with various questions in physics, the general idea being that any physical system S has a group of symmetries $G(S)$, whose study can lead to concrete conclusions about S.

Let us begin with some abstract aspects. A group is something very simple, namely a set, with a composition operation, which must satisfy what we should expect from a "multiplication". The precise definition of the groups is as follows:

Definition 9.1. *A group is a set G endowed with a multiplication operation*

$$(g,h)\to gh$$

which must satisfy the following conditions:

1. *Associativity: we have, $(gh)k=g(hk)$, for any $g,h,k\in G$.*
2. *Unit: there is an element $1\in G$ such that $g1=1g=g$, for any $g\in G$.*
3. *Inverses: for any $g\in G$ there is $g^{-1}\in G$ such that $gg^{-1}=g^{-1}g=1$.*

The multiplication law is not necessarily commutative. In the case where it is, in the sense that $gh=hg$, for any $g,h\in G$, we call G abelian, and we usually denote its multiplication, unit and inverse operation as follows:

$$(g,h)\to g+h$$

$$0\in G$$

$$g\to -g$$

However, this is not a general rule, and rather the converse is true, in the sense that if a group is denoted as above, this means that the group must be abelian.

At the level of examples, we have for instance the symmetric group S_N. There are many other examples, with typically the basic systems of numbers that we know being abelian groups, and the basic sets of matrices being non-abelian groups. Once again, this is of course not a general rule. Here are some basic examples and counterexamples:

PROPOSITION 9.2. *We have the following groups, and non-groups:*

(1) $(\mathbb{Z},+)$ *is a group.*
(2) $(\mathbb{Q},+)$, $(\mathbb{R},+)$, $(\mathbb{C},+)$ *are groups as well.*
(3) $(\mathbb{N},+)$ *is not a group.*
(4) $(\mathbb{Q}^*,\cdot)$ *is a group.*
(5) $(\mathbb{R}^*,\cdot)$, $(\mathbb{C}^*,\cdot)$ *are groups as well.*
(6) $(\mathbb{N}^*,\cdot)$, $(\mathbb{Z}^*,\cdot)$ *are not groups.*

PROOF. All this is clear from the definition of the groups, as follows:

(1) The group axioms are indeed satisfied for $\mathbb{Z}$, with the sum $g+h$ being the usual sum, 0 being the usual 0, and $-g$ being the usual $-g$.

(2) Once again, the axioms are satisfied for $\mathbb{Q},\mathbb{R},\mathbb{C}$, with the remark that for $\mathbb{Q}$ we are using here the fact that the sum of two rational numbers is rational, coming from:

$$\frac{a}{b}+\frac{c}{d}=\frac{ad+bc}{bd}$$

(3) In $\mathbb{N}$ we do not have inverses, so we do not have a group:

$$-1\notin\mathbb{N}$$

(4) The group axioms are indeed satisfied for $\mathbb{Q}^*$, with the product gh being the usual product, 1 being the usual 1, and g^{-1} being the usual g^{-1}. Observe that we must remove indeed the element $0\in\mathbb{Q}$, because in a group, any element must be invertible.

(5) Once again, the axioms are satisfied for $\mathbb{R}^*,\mathbb{C}^*$, with the remark that for $\mathbb{C}$ we are using here the fact that the nonzero complex numbers can be inverted, coming from:

$$z\bar{z}=|z|^2$$

(6) Here in $\mathbb{N}^*,\mathbb{Z}^*$ we do not have inverses, so we do not have groups, as claimed. □

There are many interesting groups coming from linear algebra, as follows:

THEOREM 9.3. *We have the following groups:*

(1) $(\mathbb{R}^N,+)$ *and* $(\mathbb{C}^N,+)$.
(2) $(M_N(\mathbb{R}),+)$ *and* $(M_N(\mathbb{C}),+)$.
(3) $(GL_N(\mathbb{R}),\cdot)$ *and* $(GL_N(\mathbb{C}),\cdot)$, *the invertible matrices.*
(4) $(SL_N(\mathbb{R}),\cdot)$ *and* $(SL_N(\mathbb{C}),\cdot)$, *with S standing for "special", meaning* $\det=1$.
(5) $(O_N,\cdot)$ *and* $(U_N,\cdot)$, *the orthogonal and unitary matrices.*
(6) $(SO_N,\cdot)$ *and* $(SU_N,\cdot)$, *with S standing as above for* $\det=1$.

PROOF. All this is clear from definitions, and from our linear algebra knowledge:

(1) The axioms are indeed clearly satisfied for $\mathbb{R}^N, \mathbb{C}^N$, with the sum being the usual sum of vectors, $-v$ being the usual $-v$, and the null vector 0 being the unit.

(2) Once again, the axioms are clearly satisfied for $M_N(\mathbb{R}), M_N(\mathbb{C})$, with the sum being the usual sum of matrices, $-M$ being the usual $-M$, and the null matrix 0 being the unit. Observe that what we have here is in fact a particular case of (1), because any $N \times N$ matrix can be regarded as a $N^2 \times 1$ vector, and so at the group level we have:

$$(M_N(\mathbb{R}), +) \simeq (\mathbb{R}^{N^2}, +)$$

$$(M_N(\mathbb{C}), +) \simeq (\mathbb{C}^{N^2}, +)$$

(3) Regarding now $GL_N(\mathbb{R}), GL_N(\mathbb{C})$, these are groups because the product of invertible matrices is invertible, according to the following formula:

$$(AB)^{-1} = B^{-1}A^{-1}$$

Observe that at $N = 1$ we obtain the groups $(\mathbb{R}^*, \cdot), (\mathbb{C}^*, \cdot)$. At $N \geq 2$ the groups $GL_N(\mathbb{R}), GL_N(\mathbb{C})$ are not abelian, because we do not have $AB = BA$ in general.

(4) The sets $SL_N(\mathbb{R}), SL_N(\mathbb{C})$ formed by the real and complex matrices of determinant 1 are subgroups of the groups in (3), because of the following formula, which shows that the matrices satisfying $\det A = 1$ are stable under multiplication:

$$\det(AB) = \det(A)\det(B)$$

(5) Regarding now O_N, U_N, here the group property is clear too from definitions, and is best seen by using the associated linear maps, because the composition of two isometries is an isometry. Equivalently, assuming $U^* = U^{-1}$ and $V^* = V^{-1}$, we have:

$$(UV)^* = V^*U^* = V^{-1}U^{-1} = (UV)^{-1}$$

(6) The sets of matrices SO_N, SU_N in the statement are obtained by intersecting the groups in (4) and (5), and so they are groups indeed:

$$SO_N = O_N \cap SL_N(\mathbb{R})$$

$$SU_N = U_N \cap SL_N(\mathbb{C})$$

Thus, all the sets in the statement are indeed groups, as claimed. □

We can see from the above that the notion of group is something extremely wide. In fact, in basically any mathematical question, there is a group somewhere around. Now back to Definition 9.1, at that level of generality, there is nothing much that can be said. Let us record, however, as our first theorem regarding the arbitrary groups:

THEOREM 9.4. *Given a group $(G, \cdot)$, we have the formula*

$$(g^{-1})^{-1} = g$$

valid for any element $g \in G$.

PROOF. This is clear from the definition of the inverses. Assume indeed that:

$$gg^{-1} = g^{-1}g = 1$$

But this shows that g is the inverse of g^{-1}, as claimed. □

As a comment here, the above result, while being something trivial, has led to a lot of controversy among mathematicians and physicists, in recent times. The point indeed is that, for the needs of quantum mechanics, the notion of group must be replaced with something more general, called "quantum group", and there are two schools here:

(1) Mathematicians believe that God is someone nasty, who created quantum mechanics by using some unfriendly quantum groups, satisfying $(g^{-1})^{-1} \neq g$.

(2) As for physicists, they are more relaxed, and prefer either to use beautiful quantum groups, satisfying $(g^{-1})^{-1} = g$, or not to use quantum groups at all.

You can check the book of Chari-Pressley [17] for a brief account of what can be done with $(g^{-1})^{-1} \neq g$, and my book [6] for what can be done with $(g^{-1})^{-1} = g$.

9b. Dihedral groups

In order to have now some theory going, we obviously have to impose some conditions on the groups that we consider. With this idea in mind, let us work out some examples, in the finite group case. The simplest possible finite group is the cyclic group $\mathbb{Z}_N$. There are many ways of picturing $\mathbb{Z}_N$, both additive and multiplicative, as follows:

DEFINITION 9.5. *The cyclic group $\mathbb{Z}_N$ is defined as follows:*

(1) *As the additive group of remainders modulo N.*
(2) *As the multiplicative group of the N-th roots of unity.*

The two definitions are equivalent, because if we set $w = e^{2\pi i/N}$, then any remainder modulo N defines a N-th root of unity, according to the following formula:

$$k \to w^k$$

We obtain in this way all the N-roots of unity, and so our correspondence is bijective. Moreover, our correspondence transforms the sum of remainders modulo N into the multiplication of the N-th roots of unity, due to the following formula:

$$w^k w^l = w^{k+l}$$

Thus, the groups defined in (1,2) above are isomorphic, via $k \to w^k$, and we agree to denote by $\mathbb{Z}_N$ the corresponding group. Observe that this group $\mathbb{Z}_N$ is abelian.

Another interesting example of a finite group, which is more advanced, and which is non-abelian this time, is the dihedral group D_N, which appears as follows:

Definition 9.6. *The dihedral group D_N is the symmetry group of the regular polygon having N vertices, which is called regular N-gon.*

Here are some basic examples of regular N-gons, at small values of the parameter $N \in \mathbb{N}$, and of their symmetry groups:

$\underline{N=2}$. Here the N-gon is just a segment, and its symmetries are the identity id and the obvious symmetry τ. Thus $D_2 = \{id, \tau\}$, and in group theory terms, $D_2 = \mathbb{Z}_2$.

$\underline{N=3}$. Here the N-gon is an equilateral triangle, and the symmetries are the $3! = 6$ possible permutations of the vertices. Thus we have $D_3 = S_3$.

$\underline{N=4}$. Here the N-gon is a square, and as symmetries we have 4 rotations, of angles $0°, 90°, 180°, 270°$, as well as 4 symmetries, with respect to the 4 symmetry axes, which are the 2 diagonals, and the 2 segments joining the midpoints of opposite sides.

$\underline{N=5}$. Here the N-gon is a regular pentagon, and as symmetries we have 5 rotations, of angles $0°, 72°, 144°, 216°, 288°$, as well as 5 symmetries, with respect to the 5 symmetry axes, which join the vertices to the midpoints of the opposite sides.

$\underline{N=6}$. Here the N-gon is a regular hexagon, and we have 6 rotations, of angles $0°, 60°, 120°, 180°, 240°, 300°$, and 6 symmetries, with respect to the 6 symmetry axes, which are the 3 diagonals, and the 3 segments joining the midpoints of opposite sides.

We can see from the above that the various dihedral groups D_N have many common features, and that there are some differences as well. In general, we have:

Proposition 9.7. *The dihedral group D_N has $2N$ elements, as follows:*

(1) *We have N rotations $R_1, \ldots, R_N$, with R_k being the rotation of angle $2k\pi/N$. When labelling the vertices of the N-gon $1, \ldots, N$, the rotation formula is:*
$$R_k : i \to k + i$$
(2) *We have N symmetries $S_1, \ldots, S_N$, with S_k being the symmetry with respect to the Ox axis rotated by $k\pi/N$. The symmetry formula is:*
$$S_k : i \to k - i$$

Proof. This is clear, indeed. To be more precise, D_N consists of:

(1) The N rotations, of angles $2k\pi/N$ with $k = 1, \ldots, N$. But these are exactly the rotations $R_1, \ldots, R_N$ from the statement.

(2) The N symmetries with respect to the N possible symmetry axes, which are the N medians of the N-gon when N is odd, and are the $N/2$ diagonals plus the $N/2$ lines connecting the midpoints of opposite edges, when N is even. But these are exactly the symmetries $S_1, \ldots, S_N$ from the statement. □

With the above description of D_N in hand, we can forget if we want about geometry and the regular N-gon, and talk about D_N abstractly, as follows:

THEOREM 9.8. *The dihedral group D_N is the group having $2N$ elements, $R_1, \ldots, R_N$ and $S_1, \ldots, S_N$, called rotations and symmetries, which multiply as follows,*

$$R_k R_l = R_{k+l}$$

$$R_k S_l = S_{k+l}$$

$$S_k R_l = S_{k-l}$$

$$S_k S_l = R_{k-l}$$

with all the indices being taken modulo N.

PROOF. With notations from Proposition 9.7, the various compositions between rotations and symmetries can be computed as follows:

$$R_k R_l \ : \ i \to l + i \to k + l + i$$

$$R_k S_l \ : \ i \to l - i \to k + l - i$$

$$S_k R_l \ : \ i \to l + i \to k - l - i$$

$$S_k S_l \ : \ i \to l - i \to k - l + i$$

But these are exactly the formulae for $R_{k+l}, S_{k+l}, S_{k-l}, R_{k-l}$, as stated. Now since a group is uniquely determined by its multiplication rules, this gives the result. $\square$

Observe that D_N has the same cardinality as $E_N = \mathbb{Z}_N \times \mathbb{Z}_2$. We obviously don't have $D_N \simeq E_N$, because D_N is not abelian, while E_N is. So, our next goal will be that of proving that D_N appears by "twisting" E_N. In order to do this, let us start with:

PROPOSITION 9.9. *The group $E_N = \mathbb{Z}_N \times \mathbb{Z}_2$ is the group having $2N$ elements, $r_1, \ldots, r_N$ and $s_1, \ldots, s_N$, which multiply according to the following rules,*

$$r_k r_l = r_{k+l}$$

$$r_k s_l = s_{k+l}$$

$$s_k r_l = s_{k+l}$$

$$s_k s_l = r_{k+l}$$

with all the indices being taken modulo N.

PROOF. With the notation $\mathbb{Z}_2 = \{1, \tau\}$, the elements of the product group $E_N = \mathbb{Z}_N \times \mathbb{Z}_2$ can be labelled $r_1, \ldots, r_N$ and $s_1, \ldots, s_N$, as follows:

$$r_k = (k, 1) \quad , \quad s_k = (k, \tau)$$

These elements multiply then according to the formulae in the statement. Now since a group is uniquely determined by its multiplication rules, this gives the result. $\square$

Let us compare now Theorem 9.8 and Proposition 9.9. In order to formally obtain D_N from E_N, we must twist some of the multiplication rules of E_N, namely:

$$s_k r_l = s_{k+l} \to s_{k-l}$$

$$s_k s_l = r_{k+l} \to r_{k-l}$$

Informally, this amounts in following the rule "τ switches the sign of what comes afterwards", and we are led in this way to the following definition:

Definition 9.10. *Given two groups A, G, with an action $A \curvearrowright G$, the crossed product*

$$P = G \rtimes A$$

is the set $G \times A$, with multiplication as follows:

$$(g, a)(h, b) = (gh^a, ab)$$

It is routine to check that P is indeed a group. Observe that when the action is trivial, $h^a = h$ for any $a \in A$ and $h \in H$, we obtain the usual product $G \times A$.

Now with this technology in hand, by getting back to the dihedral group D_N, we can improve Theorem 9.8, into a final result on the subject, as follows:

Theorem 9.11. *We have a crossed product decomposition as follows,*

$$D_N = \mathbb{Z}_N \rtimes \mathbb{Z}_2$$

with $\mathbb{Z}_2 = \{1, \tau\}$ acting on $\mathbb{Z}_N$ via switching signs, $k^\tau = -k$.

Proof. We have an action $\mathbb{Z}_2 \curvearrowright \mathbb{Z}_N$ given by the formula in the statement, namely $k^\tau = -k$, so we can consider the corresponding crossed product group:

$$P_N = \mathbb{Z}_N \rtimes \mathbb{Z}_2$$

In order to understand the structure of P_N, we follow Proposition 9.9. The elements of P_N can indeed be labelled $\rho_1, \ldots, \rho_N$ and $\sigma_1, \ldots, \sigma_N$, as follows:

$$\rho_k = (k, 1) \quad , \quad \sigma_k = (k, \tau)$$

Now when computing the products of such elements, we basically obtain the formulae in Proposition 9.9, perturbed as in Definition 9.10. To be more precise, we have:

$$\rho_k \rho_l = \rho_{k+l}$$

$$\rho_k \sigma_l = \sigma_{k+l}$$

$$\sigma_k \rho_l = \sigma_{k+l}$$

$$\sigma_k \sigma_l = \rho_{k+l}$$

But these are exactly the multiplication formulae for D_N, from Theorem 9.8. Thus, we have an isomorphism $D_N \simeq P_N$ given by $R_k \to \rho_k$ and $S_k \to \sigma_k$, as desired. $\square$

As a third basic example of a finite group, we have the symmetric group S_N. This is a group that we already met, when talking about the determinant:

THEOREM 9.12. *The permutations of* $\{1,\ldots,N\}$ *form a group, denoted* S_N*, and called symmetric group. This group has* $N!$ *elements. The signature map*

$$\varepsilon : S_N \to \mathbb{Z}_2$$

can be regarded as being a group morphism, with values in $\mathbb{Z}_2 = \{\pm 1\}$*.*

PROOF. These are things that we know from chapter 2. Indeed, the group property is clear, and the count is clear as well. As for the last assertion, recall the following formula for the signatures of the permutations, that we know as well from chapter 2:

$$\varepsilon(\sigma\tau) = \varepsilon(\sigma)\varepsilon(\tau)$$

But this tells us precisely that ε is a group morphism, as stated. $\square$

We will be back to S_N on many occasions, in what follows. At an even more advanced level now, we have the hyperoctahedral group H_N, which appears as follows:

DEFINITION 9.13. *The hyperoctahedral group* H_N *is the group of symmetries of the unit cube in* $\mathbb{R}^N$*.*

The hyperoctahedral group is a quite interesting group, whose definition, as a symmetry group, reminds that of the dihedral group D_N. So, let us start our study in the same way as we did for D_N, with a discussion at small values of $N \in \mathbb{N}$:

$\underline{N=1}$. Here the 1-cube is the segment, whose symmetries are the identity id and the flip τ. Thus, we obtain the group with 2 elements, which is a very familiar object:

$$H_1 = D_2 = S_2 = \mathbb{Z}_2$$

$\underline{N=2}$. Here the 2-cube is the square, and so the corresponding symmetry group is the dihedral group D_4, which is a group that we know well:

$$H_2 = D_4 = \mathbb{Z}_4 \rtimes \mathbb{Z}_2$$

$\underline{N=3}$. Here the 3-cube is the usual cube, and the situation is considerably more complicated, because this usual cube has no less than 48 symmetries. Identifying and counting these symmetries is actually an excellent exercise.

All this looks quite complicated, but fortunately we can count H_N, at $N = 3$, and at higher N as well, by using some tricks, the result being as follows:

THEOREM 9.14. *We have the cardinality formula*

$$|H_N| = 2^N N!$$

coming from the fact that H_N *is the symmetry group of the coordinate axes of* $\mathbb{R}^N$*.*

PROOF. This follows from some geometric thinking, as follows:

(1) Consider the standard cube in $\mathbb{R}^N$, centered at 0, and having as vertices the points having coordinates ± 1. With this picture in hand, it is clear that the symmetries of the cube coincide with the symmetries of the N coordinate axes of $\mathbb{R}^N$.

(2) In order to count now these latter symmetries, a bit as we did for the dihedral group, observe first that we have $N!$ permutations of these N coordinate axes.

(3) But each of these permutations of the coordinate axes $\sigma \in S_N$ can be further "decorated" by a sign vector $e \in \{\pm 1\}^N$, consisting of the possible ± 1 flips which can be applied to each coordinate axis, at the arrival. Thus, we have:

$$|H_N| = |S_N| \cdot |\mathbb{Z}_2^N| = N! \cdot 2^N$$

Thus, we are led to the conclusions in the statement. $\square$

As in the dihedral group case, it is possible to go beyond this, with a crossed product decomposition, of quite special type, called wreath product decomposition:

THEOREM 9.15. *We have a wreath product decomposition as follows,*

$$H_N = \mathbb{Z}_2 \wr S_N$$

which means by definition that we have a crossed product decomposition

$$H_N = \mathbb{Z}_2^N \rtimes S_N$$

with the permutations $\sigma \in S_N$ acting on the elements $e \in \mathbb{Z}_2^N$ as follows:

$$\sigma(e_1, \ldots, e_k) = (e_{\sigma(1)}, \ldots, e_{\sigma(k)})$$

PROOF. As explained in the proof of Theorem 9.14, the elements of H_N can be identified with the pairs $g = (e, \sigma)$ consisting of a permutation $\sigma \in S_N$, and a sign vector $e \in \mathbb{Z}_2^N$, so that at the level of the cardinalities, we have:

$$|H_N| = |\mathbb{Z}_2^N \times S_N|$$

To be more precise, given an element $g \in H_N$, the element $\sigma \in S_N$ is the corresponding permutation of the N coordinate axes, regarded as unoriented lines in $\mathbb{R}^N$, and $e \in \mathbb{Z}_2^N$ is the vector collecting the possible flips of these coordinate axes, at the arrival. Now observe that the product formula for two such pairs $g = (e, \sigma)$ is as follows, with the permutations $\sigma \in S_N$ acting on the elements $f \in \mathbb{Z}_2^N$ as in the statement:

$$(e, \sigma)(f, \tau) = (ef^\sigma, \sigma\tau)$$

Thus, we are precisely in the framework of Definition 9.10, and we conclude that we have a crossed product decomposition, as follows:

$$H_N = \mathbb{Z}_2^N \rtimes S_N$$

Thus, we are led to the conclusion in the statement, with the formula $H_N = \mathbb{Z}_2 \wr S_N$ being just a shorthand for the decomposition $H_N = \mathbb{Z}_2^N \rtimes S_N$ that we found. $\square$

Summarizing, we have so far many interesting examples of finite groups, and as a sequence of main examples, we have the following groups:

$$\mathbb{Z}_N \subset D_N \subset S_N \subset H_N$$

We will be back to these fundamental finite groups later on, on several occasions, with further results on them, both of algebraic and of analytic type.

9c. Cayley embeddings

At the level of the general theory now, we have the following fundamental result regarding the finite groups, due to Cayley:

THEOREM 9.16. *Given a finite group G, we have an embedding as follows,*

$$G \subset S_N \quad , \quad g \to (h \to gh)$$

with $N = |G|$. Thus, any finite group is a permutation group.

PROOF. Given a group element $g \in G$, we can associate to it the following map:

$$\sigma_g : G \to G \quad , \quad h \to gh$$

Since $gh = gh'$ implies $h = h'$, this map is bijective, and so is a permutation of G, viewed as a set. Thus, with $N = |G|$, we can view this map as a usual permutation, $\sigma_G \in S_N$. Summarizing, we have constructed so far a map as follows:

$$G \to S_N \quad , \quad g \to \sigma_g$$

Our first claim is that this is a group morphism. Indeed, this follows from:

$$\sigma_g \sigma_h(k) = \sigma_g(hk) = ghk = \sigma_{gh}(k)$$

It remains to prove that this group morphism is injective. But this follows from:

$$\begin{aligned} g \neq h \quad &\implies \quad \sigma_g(1) \neq \sigma_h(1) \\ &\implies \quad \sigma_g \neq \sigma_h \end{aligned}$$

Thus, we are led to the conclusion in the statement. $\square$

Observe that in the above statement the embedding $G \subset S_N$ that we constructed depends on a particular writing $G = \{g_1, \ldots, g_N\}$, which is needed in order to identify the permutations of G with the elements of the symmetric group S_N. This is not very good, in practice, and as an illustration, for the basic examples of groups that we know, the Cayley theorem provides us with embeddings as follows:

$$\mathbb{Z}_N \subset S_N \quad , \quad D_N \subset S_{2N} \quad , \quad S_N \subset S_{N!} \quad , \quad H_N \subset S_{2^N N!}$$

And here the first embedding is the good one, the second one is not the best possible one, but can be useful, and the third and fourth embeddings are useless. Thus, as a conclusion, the Cayley theorem remains something quite theoretical. We will be back to this later on, with a systematic study of the "representation" problem.

Getting back now to our main series of finite groups, $\mathbb{Z}_N \subset D_N \subset S_N \subset H_N$, these are of course permutation groups, according to the above. However, and perhaps even more interestingly, these are as well subgroups of the orthogonal group O_N:

$$\mathbb{Z}_N \subset D_N \subset S_N \subset H_N \subset O_N$$

Indeed, we have $H_N \subset O_N$, because any transformation of the unit cube in $\mathbb{R}^N$ must extend into an isometry of the whole $\mathbb{R}^N$, in the obvious way. Now in view of this, it makes sense to look at the finite subgroups $G \subset O_N$. With two remarks, namely:

(1) Although we do not have examples yet, following our general "complex is better than real" philosophy, it is better to look at the general subgroups $G \subset U_N$.

(2) Also, it is better to upgrade our study to the case where G is compact, and this in order to cover some interesting continuous groups, such as O_N, U_N, SO_N, SU_N.

Long story short, we are led in this way to the study of the closed subgroups $G \subset U_N$. Let us start our discussion here with the following simple fact:

PROPOSITION 9.17. *The closed subgroups $G \subset U_N$ are precisely the closed sets of matrices $G \subset U_N$ satisfying the following conditions:*

(1) $U, V \in G \implies UV \in G$.
(2) $1 \in G$.
(3) $U \in G \implies U^{-1} \in G$.

PROOF. This is clear from definitions, the only point with this statement being the fact that a subset $G \subset U_N$ can be a group or not, as indicated above. $\square$

It is possible to get beyond this, first with a result stating that any closed subgroup $G \subset U_N$ is a smooth manifold, and then with a result stating that, conversely, any smooth compact group appears as a closed subgroup $G \subset U_N$ of some unitary group. However, all this is quite advanced, and we will not need it, in what follows.

As a second result now regarding the closed subgroups $G \subset U_N$, let us prove that any finite group G appears in this way. This is something more or less clear from what we have, but let us make this precise. We first have the following key result:

THEOREM 9.18. *We have a group embedding as follows, obtained by regarding S_N as the permutation group of the N coordinate axes of $\mathbb{R}^N$,*

$$S_N \subset O_N$$

which makes $\sigma \in S_N$ correspond to the matrix having 1 on row i and column $\sigma(i)$, for any i, and having 0 entries elsewhere.

Proof. The first assertion is clear, because the permutations of the N coordinate axes of $\mathbb{R}^N$ are isometries. Regarding now the explicit formula, we have by definition:

$$\sigma(e_j) = e_{\sigma(j)}$$

Thus, the permutation matrix corresponding to σ is given by:

$$\sigma_{ij} = \begin{cases} 1 & \text{if } \sigma(j) = i \\ 0 & \text{otherwise} \end{cases}$$

Thus, we are led to the formula in the statement. □

We can combine the above result with the Cayley theorem, and we obtain the following result, which is something very nice, having theoretical importance:

Theorem 9.19. *Given a finite group G, we have an embedding as follows,*

$$G \subset O_N \quad , \quad g \to (e_h \to e_{gh})$$

with $N = |G|$. Thus, any finite group is an orthogonal matrix group.

Proof. The Cayley theorem gives an embedding as follows:

$$G \subset S_N \quad , \quad g \to (h \to gh)$$

On the other hand, Theorem 9.18 provides us with an embedding as follows:

$$S_N \subset O_N \quad , \quad \sigma \to (e_i \to e_{\sigma(i)})$$

Thus, we are led to the conclusion in the statement. □

The same remarks as for the Cayley theorem apply. First, the embedding $G \subset O_N$ that we constructed depends on a particular writing $G = \{g_1, \ldots, g_N\}$. And also, for the basic examples of groups that we know, the embeddings that we obtain are as follows:

$$\mathbb{Z}_N \subset O_N \quad , \quad D_N \subset O_{2N} \quad , \quad S_N \subset O_{N!} \quad , \quad H_N \subset O_{2^N N!}$$

As before, here the first embedding is the good one, the second one is not the best possible one, but can be useful, and the third and fourth embeddings are useless.

Summarizing, in order to advance, it is better to forget about the Cayley theorem, and build on Theorem 9.18 instead. In relation with the basic groups, we have:

Theorem 9.20. *We have the following finite groups of matrices:*

(1) $\mathbb{Z}_N \subset O_N$, *the cyclic permutation matrices.*
(2) $D_N \subset O_N$, *the dihedral permutation matrices.*
(3) $S_N \subset O_N$, *the permutation matrices.*
(4) $H_N \subset O_N$, *the signed permutation matrices.*

Proof. This is something self-explanatory, the idea being that Theorem 9.18 provides us with embeddings as follows, given by the permutation matrices:

$$\mathbb{Z}_N \subset D_N \subset S_N \subset O_N$$

In addition, looking back at the definition of H_N, this group inserts into the embedding on the right, $S_N \subset H_N \subset O_N$. Thus, we are led to the conclusion that all our 4 groups appear as groups of suitable "permutation type matrices". To be more precise:

(1) The cyclic permutation matrices are by definition the matrices as follows, with 0 entries elsewhere, and form a group, which is isomorphic to the cyclic group $\mathbb{Z}_N$:

$$U = \begin{pmatrix} & & & 1 & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \\ 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \end{pmatrix}$$

(2) The dihedral matrices are the above cyclic permutation matrices, plus some suitable symmetry permutation matrices, and form a group which is isomorphic to D_N.

(3) The permutation matrices, which by Theorem 9.18 form a group which is isomorphic to S_N, are the $0-1$ matrices having exactly one 1 on each row and column.

(4) Finally, regarding the signed permutation matrices, these are by definition the $(-1)-0-1$ matrices having exactly one nonzero entry on each row and column, and by Theorem 9.14 these matrices form a group, which is isomorphic to H_N. $\square$

The above groups are all groups of orthogonal matrices. When looking into general unitary matrices, we led to the following interesting class of groups:

Definition 9.21. *The complex reflection group $H_N^s \subset U_N$, depending on parameters*

$$N \in \mathbb{N} \quad , \quad s \in \mathbb{N} \cup \{\infty\}$$

is the group of permutation-type matrices with s-th roots of unity as entries,

$$H_N^s = M_N(\mathbb{Z}_s \cup \{0\}) \cap U_N$$

with the convention $\mathbb{Z}_\infty = \mathbb{T}$, at $s = \infty$.

Observe that at $s = 1, 2$ we obtain the following groups:

$$H_N^1 = S_N \quad , \quad H_N^2 = H_N$$

Another important particular case is $s = \infty$, where we obtain a group which is actually not finite, but is still compact, denoted as follows:

$$K_N \subset U_N$$

In general, in analogy with what we know about S_N,H_N, we first have:

PROPOSITION 9.22. *The number of elements of H_N^s with $s\in\mathbb N$ is:*

$$|H_N^s|=s^NN!$$

At $s=\infty$, the group $K_N=H_N^\infty$ that we obtain is infinite.

PROOF. This is indeed clear from our definition of H_N^s, as a matrix group as above, because there are $N!$ choices for a permutation-type matrix, and then s^N choices for the corresponding s-roots of unity, which must decorate the N nonzero entries. □

Once again in analogy with what we know at $s=1,2$, we have as well:

THEOREM 9.23. *We have a wreath product decomposition as follows,*

$$H_N^s=\mathbb Z_s\wr S_N$$

which means by definition that we have a crossed product decomposition

$$H_N^s=\mathbb Z_s^N\rtimes S_N$$

with the permutations $\sigma\in S_N$ acting on the elements $e\in\mathbb Z_s^N$ as follows:

$$\sigma(e_1,\ldots,e_k)=(e_{\sigma(1)},\ldots,e_{\sigma(k)})$$

PROOF. As explained in the proof of Proposition 9.22, the elements of H_N^s can be identified with the pairs $g=(e,\sigma)$ consisting of a permutation $\sigma\in S_N$, and a decorating vector $e\in\mathbb Z_s^N$, so that at the level of the cardinalities, we have:

$$|H_N|=|\mathbb Z_s^N\times S_N|$$

Now observe that the product formula for two such pairs $g=(e,\sigma)$ is as follows, with the permutations $\sigma\in S_N$ acting on the elements $f\in\mathbb Z_s^N$ as in the statement:

$$(e,\sigma)(f,\tau)=(ef^\sigma,\sigma\tau)$$

Thus, we are in the framework of Definition 9.10, and we obtain $H_N^s=\mathbb Z_s^N\rtimes S_N$. But this can be written, by definition, as $H_N^s=\mathbb Z_s\wr S_N$, and we are done. □

Summarizing, and by focusing now on the cases $s=1,2,\infty$, which are the most important, we have extended our series of basic unitary groups, as follows:

$$\mathbb Z_N\subset D_N\subset S_N\subset H_N\subset K_N$$

In addition to this, we have the groups H_N^s with $s\in\{3,4,\ldots,\}$. However, these will not fit well into the above series of inclusions, because we only have:

$$s|t\implies H_N^s\subset H_N^t$$

Thus, we can only extend our series of inclusions as follows:

$$\mathbb Z_N\subset D_N\subset S_N\subset H_N\subset H_N^4\subset H_N^8\subset\ldots\ldots\subset K_N$$

We will be back later to H_N^s, with more theory, and some generalizations as well.

9d. Abelian groups

We have seen so far that the basic examples of groups, even taken finite, lead us into linear algebra, and more specifically, into the study of groups of unitary matrices:

$$G \subset U_N$$

This is indeed a good idea, and we will systematically do this in this book, starting from the next chapter. Before getting into this, however, let us go back to the definition of the abstract groups, from the beginning of this chapter, and make a last attempt of developing some useful general theory there, without relation to linear algebra.

Basic common sense suggests looking into the case of the finite abelian groups, which can only be far less complicated than the arbitrary finite groups. However, as somewhat a surprise, this leads us again into linear algebra, due to the following fact:

THEOREM 9.24. *Let us call representation of a finite group G any morphism*

$$u : G \to U_N$$

to a unitary group. Then the 1-dimensional representations are the morphisms

$$\chi : G \to \mathbb{T}$$

called characters of G, and these characters form a finite abelian group $\widehat{G}$.

PROOF. Regarding the first assertion, this is just some philosophy, making the link with matrices and linear algebra, and coming from $U_1 = \mathbb{T}$. So, let us prove now the second assertion, stating that the set of characters $\widehat{G} = \{\chi : G \to \mathbb{T}\}$ is a finite abelian group. There are several things to be proved here, the idea being as follows:

(1) Our first claim is that $\widehat{G}$ is a group, with the pointwise multiplication, namely:

$$(\chi\rho)(g) = \chi(g)\rho(g)$$

Indeed, if χ, ρ are characters, so is $\chi\rho$, and so the multiplication is well-defined on $\widehat{G}$. Regarding the unit, this is the trivial character, constructed as follows:

$$1 : G \to \mathbb{T} \quad , \quad g \to 1$$

Finally, we have inverses, with the inverse of $\chi : G \to \mathbb{T}$ being its conjugate:

$$\bar{\chi} : G \to \mathbb{T} \quad , \quad g \to \overline{\chi(g)}$$

(2) Our next claim is that $\widehat{G}$ is finite. Indeed, given a group element $g \in G$, we can talk about its order, which is smallest integer $k \in \mathbb{N}$ such that $g^k = 1$. Now assuming that we have a character $\chi : G \to \mathbb{T}$, we have the following formula:

$$\chi(g)^k = 1$$

Thus $\chi(g)$ must be one of the k-th roots of unity, and in particular there are finitely many choices for $\chi(g)$. Thus, there are finitely many choices for χ, as desired.

(3) Finally, the fact that $\widehat{G}$ is abelian follows from definitions, because the pointwise multiplication of functions, and in particular of characters, is commutative. $\square$

The above construction is quite interesting, especially in the case where the starting finite group G is abelian itself, and as an illustration here, we have:

THEOREM 9.25. *The character group operation $G \to \widehat{G}$ for the finite abelian groups, called Pontrjagin duality, has the following properties:*

(1) *The dual of a cyclic group is the group itself, $\widehat{\mathbb{Z}}_N = \mathbb{Z}_N$.*
(2) *The dual of a product is the product of duals, $\widehat{G \times H} = \widehat{G} \times \widehat{H}$.*
(3) *Any product of cyclic groups $G = \mathbb{Z}_{N_1} \times \ldots \times \mathbb{Z}_{N_k}$ is self-dual, $G = \widehat{G}$.*

PROOF. We have several things to be proved, the idea being as follows:

(1) A character $\chi : \mathbb{Z}_N \to \mathbb{T}$ is uniquely determined by its value $z = \chi(g)$ on the standard generator $g \in \mathbb{Z}_N$. But this value must satisfy:

$$z^N = 1$$

Thus we must have $z \in \mathbb{Z}_N$, with the cyclic group $\mathbb{Z}_N$ being regarded this time as being the group of N-th roots of unity. Now conversely, any N-th root of unity $z \in \mathbb{Z}_N$ defines a character $\chi : \mathbb{Z}_N \to \mathbb{T}$, by setting, for any $r \in \mathbb{N}$:

$$\chi(g^r) = z^r$$

Thus we have an identification $\widehat{\mathbb{Z}}_N = \mathbb{Z}_N$, as claimed.

(2) A character of a product of groups $\chi : G \times H \to \mathbb{T}$ must satisfy:

$$\chi(g,h) = \chi\left[(g,1)(1,h)\right] = \chi(g,1)\chi(1,h)$$

Thus χ must appear as the product of its restrictions $\chi_{|G}, \chi_{|H}$, which must be both characters, and this gives the identification in the statement.

(3) This follows from (1) and (2). Alternatively, any character $\chi : G \to \mathbb{T}$ is uniquely determined by its values $\chi(g_1), \ldots, \chi(g_k)$ on the standard generators of $\mathbb{Z}_{N_1}, \ldots, \mathbb{Z}_{N_k}$, which must belong to $\mathbb{Z}_{N_1}, \ldots, \mathbb{Z}_{N_k} \subset \mathbb{T}$, and this gives $\widehat{G} = G$, as claimed. $\square$

We can get some further insight into duality by using the spectral theory methods developed in chapter 8, and we have the following result:

THEOREM 9.26. *Given a finite abelian group G, we have an isomorphism of commutative C^*-algebras as follows, obtained by linearizing/delinearizing the characters:*

$$\mathbb{C}[G] \simeq C(\widehat{G})$$

Also, the Pontrjagin duality is indeed a duality, in the sense that we have $G = \widehat{\widehat{G}}$.

PROOF. We have several assertions here, the idea being as follows:

(1) Given a finite abelian group G, consider indeed the group algebra $\mathbb{C}[G]$, having as elements the formal combinations of elements of G, and with involution given by:

$$g^* = g^{-1}$$

This $*$-algebra is then a C^*-algebra, with norm coming by acting $\mathbb{C}[G]$ on itself, and so by the Gelfand theorem we obtain an isomorphism as follows:

$$\mathbb{C}[G] = C(X)$$

To be more precise, X is the space of the $*$-algebra characters as follows:

$$\chi : \mathbb{C}[G] \to \mathbb{C}$$

The point now is that by delinearizing, such a $*$-algebra character must come from a usual group character of G, obtained by restricting to G, as follows:

$$\chi : G \to \mathbb{T}$$

Thus we have $X = \widehat{G}$, and we are led to the isomorphism in the statement, namely:

$$\mathbb{C}[G] \simeq C(\widehat{G})$$

(2) In order to prove now the second assertion, consider the following group morphism, which is available for any finite group G, not necessarily abelian:

$$G \to \widehat{\widehat{G}} \quad , \quad g \to (\chi \to \chi(g))$$

Our claim is that in the case where G is abelian, this is an isomorphism. As a first observation, we only need to prove that this morphism is injective or surjective, because the cardinalities match, according to the following formula, coming from (1):

$$|G| = \dim \mathbb{C}[G] = \dim C(\widehat{G}) = |\widehat{G}|$$

(3) We will prove that the above morphism is injective. For this purpose, let us compute its kernel. We know that $g \in G$ is in the kernel when the following happens:

$$\chi(g) = 1 \quad , \quad \forall \chi \in \widehat{G}$$

But this means precisely that $g \in \mathbb{C}[G]$ is mapped, via the isomorphism $\mathbb{C}[G] \simeq C(\widehat{G})$ constructed in (1), to the constant function $1 \in C(\widehat{G})$, and now by getting back to $\mathbb{C}[G]$ via our isomorphism, this shows that we have indeed $g = 1$, which ends the proof. $\square$

All the above is very nice, but remains something rather abstract, based on all sorts of clever algebraic manipulations, and no computations at all. So, now that we are done with this, time to get into some serious computations. For this purpose, we will need some basic abstract results, which are good to know. Let us start with:

Theorem 9.27. *Given a finite group G and a subgroup $H \subset G$, the sets*

$$G/H = \left\{gH\Big|g \in G\right\} \quad , \quad H\backslash G = \left\{Hg\Big|g \in G\right\}$$

both consist of partitions of G into subsets of size H, and we have the formula

$$|G| = |H| \cdot |G/H| = |H| \cdot |H\backslash G|$$

which shows that the order of the subgroup divides the order of the group:

$$|H| \Big| |G|$$

When $H \subset G$ is normal, $gH = Hg$ for any $g \in G$, the space $G/H = H\backslash G$ is a group.

Proof. There are several assertions here, but these are all trivial, when deduced in the precise order indicated in the statement. To be more precise, the partition claim for G/H can be deduced as follows, and the proof for $H\backslash G$ is similar:

$$gH \cap kH \neq \emptyset \iff g^{-1}k \in H \iff gH = kH$$

With this in hand, the cardinality formulae are all clear, and it remains to prove the last assertion. But here, the point is that when $H \subset G$ is normal, we have:

$$gH = kH, sH = tH \implies gsH = gtH = gHt = kHt = ktH$$

Thus $G/H = H\backslash G$ is a indeed group, with multiplication $(gH)(sH) = gsH$. □

As a main consequence of the above result, which is equally famous, we have:

Theorem 9.28. *Given a finite group G, any $g \in G$ generates a cyclic subgroup*

$$< g >= \{1, g, g^2, \ldots, g^{k-1}\}$$

with $k = ord(g)$ being the smallest number $k \in \mathbb{N}$ satisfying $g^k = 1$. Also, we have

$$ord(g) \Big| |G|$$

that is, the order of any group element divides the order of the group.

Proof. As before with Theorem 9.27, we have opted here for a long collection of statements, which are all trivial, when deduced in the above precise order. To be more precise, consider the semigroup $< g >\subset G$ formed by the sequence of powers of g:

$$< g >= \{1, g, g^2, g^3, \ldots\} \subset G$$

Since G was assumed to be finite, the sequence of powers must cycle, $g^n = g^m$ for some $n < m$, and so we have $g^k = 1$, with $k = m - n$. Thus, we have in fact:

$$< g >= \{1, g, g^2, \ldots, g^{k-1}\}$$

Moreover, we can choose $k \in \mathbb{N}$ to be minimal with this property, and with this choice, we have a set without repetitions. Thus $< g >\subset G$ is indeed a group, and more specifically a cyclic group, of order $k = ord(g)$. Finally, $ord(g) \,|\, |G|$ follows from Theorem 9.27. □

With these ingredients in hand, we can go back to the finite abelian groups. We have the following result, which is something remarkable, refining all the above:

THEOREM 9.29. *The finite abelian groups are the following groups,*

$$G = \mathbb{Z}_{N_1} \times \ldots \times \mathbb{Z}_{N_k}$$

and these groups are all self-dual, $G = \widehat{G}$.

PROOF. This is something quite tricky, the idea being as follows:

(1) In order to prove our result, assume that G is finite and abelian. For any prime number $p \in \mathbb{N}$, let us define $G_p \subset G$ to be the subset of elements having as order a power of p. Equivalently, this subset $G_p \subset G$ can be defined as follows:

$$G_p = \left\{ g \in G \Big| \exists k \in \mathbb{N}, g^{p^k} = 1 \right\}$$

(2) It is then routine to check, based on definitions, that each G_p is a subgroup. Our claim now is that we have a direct product decomposition as follows:

$$G = \prod_p G_p$$

(3) Indeed, by using the fact that our group G is abelian, we have a morphism as follows, with the order of the factors when computing $\prod_p g_p$ being irrelevant:

$$\prod_p G_p \to G \quad , \quad (g_p) \to \prod_p g_p$$

Moreover, it is routine to check that this morphism is both injective and surjective, via some simple manipulations, so we have our group decomposition, as in (2).

(4) Thus, we are left with proving that each component G_p decomposes as a product of cyclic groups, having as orders powers of p, as follows:

$$G_p = \mathbb{Z}_{p^{r_1}} \times \ldots \times \mathbb{Z}_{p^{r_s}}$$

But this is something that can be checked by recurrence on $|G_p|$, via some routine computations, and so we are led to the conclusion in the statement.

(5) Finally, the fact that the finite abelian groups are self-dual, $G = \widehat{G}$, follows from the structure result that we just proved, and from Theorem 9.25 (3). $\square$

So long for finite abelian groups. All the above was of course a bit quick, and for further details on all this, and especially on Theorem 9.29, which is something non-trivial, and for some generalizations as well, to the case of suitable non-finite abelian groups, we refer to the algebra book of Lang [62], where all this material is carefully explained.

As an application of the above, let us go back to the generalized Fourier matrices, from chapter 7. We have here the following result:

Theorem 9.30. *Given a finite abelian group G, with dual group $\widehat{G}=\{\chi:G\to\mathbb{T}\}$, consider the corresponding Fourier coupling, namely:*

$$\mathcal{F}_G:G\times\widehat{G}\to\mathbb{T}\quad,\quad(i,\chi)\to\chi(i)$$

1. *Via the standard isomorphism $G\simeq\widehat{G}$, this Fourier coupling can be regarded as a square matrix, $F_G\in M_G(\mathbb{T})$, which is a complex Hadamard matrix.*
2. *In the case of the cyclic group $G=\mathbb{Z}_N$ we obtain in this way, via the standard identification $\mathbb{Z}_N=\{1,\ldots,N\}$, the Fourier matrix F_N.*
3. *In general, when using a decomposition $G=\mathbb{Z}_{N_1}\times\ldots\times\mathbb{Z}_{N_k}$, the corresponding Fourier matrix is given by $F_G=F_{N_1}\otimes\ldots\otimes F_{N_k}$.*

Proof. This follows indeed by using the above finite abelian group theory:

(1) With the identification $G\simeq\widehat{G}$ made our matrix is given by $(F_G)_{i\chi}=\chi(i)$, and the scalar products between the rows are computed as follows:

$$\begin{aligned}<R_i,R_j> &= \sum_\chi\chi(i)\overline{\chi(j)}\\ &= \sum_\chi\chi(i-j)\\ &= |G|\cdot\delta_{ij}\end{aligned}$$

Thus, we obtain indeed a complex Hadamard matrix.

(2) This follows from the well-known and elementary fact that, via the identifications $\mathbb{Z}_N=\widehat{\mathbb{Z}_N}=\{1,\ldots,N\}$, the Fourier coupling here is as follows, with $w=e^{2\pi i/N}$:

$$(i,j)\to w^{ij}$$

(3) We use here the following formula that we know, for the duals of products:

$$\widehat{H\times K}=\widehat{H}\times\widehat{K}$$

At the level of the corresponding Fourier couplings, we obtain from this:

$$F_{H\times K}=F_H\otimes F_K$$

Now by decomposing G into cyclic groups, as in the statement, and by using (2) for the cyclic components, we obtain the formula in the statement. □

As a nice application of the above result, we have:

Proposition 9.31. *The Walsh matrix, W_N with $N=2^n$, which is given by*

$$W_N=\begin{pmatrix}1&1\\1&-1\end{pmatrix}^{\otimes n}$$

is the Fourier matrix of the finite abelian group $K_N=\mathbb{Z}_2^n$.

PROOF. We know that the first Walsh matrix is a Fourier matrix:

$$W_2 = F_2 = F_{K_2}$$

Now by taking tensor powers we obtain from this that we have, for any $N = 2^n$:

$$W_N = W_2^{\otimes n} = F_{K_2}^{\otimes n} = F_{K_2^n} = F_{K_N}$$

Thus, we are led to the conclusion in the statement. $\square$

Summarizing, we have now a better understanding of the generalized Fourier matrices, and of the complex Hadamard matrices in general, and also a new and fresh point of view on the various discrete Fourier analysis considerations from chapter 7.

9e. Exercises

There are many things that can be said about groups, especially in the matrix case, $G \subset U_N$, and we will discuss this later in this book. Our exercises here will rather focus on the abstract groups, as in the end of the present chapter, and we first have:

EXERCISE 9.32. *Develop a full theory of the order of group elements,*

$$ord(g) \in \mathbb{N} \cup \{\infty\}$$

with examples and results, both in the finite and infinite group case.

Here we are a bit vague, but there are just so many things that can be done. As an example here, try to understand when $|G|$ can be a power of a prime.

EXERCISE 9.33. *Prove with full details that any finite abelian group decomposes as*

$$G = \mathbb{Z}_{N_1} \times \ldots \times \mathbb{Z}_{N_k}$$

either by using abstract algebra, or spectral theory and matrix theory.

To be more precise, the abstract algebra approach was already outlined in the above, and the problem is that of clarifying the 2 steps mentioned there, by working out the previous exercise first, as a useful preliminary. It is possible as well to prove the above result by building on our spectral theory study, but this is more tricky.

EXERCISE 9.34. *Given a locally compact abelian group G, prove that its group characters, which must be by definition continuous,*

$$\chi : G \to \mathbb{T}$$

form a locally compact abelian group, denoted $\widehat{G}$, and called dual of G.

Here locally compact means that any group element $g \in G$ has a neighborhood which is compact, a bit in analogy with what happens for the real numbers $r \in \mathbb{R}$.

EXERCISE 9.35. *Prove that the integers are dual to the unit circle, and vice versa:*

$$\widehat{\mathbb{Z}}=\mathbb{T}\quad,\quad\widehat{\mathbb{T}}=\mathbb{Z}$$

Also, prove that the group of real numbers is self-dual, $\widehat{\mathbb{R}}=\mathbb{R}$.

To be more precise, we already know from the above that we have $\widehat{\mathbb{Z}}_N=\mathbb{Z}_N$, for any $N\in\mathbb{N}$, and the first question, regarding $\mathbb{Z}$ and $\mathbb{T}$, is a kind of "$N=\infty$" version of this. As for the second question, regarding $\mathbb{R}$, this is related to all this as well.

EXERCISE 9.36. *Prove that any abelian group G has the following property:*

$$G=\widehat{\widehat{G}}$$

Also, prove that we have an isomorphism of commutative C^*-algebras as follows,

$$C^*(G)\simeq C(\widehat{G})$$

with the integration over $\widehat{G}$ corresponding to the functional $\int_{\widehat{G}}g=\delta_{g1}$.

This is something that we already know from the above, in the case where G is finite, and the problem is that of carefully working out the extension to the compact setting, notably by clarifying what the algebra $C^*(G)$ should exactly mean. As a bonus exercise, try as well to state and prove a similar result in the general, locally compact setting.

EXERCISE 9.37. *Prove that the finitely generated abelian groups are*

$$G=\mathbb{Z}_{N_1}\times\ldots\times\mathbb{Z}_{N_k}$$

with the convention $\mathbb{Z}_\infty=\mathbb{Z}$, and that the compact matrix abelian groups are

$$H=\mathbb{Z}_{N_1}\times\ldots\times\mathbb{Z}_{N_k}$$

with this time the convention $\mathbb{Z}_\infty=\mathbb{T}$. Also, prove that $G=\widehat{H}$ and $H=\widehat{G}$.

This exercise, generalizing everything that we know, or almost, is actually something quite tricky, requiring a good knowledge of both algebra and analysis.

EXERCISE 9.38. *Clarify the relation between the dualities*

$$\widehat{\mathbb{Z}}_N=\mathbb{Z}_N$$
$$\widehat{\mathbb{Z}}=\mathbb{T}\quad,\quad\widehat{\mathbb{T}}=\mathbb{Z}$$
$$\widehat{\mathbb{R}}=\mathbb{R}$$

and the various types of Fourier transforms available.

To be more precise here, the problem is that of understanding why the above 3 dualities correspond to the main 3 types of known Fourier transforms, namely discrete Fourier transforms, usual Fourier series, and usual Fourier transforms. With the remark that this is something that we already know, for the first duality.

CHAPTER 10

Rotation groups

10a. Rotation groups

We have seen that there are many interesting examples of finite groups G, which usually appear as groups of unitary matrices, $G \subset U_N$. In this chapter we discuss similar questions, in the continuous case. Things here are quite tricky, and can quickly escalate into complicated mathematics, and we have several schools of thought, as follows:

(1) Mathematicians are usually interested in classifying everything, no matter what, and for this purpose, classifying the continuous groups $G \subset U_N$, a good method is that of looking at the tangent space at the unit, $\mathfrak{g} = T_1(G)$, called Lie algebra of G. This reduces the classification problem to a linear algebra problem, namely the classification of the Lie algebras $\mathfrak{g}$, and this latter problem can indeed be solved. Which is very nice.

(2) Applied mathematicians and physicists, on their side, already know what the groups $G \subset U_N$ that they are interested in are, and so are mildly enthusiastic about deep classification results. What they are interested in, however, are "tools" in order to deal with the groups $G \subset U_N$ that they have in mind. And here, surely the Lie algebra $\mathfrak{g}$ can be of help, but there are countless other things, that can be of help too.

(3) To be more precise, a very efficient tool in order to deal with the groups $G \subset U_N$ is representation theory, with a touch of probability, and there is a whole string of results here, which are somehow rival to the standard Lie algebra theory, that can be developed, including the theory of the Haar measure, then Peter-Weyl theory, then Tannakian duality, then Brauer algebras and easiness, and then Weingarten calculus.

(4) So, this is the situation, at least at the beginner level, with two ways to be followed, on one hand the Lie algebra way, linearizing the group in the simplest way, by looking at the tangent space at the unit, $\mathfrak{g} = T_1(G)$, and on the other hand with the representation theory way, linearizing the group in a more complicated, yet equally natural way, via representation theory invariants, such as the associated Brauer algebras.

In this book we will rather follow the representation theory way, with this being a personal choice. Let me mention too that at the advanced level the theory of Lie algebras and Brauer algebras is more or less the same thing, so in the end, things fine.

This being said, and before getting started, some references too. Algebra in general is a quite wide topic, and for the basics, of all kinds, you have the book of Lang [62]. Then, speaking algebra, you definitely need to learn some algebraic geometry, which is some extremely beautiful, and useful, and classical, with standard references here being the books of Shafarevich [81] and Harris [46]. And finally, in what regards groups, and various methods in order to deal with them, sometimes complementary to what we will be doing here, good references are the books of Humphreys [50] and Serre [79].

Back to work now, we will be mainly interested in the unitary group U_N itself, in its real version, which is the orthogonal group O_N, and in various technical versions of these basic groups O_N, U_N. So, let us start with some reminders, regarding O_N, U_N:

THEOREM 10.1. *We have the following results:*

(1) *The rotations of $\mathbb{R}^N$ form the orthogonal group O_N, which is given by:*

$$O_N = \left\{U \in M_N(\mathbb{R})\Big|U^t = U^{-1}\right\}$$

(2) *The rotations of $\mathbb{C}^N$ form the unitary group U_N, which is given by:*

$$U_N = \left\{U \in M_N(\mathbb{C})\Big|U^* = U^{-1}\right\}$$

In addition, we can restrict the attention to the rotations of the corresponding spheres.

PROOF. This is something that we already know, the idea being as follows:

(1) We have seen in chapter 1 that a linear map $T : \mathbb{R}^N \to \mathbb{R}^N$, written as $T(x) = Ux$ with $U \in M_N(\mathbb{R})$, is a rotation, in the sense that it preserves the distances and the angles, precisely when the associated matrix U is orthogonal, in the following sense:

$$U^t = U^{-1}$$

Thus, we obtain the result. As for the last assertion, this is clear as well, because an isometry of $\mathbb{R}^N$ is the same as an isometry of the unit sphere $S^{N-1}_{\mathbb{R}} \subset \mathbb{R}^N$.

(2) We have seen in chapter 3 that a linear map $T : \mathbb{C}^N \to \mathbb{C}^N$, written as $T(x) = Ux$ with $U \in M_N(\mathbb{C})$, is a rotation, in the sense that it preserves the distances and the scalar products, precisely when the associated matrix U is unitary, in the following sense:

$$U^* = U^{-1}$$

Thus, we obtain the result. As for the last assertion, this is clear as well, because an isometry of $\mathbb{C}^N$ is the same as an isometry of the unit sphere $S^{N-1}_{\mathbb{C}} \subset \mathbb{C}^N$. □

In order to introduce some further continuous groups $G \subset U_N$, we will need:

PROPOSITION 10.2. *We have the following results:*

(1) *For an orthogonal matrix $U \in O_N$ we have $\det U \in \{\pm 1\}$.*
(2) *For a unitary matrix $U \in U_N$ we have $\det U \in \mathbb{T}$.*

PROOF. This is clear from the equations defining O_N, U_N, as follows:

(1) We have indeed the following implications:

$$\begin{aligned} U \in O_N &\implies U^t = U^{-1} \\ &\implies \det U^t = \det U^{-1} \\ &\implies \det U = (\det U)^{-1} \\ &\implies \det U \in \{\pm 1\} \end{aligned}$$

(2) We have indeed the following implications:

$$\begin{aligned} U \in U_N &\implies U^* = U^{-1} \\ &\implies \det U^* = \det U^{-1} \\ &\implies \overline{\det U} = (\det U)^{-1} \\ &\implies \det U \in \mathbb{T} \end{aligned}$$

Here we have used the fact that $\bar{z} = z^{-1}$ means $z\bar{z} = 1$, and so $z \in \mathbb{T}$. $\square$

We can now introduce the subgroups $SO_N \subset O_N$ and $SU_N \subset U_N$, as being the subgroups consisting of the rotations which preserve the orientation, as follows:

THEOREM 10.3. *The following are groups of matrices,*

$$SO_N = \left\{ U \in O_N \Big| \det U = 1 \right\}$$

$$SU_N = \left\{ U \in U_N \Big| \det U = 1 \right\}$$

consisting of the rotations which preserve the orientation.

PROOF. The fact that we have indeed groups follows from the properties of the determinant, of from the property of preserving the orientation, which is clear as well. $\square$

Summarizing, we have constructed so far 4 continuous groups of matrices, consisting of various rotations, with inclusions between them, as follows:

$$\begin{tikzcd} SU_N \arrow[r] & U_N \\ SO_N \arrow[r] \arrow[u] & O_N \arrow[u] \end{tikzcd}$$

Observe that this is an intersection diagram, in the sense that:

$$SO_N = SU_N \cap O_N$$

As an illustration, let us work out what happens at $N = 1, 2$. At $N = 1$ the situation is quite trivial, and we obtain very simple groups, as follows:

Proposition 10.4. *The basic continuous groups at $N=1$, namely*

$$\begin{tikzcd} SU_1 \arrow[r] & U_1 \\ SO_1 \arrow[u] \arrow[r] & O_1 \arrow[u] \end{tikzcd}$$

are the following groups of complex numbers,

$$\begin{tikzcd} \{1\} \arrow[r] & \mathbb{T} \\ \{1\} \arrow[u] \arrow[r] & \{\pm 1\} \arrow[u] \end{tikzcd}$$

or, equivalently, are the following cyclic groups,

$$\begin{tikzcd} \mathbb{Z}_1 \arrow[r] & \mathbb{Z}_\infty \\ \mathbb{Z}_1 \arrow[u] \arrow[r] & \mathbb{Z}_2 \arrow[u] \end{tikzcd}$$

with the convention that $\mathbb{Z}_s$ is the group of s-th roots of unity.

Proof. This is clear from definitions, because for a 1×1 matrix the unitarity condition reads $\bar{U} = U^{-1}$, and so $U \in \mathbb{T}$, and this gives all the results. $\square$

At $N=2$ now, let us first discuss the real case. The result here is as follows:

Theorem 10.5. *We have the following results:*

(1) *SO_2 is the group of usual rotations in the plane, which are given by:*

$$R_t = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}$$

(2) *O_2 consists in addition of the usual symmetries in the plane, given by:*

$$S_t = \begin{pmatrix} \cos t & \sin t \\ \sin t & -\cos t \end{pmatrix}$$

(3) *Abstractly speaking, we have isomorphisms as follows:*

$$SO_2 \simeq \mathbb{T} \quad , \quad O_2 = \mathbb{T} \rtimes \mathbb{Z}_2$$

(4) *When discretizing all this, by replacing the 2-dimensional unit sphere $\mathbb{T}$ by the regular N-gon, the latter isomorphism discretizes as $D_N = \mathbb{Z}_N \rtimes \mathbb{Z}_2$.*

PROOF. This follows from some elementary computations, as follows:

(1) The first assertion is clear, because only the rotations of the plane in the usual sense preserve the orientation. As for the formula of R_t, this is something that we already know, from chapter 1, obtained by computing $R_t\binom{1}{0}$ and $R_t\binom{0}{1}$.

(2) The first assertion is clear, because rotations left aside, we are left with the symmetries of the plane, in the usual sense. As for formula of S_t, this is something that we basically know too, obtained by computing $S_t\binom{1}{0}$ and $S_t\binom{0}{1}$.

(3) The first assertion is clear, because the angles $t \in \mathbb{R}$, taken as usual modulo 2π, form the group $\mathbb{T}$. As for the second assertion, the proof here is similar to the proof of the crossed product decomposition $D_N = \mathbb{Z}_N \rtimes \mathbb{Z}_2$ for the dihedral groups.

(4) This is something more speculative, the idea here being that the isomorphism $O_2 = \mathbb{T} \rtimes \mathbb{Z}_2$ appears from $D_N = \mathbb{Z}_N \rtimes \mathbb{Z}_2$ by taking the $N \to \infty$ limit. $\square$

In general, the structure of O_N and SO_N, and the relation between them, is far more complicated than what happens at $N = 1, 2$. We will be back to this later.

10b. Euler-Rodrigues

For the moment, let us keep working out what happens at $N = 2$, this time with a study in the complex case. We first have here the following result:

THEOREM 10.6. *We have the following formula,*

$$SU_2 = \left\{ \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix} \ \Big| \ |a|^2 + |b|^2 = 1 \right\}$$

which makes SU_2 isomorphic to the unit sphere $S^1_{\mathbb{C}} \subset \mathbb{C}^2$.

PROOF. Consider an arbitrary 2×2 matrix, written as follows:

$$U = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

Assuming that we have $\det U = 1$, the inverse is then given by:

$$U^{-1} = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

On the other hand, assuming $U \in U_2$, the inverse must be the adjoint:

$$U^{-1} = \begin{pmatrix} \bar{a} & \bar{c} \\ \bar{b} & \bar{d} \end{pmatrix}$$

We are therefore led to the following equations, for the matrix entries:

$$d = \bar{a} \quad , \quad c = -\bar{b}$$

Thus our matrix must be of the following special form:

$$U=\begin{pmatrix}a&b\\-\bar{b}&\bar{a}\end{pmatrix}$$

Moreover, since the determinant is 1, we must have, as stated:

$$|a|^2+|b|^2=1$$

Thus, we are done with one direction. As for the converse, this is clear, the matrices in the statement being unitaries, and of determinant 1, and so being elements of SU_2. Finally, regarding the last assertion, recall that the unit sphere $S^1_\mathbb{C}\subset\mathbb{C}^2$ is given by:

$$S^1_\mathbb{C}=\left\{(a,b)\Big|\,|a|^2+|b|^2=1\right\}$$

Thus, we have an isomorphism of compact spaces, as follows:

$$SU_2\simeq S^1_\mathbb{C}\quad,\quad\begin{pmatrix}a&b\\-\bar{b}&\bar{a}\end{pmatrix}\to(a,b)$$

We have therefore proved our theorem. □

Regarding now the unitary group U_2, the result here is similar, as follows:

THEOREM 10.7. *We have the following formula,*

$$U_2=\left\{d\begin{pmatrix}a&b\\-\bar{b}&\bar{a}\end{pmatrix}\Big|\,|a|^2+|b|^2=1,|d|=1\right\}$$

which makes U_2 be a quotient compact space, as follows,

$$S^1_\mathbb{C}\times\mathbb{T}\to U_2$$

but with this parametrization being no longer bijective.

PROOF. In one sense, this is clear, because we have:

$$|d|=1\implies dSU_2\subset U_2$$

In the other sense, let $U\in U_2$. We have then:

$$\begin{aligned}|\det(U)|^2&=\det(U)\overline{\det(U)}\\&=\det(U)\det(U^*)\\&=\det(UU^*)\\&=\det(1)\\&=1\end{aligned}$$

Consider now the following complex number, defined up to a sign choice:

$$d=\sqrt{\det U}$$

We know from Proposition 10.2 that we have $|d| = 1$. Thus the rescaled matrix $V = U/d$ is unitary, $V \in U_2$. As for the determinant of this matrix, this is given by:

$$\begin{aligned} \det(V) &= \det(d^{-1}U) \\ &= d^{-2}\det(U) \\ &= \det(U)^{-1}\det(U) \\ &= 1 \end{aligned}$$

Thus we have $V \in SU_2$, and so we can write, with $|a|^2 + |b|^2 = 1$:

$$V = \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix}$$

Thus the matrix $U = dV$ appears as in the statement. Finally, observe that the result that we have just proved provides us with a quotient map as follows:

$$S^1_{\mathbb{C}} \times \mathbb{T} \to U_2 \quad , \quad ((a,b),d) \to d\begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix}$$

However, the parametrization is no longer bijective, because when we globally switch signs, the element $((-a,-b),-d)$ produces the same element of U_2. $\square$

Let us record now a few more results regarding SU_2, U_2, which are key groups in mathematics and physics. First, we have the following reformulation of Theorem 10.6:

THEOREM 10.8. *We have the formula*

$$SU_2 = \left\{ \begin{pmatrix} x+iy & z+it \\ -z+it & x-iy \end{pmatrix} \Big| \ x^2+y^2+z^2+t^2 = 1 \right\}$$

which makes SU_2 isomorphic to the unit real sphere $S^3_{\mathbb{R}} \subset \mathbb{R}^3$.

PROOF. We recall from Theorem 10.6 that we have:

$$SU_2 = \left\{ \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix} \Big| \ |a|^2 + |b|^2 = 1 \right\}$$

Now let us write our parameters $a, b \in \mathbb{C}$, which belong to the complex unit sphere $S^1_{\mathbb{C}} \subset \mathbb{C}^2$, in terms of their real and imaginary parts, as follows:

$$a = x + iy \quad , \quad b = z + it$$

In terms of $x, y, z, t \in \mathbb{R}$, our formula for a generic matrix $U \in SU_2$ becomes the one in the statement. As for the condition to be satisfied by the parameters $x, y, z, t \in \mathbb{R}$, this comes the condition $|a|^2 + |b|^2 = 1$ to be satisfied by $a, b \in \mathbb{C}$, which reads:

$$x^2 + y^2 + z^2 + t^2 = 1$$

Thus, we are led to the conclusion in the statement. Regarding now the last assertion, recall that the unit sphere $S^3_{\mathbb{R}} \subset \mathbb{R}^4$ is given by:

$$S^3_{\mathbb{R}} = \left\{(x,y,z,t)\ \Big|\ x^2+y^2+z^2+t^2=1\right\}$$

Thus, we have an isomorphism of compact spaces, as follows:

$$SU_2 \simeq S^3_{\mathbb{R}} \quad , \quad \begin{pmatrix}x+iy&z+it\\-z+it&x-iy\end{pmatrix} \to (x,y,z,t)$$

We have therefore proved our theorem. □

As a philosophical comment, the above parametrization of SU_2 is something very nice, because the parameters (x,y,z,t) range now over the sphere of space-time. Thus, we are probably doing some kind of physics here. More on this later.

Regarding now the group U_2, we have here a similar result, as follows:

THEOREM 10.9. *We have the following formula,*

$$U_2 = \left\{(p+iq)\begin{pmatrix}x+iy&z+it\\-z+it&x-iy\end{pmatrix}\ \Big|\ x^2+y^2+z^2+t^2=1,\ p^2+q^2=1\right\}$$

which makes U_2 be a quotient compact space, as follows,

$$S^3_{\mathbb{R}} \times S^1_{\mathbb{R}} \to U_2$$

but with this parametrization being no longer bijective.

PROOF. We recall from Theorem 10.7 that we have:

$$U_2 = \left\{d\begin{pmatrix}a&b\\-\bar{b}&\bar{a}\end{pmatrix}\ \Big|\ |a|^2+|b|^2=1,\ |d|=1\right\}$$

Now let us write our parameters $a,b \in \mathbb{C}$, which belong to the complex unit sphere $S^1_{\mathbb{C}} \subset \mathbb{C}^2$, and $d \in \mathbb{T}$, in terms of their real and imaginary parts, as follows:

$$a = x+iy \quad , \quad b = z+it \quad , \quad d = p+iq$$

In terms of these new parameters $x,y,z,t,p,q \in \mathbb{R}$, our formula for a generic matrix $U \in SU_2$, that we established before, reads:

$$U = (p+iq)\begin{pmatrix}x+iy&z+it\\-z+it&x-iy\end{pmatrix}$$

As for the condition to be satisfied by the parameters $x,y,z,t,p,q \in \mathbb{R}$, this comes the conditions $|a|^2+|b|^2=1$ and $|d|=1$ to be satisfied by $a,b,d \in \mathbb{C}$, which read:

$$x^2+y^2+z^2+t^2=1 \quad , \quad p^2+q^2=1$$

Thus, we are led to the conclusion in the statement. Regarding now the last assertion, recall that the unit spheres $S^3_\mathbb{R}\subset\mathbb{R}^4$ and $S^1_\mathbb{R}\subset\mathbb{R}^2$ are given by:

$$S^3_\mathbb{R}=\left\{(x,y,z,t)\Big|x^2+y^2+z^2+t^2=1\right\}$$

$$S^1_\mathbb{R}=\left\{(p,q)\Big|p^2+q^2=1\right\}$$

Thus, we have quotient map of compact spaces, as follows:

$$S^3_\mathbb{R}\times S^1_\mathbb{R}\to U_2$$

$$((x,y,z,t),(p,q))\to(p+iq)\begin{pmatrix}x+iy&z+it\\-z+it&x-iy\end{pmatrix}$$

However, the parametrization is no longer bijective, because when we globally switch signs, the element $((-x,-y,-z,-t),(-p,-q))$ produces the same element of U_2. $\square$

Here is now another reformulation of our main result so far, regarding SU_2, obtained by further building on the parametrization from Theorem 10.8:

THEOREM 10.10. *We have the following formula,*

$$SU_2=\left\{xc_1+yc_2+zc_3+tc_4\Big|x^2+y^2+z^2+t^2=1\right\}$$

where c_1,c_2,c_3,c_4 are the Pauli matrices, given by:

$$c_1=\begin{pmatrix}1&0\\0&1\end{pmatrix}\quad,\quad c_2=\begin{pmatrix}i&0\\0&-i\end{pmatrix}$$

$$c_3=\begin{pmatrix}0&1\\-1&0\end{pmatrix}\quad,\quad c_4=\begin{pmatrix}0&i\\i&0\end{pmatrix}$$

PROOF. We recall from Theorem 10.8 that the group SU_2 can be parametrized by the real sphere $S^3_\mathbb{R}\subset\mathbb{R}^4$, in the following way:

$$SU_2=\left\{\begin{pmatrix}x+iy&z+it\\-z+it&x-iy\end{pmatrix}\Big|x^2+y^2+z^2+t^2=1\right\}$$

Thus, the elements $U\in SU_2$ are precisely the matrices as follows, depending on parameters $x,y,z,t\in\mathbb{R}$ satisfying $x^2+y^2+z^2+t^2=1$:

$$U=x\begin{pmatrix}1&0\\0&1\end{pmatrix}+y\begin{pmatrix}i&0\\0&-i\end{pmatrix}+z\begin{pmatrix}0&1\\-1&0\end{pmatrix}+t\begin{pmatrix}0&i\\i&0\end{pmatrix}$$

But this gives the formula for SU_2 in the statement. $\square$

The above result is often the most convenient one, when dealing with SU_2. This is because the Pauli matrices have a number of remarkable properties, which are very useful when doing computations. These properties can be summarized as follows:

Proposition 10.11. *The Pauli matrices multiply according to the formulae*

$$c_2^2 = c_3^2 = c_4^2 = -1$$

$$c_2c_3 = -c_3c_2 = c_4$$

$$c_3c_4 = -c_4c_3 = c_2$$

$$c_4c_2 = -c_2c_4 = c_3$$

they conjugate according to the following rules,

$$c_1^* = c_1 \ , \ c_2^* = -c_2 \ , \ c_3^* = -c_3 \ , \ c_4^* = -c_4$$

and they form an orthonormal basis of $M_2(\mathbb{C})$, with respect to the scalar product

$$< a, b >= tr(ab^*)$$

with $tr : M_2(\mathbb{C}) \to \mathbb{C}$ being the normalized trace of 2×2 matrices, $tr = Tr/2$.

Proof. The first two assertions, regarding the multiplication and conjugation rules for the Pauli matrices, follow from some elementary computations. As for the last assertion, this follows by using these rules. Indeed, the fact that the Pauli matrices are pairwise orthogonal follows from computations of the following type, for $i \neq j$:

$$< c_i, c_j >= tr(c_ic_j^*) = tr(\pm c_ic_j) = tr(\pm c_k) = 0$$

As for the fact that the Pauli matrices have norm 1, this follows from:

$$< c_i, c_i >= tr(c_ic_i^*) = tr(\pm c_i^2) = tr(c_1) = 1$$

Thus, we are led to the conclusion in the statement. □

We should mention here that the Pauli matrices are cult objects in physics, due to the fact that they describe the spin of the electron. Remember maybe the discussion from the beginning of chapter 8, when we were talking about the wave functions $\psi : \mathbb{R}^3 \to \mathbb{C}$ of these electrons, and of the Hilbert space $H = L^2(\mathbb{R}^3)$ needed for understanding their quantum mechanics. Well, that was only half of the story, with the other half coming from the fact that, a bit like our Earth spins around its axis, the electrons spin too. And it took scientists a lot of skill in order to understand the physics and mathematics of the spin, the conclusion being that the wave function space $H = L^2(\mathbb{R}^3)$ has to be enlarged with a copy of $K = \mathbb{C}^2$, as to take into account the spin, and with this spin being described by the Pauli matrices, in some appropriate, quantum mechanical sense.

As usual, we refer to Feynman [39], Griffiths [42] or Weinberg [88] for more on all this. And with the remark that the Pauli matrices are actually subject to several possible normalizations, depending on formalism, but let us not get into all this here.

Regarding now the basic unitary groups in 3 or more dimensions, the situation here becomes fairly complicated. It is possible however to explicitely compute the rotation groups SO_3 and O_3, and explaining this result, due to Euler-Rodrigues, which is something non-trivial and very useful, for all sorts of practical purposes, will be our next goal.

The proof of the Euler-Rodrigues formula is something quite tricky. Let us start with the following construction, whose usefulness will become clear in a moment:

PROPOSITION 10.12. *The adjoint action $SU_2 \curvearrowright M_2(\mathbb{C})$, given by*

$$T_U(M) = UMU^*$$

leaves invariant the following real vector subspace of $M_2(\mathbb{C})$,

$$E = span_{\mathbb{R}}(c_1, c_2, c_3, c_4)$$

and we obtain in this way a group morphism $SU_2 \to GL_4(\mathbb{R})$.

PROOF. We have two assertions to be proved, as follows:

(1) We must first prove that, with $E \subset M_2(\mathbb{C})$ being the real vector space in the statement, we have the following implication:

$$U \in SU_2, M \in E \implies UMU^* \in E$$

But this is clear from the multiplication rules for the Pauli matrices, from Proposition 10.11. Indeed, let us write our matrices U, M as follows:

$$U = xc_1 + yc_2 + zc_3 + tc_4$$

$$M = ac_1 + bc_2 + cc_3 + dc_4$$

We know that the coefficients x, y, z, t and a, b, c, d are real, due to $U \in SU_2$ and $M \in E$. The point now is that when computing UMU^*, by using the various rules from Proposition 10.11, we obtain a matrix of the same type, namely a combination of c_1, c_2, c_3, c_4, with real coefficients. Thus, we have $UMU^* \in E$, as desired.

(2) In order to conclude, let us identify $E \simeq \mathbb{R}^4$, by using the basis c_1, c_2, c_3, c_4. The result found in (1) shows that we have a correspondence as follows:

$$SU_2 \to M_4(\mathbb{R}) \quad , \quad U \to (T_U)_{|E}$$

Now observe that for any $U \in SU_2$ and any $M \in M_2(\mathbb{C})$ we have:

$$T_{U^*}T_U(M) = U^*UMU^*U = M$$

Thus $T_{U^*} = T_U^{-1}$, and so the correspondence that we found can be written as:

$$SU_2 \to GL_4(\mathbb{R}) \quad , \quad U \to (T_U)_{|E}$$

But this a group morphism, due to the following computation:

$$T_UT_V(M) = UVMV^*U^* = T_{UV}(M)$$

Thus, we are led to the conclusion in the statement. □

The point now, which makes the link with SO_3, and which will ultimate elucidate the structure of SO_3, is that Proposition 10.12 can be improved as follows:

THEOREM 10.13. *The adjoint action $SU_2 \curvearrowright M_2(\mathbb{C})$, given by*

$$T_U(M) = UMU^*$$

leaves invariant the following real vector subspace of $M_2(\mathbb{C})$,

$$F = span_{\mathbb{R}}(c_2, c_3, c_4)$$

and we obtain in this way a group morphism $SU_2 \to SO_3$.

PROOF. We can do this in several steps, as follows:

(1) Our first claim is that the group morphism $SU_2 \to GL_4(\mathbb{R})$ constructed in Proposition 10.12 is in fact a morphism $SU_2 \to O_4$. In order to prove this, recall the following formula, valid for any $U \in SU_2$, from the proof of Proposition 10.12:

$$T_{U^*} = T_U^{-1}$$

We want to prove that the matrices $T_U \in GL_4(\mathbb{R})$ are orthogonal, and in view of the above formula, it is enough to prove that we have:

$$T_U^* = (T_U)^t$$

So, let us prove this. For any two matrices $M, N \in E$, we have:

$$\begin{aligned} < T_{U^*}(M), N > &= < U^*MU, N > \\ &= tr(U^*MUN) \\ &= tr(MUNU^*) \end{aligned}$$

On the other hand, we have as well the following formula:

$$\begin{aligned} < (T_U)^t(M), N > &= < M, T_U(N) > \\ &= < M, UNU^* > \\ &= tr(MUNU^*) \end{aligned}$$

Thus we have indeed $T_U^* = (T_U)^t$, which proves our $SU_2 \to O_4$ claim.

(2) In order now to finish, recall that we have by definition $c_1 = 1$, as a matrix. Thus, the action of SU_2 on the vector $c_1 \in E$ is given by:

$$T_U(c_1) = Uc_1U^* = UU^* = 1 = c_1$$

We conclude that $c_1 \in E$ is invariant under SU_2, and by orthogonality the following subspace of E must be invariant as well under the action of SU_2:

$$e_1^\perp = span_{\mathbb{R}}(c_2, c_3, c_4)$$

Now if we call this subspace F, and we identify $F \simeq \mathbb{R}^3$ by using the basis c_2, c_3, c_4, we obtain by restriction to F a morphism of groups as follows:

$$SU_2 \to O_3$$

But since this morphism is continuous and SU_2 is connected, its image must be connected too. Now since the target group decomposes as $O_3 = SO_3 \sqcup (-SO_3)$, and $1 \in SU_2$ gets mapped to $1 \in SO_3$, the whole image must lie inside SO_3, and we are done. $\square$

The above result is quite interesting, because we will see in a moment that the morphism $SU_2 \to SO_3$ constructed there is surjective. Thus, we will have a way of parametrizing the elements $V \in SO_3$ by elements $U \in SO_2$, and so ultimately by parameters $(x, y, z, t) \in S^3_{\mathbb{R}}$. In order to work out all this, let us start with the following result, coming as a continuation of Proposition 10.12, independently of Theorem 10.13:

PROPOSITION 10.14. *With respect to the standard basis c_1, c_2, c_3, c_4 of the vector space $\mathbb{R}^4 = span(c_1, c_2, c_3, c_4)$, the morphism $T : SU_2 \to GL_4(\mathbb{R})$ is given by:*

$$T_U = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & x^2+y^2-z^2-t^2 & 2(yz-xt) & 2(xz+yt) \\ 0 & 2(xt+yz) & x^2+z^2-y^2-t^2 & 2(zt-xy) \\ 0 & 2(yt-xz) & 2(xy+zt) & x^2+t^2-y^2-z^2 \end{pmatrix}$$

Thus, when looking at T as a group morphism $SU_2 \to O_4$, what we have in fact is a group morphism $SU_2 \to O_3$, and even $SU_2 \to SO_3$.

PROOF. With notations from Proposition 10.12 and its proof, let us first look at the action $L : SU_2 \curvearrowright \mathbb{R}^4$ by left multiplication, which is by definition given by:

$$L_U(M) = UM$$

In order to compute the matrix of this action, let us write, as usual:

$$U = xc_1 + yc_2 + zc_3 + tc_4$$

$$M = ac_1 + bc_2 + cc_3 + dc_4$$

By using the multiplication formulae in Proposition 10.11, we obtain:

$$\begin{aligned} UM &= (xc_1 + yc_2 + zc_3 + tc_4)(ac_1 + bc_2 + cc_3 + dc_4) \\ &= (xa - yb - zc - td)c_1 \\ &+ (xb + ya + zd - tc)c_2 \\ &+ (xc - yd + za + tb)c_3 \\ &+ (xd + yc - zb + ta)c_4 \end{aligned}$$

We conclude that the matrix of the left action considered above is:

$$L_U=\begin{pmatrix}x&-y&-z&-t\\y&x&-t&z\\z&t&x&-y\\t&-z&y&x\end{pmatrix}$$

Similarly, let us look now at the action $R:SU_2\curvearrowright\mathbb R^4$ by right multiplication, which is by definition given by the following formula:

$$R_U(M)=MU^*$$

In order to compute the matrix of this action, let us write, as before:

$$U=xc_1+yc_2+zc_3+tc_4$$

$$M=ac_1+bc_2+cc_3+dc_4$$

By using the multiplication formulae in Proposition 10.11, we obtain:

$$\begin{aligned}MU^*&=&&(ac_1+bc_2+cc_3+dc_4)(xc_1-yc_2-zc_3-tc_4)\\&=&&(ax+by+cz+dt)c_1\\&+&&(-ay+bx-ct+dz)c_2\\&+&&(-az+bt+cx-dy)c_3\\&+&&(-at-bz+cy+dx)c_4\end{aligned}$$

We conclude that the matrix of the right action considered above is:

$$R_U=\begin{pmatrix}x&y&z&t\\-y&x&-t&z\\-z&t&x&-y\\-t&-z&y&x\end{pmatrix}$$

Now by composing, the matrix of the adjoint matrix in the statement is:

$$\begin{aligned}T_U&=&&R_UL_U\\&=&&\begin{pmatrix}x&y&z&t\\-y&x&-t&z\\-z&t&x&-y\\-t&-z&y&x\end{pmatrix}\begin{pmatrix}x&-y&-z&-t\\y&x&-t&z\\z&t&x&-y\\t&-z&y&x\end{pmatrix}\\&=&&\begin{pmatrix}1&0&0&0\\0&x^2+y^2-z^2-t^2&2(yz-xt)&2(xz+yt)\\0&2(xt+yz)&x^2+z^2-y^2-t^2&2(zt-xy)\\0&2(yt-xz)&2(xy+zt)&x^2+t^2-y^2-z^2\end{pmatrix}\end{aligned}$$

Thus, we have indeed the formula in the statement. As for the remaining assertions, these are all clear either from this formula, or from Theorem 10.13. $\square$

We can now formulate the Euler-Rodrigues result, as follows:

THEOREM 10.15. *We have a double cover map, obtained via the adjoint representation,*

$$SU_2 \to SO_3$$

and this map produces the Euler-Rodrigues formula

$$U = \begin{pmatrix} x^2+y^2-z^2-t^2 & 2(yz-xt) & 2(xz+yt) \\ 2(xt+yz) & x^2+z^2-y^2-t^2 & 2(zt-xy) \\ 2(yt-xz) & 2(xy+zt) & x^2+t^2-y^2-z^2 \end{pmatrix}$$

for the generic elements of SO_3.

PROOF. We know from the above that we have a group morphism $SU_2 \to SO_3$, given by the formula in the statement, and the problem now is that of proving that this is a double cover map, in the sense that it is surjective, and with kernel $\{\pm 1\}$.

(1) Regarding the kernel, this is elementary to compute, as follows:

$$\begin{aligned} \ker(SU_2 \to SO_3) &= \left\{U \in SU_2 \Big| T_U(M) = M, \forall M \in E\right\} \\ &= \left\{U \in SU_2 \Big| UM = MU, \forall M \in E\right\} \\ &= \left\{U \in SU_2 \Big| Uc_i = c_iU, \forall i\right\} \\ &= \{\pm 1\} \end{aligned}$$

(2) Thus, we are done with this, and as a side remark here, this result shows that our morphism $SU_2 \to SO_3$ is ultimately a morphism as follows:

$$PU_2 \subset SO_3 \quad , \quad PU_2 = SU_2/\{\pm 1\}$$

Here P stands for "projective", and it is possible to say more about the construction $G \to PG$, which can be performed for any subgroup $G \subset U_N$. But we will not get here into this, our next goal being that of proving that we have $PU_2 = SO_3$.

(3) We must prove now that the morphism $SU_2 \to SO_3$ is surjective. This is something non-trivial, and there are several advanced proofs for this, as follows:

– A first proof is by using Lie theory. To be more precise, the tangent spaces at 1 of both SU_2 and SO_3 can be explicitely computed, by doing some linear algebra, and the morphism $SU_2 \to SO_3$ follows to be surjective around 1, and then globally.

– Another proof is via representation theory. Indeed, the representations of SU_2 and SO_3 are subject to very similar formulae, called Clebsch-Gordan rules, and this shows that $SU_2 \to SO_3$ is surjective. We will discuss this in chapter 14 below.

– Yet another advanced proof, which is actually quite bordeline for what can be called "proof", is by using the ADE/McKay classification of the subgroups $G \subset SO_3$, which shows that there is no room strictly inside SO_3 for something as big as PU_2.

(4) In short, with some good knowledge of group theory, we are done. However, this is not our case, and we will present in what follows a more pedestrian proof, which was actually the original proof, based on the fact that any rotation $U \in SO_3$ has an axis.

(5) As a first computation, let us prove that any rotation $U \in Im(SU_2 \to SO_3)$ has an axis. We must look for fixed points of such rotations, and by linearity it is enough to look for fixed points belonging to the sphere $S^2_{\mathbb{R}} \subset \mathbb{R}^3$. Now recall that in our picture for the quotient map $SU_2 \to SO_3$, the space $\mathbb{R}^3$ appears as $F = span_{\mathbb{R}}(c_2, c_3, c_4)$, naturally embedded into the space $\mathbb{R}^4$ appearing as $E = span_{\mathbb{R}}(c_1, c_2, c_3, c_4)$. Thus, we must look for fixed points belonging to the sphere $S^3_{\mathbb{R}} \subset \mathbb{R}^4$ whose first coordinate vanishes. But, in our $\mathbb{R}^4 = E$ picture, this sphere $S^3_{\mathbb{R}}$ is the group SU_2. Thus, we must look for fixed points $V \in SU_2$ whose first coordinate with respect to c_1, c_2, c_3, c_4 vanishes, which amounts in saying that the diagonal entries of V must be purely imaginary numbers.

(6) Long story short, via our various identifications, we are led into solving the equation $UV = VU$ with $U, V \in SU_2$, and with V having a purely imaginary diagonal. So, with standard notations for SU_2, we must solve the following equation, with $p \in i\mathbb{R}$:

$$\begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix} \begin{pmatrix} p & q \\ -\bar{q} & \bar{p} \end{pmatrix} = \begin{pmatrix} p & q \\ -\bar{q} & \bar{p} \end{pmatrix} \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix}$$

(7) But this is something which is routine. Indeed, by identifying coefficients we obtain the following equations, each appearing twice:

$$b\bar{q} = \bar{b}q \quad , \quad b(p - \bar{p}) = (a - \bar{a})q$$

In the case $b = 0$ the only equation which is left is $q = 0$, and reminding that we must have $p \in i\mathbb{R}$, we do have solutions, namely two of them, as follows:

$$V = \pm \begin{pmatrix} i & 0 \\ 0 & i \end{pmatrix}$$

(8) In the remaining case $b \neq 0$, the first equation reads $b\bar{q} \in \mathbb{R}$, so we must have $q = \lambda b$ with $\lambda \in \mathbb{R}$. Now with this substitution made, the second equation reads $p - \bar{p} = \lambda(a - \bar{a})$, and since we must have $p \in i\mathbb{R}$, this gives $2p = \lambda(a - \bar{a})$. Thus, our equations are:

$$q = \lambda b \quad , \quad p = \lambda \cdot \frac{a - \bar{a}}{2}$$

Getting back now to our problem about finding fixed points, assuming $|a|^2 + |b|^2 = 1$ we must find $\lambda \in \mathbb{R}$ such that the above numbers p, q satisfy $|p|^2 + |q|^2 = 1$. But:

$$\begin{aligned} |p|^2 + |q|^2 &= |\lambda b|^2 + \left| \lambda \cdot \frac{a - \bar{a}}{2} \right|^2 \\ &= \lambda^2(|b|^2 + Im(a)^2) \\ &= \lambda^2(1 - Re(a)^2) \end{aligned}$$

Thus, we have again two solutions to our fixed point problem, given by:

$$\lambda=\pm\frac{1}{\sqrt{1-Re(a)^2}}$$

(9) Summarizing, we have proved that any rotation $U\in Im(SU_2\to SO_3)$ has an axis, and with the direction of this axis, corresponding to a pair of opposite points on the sphere $S^2_\mathbb{R}\subset\mathbb{R}^3$, being given by the above formulae, via $S^2_\mathbb{R}\subset S^3_\mathbb{R}=SU_2$.

(10) In order to finish, we must argue that any rotation $U\in SO_3$ has an axis. But this follows for instance from some topology, by using the induced map $S^2_\mathbb{R}\to S^2_\mathbb{R}$. Now since $U\in SO_3$ is uniquely determined by its rotation axis, which can be regarded as a point of $S^2_\mathbb{R}/\{\pm1\}$, plus its rotation angle $t\in[0,2\pi)$, by using $S^2_\mathbb{R}\subset S^3_\mathbb{R}=SU_2$ as in (9) we are led to the conclusion that U is uniquely determined by an element of $SU_2/\{\pm1\}$, and so appears indeed via the Euler-Rodrigues formula, as desired. $\square$

So long for the Euler-Rodrigues formula. As already mentioned, all the above is just the tip of the iceberg, and there are many more things that can be said, which are all interesting, and worth learning. In what concerns us, we will be back to this in chapter 14 below, when doing representation theory, with an alternative proof for this.

Regarding now O_3, the extension from SO_3 is very simple, as follows:

THEOREM 10.16. *We have the Euler-Rodrigues formula*

$$U=\pm\begin{pmatrix}x^2+y^2-z^2-t^2&2(yz-xt)&2(xz+yt)\\2(xt+yz)&x^2+z^2-y^2-t^2&2(zt-xy)\\2(yt-xz)&2(xy+zt)&x^2+t^2-y^2-z^2\end{pmatrix}$$

for the generic elements of O_3.

PROOF. This follows from Theorem 10.15, because the determinant of an orthogonal matrix $U\in O_3$ must satisfy $\det U=\pm1$, and in the case $\det U=-1$, we have:

$$\det(-U)=(-1)^3\det U=-\det U=1$$

Thus, assuming $\det U=-1$, we can therefore rescale U into an element $-U\in SO_3$, and this leads to the conclusion in the statement. $\square$

With the above small N examples worked out, let us discuss now the general theory, at arbitrary values of $N\in\mathbb{N}$. In the real case, we have the following result:

PROPOSITION 10.17. *We have a decomposition as follows, with SO_N^{-1} consisting by definition of the orthogonal matrices having determinant -1:*

$$O_N=SO_N\cup SO_N^{-1}$$

Moreover, when N is odd the set SO_N^{-1} is simply given by $SO_N^{-1}=-SO_N$.

PROOF. The first assertion is clear from definitions, because the determinant of an orthogonal matrix must be ± 1. The second assertion is clear too, and we have seen this already at $N=3$, in the proof of Theorem 10.16. Finally, when N is even the situation is more complicated, and requires complex numbers. We will be back to this. $\square$

In the complex case now, the result is simpler, as follows:

PROPOSITION 10.18. *We have a decomposition as follows, with SU_N^d consisting by definition of the unitary matrices having determinant $d\in\mathbb T$:*

$$O_N=\bigcup_{d\in\mathbb T}SU_N^d$$

Moreover, the components are $SU_N^d=f\cdot SU_N$, where $f\in\mathbb T$ is such that $f^N=d$.

PROOF. This is clear from definitions, and from the fact that the determinant of a unitary matrix belongs to $\mathbb T$, by extracting a suitable square root of the determinant. $\square$

It is possible to use the decomposition in Proposition 10.18 in order to say more about what happens in the real case, in the context of Proposition 10.17, but we will not get into this. We will basically stop here with our study of O_N,U_N, and of their versions SO_N,SU_N. As a last result on the subject, however, let us record:

THEOREM 10.19. *We have subgroups of O_N,U_N constructed via the condition*

$$(\det U)^d=1$$

with $d\in\mathbb N\cup\{\infty\}$, which generalize both O_N,U_N and SO_N,SU_N.

PROOF. This is indeed from definitions, and from the multiplicativity property of the determinant. We will be back to these groups, which are quite specialized, later on. $\square$

10c. Symplectic groups

At a more specialized level now, we first have the groups B_N,C_N, consisting of the orthogonal and unitary bistochastic matrices. Let us start with:

DEFINITION 10.20. *A square matrix $M\in M_N(\mathbb C)$ is called bistochastic if each row and each column sum up to the same number:*

$$\begin{matrix}M_{11}&\ldots&M_{1N}&\to&\lambda\\ \vdots&&\vdots\\ M_{N1}&\ldots&M_{NN}&\to&\lambda\\ \downarrow&&\downarrow\\ \lambda&&\lambda\end{matrix}$$

If this happens only for the rows, or only for the columns, the matrix is called row-stochastic, respectively column-stochastic.

In what follows we will be interested in the unitary bistochastic matrices, which are quite interesting objects. As a first result, regarding such matrices, we have:

Proposition 10.21. *For a unitary matrix $U \in U_N$, the following are equivalent:*
(1) *H is bistochastic, with sums λ.*
(2) *H is row stochastic, with sums λ, and $|\lambda| = 1$.*
(3) *H is column stochastic, with sums λ, and $|\lambda| = 1$.*

Proof. This is something that we know from chapter 7, with (1) $\iff$ (2) being elementary, and with the further equivalence with (3) coming by symmetry. □

The unitary bistochastic matrices are stable under a number of operations, and in particular under taking products. Thus, these matrices form a group. We have:

Theorem 10.22. *The real and complex bistochastic groups, which are the sets*

$$B_N \subset O_N \quad , \quad C_N \subset U_N$$

consisting of matrices which are bistochastic, are isomorphic to O_{N-1}, U_{N-1}.

Proof. This is something that we know too from chapter 7. To be more precise, let us pick a matrix $F \in U_N$, such as the Fourier matrix F_N, satisfying the following condition, where $e_0, \ldots, e_{N-1}$ is the standard basis of $\mathbb{C}^N$, and where ξ is the all-one vector:

$$Fe_0 = \frac{1}{\sqrt{N}}\xi$$

We have then, by using the above property of F:

$$\begin{aligned} u\xi = \xi \quad &\iff \quad uFe_0 = Fe_0 \\ &\iff \quad F^*uFe_0 = e_0 \\ &\iff \quad F^*uF = diag(1, w) \end{aligned}$$

Thus we have isomorphisms as in the statement, given by $w_{ij} \to (F^*uF)_{ij}$. □

We will be back to B_N, C_N later. Moving ahead now, as yet another basic example of a continuous group, we have the symplectic group Sp_N. Let us begin with:

Definition 10.23. *The "super-space" $\bar{\mathbb{C}}^N$ is the usual space $\mathbb{C}^N$, with its standard basis $\{e_1, \ldots, e_N\}$, with a chosen sign $\varepsilon = \pm 1$, and a chosen involution on the indices:*

$$i \to \bar{i}$$

The "super-identity" matrix is $J_{ij} = \delta_{i\bar{j}}$ for $i \leq j$ and $J_{ij} = \varepsilon\delta_{i\bar{j}}$ for $i \geq j$.

Up to a permutation of the indices, we have a decomposition $N = 2p + q$, such that the involution is, in standard permutation notation:

$$(12) \ldots (2p-1, 2p)(2p+1) \ldots (q)$$

Thus, up to a base change, the super-identity is as follows, where $N = 2p + q$ and $\varepsilon = \pm 1$, with the 1_q block at right disappearing if $\varepsilon = -1$:

$$J = \begin{pmatrix} 0 & 1 &&&&&&& \\ \varepsilon 1 & 0_{(0)} &&&&&&& \\ && \ddots &&&&&& \\ &&& 0 & 1 &&&& \\ &&& \varepsilon 1 & 0_{(p)} &&&& \\ &&&&& 1_{(1)} && \\ &&&&&& \ddots & \\ &&&&&&& 1_{(q)} \end{pmatrix}$$

In the case $\varepsilon = 1$, the super-identity is the following matrix:

$$J_+(p,q) = \begin{pmatrix} 0 & 1 &&&&&&& \\ 1 & 0_{(1)} &&&&&&& \\ && \ddots &&&&&& \\ &&& 0 & 1 &&&& \\ &&& 1 & 0_{(p)} &&&& \\ &&&&& 1_{(1)} && \\ &&&&&& \ddots & \\ &&&&&&& 1_{(q)} \end{pmatrix}$$

In the case $\varepsilon = -1$ now, the diagonal terms vanish, and the super-identity is:

$$J_-(p,0) = \begin{pmatrix} 0 & 1 &&& \\ -1 & 0_{(1)} &&& \\ && \ddots && \\ &&& 0 & 1 \\ &&& -1 & 0_{(p)} \end{pmatrix}$$

With the above notions in hand, we have the following result:

THEOREM 10.24. *The super-orthogonal group, which is by definition*

$$\bar{O}_N = \left\{ U \in U_N \middle| U = J\bar{U}J^{-1} \right\}$$

with J being the super-identity matrix, is as follows:

(1) *At $\varepsilon = 1$ we have $\bar{O}_N = O_N$.*
(2) *At $\varepsilon = -1$ we have $\bar{O}_N = Sp_N$.*

PROOF. These results are both elementary, as follows:

(1) At $\varepsilon = -1$ this follows from definitions.

(2) At $\varepsilon = 1$ now, consider the root of unity $\rho = e^{\pi i/4}$, and let:

$$\Gamma = \frac{1}{\sqrt{2}}\begin{pmatrix}\rho&\rho^7\\ \rho^3&\rho^5\end{pmatrix}$$

Then this matrix Γ is unitary, and we have the following formula:

$$\Gamma\begin{pmatrix}0&1\\1&0\end{pmatrix}\Gamma^t = 1$$

Thus the following matrix is unitary as well, and satisfies $CJC^t = 1$:

$$C = \begin{pmatrix}\Gamma^{(1)}\\ &\ddots\\ &&\Gamma^{(p)}\\ &&&1_q\end{pmatrix}$$

Thus in terms of $V = CUC^*$ the relations $U = J\bar{U}J^{-1} =$ unitary simply read:

$$V = \bar{V} = \text{unitary}$$

Thus we obtain an isomorphism $\bar{O}_N = O_N$ as in the statement. $\square$

Regarding now Sp_N, we have the following result:

THEOREM 10.25. *The symplectic group $Sp_N \subset U_N$, which is by definition*

$$Sp_N = \left\{U \in U_N \Big| U = J\bar{U}J^{-1}\right\}$$

consists of the SU_2 patterned matrices,

$$U = \begin{pmatrix}a&b&\ldots\\ -\bar{b}&\bar{a}\\ \vdots&&\ddots\end{pmatrix}$$

which are unitary, $U \in U_N$. In particular, we have $Sp_2 = SU_2$.

PROOF. This follows indeed from definitions, because the condition $U = J\bar{U}J^{-1}$ corresponds precisely to the fact that U must be a SU_2-patterned matrix. $\square$

We will be back later to the symplectic groups, with more results about them.

10d. Reflection groups

We have now a quite good understanding of the main continuous groups of unitary matrices $G \subset U_N$, especially at small values of N. So, let us go back to the finite groups from the previous chapter, and make a link with the material there. We first have:

THEOREM 10.26. *The full complex reflection group $K_N \subset U_N$, given by*

$$K_N = M_N(\mathbb{T} \cup \{0\}) \cap U_N$$

has a wreath product decomposition as follows,

$$K_N = \mathbb{T} \wr S_N$$

with S_N acting on $\mathbb{T}^N$ in the standard way, by permuting the factors.

PROOF. This is something that we know from chapter 9, as the $s = \infty$ particular case of the results established for the complex reflection groups H_N^s. $\square$

By using the above full complex reflection group K_N, we can talk in fact about the reflection subgroup of any compact group $G \subset U_N$, as follows:

DEFINITION 10.27. *Given $G \subset U_N$, we define its reflection subgroup to be*

$$K = G \cap K_N$$

with the intersection taken inside U_N.

This notion is something quite interesting, leading us into the question of understanding what the subgroups of K_N are. We have here the following construction:

THEOREM 10.28. *We have subgroups of the basic complex reflection groups,*

$$H_N^{sd} \subset H_N^s$$

constructed via the following condition, with $d \in \mathbb{N} \cup \{\infty\}$,

$$(\det U)^d = 1$$

which generalize all the complex reflection groups that we have so far.

PROOF. Here the first assertion is clear from definitions, and from the multiplicativity of the determinant. As for the second assertion, this is rather a remark, coming from the fact that the alternating group A_N, which is the only finite group so far not fitting into the series $\{H_N^s\}$, is indeed of this type, obtained from $H_N^1 = S_N$ by using $d = 1$. $\square$

The point now is that, by a well-known and deep result in group theory, the complex reflection groups consist of the series $\{H_N^{sd}\}$ constructed above, and of a number of exceptional groups, which can be fully classified. To be more precise, we have:

THEOREM 10.29. *The irreducible complex reflection groups are*

$$H_N^{sd} = \left\{ U \in H_N^s \Big| (\det U)^d = 1 \right\}$$

along with 34 exceptional examples.

PROOF. This is something quite advanced, and we refer here to the paper of Shephard and Todd [80], and to the subsequent literature on the subject. $\square$

Let us discuss now a number of more specialized questions. Consider the following diagram, formed by the main rotation and reflection groups:

$$\begin{tikzcd}
K_N \arrow[r] & U_N \\
H_N \arrow[r] \arrow[u] & O_N \arrow[u]
\end{tikzcd}$$

We know from the above that this is an intersection and generation diagram. Now assume that we have an intermediate compact group, as follows:

$$H_N \subset G_N \subset U_N$$

The point is that we can think of our group G_N as living inside the above square, and so project it on the edges, as to obtain information about it. Let us start with:

DEFINITION 10.30. *Associated to any closed subgroup $G_N \subset U_N$ are its discrete and real versions, given by*

$$G_N^d = G_N \cap K_N \quad , \quad G_N^r = G_N \cap O_N$$

as well as its smooth and unitary versions, given by

$$G_N^s = < G_N, O_N > \quad , \quad G_N^u = < G_N, K_N >$$

where $<,>$ is the topological generation operation.

Assuming now that we have an intermediate compact group $H_N \subset G_N \subset U_N$, as above, we are led in this way to the following notion:

DEFINITION 10.31. *A compact group $H_N \subset G_N \subset U_N$ is called oriented if*

$$\begin{tikzcd}
K_N \arrow[r] & G_N^u \arrow[r] & U_N \\
G_N^d \arrow[r] \arrow[u] & G_N \arrow[r] \arrow[u] & G_N^s \arrow[u] \\
H_N \arrow[r] \arrow[u] & G_N^r \arrow[r] \arrow[u] & O_N \arrow[u]
\end{tikzcd}$$

is an intersection and generation diagram.

Most of our examples of compact groups $G_N \subset U_N$, finite or continuous, usually come in series, depending uniformly on $N \in \mathbb{N}$. This suggests the following definition:

DEFINITION 10.32. *A family of compact groups $G = (G_N)$, with $S_N \subset G_N \subset U_N$, is called uniform if it satisfies the following equivalent conditions:*

(1) *We have $G_{N-1} = G_N \cap U_{N-1}$, via the embedding $U_{N-1} \subset U_N$ given by:*

$$u \to \begin{pmatrix} u & 0 \\ 0 & 1 \end{pmatrix}$$

(2) *We have $G_{N-1} = G_N \cap U_{N-1}$, via the N possible diagonal embeddings*

$$U_{N-1} \subset U_N$$

obtained as above, by inserting 1 somewhere on the diagonal.

We have many examples of uniform oriented groups, and in principle some classification results are possible. Indeed, the uniform oriented groups should be the obvious examples of such groups, that we already know from the above. We will be back to this in chapters 13-16 below, under a number of supplementary assumptions on the groups that we consider, which will allow us to derive a number of classification results.

10e. Exercises

There has been a lot of theory in this chapter, and this is just the tip of the iceberg, on what can be said about the continuous groups. As a first exercise, we have:

EXERCISE 10.33. *Work out all the details of the Euler-Rodrigues formula, by using the fact that any rotation in $\mathbb{R}^3$ has a rotation axis.*

Here the problem, once the rotation axis found, is that of drawing the picture, identifying the relevant angles, and then doing the math in terms of these angles.

EXERCISE 10.34. *Work out the theory of the subgroups of O_N, U_N constructed via*

$$(\det U)^d = 1$$

with $d \in \mathbb{N} \cup \{\infty\}$, which generalize both O_N, U_N and SO_N, SU_N.

There are many things that can be done here, and the more, the better.

EXERCISE 10.35. *Look up the literature, and find the relevance of the symplectic groups, and of symplectic geometry in general, to questions in classical mechanics.*

As before with the previous exercise, many things that can be learned and done here, especially from classical mechanics books, and the more you learn, the better.

EXERCISE 10.36. *Find and then write down a brief account of the Shephard-Todd theorem, stating that the irreducible complex reflection groups are*

$$H_N^{sd} = \left\{ U \in H_N^s \Big| (\det U)^d = 1 \right\}$$

along with a number of exceptional examples, more precisely 34 of them.

As before with the previous exercises, the more here, the better.

CHAPTER 11

Symmetric groups

11a. Character laws

We develop in what follows some general theory for the compact subgroups $G \subset U_N$, usually taken finite, with our main example being the symmetric group $S_N \subset O_N$. Let us start with a definition that we have already met in chapter 9, namely:

DEFINITION 11.1. *A representation of a finite group G is a group morphism*

$$u : G \to U_N$$

into a unitary group. The character of such a representation is the function

$$\chi : G \to \mathbb{C} \quad , \quad g \to Tr(u_g)$$

where Tr is the usual, unnormalized trace of the $N \times N$ matrices.

As explained in chapter 9, the simplest case of all this, namely $N = 1$, is of particular interest. Here the representations coincide with their characters, and are by definition the group morphisms as follows, called characters of the group:

$$\chi : G \to \mathbb{T}$$

These characters from an abelian group $\widehat{G}$, and when G itself is abelian, the correspondence $G \to \widehat{G}$ is a duality, in the sense that it maps $\widehat{G} \to G$ as well. Moreover, a more detailed study shows that we have in fact an isomorphism $G \simeq \widehat{G}$, with this being something quite subtle, related at the same time to the structure theorem for the finite abelian groups, $G \simeq \mathbb{Z}_{N_1} \times \ldots \times \mathbb{Z}_{N_k}$, and to the Fourier transforms over such groups.

In what follows we will be interested in the general case, $N \in \mathbb{N}$. It is technically convenient to assume that the representation $\pi : G \to U_N$ is faithful, by replacing if necessary G with its image. Thus, we are led to the following definition:

DEFINITION 11.2. *The main character of a compact group $G \subset U_N$ is the map*

$$\chi : G \to \mathbb{C} \quad , \quad g \to Tr(g)$$

which associates to the group elements, viewed as unitary matrices, their trace.

We will see in a moment some motivations for the study of these characters. From a naive viewpoint, which is ours at the present stage, we want to do some linear algebra with our group elements $g \in U_N$, and we have several choices here, as follows:

(1) A first idea would be to look at the determinant, $\det g \in \mathbb{T}$. However, this is usually not a very interesting quantity, for instance because $g \in O_N$ implies $\det g = \pm 1$. Also, for groups like SO_N, SU_N, this determinant is by definition 1.

(2) A second idea would be to try to compute eigenvalues and eigenvectors for the group elements $g \in G$, and then solve diagonalization questions for these elements. However, all this is quite complicated, so this idea is not good either.

(3) Thus, we are left with looking at the trace, $Tr(g) \in \mathbb{C}$. We will see soon that this is a very reasonable choice, with the mathematics being at the same time non-trivial, doable, and also interesting, for a whole number of reasons.

Before starting our study, let us briefly discuss as well some more advanced reasons, leading to the study of characters. The idea here is that a given finite or compact group G can have several representations $\pi : G \to U_N$, and these representations can be studied via their characters $\chi_\pi : G \to \mathbb{C}$, with a well-known and deep theorem basically stating that π can be recovered from its character χ_π. We will be back to this later.

As a basic result regarding the characters, we have:

THEOREM 11.3. *Given a compact group $G \subset U_N$, its main character $\chi : G \to \mathbb{C}$ is a central function, in the sense that it satisfies the following condition:*

$$\chi(gh) = \chi(hg)$$

Equivalently, χ is constant on the conjugacy classes of G.

PROOF. This is clear from the fact that the trace of matrices satisfies:

$$Tr(AB) = Tr(BA)$$

Thus, we are led to the conclusion in the statement. $\square$

As before, there is some interesting mathematics behind all this. We will prove later, when doing representation theory, that any central function $f : G \to \mathbb{C}$ appears as a linear combination of characters $\chi_\pi : G \to \mathbb{C}$ of representations $\pi : G \to U_N$.

In order to work out now some examples, let us get back now to our main examples of finite groups, consstructed in chapter 9 above, namely:

$$\mathbb{Z}_N \subset D_N \subset S_N \subset H_N$$

We will do in what follows some character computations for these groups, which are all quite elementary, or at least not requiring very advanced theory.

Let us start with the following result, which covers $\mathbb{Z}_N \subset D_N \subset S_N$, or rather tells us what is to be done with these groups, in relation with their main characters:

PROPOSITION 11.4. *For the symmetric group, regarded as group of permutation matrices, $S_N \subset O_N$, the main character counts the number of fixed points:*

$$\chi(g) = \#\left\{i \in \{1, \ldots, N\}\Big|\sigma(i) = i\right\}$$

The same goes for any $G \subset S_N$, regarded as a matrix group via $G \subset S_N \subset O_N$.

PROOF. This is indeed clear from definitions, because the diagonal entries of the permutation matrices correspond to the fixed points of the permutation. □

Summarizing, we are left with counting fixed points. For the simplest possible group, namely the cyclic group $\mathbb{Z}_N \subset S_N$, the computation is as follows:

PROPOSITION 11.5. *The character of $\mathbb{Z}_N \subset O_N$ is given by:*

$$\chi(g) = \begin{cases} 0 & \text{if } g \neq 1 \\ N & \text{if } g = 1 \end{cases}$$

PROOF. This is clear from definitions, because the cyclic permutation matrices have 0 on the diagonal, and so 0 as trace, unless the matrix is the identity, having trace N. □

Let us record as well a probabilistic version of the above result. In probabilistic terms, the result states that the corresponding distribution is a Bernoulli law:

$$law(\chi) = \left(1 - \frac{1}{N}\right)\delta_0 + \frac{1}{N}\delta_N$$

For the dihedral group now, which is the next one in our hierarchy, the computation is more interesting, and the final answer is not uniform in N, as follows:

PROPOSITION 11.6. *For the dihedral group $D_N \subset S_N$ we have:*

$$law(\chi) = \begin{cases} \left(\frac{3}{4} - \frac{1}{2N}\right)\delta_0 + \frac{1}{4}\delta_2 + \frac{1}{2N}\delta_N & (N\ even) \\ \\ \left(\frac{1}{2} - \frac{1}{2N}\right)\delta_0 + \frac{1}{2}\delta_1 + \frac{1}{2N}\delta_N & (N\ odd) \end{cases}$$

PROOF. The dihedral group D_N consists indeed of:

(1) N symmetries, having each 1 fixed point when N is odd, and having 0 or 2 fixed points, distributed $50 - 50$, when N is even.

(2) N rotations, each having 0 fixed points, except for the identity, which is technically a rotation too, and which has N fixed points.

Thus, we are led to the formulae in the statement. □

Regarding now the symmetric group S_N itself, the permutations having no fixed points at all are called derangements, and the first question which appears, which is a classical question in combinatorics, is that of counting these derangements. And the result here, which is something remarkable, and very beautiful, is as follows:

THEOREM 11.7. *The probability for a random permutation $\sigma \in S_N$ to be a derangement is given by the following formula:*

$$P = 1 - \frac{1}{1!} + \frac{1}{2!} - \ldots + (-1)^{N-1}\frac{1}{(N-1)!} + (-1)^N \frac{1}{N!}$$

Thus we have the following asymptotic formula, in the $N \to \infty$ limit,

$$P \simeq \frac{1}{e}$$

where $e = 2.7182\ldots$ is the usual constant from analysis.

PROOF. This is something very classical, which is best viewed by using the inclusion-exclusion principle. Consider indeed the following sets:

$$S_N^i = \left\{\sigma \in S_N \Big| \sigma(i) = i\right\}$$

The set of permutations having no fixed points is then:

$$X_N = \left(\bigcup_i S_N^i\right)^c$$

In order to compute now the cardinality $|X_N|$, consider as well the following sets, depending on indices $i_1 < \ldots < i_k$, obtained by taking intersections:

$$S_N^{i_1\ldots i_k} = S_N^{i_1} \bigcap \ldots \bigcap S_N^{i_k}$$

Observe that we have the following formula:

$$S_N^{i_1\ldots i_k} = \left\{\sigma \in S_N \Big| \sigma(i_1) = i_1, \ldots, \sigma(i_k) = i_k\right\}$$

The inclusion-exclusion principle tells us that we have:

$$\begin{aligned}
&|X_N| \\
= \ & |S_N| - \sum_i |S_N^i| + \sum_{i<j} |S_N^i \cap S_N^j| - \ldots + (-1)^N \sum_{i_1<\ldots<i_N} |S_N^{i_1} \cup \ldots \cup S_N^{i_N}| \\
= \ & |S_N| - \sum_i |S_N^i| + \sum_{i<j} |S_N^{ij}| - \ldots + (-1)^N \sum_{i_1<\ldots<i_N} |S_N^{i_1\ldots i_N}|
\end{aligned}$$

Thus, the probability that we are interested in is given by:

$$\begin{aligned}
P &= \frac{1}{N!}\left(|S_N|-\sum_i|S_N^i|+\sum_{i<j}|S_N^{ij}|-\ldots+(-1)^N\sum_{i_1<\ldots<i_N}|S_N^{i_1\ldots i_N}|\right)\\
&= \frac{1}{N!}\sum_{k=0}^N(-1)^k\sum_{i_1<\ldots<i_k}|S_N^{i_1\ldots i_k}|\\
&= \frac{1}{N!}\sum_{k=0}^N(-1)^k\sum_{i_1<\ldots<i_k}(N-k)!\\
&= \frac{1}{N!}\sum_{k=0}^N(-1)^k\binom{N}{k}(N-k)!\\
&= \sum_{k=0}^N\frac{(-1)^k}{k!}\\
&= 1-\frac{1}{1!}+\frac{1}{2!}-\ldots+(-1)^{N-1}\frac{1}{(N-1)!}+(-1)^N\frac{1}{N!}
\end{aligned}$$

Since on the right we have the expansion of $\frac{1}{e}$, we obtain:

$$P\simeq\frac{1}{e}$$

Thus, we are led to the conclusion in the statement. $\square$

The above result is something remarkable, and there are many versions and generalizations of it. We will discuss this gradually, in what follows, all this being key material. To start with, in terms of characters, the above result reformulates as follows:

THEOREM 11.8. *For the symmetric group, the probability for main character*

$$\chi:S_N\to\mathbb{N}$$

to vanish is given by the following formula:

$$\mathbb{P}(\chi=0)=1-\frac{1}{1!}+\frac{1}{2!}-\ldots+(-1)^{N-1}\frac{1}{(N-1)!}+(-1)^N\frac{1}{N!}$$

Thus we have the following asymptotic formula, in the $N\to\infty$ limit

$$\mathbb{P}(\chi=0)\simeq\frac{1}{e}$$

where $e=2.7182\ldots$ is the usual constant from analysis.

PROOF. This follows indeed by combining Proposition 11.4, which tells us that χ counts the number of fixed points, and Theorem 11.7 above. $\square$

Let us discuss now, more generally, what happens when counting permutations having exactly k fixed points. The result here, extending Theorem 11.7, is as follows:

THEOREM 11.9. *The probability for a random permutation $\sigma \in S_N$ to have exactly k fixed points, with $k \in \mathbb{N}$, is given by the following formula:*

$$P = \frac{1}{k!}\left(1 - \frac{1}{1!} + \frac{1}{2!} - \ldots + (-1)^{N-1}\frac{1}{(N-1)!} + (-1)^N\frac{1}{N!}\right)$$

Thus we have the following asymptotic formula, in the $N \to \infty$ limit

$$P \simeq \frac{1}{ek!}$$

where $e = 2.7182\ldots$ is the usual constant from analysis.

PROOF. We already know, from Theorem 11.7, that this formula holds at $k = 0$. In the general case now, we have to count the permutations $\sigma \in S_N$ having exactly k points. Since having such a permutation amounts in choosing k points among $1, \ldots, N$, and then permuting the $N - k$ points left, without fixed points allowed, we have:

$$\begin{aligned}
\#\left\{\sigma \in S_N\Big|\chi(\sigma) = k\right\} &= \binom{N}{k}\#\left\{\sigma \in S_{N-k}\Big|\chi(\sigma) = 0\right\}\\
&= \frac{N!}{k!(N-k)!}\#\left\{\sigma \in S_{N-k}\Big|\chi(\sigma) = 0\right\}\\
&= N! \times \frac{1}{k!} \times \frac{\#\left\{\sigma \in S_{N-k}\Big|\chi(\sigma) = 0\right\}}{(N-k)!}
\end{aligned}$$

Now by dividing everything by $N!$, we obtain from this the following formula:

$$\frac{\#\left\{\sigma \in S_N\Big|\chi(\sigma) = k\right\}}{N!} = \frac{1}{k!} \times \frac{\#\left\{\sigma \in S_{N-k}\Big|\chi(\sigma) = 0\right\}}{(N-k)!}$$

By using now the computation at $k = 0$, that we already have, from Theorem 11.7 above, it follows that with $N \to \infty$ we have the following estimate:

$$\begin{aligned}
P(\chi = k) &\simeq \frac{1}{k!} \cdot P(\chi = 0)\\
&\simeq \frac{1}{k!} \cdot \frac{1}{e}
\end{aligned}$$

Thus, we are led to the conclusion in the statement. □

As before, in regards with derangements, we can reformulate what we found in terms of the main character, and we obtain in this way the following statement:

THEOREM 11.10. *For the symmetric group, the distribution of the main character*

$$\chi : S_N \to \mathbb{N}$$

is given by the following formula:

$$\mathbb{P}(\chi = k) = \frac{1}{k1}\left(1 - \frac{1}{1!} + \frac{1}{2!} - \ldots + (-1)^{N-1}\frac{1}{(N-1)!} + (-1)^N\frac{1}{N!}\right)$$

Thus we have the following asymptotic formula, in the $N \to \infty$ limit,

$$\mathbb{P}(\chi = k) \simeq \frac{1}{ek!}$$

where $e = 2.7182\ldots$ is the usual constant from analysis.

PROOF. This follows indeed by combining Proposition 11.4, which tells us that χ counts the number of fixed points, and Theorem 11.9 above. $\square$

11b. Poisson limits

In order to best interpret the above results, and get some advanced insight into the structure of S_N, we will need some probability theory, coming as the "discrete" counterpart of the theory developed in chapter 5, for the Gaussian laws. We first have:

DEFINITION 11.11. *The Poisson law of parameter 1 is the following measure,*

$$p_1 = \frac{1}{e}\sum_k \frac{\delta_k}{k!}$$

and the Poisson law of parameter $t > 0$ is the following measure,

$$p_t = e^{-t}\sum_k \frac{t^k}{k!}\,\delta_k$$

with the letter "p" standing for Poisson.

Observe that these laws have indeed mass 1, as they should, and this due to the following well-known formula, which is the foundational formula of calculus:

$$e^t = \sum_k \frac{t^k}{k!}$$

We will see in the moment why these measures appear a bit everywhere, in discrete contexts, the reasons behind this coming from the Poisson Limit Theorem (PLT). Let us first develop some general theory. We first have:

THEOREM 11.12. *We have the following formula, for any $s, t > 0$,*

$$p_s * p_t = p_{s+t}$$

so the Poisson laws form a convolution semigroup.

Proof. The convolution of Dirac masses is given by:

$$\delta_k * \delta_l = \delta_{k+l}$$

By using this formula and the binomial formula, we obtain:

$$\begin{aligned}
p_s * p_t &= e^{-s}\sum_k \frac{s^k}{k!}\,\delta_k * e^{-t}\sum_l \frac{t^l}{l!}\,\delta_l\\
&= e^{-s-t}\sum_{kl}\frac{s^kt^l}{k!l!}\delta_{k+l}\\
&= e^{-s-t}\sum_n \delta_n \sum_{k+l=n}\frac{s^kt^l}{k!l!}\\
&= e^{-s-t}\sum_n \frac{\delta_n}{n!}\sum_{k+l=n}\frac{n!}{k!l!}s^kt^l\\
&= e^{-s-t}\sum_n \frac{(s+t)^n}{n!}\,\delta_n\\
&= p_{s+t}
\end{aligned}$$

Thus, we are led to the conclusion in the statement. □

We have as well the following result:

Theorem 11.13. *The Poisson laws appear as exponentials*

$$p_t = \sum_k \frac{t^k(\delta_1-\delta_0)^{*k}}{k!}$$

with respect to the convolution of measures $*$.

Proof. By using the binomial formula, the measure at right is:

$$\begin{aligned}
\mu &= \sum_k \frac{t^k}{k!}\sum_{p+q=k}(-1)^q\frac{k!}{p!q!}\delta_p\\
&= \sum_k t^k \sum_{p+q=k}(-1)^q\frac{\delta_p}{p!q!}\\
&= \sum_p \frac{t^p\delta_p}{p!}\sum_q\frac{(-1)^q}{q!}\\
&= \frac{1}{e}\sum_p\frac{t^p\delta_p}{p!}\\
&= p_t
\end{aligned}$$

Thus, we are led to the conclusion in the statement. □

As in the continuous case, for the normal laws, our main tool for dealing with the Poisson laws will be the Fourier transform. The formula here is as follows:

THEOREM 11.14. *The Fourier transform of p_t is given by*

$$F_{p_t}(x)=\exp\left((e^{ix}-1)t\right)$$

for any $t>0$.

PROOF. The Fourier transform is given by definition by the following formula:

$$F_f(x)=\mathbb{E}(e^{ixf})$$

We therefore obtain the following formula:

$$\begin{aligned}
F_{p_t}(x) &= e^{-t}\sum_k\frac{t^k}{k!}F_{\delta_k}(x)\\
&= e^{-t}\sum_k\frac{t^k}{k!}\,e^{ikx}\\
&= e^{-t}\sum_k\frac{(e^{ix}t)^k}{k!}\\
&= \exp(-t)\exp(e^{ix}t)\\
&= \exp\left((e^{ix}-1)t\right)
\end{aligned}$$

Thus, we obtain the formula in the statement. □

Observe that we obtain in this way another proof for the convolution semigroup property of the Poisson laws, that we established above, by using the fact, that we know from chapter 5, that the logarithm of the Fourier transform linearizes the convolution.

We can now establish the Poisson Limit Theorem (PLT), as follows:

THEOREM 11.15. *We have the following convergence, in moments,*

$$\left(\left(1-\frac{t}{n}\right)\delta_0+\frac{t}{n}\delta_1\right)^{*n}\to p_t$$

for any $t>0$.

PROOF. Let us denote by μ_n the measure under the convolution sign:

$$\mu_n=\left(1-\frac{t}{n}\right)\delta_0+\frac{t}{n}\delta_1$$

We have then the following computation:

$$\begin{aligned}
F_{\delta_r}(x)=e^{irx} \quad & \Longrightarrow \quad F_{\mu_n}(x)=\left(1-\frac{t}{n}\right)+\frac{t}{n}e^{ix} \\
& \Longrightarrow \quad F_{\mu_n^{*n}}(x)=\left(\left(1-\frac{t}{n}\right)+\frac{t}{n}e^{ix}\right)^n \\
& \Longrightarrow \quad F_{\mu_n^{*n}}(x)=\left(1+\frac{(e^{ix}-1)t}{n}\right)^n \\
& \Longrightarrow \quad F(x)=\exp\left((e^{ix}-1)t\right)
\end{aligned}$$

Thus, we obtain the Fourier transform of p_t, as desired. □

There are of course many other things that can be said about the PLT, including examples and illustrations, and more technical results regarding the convergence, and we refer here to any standard probability book, such as Feller [35] or Durrett [33]. In what follows, we will be rather doing more combinatorics. To start with, we have:

THEOREM 11.16. *The moments of p_1 are the Bell numbers,*

$$M_k(p_1)=|P(k)|$$

where $P(k)$ is the set of partitions of $\{1,\ldots,k\}$.

PROOF. The moments of p_1 are given by the following formula:

$$M_k=\frac{1}{e}\sum_s\frac{s^k}{s!}$$

We have the following recurrence formula for these moments:

$$\begin{aligned}
M_{k+1} &= \frac{1}{e}\sum_s\frac{(s+1)^{k+1}}{(s+1)!} \\
&= \frac{1}{e}\sum_s\frac{(s+1)^k}{s!} \\
&= \frac{1}{e}\sum_s\frac{s^k}{s!}\left(1+\frac{1}{s}\right)^k \\
&= \frac{1}{e}\sum_s\frac{s^k}{s!}\sum_r\binom{k}{r}s^{-r} \\
&= \sum_r\binom{k}{r}\cdot\frac{1}{e}\sum_s\frac{s^{k-r}}{s!} \\
&= \sum_r\binom{k}{r}M_{k-r}
\end{aligned}$$

Let us try now to find a recurrence for the Bell numbers:

$$B_k = |P(k)|$$

A partition of $\{1,\ldots,k+1\}$ appears by choosing r neighbors for 1, among the k numbers available, and then partitioning the $k-r$ elements left. Thus, we have:

$$B_{k+1} = \sum_r \binom{k}{r} B_{k-r}$$

Thus, the numbers M_k satisfy the same recurrence as the numbers B_k. Regarding now the initial values, for the moments of p_1, these are:

$$M_0 = 1 \quad , \quad M_1 = 1$$

Indeed, the formula $M_0 = 1$ is clear, and the formula $M_1 = 1$ follows from:

$$\begin{aligned} M_1 &= \frac{1}{e}\sum_s \frac{s}{s!} \\ &= \frac{1}{e}\sum_s \frac{1}{(s-1)!} \\ &= \frac{1}{e} \times e \\ &= 1 \end{aligned}$$

Now by using the above recurrence we obtain from this:

$$\begin{aligned} M_2 &= \sum_r \binom{1}{r} M_{k-r} \\ &= 1+1 \\ &= 2 \end{aligned}$$

Thus, we can say that the initial values for the moments are:

$$M_1 = 1 \quad , \quad M_2 = 2$$

As for the Bell numbers, here the initial values are as follows:

$$B_1 = 1 \quad , \quad B_2 = 2$$

Thus the initial values coincide, and so these numbers are equal, as stated. □

More generally, we have the following result:

THEOREM 11.17. *The moments of p_t are given by*

$$M_k(p_t) = \sum_{\pi \in P(k)} t^{|\pi|}$$

where $|.|$ is the number of blocks.

PROOF. Observe first that the formula in the statement generalizes the one in Theorem 11.16 above, because at $t=1$ we obtain, as we should:

$$\begin{aligned} M_k(p_1) &= \sum_{\pi\in P(k)} 1^{|\pi|} \\ &= |P(k)| \\ &= B_k \end{aligned}$$

In general now, the moments of p_t with $t>0$ are given by:

$$M_k = e^{-t}\sum_s \frac{t^s s^k}{s!}$$

We have the following recurrence formula for the moments of p_t:

$$\begin{aligned} M_{k+1} &= e^{-t}\sum_s \frac{t^{s+1}(s+1)^{k+1}}{(s+1)!} \\ &= e^{-t}\sum_s \frac{t^{s+1}(s+1)^k}{s!} \\ &= e^{-t}\sum_s \frac{t^{s+1}s^k}{s!}\left(1+\frac{1}{s}\right)^k \\ &= e^{-t}\sum_s \frac{t^{s+1}s^k}{s!}\sum_r \binom{k}{r} s^{-r} \\ &= \sum_r \binom{k}{r}\cdot e^{-t}\sum_s \frac{t^{s+1}s^{k-r}}{s!} \\ &= t\sum_r \binom{k}{r} M_{k-r} \end{aligned}$$

As for the initial values of these moments, these are as follows:

$$M_1 = t \quad , \quad M_2 = t + t^2$$

On the other hand, consider the numbers in the statement, namely:

$$S_k = \sum_{\pi\in P(k)} t^{|\pi|}$$

Since a partition of $\{1,\ldots,k+1\}$ appears by choosing r neighbors for 1, among the k numbers available, and then partitioning the $k-r$ elements left, we have:

$$S_{k+1} = t\sum_r \binom{k}{r} S_{k-r}$$

As for the initial values of these numbers, these are:

$$S_1 = t \quad , \quad S_2 = t + t^2$$

Thus the initial values coincide, so these numbers are the moments, as stated. □

Observe the analogy with the moment formulae for g_t and G_t, discussed before. To be more precise, the moments for the main laws come from partitions, as follows:

$$p_t \to P \quad , \quad g_t \to P_2 \quad , \quad G_t \to \mathcal{P}_2$$

We will be back later with some more conceptual explanations for these results.

11c. Truncated characters

With the above probabilistic preliminaries done, let us get back now to finite groups, and compute laws of characters. As a first piece of good news, our main result so far, namely Theorem 11.10, reformulates into something very simple, as follows:

THEOREM 11.18. *For the symmetric group $S_N \subset O_N$ we have*

$$\chi \sim p_1$$

in the $N \to \infty$ limit.

PROOF. This is indeed a reformulation of Theorem 11.10 above, which tells us that with $N \to \infty$ we have the following estimate:

$$\mathbb{P}(\chi = k) \simeq \frac{1}{ek!}$$

But, according to our definition of the Poisson laws, this tells us precisely that the asymptotic law of the main character χ is Poisson (1), as stated. □

An interesting question now is that of recovering all the Poisson laws p_t, by using group theory. In order to do this, let us formulate the following definition:

DEFINITION 11.19. *Given a closed subgroup $G \subset U_N$, the function*

$$\chi : G \to \mathbb{C}$$

$$\chi_t(g) = \sum_{i=1}^{[tN]} g_{ii}$$

is called main truncated character of G, of parameter $t \in (0, 1]$.

As before with the plain characters, there is some general theory behind this definition, and we will discuss this later on, systematically, in chapters 13-16 below.

Getting back now to the symmetric groups, we first have the following result:

PROPOSITION 11.20. *Consider the symmetric group S_N, regarded as the permutation group of the N coordinate axes of $\mathbb{R}^N$. This picture provides us with an embedding*

$$S_N \subset O_N$$

the coordinate functions for the permutations being as folows:

$$g_{ij} = \chi\left(\sigma \in S_N \middle| \sigma(j) = i\right)$$

In this picture, the truncated characters count the number of partial fixed points

$$\chi_t(\sigma) = \#\left\{i \in \{1, \ldots, [tN]\} \middle| \sigma(i) = i\right\}$$

with respect to the truncation parameter $t \in (0, 1]$.

PROOF. All this is clear from definitions, with the formula for the coordinates being clear from the definition of the embedding $S_N \subset O_N$, and with the character formulae following from it, by summing over $i = j$. To be more precise, we have:

$$\begin{aligned}
\chi_t(\sigma) &= \sum_{i=1}^{[tN]} \sigma_{ii} \\
&= \sum_{i=1}^{[tN]} \delta_{\sigma(i)i} \\
&= \#\left\{i \in \{1, \ldots, [tN]\} \middle| \sigma(i) = i\right\}
\end{aligned}$$

Thus, we are led to the conclusions in the statement. □

Regarding now the asymptotic laws of the truncated characters, the result here, generalizing everything that we have so far, is as follows:

THEOREM 11.21. *For the symmetric group $S_N \subset O_N$ we have*

$$\chi_t \sim p_t$$

in the $N \to \infty$ limit, for any $t \in (0, 1]$.

PROOF. We already know from Theorem 11.18 that the result holds at $t = 1$. In general, the proof is similar, the idea being as follows:

(1) Consider indeed the following sets, as in the proof of Theorem 11.18, or rather as in the proof of Theorem 11.7, leaading to Theorem 11.18:

$$S_N^i = \left\{\sigma \in S_N \middle| \sigma(i) = i\right\}$$

The set of permutations having no fixed points among $1,\ldots,[tN]$ is then:

$$X_N=\left(\bigcup_{i\leq[tN]}S_N^i\right)^c$$

In order to compute now the cardinality $|X_N|$, consider as well the following sets, depending on indices $i_1<\ldots<i_k$, obtained by taking intersections:

$$S_N^{i_1\ldots i_k}=S_N^{i_1}\bigcap\ldots\bigcap S_N^{i_k}$$

As before in the proof of Theorem 11.18, we obtain by inclusion-exclusion that:

$$\begin{aligned}
P(\chi_t=0) &= \frac{1}{N!}\sum_{k=0}^{[tN]}(-1)^k\sum_{i_1<\ldots<i_k<[tN]}|S_N^{i_1\ldots i_k}|\\
&= \frac{1}{N!}\sum_{k=0}^{[tN]}(-1)^k\sum_{i_1<\ldots<i_k<[tN]}(N-k)!\\
&= \frac{1}{N!}\sum_{k=0}^{[tN]}(-1)^k\binom{[tN]}{k}(N-k)!\\
&= \sum_{k=0}^{[tN]}\frac{(-1)^k}{k!}\cdot\frac{[tN]!(N-k)!}{N!([tN]-k)!}
\end{aligned}$$

With $N\to\infty$, we obtain from this the following estimate:

$$\begin{aligned}
P(\chi_t=0) &\simeq \sum_{k=0}^{[tN]}\frac{(-1)^k}{k!}\cdot t^k\\
&= \sum_{k=0}^{[tN]}\frac{(-t)^k}{k!}\\
&\simeq e^{-t}
\end{aligned}$$

(2) More generally now, by counting the permutations $\sigma\in S_N$ having exactly k fixed points among $1,\ldots,[tN]$, as in the proof of Theorem 11.9, our claim is that we get:

$$P(\chi_t=k)\simeq\frac{t^k}{k!e^t}$$

We already know from (1) that this formula holds at $k=0$. In the general case now, we have to count the permutations $\sigma\in S_N$ having exactly k fixed points among $1,\ldots,[tN]$. Since having such a permutation amounts in choosing k points among $1,\ldots,[tN]$, and then permuting the $N-k$ points left, without fixed points among $1,\ldots,[tN]$ allowed, we obtain the following formula, where $s\in(0,1]$ is such that $[s(N-k)]=[tN]-k$:

$$\begin{aligned}
\#\left\{\sigma\in S_N\Big|\chi_t(\sigma)=k\right\} &= \binom{[tN]}{k}\#\left\{\sigma\in S_{N-k}\Big|\chi_s(\sigma)=0\right\}\\
&= \frac{[tN]!}{k!([tN]-k)!}\#\left\{\sigma\in S_{N-k}\Big|\chi_s(\sigma)=0\right\}\\
&= \frac{1}{k!}\times\frac{[tN]!(N-k)!}{([tN]-k)!}\times\frac{\#\left\{\sigma\in S_{N-k}\Big|\chi_s(\sigma)=0\right\}}{(N-k)!}
\end{aligned}$$

Now by dividing everything by $N!$, we obtain from this the following formula:

$$\frac{\#\left\{\sigma\in S_N\Big|\chi_t(\sigma)=k\right\}}{N!}=\frac{1}{k!}\times\frac{[tN]!(N-k)!}{N!([tN]-k)!}\times\frac{\#\left\{\sigma\in S_{N-k}\Big|\chi_s(\sigma)=0\right\}}{(N-k)!}$$

By using now the computation at $k=0$, that we already have, from (1) above, it follows that with $N\to\infty$ we have the following estimate:

$$\begin{aligned}
P(\chi_t=k) &\simeq \frac{1}{k!}\times\frac{[tN]!(N-k)!}{N!([tN]-k)!}\cdot P(\chi_s=0)\\
&\simeq \frac{t^k}{k!}\cdot P(\chi_s=0)\\
&\simeq \frac{t^k}{k!}\cdot\frac{1}{e^s}
\end{aligned}$$

Now recall that the parameter $s\in(0,1]$ was chosen in the above such that:

$$[s(N-k)]=[tN]-k$$

Thus in the $N\to\infty$ limit we have $s=t$, and so we obtain, as claimed:

$$P(\chi_t=k)\simeq\frac{t^k}{k!}\cdot\frac{1}{e^t}$$

It follows that we obtain in the limit a Poisson law of parameter t, as stated. □

11d. Further results

All the above is quite interesting, and is at the core of the theory that we want to develop, so let us further build on all this, with a number of more specialized results on the subject, which will be sometimes research-grade. We will be following [**10**].

To start with, let us first present a new, instructive proof for the above character results. The point indeed is that we can approach the problems as well directly, by integrating over S_N, and in order to do so, we can use the following result:

THEOREM 11.22. *Consider the symmetric group S_N, with its standard coordinates:*

$$g_{ij} = \chi\left(\sigma \in S_N \middle| \sigma(j) = i\right)$$

The products of these coordinates span the algebra $C(S_N)$, and the arbitrary integrals over S_N are given, modulo linearity, by the formula

$$\int_{S_N} g_{i_1j_1}\ldots g_{i_kj_k} = \begin{cases}\frac{(N-|\ker i|)!}{N!} & \text{if } \ker i = \ker j\\ 0 & \text{otherwise}\end{cases}$$

where $\ker i$ denotes as usual the partition of $\{1,\ldots,k\}$ whose blocks collect the equal indices of i, and where $|.|$ denotes the number of blocks.

PROOF. The first assertion follows from the Stone-Weierstrass theorem, because the standard coordinates g_{ij} separate the points of S_N, and so the algebra $< g_{ij} >$ that they generate must be equal to the whole function algebra $C(S_N)$:

$$< g_{ij} >= C(S_N)$$

Regarding now the second assertion, according to the definition of the matrix coordinates g_{ij}, the integrals in the statement are given by:

$$\int_{S_N} g_{i_1j_1}\ldots g_{i_kj_k} = \frac{1}{N!}\#\left\{\sigma \in S_N \middle| \sigma(j_1) = i_1,\ldots,\sigma(j_k) = i_k\right\}$$

Now observe that the existence of $\sigma \in S_N$ as above requires:

$$i_m = i_n \iff j_m = j_n$$

Thus, the above integral vanishes when:

$$\ker i \neq \ker j$$

Regarding now the case $\ker i = \ker j$, if we denote by $b \in \{1,\ldots,k\}$ the number of blocks of this partition $\ker i = \ker j$, we have $N-b$ points to be sent bijectively to $N-b$ points, and so $(N-b)!$ solutions, and the integral is $\frac{(N-b)!}{N!}$, as claimed. $\square$

As an illustration for the above formula, we can recover the computation of the asymptotic laws of the truncated characters χ_t. We have indeed:

THEOREM 11.23. *For the symmetric group $S_N \subset O_N$, regarded as a compact group of matrices, $S_N \subset O_N$, via the standard permutation matrices, the truncated character*

$$\chi_t(g) = \sum_{i=1}^{[tN]} g_{ii}$$

counts the number of fixed points among $\{1,\ldots,[tN]\}$, and its law with respect to the counting measure becomes, with $N \to \infty$, a Poisson law of parameter t.

PROOF. The first assertion comes from the following formula:

$$g_{ij}=\chi\left(\sigma\middle|\sigma(j)=i\right)$$

Regarding now the second assertion, we can use here the integration formula in Theorem 11.22 above. With S_{kb} being the Stirling numbers, counting the partitions of $\{1,\ldots,k\}$ having exactly b blocks, we have the following formula:

$$\begin{aligned}
\int_{S_N}\chi_t^k &= \sum_{i_1\ldots i_k=1}^{[tN]}\int_{S_N} g_{i_1i_1}\ldots g_{i_ki_k}\\
&= \sum_{\pi\in P(k)}\frac{[tN]!}{([tN]-|\pi|!)}\cdot\frac{(N-|\pi|!)}{N!}\\
&= \sum_{b=1}^{[tN]}\frac{[tN]!}{([tN]-b)!}\cdot\frac{(N-b)!}{N!}\cdot S_{kb}
\end{aligned}$$

In particular with $N\to\infty$ we obtain the following formula:

$$\lim_{N\to\infty}\int_{S_N}\chi_t^k=\sum_{b=1}^k S_{kb}t^b$$

But this is the k-th moment of the Poisson law p_t, and so we are done. $\square$

Summarizing, we have a good understanding of our main result so far, involving the characters of the symmetric group S_N and the Poisson laws of parameter $t\in(0,1]$, by using 2 different methods. We will see in a moment a third proof as well, and we will be actually back to this in chapters 13-16 too, with a fourth method as well.

As another result now regarding S_N, here is a useful related formula:

THEOREM 11.24. *We have the law formula*

$$\mathrm{law}(g_{11}+\ldots+g_{ss})=\frac{s!}{N!}\sum_{p=0}^s\frac{(N-p)!}{(s-p)!}\cdot\frac{(\delta_1-\delta_0)^{*p}}{p!}$$

where g_{ij} are the standard coordinates of $S_N\subset O_N$.

PROOF. We have the following moment formula, where m_f is the number of permutations of $\{1,\ldots,N\}$ having exactly f fixed points in the set $\{1,\ldots,s\}$:

$$\int_{S_N}(u_{11}+\ldots+u_{ss})^k=\frac{1}{N!}\sum_{f=0}^s m_f f^k$$

Thus the law in the statement, say ν_{sN}, is the following average of Dirac masses:

$$\nu_{sN}=\frac{1}{N!}\sum_{f=0}^s m_f\,\delta_f$$

Now observe that the permutations contributing to m_f are obtained by choosing f points in the set $\{1,\ldots,s\}$, then by permuting the remaining $N-f$ points in $\{1,\ldots,n\}$ in such a way that there is no fixed point in $\{1,\ldots,s\}$.

But these latter permutations are counted as follows: we start with all permutations, we substract those having one fixed point, we add those having two fixed points, and so on. We obtain in this way the following formula:

$$\begin{aligned}
\nu_{sN} &= \frac{1}{N!}\sum_{f=0}^s\binom{s}{f}\left(\sum_{k=0}^{s-f}(-1)^k\binom{s-f}{k}(N-f-k)!\right)\delta_f\\
&= \sum_{f=0}^s\sum_{k=0}^{s-f}(-1)^k\frac{1}{N!}\cdot\frac{s!}{f!(s-f)!}\cdot\frac{(s-f)!(N-f-k)!}{k!(s-f-k)!}\,\delta_f\\
&= \frac{s!}{N!}\sum_{f=0}^s\sum_{k=0}^{s-f}\frac{(-1)^k(N-f-k)!}{f!k!(s-f-k)!}\,\delta_f
\end{aligned}$$

We can proceed as follows, by using the new index $p=f+k$:

$$\begin{aligned}
\nu_{sN} &= \frac{s!}{N!}\sum_{p=0}^s\sum_{k=0}^p\frac{(-1)^k(N-p)!}{(p-k)!k!(s-p)!}\,\delta_{p-k}\\
&= \frac{s!}{N!}\sum_{p=0}^s\frac{(N-p)!}{(s-p)!p!}\sum_{k=0}^p(-1)^k\binom{p}{k}\,\delta_{p-k}\\
&= \frac{s!}{N!}\sum_{p=0}^s\frac{(N-p)!}{(s-p)!}\cdot\frac{(\delta_1-\delta_0)^{*p}}{p!}
\end{aligned}$$

Here $*$ is convolution of real measures, and the assertion follows. $\square$

Observe that the above formula is finer than most of our previous formulae, which were asymptotic, because it is valid at any $N\in\mathbb{N}$. We can use this formula as follows:

THEOREM 11.25. *Let g_{ij} be the standard coordinates of $C(S_N)$.*

(1) *$u_{11}+\ldots+u_{ss}$ with $s=o(N)$ is a projection of trace s/N.*
(2) *$u_{11}+\ldots+u_{ss}$ with $s=tN+o(N)$ is Poisson of parameter t.*

PROOF. We use the formula in Theorem 11.24 above.

(1) With s fixed and $N \to \infty$ we have the following estimate:

$$\begin{aligned}
&\mathrm{law}(u_{11} + \ldots + u_{ss})\\
&= \sum_{p=0}^{s} \frac{(N-p)!}{N!} \cdot \frac{s!}{(s-p)!} \cdot \frac{(\delta_1 - \delta_0)^{*p}}{p!}\\
&= \delta_0 + \frac{s}{N}(\delta_1 - \delta_0) + O(N^{-2})
\end{aligned}$$

But the law on the right is that of a projection of trace s/N, as desired.

(2) We have a law formula of the following type:

$$\mathrm{law}(u_{11} + \ldots + u_{ss}) = \sum_{p=0}^{s} c_p \cdot \frac{(\delta_1 - \delta_0)^{*p}}{p!}$$

The coefficients c_p can be estimated by using the Stirling formula, as follows:

$$\begin{aligned}
c_p &= \frac{(tN)!}{N!} \cdot \frac{(N-p)!}{(tN-p)!}\\
&\simeq \frac{(tN)^{tN}}{N^N} \cdot \frac{(N-p)^{N-p}}{(tN-p)^{tN-p}}\\
&= \left(\frac{tN}{tN-p}\right)^{tN-p} \left(\frac{N-p}{N}\right)^{N-p} \left(\frac{tN}{N}\right)^{p}
\end{aligned}$$

The last expression can be estimated by using the definition of exponentials, and we obtain the following estimate:

$$\begin{aligned}
c_p &\simeq e^p e^{-p} t^p\\
&= t^p
\end{aligned}$$

We can now compute the Fourier transform with respect to a variable y:

$$\begin{aligned}
\mathcal{F}\left(\mathrm{law}(u_{11} + \ldots + u_{ss})\right) &\simeq \sum_{p=0}^{s} t^p \cdot \frac{(e^y - 1)^p}{p!}\\
&= e^{t(e^y - 1)}
\end{aligned}$$

But this is the Fourier transform of the Poisson law p_t, as explained in Theorem 11.14 above, and this gives the second assertion. □

Let us discuss now, as an instructive variation of the above, the computation for the alternating group $A_N \subset S_N$. We will see that with $N \to \infty$ nothing changes, and with this being part of a more general phenomenon, regarding more general types of reflection groups and subgroups, that we will further discuss in the next chapter.

Let us start with some algebraic considerations. We first have:

Proposition 11.26. *For the symmetric group, regarded as group of permutations of the N coordinate axes of $\mathbb{R}^N$, and so as group of permutation matrices,*

$$S_N \subset O_N$$

the determinant is the signature. The subgroup $A_N \subset S_N$ given by

$$A_N = S_N \cap SO_N$$

and called alternating group, consists of the even permutations.

Proof. In this statement the first assertion is clear from the definition of the determinant, and of the permutation matrices, and all the rest is standard. □

Regarding now character computations, the best here is to use an analogue of Theorem 11.22 above. To be more precise, we have here the following result:

Theorem 11.27. *Consider the alternating group A_N, regarded as group of permutation matrices, with its standard coordinates:*

$$g_{ij} = \chi\left(\sigma \in A_N \Big| \sigma(j) = i\right)$$

The products of these coordinates span the algebra $C(A_N)$, and the arbitrary integrals over A_N are given, modulo linearity, by the formula

$$\int_{A_N} g_{i_1 j_1} \ldots g_{i_k j_k} \simeq \begin{cases} \frac{(N-|\ker i|)!}{N!} & \text{if } \ker i = \ker j \\ 0 & \text{otherwise} \end{cases}$$

with $N \to \infty$, where $\ker i$ denotes as usual the partition of $\{1, \ldots, k\}$ whose blocks collect the equal indices of i, and where $|.|$ denotes the number of blocks.

Proof. The first assertion follows from the Stone-Weierstrass theorem, because the standard coordinates g_{ij} separate the points of A_N, and so the algebra $< g_{ij} >$ that they generate must be equal to the whole function algebra $C(A_N)$:

$$< g_{ij} >= C(A_N)$$

Regarding now the second assertion, according to the definition of the standard coordinates g_{ij}, the integrals in the statement are given by:

$$\int_{A_N} g_{i_1 j_1} \ldots g_{i_k j_k} = \frac{1}{N!/2} \# \left\{\sigma \in A_N \Big| \sigma(j_1) = i_1, \ldots, \sigma(j_k) = i_k\right\}$$

Now observe that the existence of $\sigma \in A_N$ as above requires:

$$i_m = i_n \iff j_m = j_n$$

Thus, the above integral vanishes when:

$$\ker i \neq \ker j$$

Regarding now the case $\ker i = \ker j$, if we denote by $b \in \{1, \ldots, k\}$ the number of blocks of this partition $\ker i = \ker j$, we have $N - b$ points to be sent bijectively to $N - b$ points. But when assuming $N >> 0$, and more specifically $N > k$, half of these bijections will be alternating, and so we have $(N - b)!/2$ solutions. Thus, the integral is:

$$\begin{aligned}\int_{A_N} g_{i_1 j_1} \ldots g_{i_k j_k} &= \frac{1}{N!/2} \# \left\{ \sigma \in A_N \middle| \sigma(j_1) = i_1, \ldots, \sigma(j_k) = i_k \right\} \\ &= \frac{(N-b)!/2}{N!/2} \\ &= \frac{(N-b)!}{N!}\end{aligned}$$

Thus, we are led to the conclusion in the statement. □

As an illustration for the above formula, we can recover the computation of the asymptotic laws of the truncated characters χ_t. We have indeed:

THEOREM 11.28. *For the alternating group $A_N \subset O_N$, regarded as a compact group of matrices, $A_N \subset O_N$, via the standard permutation matrices, the truncated character*

$$\chi_t(g) = \sum_{i=1}^{[tN]} g_{ii}$$

counts the number of fixed points among $\{1, \ldots, [tN]\}$, and its law with respect to the counting measure becomes, with $N \to \infty$, a Poisson law of parameter t.

PROOF. The first assertion comes from the following formula:

$$g_{ij} = \chi \left(\sigma \middle| \sigma(j) = i \right)$$

Regarding now the second assertion, we can use here the integration formula in Theorem 11.27 above. With S_{kb} being the Stirling numbers, counting the partitions of $\{1, \ldots, k\}$ having exactly b blocks, we have the following formula:

$$\begin{aligned}\int_{A_N} \chi_t^k &= \sum_{i_1 \ldots i_k = 1}^{[tN]} \int_{A_N} g_{i_1 i_1} \ldots g_{i_k i_k} \\ &\simeq \sum_{\pi \in P(k)} \frac{[tN]!}{([tN] - |\pi|!)} \cdot \frac{(N - |\pi|!)}{N!} \\ &= \sum_{b=1}^{[tN]} \frac{[tN]!}{([tN] - b)!} \cdot \frac{(N-b)!}{N!} \cdot S_{kb}\end{aligned}$$

In particular with $N\to\infty$ we obtain the following formula:

$$\lim_{N\to\infty}\int_{S_N}\chi_t^k=\sum_{b=1}^kS_{kb}t^b$$

But this is the k-th moment of the Poisson law p_t, and so we are done. $\square$

As a conclusion to this, when passing from the symmetric group S_N to its subgroup $A_N\subset S_N$, in what concerns character computations, with $N\to\infty$ nothing changes. This is actually part of a more general phenomenon, regarding more general types of reflection groups and subgroups, that we will further discuss in the next chapter.

As a conclusion to all this, we have seen that the truncated characters χ_t of the symmetric group S_N have the Poisson laws p_t as limiting distributions, with $N\to\infty$. Moreover, we have seen several proofs for this fundamental fact, using inclusion-exclusion, direct integration, and convolution exponentials and Fourier transforms as well.

We will keep building on all this in the next chapter, by stating and proving similar results for more general reflection groups $G\subset U_N$. Also, we will be back to the symmetric group S_N and to the Poisson laws in chapters 13-16, with a fourth proof for our results, using representation theory, and a property of S_N called easiness. More on this later.

11e. Exercises

There are many interesting exercises in connection with the above. First, in relation with derangements and fixed points, we have:

Exercise 11.29. *Compute the number of derangements in S_4, by explicitely listing them, and then comment on the estimate of*

$$e=2.7182\ldots$$

that you obtain in this way.

Here the first question is of course elementary, but the problem is that of finding out what the best notation for permutations is, in order to solve this problem quickly. As for the second question, that you can investigate at higher N too, based on the various formulae established in this chapter, this is something quite instructive too.

Exercise 11.30. *Show that the probability for a length 1 needle to intersect, when thrown, a 1-spaced grid is $2/\pi$, and then comment on the estimate on*

$$\pi=3.1415\ldots$$

that you obtain in this way.

Here the first question is quite tricky, because there are several possible ways of modelling the problem, but only one of them gives the correct, real-life answer. As for the second question, this is a good introduction to applied mathematics too.

EXERCISE 11.31. *Interpret the abstract PLT formula established above, namely*

$$\left(\left(1-\frac{t}{n}\right)\delta_0+\frac{t}{n}\delta_1\right)^{*n}\to p_t$$

as a Poisson Limit Theorem, with full probabilistic details.

Here the problem is to understand why the above formula, that we already know, makes the Poisson law appears everywhere, in real life.

EXERCISE 11.32. *Find some formulae for the Bell numbers B_k, or rather for their generating series, or suitable transforms of that series, and the more the better.*

There is a lot of interesting mathematics here, and after solving the exercise, you can check the internet, and complete your knowledge with more things.

EXERCISE 11.33. *Show that the truncated characters of S_N, suitably moved over the diagonal, as to not overlap, become independent with $N\to\infty$.*

Here the formulation is of course a bit loose, but this is intentional, and finding the precise formulation is part of the exercise. As for the proof, this can only come by using the various integration formulae over S_N established in the above.

EXERCISE 11.34. *Find some alternative proofs for the fact, that we already know, that the truncated charcters for $A_N\subset O_N$ become Poisson, with $N\to\infty$.*

This is a bit technical, the problem being that of picking the best alternative proof for S_N, from the above, and then extending it to A_N. As a bonus exercise, you can work out as well independence aspects for A_N, in the spirit of the previous exercise.

CHAPTER 12

Reflection groups

12a. Real reflections

We have seen in the previous chapter that some interesting computations, in relation with the law of the main character, happen for the symmetric group S_N, in the $N\to\infty$ limit. Moreover, we have seen as well that the same kind of conclusions, with the law of the main character becoming Poisson (1), and the laws of the truncated characters becoming Poisson (t), in the $N\to\infty$ limit, happen for the alternating subgroup $A_N\subset S_N$.

All this suggests systematically looking at the general series of complex reflection groups H_N^{sd}. However, things are quite technical here, and we will do this slowly. First we will study the hyperoctahedral group H_N. Then we will develop some more general probabilistic theory, in relation with the Poisson laws and their versions. Then we will discuss the groups H_N^s, generalizing both S_N,H_N. And finally, we will comment on the groups H_N^{sd}, and we will do some new computations for the continuous groups as well.

Summarizing, a lot of things to be done. And you might perhaps ask, is this obsession with computing laws of characters justified? And our answer is that yes it is, the precise reasons behind our obsession, and motivations in general, being as follows:

(1) First, we will do this for the fun. Most of the material here will be research-grade, based on the papers [8], [9], [10], written in the late 00s. And isn't this nice, to read about what researchers are doing, not long after learning what a 2×2 matrix is.

(2) All these character computations will be an excellent introduction to the somewhat heavy representation theory methods to be developed in chapters 13-16 below, following Weyl [92], [93], [94], and then Brauer [16], Weingarten [91] and many others.

(3) And finally, again back to research, all this will be as well an introduction to all sorts of exciting things, such as random matrices and free probability following Wigner [95], Marchenko-Pastur [69] and Voiculescu [85], quantum groups, and many more.

In short, trust us, we will be doing some first-class mathematics in this chapter, of rather applied and computational type. And for abstractions and everything, don't worry, they will come back in chapters 13-16 below, inspired by what we will be doing here.

Getting started now, let us begin by discussing the hyperoctahedral group H_N. We first recall, from chapter 9 above, that we have:

THEOREM 12.1. *Consider the hyperoctahedral group H_N, which appears as the symmetry group of the N-cube, or the symmetry group of the N coordinate axes of $\mathbb{R}^N$:*

$$S_N \subset H_N \subset O_N$$

In matrix terms, H_N consists of the permutation-type matrices having ± 1 as nonzero entries, and we have a wreath product decomposition as follows:

$$H_N = \mathbb{Z}_2 \wr S_N$$

In this picture, the main character counts the signed number of fixed points, among the coordinate axes, and its truncations count the truncations of such numbers.

PROOF. This is something that was discussed before, the idea being that the first assertions are clear, and that the wreath product decomposition in the statement comes from a crossed product decomposition $H_N = \mathbb{Z}_2^N \rtimes S_N$. As for the assertions regarding the main character and its truncations, once again these are clear, as for S_N. $\square$

Regarding now the character laws, we can compute them by using the same method as for the symmetric group S_N, namely inclusion-exclusion, and we have:

THEOREM 12.2. *For the hyperoctahedral group $H_N \subset O_N$, the law of the variable*

$$\chi_t = \sum_{i=1}^{[tN]} g_{ii}$$

is in the $N \to \infty$ limit the measure

$$b_t = e^{-t} \sum_{k=-\infty}^{\infty} \delta_k \sum_{p=0}^{\infty} \frac{(t/2)^{|k|+2p}}{(|k|+p)!p!}$$

where δ_k is the Dirac mass at $k \in \mathbb{Z}$.

PROOF. We follow [9]. We regard H_N as being the symmetry group of the graph $I_N = \{I^1, \ldots, I^N\}$ formed by N segments. The diagonal coefficients are given by:

$$u_{ii}(g) = \begin{cases} 0 & \text{if } g \text{ moves } I^i \\ 1 & \text{if } g \text{ fixes } I^i \\ -1 & \text{if } g \text{ returns } I^i \end{cases}$$

We denote by $\uparrow g, \downarrow g$ the number of segments among $\{I^1, \ldots, I^s\}$ which are fixed, respectively returned by an element $g \in H_N$. With this notation, we have:

$$u_{11} + \ldots + u_{ss} = \uparrow g - \downarrow g$$

We denote by P_N probabilities computed over the group H_N. The density of the law of $u_{11}+\ldots+u_{ss}$ at a point $k\geq 0$ is given by the following formula:

$$\begin{aligned} D(k) &= P_N(\uparrow g - \downarrow g = k) \\ &= \sum_{p=0}^{\infty} P_N(\uparrow g = k+p, \downarrow g = p) \end{aligned}$$

Assume first that we have $t=1$. We use the fact that the probability of $\sigma\in S_N$ to have no fixed points is asymptotically:

$$P_0 = 1/e$$

Thus the probability of $\sigma\in S_N$ to have m fixed points is asymptotically:

$$P_m = 1/(em!)$$

In terms of probabilities over H_N, we get:

$$\begin{aligned} \lim_{N\to\infty} D(k) &= \lim_{N\to\infty}\sum_{p=0}^{\infty}(1/2)^{k+2p}\binom{k+2p}{k+p} P_N(\uparrow g + \downarrow g = k+2p) \\ &= \sum_{p=0}^{\infty}(1/2)^{k+2p}\binom{k+2p}{k+p}\frac{1}{e(k+2p)!} \\ &= \frac{1}{e}\sum_{p=0}^{\infty}\frac{(1/2)^{k+2p}}{(k+p)!p!} \end{aligned}$$

The general case $0<t\leq 1$ follows by performing some modifications in the above computation. The asymptotic density is computed as follows:

$$\begin{aligned} \lim_{N\to\infty} D(k) &= \lim_{N\to\infty}\sum_{p=0}^{\infty}(1/2)^{k+2p}\binom{k+2p}{k+p} P_N(\uparrow g + \downarrow g = k+2p) \\ &= \sum_{p=0}^{\infty}(1/2)^{k+2p}\binom{k+2p}{k+p}\frac{t^{k+2p}}{e^t(k+2p)!} \\ &= e^{-t}\sum_{p=0}^{\infty}\frac{(t/2)^{k+2p}}{(k+p)!p!} \end{aligned}$$

Together with $D(-k)=D(k)$, this gives the formula in the statement. $\square$

12b. Bessel laws

The above result is quite interesting, because the densities there are the Bessel functions of the first kind. Due to this fact, the limiting measures are called Bessel laws:

DEFINITION 12.3. *The Bessel law of parameter $t > 0$ is the measure*

$$b_t = e^{-t} \sum_{k=-\infty}^{\infty} \delta_k \, f_k(t/2)$$

with the density being the Bessel function of the first kind:

$$f_k(t) = \sum_{p=0}^{\infty} \frac{t^{|k|+2p}}{(|k|+p)!p!}$$

Let us study now these Bessel laws. We first have the following result:

THEOREM 12.4. *The Bessel laws b_t have the property*

$$b_s * b_t = b_{s+t}$$

so they form a truncated one-parameter semigroup with respect to convolution.

PROOF. We follow [9]. We use the formula in Definition 12.3, namely:

$$b_t = e^{-t} \sum_{k=-\infty}^{\infty} \delta_k \, f_k(t/2)$$

The Fourier transform of this measure is given by:

$$Fb_t(y) = e^{-t} \sum_{k=-\infty}^{\infty} e^{ky} \, f_k(t/2)$$

We compute now the derivative with respect to t:

$$Fb_t(y)' = -Fb_t(y) + \frac{e^{-t}}{2} \sum_{k=-\infty}^{\infty} e^{ky} \, f_k'(t/2)$$

On the other hand, the derivative of f_k with $k \geq 1$ is given by:

$$\begin{aligned}
f_k'(t) &= \sum_{p=0}^{\infty} \frac{(k+2p)t^{k+2p-1}}{(k+p)!p!} \\
&= \sum_{p=0}^{\infty} \frac{(k+p)t^{k+2p-1}}{(k+p)!p!} + \sum_{p=0}^{\infty} \frac{p \, t^{k+2p-1}}{(k+p)!p!} \\
&= \sum_{p=0}^{\infty} \frac{t^{k+2p-1}}{(k+p-1)!p!} + \sum_{p=1}^{\infty} \frac{t^{k+2p-1}}{(k+p)!(p-1)!} \\
&= \sum_{p=0}^{\infty} \frac{t^{(k-1)+2p}}{((k-1)+p)!p!} + \sum_{p=1}^{\infty} \frac{t^{(k+1)+2(p-1)}}{((k+1)+(p-1))!(p-1)!} \\
&= f_{k-1}(t) + f_{k+1}(t)
\end{aligned}$$

This computation works in fact for any k, so we get:

$$\begin{aligned} Fb_t(y)' &= -Fb_t(y)+\frac{e^{-t}}{2}\sum_{k=-\infty}^\infty e^{ky}(f_{k-1}(t/2)+f_{k+1}(t/2))\\ &= -Fb_t(y)+\frac{e^{-t}}{2}\sum_{k=-\infty}^\infty e^{(k+1)y}f_k(t/2)+e^{(k-1)y}f_k(t/2)\\ &= -Fb_t(y)+\frac{e^y+e^{-y}}{2}\,Fb_t(y)\\ &= \left(\frac{e^y+e^{-y}}{2}-1\right)Fb_t(y) \end{aligned}$$

Thus the log of the Fourier transform is linear in t, and we get the assertion. □

12c. Complex reflections

In order to further discuss all this, we will need a number of probabilistic preliminaries. So, let us pause now from our reflection group study, and do some theoretical probability theory, following [8], as a continuation of the material from chapter 11 above.

We recall that, conceptually speaking, the Poisson laws are the laws appearing via the Poisson Limit Theorem (PLT). In order to generalize this construction, as to cover for instance for Bessel laws that we found in connection with the hyperoctahedral group H_N, we have the following notion, extending the Poisson limit theory from chapter 11:

Definition 12.5. *Associated to any compactly supported positive measure ν on $\mathbb R$ is the probability measure*

$$p_\nu=\lim_{n\to\infty}\left(\left(1-\frac{c}{n}\right)\delta_0+\frac{1}{n}\nu\right)^{*n}$$

where $c=mass(\nu)$, called compound Poisson law.

In other words, what we are doing here is to generalize the construction in the Poisson Limit Theorem, by allowing the only parameter there, which was the positive real number $t>0$, to be replaced by a certain probability measure ν, or arbitrary mass $c>0$.

In what follows we will be interested in the case where ν is discrete, as is for instance the case for the measure $\nu=t\delta_1$ with $t>0$, which produces via the above procedure the Poisson laws. To be more precise, we will be mainly interested in the case where ν is a multiple of the uniform measure on the s-th roots of unity. More on this later.

The following result allows one to detect compound Poisson laws:

Proposition 12.6. *For a discrete measure, written as*

$$\nu=\sum_{i=1}^s c_i\delta_{z_i}$$

with $c_i>0$ and $z_i\in\mathbb R$, we have

$$F_{p_\nu}(y)=\exp\left(\sum_{i=1}^s c_i(e^{iyz_i}-1)\right)$$

where F denotes the Fourier transform.

Proof. Let μ_n be the measure appearing in Definition 12.5, under the convolution signs, namely:

$$\mu_n=\left(1-\frac{c}{n}\right)\delta_0+\frac{1}{n}\nu$$

We have the following computation:

$$\begin{aligned}
F_{\mu_n}(y)&=\left(1-\frac{c}{n}\right)+\frac{1}{n}\sum_{i=1}^s c_ie^{iyz_i}\\
\implies\quad F_{\mu_n^{*n}}(y)&=\left(\left(1-\frac{c}{n}\right)+\frac{1}{n}\sum_{i=1}^s c_ie^{iyz_i}\right)^n\\
\implies\quad F_{p_\nu}(y)&=\exp\left(\sum_{i=1}^s c_i(e^{iyz_i}-1)\right)
\end{aligned}$$

Thus, we have obtained the formula in the statement. □

We have as well the following result, providing an alternative to Definition 12.5:

Theorem 12.7. *For a discrete measure, written as*

$$\nu=\sum_{i=1}^s c_i\delta_{z_i}$$

with $c_i>0$ and $z_i\in\mathbb R$, we have

$$p_\nu=\mathrm{law}\left(\sum_{i=1}^s z_i\alpha_i\right)$$

where the variables α_i are Poisson (c_i), independent.

Proof. Let α be the sum of Poisson variables in the statement:

$$\alpha=\sum_{i=1}^s z_i\alpha_i$$

By using some well-known Fourier transform formulae, we have:

$$\begin{aligned} F_{\alpha_i}(y) = \exp(c_i(e^{iy} - 1) \quad &\implies \quad F_{z_i\alpha_i}(y) = \exp(c_i(e^{iyz_i} - 1)) \\ &\implies \quad F_\alpha(y) = \exp\left(\sum_{i=1}^s c_i(e^{iyz_i} - 1)\right) \end{aligned}$$

Thus we have indeed the same formula as in Proposition 12.6. ☐

Getting back now to the Bessel laws, we have:

THEOREM 12.8. *The Bessel laws b_t are compound Poisson laws, given by*

$$b_t = p_{t\varepsilon}$$

where $\varepsilon = \frac{1}{2}(\delta_{-1} + \delta_1)$ is the uniform measure on $\mathbb{Z}_2$.

PROOF. This follows indeed by comparing the formula of the Fourier transform of b_t, from the proof of Theorem 12.2 above, with the formula in Proposition 12.7. ☐

Our next task will be that of generalizing the results that we have for S_N, H_N. For this purpose, let us consider the following family of groups:

DEFINITION 12.9. *The complex reflection group $H_N^s \subset U_N$, depending on parameters*

$$N \in \mathbb{N} \quad , \quad s \in \mathbb{N} \cup \{\infty\}$$

are the groups of permutation-type matrices with s-th roots of unity as entries,

$$H_N^s = M_N(\mathbb{Z}_s \cup \{0\}) \cap U_N$$

with the convention $\mathbb{Z}_\infty = \mathbb{T}$, at $s = \infty$.

Observe that at $s = 1, 2$ we obtain the symmetric and hyperoctahedral groups:

$$H_N^1 = S_N \quad , \quad H_N^2 = H_N$$

Another important particular case is $s = \infty$, where we obtain a group which is actually not finite, denoted as follows:

$$H_N^\infty = K_N$$

In order to do now the character computations for H_N^s, in general, we need a number of further probabilistic preliminaries. Let us start with the following definition:

DEFINITION 12.10. *The Bessel law of level $s \in \mathbb{N} \cup \{\infty\}$ and parameter $t > 0$ is*

$$b_t^s = p_{t\varepsilon_s}$$

with ε_s being the uniform measure on the s-th roots of unity.

Observe that at $s = 1, 2$ we obtain the Poisson and real Bessel laws:

$$b_t^1 = p_t \quad , \quad b_t^2 = b_t$$

Another important particular case is $s = \infty$, where we obtain a measure which is actually not discrete, denoted as follows:

$$b_t^\infty = B_t$$

As a basic result on these laws, generalizing those about p_t, b_t, we have:

THEOREM 12.11. *The generalized Bessel laws b_t^s have the property*

$$b_t^s * b_{t'}^s = b_{t+t'}^s$$

so they form a truncated one-parameter semigroup with respect to convolution.

PROOF. This follows indeed from the Fourier transform formulae from Proposition 12.6, because the log of these Fourier transforms are linear in t. □

Regarding now the moments, the result here is as follows:

THEOREM 12.12. *The moments of the Bessel law b_t^s are the numbers*

$$M_k = |P^s(k)|$$

where $P^s(k)$ is the set of partitions of $\{1, \ldots, k\}$ satisfying

$$\# \circ = \# \bullet \, (s)$$

as a weighted sum, in each block.

PROOF. We already know that the formula in the statement holds indeed at $s = 1$, where $b_t^1 = p_t$ is the Poisson law of parameter $t > 0$, and where $P^1 = P$ is the set of all partitions. At $s = 2$ we have $P^2 = P_{even}$, and the result is elementary as well, and already proved in the above. In general, this follows by doing some combinatorics. See [8]. □

We can go back now to the reflection groups, and we have:

THEOREM 12.13. *For the complex reflection group*

$$H_N^s = \mathbb{Z}_s \wr S_N$$

we have, with $N \to \infty$, the estimate

$$\chi_t \sim b_t^s$$

where $b_t^s = p_{t\varepsilon_s}$, with ε_s being the uniform measure on the s-th roots of unity.

PROOF. We denote by ρ the uniform measure on the s-roots of unity. The best is to proceed in two steps, as follows:

(1) We work out first the case $t = 1$. Since the limit probability for a random permutation to have exactly k fixed points is $e^{-1}/k!$, we get:

$$\lim_{N\to\infty} \mathrm{law}(\chi_1) = e^{-1}\sum_{k=0}^{\infty}\frac{1}{k!}\,\rho^{*k}$$

On the other hand, we get from the definition of the Bessel law b_1^s:

$$\begin{aligned}
b_1^s &= \lim_{N\to\infty}\left(\left(1-\frac{1}{N}\right)\delta_0+\frac{1}{N}\,\rho\right)^{*N}\\
&= \lim_{N\to\infty}\sum_{k=0}^{N}\binom{N}{k}\left(1-\frac{1}{N}\right)^{N-k}\frac{1}{N^k}\,\rho^{*k}\\
&= e^{-1}\sum_{k=0}^{\infty}\frac{1}{k!}\,\rho^{*k}
\end{aligned}$$

But this gives the assertion for $t = 1$.

(2) Now in the case $t > 0$ arbitrary, we can use the same method, by performing the following modifications to the above computation:

$$\begin{aligned}
\lim_{N\to\infty}\mathrm{law}(\chi_t) &= e^{-t}\sum_{k=0}^{\infty}\frac{t^k}{k!}\,\rho^{*k}\\
&= \lim_{N\to\infty}\left(\left(1-\frac{1}{N}\right)\delta_0+\frac{1}{N}\,\rho\right)^{*[tN]}\\
&= b_t^s
\end{aligned}$$

Thus, we are led to the conclusion in the statement. □

Here is now some more theory for the Bessel laws, following [8]. First, it is convenient to introduce as well "modified" versions of the Bessel laws, as follows:

Definition 12.14. *The Bessel and modified Bessel laws are given by*

$$b_t^s = \mathrm{law}\left(\sum_{k=1}^{s} w^k a_k\right)$$

$$\tilde{b}_t^s = \mathrm{law}\left(\sum_{k=1}^{s} w^k a_k\right)^s$$

where $a_1,\ldots,a_s$ are independent random variables, each of them following the Poisson law of parameter t/s, and $w = e^{2\pi i/s}$.

As a first remark, at $s=1$ we get the Poisson law of parameter t:

$$b_t^1=\tilde{b}_t^1=e^{-t}\sum_{r=0}^\infty\frac{t^r}{r!}\,\delta_r$$

Consider now the level s exponential function, given by:

$$\exp_s z=\sum_{k=0}^\infty\frac{z^{sk}}{(sk)!}$$

We have the following formula, in terms of $w=e^{2\pi i/s}$:

$$\exp_s z=\frac{1}{s}\sum_{k=1}^s\exp(w^kz)$$

Observe that we have the following formulae:

$$\exp_1=\exp\quad,\quad\exp_2=\cosh$$

We have the following result, regarding both the plain and modified Bessel laws, which is a more explicit version of Proposition 12.6 above, for the Bessel laws:

THEOREM 12.15. *The Fourier transform of b_t^s is given by*

$$\log F_t^s(z)=t\,(\exp_s z-1)$$

so in particular the measures b_t^s are additive with respect to t.

PROOF. Consider, as in Definition 12.14, the following variable:

$$a=\sum_{k=1}^s w^ka_k$$

We have then the following computation:

$$\begin{aligned}\log F_a(z)&=\sum_{k=1}^s\log F_{a_k}(w^kz)\\&=\sum_{k=1}^s\frac{t}{s}\big(\exp(w^kz)-1\big)\end{aligned}$$

This gives the following formula:

$$\begin{aligned}\log F_a(z)&=t\left(\left(\frac{1}{s}\sum_{k=1}^s\exp(w^kz)\right)-1\right)\\&=t\,(\exp_s(z)-1)\end{aligned}$$

Now since b_t^s is the law of a, this gives the formula in the statement. $\square$

Let us study now the densities of $b_t^s,\tilde{b}_t^s$. We have here the following result:

THEOREM 12.16. *We have the formulae*

$$b_t^s = e^{-t}\sum_{p_1=0}^\infty\ldots\sum_{p_s=0}^\infty\frac{1}{p_1!\ldots p_s!}\left(\frac{t}{s}\right)^{p_1+\ldots+p_s}\delta\left(\sum_{k=1}^s w^kp_k\right)$$

$$\tilde{b}_t^s = e^{-t}\sum_{p_1=0}^\infty\ldots\sum_{p_s=0}^\infty\frac{1}{p_1!\ldots p_s!}\left(\frac{t}{s}\right)^{p_1+\ldots+p_s}\delta\left(\sum_{k=1}^s w^kp_k\right)^s$$

where $w=e^{2\pi i/s}$, and the δ symbol is a Dirac mass.

PROOF. It is enough to prove the formula for b_t^s. For this purpose, we compute the Fourier transform of the measure on the right. This is given by:

$$\begin{aligned}
F(z) &= e^{-t}\sum_{p_1=0}^\infty\ldots\sum_{p_s=0}^\infty\frac{1}{p_1!\ldots p_s!}\left(\frac{t}{s}\right)^{p_1+\ldots+p_s}F\delta\left(\sum_{k=1}^s w^kp_k\right)(z)\\
&= e^{-t}\sum_{p_1=0}^\infty\ldots\sum_{p_s=0}^\infty\frac{1}{p_1!\ldots p_s!}\left(\frac{t}{s}\right)^{p_1+\ldots+p_s}\exp\left(\sum_{k=1}^s w^kp_kz\right)\\
&= e^{-t}\sum_{r=0}^\infty\left(\frac{t}{s}\right)^r\sum_{\Sigma p_i=r}\frac{\exp\left(\sum_{k=1}^s w^kp_kz\right)}{p_1!\ldots p_s!}
\end{aligned}$$

We multiply by e^t, and we compute the derivative with respect to t:

$$\begin{aligned}
(e^tF(z))' &= \sum_{r=1}^\infty\frac{r}{s}\left(\frac{t}{s}\right)^{r-1}\sum_{\Sigma p_i=r}\frac{\exp\left(\sum_{k=1}^s w^kp_kz\right)}{p_1!\ldots p_s!}\\
&= \frac{1}{s}\sum_{r=1}^\infty\left(\frac{t}{s}\right)^{r-1}\sum_{\Sigma p_i=r}\left(\sum_{l=1}^s p_l\right)\frac{\exp\left(\sum_{k=1}^s w^kp_kz\right)}{p_1!\ldots p_s!}\\
&= \frac{1}{s}\sum_{r=1}^\infty\left(\frac{t}{s}\right)^{r-1}\sum_{\Sigma p_i=r}\sum_{l=1}^s\frac{\exp\left(\sum_{k=1}^s w^kp_kz\right)}{p_1!\ldots p_{l-1}!(p_l-1)!p_{l+1}!\ldots p_s!}
\end{aligned}$$

By using the variable $u=r-1$, we get:

$$\begin{aligned}
(e^tF(z))' &= \frac{1}{s}\sum_{u=0}^\infty\left(\frac{t}{s}\right)^u\sum_{\Sigma q_i=u}\sum_{l=1}^s\frac{\exp\left(w^lz+\sum_{k=1}^s w^kq_kz\right)}{q_1!\ldots q_s!}\\
&= \left(\frac{1}{s}\sum_{l=1}^s\exp(w^lz)\right)\left(\sum_{u=0}^\infty\left(\frac{t}{s}\right)^u\sum_{\Sigma q_i=u}\frac{\exp\left(\sum_{k=1}^s w^kq_kz\right)}{q_1!\ldots q_s!}\right)\\
&= (\exp_s z)(e^tF(z))
\end{aligned}$$

On the other hand, consider the following function:

$$\Phi(t) = \exp(t \exp_s z)$$

This function satisfies as well the equation found above, namely:

$$\Phi'(t) = (\exp_s z)\Phi(t)$$

Thus, we have the following equality of functions:

$$e^t F(z) = \Phi(t)$$

But this gives the following formula:

$$\begin{aligned}\log F &= \log(e^{-t} \exp(t \exp_s z)) \\ &= \log(\exp(t(\exp_s z - 1))) \\ &= t(\exp_s z - 1)\end{aligned}$$

Thus, we obtain the formulae in the statement. □

Let us discuss now the $s = \infty$ particular case of all the above, which will be of interest in what follows, along with the particular cases $s = 1, 2$. First, we have:

THEOREM 12.17. *The full complex reflection group, denoted*

$$K_N \subset U_N$$

and formed of the permutation-type matrices with numbers in $\mathbb{T}$ as nonzero entries,

$$K_N = M_N(\mathbb{T} \cup \{0\}) \cap U_N$$

has a wreath product decomposition, as follows,

$$K_N = \mathbb{T} \wr S_N$$

with S_N acting on $\mathbb{T}^N$ in the standard way, by permuting the factors.

PROOF. This is something that we already know from the above, at the $s = \infty$ particular case of the results established for the complex reflection groups H_N^s. □

At the probabilistic level, we have the following result:

THEOREM 12.18. *For the full reflection group $K_N \subset U_N$, the law of the variable*

$$\chi_t = \sum_{i=1}^{[tN]} g_{ii}$$

is in the $N \to \infty$ limit the purely complex Bessel law,

$$B_t = p_{t\varepsilon}$$

where ε is the uniform measure on $\mathbb{T}$.

PROOF. Once again, this is something that we already know from the above, at the $s=\infty$ particular case of the results established for the complex reflection groups H_N^s. $\square$

Let us end this presentation with some philosophical considerations, in connection with abstract probability theory. We have 4 main limiting results in probability, namely discrete and continuous, and real and complex, which are as follows:

$$\begin{tikzcd} CCPLT \arrow[r, dash] \arrow[d, dash] & CCLT \arrow[d, dash] \\ RCPLT \arrow[r, dash] & CLT \end{tikzcd}$$

We have seen that the limiting laws in these main limiting theorems are the real and complex Gaussian and Bessel laws, which are as follows:

$$\begin{tikzcd} B_t \arrow[r, dash] \arrow[d, dash] & G_t \arrow[d, dash] \\ b_t \arrow[r, dash] & g_t \end{tikzcd}$$

Moreover, we have seen that at the level of the moments, these come from certain collections of partitions, as follows:

$$\begin{tikzcd} \mathcal{P}_{even} \arrow[r, dash] \arrow[d, dash] & \mathcal{P}_2 \arrow[d, dash] \\ P_{even} \arrow[r, dash] & P_2 \end{tikzcd}$$

Finally, we have seen that there are some Lie groups behind all this, namely the basic real and complex rotation and reflection groups, as follows:

$$\begin{tikzcd} K_N \arrow[r, dash] \arrow[d, dash] & U_N \arrow[d, dash] \\ H_N \arrow[r, dash] & O_N \end{tikzcd}$$

All this is quite interesting, and we will be back to this, with more explanations, in chapters 13-16 below, by using some advanced representation theory tools.

Finally, in order for our discussion to be complete, we still have to talk about the general series of complex reflection groups H_N^{sd}. Here the determinant does not really contribute, in the $N \to \infty$ limit, and we obtain the same asymptotic laws as for the groups H_N^s. This is something that we have already seen in chapter 11 above, in relation with the alternating group $A_N \subset S_N$, and the proof in general is similar.

As before with other more advanced questions, we will leave all this for further discussion in chapters 13-16 below, after developing more powerful tools for dealing with such questions, and more specifically, advanced representation theory tools.

12d. Wigner laws

In the continuous group case now, as a continuation of the above investigations, an interesting input comes from the various computations from chapter 6 above. In order to discuss all this, let us first recall some useful formulae from that chapter 6. One of the key results there, which is very useful in practice, was as follows:

PROPOSITION 12.19. *We have the following formula,*

$$\int_0^{\pi/2}\cos^p t\sin^q t\,dt=\left(\frac{\pi}{2}\right)^{\varepsilon(p)\varepsilon(q)}\frac{p!!q!!}{(p+q+1)!!}$$

where $\varepsilon(p)=1$ if p is even, and $\varepsilon(p)=0$ if p is odd, and where

$$m!!=(m-1)(m-3)(m-5)\ldots$$

with the product ending at 2 if m is odd, and ending at 1 if m is even.

PROOF. This is something that follows by partial integration, and then a double recurrence on $p,q\in\mathbb{N}$, worked out in chapter 6 above. $\square$

More generally now, we can in fact compute the polynomial integrals over the unit sphere in arbitrary dimensions, the result here being as follows:

THEOREM 12.20. *The polynomial integrals over the unit sphere $S^{N-1}_{\mathbb{R}}\subset\mathbb{R}^N$, with respect to the normalized, mass 1 measure, are given by the following formula,*

$$\int_{S^{N-1}_{\mathbb{R}}}x_1^{k_1}\ldots x_N^{k_N}\,dx=\frac{(N-1)!!k_1!!\ldots k_N!!}{(N+\Sigma k_i-1)!!}$$

valid when all exponents k_i are even. If an exponent is odd, the integral vanishes.

PROOF. As before this is something that we know from chapter 6, the idea being that the $N=2$ case is solved by Proposition 12.19 above, and that the general case, $N\in\mathbb{N}$, follows from this, by using spherical coordinates and the Fubini theorem. $\square$

As an application of the above formula, we can investigate the hyperspherical laws and their asymptotics, and we have the following result:

THEOREM 12.21. *The moments of the hyperspherical variables are*

$$\int_{S^{N-1}_\mathbb{R}}x_i^kdx=\frac{(N-1)!!k!!}{(N+k-1)!!}$$

and the normalized hyperspherical variables

$$y_i=\frac{x_i}{\sqrt{N}}$$

become normal and independent with $N\to\infty$.

PROOF. We have two things to be proved, the idea being as follows:

(1) The formula in the statement follows from the general integration formula over the sphere, established above, which is as follows:

$$\int_{S^{N-1}_\mathbb{R}}x_{i_1}\ldots x_{i_k}\,dx=\frac{(N-1)!!l_1!!\ldots l_N!!}{(N+\sum l_i-1)!!}$$

Indeed, with $i_1=\ldots=i_k=i$, we obtain from this:

$$\int_{S^{N-1}_\mathbb{R}}x_i^kdx=\frac{(N-1)!!k!!}{(N+k-1)!!}$$

Now observe that with $N\to\infty$ we have the following estimate:

$$\begin{aligned}\int_{S^{N-1}_\mathbb{R}}x_i^kdx&=\frac{(N-1)!!}{(N+k-1)!!}\times k!!\\&\simeq N^{k/2}k!!\\&=N^{k/2}M_k(g_1)\end{aligned}$$

Thus, the variables $y_i=\frac{x_i}{\sqrt{N}}$ become normal with $N\to\infty$.

(2) As for the asymptotic independence result, this is standard as well, once again by using Theorem 12.20, for computing mixed moments, and taking the $N\to\infty$ limit. □

We can talk as well about rotations, as follows:

THEOREM 12.22. *We have the integration formula*

$$\int_{O_N}U_{ij}^kdU=\frac{(N-1)!!k!!}{(N+k-1)!!}$$

and the normalized coordinates on O_N, constructed as follows,

$$V_{ij}=\frac{U_{ij}}{\sqrt{N}}$$

become normal and independent with $N\to\infty$.

PROOF. We use the well-known fact that we have an embedding as follows, for any i, which makes correspond the respective integration functionals:

$$C(S^{N-1}_{\mathbb{R}}) \subset C(O_N)$$
$$x_i \to U_{1i}$$

With this identification made, the result follows from Theorem 12.21. ☐

We have similar results in the unitary case. First, we have:

THEOREM 12.23. *We have the following integration formula over the complex sphere $S^{N-1}_{\mathbb{C}} \subset \mathbb{R}^N$, with respect to the normalized measure,*

$$\int_{S^{N-1}_{\mathbb{C}}} |z_1|^{2l_1} \ldots |z_N|^{2l_N} \, dz = 4^{\sum l_i} \frac{(2N-1)! l_1! \ldots l_n!}{(2N + \sum l_i - 1)!}$$

valid for any exponents $l_i \in \mathbb{N}$. As for the other polynomial integrals in $z_1, \ldots, z_N$ and their conjugates $\bar{z}_1, \ldots, \bar{z}_N$, these all vanish.

PROOF. As before, this is something that we know from chapter 6 above, and which can be proved either directly, or by using the formula in Theorem 12.20 above. ☐

We can talk about complex hyperspherical laws, and we have:

THEOREM 12.24. *The rescaled coordinates on the complex sphere $S^{N-1}_{\mathbb{C}}$,*

$$w_i = \frac{z_i}{\sqrt{N}}$$

become complex Gaussian and independent with $N \to \infty$.

PROOF. This follows indeed by using Theorem 12.22 and Theorem 12.23. ☐

In relation now with Lie groups, the result that we obtain is as follows:

THEOREM 12.25. *For the unitary group U_N, the normalized coordinates*

$$V_{ij} = \frac{U_{ij}}{\sqrt{N}}$$

become complex Gaussian and independent with $N \to \infty$.

PROOF. We use the well-known fact that we have an embedding as follows, for any i, which makes correspond the respective integration functionals:

$$C(S^{N-1}_{\mathbb{C}}) \subset C(U_N)$$
$$x_i \to U_{1i}$$

With this identification made, the result follows from Theorem 12.21. ☐

Our claim now is that the above results can be reformulated in terms of the truncated characters introduced in chapter 11. Let us recall indeed from there that we have:

Definition 12.26. *Given a closed subgroup $G\subset U_N$, the function*

$$\chi:G\to\mathbb C\quad,\quad \chi_t(g)=\sum_{i=1}^{[tN]}g_{ii}$$

is called main truncated character of G, of parameter $t\in(0,1]$.

We refer to chapter 11 above for more on this notion. Also, we will be back to all this in chapters 13-16 below, with more details about all this, and motivations.

In connection with the present considerations, the point is that with the above notion in hand, our above results reformulate as follows:

Theorem 12.27. *For the orthogonal and unitary groups O_N,U_N, the rescalings*

$$\chi=\frac{\chi_{1/N}}{\sqrt N}$$

become respectively real and complex Gaussian, in the $N\to\infty$ limit.

Proof. According to our conventions, given a closed subgroup $G\subset U_N$, the main character truncated at $t=1/N$ is simply the first coordinate:

$$\chi_{1/N}(g)=g_{11}$$

With this remark made, the conclusions from the statement follow from the computations performed above, for the laws of coordinates on O_N,U_N. □

It is possible to get beyond such results, by using advanced representation theory methods, with full results about all the truncated characters, and in particular about the main characters. We will be back to this in chapters 13-16 below.

Also, it is possible to compute as well the laws of individual coordinates for some of the remaining continuous groups. To be more precise, it is possible to work out results for the bistochastic groups B_N,C_N, as well as for the symplectic groups Sp_N, the computations being not very complicated. However, we prefer here to defer the discussion to chapters 13-16 below, after developing some appropriate tools, for dealing with such questions.

As a last topic now, let us discuss the case where N is fixed. Things are quite complicated here, and as a main goal, we would like to find the law of the main character for our favorite rotation groups, namely SU_2 and SO_3.

In order to do so, we will need some combinatorial preliminaries. We first have the following well-known result, which is the cornerstone of all modern combinatorics:

THEOREM 12.28. *The Catalan numbers, which are by definition given by*

$$C_k = |NC_2(2k)|$$

satisfy the following recurrence formula,

$$C_{k+1} = \sum_{a+b=k} C_a C_b$$

and their generating series, given by

$$f(z) = \sum_{k\geq 0} C_k z^k$$

satisfies the following degree 2 equation,

$$zf^2 - f + 1 = 0$$

and we have the following explicit formula for these numbers:

$$C_k = \frac{1}{k+1}\binom{2k}{k}$$

PROOF. We must count the noncrossing pairings of $\{1,\ldots,2k\}$. But such a pairing appears by pairing 1 to an odd number, $2a+1$, and then inserting a noncrossing pairing of $\{2,\ldots,2a\}$, and a noncrossing pairing of $\{2a+2,\ldots,2l\}$. We conclude from this that we have the following recurrence formula for the Catalan numbers:

$$C_k = \sum_{a+b=k-1} C_a C_b$$

In terms of the generating series f, the above recurrence gives:

$$\begin{aligned} zf^2 &= \sum_{a,b\geq 0} C_a C_b z^{a+b+1} \\ &= \sum_{k\geq 1}\sum_{a+b=k-1} C_a C_b z^k \\ &= \sum_{k\geq 1} C_k z^k \\ &= f - 1 \end{aligned}$$

Thus the generating series f satisfies the following degree 2 equation:

$$zf^2 - f + 1 = 0$$

By choosing the solution which is bounded at $z=0$, we obtain:

$$f(z) = \frac{1-\sqrt{1-4z}}{2z}$$

By using now the Taylor formula for $\sqrt{x}$, we obtain the following formula:

$$f(z)=\sum_{k\geq 0}\frac{1}{k+1}\binom{2k}{k}z^k$$

It follows that the Catalan numbers are given by the formula the statement. $\square$

The Catalan numbers are central objects in probability as well, and we have the following key result here, complementing the formulae from Theorem 12.28:

THEOREM 12.29. *The normalized Wigner semicircle law, which is by definition*

$$\gamma_1=\frac{1}{2\pi}\sqrt{4-x^2}dx$$

has the Catalan numbers as even moments. As for the odd moments, these all vanish.

PROOF. The even moments of the Wigner law can be computed with the change of variable $x=2\cos t$, and we are led to the following formula:

$$\begin{aligned}
M_{2k} &= \frac{1}{\pi}\int_0^2\sqrt{4-x^2}x^{2k}dx\\
&= \frac{1}{\pi}\int_0^{\pi/2}\sqrt{4-4\cos^2 t}\,(2\cos t)^{2k}2\sin t\,dt\\
&= \frac{4^{k+1}}{\pi}\int_0^{\pi/2}\cos^{2k}t\sin^2 t\,dt\\
&= \frac{4^{k+1}}{\pi}\cdot\frac{\pi}{2}\cdot\frac{(2k)!!2!!}{(2k+3)!!}\\
&= 2\cdot 4^k\cdot\frac{(2k)!/2^k k!}{2^{k+1}(k+1)!}\\
&= C_k
\end{aligned}$$

As for the odd moments, these all vanish, because the density of γ_1 is an even function. Thus, we are led to the conclusion in the statement. $\square$

We can now formulate our result regarding SU_2, as follows:

THEOREM 12.30. *The main character of SU_2, given by*

$$\chi\begin{pmatrix}a&b\\-\bar{b}&\bar{a}\end{pmatrix}=2Re(a)$$

follows a Wigner semicircle law γ_1.

PROOF. This follows by identifying SU_2 with the sphere $S^3_{\mathbb{R}}\subset\mathbb{R}^4$, and the uniform measure on SU_2 with the uniform measure on this sphere.

Indeed, in real notation for the standard parametrization of SU_2, from chapter 9 above, we have the following formula, for the main character of SU_2:

$$\chi\begin{pmatrix}x+iy&z+it\\-z+it&x-iy\end{pmatrix}=2x$$

We are therefore left with computing the law of the following variable:

$$x\in C(S^3_\mathbb{R})$$

But for this purpose, we can use the moment method. Indeed, by using the trigonometric integral formulae from chapter 6 above, we obtain:

$$\begin{aligned}\int_{S^3_\mathbb{R}}x^{2k}&=\frac{3!!(2k)!!}{(2k+3)!!}\\&=2\cdot\frac{3\cdot5\cdot7\ldots(2k-1)}{2\cdot4\cdot6\ldots(2k+2)}\\&=2\cdot\frac{(2k)!}{2^kk!2^{k+1}(k+1)!}\\&=\frac{1}{4^k}\cdot\frac{1}{k+1}\binom{2k}{k}\\&=\frac{C_k}{4^k}\end{aligned}$$

Thus the variable $2x\in C(S^3_\mathbb{R})$ has the Catalan numbers as even moments, and so by Theorem 12.28 its distribution is the Wigner semicircle law γ_1, as claimed. $\square$

In order to do the computation for SO_3, we will need some more probabilistic preliminaries, which are standard random matrix theory material. Let us start with:

Proposition 12.31. *We have a bijection between noncrossing partitions and pairings*

$$NC(k)\simeq NC_2(2k)$$

constructed as follows:

1. *The application $NC(k)\to NC_2(2k)$ is the "fattening" one, obtained by doubling all the legs, and doubling all the strings as well.*
2. *Its inverse $NC_2(2k)\to NC(k)$ is the "shrinking" application, obtained by collapsing pairs of consecutive neighbors.*

Proof. The fact that the two operations in the statement are indeed inverse to each other is clear, by computing the corresponding two compositions, with the remark that the construction of the fattening operation requires the partitions to be noncrossing. $\square$

As a consequence of the above result, we have a new look on the Catalan numbers, which is more adapted to our present SO_3 considerations, as follows:

PROPOSITION 12.32. *The Catalan numbers $C_k = |NC_2(2k)|$ appear as well as*

$$C_k = |NC(k)|$$

where $NC(k)$ is the set of all noncrossing partitions of $\{1, \ldots, k\}$.

PROOF. This follows indeed from Proposition 12.31 above. □

Let us formulate now the following definition:

DEFINITION 12.33. *The standard Marchenko-Pastur law π_1 is given by:*

$$f \sim \gamma_1 \implies f^2 \sim \pi_1$$

That is, π_1 is the law of the square of a variable following the semicircle law γ_1.

Here the fact that π_1 is indeed well-defined comes from the fact that a measure is uniquely determined by its moments. More explicitely now, we have:

PROPOSITION 12.34. *The density of the Marchenko-Pastur law is*

$$\pi_1 = \frac{1}{2\pi}\sqrt{4x^{-1} - 1}\, dx$$

and the moments of this measure are the Catalan numbers.

PROOF. There are several proofs here, either by using Definition 12.33, or by Stieltjes inversion, of just by cheating. Whis this latter method, the point is that the moments of the law in the statement can be computed with the following change of variables:

$$x = 4\cos^2 t$$

To be more precise, the moments of the law in the statement can be computed with the change of variable $x = 4\cos^2 t$, and we are led to the following formula:

$$\begin{aligned}
M_k &= \frac{1}{2\pi}\int_0^4 \sqrt{4x^{-1} - 1}\, x^k dx \\
&= \frac{1}{2\pi}\int_0^{\pi/2} \frac{\sin t}{\cos t} \cdot (4\cos^2 t)^k \cdot 2\cos t \sin t\, dt \\
&= \frac{4^{k+1}}{\pi}\int_0^{\pi/2} \cos^{2k} t \sin^2 t\, dt \\
&= \frac{4^{k+1}}{\pi} \cdot \frac{\pi}{2} \cdot \frac{(2k)!!2!!}{(2k+3)!!} \\
&= 2 \cdot 4^k \cdot \frac{(2k)!/2^k k!}{2^{k+1}(k+1)!} \\
&= C_k
\end{aligned}$$

Thus, we are led to the conclusion in the statement. □

We can do now the character computation for SO_3, as follows:

THEOREM 12.35. *The main character of SO_3, modified by adding 1 to it, given in standard Euler-Rodrigues coordinates by*

$$\chi = 3x^2 - y^2 - z^2 - t^2$$

follows a squared semicircle law, or Marchenko-Pastur law π_1.

PROOF. This follows by using the canonical quotient map $SU_2 \to SO_3$, and the result for SU_2 from Theorem 12.30. To be more precise, let us recall from chapter 9 above that the elements of SU_2 can be parametrized as follows:

$$U = \begin{pmatrix} x+iy & z+it \\ -z+it & x-iy \end{pmatrix}$$

As for the elements of SO_3, these can be parametrized as follows:

$$V = \begin{pmatrix} x^2+y^2-z^2-t^2 & 2(yz-xt) & 2(xz+yt) \\ 2(xt+yz) & x^2+z^2-y^2-t^2 & 2(zt-xy) \\ 2(yt-xz) & 2(xy+zt) & x^2+t^2-y^2-z^2 \end{pmatrix}$$

The point now is that, by using these formulae, in the context of Theorem 12.30, the main character of SO_3 is then given by:

$$\chi = 4Re(a)^2$$

Now recall from the proof of Theorem 12.30 above that we have:

$$2Re(a) \sim \gamma_1$$

On the other hand, a quick comparison between the moment formulae for the Wigner and Marchenko-Pastur laws, which are very similar, shows that we have:

$$f \sim \gamma_1 \implies f^2 \sim \pi_1$$

Thus, with $f = 2Re(a)$, we obtain the result in the statement. □

As an interesting question now, appearing from the above, and which is quite philosophical, we have the problem of understanding how the Wigner and Marchenko-Pastur laws γ_t, π_t fit in regards with the main limiting laws from classical probability.

The answer here is quite tricky, the idea being that, with a suitable formalism for freeness, γ_t, π_t can be thought of as being "free analogues" of the Gaussian and Poisson laws g_t, p_t. This is something quite subtle, requiring some further knowledge, and we will be back to this in chapters 13-16 below, when doing representation theory.

In any case, as a conclusion, we have now a pretty decent level in probability. And if you want to learn right away a bit more about the modern topics here, such as random matrices and free probability, from Marchenko-Pastur [69], Mehta [70], Voiculescu [85], Wigner [95], you can go for it. As for more groups and algebra, stay with us.

12e. Exercises

There has been a lot of technical material in this chapter, and technical as well will be most of our exercises. First, we have:

EXERCISE 12.36. *Work out an alternative proof for the main result regarding the truncated characters of the hyperoctahedral group H_N, namely*

$$\chi_t \sim e^{-t} \sum_{k=-\infty}^{\infty} \delta_k \sum_{p=0}^{\infty} \frac{(t/2)^{|k|+2p}}{(|k|+p)!p!}$$

with $N \to \infty$, by working our first an explicit formula for the polynomials integrals over H_N, and then using that for computing the laws of truncated characters.

The idea here is of course that of adapting the computation that we have for the symmetric group S_N, from the previous chapter.

EXERCISE 12.37. *Work out all the details for the moment formula for the Bessel laws,*

$$M_k = |P^s(k)|$$

where $P^s(k)$ are the partitions satisfying the formula

$$\#\circ = \#\bullet\,(s)$$

as a weighted sum, in each block.

This is something that we briefly discussed in the above, and the problem is now that of working out all the details, first as $s = 1, 2, \infty$, and then in general.

EXERCISE 12.38. *Work out all the details for the truncated character formula for H_N^s,*

$$\chi_t \sim b_t^s$$

where $b_t^s = p_{t\varepsilon_s}$, with ε_s being the uniform measure on the s-th roots of unity.

As before, this is something that we briefly discussed in the above, and the problem is now that of working out all the details, first as $s = 1, 2, \infty$, and then in general.

EXERCISE 12.39. *Show that the passage from H_N^s to H_N^{sd} does not change the asymptotic laws of the truncated characters.*

This is something that we discussed in the previous chapter, in a particular case, namely for the passage from the symmetric group S_N to the alternating group A_N. The problem is that of having this done in general, by using a similar method.

EXERCISE 12.40. *Compute the asymptotic laws of characters and coordinates for the bistochastic groups B_N and C_N, as well as for the symplectic group $Sp_N \subset U_N$.*

These computations are all quite standard, the idea being that, by using the various group theory considerations from chapter 10, the computation for B_N is quite similar to that for O_{N-1}, the computation for C_N is quite similar to that for U_{N-1}, and the computation for Sp_N is quite similar to that for O_{N-1}.

EXERCISE 12.41. *Compute the character laws for the groups O_1, SO_1, then for the groups U_1, SU_1, and then for the groups O_2, SO_2.*

As before with the previous exercise, the computations here are quite standard. In fact, the more difficult questions of this type concern the next groups in the above series, namely SU_2 and SO_3, which were discussed in the above.

EXERCISE 12.42. *Work out all the combinatorics and calculus details in relation with the Wigner and Marchenko-Pastur laws, and their moments, the Catalan numbers.*

This is a very instructive exercise, with lots of nice combinatorics involved. Most of this combinatorics was actually already discussed in the above.

EXERCISE 12.43. *Look up the Wigner and Marchenko-Pastur laws, given by*

$$\gamma_1 = \frac{1}{2\pi}\sqrt{4-x^2}dx \quad , \quad \pi_1 = \frac{1}{2\pi}\sqrt{4x^{-1}-1}\,dx$$

and write down a brief account of what you found, and understood.

Here the solution is not unique, the Wigner and Marchenko-Pastur laws appearing in random matrix theory, or free probability, or both. All nice mathematics, enjoy.

Part IV

Haar integration

And the band plays Waltzing Matilda
And the old men still answer the call
But year after year, their numbers get fewer
Someday, no one will march there at all

CHAPTER 13

Representations

13a. Basic theory

We have seen so far that some algebraic and probabilistic theory for the finite subgroups $G \subset U_N$, ranging from elementary to quite advanced, can be developed. We have seen as well a few computations for the continuous compact subgroups $G \subset U_N$. In what follows we develop some systematic theory for the arbitrary closed subgroups $G \subset U_N$, covering both the finite and the infinite case. The main examples that we have in mind, and the questions that we would like to solve for them, are as follows:

(1) The orthogonal and unitary groups O_N, U_N. Here we would like to have an integration formula, and results about character laws, in the $N \to \infty$ limit.

(2) Various versions of O_N, U_N, such as the bistochastic groups B_N, C_N, or the symplectic groups Sp_N, with similar questions to be solved.

(3) The reflection groups $H_N^{sd} \subset U_N$, with results about characters extending, or at least putting in a more conceptual framework, what we already have.

There is a lot of theory to be developed, and we will do this gradually. To be more precise, in this chapter and in the next one we will work out algebraic aspects, and then in the chapter afterwards and in the last one we will use these algebraic techniques, in order to work out probabilistic results, and in particular to answer the above questions. As before, the main notion that we will be interested in is that of a representation:

Definition 13.1. *A representation of a compact group G is a continuous group morphism, which can be faithful or not, into a unitary group:*

$$u : G \to U_N$$

The character of such a representation is the function $\chi : G \to \mathbb{C}$ given by

$$g \to Tr(u_g)$$

where Tr is the usual trace of the $N \times N$ matrices, $Tr(M) = \sum_i M_{ii}$.

As a basic example here, for any compact group we always have available the trivial 1-dimensional representation, which is by definition as follows:

$$u : G \to U_1 \quad , \quad g \to (1)$$

At the level of non-trivial examples now, most of the compact groups that we met so far, finite or continuous, naturally appear as closed subgroups $G\subset U_N$. In this case, the embedding $G\subset U_N$ is of course a representation, called fundamental representation:

$$u:G\subset U_N\quad,\quad g\to g$$

In this situation, there are many other representations of G, which are equally interesting. For instance, we can define the representation conjugate to u, as being:

$$\bar{u}:G\subset U_N\quad,\quad g\to\bar{g}$$

In order to clarify all this, and see which representations are available, let us first discuss the various operations on the representations. The result here is as follows:

Proposition 13.2. *The representations of a given compact group G are subject to the following operations:*

1. *Making sums. Given representations u,v, having dimensions N,M, their sum is the $N+M$-dimensional representation $u+v=diag(u,v)$.*
2. *Making products. Given representations u,v, having dimensions N,M, their tensor product is the NM-dimensional representation $(u\otimes v)_{ia,jb}=u_{ij}v_{ab}$.*
3. *Taking conjugates. Given a representation u, having dimension N, its complex conjugate is the N-dimensional representation $(\bar{u})_{ij}=\bar{u}_{ij}$.*
4. *Spinning by unitaries. Given a representation u, having dimension N, and a unitary $V\in U_N$, we can spin u by this unitary, $u\to VuV^*$.*

Proof. The fact that the operations in the statement are indeed well-defined, among maps from G to unitary groups, can be checked as follows:

(1) This follows from the trivial fact that if $g\in U_N$ and $h\in U_M$ are two unitaries, then their diagonal sum is a unitary too, as follows:

$$\begin{pmatrix}g&0\\0&h\end{pmatrix}\in U_{N+M}$$

(2) This follows from the fact that if $g\in U_N$ and $h\in U_M$ are two unitaries, then $g\otimes h\in U_{NM}$ is a unitary too. Given unitaries g,h, let us set indeed:

$$(g\otimes h)_{ia,jb}=g_{ij}h_{ab}$$

This matrix is then a unitary too, due to the following computation:

$$\begin{aligned}[(g\otimes h)(g\otimes h)^*]_{ia,jb} &= \sum_{kc}(g\otimes h)_{ia,kc}((g\otimes h)^*)_{kc,jb}\\ &= \sum_{kc}(g\otimes h)_{ia,kc}\overline{(g\otimes h)_{jb,kc}}\\ &= \sum_{kc} g_{ik}h_{ac}\bar{g}_{jk}\bar{h}_{bc}\\ &= \sum_k g_{ik}\bar{g}_{jk}\sum_c h_{ac}\bar{h}_{bc}\\ &= \delta_{ij}\delta_{ab}\end{aligned}$$

(3) This simply follows from the fact that if $g\in U_N$ is unitary, then so is its complex conjugate, $\bar{g}\in U_N$, and this due to the following formula, obtained by conjugating:

$$g^*=g^{-1}\implies g^t=\bar{g}^{-1}$$

(4) This is clear as well, because if $g\in U_N$ is unitary, and $V\in U_N$ is another unitary, then we can spin g by this unitary, and we obtain a unitary as follows:

$$VgV^*\in U_N$$

Thus, our operations are well-defined, and this leads to the above conclusions. □

In relation now with characters, we have the following result:

Proposition 13.3. *We have the following formulae, regarding characters*

$$\chi_{u+v}=\chi_u+\chi_v\quad,\quad\chi_{u\otimes v}=\chi_u\chi_v\quad,\quad\chi_{\bar{u}}=\bar{\chi}_u\quad,\quad\chi_{VuV^*}=\chi_u$$

in relation with the basic operations for the representations.

Proof. All these assertions are elementary, by using the following well-known trace formulae, valid for any two square matrices g,h, and any unitary V:

$$Tr(diag(g,h))=Tr(g)+Tr(h)$$

$$Tr(g\otimes h)=Tr(g)Tr(h)$$

$$Tr(\bar{g})=\overline{Tr(g)}$$

$$Tr(VgV^*)=Tr(g)$$

To be more precise, the first formula is clear from definitions. Regarding now the second formula, the computation here is immediate too, as follows:

$$\begin{aligned} Tr(g \otimes h) &= \sum_{ia} (g \otimes h)_{ia,ia} \\ &= \sum_{ia} g_{ii} h_{aa} \\ &= Tr(g)Tr(h) \end{aligned}$$

Regarding now the third formula, this is clear from definitions, by conjugating. Finally, regarding the fourth formula, this can be established as follows:

$$Tr(VgV^*) = Tr(gV^*V) = Tr(g)$$

Thus, we are led to the conclusions in the statement. □

Assume now that we are given a closed subgroup $G \subset U_N$. By using the above operations, we can construct a whole family of representations of G, as follows:

Definition 13.4. *Given a closed subgroup $G \subset U_N$, its Peter-Weyl representations are the tensor products between the fundamental representation and its conjugate:*

$$u : G \subset U_N \quad , \quad \bar{u} : G \subset U_N$$

We denote these tensor products $u^{\otimes k}$, with $k = \circ \bullet \bullet \circ \ldots$ being a colored integer, with the colored tensor powers being defined according to the rules

$$u^{\otimes \circ} = u \quad , \quad u^{\otimes \bullet} = \bar{u} \quad , \quad u^{\otimes kl} = u^{\otimes k} \otimes u^{\otimes l}$$

and with the convention that $u^{\otimes \emptyset}$ is the trivial representation $1 : G \to U_1$.

Here are a few examples of such Peter-Weyl representations, namely those coming from the colored integers of length 2, to be often used in what follows:

$$u^{\otimes \circ\circ} = u \otimes u \quad , \quad u^{\otimes \circ\bullet} = u \otimes \bar{u}$$

$$u^{\otimes \bullet\circ} = \bar{u} \otimes u \quad , \quad u^{\otimes \bullet\bullet} = \bar{u} \otimes \bar{u}$$

In relation now with characters, we have the following result:

Proposition 13.5. *The characters of Peter-Weyl representations are given by*

$$\chi_{u^{\otimes k}} = (\chi_u)^k$$

with the colored powers of a variable χ being by definition given by

$$\chi^\circ = \chi \quad , \quad \chi^\bullet = \bar{\chi} \quad , \quad \chi^{kl} = \chi^k \chi^l$$

and with the convention that $\chi^\emptyset$ equals by definition 1.

Proof. This follows indeed from the additivity, multiplicativity and conjugation formulae established in Proposition 13.3, via the conventions in Definition 13.4. □

Getting back now to our motivations, we can see the interest in the above constructions. Indeed, the joint moments of the main character $\chi = \chi_u$ and its adjoint $\bar{\chi} = \chi_{\bar{u}}$ are simply the expectations of the characters of various Peter-Weyl representations:

$$\int_G \chi^k = \int_G \chi_{u^{\otimes k}}$$

Summarizing, given a closed subgroup $G \subset U_N$, we would like to understand its Peter-Weyl representations, and compute the expectations of the characters of these representations. In order to do so, let us formulate the following key definition:

Definition 13.6. *Given a compact group G, and two of its representations,*

$$u : G \to U_N \quad , \quad v : G \to U_M$$

we define the linear space of intertwiners between these representations as being

$$Hom(u,v) = \left\{ T \in M_{M \times N}(\mathbb{C}) \Big| Tu_g = v_g T, \forall g \in G \right\}$$

and we use the following conventions:

1. *We use the notations $Fix(u) = Hom(1,u)$, and $End(u) = Hom(u,u)$.*
2. *We write $u \sim v$ when $Hom(u,v)$ contains an invertible element.*
3. *We say that u is irreducible, and write $u \in Irr(G)$, when $End(u) = \mathbb{C}1$.*

The terminology here is standard, with Hom and End standing for "homomorphisms" and "endomorphisms", and with Fix standing for "fixed points". In practice, it is useful to think of the representations of G as being the objects of some kind of abstract combinatorial structure associated to G, and of the intertwiners between these representations as being the "arrows" between these objects. We have in fact the following result:

Theorem 13.7. *The following happen:*

1. *The intertwiners are stable under composition:*
$$T \in Hom(u,v) \ , \ S \in Hom(v,w) \implies ST \in Hom(u,w)$$
2. *The intertwiners are stable under taking tensor products:*
$$S \in Hom(u,v) \ , \ T \in Hom(w,t) \implies S \otimes T \in Hom(u \otimes w, v \otimes t)$$
3. *The intertwiners are stable under taking adjoints:*
$$T \in Hom(u,v) \implies T^* \in Hom(v,u)$$
4. *Thus, the Hom spaces form a tensor $*$-category.*

Proof. All this is clear from definitions, the verifications being as follows:

(1) This follows indeed from the following computation, valid for any $g \in G$:

$$STu_g = Sv_g T = w_g ST$$

(2) Again, this is clear, because we have the following computation:

$$(S \otimes T)(u_g \otimes w_g) = Su_g \otimes Tw_g = v_g S \otimes t_g T = (v_g \otimes t_g)(S \otimes T)$$

(3) This follows from the following computation, valid for any $g \in G$:

$$Tu_g = v_g T \implies u_g^* T^* = T^* v_g^* \implies T^* v_g = u_g T^*$$

(4) This is just a conclusion of (1,2,3), with a tensor $*$-category being by definition an abstract beast satisfying these conditions (1,2,3). We will be back to tensor categories later on, in chapter 14 below, with more details on all this. □

The above result is quite interesting, because it shows that the combinatorics of a compact group G is described by a certain collection of linear spaces, which can be in principle investigated by using tools from linear algebra. Thus, what we have here is a "linearization" idea. We will heavily use this idea, in what follows.

13b. Peter-Weyl theory

In what follows we develop a systematic theory of the representations of the compact groups G, with emphasis on the Peter-Weyl representations, in the closed subgroup case $G \subset U_N$, that we are mostly interested in. Before starting, some comments:

(1) First of all, all this goes back to Hermann Weyl, who along with Einstein, Poincaré and a few others was part of the last generation of great mathematicians and physicists, knowing everything about mathematics, and everything about physics too. Get to know more about him, and have a look at his books [**92**], [**93**], [**94**].

(2) In what concerns Peter-Weyl theory, which is something quite tricky, this is perhaps best learned for the finite groups first, and for more general groups afterwards. This is how I learned it myself, back when I was a student, by reading Serre [**79**] for finite groups, and then Woronowicz [**97**], [**98**] directly for the compact quantum groups.

(3) Getting back now to the present book, we will be a bit in a hurry with this, because we have so many other things to talk about. So, we will skip the finite group preliminaries, and deal directly with the compact groups. And by using a somewhat rough functional analysis viewpoint. For a complement to all this, we recommend Serre [**79**].

As a starting point, as a main consequence of Theorem 13.7, we have:

Theorem 13.8. *Given a representation of a compact group $u : G \to U_N$, the corresponding linear space of self-intertwiners*

$$End(u) \subset M_N(\mathbb{C})$$

is a $*$-algebra, with respect to the usual involution of the matrices.

PROOF. By definition, the space $End(u)$ is a linear subspace of $M_N(\mathbb{C})$. We know from Theorem 13.7 (1) that this subspace $End(u)$ is a subalgebra of $M_N(\mathbb{C})$, and then we know as well from Theorem 13.7 (3) that this subalgebra is stable under the involution $*$. Thus, what we have here is a $*$-subalgebra of $M_N(\mathbb{C})$, as claimed. $\square$

The above result is quite interesting, because it gets us into linear algebra. To be more precise, associated to any group representation $u : G \to U_N$ is now a quite familiar object, namely the algebra $End(u) \subset M_N(\mathbb{C})$. In order to exploit this fact, we will need a well-known result, complementing the theory developed in chapter 8, as follows:

THEOREM 13.9. *Let $A \subset M_N(\mathbb{C})$ be a $*$-algebra.*

(1) *The unit decomposes as follows, with $p_i \in A$ being central minimal projections:*

$$1 = p_1 + \ldots + p_k$$

(2) *Each of the following linear spaces is a non-unital $*$-subalgebra of A:*

$$A_i = p_i A p_i$$

(3) *We have a non-unital $*$-algebra sum decomposition, as follows:*

$$A = A_1 \oplus \ldots \oplus A_k$$

(4) *We have unital $*$-algebra isomorphisms as follows, with $N_i = rank(p_i)$:*

$$A_i \simeq M_{N_i}(\mathbb{C})$$

(5) *Thus, we have a $*$-algebra isomorphism as follows:*

$$A \simeq M_{N_1}(\mathbb{C}) \oplus \ldots \oplus M_{N_k}(\mathbb{C})$$

PROOF. Consider indeed an arbitrary $*$-algebra of the $N \times N$ matrices, $A \subset M_N(\mathbb{C})$. Let us first look at the center of this algebra, $Z(A) = A \cap A'$. It is elementary to prove that this center, as an algebra, is of the following form:

$$Z(A) \simeq \mathbb{C}^k$$

Consider now the standard basis $e_1, \ldots, e_k \in \mathbb{C}^k$, and let $p_1, \ldots, p_k \in Z(A)$ be the images of these vectors via the above identification. In other words, these elements $p_1, \ldots, p_k \in A$ are central minimal projections, summing up to 1:

$$p_1 + \ldots + p_k = 1$$

The idea is then that this partition of the unity will eventually lead to the block decomposition of A, as in the statement. We prove this in 4 steps, as follows:

Step 1. We first construct the matrix blocks, our claim here being that each of the following linear subspaces of A are non-unital $*$-subalgebras of A:

$$A_i = p_i A p_i$$

But this is clear, with the fact that each A_i is closed under the various non-unital $*$-subalgebra operations coming from the projection equations $p_i^2 = p_i = p_i^*$.

Step 2. We prove now that the above algebras $A_i \subset A$ are in a direct sum position, in the sense that we have a non-unital $*$-algebra sum decomposition, as follows:

$$A = A_1 \oplus \ldots \oplus A_k$$

As with any direct sum question, we have two things to be proved here. First, by using the formula $p_1 + \ldots + p_k = 1$ and the projection equations $p_i^2 = p_i = p_i^*$, we conclude that we have the needed generation property, namely:

$$A_1 + \ldots + A_k = A$$

As for the fact that the sum is indeed direct, this follows as well from the formula $p_1 + \ldots + p_k = 1$, and from the projection equations $p_i^2 = p_i = p_i^*$.

Step 3. Our claim now, which will finish the proof, is that each of the $*$-subalgebras $A_i = p_i A p_i$ constructed above is a full matrix algebra. To be more precise here, with $r_i = rank(p_i)$, our claim is that we have isomorphisms, as follows:

$$A_i \simeq M_{r_i}(\mathbb{C})$$

In order to prove this claim, recall that the projections $p_i \in A$ were chosen central and minimal. Thus, the center of each of the algebras A_i reduces to the scalars:

$$Z(A_i) = \mathbb{C}$$

But this shows, either via a direct computation, or via the bicommutant theorem, that the each of the algebras A_i is a full matrix algebra, as claimed.

Step 4. We can now obtain the result, by putting together what we have. Indeed, by using the results from Step 2 and Step 3, we obtain an isomorphism as follows:

$$A \simeq M_{r_1}(\mathbb{C}) \oplus \ldots \oplus M_{r_k}(\mathbb{C})$$

In addition to this, a careful look at the isomorphisms established in Step 3 shows that at the global level, of the algebra A itself, the above isomorphism simply comes by twisting the following standard multimatrix embedding, discussed in the beginning of the proof, (1) above, by a certain unitary matrix $U \in U_N$:

$$M_{r_1}(\mathbb{C}) \oplus \ldots \oplus M_{r_k}(\mathbb{C}) \subset M_N(\mathbb{C})$$

Now by putting everything together, we obtain the result. $\square$

Observe that, in terms of the basic spectral theory notions developed in chapter 8, the above result basically tells us that the finite dimensional C^*-algebras are exactly the direct sums of matrix algebras. We will use this several times, in what follows.

We can now formulate our first Peter-Weyl theorem, as follows:

THEOREM 13.10 (PW1). *Let $u : G \to U_N$ be a group representation, consider the algebra $A = End(u)$, and write its unit as above, as follows:*

$$1 = p_1 + \ldots + p_k$$

The representation u decomposes then as a direct sum, as follows,

$$u = u_1 + \ldots + u_k$$

with each u_i being an irreducible representation, obtained by restricting u to $Im(p_i)$.

PROOF. This basically follows from Theorem 13.8 and Theorem 13.9, as follows:

(1) As a first observation, by replacing G with its image $u(G) \subset U_N$, we can assume if we want that our representation u is faithful, $G \subset_u U_N$. However, this replacement will not be really needed, and we will keep using $u : G \to U_N$, as above.

(2) In order to prove the result, we will need some preliminaries. We first associate to our representation $u : G \to U_N$ the corresponding action map on $\mathbb{C}^N$. If a linear subspace $V \subset \mathbb{C}^N$ is invariant, the restriction of the action map to V is an action map too, which must come from a subrepresentation $v \subset u$. This is clear indeed from definitions, and with the remark that the unitaries, being isometries, restrict indeed into unitaries.

(3) Consider now a projection $p \in End(u)$. From $pu = up$ we obtain that the linear space $V = Im(p)$ is invariant under u, and so this space must come from a subrepresentation $v \subset u$. It is routine to check that the operation $p \to v$ maps subprojections to subrepresentations, and minimal projections to irreducible representations.

(4) To be more precise here, the condition $p \in End(u)$ reformulates as follows:

$$pu_g = u_g p \quad , \quad \forall g \in G$$

As for the condition that $V = Im(p)$ is invariant, this reformulates as follows:

$$pu_g p = u_g p \quad , \quad \forall g \in G$$

Thus, we are in need of a technical linear algebra result, stating that for a projection $P \in M_N(\mathbb{C})$ and a unitary $U \in U_N$, the following happens:

$$PUP = UP \implies PU = UP$$

(5) But this can be established with some C^*-algebra know-how, as follows:

$$\begin{aligned}
tr[(PU - UP)(PU - UP)^*] &= tr[(PU - UP)(U^*P - PU^*)] \\
&= tr[P - PUPU^* - UPU^*P + UPU^*] \\
&= tr[P - UPU^* - UPU^* + UPU^*] \\
&= tr[P - UPU^*] \\
&= 0
\end{aligned}$$

Indeed, by positivity this gives $PU - UP = 0$, as desired.

(6) With these preliminaries in hand, let us decompose the algebra $End(u)$ as in Theorem 13.9, by using the decomposition $1 = p_1 + \ldots + p_k$ into minimal projections. If we denote by $u_i \subset u$ the subrepresentation coming from the vector space $V_i = Im(p_i)$, then we obtain in this way a decomposition $u = u_1 + \ldots + u_k$, as in the statement. $\square$

In order to formulate our second Peter-Weyl theorem, we need to talk about coefficients, and smoothness. Things here are quite tricky, and best is to proceed as follows:

Definition 13.11. *Given a closed subgroup $G \subset U_N$, and a unitary representation $v : G \to U_M$, the space of coefficients of this representation is:*

$$C_v = \left\{ f \circ v \Big| f \in M_M(\mathbb{C})^* \right\}$$

In other words, by delinearizing, $C_\nu \subset C(G)$ is the following linear space:

$$C_v = span\Big[g \to (v_g)_{ij} \Big]$$

We say that v is smooth if its matrix coefficients $g \to (v_g)_{ij}$ appear as polynomials in the standard matrix coordinates $g \to g_{ij}$, and their conjugates $g \to \overline{g}_{ij}$.

As a basic example of coefficient we have, besides the matrix coefficients $g \to (v_g)_{ij}$, the character, which appears as the diagonal sum of these coefficients:

$$\chi_v(g) = \sum_i (v_g)_{ii}$$

Regarding the notion of smoothness, things are quite tricky here, the idea being that any closed subgroup $G \subset U_N$ can be shown to be a Lie group, and that, with this result in hand, a representation $v : G \to U_M$ is smooth precisely when the condition on coefficients from the above definition is satisfied. All this is quite technical, and we will not get into it. We will simply use Definition 13.11 as such, and further comment on this later on. Here is now our second Peter-Weyl theorem, complementing Theorem 13.10:

Theorem 13.12 (PW2). *Given a closed subgroup $G \subset_u U_N$, any of its irreducible smooth representations*

$$v : G \to U_M$$

appears inside a tensor product of the fundamental representation u and its adjoint $\bar{u}$.

Proof. In order to prove the result, we will use the following three elementary facts, regarding the spaces of coefficients introduced above:

(1) The construction $v \to C_v$ is functorial, in the sense that it maps subrepresentations into linear subspaces. This is indeed something which is routine to check.

(2) Our smoothness assumption on $v : G \to U_M$, as formulated in Definition 13.11, means that we have an inclusion of linear spaces as follows:

$$C_v \subset < g_{ij} >$$

(3) By definition of the Peter-Weyl representations, as arbitrary tensor products between the fundamental representation u and its conjugate $\bar{u}$, we have:

$$< g_{ij} >= \sum_k C_{u^{\otimes k}}$$

(4) Now by putting together the observations (2,3) we conclude that we must have an inclusion as follows, for certain exponents $k_1, \ldots, k_p$:

$$C_v \subset C_{u^{\otimes k_1} \oplus \ldots \oplus \pi^{\otimes k_p}}$$

By using now the functoriality result from (1), we deduce from this that we have an inclusion of representations, as follows:

$$v \subset u^{\otimes k_1} \oplus \ldots \oplus u^{\otimes k_p}$$

Together with Theorem 13.10, this leads to the conclusion in the statement. □

As a conclusion to what we have so far, the problem to be solved is that of splitting the Peter-Weyl representations into sums of irreducible representations.

13c. Haar integration

In order to further advance, and complete the Peter-Weyl theory, we need to talk about integration over G. In the finite group case the situation is trivial, as follows:

Proposition 13.13. *Any finite group G has a unique probability measure which is invariant under left and right translations,*

$$\mu(E) = \mu(gE) = \mu(Eg)$$

and this is the normalized counting measure on G, given by $\mu(E) = |E|/|G|$.

Proof. The uniformity condition in the statement gives, with $E = \{h\}$:

$$\mu\{h\} = \mu\{gh\} = \mu\{hg\}$$

Thus μ must be the usual counting measure, normalized as to have mass 1. □

In the continuous group case now, the simplest examples, to be studied first, are the compact abelian groups. Here things are standard again, as follows:

Theorem 13.14. *Given a compact abelian group G, with dual group denoted $\Gamma = \widehat{G}$, we have an isomorphism of commutative algebras*

$$C(G) \simeq C^*(\Gamma)$$

and via this isomorphism, the functional defined by linearity and the following formula,

$$\int_G g = \delta_{g1}$$

for any $g \in \Gamma$, is the integration with respect to the unique uniform measure on G.

Proof. This is something that we basically know, from chapters 8 and 9, coming as a consequence of the general results regarding the abelian groups and the commutative algebras developed there. To be more precise, and skipping some details here, the conclusions in the statement can be deduced as follows:

(1) We can either apply the Gelfand theorem, from chapter 8 above, to the group algebra $C^*(\Gamma)$, which is commutative, and this gives all the results.

(2) Or, we can use decomposition results for the compact abelian groups from chapter 9, and by reducing things to summands, once again we obtain the results. $\square$

Summarizing, we have results in the finite case, and in the compact abelian case. With the remark that the proof in the compact abelian case was quite brief, but this result, coming as an illustration for more general things to follow, is not crucial for us. Let us discuss now the construction of the uniform probability measure in general. This is something quite technical, the idea being that the uniform measure μ over G can be constructed by starting with an arbitrary probability measure ν, and setting:

$$\mu=\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^n\nu^{*k}$$

Thus, our next task will be that of proving this result. It is convenient, for this purpose, to work with the integration functionals with respect to the various measures on G, instead of the measures themselves. Let us begin with the following key result:

Proposition 13.15. *Given a unital positive linear form $\varphi:C(G)\to\mathbb C$, the limit*

$$\int_\varphi f=\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^n\varphi^{*k}(f)$$

exists, and for a coefficient of a representation $f=(\tau\otimes id)v$ we have

$$\int_\varphi f=\tau(P)$$

where P is the orthogonal projection onto the 1-eigenspace of $(id\otimes\varphi)v$.

Proof. By linearity it is enough to prove the first assertion for functions of the following type, where v is a Peter-Weyl representation, and τ is a linear form:

$$f=(\tau\otimes id)v$$

Thus we are led into the second assertion, and more precisely we can have the whole result proved if we can establish the following formula, with $f=(\tau\otimes id)v$:

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^n\varphi^{*k}(f)=\tau(P)$$

In order to prove this latter formula, observe that we have:

$$\varphi^{*k}(f) = (\tau \otimes \varphi^{*k})v = \tau((id \otimes \varphi^{*k})v)$$

Let us set $M = (id \otimes \varphi)v$. In terms of this matrix, we have:

$$((id \otimes \varphi^{*k})v)_{i_0 i_{k+1}} = \sum_{i_1 \ldots i_k} M_{i_0 i_1} \ldots M_{i_k i_{k+1}} = (M^k)_{i_0 i_{k+1}}$$

Thus we have the following formula, for any $k \in \mathbb{N}$:

$$(id \otimes \varphi^{*k})v = M^k$$

It follows that our Cesàro limit is given by:

$$\begin{aligned}
\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^n \varphi^{*k}(f) &= \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^n \tau(M^k) \\
&= \tau\left(\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^n M^k\right)
\end{aligned}$$

Now since v is unitary we have $||v|| = 1$, and so $||M|| \leq 1$. Thus the last Cesàro limit converges, and equals the orthogonal projection onto the 1-eigenspace of M:

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^n M^k = P$$

Thus our initial Cesàro limit converges as well, to $\tau(P)$, as desired. $\square$

The point now is that when the linear form $\varphi \in C(G)^*$ from the above result is chosen to be faithful, we obtain the following finer result:

Proposition 13.16. *Given a faithful unital linear form $\varphi \in C(G)^*$, the limit*

$$\int_\varphi f = \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^n \varphi^{*k}(f)$$

exists, and is independent of φ, given on coefficients of representations by

$$\left(id \otimes \int_\varphi\right) v = P$$

where P is the orthogonal projection onto the following space:

$$Fix(v) = \left\{\xi \in \mathbb{C}^n \middle| v\xi = \xi\right\}$$

Proof. In view of Proposition 13.15, it remains to prove that when φ is faithful, the 1-eigenspace of the matrix $M = (id \otimes \varphi)v$ equals the space $Fix(v)$.

"$\supset$" This is clear, and for any φ, because we have:

$$v\xi = \xi \implies M\xi = \xi$$

"$\subset$" Here we must prove that, when φ is faithful, we have:

$$M\xi = \xi \implies v\xi = \xi$$

For this purpose, assume that we have $M\xi = \xi$, and consider the following function:

$$f = \sum_i \left(\sum_j v_{ij}\xi_j - \xi_i \right) \left(\sum_k v_{ik}\xi_k - \xi_i \right)^*$$

We must prove that we have $f = 0$. Since v is unitary, we have:

$$\begin{aligned} f &= \sum_{ijk} v_{ij}v_{ik}^*\xi_j\bar{\xi}_k - \frac{1}{N}v_{ij}\xi_j\bar{\xi}_i - \frac{1}{N}v_{ik}^*\xi_i\bar{\xi}_k + \frac{1}{N^2}\xi_i\bar{\xi}_i \\ &= \sum_j |\xi_j|^2 - \sum_{ij} v_{ij}\xi_j\bar{\xi}_i - \sum_{ik} v_{ik}^*\xi_i\bar{\xi}_k + \sum_i |\xi_i|^2 \\ &= ||\xi||^2 - < v\xi, \xi > - \overline{< v\xi, \xi >} + ||\xi||^2 \\ &= 2(||\xi||^2 - Re(< v\xi, \xi >)) \end{aligned}$$

By using now our assumption $M\xi = \xi$, we obtain from this:

$$\begin{aligned} \varphi(f) &= 2\varphi(||\xi||^2 - Re(< v\xi, \xi >)) \\ &= 2(||\xi||^2 - Re(< M\xi, \xi >)) \\ &= 2(||\xi||^2 - ||\xi||^2) \\ &= 0 \end{aligned}$$

Now since φ is faithful, this gives $f = 0$, and so $v\xi = \xi$, as claimed. $\square$

We can now formulate a main result, as follows:

THEOREM 13.17. *Any compact group G has a unique Haar integration, which can be constructed by starting with any faithful positive unital state $\varphi \in C(G)^*$, and setting:*

$$\int_G = \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^n \varphi^{*k}$$

Moreover, for any representation v we have the formula

$$\left(id \otimes \int_G \right) v = P$$

where P is the orthogonal projection onto $Fix(v) = \left\{ \xi \in \mathbb{C}^n \middle| v\xi = \xi \right\}$.

PROOF. We can prove this from what we have, in several steps, as follows:

(1) Let us first go back to the general context of Proposition 13.15. Since convolving one more time with φ will not change the Cesàro limit appearing there, the functional $\int_\varphi \in C(G)^*$ constructed there has the following invariance property:

$$\int_\varphi * \varphi = \varphi * \int_\varphi = \int_\varphi$$

In the case where φ is assumed to be faithful, as in Proposition 13.16, our claim is that we have the following formula, valid this time for any $\psi \in C(G)^*$:

$$\int_\varphi * \psi = \psi * \int_\varphi = \psi(1) \int_\varphi$$

It is enough to prove this formula on a coefficient of a representation:

$$f = (\tau \otimes id)v$$

In order to do so, consider the following two matrices:

$$P = \left(id \otimes \int_\varphi\right) v \quad , \quad Q = (id \otimes \psi)v$$

We have then the following computation:

$$\left(\int_\varphi * \psi\right) f = \left(\tau \otimes \int_\varphi \otimes \psi\right) (v_{12}v_{13}) = \tau(PQ)$$

Similarly, we have the following computation:

$$\left(\psi * \int_\varphi\right) f = \left(\tau \otimes \psi \otimes \int_\varphi\right) (v_{12}v_{13}) = \tau(QP)$$

Finally, regarding the term on the right, this is given by:

$$\psi(1) \int_\varphi f = \psi(1)\tau(P)$$

Thus, our claim is equivalent to the following equality:

$$PQ = QP = \psi(1)P$$

But this follows from the fact, coming from Proposition 13.16, that $P = (id \otimes \int_\varphi)v$ equals the orthogonal projection onto $Fix(v)$. Thus, we have proved our claim.

(2) In order to finish now, it is convenient to introduce the following abstract operation, on the continuous functions $f, f' : C(G) \to \mathbb{C}$ on our group:

$$\Delta(f \otimes f')(g \otimes h) = f(g)f'(h)$$

With this convention, the formula that we established above can be written as:

$$\psi\left(\int_\varphi\otimes id\right)\Delta=\psi\left(id\otimes\int_\varphi\right)\Delta=\psi\int_\varphi(.)1$$

This formula being true for any $\psi\in C(G)^*$, we can simply delete ψ. We conclude that the following invariance formula holds indeed, with $\int_G=\int_\varphi$:

$$\left(\int_G\otimes id\right)\Delta=\left(id\otimes\int_G\right)\Delta=\int_G(.)1$$

But this is exactly the left and right invariance formula we were looking for.

(3) Finally, in order to prove the uniqueness assertion, assuming that we have two invariant integrals $\int_G,\int_G'$, we have, according to the above invariance formula:

$$\left(\int_G\otimes\int_G'\right)\Delta=\left(\int_G'\otimes\int_G\right)\Delta=\int_G(.)1=\int_G'(.)1$$

Thus we have $\int_G=\int_G'$, and this finishes the proof. □

Summarizing, we can now integrate over G. As a first application, we have:

THEOREM 13.18. *Given a compact group G, we have the following formula, valid for any unitary group representation $v:G\to U_M$:*

$$\int_G\chi_v=\dim(Fix(v))$$

In particular, in the unitary matrix group case, $G\subset_u U_N$, the moments of the main character $\chi=\chi_u$ are given by the following formula:

$$\int_G\chi^k=\dim(Fix(u^{\otimes k}))$$

Thus, knowing the law of χ is the same as knowing the dimensions on the right.

PROOF. We have three assertions here, the idea being as follows:

(1) Given a unitary representation $v:G\to U_M$ as in the statement, its character χ_v is a coefficient, so we can use the integration formula for coefficients in Theorem 13.17. If we denote by P the projection onto $Fix(v)$, this formula gives, as desired:

$$\begin{aligned}\int_G\chi_v&=Tr(P)\\&=\dim(Im(P))\\&=dim(Fix(v))\end{aligned}$$

(2) This follows from (1), applied to the Peter-Weyl representations, as follows:

$$\begin{aligned}\int_G \chi^k &= \int_G \chi_u^k \\ &= \int_G \chi_{u^{\otimes k}} \\ &= \dim(Fix(u^{\otimes k}))\end{aligned}$$

(3) This follows from (2), and from the standard fact, which follows from definitions, that a probability measure is uniquely determined by its moments. □

As a key remark now, the integration formula in Theorem 13.17 allows the computation for the truncated characters too, because these truncated characters are coefficients as well. To be more precise, all the probabilistic questions about G, regarding characters, or truncated characters, or more complicated variables, require a good knowledge of the integration over G, and more precisely, of the various polynomial integrals over G:

Definition 13.19. *Given a closed subgroup $G \subset U_N$, the quantities*

$$I_k = \int_G g_{i_1 j_1}^{e_1} \ldots g_{i_k j_k}^{e_k} \, dg$$

depending on a colored integer $k = e_1 \ldots e_k$, are called polynomial integrals over G.

As a first observation, the knowledge of these integrals is the same as the knowledge of the integration functional over G. Indeed, since the coordinate functions $g \to g_{ij}$ separate the points of G, we can apply the Stone-Weierstrass theorem, and we obtain:

$$C(G) = < g_{ij} >$$

Thus, by linearity, the computation of any functional $f : C(G) \to \mathbb{C}$, and in particular of the integration functional, reduces to the computation of this functional on the polynomials of the coordinate functions $g \to g_{ij}$ and their conjugates $g \to \bar{g}_{ij}$.

By using now Peter-Weyl theory, everything reduces to algebra, as follows:

Theorem 13.20. *The Haar integration over a closed subgroup $G \subset_u U_N$ is given on the dense subalgebra of smooth functions by the Weingarten formula*

$$\int_G g_{i_1 j_1}^{e_1} \ldots g_{i_k j_k}^{e_k} \, dg = \sum_{\pi, \sigma \in D_k} \delta_\pi(i) \delta_\sigma(j) W_k(\pi, \sigma)$$

valid for any colored integer $k = e_1 \ldots e_k$ and any multi-indices i, j, where D_k is a linear basis of $Fix(u^{\otimes k})$, the associated generalized Kronecker symbols are given by

$$\delta_\pi(i) = < \pi, e_{i_1} \otimes \ldots \otimes e_{i_k} >$$

and $W_k = G_k^{-1}$ is the inverse of the Gram matrix, $G_k(\pi, \sigma) = < \pi, \sigma >$.

PROOF. We know from Peter-Weyl theory that the integrals in the statement form altogether the orthogonal projection P^k onto the following space:

$$Fix(u^{\otimes k}) = span(D_k)$$

Consider now the following linear map, with $D_k = \{\xi_k\}$ being as in the statement:

$$E(x) = \sum_{\pi \in D_k} < x, \xi_\pi > \xi_\pi$$

By a standard linear algebra computation, it follows that we have $P = WE$, where W is the inverse of the restriction of E to the following space:

$$K = span\left(T_\pi \Big| \pi \in D_k\right)$$

But this restriction is precisely the linear map given by the matrix G_k, and so W itself is the linear map given by the matrix W_k, and this gives the result. □

We will be back to this in chapter 16 below, with some concrete applications.

13d. More Peter-Weyl

In order to further develop now the Peter-Weyl theory, which is something very useful, we will need the following result, which is of independent interest:

PROPOSITION 13.21. *We have a Frobenius type isomorphism*

$$Hom(v, w) \simeq Fix(v \otimes \bar{w})$$

valid for any two representations v, w.

PROOF. According to the definitions, we have the following equivalences:

$$\begin{aligned} T \in Hom(v, w) &\iff Tv = wT \\ &\iff \sum_j T_{aj} v_{ji} = \sum_b w_{ab} T_{bi}, \forall a, i \end{aligned}$$

On the other hand, we have as well the following equivalences:

$$\begin{aligned} T \in Fix(v \otimes \bar{w}) &\iff (v \otimes \bar{w})T = \xi \\ &\iff \sum_{jb} v_{ij} w_{ab}^* T_{bj} = T_{ai} \forall a, i \end{aligned}$$

With these formulae in hand, both inclusions follow from the unitarity of v, w. □

We can now formulate our third Peter-Weyl theorem, as follows:

THEOREM 13.22 (PW3). *The norm dense $*$-subalgebra*

$$\mathcal{C}(G) \subset C(G)$$

generated by the coefficients of the fundamental representation decomposes as a direct sum

$$\mathcal{C}(G) = \bigoplus_{v \in Irr(G)} M_{\dim(v)}(\mathbb{C})$$

with the summands being pairwise orthogonal with respect to the scalar product

$$< a, b >= \int_G ab^*$$

where $\int_G$ is the Haar integration over G.

PROOF. By combining the previous two Peter-Weyl results, we deduce that we have a linear space decomposition as follows:

$$\mathcal{C}(G) = \sum_{v \in Irr(G)} C_v = \sum_{v \in Irr(G)} M_{\dim(v)}(\mathbb{C})$$

Thus, in order to conclude, it is enough to prove that for any two irreducible corepresentations $v, w \in Irr(A)$, the corresponding spaces of coefficients are orthogonal:

$$v \not\sim w \implies C_v \perp C_w$$

But this follows from Theorem 13.17, via Proposition 13.21. Let us set indeed:

$$P_{ia,jb} = \int_G v_{ij} w_{ab}^*$$

Then P is the orthogonal projection onto the following vector space:

$$Fix(v \otimes \bar{w}) \simeq Hom(v, w) = \{0\}$$

Thus we have $P = 0$, and this gives the result. □

Finally, we have the following result, completing the Peter-Weyl theory:

THEOREM 13.23 (PW4). *The characters of irreducible representations belong to*

$$\mathcal{C}(G)_{central} = \left\{ f \in \mathcal{C}(G) \Big| f(gh) = f(hg), \forall g, h \in G \right\}$$

called algebra of smooth central functions on G, and form an orthonormal basis of it.

PROOF. We have several things to be proved, the idea being as follows:

(1) Observe first that $\mathcal{C}(G)_{central}$ is indeed an algebra, which contains all the characters. Conversely, consider a function $f \in \mathcal{C}(G)$, written as follows:

$$f = \sum_{v \in Irr(G)} f_v$$

The condition $f \in \mathcal{C}(G)_{central}$ states then that for any $v \in Irr(G)$, we must have:

$$f_v \in \mathcal{C}(G)_{central}$$

But this means precisely that the coefficient f_v must be a scalar multiple of χ_v, and so the characters form a basis of $\mathcal{C}(G)_{central}$, as stated.

(2) The fact that we have an orthogonal basis follows from Theorem 13.22.

(3) As for the fact that the characters have norm 1, this follows from:

$$\begin{aligned}\int_G \chi_v \chi_v^* &= \sum_{ij} \int_G v_{ii} v_{jj}^* \\ &= \sum_i \frac{1}{N} \\ &= 1\end{aligned}$$

Here we have used the fact, coming from Theorem 13.22, that the integrals $\int_G v_{ij} v_{kl}^*$ form the orthogonal projection onto the following vector space:

$$Fix(v \otimes \bar{v}) \simeq End(v) = \mathbb{C}1$$

Thus, the proof of our theorem is now complete. □

As a key observation here, complementing Theorem 13.23, observe that a function $f : G \to \mathbb{C}$ is central, in the sense that it satisfies $f(gh) = f(hg)$, precisely when it satisfies the following condition, saying that it must be constant on conjugacy classes:

$$f(ghg^{-1}) = f(h), \forall g, h \in G$$

Thus, in the finite group case for instance, the algebra of central functions is something which is very easy to compute, and this gives useful information about $Rep(G)$. We will not get into this here, but some of our exercises will be about this.

So long for Peter-Weyl theory. As a basic illustration for all this, which clarifies some previous considerations from chapter 9, we have the following result:

THEOREM 13.24. *For a compact abelian group G the irreducible representations are all 1-dimensional, and form the dual discrete abelian group $\widehat{G}$.*

PROOF. This is clear from the Peter-Weyl theory, because when G is abelian any function $f : G \to \mathbb{C}$ is central, and so the algebra of central functions is $\mathcal{C}(G)$ itself, and so the irreducible representations $u \in Irr(G)$ coincide with their characters $\chi_u \in \widehat{G}$. □

There are also many things that can be said in the finite group case, in relation with central functions, and conjugacy classes. For more here, we recommend Serre [79].

13e. Exercises

There has been a lot of theory on this chapter, which was often quite abstract, and as basic exercise on the subject, which is quite elementary, we have:

EXERCISE 13.25. *Work out the details for the Peter-Weyl decomposition formula*

$$C(G)=\bigoplus_{v\in Irr(G)}M_{\dim(v)}(\mathbb{C})$$

in an explicit way, for some finite non-abelian groups, of your choice.

There are of course many possible choices for G. As a bonus exercise here, you can try do the same for the subalgebra of central functions $C(G)_{central}$ too.

EXERCISE 13.26. *Given a finite group G, setting $A=C(G)$, prove that the maps*

$$\Delta:A\to A\otimes A$$

$$\varepsilon:A\to\mathbb{C}$$

$$S:A\to A$$

which are transpose to the multiplication $m:G\times G\to G$, unit $u:\{.\}\to G$ and inverse map $i:G\to G$, are subject to the following conditions

$$(\varepsilon\otimes id)\Delta=(id\otimes\varepsilon)\Delta=id$$

$$m(S\otimes id)\Delta=m(id\otimes S)\Delta=\varepsilon(.)1$$

in usual tensor product notation, along with the extra condition $S^2=id$.

This does not look difficult, with the conditions in the statement reminding the usual group axioms, satisfied by m,u,i. Up to you to prove this now, with full details.

EXERCISE 13.27. *Given a finite group H, setting $A=C^*(H)$, prove that the maps*

$$\Delta:A\to A\otimes A$$

$$\varepsilon:A\to\mathbb{C}$$

$$S:A\to A^{opp}$$

given by the formulae $\Delta(g)=g\otimes g$, $\varepsilon(g)=1$, $S(g)=g^{-1}$ and linearity, are subject to the same conditions as above, namely

$$(\varepsilon\otimes id)\Delta=(id\otimes\varepsilon)\Delta=id$$

$$m(S\otimes id)\Delta=m(id\otimes S)\Delta=\varepsilon(.)1$$

in usual tensor product notation, along with the extra condition $S^2=id$.

As before with the previous exercise, this does not look very difficult, with most likely only some elementary algebraic computations involved.

EXERCISE 13.28. *Let us call finite Hopf algebra a finite dimensional C^*-algebra, with maps as follows, called comultiplication, counit and antipode,*

$$\Delta : A \to A \otimes A \quad , \quad \varepsilon : A \to \mathbb{C} \quad , \quad S : A \to A^{opp}$$

satisfying the following conditions, which are those found above,

$$(\varepsilon \otimes id)\Delta = (id \otimes \varepsilon)\Delta = id$$

$$m(S \otimes id)\Delta = m(id \otimes S)\Delta = \varepsilon(.)1$$

along with $S^2 = id$. Prove that if G, H are finite abelian groups, dual to each other via Pontrjagin duality, then we have an identification of Hopf algebras as follows,

$$C(G) = C^*(H)$$

and based on this, go ahead and formally write any finite Hopf algebra as

$$A = C(G) = C^*(H)$$

and call G, H finite quantum groups, dual to each other.

Here the thing to be done, namely to establish the identification in the statement, looks like something quite routine, related to many things that we already know. As for the last part, there is nothing to be done here, just enjoying that definition.

EXERCISE 13.29. *Given a finite dimensional Hopf algebra A, prove that its dual A^* is a Hopf algebra too, with structural maps as follows:*

$$\Delta^t : A^* \otimes A^* \to A^*$$

$$\varepsilon^t : \mathbb{C} \to A^*$$

$$m^t : A^* \to A^* \otimes A^*$$

$$u^t : A^* \to \mathbb{C}$$

$$S^t : A^* \to A^*$$

Also, check that A is commutative if and only if A^* is cocommutative, and also discuss what happens in the cases $A = C(G)$ and $A = C^*(H)$, with G, H being finite groups.

This exercise is actually to be best worked out at the same time with the previous one. In this way you'll fully learn what the finite quantum groups are.

EXERCISE 13.30. *Develop a theory of compact and discrete quantum groups, generalizing at the same time the theory of usual compact and discrete groups, and the Pontrjagin duality for them, in the abelian case, and the theory of finite quantum groups, and the abstract duality for them, developed in the above series of exercises.*

This latter exercise is probably quite difficult, better look it up.

CHAPTER 14

Tannakian duality

14a. Tensor categories

We have seen that the representations of a closed subgroup $G \subset U_N$ are subject to a number of non-trivial results, collectively known as Peter-Weyl theory. To be more precise, the main ideas of Peter-Weyl theory were as follows:

(1) The representations of G split as sums of irreducibles, and the irreducibles can be found inside the tensor products $u^{\otimes k}$ between the fundamental representation $u : G \subset U_N$ and its adjoint $\bar{u} : G \subset U_N$, called Peter-Weyl representations.

(2) The main problem is therefore that of splitting the Peter-Weyl representations $u^{\otimes k}$ into irreducibles. Technically speaking, this leads to the question of explicitely computing the corresponding fixed point spaces $Fix(u^{\otimes k})$.

(3) From a probabilistic perspective, in connection with characters and truncated characters, which require the explicit knowledge of $\int_G$, we are led into the same fundamental question, namely the computation of the spaces $Fix(u^{\otimes k})$.

Summarizing, no matter what we want to do with G, we must compute the spaces $Fix(u^{\otimes k})$. As a first idea now, it is technically convenient to slightly enlarge the class of spaces to be computed, by talking about Tannakian categories, as follows:

Definition 14.1. *The Tannakian category associated to a closed subgroup $G \subset_u U_N$ is the collection $C = (C(k,l))$ of vector spaces*

$$C(k,l) = Hom(u^{\otimes k}, u^{\otimes l})$$

where the representations $u^{\otimes k}$ with $k = \circ \bullet \bullet \circ \ldots$ colored integer, defined by

$$u^{\otimes \emptyset} = 1 \quad , \quad u^{\otimes \circ} = u \quad , \quad u^{\otimes \bullet} = \bar{u}$$

and multiplicativity, $u^{\otimes kl} = u^{\otimes k} \otimes u^{\otimes l}$, are the Peter-Weyl representations.

Here are a few examples of such representations, namely those coming from the colored integers of length 2, to be often used in what follows:

$$u^{\otimes \circ\circ} = u \otimes u \quad , \quad u^{\otimes \circ\bullet} = u \otimes \bar{u}$$

$$u^{\otimes \bullet\circ} = \bar{u} \otimes u \quad , \quad u^{\otimes \bullet\bullet} = \bar{u} \otimes \bar{u}$$

As a first observation, the knowledge of the Tannakian category is more or less the same thing as the knowledge of the fixed point spaces, which appear as:

$$Fix(u^{\otimes k}) = C(0,k)$$

Indeed, these latter spaces fully determine all the spaces $C(k,l)$, because of the Frobenius isomorphisms, which for the Peter-Weyl representations read:

$$\begin{aligned} C(k,l) &= Hom(u^{\otimes k}, u^{\otimes l}) \\ &\simeq Hom(1, \bar{u}^{\otimes k} \otimes u^{\otimes l}) \\ &= Hom(1, u^{\otimes \bar{k}l}) \\ &= Fix(u^{\otimes \bar{k}l}) \end{aligned}$$

In order to get started now, let us make a summary of what we have so far, regarding these spaces $C(k,l)$, coming from the general theory developed in chapter 13. In order to formulate our result, let us start with an abstract definition, as follows:

Definition 14.2. *Let H be a finite dimensional Hilbert space. A tensor category over H is a collection $C = (C(k,l))$ of linear spaces*

$$C(k,l) \subset \mathcal{L}(H^{\otimes k}, H^{\otimes l})$$

satisfying the following conditions:

1. *$S, T \in C$ implies $S \otimes T \in C$.*
2. *If $S, T \in C$ are composable, then $ST \in C$.*
3. *$T \in C$ implies $T^* \in C$.*
4. *Each $C(k,k)$ contains the identity operator.*
5. *$C(\emptyset, k)$ with $k = \circ\bullet, \bullet\circ$ contain the operator $R : 1 \to \sum_i e_i \otimes e_i$.*
6. *$C(kl, lk)$ with $k, l = \circ, \bullet$ contain the flip operator $\Sigma : a \otimes b \to b \otimes a$.*

Here the tensor powers $H^{\otimes k}$, which are Hilbert spaces depending on a colored integer $k = \circ\bullet\bullet\circ\ldots$, are defined by the following formulae, and multiplicativity:

$$H^{\otimes \emptyset} = \mathbb{C} \quad , \quad H^{\otimes \circ} = H \quad , \quad H^{\otimes \bullet} = \bar{H} \simeq H$$

With these conventions, we have the following result, summarizing our knowledge on the subject, coming from the results from the previous chapter:

Theorem 14.3. *For a closed subgroup $G \subset_u U_N$, the associated Tannakian category*

$$C(k,l) = Hom(u^{\otimes k}, u^{\otimes l})$$

is a tensor category over the Hilbert space $H = \mathbb{C}^N$.

Proof. We know that the fundamental representation u acts on the Hilbert space $H = \mathbb{C}^N$, and that its conjugate $\bar{u}$ acts on the Hilbert space $\bar{H} = \mathbb{C}^N$. Now by multiplicativity we conclude that any Peter-Weyl representation $u^{\otimes k}$ acts on the Hilbert space

$H^{\otimes k}$, so that we have embeddings as in Definition 14.2, as follows:

$$C(k,l)\subset\mathcal{L}(H^{\otimes k},H^{\otimes l})$$

Regarding now the fact that the axioms (1-6) in Definition 14.2 are indeed satisfied, this is something that we basically already know, as follows:

(1,2,3) These results follow from definitions, and were explained in chapter 13.

(4) This is something trivial, coming from definitions.

(5) This follows from the fact that each element $g\in G$ is a unitary, which can be reformulated as follows, with $R:1\to\sum_i e_i\otimes e_i$ being the map in Definition 14.2:

$$R\in Hom(1,g\otimes\bar{g})\quad,\quad R\in Hom(1,\bar{g}\otimes g)$$

Indeed, given an arbitrary matrix $g\in M_N(\mathbb{C})$, we have the following computation:

$$\begin{aligned}
(g\otimes\bar{g})(R(1)\otimes 1)
&=\left(\sum_{ijkl}e_{ij}\otimes e_{kl}\otimes g_{ij}\bar{g}_{kl}\right)\left(\sum_a e_a\otimes e_a\otimes 1\right)\\
&=\sum_{ika}e_i\otimes e_k\otimes g_{ia}\bar{g}^*_{ka}\\
&=\sum_{ik}e_i\otimes e_k\otimes (gg^*)_{ik}
\end{aligned}$$

We conclude from this that we have the following equivalence:

$$R\in Hom(1,g\otimes\bar{g})\iff gg^*=1$$

By replacing g with its conjugate matrix $\bar{g}$, we have as well:

$$R\in Hom(1,\bar{g}\otimes g)\iff \bar{g}g^t=1$$

Thus, the two intertwining conditions in Definition 14.2 (5) are both equivalent to the fact that g is unitary, and so these conditions are indeed satisfied, as desired.

(6) This is again something elementary, coming from the fact that the various matrix coefficients $g\to g_{ij}$ and their complex conjugates $g\to\bar{g}_{ij}$ commute with each other. To be more precise, with $\Sigma:a\otimes b\to b\otimes a$ being the flip operator, we have:

$$\begin{aligned}
(g\otimes h)(\Sigma\otimes id)(e_a\otimes e_b\otimes 1)
&=\left(\sum_{ijkl}e_{ij}\otimes e_{kl}\otimes g_{ij}h_{kl}\right)(e_b\otimes e_a\otimes 1)\\
&=\sum_{ik}e_i\otimes e_k\otimes g_{ib}h_{ka}
\end{aligned}$$

On the other hand, we have as well the following computation:

$$\begin{aligned}(\Sigma\otimes id)(h\otimes g)(e_a\otimes e_b\otimes 1) &= (\Sigma\otimes id)\left(\sum_{ijkl}e_{ij}\otimes e_{kl}\otimes h_{ij}g_{kl}\right)(e_a\otimes e_b\otimes 1)\\ &= (\Sigma\otimes id)\left(\sum_{ik}e_i\otimes e_k\otimes h_{ia}g_{kb}\right)\\ &= \sum_{ik}e_k\otimes e_i\otimes h_{ia}g_{kb}\\ &= \sum_{ik}e_i\otimes e_k\otimes h_{ka}g_{ib}\end{aligned}$$

Now since functions commute, $g_{ib}h_{ka}=h_{ka}g_{ib}$, this gives the result. □

With the above in hand, our purpose now will be that of showing that any closed subgroup $G\subset U_N$ is uniquely determined by its Tannakian category $C=(C(k,l))$:

$$G\leftrightarrow C$$

This result, known as Tannakian duality, is something quite deep, and very useful. Indeed, the idea is that what we would have here is a "linearization" of G, allowing us to do combinatorics, and ultimately reach to very concrete and powerful results, regarding G itself. And as a consequence, solve our probability questions left.

Speaking linearization, there is also a comment to be made here, in relation with Lie algebras. Remember a discussion that we had long time ago, in the beginning of chapter 10, with me arguing that yes, a good idea for the study of the closed subgroups $G\subset U_N$ would be the consideration of the tangent space at the origin $\mathfrak{g}=T_1(G)$, called Lie algebra of G, which is a very good "linearization" of G, but no need to get head-first into that, because there are perhaps some other methods for linearizing G? Well, time now to justify this claim, with some general theory of the correspondence $G\leftrightarrow C$.

Getting started now, we want to construct a correspondence $G\leftrightarrow C$, and we already know from Theorem 14.3 how the correspondence $G\to C$ appears, namely via:

$$C(k,l)=Hom(u^{\otimes k},u^{\otimes l})$$

Regarding now the construction in the other sense, $C\to G$, this is something very simple as well, coming from the following elementary result:

Theorem 14.4. *Given a tensor category $C=(C(k,l))$ over the space $H\simeq\mathbb{C}^N$,*

$$G=\left\{g\in U_N\Big|Tg^{\otimes k}=g^{\otimes l}T\ ,\ \forall k,l,\forall T\in C(k,l)\right\}$$

is a closed subgroup $G\subset U_N$.

PROOF. Consider indeed the closed subset $G \subset U_N$ constructed in the statement. We want to prove that G is indeed a group, and the verifications here go as follows:

(1) Given two matrices $g, h \in G$, their product satisfies $gh \in G$, due to the following computation, valid for any k, l and any $T \in C(k, l)$:

$$\begin{aligned} T(gh)^{\otimes k} &= Tg^{\otimes k}h^{\otimes k} \\ &= g^{\otimes l}Th^{\otimes k} \\ &= g^{\otimes l}h^{\otimes l}T \\ &= (gh)^{\otimes l}T \end{aligned}$$

(2) Also, we have $1 \in G$, trivially. Finally, for $g \in G$ and $T \in C(k, l)$, we have:

$$\begin{aligned} T(g^{-1})^{\otimes k} &= (g^{-1})^{\otimes l}[g^{\otimes l}T](g^{-1})^{\otimes k} \\ &= (g^{-1})^{\otimes l}[Tg^{\otimes k}](g^{-1})^{\otimes k} \\ &= (g^{-1})^{\otimes l}T \end{aligned}$$

Thus we have $g^{-1} \in G$, and so G is a group, as claimed. $\square$

Summarizing, we have so far precise axioms for the tensor categories $C = (C(k, l))$, given in Definition 14.2, as well as correspondences as follows:

$$G \to C \quad , \quad C \to G$$

We will show in what follows that these correspondences are inverse to each other. In order to get started, we first have the following technical result:

THEOREM 14.5. *If we denote the correspondences in Theorem 14.3 and 14.4, between closed subgroups $G \subset U_N$ and tensor categories $C = (C(k, l))$ over $H = \mathbb{C}^N$, as*

$$G \to C_G \quad , \quad C \to G_C$$

then we have embeddings as follows, for any G and C respectively,

$$G \subset G_{C_G} \quad , \quad C \subset C_{G_C}$$

and proving that these correspondences are inverse to each other amounts in proving

$$C_{G_C} \subset C$$

for any tensor category $C = (C(k, l))$ over the space $H = \mathbb{C}^N$.

PROOF. This is something trivial, with the embeddings $G \subset G_{C_G}$ and $C \subset C_{G_C}$ being both clear from definitions, and with the last assertion coming from this. $\square$

In order to establish Tannakian duality, we will need some abstract constructions. Following Malacarne [68], let us start with the following elementary fact:

Proposition 14.6. *Given a tensor category $C=C((k,l))$ over a Hilbert space H,*

$$E_C^{(s)}=\bigoplus_{|k|,|l|\leq s}C(k,l)\subset\bigoplus_{|k|,|l|\leq s}B(H^{\otimes k},H^{\otimes l})=B\left(\bigoplus_{|k|\leq s}H^{\otimes k}\right)$$

is a finite dimensional $*$-subalgebra. Also,

$$E_C=\bigoplus_{k,l}C(k,l)\subset\bigoplus_{k,l}B(H^{\otimes k},H^{\otimes l})\subset B\left(\bigoplus_{k}H^{\otimes k}\right)$$

is a closed $*$-subalgebra.

Proof. This is clear indeed from the categorical axioms from Definition 14.2, which, since satisfied, prove that the various linear spaces in the statement are stable under both the multiplication operation, and under taking the adjoints. □

Now back to our reconstruction question, we want to prove $C=C_{G_C}$, which is the same as proving $E_C=E_{C_{G_C}}$. We will use a standard commutant trick, as follows:

Theorem 14.7. *For any $*$-algebra $A\subset M_n(\mathbb{C})$ we have the equality*

$$A=A''$$

where prime denotes the commutant, $X'=\left\{T\in M_n(\mathbb{C})\Big|Tx=xT,\forall x\in X\right\}$.

Proof. This is a particular case of von Neumann's bicommutant theorem, which follows from the explicit description of A given in chapter 13, namely:

$$A=M_{r_1}(\mathbb{C})\oplus\ldots\oplus M_{r_k}(\mathbb{C})$$

Indeed, the center of each matrix algebra being reduced to the scalars, the commutant of this algebra is as follows, with each copy of $\mathbb{C}$ corresponding to a matrix block:

$$A'=\mathbb{C}\oplus\ldots\oplus\mathbb{C}$$

Now when taking once again the commutant, things are trivial, and we obtain in this way A itself, and this leads to the conclusion in the statement. □

By using now the bicommutant theorem, we have:

Proposition 14.8. *Given a Tannakian category C, the following are equivalent:*

(1) $C=C_{G_C}$.
(2) $E_C=E_{C_{G_C}}$.
(3) $E_C^{(s)}=E_{C_{G_C}}^{(s)}$, *for any $s\in\mathbb{N}$.*
(4) $E_C^{(s)'}=E_{C_{G_C}}^{(s)'}$, *for any $s\in\mathbb{N}$.*

In addition, the inclusions $\subset$, $\subset$, $\subset$, $\supset$ are automatically satisfied.

PROOF. This follows from the above results, as follows:

(1) $\iff$ (2) This is clear from definitions.

(2) $\iff$ (3) This is clear from definitions as well.

(3) $\iff$ (4) This comes from the bicommutant theorem. As for the last assertion, we have indeed $C \subset C_{G_C}$ from Theorem 14.5, and this shows that we have as well:

$$E_C \subset E_{C_{G_C}}$$

We therefore obtain by truncating $E_C^{(s)} \subset E_{C_{G_C}}^{(s)}$, and by taking the commutants, this gives $E_C^{(s)} \supset E_{C_{G_C}}^{(s)}$. Thus, we are led to the conclusion in the statement. □

Summarizing, we would like to prove that we have $E_C^{(s)'} \subset E_{C_{G_C}}^{(s)'}$. Let us first study the commutant on the right. As a first observation, we have:

PROPOSITION 14.9. *We have the following equality,*

$$E_{C_G}^{(s)} = End\left(\bigoplus_{|k|\leq s} u^{\otimes k}\right)$$

between subalgebras of $B\left(\bigoplus_{|k|\leq s} H^{\otimes k}\right)$.

PROOF. We know that the category C_G is by definition given by:

$$C_G(k,l) = Hom(u^{\otimes k}, u^{\otimes l})$$

Thus, the corresponding algebra $E_{C_G}^{(s)}$ appears as follows:

$$E_{C_G}^{(s)} = \bigoplus_{|k|,|l|\leq s} Hom(u^{\otimes k}, u^{\otimes l}) \subset \bigoplus_{|k|,|l|\leq s} B(H^{\otimes k}, H^{\otimes l}) = B\left(\bigoplus_{|k|\leq s} H^{\otimes k}\right)$$

On the other hand, the algebra of intertwiners of $\bigoplus_{|k|\leq s} u^{\otimes k}$ is given by:

$$End\left(\bigoplus_{|k|\leq s} u^{\otimes k}\right) = \bigoplus_{|k|,|l|\leq s} Hom(u^{\otimes k}, u^{\otimes l}) \subset \bigoplus_{|k|,|l|\leq s} B(H^{\otimes k}, H^{\otimes l}) = B\left(\bigoplus_{|k|\leq s} H^{\otimes k}\right)$$

Thus we have indeed the same algebra, and we are done. □

We have to compute the commutant of the above algebra. For this purpose, we can use the following general result, valid for any representation of a compact group:

Proposition 14.10. *Given a unitary group representation $v : G \to U_n$ we have an algebra representation as follows,*

$$\pi_v : C(G)^* \to M_n(\mathbb{C}) \quad , \quad \varphi \to (\varphi(v_{ij}))_{ij}$$

whose image is given by $Im(\pi_v) = End(v)'$.

Proof. The first assertion is clear, with the multiplicativity claim for π_v coming from the following computation, where $\Delta : C(G) \to C(G) \otimes C(G)$ is the comultiplication:

$$\begin{aligned}
(\pi_v(\varphi * \psi))_{ij} &= (\varphi \otimes \psi)\Delta(v_{ij}) \\
&= \sum_k \varphi(v_{ik})\psi(v_{kj}) \\
&= \sum_k (\pi_v(\varphi))_{ik}(\pi_v(\psi))_{kj} \\
&= (\pi_v(\varphi)\pi_v(\psi))_{ij}
\end{aligned}$$

Let us establish now the equality in the statement, namely:

$$Im(\pi_v) = End(v)'$$

Let us first prove the inclusion $\subset$. Given $\varphi \in C(G)^*$ and $T \in End(v)$, we have:

$$\begin{aligned}
[\pi_v(\varphi), T] = 0 &\iff \sum_k \varphi(v_{ik})T_{kj} = \sum_k T_{ik}\varphi(v_{kj}), \forall i, j \\
&\iff \varphi\left(\sum_k v_{ik}T_{kj}\right) = \varphi\left(\sum_k T_{ik}v_{kj}\right), \forall i, j \\
&\iff \varphi((vT)_{ij}) = \varphi((Tv)_{ij}), \forall i, j
\end{aligned}$$

But this latter formula is true, because $T \in End(v)$ means that we have:

$$vT = Tv$$

As for the converse inclusion $\supset$, the proof is quite similar. Indeed, by using the bicommutant theorem, this is the same as proving that we have:

$$Im(\pi_v)' \subset End(v)$$

But, by using the above equivalences, we have the following computation:

$$\begin{aligned}
T \in Im(\pi_v)' &\iff [\pi_v(\varphi), T] = 0, \forall \varphi \\
&\iff \varphi((vT)_{ij}) = \varphi((Tv)_{ij}), \forall \varphi, i, j \\
&\iff vT = Tv
\end{aligned}$$

Thus, we have obtained the desired inclusion, and we are done. $\square$

By combining the above results, we obtain the following technical statement:

THEOREM 14.11. *We have the following equality,*

$$E_{C_G}^{(s)'} = Im(\pi_v)$$

where the representation v is the following direct sum,

$$v = \bigoplus_{|k| \leq s} u^{\otimes k}$$

and where the algebra representation $\pi_v : C(G)^* \to M_n(\mathbb{C})$ is given by $\varphi \to (\varphi(v_{ij}))_{ij}$.

PROOF. This follows indeed by combining the above results, and more precisely by combining Proposition 14.9 and Proposition 14.10. □

We recall that we want to prove that we have $E_C^{(s)'} \subset E_{C_{G_C}}^{(s)'}$, for any $s \in \mathbb{N}$. For this purpose, we must first refine Theorem 14.11, in the case $G = G_C$.

14b. The correspondence

Generally speaking, in order to prove anything about G_C, we are in need of an explicit model for this group. In order to construct such a model, let $< u_{ij} >$ be the free $*$-algebra over $\dim(H)^2$ variables, with comultiplication and counit as follows:

$$\Delta(u_{ij}) = \sum_k u_{ik} \otimes u_{kj} \quad , \quad \varepsilon(u_{ij}) = \delta_{ij}$$

Following [68], we can model this $*$-bialgebra, in the following way:

PROPOSITION 14.12. *Consider the following pair of dual vector spaces,*

$$F = \bigoplus_k B\left(H^{\otimes k}\right) \quad , \quad F^* = \bigoplus_k B\left(H^{\otimes k}\right)^*$$

and let $f_{ij}, f_{ij}^* \in F^*$ be the standard generators of $B(H)^*, B(\bar{H})^*$.

1. *F^* is a $*$-algebra, with multiplication $\otimes$ and involution as follows:*
$$f_{ij} \leftrightarrow f_{ij}^*$$
2. *F^* is a $*$-bialgebra, with $*$-bialgebra operations as follows:*
$$\Delta(f_{ij}) = \sum_k f_{ik} \otimes f_{kj} \quad , \quad \varepsilon(f_{ij}) = \delta_{ij}$$
3. *We have a $*$-bialgebra isomorphism $< u_{ij} > \simeq F^*$, given by $u_{ij} \to f_{ij}$.*

PROOF. Since F^* is spanned by the various tensor products between the variables f_{ij}, f_{ij}^*, we have a vector space isomorphism as follows:

$$< u_{ij} > \simeq F^* \quad , \quad u_{ij} \to f_{ij} \quad , \quad u_{ij}^* \to f_{ij}^*$$

The corresponding $*$-bialgebra structure induced on the vector space F^* is then the one in the statement, and this gives the result. □

Now back to our group G_C, we have the following modelling result for it:

PROPOSITION 14.13. *The smooth part of the algebra $A_C = C(G_C)$ is given by*

$$\mathcal{A}_C \simeq F^*/J$$

where $J \subset F^*$ is the ideal coming from the following relations, for any i, j,

$$\sum_{p_1,\ldots,p_k} T_{i_1\ldots i_l,p_1\ldots p_k} f_{p_1 j_1} \otimes \ldots \otimes f_{p_k j_k} = \sum_{q_1,\ldots,q_l} T_{q_1\ldots q_l,j_1\ldots j_k} f_{i_1 q_1} \otimes \ldots \otimes f_{i_l q_l}$$

one for each pair of colored integers k, l, and each $T \in C(k, l)$.

PROOF. As a first observation, A_C appears as enveloping C^*-algebra of the following universal $*$-algebra, where $u = (u_{ij})$ is regarded as a formal corepresentation:

$$\mathcal{A}_C = \left\langle (u_{ij})_{i,j=1,\ldots,N} \middle| T \in Hom(u^{\otimes k}, u^{\otimes l}), \forall k, l, \forall T \in C(k, l) \right\rangle$$

With this observation in hand, the conclusion is that we have a formula as follows, where I is the ideal coming from the relations $T \in Hom(u^{\otimes k}, u^{\otimes l})$, with $T \in C(k, l)$:

$$\mathcal{A}_C =< u_{ij} > /I$$

Now if we denote by $J \subset F^*$ the image of the ideal I via the $*$-algebra isomorphism $< u_{ij} >\simeq F^*$ from Proposition 14.15, we obtain an identification as follows:

$$\mathcal{A}_C \simeq F^*/J$$

With standard multi-index notations, and by assuming now that $k, l \in \mathbb{N}$ are usual integers, for simplifying the presentation, the general case being similar, a relation of type $T \in Hom(u^{\otimes k}, u^{\otimes l})$ inside $< u_{ij} >$ is equivalent to the following conditions:

$$\sum_{p_1,\ldots,p_k} T_{i_1\ldots i_l,p_1\ldots p_k} u_{p_1 j_1} \ldots u_{p_k j_k} = \sum_{q_1,\ldots,q_l} T_{q_1\ldots q_l,j_1\ldots j_k} u_{i_1 q_1} \ldots u_{i_l q_l}$$

Now by recalling that the isomorphism of $*$-algebras $< u_{ij} >\to F^*$ is given by $u_{ij} \to f_{ij}$, and that the multiplication operation of F^* corresponds to the tensor product operation $\otimes$, we conclude that $J \subset F^*$ is the ideal from the statement. $\square$

With the above result in hand, let us go back to Theorem 14.11. We have:

PROPOSITION 14.14. *The linear space $\mathcal{A}_C^*$ is given by the formula*

$$\mathcal{A}_C^* = \left\{ a \in F \middle| T a_k = a_l T, \forall T \in C(k, l) \right\}$$

and the representation

$$\pi_v : \mathcal{A}_C^* \to B\left(\bigoplus_{|k|\leq s} H^{\otimes k}\right)$$

appears diagonally, by truncating, $\pi_v : a \to (a_k)_{kk}$.

PROOF. We know from Proposition 14.13 above that we have an identification of $*$-bialgebras $\mathcal{A}_C \simeq F^*/J$. But this gives a quotient map, as follows:

$$F^* \to \mathcal{A}_C$$

At the dual level, this gives $\mathcal{A}_C^* \subset F$. To be more precise, we have:

$$\mathcal{A}_C^* = \left\{a \in \ F \middle| f(a) = 0, \forall f \in J\right\}$$

Now since $J = < f_T >$, where f_T are the relations in Proposition 14.13, we obtain:

$$\mathcal{A}_C^* = \left\{a \in F \middle| f_T(a) = 0, \forall T \in C\right\}$$

Given $T \in C(k,l)$, for an arbitrary element $a = (a_k)$, we have:

$$\begin{aligned}
& f_T(a) = 0 \\
\iff & \sum_{p_1,\ldots,p_k} T_{i_1\ldots i_l,p_1\ldots p_k}(a_k)_{p_1\ldots p_k,j_1\ldots j_k} = \sum_{q_1,\ldots,q_l} T_{q_1\ldots q_l,j_1\ldots j_k}(a_l)_{i_1\ldots i_l,q_1\ldots q_l}, \forall i,j \\
\iff & (Ta_k)_{i_1\ldots i_l,j_1\ldots j_k} = (a_l T)_{i_1\ldots i_l,j_1\ldots j_k}, \forall i,j \\
\iff & Ta_k = a_l T
\end{aligned}$$

Thus, $\mathcal{A}_C^*$ is given by the formula in the statement. It remains to compute π_v:

$$\pi_v : \mathcal{A}_C^* \to B\left(\bigoplus_{|k|\leq s} H^{\otimes k}\right)$$

With $a = (a_k)$, we have the following computation:

$$\begin{aligned}
\pi_v(a)_{i_1\ldots i_k,j_1\ldots j_k} &= a(v_{i_1\ldots i_k,j_1\ldots j_k}) \\
&= (f_{i_1 j_1} \otimes \ldots \otimes f_{i_k j_k})(a) \\
&= (a_k)_{i_1\ldots i_k,j_1\ldots j_k}
\end{aligned}$$

Thus, our representation π_v appears diagonally, by truncating, as claimed. □

In order to further advance, consider the following vector spaces:

$$F_s = \bigoplus_{|k|\leq s} B\left(H^{\otimes k}\right) \quad , \quad F_s^* = \bigoplus_{|k|\leq s} B\left(H^{\otimes k}\right)^*$$

We denote by $a \to a_s$ the truncation operation $F \to F_s$. We have:

PROPOSITION 14.15. *The following hold:*

(1) $E_C^{(s)'} \subset F_s$.
(2) $E_C' \subset F$.
(3) $\mathcal{A}_C^* = E_C'$.
(4) $Im(\pi_v) = (E_C')_s$.

Proof. These results basically follow from what we have, as follows:

(1) We have an inclusion as follows, as a diagonal subalgebra:

$$F_s \subset B\left(\bigoplus_{|k|\leq s} H^{\otimes k}\right)$$

The commutant of this algebra is then given by:

$$F_s' = \left\{b \in F_s \Big| b = (b_k), b_k \in \mathbb{C}, \forall k\right\}$$

On the other hand, we know from the identity axiom for the category C that we have $F_s' \subset E_C^{(s)}$. Thus, our result follows from the bicommutant theorem, as follows:

$$F_s' \subset E_C^{(s)} \implies F_s \supset E_C^{(s)'}$$

(2) This follows from (1), by taking inductive limits.

(3) With the present notations, the formula of $\mathcal{A}_C^*$ from Proposition 14.14 reads $\mathcal{A}_C^* = F \cap E_C'$. Now since by (2) we have $E_C' \subset F$, we obtain from this $\mathcal{A}_C^* = E_C'$.

(4) This follows from (3), and from the formula of π_ν in Proposition 14.14. □

Following [68], we can now state and prove our main result, as follows:

Theorem 14.16. *The Tannakian duality constructions*

$$C \to G_C \quad , \quad G \to C_G$$

are inverse to each other.

Proof. According to our various results above, we have to prove that, for any Tannakian category C, and any $s \in \mathbb{N}$, we have an inclusion as follows:

$$E_C^{(s)'} \subset (E_C')_s$$

By taking duals, this is the same as proving that we have:

$$\left\{f \in F_s^* \Big| f_{|(E_C')_s} = 0\right\} \subset \left\{f \in F_s^* \Big| f_{|E_C^{(s)'}} = 0\right\}$$

In order to do so, we use the following formula, from Proposition 14.15:

$$\mathcal{A}_C^* = E_C'$$

We know that we have an identification as follows:

$$\mathcal{A}_C = F^*/J$$

We conclude that the ideal J is given by the following formula:

$$J = \left\{f \in F^* \Big| f_{|E_C'} = 0\right\}$$

Our claim is that we have the following formula, for any $s\in\mathbb N$:

$$J\cap F_s^*=\left\{f\in F_s^*\Big|f_{|E_C^{(s)\prime}}=0\right\}$$

Indeed, let us denote by X_s the spaces on the right. The axioms for C show that these spaces are increasing, that their union $X=\cup_sX_s$ is an ideal, and that:

$$X_s=X\cap F_s^*$$

We must prove that we have $J=X$, and this can be done as follows:

"$\subset$" This follows from the following fact, for any $T\in C(k,l)$ with $|k|,|l|\leq s$:

$$\begin{aligned}(f_T)_{|\{T\}'}=0\quad&\implies\quad(f_T)_{|E_C^{(s)\prime}}=0\\&\implies\quad f_T\in X_s\end{aligned}$$

"$\supset$" This follows from our description of J, because from $E_C^{(s)}\subset E_C$ we obtain:

$$f_{|E_C^{(s)\prime}}=0\implies f_{|E_C'}=0$$

Summarizing, we have proved our claim. On the other hand, we have:

$$\begin{aligned}J\cap F_s^*&=\left\{f\in F^*\Big|f_{|E_C'}=0\right\}\cap F_s^*\\&=\left\{f\in F_s^*\Big|f_{|E_C'}=0\right\}\\&=\left\{f\in F_s^*\Big|f_{|(E_C')_s}=0\right\}\end{aligned}$$

Thus, our claim is exactly the inclusion that we wanted to prove, and we are done. $\square$

Summarizing, we have proved Tannakian duality. We should mention that there are many other versions of this duality, the story being as follows:

(1) The original version of the duality, due to Tannaka and Krein, states that any compact group G can be recovered from the knowledge of its category of representations $\mathcal R_G$, viewed as subcategory of the category $\mathcal H$ of the finite dimensional Hilbert spaces, with each $v\in\mathcal R_G$ corresponding to its Hilbert space $H_v\in\mathcal H$.

(2) Regarding the proof of this fact, this is something established long ago by Tannaka and Krein. From a more modern perspective, coming from the work of Grothendieck, Saavedra and others, the group G simply appears as the group of endomorphisms of the embedding functor $\mathcal R_G\subset\mathcal H$. See Chari-Pressley [17].

(3) With this understood, comes now a non-trivial result, due independently to Deligne [25] and Doplicher-Roberts [31], stating that the compact group G can be recovered from the sole knowledge of the category $\mathcal R_G$. This is obviously something more advanced, and the proof is quite tricky.

(4) Getting back to Earth now, as already explained in the above, for concrete applications it is technically convenient to replace the full category $\mathcal{R}_G$ by its subcategory $\mathcal{R}_G^\circ$ consisting of the Peter-Weyl representations. And here, we have the analogue of the original result of Tannaka-Krein, explained in the above.

(5) For more specialized questions, it is convenient to further shrink the Peter-Weyl category $\mathcal{R}_G^\circ$, consisting of the spaces $C(k,l) = Hom(u^{\otimes k}, u^{\otimes l})$, to its diagonal algebra $P_G = \Delta\mathcal{R}_G$, consisting of the $*$-algebras $C(k,k) = End(u^{\otimes k})$. This is the idea behind Jones' planar algebras [**56**], with P_G being such an algebra.

(6) And this is not the end of the story, because one can still try to have Doplicher-Roberts and Deligne type results in the Peter-Weyl category setting, or in the planar algebra setting. We refer here to the modern quantum algebra literature, where Tannakian duality, in all its forms, is something highly valued.

Importantly, Tannakian duality of all sorts is just a way of "linearizing" the group, and there is as well a second method, namely considering the associated Lie algebra. Again, we refer here to the quantum algebra literature, for instance Chari-Pressley [**17**].

14c. Brauer theorems

As a basic illustration for the Tannakian correspondence, we will work out Brauer theorems for O_N, U_N. These are very classical results, and there are many possible proofs for them. We will follow here the modern approach from [**13**]. Let us start with:

DEFINITION 14.17. *Given a pairing $\pi \in P_2(k,l)$ and an integer $N \in \mathbb{N}$, we can construct a linear map between tensor powers of $\mathbb{C}^N$,*

$$T_\pi : (\mathbb{C}^N)^{\otimes k} \to (\mathbb{C}^N)^{\otimes l}$$

by the following formula, with $e_1, \ldots, e_N$ being the standard basis of $\mathbb{C}^N$,

$$T_\pi(e_{i_1} \otimes \ldots \otimes e_{i_k}) = \sum_{j_1 \ldots j_l} \delta_\pi \begin{pmatrix} i_1 & \ldots & i_k \\ j_1 & \ldots & j_l \end{pmatrix} e_{j_1} \otimes \ldots \otimes e_{j_l}$$

and with the coefficients on the right being Kronecker type symbols,

$$\delta_\pi \begin{pmatrix} i_1 & \ldots & i_k \\ j_1 & \ldots & j_l \end{pmatrix} \in \{0,1\}$$

whose values depend on whether the indices fit or not.

To be more precise here, we put the multi-indices $i = (i_1, \ldots, i_k)$ and $j = (j_1, \ldots, j_l)$ on the legs of our pairing π, in the obvious way. In the case where all strings of π join pairs of equal indices of i, j, we set $\delta_\pi\binom{i}{j} = 1$. Otherwise, we set $\delta_\pi\binom{i}{j} = 0$.

The point with the above definition comes from the fact that most of the "familiar" maps, in the Tannakian context, are of the above form. Here are some examples:

PROPOSITION 14.18. *The correspondence $\pi \to T_\pi$ has the following properties:*

(1) $T_\cap = (1 \to \sum_i e_i \otimes e_i)$.
(2) $T_\cup = (e_i \otimes e_j \to \delta_{ij})$.
(3) $T_{||\ldots||} = id$.
(4) $T_{\slash\!\!\!\backslash} = (e_a \otimes e_b \to e_b \otimes e_a)$.

PROOF. We can assume that all legs of π are colored $\circ$, and then:

(1) We have $\cap \in P_2(\emptyset, \circ\circ)$, so the corresponding linear map is as follows:

$$T_\cap : \mathbb{C} \to \mathbb{C}^N \otimes \mathbb{C}^N$$

The formula of this map is then, as claimed:

$$\begin{aligned} T_\cap(1) &= \sum_{ij} \delta_\cap(i\ j) e_i \otimes e_j \\ &= \sum_{ij} \delta_{ij} e_i \otimes e_j \\ &= \sum_i e_i \otimes e_i \end{aligned}$$

(2) Here we have $\cup \in P_2(\circ\circ, \emptyset)$, so the corresponding linear map is as follows:

$$T_\cap : \mathbb{C}^N \otimes \mathbb{C}^N \to \mathbb{C}$$

The formula of this linear form is then as follows:

$$T_\cap(e_i \otimes e_j) = \delta_\cap(i\ j) = \delta_{ij}$$

(3) Consider indeed the "identity" pairing $|| \ldots || \in P_2(k,k)$, with $k = \circ\circ \ldots \circ\circ$. The corresponding linear map is then the identity, because we have:

$$\begin{aligned} T_{||\ldots||}(e_{i_1} \otimes \ldots \otimes e_{i_k}) &= \sum_{j_1 \ldots j_k} \delta_{||\ldots||} \begin{pmatrix} i_1 & \ldots & i_k \\ j_1 & \ldots & j_k \end{pmatrix} e_{j_1} \otimes \ldots \otimes e_{j_k} \\ &= \sum_{j_1 \ldots j_k} \delta_{i_1 j_1} \ldots \delta_{i_k j_k} e_{j_1} \otimes \ldots \otimes e_{j_k} \\ &= e_{i_1} \otimes \ldots \otimes e_{i_k} \end{aligned}$$

(4) For the basic crossing $\slash\!\!\!\backslash \in P_2(\circ\circ, \circ\circ)$, the corresponding linear map is as follows:

$$T_{\slash\!\!\!\backslash} : \mathbb{C}^N \otimes \mathbb{C}^N \to \mathbb{C}^N \otimes \mathbb{C}^N$$

This map can be computed as follows:

$$\begin{aligned}
T_{\chi}(e_i \otimes e_j) &= \sum_{kl} \delta_{\chi}\begin{pmatrix} i & j \\ k & l \end{pmatrix} e_k \otimes e_l \\
&= \sum_{kl} \delta_{il}\delta_{jk} e_k \otimes e_l \\
&= e_j \otimes e_i
\end{aligned}$$

Thus we obtain the flip operator $\Sigma(a \otimes b) = b \otimes a$, as claimed. $\square$

The relation with the Tannakian categories comes from the following key result:

PROPOSITION 14.19. *The assignement $\pi \to T_\pi$ is categorical, in the sense that*

$$T_\pi \otimes T_\sigma = T_{[\pi\sigma]} \quad , \quad T_\pi T_\sigma = N^{c(\pi,\sigma)} T_{[^\sigma_\pi]} \quad , \quad T_\pi^* = T_{\pi^*}$$

where $c(\pi,\sigma)$ are certain integers, coming from the erased components in the middle.

PROOF. The concatenation axiom follows from the following computation:

$$\begin{aligned}
&(T_\pi \otimes T_\sigma)(e_{i_1} \otimes \ldots \otimes e_{i_p} \otimes e_{k_1} \otimes \ldots \otimes e_{k_r}) \\
= &\sum_{j_1\ldots j_q} \sum_{l_1\ldots l_s} \delta_\pi \begin{pmatrix} i_1 & \ldots & i_p \\ j_1 & \ldots & j_q \end{pmatrix} \delta_\sigma \begin{pmatrix} k_1 & \ldots & k_r \\ l_1 & \ldots & l_s \end{pmatrix} e_{j_1} \otimes \ldots \otimes e_{j_q} \otimes e_{l_1} \otimes \ldots \otimes e_{l_s} \\
= &\sum_{j_1\ldots j_q} \sum_{l_1\ldots l_s} \delta_{[\pi\sigma]} \begin{pmatrix} i_1 & \ldots & i_p & k_1 & \ldots & k_r \\ j_1 & \ldots & j_q & l_1 & \ldots & l_s \end{pmatrix} e_{j_1} \otimes \ldots \otimes e_{j_q} \otimes e_{l_1} \otimes \ldots \otimes e_{l_s} \\
= &T_{[\pi\sigma]}(e_{i_1} \otimes \ldots \otimes e_{i_p} \otimes e_{k_1} \otimes \ldots \otimes e_{k_r})
\end{aligned}$$

The composition axiom follows from the following computation:

$$\begin{aligned}
&T_\pi T_\sigma(e_{i_1} \otimes \ldots \otimes e_{i_p}) \\
= &\sum_{j_1\ldots j_q} \delta_\sigma \begin{pmatrix} i_1 & \ldots & i_p \\ j_1 & \ldots & j_q \end{pmatrix} \sum_{k_1\ldots k_r} \delta_\pi \begin{pmatrix} j_1 & \ldots & j_q \\ k_1 & \ldots & k_r \end{pmatrix} e_{k_1} \otimes \ldots \otimes e_{k_r} \\
= &\sum_{k_1\ldots k_r} N^{c(\pi,\sigma)} \delta_{[^\sigma_\pi]} \begin{pmatrix} i_1 & \ldots & i_p \\ k_1 & \ldots & k_r \end{pmatrix} e_{k_1} \otimes \ldots \otimes e_{k_r} \\
= &N^{c(\pi,\sigma)} T_{[^\sigma_\pi]}(e_{i_1} \otimes \ldots \otimes e_{i_p})
\end{aligned}$$

Finally, the involution axiom follows from the following computation:

$$\begin{aligned}
& T_\pi^*(e_{j_1}\otimes\ldots\otimes e_{j_q})\\
= & \sum_{i_1\ldots i_p} < T_\pi^*(e_{j_1}\otimes\ldots\otimes e_{j_q}), e_{i_1}\otimes\ldots\otimes e_{i_p} > e_{i_1}\otimes\ldots\otimes e_{i_p}\\
= & \sum_{i_1\ldots i_p} \delta_\pi\begin{pmatrix} i_1 & \ldots & i_p\\ j_1 & \ldots & j_q\end{pmatrix} e_{i_1}\otimes\ldots\otimes e_{i_p}\\
= & T_{\pi^*}(e_{j_1}\otimes\ldots\otimes e_{j_q})
\end{aligned}$$

Summarizing, our correspondence is indeed categorical. $\square$

The above result suggests the following general definition, from [13]:

DEFINITION 14.20. *Let $P_2(k,l)$ be the set of pairings between an upper colored integer k, and a lower colored integer l. A collection of subsets*

$$D = \bigsqcup_{k,l} D(k,l)$$

with $D(k,l)\subset P_2(k,l)$ is called a category of pairings when it has the following properties:

1. *Stability under the horizontal concatenation, $(\pi,\sigma)\to[\pi\sigma]$.*
2. *Stability under vertical concatenation $(\pi,\sigma)\to[^\sigma_\pi]$, with matching middle symbols.*
3. *Stability under the upside-down turning $*$, with switching of colors, $\circ\leftrightarrow\bullet$.*
4. *Each set $P(k,k)$ contains the identity partition $||\ldots||$.*
5. *The sets $P(\emptyset,\circ\bullet)$ and $P(\emptyset,\bullet\circ)$ both contain the semicircle $\cap$.*
6. *The sets $P(k,\bar{k})$ with $|k|=2$ contain the crossing partition $\slash\!\!\!\backslash$.*

Observe the similarity with the axioms for Tannakian categories, from the beginning of this chapter. We will see in a moment that this similarity can be turned into something very precise, with the categories of pairings producing Tannakian categories.

As basic examples of such categories of pairings, that we have already met in the above, we have the categories $P_2,\mathcal{P}_2$ of pairings, and of matching pairings. There are many other such categories, and we will discuss this gradually, in what follows.

In relation with the compact groups, we have the following result:

THEOREM 14.21. *Each category of pairings, in the above sense,*

$$D=(D(k,l))$$

produces a family of compact groups $G=(G_N)$, one for each $N\in\mathbb{N}$, via the formula

$$Hom(u^{\otimes k},u^{\otimes l}) = span\left(T_\pi\Big|\pi\in D(k,l)\right)$$

and the Tannakian duality correspondence.

PROOF. Given an integer $N \in \mathbb{N}$, consider the correspondence $\pi \to T_\pi$ constructed in Definition 14.17, and then the collection of linear spaces in the statement, namely:

$$C_{kl} = span\left(T_\pi \Big| \pi \in D(k,l)\right)$$

According to Proposition 14.19, and to our axioms for the categories of partitions, from Definition 14.20, this collection of spaces $C = (C_{kl})$ satisfies the axioms for the Tannakian categories, from the beginning of this chapter. Thus the Tannakian duality result there applies, and provides us with a closed subgroup $G_N \subset U_N$ such that:

$$C_{kl} = Hom(u^{\otimes k}, u^{\otimes l})$$

Thus, we are led to the conclusion in the statement. □

The above result is something fundamental, and suggests formulating:

DEFINITION 14.22. *A compact group $G \subset_u U_N$ having the property*

$$Hom(u^{\otimes k}, u^{\otimes l}) = span\left(T_\pi \Big| \pi \in D(k,l)\right)$$

for a cartain category of pairings $D = (D(k,l))$ is called easy.

This definition, from [13], is motivated by the fact that, from the point of view of Tannakian duality, the above groups are indeed the "easiest" possible ones. Of course, this might sound a bit offending, after all the complicated things that we did in this chapter. But hey, there is a beginning of everything. We will get to know better Tannakian duality and easiness, and their applications, in what follows, and please believe me, you will reach too to the conclusion that Definition 14.22 is justified.

As another comment, it is possible to talk about more general easy groups, by using general categories of partitions, instead of just categories of pairings. We will be back to all this, with a systematic study of easiness, in chapter 15 below.

As a technical remark now, to be always kept in mind, when dealing with easiness, the category of pairings producing an easy group is not unique, for instance because at $N = 1$ all the possible categories of pairings produce the same easy group, namely the trivial group $G = \{1\}$. Thus, some subtleties are going on here. More on this later.

Getting back now to concrete things, the point now is that with the above ingredients in hand, and as a first application of Tannakian duality, we can establish a useful result, namely the Brauer theorem for the unitary group U_N. The statement is a follows:

THEOREM 14.23. *For the unitary group U_N we have*

$$Hom(u^{\otimes k}, u^{\otimes l}) = span\left(T_\pi \Big| \pi \in \mathcal{P}_2(k,l)\right)$$

where $\mathcal{P}_2$ denotes as usual the category of all matching pairings.

Proof. This is something very old and classical, due to Brauer [**16**], and in what follows we will present a simplified proof for it, based on the easiness technology developed above. Consider the spaces on the right in the statement, namely:

$$C_{kl}=span\left(T_\pi\Big|\pi\in\mathcal{P}_2(k,l)\right)$$

According to Proposition 14.19 these spaces form a tensor category. Thus, by Tannakian duality, these spaces must come from a certain closed subgroup $G\subset U_N$. To be more precise, if we denote by v the fundamental representation of G, then:

$$C_{kl}=Hom(v^{\otimes k},v^{\otimes l})$$

We must prove that we have $G=U_N$. For this purpose, let us recall that the unitary group U_N is defined via the following relations:

$$u^*=u^{-1}\quad,\quad u^t=\bar{u}^{-1}$$

But these relations tell us precisely that the following two operators must be in the associated Tannakian category C:

$$T_\pi\quad:\quad\pi={}^{\cap}_{\circ\bullet}\ ,\ {}^{\cap}_{\bullet\circ}$$

Thus the associated Tannakian category is $C=span(T_\pi|\pi\in D)$, with:

$$D=<{}^{\cap}_{\circ\bullet}\ ,\ {}^{\cap}_{\bullet\circ}>=\mathcal{P}_2$$

Thus, we are led to the conclusion in the statement. □

Regarding the orthogonal group O_N, we have here a similar result, as follows:

Theorem 14.24. *For the orthogonal group O_N we have*

$$Hom(u^{\otimes k},u^{\otimes l})=span\left(T_\pi\Big|\pi\in P_2(k,l)\right)$$

where P_2 denotes as usual the category of all pairings.

Proof. As before with Theorem 14.23, regarding U_N, this is something very old and classical, due to Brauer [**16**], that we can now prove by using the easiness technology developed above. Consider the spaces on the right in the statement, namely:

$$C_{kl}=span\left(T_\pi\Big|\pi\in P_2(k,l)\right)$$

According to Proposition 14.19 these spaces form a tensor category. Thus, by Tannakian duality, these spaces must come from a certain closed subgroup $G\subset U_N$. To be more precise, if we denote by v the fundamental representation of G, then:

$$C_{kl}=Hom(v^{\otimes k},v^{\otimes l})$$

We must prove that we have $G=O_N$. For this purpose, let us recall that the orthogonal group $O_N\subset U_N$ is defined by imposing the following relations:

$$u_{ij}=\bar{u}_{ij}$$

But these relations tell us precisely that the following two operators must be in the associated Tannakian category C:

$$T_\pi \quad : \quad \pi = {}^{\circ}_{\bullet}\ ,\ {}^{\bullet}_{\circ}$$

Thus the associated Tannakian category is $C = span(T_\pi|\pi \in D)$, with:

$$D = < \mathcal{P}_2, {}^{\circ}_{\bullet}, {}^{\bullet}_{\circ} > = P_2$$

Thus, we are led to the conclusion in the statement. □

We will see later, in chapter 16 below, applications of the above results, to integration problems over O_N, U_N, by using the Peter-Weyl methods from chapter 13.

14d. Clebsch-Gordan rules

As a last piece of pure algebra, we are now in position of dealing, in a conceptual way, with SU_2 and SO_3. Regarding SU_2, the result here is as follows:

Theorem 14.25. *The irreducible representations of SU_2 are all self-adjoint, and can be labelled by positive integers, with their fusion rules being as follows,*

$$r_k \otimes r_l = r_{|k-l|} + r_{|k-l|+2} + \ldots + r_{k+l}$$

called Clebsch-Gordan rules. The corresponding dimensions are $\dim r_k = k + 1$.

Proof. There are several proofs for this fact, the simplest one, with the knowledge that we have, being via purely algebraic methods, as follows:

(1) Our first claim is that we have the following estimate, telling us that the even moments of the main character are smaller than the Catalan numbers:

$$\int_{SU_2} \chi^{2k} \leq C_k$$

But this is something that we know from chapter 8, obtained by using $SU_2 \simeq S^3_{\mathbb{R}}$ and spherical integrals, and with the stronger statement that we have in fact equality $=$. However, for the purposes of what follows, the above $\leq$ estimate will do.

(2) Alternatively, the above estimate can be deduced with purely algebraic methods, by using an easiness type argument for SU_2, as follows:

$$\begin{aligned}
\int_{SU_2} \chi^{2k} &= \dim(Fix(u^{\otimes 2k})) \\
&= \dim\left(span\left(T'_\pi\Big|\pi \in NC_2(2k)\right)\right) \\
&\leq |NC_2(2k)| \\
&= C_k
\end{aligned}$$

To be more precise, SU_2 is not exactly easy, but rather "super-easy", coming from a different implementation $\pi \to T'_\pi$ of the pairings, involving some signs. And with this being proved exactly as the Brauer theorem for O_N, with modifications where needed.

(3) Long story short, we have our estimate in (1), and this is all that we need. Our claim is that we can construct, by recurrence on $k \in \mathbb{N}$, a sequence r_k of irreducible, self-adjoint and distinct representations of SU_2, satisfying:

$$r_0 = 1 \quad , \quad r_1 = u \quad , \quad r_k + r_{k-2} = r_{k-1} \otimes r_1$$

Indeed, assume that $r_0, \ldots, r_{k-1}$ are constructed, and let us construct r_k. We have:

$$r_{k-1} + r_{k-3} = r_{k-2} \otimes r_1$$

Thus $r_{k-1} \subset r_{k-2} \otimes r_1$, and since r_{k-2} is irreducible, by Frobenius we have:

$$r_{k-2} \subset r_{k-1} \otimes r_1$$

We conclude there exists a certain representation r_k such that:

$$r_k + r_{k-2} = r_{k-1} \otimes r_1$$

(4) By recurrence, r_k is self-adjoint. Now observe that according to our recurrence formula, we can split $u^{\otimes k}$ as a sum of the following type, with positive coefficients:

$$u^{\otimes k} = c_k r_k + c_{k-2} r_{k-2} + \ldots$$

We conclude by Peter-Weyl that we have an inequality as follows, with equality precisely when r_k is irreducible, and non-equivalent to the other summands r_i:

$$\sum_i c_i^2 \leq \dim(End(u^{\otimes k}))$$

(5) But by (1) the number on the right is $\leq C_k$, and some straightforward combinatorics, based on the fusion rules, shows that the number on the left is C_k as well:

$$C_k = \sum_i c_i^2 \leq \dim(End(u^{\otimes k})) = \int_{SU_2} \chi^{2k} \leq C_k$$

Thus we have equality in our estimate, so our representation r_k is irreducible, and non-equivalent to $r_{k-2}, r_{k-4}, \ldots$ Moreover, this representation r_k is not equivalent to $r_{k-1}, r_{k-3}, \ldots$ either, with this coming from $r_p \subset u^{\otimes p}$ for any p, and from:

$$\dim(Fix(u^{\otimes 2s+1})) = \int_{SU_2} \chi^{2s+1} = 0$$

(6) Thus, we proved our claim. Now since each irreducible representation of SU_2 appears into some $u^{\otimes k}$, and we know how to decompose each $u^{\otimes k}$ into sums of representations r_k, these representations r_k are all the irreducible representations of SU_2, and we are done with the main assertion. As for the dimension formula, this is clear. $\square$

Regarding now SO_3, we have here a similar result, as follows:

THEOREM 14.26. *The irreducible representations of SO_3 are all self-adjoint, and can be labelled by positive integers, with their fusion rules being as follows,*

$$r_k \otimes r_l = r_{|k-l|} + r_{|k-l|+1} + \ldots + r_{k+l}$$

also called Clebsch-Gordan rules. The corresponding dimensions are $\dim r_k = 2k+1$.

PROOF. As before with SU_2, there are many possible proofs here, which are all instructive. Here is our take on the subject, in the spirit of our proof for SU_2:

(1) Our first claim is that we have the following formula, telling us that the moments of the main character equal the Catalan numbers:

$$\int_{SO_3} \chi^k = C_k$$

But this is something that we know from chapter 8, coming from Euler-Rodrigues. Alternatively, this can be deduced as well from Tannakian duality, a bit as for SU_2.

(2) Our claim now is that we can construct, by recurrence on $k \in \mathbb{N}$, a sequence r_k of irreducible, self-adjoint and distinct representations of SO_3, satisfying:

$$r_0 = 1 \quad , \quad r_1 = u - 1 \quad , \quad r_k + r_{k-1} + r_{k-2} = r_{k-1} \otimes r_1$$

Indeed, assume that $r_0, \ldots, r_{k-1}$ are constructed, and let us construct r_k. The Frobenius trick from the proof for SU_2 will no longer work, due to some technical reasons, so we have to invoke (1). To be more precise, by integrating characters we obtain:

$$r_{k-1}, r_{k-2} \subset r_{k-1} \otimes r_1$$

Thus there exists a representation r_k such that:

$$r_{k-1} \otimes r_1 = r_k + r_{k-1} + r_{k-2}$$

(3) Once again by integrating characters, we conclude that r_k is irreducible, and non-equivalent to $r_1, \ldots, r_{k-1}$, and this proves our claim. Also, since any irreducible representation of SO_3 must appear in some tensor power of u, and we can decompose each $u^{\otimes k}$ into sums of representations r_p, we conclude that these representations r_p are all the irreducible representations of SO_3. Finally, the dimension formula is clear. $\square$

There are of course many other things that can be said about SU_2 and SO_3. For instance, with the proof of Theorem 14.25 and Theorem 14.26 done in a purely algebraic fashion, by using the super-easiness property of SU_2 and SO_3, the Euler-Rodrigues formula can be deduced afterwards from this, without any single computation, the argument being that by Peter-Weyl the embedding $PU_2 \subset SO_3$ must be indeed an equality.

As a conclusion to all this, you have now a decent level in group theory, and algebra in general, and you can start if you want exploring all sorts of other things, such as:

(1) Quantum groups. This is something modern and interesting, inspired by quantum mechanics, and the algebra is not that much complicated than what we did in the above. The must-read papers here are those of Drinfeld [32], Jimbo [52] on one hand, and of Woronowicz [97], [98] on the other. As for books, you have Chari-Pressley [17] for the Drinfeld-Jimbo quantum groups, and my book [6] for Woronowicz quantum groups.

(2) Planar algebras. This is something very related to the quantum groups, and perhaps even more exciting than them, due to all sorts of pictures, and relations with modern physics, developed by Jones in [53], [54], [55], [56]. With all sorts of interesting ramifications, and you can check here too the classical books or papers of Atiyah [5], Di Francesco [26], Temperley-Lieb [84], Witten [96] and Zwiebach [100].

But since we are now towards the end of the present book, better stay with us, and you can look into all this afterwards. We still have all sorts of interesting things to be done, including applying all the algebra that we learned, to questions in probability.

14e. Exercises

With the technology presented above, however, we can work out a few interesting particular cases of the Tannakian duality, and this will be the purpose of the first few exercises that we have here. Let us start with something quite elementary:

EXERCISE 14.27. *Work out the Tannakian duality for the closed subgroups*

$$G\subset O_N$$

first as a consequence of the general results that we have, regarding the closed subgroups

$$G\subset U_N$$

and then independently, by pointing out the simplifications that appear in the real case.

Regarding the first question, this is normally something quite quick, obtained by adding the assumption $u=\bar{u}$ to the Tannakian statement that we have, and then working out the details. Regarding the second question, the idea here is basically that the colored exponents $k,l=\circ\bullet\bullet\circ\ldots$ will become in this way usual exponents, $k,l\in\mathbb{N}$, and this brings a number of simplifications in the proof, which are to be found.

EXERCISE 14.28. *Work out the Tannakian duality for the closed subgroups*

$$G\subset U_N$$

whose fundamental representation is self-adjoint, up to equivalence,

$$u\sim\bar{u}$$

first as a consequence of the results that we have, and then independently.

Here are there are several possible paths, either by proceeding a bit as for the previous exercise, but with the condition $u = \bar{u}$ there replaced by the more general condition $u \sim \bar{u}$, or by using what was done in the previous exercise, and generalizing, from $u = \bar{u}$ to $u \sim \bar{u}$. In any case, regardless of the method which is chosen, the problem is that understanding what the condition $u \sim \bar{u}$ really means, categorically speaking.

EXERCISE 14.29. *Check the Brauer theorems for O_N, U_N, which are both of type*

$$Hom(u^{\otimes k}, u^{\otimes l}) = span\left(T_\pi\Big|\pi \in D(k,l)\right)$$

for small values of the global length parameter, $k + l \in \{1, 2, 3\}$.

The idea here is to prove these results that we already know directly, by double inclusion, with the inclusion in one sense being normally something quite elementary, and with the inclusion in the other sense being probably somehing quite tricky.

EXERCISE 14.30. *Write down Brauer theorems for the bistochastic groups*

$$B_N \subset O_N \quad , \quad C_N \subset U_N$$

by identifying first the partition which produces them, as subgroups of O_N, U_N.

This is actually something that will be discussed later on in this book, but without too much details, so the answer "done in the book" will not do.

EXERCISE 14.31. *Look up the original version of Tannakian duality, stating that G can be recovered from the knowledge of its full category of representations $\mathcal{R}_G$, viewed as subcategory of the category $\mathcal{H}$ of the finite dimensional Hilbert spaces, with each $\pi \in \mathcal{R}_G$ corresponding to its Hilbert space $H_\pi \in \mathcal{H}$, and write down a brief account of this.*

As already mentioned in the above, the idea is that the group G appears as the group of endomorphisms of the embedding functor $\mathcal{R}_G \subset \mathcal{H}$. Time to understand this.

EXERCISE 14.32. *Look up the Doplicher-Roberts and Deligne theorems, stating that the compact group G can be in fact recovered from the sole knowledge of the category $\mathcal{R}_G$, with no need for the embedding into $\mathcal{H}$, and write down a brief account of this.*

This is obviously something more advanced, and the proof is quite tricky. Try however to undertand the main ideas behid the proof, which are very instructive.

EXERCISE 14.33. *Given a closed subgroup $G \subset_u U_N$, understand and then briefly explain, in a short piece of writing, why the $*$-algebras*

$$C(k,k) = End(u^{\otimes k})$$

form a planar algebra in the sense of Jones, and then comment as well on the various formulations of Tannakian duality, in the planar algebra setting.

This is actually quite difficult. And as a final, bonus exercise, try learning as well some Lie algebras, and their relation with the above, and write down what you learned.

CHAPTER 15

Diagrams, easiness

15a. Easy groups

We have seen in the previous chapter that the Tannakian duals of the groups O_N, U_N are very simple objects. To be more precise, the Brauer theorem for these two groups states that we have equalities as follows, with $D = P_2, \mathcal{P}_2$ respectively:

$$Hom(u^{\otimes k}, u^{\otimes l}) = span\left(T_\pi\Big|\pi \in D(k,l)\right)$$

Our goal here will be that of axiomatizing and studying the closed subgroups $G \subset U_N$ which are of this type, that we will call "easy". Our results will be as follows:

(1) At the level of the examples, we will see that besides O_N, U_N, we have the bistochastic groups B_N, C_N, the symmetric group S_N, the hyperoctahedral group H_N, and more generally the series of complex reflection groups H_N^s.

(2) Also at the level of the basic examples, some key groups such as the symplectic group Sp_N are not easy, but we will show here that these are covered by a suitable "super-easiness" version of the easiness, as defined above.

(3) At the level of the general theory, we will develop some algebraic theory in this chapter, for the most in relation with various product operations, the idea being that in the easy case, everything eventually reduces to computations with partitions.

(4) Also at the level of the general theory, we will develop as well some analytic theory, in the next chapter, based on the same idea, namely that in the easy case, everything eventually reduces to some elementary computations with partitions.

In order to get started, let us formulate the following key definition, extending to the case of arbitrary partitions what we already know about pairings:

Definition 15.1. *Given a partition $\pi \in P(k,l)$ and an integer $N \in \mathbb{N}$, we define*

$$T_\pi : (\mathbb{C}^N)^{\otimes k} \to (\mathbb{C}^N)^{\otimes l}$$

by the following formula, with $e_1, \ldots, e_N$ being the standard basis of $\mathbb{C}^N$,

$$T_\pi(e_{i_1} \otimes \ldots \otimes e_{i_k}) = \sum_{j_1 \ldots j_l} \delta_\pi \begin{pmatrix} i_1 & \ldots & i_k \\ j_1 & \ldots & j_l \end{pmatrix} e_{j_1} \otimes \ldots \otimes e_{j_l}$$

and with the coefficients on the right being Kronecker type symbols.

To be more precise here, in order to compute the Kronecker type symbols $\delta_\pi\binom{i}{j} \in \{0,1\}$, we proceed exactly as in the pairing case, namely by putting the multi-indices $i = (i_1, \ldots, i_k)$ and $j = (j_1, \ldots, j_l)$ on the legs of π, in the obvious way. In case all the blocks of π contain equal indices of i, j, we set $\delta_\pi\binom{i}{j} = 1$. Otherwise, we set $\delta_\pi\binom{i}{j} = 0$.

With the above notion in hand, we can now formulate the following key definition, from [**13**], motivated by the Brauer theorems for O_N, U_N, as indicated before:

DEFINITION 15.2. *A closed subgroup $G \subset U_N$ is called easy when*

$$Hom(u^{\otimes k}, u^{\otimes l}) = span\left(T_\pi \Big| \pi \in D(k,l)\right)$$

for any two colored integers $k, l = \circ\bullet\circ\bullet\ldots$, for certain sets of partitions

$$D(k,l) \subset P(k,l)$$

where $\pi \to T_\pi$ is the standard implementation of the partitions, as linear maps.

In other words, we call a group G easy when its Tannakian category appears in the simplest possible way: from the linear maps associated to partitions. The terminology is quite natural, because Tannakian duality is basically our only serious tool.

As basic examples, the orthogonal and unitary groups O_N, U_N are both easy, coming respectively from the following collections of sets of partitions:

$$P_2 = \bigsqcup_{k,l} P_2(k,l) \quad , \quad \mathcal{P}_2 = \bigsqcup_{k,l} \mathcal{P}_2(k,l)$$

In the general case now, as a theoretical remark, in the context of Definition 15.2 above, consider the following collection of sets of partitions:

$$D = \bigsqcup_{k,l} D(k,l)$$

This collection of sets D obviously determines G, but the converse is not true. Indeed, at $N = 1$ for instance, both the choices $D = P_2, \mathcal{P}_2$ produce the same easy group, namely $G = \{1\}$. We will be back to this issue on several occasions, with results about it.

In order to advance, our first goal will be that of establishing a duality between easy groups and certain special classes of collections of sets as above, namely:

$$D = \bigsqcup_{k,l} D(k,l)$$

Let us begin with a general definition, from [**13**], as follows:

DEFINITION 15.3. *Let $P(k,l)$ be the set of partitions between an upper colored integer k, and a lower colored integer l. A collection of subsets*

$$D = \bigsqcup_{k,l} D(k,l)$$

with $D(k,l) \subset P(k,l)$ is called a category of partitions when it has the following properties:

(1) *Stability under the horizontal concatenation, $(\pi,\sigma) \to [\pi\sigma]$.*
(2) *Stability under vertical concatenation $(\pi,\sigma) \to [^\sigma_\pi]$, with matching middle symbols.*
(3) *Stability under the upside-down turning $*$, with switching of colors, $\circ \leftrightarrow \bullet$.*
(4) *Each set $P(k,k)$ contains the identity partition $||\ldots||$.*
(5) *The sets $P(\emptyset, \circ\bullet)$ and $P(\emptyset, \bullet\circ)$ both contain the semicircle $\cap$.*
(6) *The sets $P(k,\bar{k})$ with $|k|=2$ contain the crossing partition $\slash\!\!\!\backslash$.*

As before, this is something that we already met in chapter 14, but for the pairings only. Observe the similarity with the axioms for Tannakian categories, also from chapter 14. We will see in a moment that this similarity can be turned into something very precise, the idea being that such a category produces a family of easy quantum groups $(G_N)_{N\in\mathbb{N}}$, one for each $N \in \mathbb{N}$, via the formula in Definition 15.1, and Tannakian duality.

As basic examples, that we have already met in chapter 14, in connection with the representation theory of O_N, U_N, we have the categories $P_2, \mathcal{P}_2$ of pairings, and of matching pairings. Further basic examples include the categories P, P_{even} of all partitions, and of all partitions whose blocks have even size. We will see later that these latter categories are related to the symmetric and hyperoctahedral groups S_N, H_N.

The relation with the Tannakian categories comes from the following result:

PROPOSITION 15.4. *The assignement $\pi \to T_\pi$ is categorical, in the sense that*

$$T_\pi \otimes T_\sigma = T_{[\pi\sigma]} \quad , \quad T_\pi T_\sigma = N^{c(\pi,\sigma)} T_{[^\sigma_\pi]} \quad , \quad T_\pi^* = T_{\pi^*}$$

where $c(\pi,\sigma)$ are certain integers, coming from the erased components in the middle.

PROOF. This is something that we already know for the pairings, from chapter 14 above, and the proof in general is similar, the computations being as follows:

(1) The concatenation axiom follows from the following computation:

$$\begin{aligned}
&(T_\pi \otimes T_\sigma)(e_{i_1} \otimes \ldots \otimes e_{i_p} \otimes e_{k_1} \otimes \ldots \otimes e_{k_r})\\
=&\sum_{j_1\ldots j_q}\sum_{l_1\ldots l_s} \delta_\pi\begin{pmatrix} i_1 & \ldots & i_p\\ j_1 & \ldots & j_q\end{pmatrix} \delta_\sigma\begin{pmatrix} k_1 & \ldots & k_r\\ l_1 & \ldots & l_s\end{pmatrix} e_{j_1} \otimes \ldots \otimes e_{j_q} \otimes e_{l_1} \otimes \ldots \otimes e_{l_s}\\
=&\sum_{j_1\ldots j_q}\sum_{l_1\ldots l_s} \delta_{[\pi\sigma]}\begin{pmatrix} i_1 & \ldots & i_p & k_1 & \ldots & k_r\\ j_1 & \ldots & j_q & l_1 & \ldots & l_s\end{pmatrix} e_{j_1} \otimes \ldots \otimes e_{j_q} \otimes e_{l_1} \otimes \ldots \otimes e_{l_s}\\
=&\ T_{[\pi\sigma]}(e_{i_1} \otimes \ldots \otimes e_{i_p} \otimes e_{k_1} \otimes \ldots \otimes e_{k_r})
\end{aligned}$$

(2) The composition axiom follows from the following computation:

$$\begin{aligned}
& T_\pi T_\sigma(e_{i_1}\otimes\ldots\otimes e_{i_p})\\
= & \sum_{j_1\ldots j_q}\delta_\sigma\begin{pmatrix}i_1&\ldots&i_p\\ j_1&\ldots&j_q\end{pmatrix}\sum_{k_1\ldots k_r}\delta_\pi\begin{pmatrix}j_1&\ldots&j_q\\ k_1&\ldots&k_r\end{pmatrix}e_{k_1}\otimes\ldots\otimes e_{k_r}\\
= & \sum_{k_1\ldots k_r}N^{c(\pi,\sigma)}\delta_{[^\sigma_\pi]}\begin{pmatrix}i_1&\ldots&i_p\\ k_1&\ldots&k_r\end{pmatrix}e_{k_1}\otimes\ldots\otimes e_{k_r}\\
= & N^{c(\pi,\sigma)}T_{[^\sigma_\pi]}(e_{i_1}\otimes\ldots\otimes e_{i_p})
\end{aligned}$$

(3) Finally, the involution axiom follows from the following computation:

$$\begin{aligned}
& T_\pi^*(e_{j_1}\otimes\ldots\otimes e_{j_q})\\
= & \sum_{i_1\ldots i_p}< T_\pi^*(e_{j_1}\otimes\ldots\otimes e_{j_q}),e_{i_1}\otimes\ldots\otimes e_{i_p}> e_{i_1}\otimes\ldots\otimes e_{i_p}\\
= & \sum_{i_1\ldots i_p}\delta_\pi\begin{pmatrix}i_1&\ldots&i_p\\ j_1&\ldots&j_q\end{pmatrix}e_{i_1}\otimes\ldots\otimes e_{i_p}\\
= & T_{\pi^*}(e_{j_1}\otimes\ldots\otimes e_{j_q})
\end{aligned}$$

Summarizing, our correspondence is indeed categorical. □

It is time now to put everyting together. All the above was pure combinatorics, and in relation with the compact groups, we have the following result:

THEOREM 15.5. *Each category of partitions $D=(D(k,l))$ produces a family of compact groups $G=(G_N)$, one for each $N\in\mathbb N$, via the formula*

$$Hom(u^{\otimes k},u^{\otimes l})=span\left(T_\pi\Big|\pi\in D(k,l)\right)$$

and the Tannakian duality correspondence.

PROOF. Given an integer $N\in\mathbb N$, consider the correspondence $\pi\to T_\pi$ constructed in Definition 15.1, and then the collection of linear spaces in the statement, namely:

$$C_{kl}=span\left(T_\pi\Big|\pi\in D(k,l)\right)$$

According to the formulae in Proposition 15.4, and to our axioms for the categories of partitions, from Definition 15.3, this collection of spaces $C=(C_{kl})$ satisfies the axioms for the Tannakian categories, from chapter 14. Thus the Tannakian duality result there applies, and provides us with a closed subgroup $G_N\subset U_N$ such that:

$$C_{kl}=Hom(u^{\otimes k},u^{\otimes l})$$

Thus, we are led to the conclusion in the statement. □

In relation with the easiness property, we can now formulate a key result, which can serve as an alternative definition for the easy groups, as follows:

THEOREM 15.6. *A closed subgroup $G\subset U_N$ is easy precisely when*

$$Hom(u^{\otimes k},u^{\otimes l})=span\left(T_\pi\Big|\pi\in D(k,l)\right)$$

for any colored integers k,l, for a certain category of partitions $D\subset P$.

PROOF. This basically follows from Theorem 15.5, as follows:

(1) In one sense, we know from Theorem 15.5 that any category of partitions $D\subset P$ produces a family of closed groups $G\subset U_N$, one for each $N\in\mathbb N$, according to Tannakian duality and to the Hom space formula there, namely:

$$Hom(u^{\otimes k},u^{\otimes l})=span\left(T_\pi\Big|\pi\in D(k,l)\right)$$

But these groups $G\subset U_N$ are indeed easy, in the sense of Definition 15.2.

(2) In the other sense now, assume that $G\subset U_N$ is easy, in the sense of Definition 15.2, coming via the above Hom space formula, from a collection of sets as follows:

$$D=\bigsqcup_{k,l}D(k,l)$$

Consider now the category of partitions $\widetilde{D}=<D>$ generated by this family. This is by definition the smallest category of partitions containing D, whose existence follows by starting with D, and performing the various categorical operations, namely horizontal and vertical concatenation, and upside-down turning. It follows then, via another application of Tannakian duality, that we have the following formula, for any k,l:

$$Hom(u^{\otimes k},u^{\otimes l})=span\left(T_\pi\Big|\pi\in \widetilde{D}(k,l)\right)$$

Thus, our group $G\subset U_N$ can be viewed as well as coming from $\widetilde{D}$, and so appearing as particular case of the construction in Theorem 15.6, and this gives the result. □

As already mentioned above, Theorem 15.6 can be regarded as an alternative definition for easiness, with the assumption that $D\subset P$ must be a category of partitions being added. In what follows we will rather use this new definition, which is more precise.

Generally speaking, the same comments as before apply. First, G is easy when its Tannakian category appears in the simplest possible way: from a category of partitions. The terminology is quite natural, because Tannakian duality is our only serious tool.

Also, the category of partitions D is not unique, for instance because at $N=1$ all the categories of partitions produce the same easy group, namely $G=\{1\}$. We will be back to this issue on several occasions, with various results about it.

We will see in what follows that many interesting examples of compact quantum groups are easy. Moreover, most of the known series of "basic" compact quantum groups, $G=(G_N)$ with $N\in\mathbb{N}$, can be in principle made fit into some suitable extensions of the easy quantum group formalism. We will discuss this too, in what follows.

The notion of easiness goes back to the results of Brauer in [16] regarding the orthogonal group O_N, and the unitary group U_N, which reformulate as follows:

THEOREM 15.7. *We have the following results:*

(1) *The unitary group U_N is easy, coming from the category $\mathcal{P}_2$.*
(2) *The orthogonal group O_N is easy as well, coming from the category P_2.*

PROOF. This is something that we already know, from chapter 14, based on Tannakian duality, the idea of the proof being as follows:

(1) The group U_N being defined via the relations $u^*=u^{-1}$, $u^t=\bar{u}^{-1}$, the associated Tannakian category is $C=span(T_\pi|\pi\in D)$, with:

$$D=<\overset{\cap}{\circ\bullet}\,,\,\overset{\cap}{\bullet\circ}>=\mathcal{P}_2$$

(2) The group $O_N\subset U_N$ being defined by imposing the relations $u_{ij}=\bar{u}_{ij}$, the associated Tannakian category is $C=span(T_\pi|\pi\in D)$, with:

$$D=<\mathcal{P}_2,\,{}^{\circ}_{\bullet}\!\!\mid,\,{}^{\bullet}_{\circ}\!\!\mid>=P_2$$

Thus, we are led to the conclusion in the statement. $\square$

As already mentioned in the beginning of this chapter, there are many other examples of easy groups, and we will gradually explore this. To start with, we have the following result, dealing with the orthogonal and unitary bistochastic groups B_N,C_N:

THEOREM 15.8. *We have the following results:*

(1) *The unitary bistochastic group C_N is easy, coming from the category $\mathcal{P}_{12}$ of matching singletons and pairings.*
(2) *The orthogonal bistochastic group B_N is easy, coming from the category P_{12} of singletons and pairings.*

PROOF. The proof here is similar to the proof of Theorem 15.7. To be more precise, we can use the results there, and the proof goes as follows:

(1) The group $C_N\subset U_N$ is defined by imposing the following relations, with ξ being the all-one vector, which correspond to the bistochasticity condition:

$$u\xi=\xi\quad,\quad\bar{u}\xi=\xi$$

But these relations tell us precisely that the following two operators, with the partitions on the right being singletons, must be in the associated Tannakian category C:

$$T_\pi\quad:\quad\pi=\underset{\circ}{|}\,,\,\underset{\bullet}{|}$$

Thus the associated Tannakian category is $C = span(T_\pi|\pi \in D)$, with:

$$D =< \mathcal{P}_2, \overset{\circ}{|}, \overset{\bullet}{|} >= \mathcal{P}_{12}$$

Thus, we are led to the conclusion in the statement.

(2) In order to deal now with the real bistochastic group B_N, we can either use a similar argument, or simply use the following intersection formula:

$$B_N = C_N \cap O_N$$

Indeed, at the categorical level, this intersection formula tells us that the associated Tannakian category is given by $C = span(T_\pi|\pi \in D)$, with:

$$D =< \mathcal{P}_{12}, P_2 >= P_{12}$$

Thus, we are led to the conclusion in the statement. $\square$

As a comment here, we have used in the above the fact, which is something quite trivial, that the category of partitions associated to an intersection of easy quantum groups is the intersection of the corresponding categories of partitions. We will be back to this, and to some other product operations as well, with similar results, later on.

We can put now the results that we have together, as follows:

Theorem 15.9. *The basic unitary and bistochastic groups,*

$$\begin{tikzcd} C_N \arrow[r] & U_N \\ B_N \arrow[r] \arrow[u] & O_N \arrow[u] \end{tikzcd}$$

are all easy, coming from the various categories of singletons and pairings.

Proof. We know from the above that the groups in the statement are indeed easy, the corresponding diagram of categories of partitions being as follows:

$$\begin{tikzcd} \mathcal{P}_{12} \arrow[d] & \mathcal{P}_2 \arrow[l] \arrow[d] \\ P_{12} & P_2 \arrow[l] \end{tikzcd}$$

Thus, we are led to the conclusion in the statement. $\square$

Summarizing, what we have so far is a general notion of "easiness", coming from the Brauer theorems for O_N, U_N, and their straightforward extensions to B_N, C_N.

15b. Reflection groups

In view of the above, the notion of easiness is a quite interesting one, deserving a full, systematic investigation. As a first natural question that we would like to solve, we would like to compute the easy group associated to the category of all partitions P itself. And here, no surprise, we are led to the most basic, but non-trivial, classical group that we know, namely the symmetric group S_N. To be more precise, we have the following Brauer type theorem for S_N, which answers our question formulated above:

THEOREM 15.10. *The symmetric group S_N, regarded as group of unitary matrices,*

$$S_N \subset O_N \subset U_N$$

via the permutation matrices, is easy, coming from the category of all partitions P.

PROOF. Consider indeed the group S_N, regarded as a group of unitary matrices, with each permutation $\sigma \in S_N$ corresponding to the associated permutation matrix:

$$\sigma(e_i) = e_{\sigma(i)}$$

Consider as well the easy group $G \subset O_N$ coming from the category of all partitions P. Since P is generated by the one-block "fork" partition $\mu \in P(2,1)$, we have:

$$C(G) = C(O_N) \Big/ \Big\langle T_\mu \in Hom(u^{\otimes 2}, u) \Big\rangle$$

The linear map associated to μ is given by the following formula:

$$T_\mu(e_i \otimes e_j) = \delta_{ij} e_i$$

In order to do the computations, we use the following formulae:

$$u = (u_{ij})_{ij} \quad , \quad u^{\otimes 2} = (u_{ij} u_{kl})_{ik,jl} \quad , \quad T_\mu = (\delta_{ijk})_{i,jk}$$

We therefore obtain the following formula:

$$(T_\mu u^{\otimes 2})_{i,jk} = \sum_{lm} (T_\mu)_{i,lm} (u^{\otimes 2})_{lm,jk} = u_{ij} u_{ik}$$

On the other hand, we have as well the following formula:

$$(u T_\mu)_{i,jk} = \sum_l u_{il} (T_\mu)_{l,jk} = \delta_{jk} u_{ij}$$

Thus, the relation defining $G \subset O_N$ reformulates as follows:

$$T_\mu \in Hom(u^{\otimes 2}, u) \iff u_{ij} u_{ik} = \delta_{jk} u_{ij}, \forall i, j, k$$

In other words, the elements u_{ij} must be projections, which must be pairwise orthogonal on the rows of $u = (u_{ij})$. We conclude that $G \subset O_N$ is the subgroup of matrices $g \in O_N$ having the property $g_{ij} \in \{0, 1\}$. Thus we have $G = S_N$, as desired. $\square$

As a continuation of this, and following the hierarchy of finite groups explained in chapters 9-12, with the original motivations there being actually probabilistic, but with these motivations leading to the same hierarchy that we need here, let us discuss now the hyperoctahedral group H_N. The result here, from [9], is quite similar, as follows:

THEOREM 15.11. *The hyperoctahedral group H_N, regarded as a group of matrices,*

$$S_N \subset H_N \subset O_N$$

is easy, coming from the category of partitions with even blocks P_{even}.

PROOF. This follows as usual from Tannakian duality. To be more precise, consider the following one-block partition, which, as the name indicates, looks like a H letter:

$$H \in P(2,2)$$

The linear map associated to this partition is then given by:

$$T_H(e_i \otimes e_j) = \delta_{ij} e_i \otimes e_i$$

By using this formula, we have the following computation:

$$\begin{aligned}
(T_H \otimes id)u^{\otimes 2}(e_a \otimes e_b) &= (T_H \otimes id)\left(\sum_{ijkl} e_{ij} \otimes e_{kl} \otimes u_{ij}u_{kl}\right)(e_a \otimes e_b)\\
&= (T_H \otimes id)\left(\sum_{ik} e_i \otimes e_k \otimes u_{ia}u_{kb}\right)\\
&= \sum_i e_i \otimes e_i \otimes u_{ia}u_{ib}
\end{aligned}$$

On the other hand, we have as well the following computation:

$$\begin{aligned}
u^{\otimes 2}(T_H \otimes id)(e_a \otimes e_b) &= \delta_{ab}\left(\sum_{ijkl} e_{ij} \otimes e_{kl} \otimes u_{ij}u_{kl}\right)(e_a \otimes e_a)\\
&= \delta_{ab}\sum_{ij} e_i \otimes e_k \otimes u_{ia}u_{ka}
\end{aligned}$$

We conclude from this that we have the following equivalence:

$$T_H \in End(u^{\otimes 2}) \iff \delta_{ik}u_{ia}u_{ib} = \delta_{ab}u_{ia}u_{ka}, \forall i,k,a,b$$

But the relations on the right tell us that the entries of $u = (u_{ij})$ must satisfy $\alpha\beta = 0$ on each row and column of u, and so that the corresponding closed subgroup $G \subset O_N$ consists of the matrices $g \in O_N$ which are permutation-like, with ± 1 nonzero entries. Thus, the corresponding group is $G = H_N$, and as a conclusion to this, we have:

$$C(H_N) = C(O_N)\Big/\Big\langle T_H \in End(u^{\otimes 2})\Big\rangle$$

According to our conventions, this means that the hyperoctahedral group H_N is easy, coming from the following category of partitions:

$$D=<H>$$

But the category on the right can be computed by drawing pictures, and we have:

$$<H>=P_{even}$$

Thus, we are led to the conclusion in the statement. □

More generally now, we have in fact the following result, from [8], regarding the series of complex reflection groups H_N^s, which covers both the groups S_N, H_N:

THEOREM 15.12. *The complex reflection group $H_N^s=\mathbb{Z}_s\wr S_N$ is easy, the corresponding category P^s consisting of the partitions satisfying the condition*

$$\#\circ=\#\bullet\,(s)$$

as a weighted sum, in each block. In particular, we have the following results:

(1) *S_N is easy, coming from the category P.*
(2) *$H_N=\mathbb{Z}_2\wr S_N$ is easy, coming from the category P_{even}.*
(3) *$K_N=\mathbb{T}\wr S_N$ is easy, coming from the category $\mathcal{P}_{even}$.*

PROOF. This is something that we already know at $s=1,2$, from Theorems 15.10 and 15.11. In general, the proof is similar, based on Tannakian duality. To be more precise, in what regards the main assertion, the idea here is that the one-block partition $\pi\in P(s)$, which generates the category of partitions P^s in the statement, implements the relations producing the subgroup $H_N^s\subset S_N$. As for the last assertions, these are all elementary:

(1) At $s=1$ we know that we have $H_N^1=S_N$. Regarding now the corresponding category, here the condition $\#\circ=\#\bullet\,(1)$ is automatic, and so $P^1=P$.

(2) At $s=2$ we know that we have $H_N^2=H_N$. Regarding now the corresponding category, here the condition $\#\circ=\#\bullet\,(2)$ reformulates as follows:

$$\#\circ+\#\bullet=0(2)$$

Thus each block must have even size, and we obtain, as claimed, $P^2=P_{even}$.

(3) At $s=\infty$ we know that we have $H_N^\infty=K_N$. Regarding now the corresponding category, here the condition $\#\circ=\#\bullet\,(\infty)$ reads:

$$\#\circ=\#\bullet$$

But this is the condition defining $\mathcal{P}_{even}$, and so $P^\infty=\mathcal{P}_{even}$, as claimed. □

Summarizing, we have many examples. In fact, our list of easy groups has currently become quite big, and here is a selection of the main results that we have so far:

THEOREM 15.13. *We have a diagram of compact groups as follows,*

$$\begin{tikzcd} K_N \arrow[r] & U_N \\ H_N \arrow[u] \arrow[r] & O_N \arrow[u] \end{tikzcd}$$

where $H_N = \mathbb{Z}_2 \wr S_N$ and $K_N = \mathbb{T} \wr S_N$, and all these groups are easy.

PROOF. This follows from the above results. To be more precise, we know that the above groups are all easy, the corresponding categories of partitions being as follows:

$$\begin{tikzcd} \mathcal{P}_{even} \arrow[d] & \mathcal{P}_2 \arrow[l] \arrow[d] \\ P_{even} & P_2 \arrow[l] \end{tikzcd}$$

Thus, we are led to the conclusion in the statement. □

Summarizing, most of the groups that we investigated in this book are covered by the easy group formalism. One exception is the symplectic group Sp_N, but we will see later that this group is covered as well, by a certain extension of the easy group formalism.

15c. Basic operations

We develop now some general abstract theory for the easy groups. Let us first discuss some basic composition operations. We will be mainly interested in:

DEFINITION 15.14. *The closed subgroups of U_N are subject to intersection and generation operations, constructed as follows:*

(1) *Intersection: $H \cap K$ is the usual intersection of H, K.*
(2) *Generation: $< H, K >$ is the closed subgroup generated by H, K.*

Alternatively, we can define these operations at the function algebra level, by performing certain operations on the associated ideals, as follows:

PROPOSITION 15.15. *Assuming that we have presentation results as follows,*

$$C(H) = C(U_N)/I \quad , \quad C(K) = C(U_N)/J$$

the groups $H \cap K$ and $< H, K >$ are given by the following formulae,

$$C(H \cap K) = C(U_N)/ < I, J >$$

$$C(< H, K >) = C(U_N)/(I \cap J)$$

at the level of the associated algebras of functions.

Proof. This is indeed clear from the definition of the operations $\cap$ and $<,>$, as formulated above, and from the Stone-Weierstrass theorem. □

In what follows we will need Tannakian formulations of the above two operations. The result here, coming from the duality result established in chapter 14, is as follows:

Theorem 15.16. *The intersection and generation operations $\cap$ and $<,>$ can be constructed via the Tannakian correspondence $G \to C_G$, as follows:*

1. *Intersection: defined via $C_{G\cap H} = < C_G, C_H >$.*
2. *Generation: defined via $C_{<G,H>} = C_G \cap C_H$.*

Proof. This follows from Proposition 15.15, and from Tannakian duality. Indeed, it follows from Tannakian duality that given a closed subgroup $G \subset U_N$, with fundamental representation v, the algebra of functions $C(G)$ has the following presentation:

$$C(G) = C(U_N) \Big/ \left\langle T \in Hom(u^{\otimes k}, u^{\otimes l}) \Big| \forall k, \forall l, \forall T \in Hom(v^{\otimes k}, v^{\otimes l}) \right\rangle$$

In other words, given a closed subgroup $G \subset U_N$, we have a presentation of the following type, with I_G being the ideal coming from the Tannakian category of G:

$$C(G) = C(U_N)/I_G$$

But this leads to the conclusion in the statement. □

In relation now with our easiness questions, we first have the following result:

Proposition 15.17. *Assuming that H, K are easy, then so is $H \cap K$, and we have*

$$D_{H\cap K} = < D_H, D_K >$$

at the level of the corresponding categories of partitions.

Proof. We have indeed the following computation:

$$\begin{aligned} C_{H\cap K} &= < C_H, C_K > \\ &= < span(D_H), span(D_K) > \\ &= span(< D_H, D_K >) \end{aligned}$$

Thus, by Tannakian duality we obtain the result. □

Regarding now the generation operation, the situation here is more complicated, due to a number of technical reasons, and we have the following statement:

Proposition 15.18. *Assuming that H, K are easy, we have an inclusion*

$$< H, K > \subset \{H, K\}$$

coming from an inclusion of Tannakian categories as follows,

$$C_H \cap C_K \supset span(D_H \cap D_K)$$

where $\{H, K\}$ is the easy group having as category of partitions $D_H \cap D_K$.

PROOF. This follows from the definition and properties of the generation operation, explained above, and from the following computation:

$$\begin{aligned} C_{<H,K>} &= C_H \cap C_K \\ &= span(D_H) \cap span(D_K) \\ &\supset span(D_H \cap D_K) \end{aligned}$$

Indeed, by Tannakian duality we obtain from this all the assertions. □

It is not clear if the inclusions in Proposition 15.18 are isomorphisms or not, and this even under a supplementary $N >> 0$ assumption. Technically speaking, the problem comes from the fact that the operation $\pi \to T_\pi$ does not produce linearly independent maps, and so all that we are doing is sensitive to the value of $N \in \mathbb{N}$. The subject here is quite technical, to be further developed in chapter 16 below, with probabilistic motivations in mind, without however solving the present algebraic questions.

Summarizing, we have some problems here, and we must proceed as follows:

THEOREM 15.19. *The intersection and easy generation operations $\cap$ and $\{\,,\}$ can be constructed via the Tannakian correspondence $G \to D_G$, as follows:*

(1) *Intersection: defined via $D_{G\cap H} =< D_G, D_H >$.*
(2) *Easy generation: defined via $D_{\{G,H\}} = D_G \cap D_H$.*

PROOF. Here the situation is as follows:

(1) This is a true and honest result, coming from Proposition 15.17.

(2) This is more of an empty statement, coming from Proposition 15.18. □

As already mentioned, there is some interesting mathematics still to be worked out, in relation with all this, and we will be back to this later, with further details. With the above notions in hand, however, even if not fully satisfactory, we can formulate a nice result, which improves our main result so far, namely Theorem 15.13, as follows:

THEOREM 15.20. *The basic unitary and reflection groups, namely*

$$\begin{tikzcd} K_N \arrow[r] & U_N \\ H_N \arrow[u] \arrow[r] & O_N \arrow[u] \end{tikzcd}$$

are all easy, and they form an intersection and easy generation diagram, in the sense that the above square diagram satisfies $U_N = \{K_N, O_N\}$, and $H_N = K_N \cap O_N$.

PROOF. We know from Theorem 15.13 that the groups in the statement are easy, the corresponding categories of partitions being as follows:

$$\begin{tikzcd}
\mathcal{P}_{even} \arrow[d] & \mathcal{P}_2 \arrow[l] \arrow[d] \\
P_{even} & P_2 \arrow[l]
\end{tikzcd}$$

Now observe that this latter diagram is an intersection and generation diagram. By using Theorem 15.19, this reformulates into the fact that the diagram of quantum groups is an intersection and easy generation diagram, as claimed. □

It is possible to further improve the above result, by proving that the diagram there is actually a plain generation diagram. However, this is something more technical, and for a discussion here, we refer for instance to the quantum group book [6].

Let us also mention that it is possible to develop some more general theory, at this level. Given a closed subgroup $G \subset U_N$, we can talk about its "easy envelope", which is the smallest easy quantum group $\widetilde{G}$ containing G, which is easy. This easy envelope appears by definition as an intermediate closed subgroup, as follows:

$$G \subset \widetilde{G} \subset U_N$$

With this notion in hand, Proposition 15.18 can be refined into a result stating that given two easy groups H, K, we have inclusions as follows:

$$< H, K > \subset < \widetilde{H, K} > \subset \{H, K\}$$

In order to discuss all this, let us start with the following definition:

DEFINITION 15.21. *A closed subgroup $G \subset U_N$ is called homogeneous when*

$$S_N \subset G \subset U_N$$

with $S_N \subset U_N$ being the standard embedding, via permutation matrices.

We will be interested in such groups, which cover for instance all the easy groups, and many more. At the Tannakian level, we have the following result:

THEOREM 15.22. *The homogeneous groups $S_N \subset G \subset U_N$ are in one-to-one correspondence with the intermediate tensor categories*

$$span\left(T_\pi\Big|\pi \in \mathcal{P}_2\right) \subset C \subset span\left(T_\pi\Big|\pi \in P\right)$$

where P is the category of all partitions, $\mathcal{P}_2$ is the category of the matching pairings, and $\pi \to T_\pi$ is the standard implementation of partitions, as linear maps.

Proof. This follows from Tannakian duality, and from the Brauer type results for S_N, U_N. To be more precise, we know from Tannakian duality that each closed subgroup $G\subset U_N$ can be reconstructed from its Tannakian category $C=(C(k,l))$, as follows:

$$C(G)=C(U_N)\Big/\left\langle T\in Hom(u^{\otimes k},u^{\otimes l})\Big|\forall k,l,\forall T\in C(k,l)\right\rangle$$

Thus we have a one-to-one correspondence $G\leftrightarrow C$, given by Tannakian duality, and since the endpoints $G=S_N,U_N$ are both easy, corresponding to the categories $C=span(T_\pi|\pi\in D)$ with $D=P,\mathcal{P}_2$, this gives the result. □

Our purpose now will be that of using the Tannakian result in Theorem 15.22, in order to introduce and study a combinatorial notion of "easiness level", for the arbitrary intermediate groups $S_N\subset G\subset U_N$. Let us begin with the following simple fact:

Proposition 15.23. *Given a homogeneous group $S_N\subset G\subset U_N$, with associated Tannakian category $C=(C(k,l))$, the sets*

$$D^1(k,l)=\left\{\pi\in P(k,l)\Big|T_\pi\in C(k,l)\right\}$$

form a category of partitions, in the sense of Definition 15.3.

Proof. We use the basic categorical properties of the correspondence $\pi\to T_\pi$ between partitions and linear maps, that we established in the above, namely:

$$T_{[\pi\sigma]}=T_\pi\otimes T_\sigma\quad,\quad T_{[^\sigma_\pi]}\sim T_\pi T_\sigma\quad,\quad T_{\pi^*}=T_\pi^*$$

Together with the fact that C is a tensor category, we deduce from these formulae that we have the following implication:

$$\begin{aligned}\pi,\sigma\in D^1&\implies T_\pi,T_\sigma\in C\\&\implies T_\pi\otimes T_\sigma\in C\\&\implies T_{[\pi\sigma]}\in C\\&\implies [\pi\sigma]\in D^1\end{aligned}$$

We have as well the following implication:

$$\begin{aligned}\pi,\sigma\in D^1&\implies T_\pi,T_\sigma\in C\\&\implies T_\pi T_\sigma\in C\\&\implies T_{[^\sigma_\pi]}\in C\\&\implies [^\sigma_\pi]\in D^1\end{aligned}$$

Finally, we have as well the following implication:

$$\begin{aligned}\pi \in D^1 &\implies T_\pi \in C\\ &\implies T_\pi^* \in C\\ &\implies T_{\pi^*} \in C\\ &\implies \pi^* \in D^1\end{aligned}$$

Thus D^1 is indeed a category of partitions, as claimed. □

We can further refine the above observation, in the following way:

PROPOSITION 15.24. *Given a compact group $S_N \subset G \subset U_N$, construct $D^1 \subset P$ as above, and let $S_N \subset G^1 \subset U_N$ be the easy group associated to D^1. Then:*

(1) *We have $G \subset G^1$, as subgroups of U_N.*
(2) *G^1 is the smallest easy group containing G.*
(3) *G is easy precisely when $G \subset G^1$ is an isomorphism.*

PROOF. All this is elementary, the proofs being as follows:

(1) We know that the Tannakian category of G^1 is given by:

$$C_{kl}^1 = span\left(T_\pi \Big| \pi \in D^1(k,l)\right)$$

Thus we have $C^1 \subset C$, and so $G \subset G^1$, as subgroups of U_N.

(2) Assuming that we have $G \subset G'$, with G' easy, coming from a Tannakian category $C' = span(D')$, we must have $C' \subset C$, and so $D' \subset D^1$. Thus, $G^1 \subset G'$, as desired.

(3) This is a trivial consequence of (2). □

Summarizing, we have now a notion of "easy envelope", as follows:

DEFINITION 15.25. *The easy envelope of a homogeneous group $S_N \subset G \subset U_N$ is the easy group $S_N \subset G^1 \subset U_N$ associated to the category of partitions*

$$D^1(k,l) = \left\{\pi \in P(k,l) \Big| T_\pi \in C(k,l)\right\}$$

where $C = (C(k,l))$ is the Tannakian category of G.

At the level of the examples, most of the known homogeneous groups $S_N \subset G \subset U_N$ are in fact easy. However, there are many non-easy examples as well, and we will compute the corresponding easy envelopes in several cases of interest, later on.

As a technical observation now, we can in fact generalize the above construction to any closed subgroup $G \subset U_N$, and we have the following result:

PROPOSITION 15.26. *Given a closed subgroup $G \subset U_N$, construct $D^1 \subset P$ as above, and let $S_N \subset G^1 \subset U_N$ be the easy group associated to D^1. We have then*

$$G^1 = (< G, S_N >)^1$$

where $< G, S_N > \subset U_N$ is the smallest closed subgroup containing G, S_N.

PROOF. It is well-known, and elementary to show, using Tannakian duality, that the smallest subgroup $< G, S_N > \subset U_N$ from the statement exists indeed, and can be obtained by intersecting the Tannakian categories of G, S_N, as follows:

$$C_{<G,S_N>} = C_G \cap C_{S_N}$$

We conclude from this that for any $\pi \in P(k,l)$ we have:

$$T_\pi \in C_{<G,S_N>}(k,l) \iff T_\pi \in C_G(k,l)$$

It follows that the D^1 categories for the groups $< G, S_N >$ and G coincide, and so the easy envelopes $(< G, S_N >)^1$ and G^1 coincide as well, as stated. $\square$

In order now to fine-tune all this, by using an arbitrary parameter $p \in \mathbb{N}$, which can be thought of as being an "easiness level", we can proceed as follows:

DEFINITION 15.27. *Given a compact group $S_N \subset G \subset U_N$, and an integer $p \in \mathbb{N}$, we construct the family of linear spaces*

$$E^p(k,l) = \left\{ \alpha_1 T_{\pi_1} + \ldots + \alpha_p T_{\pi_p} \in C(k,l) \Big| \alpha_i \in \mathbb{C}, \pi_i \in P(k,l) \right\}$$

and we denote by C^p the smallest tensor category containing $E^p = (E^p(k,l))$, and by $S_N \subset G^p \subset U_N$ the compact group corresponding to this category C^p.

As a first observation, at $p = 1$ we have $C^1 = E^1 = span(D^1)$, where D^1 is the category of partitions constructed in Proposition 15.24. Thus the group G^1 constructed above coincides with the "easy envelope" of G, from Definition 15.25 above.

In the general case, $p \in \mathbb{N}$, the family $E^p = (E^p(k,l))$ constructed above is not necessarily a tensor category, but we can of course consider the tensor category C^p generated by it, as indicated. Finally, in the above definition we have used of course the Tannakian duality results, in order to perform the operation $C^p \to G^p$.

In practice, the construction in Definition 15.27 is often something quite complicated, and it is convenient to use the following observation:

PROPOSITION 15.28. *The category C^p constructed above is generated by the spaces*

$$E^p(l) = \left\{ \alpha_1 T_{\pi_1} + \ldots + \alpha_p T_{\pi_p} \in C(l) \Big| \alpha_i \in \mathbb{C}, \pi_i \in P(l) \right\}$$

where $C(l) = C(0,l), P(l) = P(0,l)$, with l ranging over the colored integers.

PROOF. We use the well-known fact that given a closed subgroup $G \subset U_N$, we have a Frobenius type isomorphism, as follows:

$$Hom(u^{\otimes k}, u^{\otimes l}) \simeq Fix(u^{\otimes \bar{k}l})$$

If we apply this to the group G^p, we obtain an isomorphism as follows:

$$C(k,l) \simeq C(\bar{k}l)$$

On the other hand, we have as well an isomorphism $P(k,l) \simeq P(\bar{k}l)$, obtained by performing a counterclockwise rotation to the partitions $\pi \in P(k,l)$. According to the above definition of the spaces $E^p(k,l)$, this induces an isomorphism as follows:

$$E^p(k,l) \simeq E^p(\bar{k}l)$$

We deduce from this that for any partitions $\pi_1, \ldots, \pi_p \in C(k,l)$, having rotated versions $\rho_1, \ldots, \rho_p \in C(\bar{k}l)$, and for any scalars $\alpha_1, \ldots, \alpha_p \in \mathbb{C}$, we have:

$$\alpha_1 T_{\pi_1} + \ldots + \alpha_p T_{\pi_p} \in C(k,l) \iff \alpha_1 T_{\rho_1} + \ldots + \alpha_p T_{\rho_p} \in C(\bar{k}l)$$

But this gives the conclusion in the statement, and we are done. □

The main properties of the construction $G \to G^p$ can be summarized as follows:

THEOREM 15.29. *Given a compact group $S_N \subset G \subset U_N$, the compact groups G^p constructed above form a decreasing family, whose intersection is G:*

$$G = \bigcap_{p \in \mathbb{N}} G^p$$

Moreover, G is easy when this decreasing limit is stationary, $G = G^1$.

PROOF. By definition of $E^p(k,l)$, and by using Proposition 15.28, these linear spaces form an increasing filtration of $C(k,l)$. The same remains true when completing into tensor categories, and so we have an increasing filtration, as follows:

$$C = \bigcup_{p \in \mathbb{N}} C^p$$

At the compact group level now, we obtain the decreasing intersection in the statement. Finally, the last assertion is clear from Proposition 15.28. □

As a main consequence of the above results, we can now formulate:

DEFINITION 15.30. *We say that a homogeneous compact group*

$$S_N \subset G \subset U_N$$

is easy at order p when $G = G^p$, with p being chosen minimal with this property.

Observe that the order 1 notion corresponds to the usual easiness. In general, all this is quite abstract, but there are several explicit examples, that can be worked out. For more on all this, we refer to the quantum group book [6].

15d. Classification results

Let us go back now to plain easiness, and discuss some classification results, following [13], and then the papers of Raum-Weber [72] and Tarrago-Weber [83]. In order to cut from the complexity, we must impose an extra axiom, and we will use here:

THEOREM 15.31. *For an easy group $G = (G_N)$, coming from a category of partitions $D \subset P$, the following conditions are equivalent:*

1. *$G_{N-1} = G_N \cap U_{N-1}$, via the embedding $U_{N-1} \subset U_N$ given by $u \to diag(u, 1)$.*
2. *$G_{N-1} = G_N \cap U_{N-1}$, via the N possible diagonal embeddings $U_{N-1} \subset U_N$.*
3. *D is stable under the operation which consists in removing blocks.*

If these conditions are satisfied, we say that $G = (G_N)$ is uniform.

PROOF. We use the general easiness theory explained above, as follows:

(1) $\iff$ (2) This is something standard, coming from the inclusion $S_N \subset G_N$, which makes everything S_N-invariant. The result follows as well from the proof of (1) $\iff$ (3) below, which can be converted into a proof of (2) $\iff$ (3), in the obvious way.

(1) $\iff$ (3) Given a subgroup $K \subset U_{N-1}$, with fundamental representation u, consider the $N \times N$ matrix $v = diag(u, 1)$. Our claim is that for any $\pi \in P(k)$ we have:

$$\xi_\pi \in Fix(v^{\otimes k}) \iff \xi_{\pi'} \in Fix(v^{\otimes k'}), \forall \pi' \in P(k'), \pi' \subset \pi$$

In order to prove this, we must study the condition on the left. We have:

$$\begin{aligned}
\xi_\pi \in Fix(v^{\otimes k}) &\iff (v^{\otimes k}\xi_\pi)_{i_1\dots i_k} = (\xi_\pi)_{i_1\dots i_k}, \forall i \\
&\iff \sum_j (v^{\otimes k})_{i_1\dots i_k, j_1\dots j_k}(\xi_\pi)_{j_1\dots j_k} = (\xi_\pi)_{i_1\dots i_k}, \forall i \\
&\iff \sum_j \delta_\pi(j_1, \dots, j_k) v_{i_1 j_1} \dots v_{i_k j_k} = \delta_\pi(i_1, \dots, i_k), \forall i
\end{aligned}$$

Now let us recall that our representation has the special form $v = diag(u, 1)$. We conclude from this that for any index $a \in \{1, \dots, k\}$, we must have:

$$i_a = N \implies j_a = N$$

With this observation in hand, if we denote by i', j' the multi-indices obtained from i, j obtained by erasing all the above $i_a = j_a = N$ values, and by $k' \leq k$ the common length of these new multi-indices, our condition becomes:

$$\sum_{j'} \delta_\pi(j_1, \dots, j_k)(v^{\otimes k'})_{i'j'} = \delta_\pi(i_1, \dots, i_k), \forall i$$

Here the index j is by definition obtained from j' by filling with N values. In order to finish now, we have two cases, depending on i, as follows:

Case 1. Assume that the index set $\{a|i_a = N\}$ corresponds to a certain subpartition $\pi' \subset \pi$. In this case, the N values will not matter, and our formula becomes:

$$\sum_{j'} \delta_\pi(j'_1, \ldots, j'_{k'})(v^{\otimes k'})_{i'j'} = \delta_\pi(i'_1, \ldots, i'_{k'})$$

Case 2. Assume now the opposite, namely that the set $\{a|i_a = N\}$ does not correspond to a subpartition $\pi' \subset \pi$. In this case the indices mix, and our formula reads:

$$0 = 0$$

Thus, we are led to $\xi_{\pi'} \in Fix(v^{\otimes k'})$, for any subpartition $\pi' \subset \pi$, as claimed.

Now with this claim in hand, the result follows from Tannakian duality. $\square$

We can now formulate a first classification result, as follows:

THEOREM 15.32. *The uniform orthogonal easy groups are as follows,*

$$\begin{tikzcd} B_N \arrow[r] & O_N \\ S_N \arrow[u] \arrow[r] & H_N \arrow[u] \end{tikzcd}$$

and this diagram is an intersection and easy generation diagram.

PROOF. We know that the quantum groups in the statement are indeed easy and uniform, the corresponding categories of partitions being as follows:

$$\begin{tikzcd} P_{12} \arrow[d] & P_2 \arrow[l] \arrow[d] \\ P & P_{even} \arrow[l] \end{tikzcd}$$

Since this latter diagram is an intersection and generation diagram, we conclude that we have an intersection and easy generation diagram of quantum groups, as stated. Regarding now the classification, consider an arbitrary easy group, as follows:

$$S_N \subset G_N \subset O_N$$

This group must then come from a category of partitions, as follows:

$$P_2 \subset D \subset P$$

Now if we assume $G = (G_N)$ to be uniform, this category of partitions D is uniquely determined by the subset $L \subset \mathbb{N}$ consisting of the sizes of the blocks of the partitions in D. Following [13], our claim is that the admissible sets are as follows:

(1) $L=\{2\}$, producing O_N.

(2) $L=\{1,2\}$, producing B_N.

(3) $L=\{2,4,6,\ldots\}$, producing H_N.

(4) $L=\{1,2,3,\ldots\}$, producing S_N.

Indeed, in one sense, this follows from our easiness results for O_N,B_N,H_N,S_N. In the other sense now, assume that $L\subset\mathbb N$ is such that the set P_L consisting of partitions whose sizes of the blocks belong to L is a category of partitions. We know from the axioms of the categories of partitions that the semicircle $\cap$ must be in the category, so we have $2\in L$. Our claim is that the following conditions must be satisfied as well:

$$k,l\in L,\ k>l\implies k-l\in L$$

$$k\in L,\ k\geq 2\implies 2k-2\in L$$

Indeed, we will prove that both conditions follow from the axioms of the categories of partitions. Let us denote by $b_k\in P(0,k)$ the one-block partition, as follows:

$$b_k=\begin{Bmatrix}\sqcap\!\!\sqcap&\ldots&\sqcap\\1\ 2&\ldots&k\end{Bmatrix}$$

For $k>l$, we can write b_{k-l} in the following way:

$$b_{k-l}=\left\{\begin{matrix}\sqcap\!\!\sqcap&\ldots&\ldots&\ldots&\ldots&\sqcap\\1\ 2&\ldots&l&l+1&\ldots&k\\\sqcup\!\!\sqcup&\ldots&\sqcup&|&\ldots&|\\&&&1&\ldots&k-l\end{matrix}\right\}$$

In other words, we have the following formula:

$$b_{k-l}=(b_l^*\otimes|^{\otimes k-l})b_k$$

Since all the terms of this composition are in P_L, we have $b_{k-l}\in P_L$, and this proves our first formula. As for the second formula, this can be proved in a similar way, by capping two adjacent k-blocks with a 2-block, in the middle.

With the above two formulae in hand, we can conclude in the following way:

<u>Case 1</u>. Assume $1\in L$. By using the first formula with $l=1$ we get:

$$k\in L\implies k-1\in L$$

This condition shows that we must have $L=\{1,2,\ldots,m\}$, for a certain number $m\in\{1,2,\ldots,\infty\}$. On the other hand, by using the second formula we get:

$$\begin{aligned}m\in L&\implies 2m-2\in L\\&\implies 2m-2\leq m\\&\implies m\in\{1,2,\infty\}\end{aligned}$$

The case $m = 1$ being excluded by the condition $2 \in L$, we reach to one of the two sets producing the groups S_N, B_N.

Case 2. Assume $1 \notin L$. By using the first formula with $l = 2$ we get:

$$k \in L \implies k - 2 \in L$$

This condition shows that we must have $L = \{2, 4, \ldots, 2p\}$, for a certain number $p \in \{1, 2, \ldots, \infty\}$. On the other hand, by using the second formula we get:

$$\begin{aligned} 2p \in L &\implies 4p - 2 \in L \\ &\implies 4p - 2 \leq 2p \\ &\implies p \in \{1, \infty\} \end{aligned}$$

Thus L must be one of the two sets producing O_N, H_N, and we are done. $\square$

When lifting the uniformity assumption in all the above, things become more complicated, and the final classification results become more technical, due to the presence of various copies of $\mathbb{Z}_2$, that can be added, as for keeping the easiness property true. To be more precise, in the real case, as explained in [**13**], we have exactly 6 solutions, which are as follows, with the convention $G'_N = G_N \times \mathbb{Z}_2$:

$$\begin{tikzcd} B_N \arrow[r] & B'_N \arrow[r] & O_N \\ S_N \arrow[u] \arrow[r] & S'_N \arrow[u] \arrow[r] & H_N \arrow[u] \end{tikzcd}$$

In the unitary case now, the classification is quite similar, but a bit more complicated, as explained in the paper of Tarrago-Weber [**83**]. In particular we have:

THEOREM 15.33. *The uniform easy groups which are purely unitary, in the sense that they appear as complexifications of real easy groups, are as follows,*

$$\begin{tikzcd} C_N \arrow[r] & U_N \\ S_N \arrow[u] \arrow[r] & K_N \arrow[u] \end{tikzcd}$$

and this diagram is an intersection and easy generation diagram.

PROOF. We know from the above that the groups in the statement are indeed easy and uniform, the corresponding categories of partitions being as follows:

$$\begin{tikzcd}
\mathcal{P}_{12} \arrow[d] & \mathcal{P}_2 \arrow[l] \arrow[d] \\
P & \mathcal{P}_{even} \arrow[l]
\end{tikzcd}$$

Since this latter diagram is an intersection and generation diagram, we conclude that we have an intersection and easy generation diagram of groups, as stated. As for the uniqueness result, the proof here is similar to the proof from the real case, from Theorem 15.32, by examining the possible sizes of the blocks of the partitions in the category, and doing some direct combinatorics. See Tarrago-Weber [83]. □

Finally, let us mention that the easy quantum group formalism can be extended into a "super-easy" group formalism, covering as well the symplectic group Sp_N. This is something a bit technical, and we refer here to the paper of Collins-Śniady [23].

15e. Exercises

In relation with the notion of easy envelope, which was introduced in the above, we have some explicit computations to be done, as follows:

EXERCISE 15.34. *Compute the easy envelope of the general complex reflection groups*

$$H_N^{sd} = \left\{ U \in H_N^s \Big| (\det U)^d = 1 \right\}$$

and of other closed subgroups of U_N that you know, that are not easy.

This is something which does not look very difficult, and you have the choice here, either by using combinatorics, or the universality property of the easy envelope.

EXERCISE 15.35. *Work out the super-easiness property of the symplectic group*

$$Sp_N \subset U_N$$

defined for $N \in \mathbb{N}$ even, then try as well the groups SU_2 and SO_3.

This is actually a quite difficult exercise, and for Sp_N you can of course look it up, if you don't find the answer. As for SU_2 and SO_3, you're into advanced mathematics here, and good luck with this. Also, as a bonus exercise, which brings you into hot research topics, you can learn as well about the exceptional Lie groups, and then try to understand what is the suitable extension of the easiness property, covering them.

Exercise 15.36. *Prove that when lifting the uniformity assumption, the groups*

$$\begin{tikzcd}
B_N \arrow[r] & B'_N \arrow[r] & O_N \\
S_N \arrow[r] \arrow[u] & S'_N \arrow[r] \arrow[u] & H_N \arrow[u]
\end{tikzcd}$$

with the convention $G'_N = G_N \times \mathbb{Z}_2$, are the only easy real groups.

This is something quite standard, briefly discussed in the above, the idea being that of adapting the proof of the classification from the real uniform case.

Exercise 15.37. *Prove that the uniform, purely unitary easy groups are*

$$\begin{tikzcd}
C_N \arrow[r] & U_N \\
S_N \arrow[r] \arrow[u] & K_N \arrow[u]
\end{tikzcd}$$

with a suitable definition for the notion of pure unitarity.

As before, this is something quite standard, briefly discussed in the above, the idea being that of adapting the proof of the classification from the real uniform case.

Exercise 15.38. *Learn a bit about quantum groups, and about easy quantum groups, as to understand the definition of the main* 8 *easy quantum groups, namely*

$$\begin{tikzcd}
 & K_N^+ \arrow[rr] & & U_N^+ \\
H_N^+ \arrow[rr] \arrow[ur] & & O_N^+ \arrow[ur] & \\
 & K_N \arrow[rr] \arrow[uu] & & U_N \arrow[uu] \\
H_N \arrow[rr] \arrow[ur] \arrow[uu] & & O_N \arrow[ur] \arrow[uu] &
\end{tikzcd}$$

and the Ground Zero theorem in quantum groups, stating that under suitable, strong combinatorial assumptions, these are the only 8 *quantum groups.*

This is something quite interesting, providing us with some “orientation” inside quantum algebra, with the 3 coordinate directions corresponding to the real/complex, discrete/continuous and classical/free dichotomies.

CHAPTER 16

Weingarten calculus

16a. Weingarten formula

Time to put everything together. We discuss here applications of the theory developed in chapters 13-15, to the computation of the laws of characters, and truncated characters, as to solve the various questions left open in chapters 9-12, for the continuous groups. Generally speaking, all these questions require a good knowledge of the integration over G, and more precisely, of the various polynomial integrals over G, defined as follows:

Definition 16.1. *Given a closed subgroup $G \subset U_N$, the quantities*

$$I_k = \int_G g_{i_1 j_1}^{e_1} \ldots g_{i_k j_k}^{e_k} \, dg$$

depending on a colored integer $k = e_1 \ldots e_k$, are called polynomial integrals over G.

As a first observation, the knowledge of these integrals is the same as the full knowledge of the integration functional over G. Indeed, since the coordinate functions $g \to g_{ij}$ separate the points of G, we can apply the Stone-Weierstrass theorem, and we obtain:

$$C(G) = < g_{ij} >$$

Thus, by linearity, the computation of any functional $f : C(G) \to \mathbb{C}$, and in particular of the integration functional, reduces to the computation of this functional on the polynomials of the coordinate functions $g \to g_{ij}$ and their conjugates $g \to \bar{g}_{ij}$.

The point now is that, by using Peter-Weyl, everything reduces to linear algebra, and more specifically to a matrix inversion question, due to the following result:

Theorem 16.2. *The Haar integration over a closed subgroup $G \subset U_N$ is given on the dense subalgebra of smooth functions by the Weingarten type formula*

$$\int_G g_{i_1 j_1}^{e_1} \ldots g_{i_k j_k}^{e_k} \, dg = \sum_{\pi,\sigma \in D_k} \delta_\pi(i) \delta_\sigma(j) W_k(\pi, \sigma)$$

valid for any colored integer $k = e_1 \ldots e_k$ and any multi-indices i, j, where D_k is a linear basis of $Fix(u^{\otimes k})$, the associated generalized Kronecker symbols are given by

$$\delta_\pi(i) = < \pi, e_{i_1} \otimes \ldots \otimes e_{i_k} >$$

and $W_k = G_k^{-1}$ is the inverse of the Gram matrix, $G_k(\pi, \sigma) = < \pi, \sigma >$.

PROOF. We know from Peter-Weyl theory that the integrals in the statement form altogether the orthogonal projection P^k onto the following space:

$$Fix(u^{\otimes k}) = span(D_k)$$

Consider now the following linear map, with $D_k = \{\xi_k\}$ being as in the statement:

$$E(x) = \sum_{\pi\in D_k} < x, \xi_\pi > \xi_\pi$$

By a standard linear algebra computation, it follows that we have $P = WE$, where W is the inverse of the restriction of E to the following space:

$$K = span\left(T_\pi\Big|\pi \in D_k\right)$$

But this restriction is the linear map given by the matrix G_k, and so W is the linear map given by the inverse matrix $W_k = G_k^{-1}$, and this gives the result. □

In the easy case, we have the following more precise result:

THEOREM 16.3. *For an easy group $G \subset U_N$, coming from a category of partitions $D = (D(k,l))$, we have the Weingarten integration formula*

$$\int_G u_{i_1j_1}^{e_1}\ldots u_{i_kj_k}^{e_k} = \sum_{\pi,\sigma\in D(k)} \delta_\pi(i)\delta_\sigma(j)W_{kN}(\pi,\sigma)$$

for any multi-indices i, j and any exponent $k = e_1 \ldots e_k$, where $D(k) = D(\emptyset, k)$, the δ numbers are the usual Kronecker type symbols, and $W_{kN} = G_{kN}^{-1}$, with

$$G_{kN}(\pi,\sigma) = N^{|\pi\vee\sigma|}$$

where $|.|$ is the number of blocks.

PROOF. We use the abstract Weingarten formula, from Theorem 16.2. According to our easiness conventions, the Kronecker symbols are given by:

$$\begin{aligned}
\delta_{\xi_\pi}(i) &= < \xi_\pi, e_{i_1} \otimes \ldots \otimes e_{i_k} > \\
&= \left\langle \sum_j \delta_\pi(j_1,\ldots,j_k)e_{j_1} \otimes \ldots \otimes e_{j_k}, e_{i_1} \otimes \ldots \otimes e_{i_k} \right\rangle \\
&= \delta_\pi(i_1,\ldots,i_k)
\end{aligned}$$

The Gram matrix being as well the correct one, we obtain the result. □

Generally speaking, the above result is something quite powerful, because the main computation there, that of the inverse matrix $W_{kN} = G_{kN}^1$, can be run on an ordinary laptop, after implementing the formula of the Gram matrix, namely $G_{kN}(\pi,\sigma) = N^{|\pi\vee\sigma|}$, which is something quite easy to do. Thus, you can prove theorems about integrals over easy groups just by smoking cigars, and letting your computer do the work.

Let us also mention that there is a long story behind the above results. Generally speaking, such things have been known since ever, and more precisely, since the old work of Weyl [93] and Brauer [16]. However, in what regards the applications of the Weingarten formula, to various questions in mathematics or physics, and the interest in this formula in general, things here have evolved over the time with several ups and lows:

(1) In modern times, this formula has been quite popular among physicists since the 1978 paper of Weingarten [91], who was motivated by physics, and among mathematicians, since the 2003 paper of Collins [18], who was motivated by physics too.

(2) A key step was the 2006 paper of Collins-Śniady [23], with this formula clearly explained, for the unitary, orthogonal, and symplectic groups as well, and made ready to use, for everyone willing to do so, be them mathematicians or physicists.

(3) This technology has always been something rival to the Lie algebra theory, and a further increase in popularity came from the series of papers [8], [9], [10], [13], extending this formula to the quantum group setting, where no Lie theory is available.

(4) Finally, at the level of the applications, there are many of them, but probably the most popular ones, in recent times, came from the series of quantum information theory papers of Collins-Nechita [20], [21], [22], heavily relying on this formula.

Back to work now, as a first illustration for Theorem 16.3, let us discuss the computation of the Weingarten function for S_N. For this purpose, we can use the following result, which actually shows that the Weingarten formula is not really needed for S_N:

Theorem 16.4. *Consider the symmetric group $S_N\subset O_N$, with coordinates given by:*

$$g_{ij}=\chi\left(\sigma\in S_N\Big|\sigma(j)=i\right)$$

The products of these coordinates span then the algebra of functions $C(S_N)$, and the arbitrary integrals over S_N are given, modulo linearity, by the formula

$$\int_{S_N}g_{i_1j_1}\ldots g_{i_kj_k}=\begin{cases}\frac{(N-|\ker i|)!}{N!}&\text{if }\ \ker i=\ker j\\ 0&\text{otherwise}\end{cases}$$

where $\ker i$ *denotes as usual the partition of* $\{1,\ldots,k\}$ *whose blocks collect the equal indices of i, and where $|.|$ denotes the number of blocks.*

Proof. The first assertion follows from the Stone-Weierstrass theorem, because the standard coordinates g_{ij} separate the points of S_N, and so the algebra $<g_{ij}>$ that they generate must be equal to the whole function algebra $C(S_N)$:

$$<g_{ij}>=C(S_N)$$

Regarding now the second assertion, according to the definition of coordinates g_{ij}, the integrals in the statement are given by:

$$\int_{S_N} g_{i_1j_1}\ldots g_{i_kj_k} = \frac{1}{N!}\#\left\{\sigma\in S_N\Big|\sigma(j_1)=i_1,\ldots,\sigma(j_k)=i_k\right\}$$

Now observe that the existence of $\sigma\in S_N$ as above requires:

$$i_m=i_n\iff j_m=j_n$$

Thus, the above integral vanishes when the following happens:

$$\ker i\neq\ker j$$

Regarding now the case $\ker i=\ker j$, if we denote by $b\in\{1,\ldots,k\}$ the number of blocks of this partition $\ker i=\ker j$, we have $N-b$ points to be sent bijectively to $N-b$ points, and so $(N-b)!$ solutions, and the integral is $\frac{(N-b)!}{N!}$, as claimed. $\square$

The above result shows that the integration over S_N is something quite trivial, and so the computation of the Weingarten function should be something quite trivial too. In practice now, in order to compute the Weingarten function for S_N, by using the above result, we will need some combinatorics, and more specifically the Möbius inversion formula. Let us begin with some standard definitions, as follows:

Definition 16.5. *Let $P(k)$ be the set of partitions of $\{1,\ldots,k\}$, and let $\pi,\sigma\in P(k)$.*

1. *We write $\pi\leq\sigma$ if each block of π is contained in a block of σ.*
2. *We let $\pi\vee\sigma\in P(k)$ be the partition obtained by superposing π,σ.*

As an illustration here, at $k=2$ we have $P(2)=\{||,\sqcap\}$, and we have:

$$||\leq\sqcap$$

Also, at $k=3$ we have $P(3)=\{|||,\sqcap|,\sqcap\!\!\!\!|\,,|\sqcap,\sqcap\!\!\sqcap\}$, and the order relation is as follows:

$$|||\leq\sqcap|,\sqcap\!\!\!\!|\,,|\sqcap\leq\sqcap\!\!\sqcap$$

Observe also that we have the following inequalities:

$$\pi,\sigma\leq\pi\vee\sigma$$

In fact, the partition $\pi\vee\sigma$ is by construction the smallest possible one with this property. Due to this fact, this partition $\pi\vee\sigma$ is called supremum of π,σ.

We can now introduce the Möbius function, as follows:

Definition 16.6. *The Möbius function of any lattice, and so of P, is given by*

$$\mu(\pi,\sigma)=\begin{cases}1 & \text{if } \pi=\sigma\\ -\sum_{\pi\leq\tau<\sigma}\mu(\pi,\tau) & \text{if } \pi<\sigma\\ 0 & \text{if } \pi\not\leq\sigma\end{cases}$$

with the construction being performed by recurrence.

As an illustration here, let us go back to the set of 2-point partitions, $P(2)=\{||,\sqcap\}$. We have here, by definition of the Möbius function:

$$\mu(||,||)=\mu(\sqcap,\sqcap)=1$$

Also, we know that we have $||<\sqcap$, with no intermediate partition in between, and so the above recurrence procedure gives the following formulae:

$$\mu(||,\sqcap)=-\mu(||,||)=-1$$

Finally, we have $\sqcap\not\leq||$, and so $\mu(\sqcap,||)=0$. Thus, as a conclusion, the Möbius matrix $M_{\pi\sigma}=\mu(\pi,\sigma)$ of the lattice $P(2)=\{||,\sqcap\}$ is as follows:

$$M=\begin{pmatrix}1&-1\\0&1\end{pmatrix}$$

The interest in the Möbius function comes from the Möbius inversion formula:

$$f(\sigma)=\sum_{\pi\leq\sigma}g(\pi)\implies g(\sigma)=\sum_{\pi\leq\sigma}\mu(\pi,\sigma)f(\pi)$$

In linear algebra terms, the statement and proof of this formula are as follows:

THEOREM 16.7. *The inverse of the adjacency matrix of P, given by*

$$A_{\pi\sigma}=\begin{cases}1&\text{if }\pi\leq\sigma\\0&\text{if }\pi\not\leq\sigma\end{cases}$$

is the Möbius matrix of P, given by $M_{\pi\sigma}=\mu(\pi,\sigma)$.

PROOF. This is well-known, coming for instance from the fact that A is upper triangular. Indeed, when inverting, we are led into the recurrence from Definition 16.6. $\square$

As a first illustration, for $P(2)$ the formula $M=A^{-1}$ appears as follows:

$$\begin{pmatrix}1&-1\\0&1\end{pmatrix}=\begin{pmatrix}1&1\\0&1\end{pmatrix}^{-1}$$

Also, for $P(3)=\{|||,\sqcap|,\sqcap\!\!\!\!\!\;\,|\,,|\sqcap,\sqcap\!\sqcap\}$ the formula $M=A^{-1}$ reads:

$$\begin{pmatrix}1&-1&-1&-1&2\\0&1&0&0&-1\\0&0&1&0&-1\\0&0&0&1&-1\\0&0&0&0&1\end{pmatrix}=\begin{pmatrix}1&1&1&1&1\\0&1&0&0&1\\0&0&1&0&1\\0&0&0&1&1\\0&0&0&0&1\end{pmatrix}^{-1}$$

With the above results in hand, we can now compute the Weingarten function of S_N, and also find a precise estimate for it, as follows:

Theorem 16.8. *For S_N the Weingarten function is given by*

$$W_{kN}(\pi,\sigma)=\sum_{\tau\leq\pi\wedge\sigma}\mu(\tau,\pi)\mu(\tau,\sigma)\frac{(N-|\tau|)!}{N!}$$

and satisfies the folowing estimate,

$$W_{kN}(\pi,\sigma)=N^{-|\pi\wedge\sigma|}(\mu(\pi\wedge\sigma,\pi)\mu(\pi\wedge\sigma,\sigma)+O(N^{-1}))$$

with μ being the Möbius function of $P(k)$.

Proof. The first assertion follows from the Weingarten formula, namely:

$$\int_{S_N}u_{i_1j_1}\ldots u_{i_kj_k}=\sum_{\pi,\sigma\in P(k)}\delta_\pi(i)\delta_\sigma(j)W_{kN}(\pi,\sigma)$$

Indeed, in this formula the integrals on the left are known, from the explicit integration formula over S_N that we established above, namely:

$$\int_{S_N}g_{i_1j_1}\ldots g_{i_kj_k}=\begin{cases}\frac{(N-|\ker i|)!}{N!}&\text{if }\ker i=\ker j\\0&\text{otherwise}\end{cases}$$

But this allows the computation of the right term, via the Möbius inversion formula, explained above. As for the second assertion, this follows from the first one. See [12]. $\square$

As an illustration, let us record the formulae at $k=2,3$. At $k=2$, with indices $||,\sqcap$, and with the convention that $\approx$ means componentwise dominant term, we have:

$$W_{2N}\approx\begin{pmatrix}N^{-2}&-N^{-2}\\-N^{-2}&N^{-1}\end{pmatrix}$$

At $k=3$ now, with indices $|||,|\sqcap,\sqcap|,\sqcap\!\!\!|\,,\sqcap\!\sqcap$, and same meaning for $\approx$, we have:

$$W_{3N}\approx\begin{pmatrix}N^{-3}&-N^{-3}&-N^{-3}&-N^{-3}&2N^{-3}\\-N^{-3}&N^{-2}&N^{-3}&N^{-3}&-N^{-2}\\-N^{-3}&N^{-3}&N^{-2}&N^{-3}&-N^{-2}\\-N^{-3}&N^{-3}&N^{-3}&N^{-2}&-N^{-2}\\2N^{-3}&-N^{-2}&-N^{-2}&-N^{-2}&N^{-1}\end{pmatrix}$$

We will be back to all this later, with results about the orthogonal group O_N and about some other easy groups as well, where the Weingarten function is in general not explicitely computable, but where some useful estimates are still possible.

16b. Laws of characters

As a first concrete application of the above, let us discuss now the computation of the asymptotic laws of truncated characters. We have the following result, to start with:

THEOREM 16.9. *Assuming that $G\subset U_N$ is easy, coming from a category of partitions*
$$D=(D(k,l))$$
the moments of the main character are given by the formula
$$\int_G\chi^k=\dim\left(span\left(\xi_\pi\Big|\pi\in D(k)\right)\right)$$
where $D(k)=D(\emptyset,k)$, and where for $\pi\in D(k)$ we use the notation $\xi_\pi=T_\pi$.

PROOF. We recall that for an easy group $G\subset U_N$, coming from a category of partitions $D=(D(k,l))$, we have by definition equalities as follows:
$$Hom(u^{\otimes k},u^{\otimes l})=span\left(T_\pi\Big|\pi\in D(k,l)\right)$$

By interchanging $k\leftrightarrow l$ in this formula, and then setting $l=\emptyset$, we obtain:
$$Fix(u^{\otimes k})=span\left(\xi_\pi\Big|\pi\in D(k)\right)$$

Now since by the Peter-Weyl theory integrating a character amounts in counting the fixed points, we are led to the conclusion in the statement. □

In order to investigate the linear independence questions for the vectors ξ_π, we will use the Gram matrix of these vectors. We have the following result, to start with:

PROPOSITION 16.10. *The Gram matrix $G_{kN}(\pi,\sigma)=<\xi_\pi,\xi_\sigma>$ is given by*
$$G_{kN}(\pi,\sigma)=N^{|\pi\vee\sigma|}$$
where $|.|$ is the number of blocks.

PROOF. According to the formula of the vectors ξ_π, we have:
$$\begin{aligned}<\xi_\pi,\xi_\sigma> &= \sum_{i_1\ldots i_k}\delta_\pi(i_1,\ldots,i_k)\delta_\sigma(i_1,\ldots,i_k)\\ &= \sum_{i_1\ldots i_k}\delta_{\pi\vee\sigma}(i_1,\ldots,i_k)\\ &= N^{|\pi\vee\sigma|}\end{aligned}$$

Thus, we have obtained the formula in the statement. □

Next in line, we have the following key result:

Proposition 16.11. *The Gram matrix is given by $G_{kN}=AL$, where*

$$L(\pi,\sigma)=\begin{cases}N(N-1)\ldots(N-|\pi|+1)&\text{if }\sigma\leq\pi\\ 0&\text{otherwise}\end{cases}$$

and where $A=M^{-1}$ is the adjacency matrix of $P(k)$.

Proof. We have indeed the following computation:

$$\begin{aligned}N^{|\pi\vee\sigma|}&=\#\left\{i_1,\ldots,i_k\in\{1,\ldots,N\}\Big|\ker i\geq\pi\vee\sigma\right\}\\ &=\sum_{\tau\geq\pi\vee\sigma}\#\left\{i_1,\ldots,i_k\in\{1,\ldots,N\}\Big|\ker i=\tau\right\}\\ &=\sum_{\tau\geq\pi\vee\sigma}N(N-1)\ldots(N-|\tau|+1)\end{aligned}$$

According to Proposition 16.10 and to the definition of A,L, this formula reads:

$$\begin{aligned}(G_{kN})_{\pi\sigma}&=\sum_{\tau\geq\pi}L_{\tau\sigma}\\ &=\sum_\tau A_{\pi\tau}L_{\tau\sigma}\\ &=(AL)_{\pi\sigma}\end{aligned}$$

Thus, we obtain in this way the formula in the statement. $\square$

As an illustration for the above result, at $k=2$ we have $P(2)=\{||,\sqcap\}$, and the above formula $G_{kN}=AL$ appears as follows:

$$\begin{pmatrix}N^2&N\\ N&N\end{pmatrix}=\begin{pmatrix}1&1\\ 0&1\end{pmatrix}\begin{pmatrix}N^2-N&0\\ N&N\end{pmatrix}$$

At $k=3$ now, we have $P(3)=\{|||,\sqcap|,\sqcap\hskip-3.2mm{\ }_|\,,|\sqcap,\sqcap\sqcap\}$, and the Gram matrix is:

$$G_3=\begin{pmatrix}N^3&N^2&N^2&N^2&N\\ N^2&N^2&N&N&N\\ N^2&N&N^2&N&N\\ N^2&N&N&N^2&N\\ N&N&N&N&N\end{pmatrix}$$

Regarding L_3, this can be computed by writing down the matrix $E_3(\pi,\sigma)=\delta_{\sigma\leq\pi}|\pi|$, and then replacing each entry by the corresponding polynomial in N. We reach to the conclusion that the product A_3L_3 is as follows, producing the above matrix G_3:

$$A_3L_3=\begin{pmatrix}1&1&1&1&1\\0&1&0&0&1\\0&0&1&0&1\\0&0&0&1&1\\0&0&0&0&1\end{pmatrix}\begin{pmatrix}N^3-3N^2+2N&0&0&0&0\\N^2-N&N^2-N&0&0&0\\N^2-N&0&N^2-N&0&0\\N^2-N&0&0&N^2-N&0\\N&N&N&N&N\end{pmatrix}$$

In general, the formula $G_k=A_kL_k$ appears a bit in the same way, with A_k being binary and upper triangular, and with L_k depending on N, and being lower triangular.

With the above result in hand, we can now investigate the linear independence properties of the vectors ξ_π. We have here the following result of Lindstöm [67]:

THEOREM 16.12. *The determinant of the Gram matrix G_{kN} is given by*

$$\det(G_{kN})=\prod_{\pi\in P(k)}\frac{N!}{(N-|\pi|)!}$$

and in particular, for $N\geq k$, the vectors $\{\xi_\pi|\pi\in P(k)\}$ are linearly independent.

PROOF. According to the formula in Proposition 16.11 above, we have:

$$\det(G_{kN})=\det(A)\det(L)$$

Now if we order $P(k)$ as above, with respect to the number of blocks, and then lexicographically, we see that A is upper triangular, and that L is lower triangular. Thus $\det(A)$ can be computed simply by making the product on the diagonal, and we obtain 1. As for $\det(L)$, this can computed as well by making the product on the diagonal, and we obtain the number in the statement, with the technical remark that in the case $N<k$ the convention is that we obtain a vanishing determinant. ∎

Now back to the laws of characters, we can formulate:

THEOREM 16.13. *For an easy group $G=(G_N)$, coming from a category of partitions $D=(D(k,l))$, the asymptotic moments of the main character are given by*

$$\lim_{N\to\infty}\int_{G_N}\chi^k=\#D(k)$$

where $D(k)=D(\emptyset,k)$, with the limiting sequence on the left consisting of certain integers, and being stationary at least starting from the k-th term.

PROOF. This follows indeed from the general formula from Theorem 16.9, by using the linear independence result from Theorem 16.12. ∎

Our next purpose will be that of understanding what happens for the basic classes of easy groups. We have here the following result, to start with:

Theorem 16.14. *In the $N \to \infty$ limit, the law of the main character*

$$\chi_u = \sum_{i=1}^N u_{ii}$$

for the orthogonal and unitary groups is as follows:

(1) *For O_N we obtain a real Gaussian law g_1.*
(2) *For U_N we obtain a complex Gaussian law G_1.*

Proof. These results follow from the general formula from Theorem 16.13 above, by using the knowledge of the associated categories of partitions, as follows:

(1) For O_N the associated category of partitions is P_2, so the asymptotic moments of the main character are as follows, with the convention $k!! = 0$ when k is odd:

$$M_k = \#P_2(k) = k!!$$

Thus, we obtain the real Gaussian law, as stated.

(2) For U_N the associated category of partitions is $\mathcal{P}_2$, so the asymptotic moments of the main character, with respect to the colored integers, are as follows:

$$M_k = \#\mathcal{P}_2(k)$$

Thus, we obtain the complex Gaussian law, as stated. □

More generally now, we have the following result:

Theorem 16.15. *With $N \to \infty$, the laws of main character is as follows:*

(1) *For O_N we obtain the Gaussian law g_1.*
(2) *For U_N we obtain the complex Gaussian law G_1.*
(3) *For S_N we obtain the Poisson law p_1.*
(4) *For H_N we obtain the Bessel law b_1.*
(5) *For H_N^s we obtain the generalized Bessel law b_1^s.*
(6) *For K_N we obtain the complex Bessel law B_1.*

Also, for B_N, C_N and for Sp_N we obtain modified Gaussian laws.

Proof. We already know the results for O_N and for U_N, from Theorem 16.14. In general, the proof is similar, by counting the partitions in the associated category of partitions, and then doing some calculus, based on the various moment results for the laws in the statement, coming from the general theory developed in the above. All this is of course a bit technical, and for details we refer to [8], [23] and related papers. □

16c. Truncated characters

In order to fully solve the questions left open in chapters 9-12, we have to discuss now the more advanced question of computing the laws of truncated characters. First, we have the following formula, in the general easy group setting:

PROPOSITION 16.16. *The moments of truncated characters are given by the formula*

$$\int_G (g_{11}+\ldots+g_{ss})^k = Tr(W_{kN}G_{ks})$$

where G_{kN} and $W_{kN}=G_{kN}^{-1}$ are the associated Gram and Weingarten matrices.

PROOF. We have indeed the following computation:

$$\begin{aligned}
\int_G (g_{11}+\ldots+g_{ss})^k &= \sum_{i_1=1}^s\ldots\sum_{i_k=1}^s\int_G g_{i_1i_1}\ldots g_{i_ki_k}\\
&= \sum_{\pi,\sigma\in D(k)} W_{kN}(\pi,\sigma)\sum_{i_1=1}^s\ldots\sum_{i_k=1}^s\delta_\pi(i)\delta_\sigma(i)\\
&= \sum_{\pi,\sigma\in D(k)} W_{kN}(\pi,\sigma)G_{ks}(\sigma,\pi)\\
&= Tr(W_{kN}G_{ks})
\end{aligned}$$

Thus, we have obtained the formula in the statement. □

In order to process now the above formula, and reach to concrete results, we can impose the uniformity condition from chapter 15, originally used there for some technical classification purposes. Let us recall indeed from there that we have:

DEFINITION 16.17. *An easy group $G=(G_N)$, coming from a category of partitions $D\subset P$, is called uniform if it satisfies the following equivalent conditions:*

1. *$G_{N-1}=G_N\cap U_{N-1}$, via the embedding $U_{N-1}\subset U_N$ given by $u\to diag(u,1)$.*
2. *$G_{N-1}=G_N\cap U_{N-1}$, via the N possible diagonal embeddings $U_{N-1}\subset U_N$.*
3. *D is stable under the operation which consists in removing blocks.*

Here the equivalence between the above three conditions is something standard, obtained by doing some combinatorics, and this was discussed in chapter 15. We refer as well to chapter 15 for examples and counterexamples of such groups, the idea here being that the most familiar easy groups $G=(G_N)$ that we know are indeed uniform.

In what follows we will be mostly interested in the condition (3) above, which makes the link with our computations for truncated characters, and simplifies them. To be more precise, by imposing the uniformity condition we obtain:

Theorem 16.18. *For a uniform easy group $G = (G_N)$, we have the formula*

$$\lim_{N\to\infty}\int_{G_N}\chi_t^k = \sum_{\pi\in D(k)} t^{|\pi|}$$

with $D \subset P$ being the associated category of partitions.

Proof. We use the general moment formula from Proposition 16.18, namely:

$$\int_G (g_{11}+\ldots+g_{ss})^k = Tr(W_{kN}G_{ks})$$

By setting $s = [tN]$, with $t > 0$ being a given parameter, this formula becomes:

$$\int_{G_N}\chi_t^k = Tr(W_{kN}G_{k[tN]})$$

The point now is that in the uniform case the Gram and Weingarten matrices are asymptotically diagonal, and this leads to the formula in the statement. See [10]. □

We can now improve our character results, as follows:

Theorem 16.19. *With $N \to \infty$, the laws of truncated characters are as follows:*

(1) *For O_N we obtain the Gaussian law g_t.*
(2) *For U_N we obtain the complex Gaussian law G_t.*
(3) *For S_N we obtain the Poisson law p_t.*
(4) *For H_N we obtain the Bessel law b_t.*
(5) *For H_N^s we obtain the generalized Bessel law b_t^s.*
(6) *For K_N we obtain the complex Bessel law B_t.*

Also, for B_N, C_N and for Sp_N we obtain modified normal laws.

Proof. We use the formula that we found in Theorem 16.18, namely:

$$\lim_{N\to\infty}\int_{G_N}\chi_t^k = \sum_{\pi\in D(k)} t^{|\pi|}$$

By doing now some combinatorics, for instance in relation with the cumulants, this gives the results. We refer here to [10] and various related papers. □

All the above is quite interesting in relation with theoretical probability. Let us recall indeed from the first part of this book that we have 4 main limiting results in probability, namely discrete and continuous, and real and complex, which are as follows:

$$\begin{tikzcd}
CCPLT \arrow[r, dash] \arrow[d, dash] & CCLT \arrow[d, dash] \\
RCPLT \arrow[r, dash] & CLT
\end{tikzcd}$$

We also know that the limiting laws in these main limiting theorems are the real and complex Gaussian and Bessel laws, which are as follows:

$$\begin{array}{ccc} B_t & - & G_t \\ | & & | \\ b_t & - & g_t \end{array}$$

Moreover, we have also seen in the above that at the level of the moments, these come from certain collections of partitions, as follows:

$$\begin{array}{ccc} \mathcal{P}_{even} & - & \mathcal{P}_2 \\ | & & | \\ P_{even} & - & P_2 \end{array}$$

The point now is that, according to our general easiness philosophy, and also to Theorem 16.19, there are some Lie groups behind all this probability theory, namely the basic real and complex rotation and reflection groups, which as follows:

$$\begin{array}{ccc} K_N & - & U_N \\ | & & | \\ H_N & - & O_N \end{array}$$

To be more precise, these Lie groups correspond via easiness to the categories of partitions given above, and the corresponding measures can be recaptured as well, as being the asymptotic laws of the corresponding truncated characters, as explained in Theorem 16.19. As for the main probabilistic limiting results themselves, these are of course related too to these Lie groups, but this is something a bit more technical.

All this is very nice. With all this in hand, we are now at a rather advanced level in theoretical probability, and with this knowledge, you can virtually read any article and book in theoretical probability, that you might want to. With our recommendations here being the article of Diaconis-Shahshahani [27], and other texts by Diaconis, which are all quite magic, and no wonder here, because Diaconis used to be a professional magician before doing mathematics, then the classical and lovely random matrix book by Mehta [70], and then some fancy theoretical physics from Collins-Nechita [20], [21], [22].

16d. Standard estimates

We have seen in the above that the Weingarten calculus is something very efficient in dealing with various probability questions over the easy groups $G \subset U_N$. We discuss now, as a continuation of this, a number of more advanced aspects of the Weingarten function combinatorics. We will be mostly interested in the case $G = O_N$. To be more precise, we will be interested in the computation of the polynomial integrals over O_N. These polynomial integrals are best introduced in a "rectangular way", as follows:

DEFINITION 16.20. *Associated to any matrix $a \in M_{p\times q}(\mathbb{N})$ is the integral*

$$I(a) = \int_{O_N} \prod_{i=1}^{p} \prod_{j=1}^{q} u_{ij}^{a_{ij}}\, du$$

with respect to the Haar measure of O_N, where $N \geq p, q$.

We can of course complete our matrix with 0 values, as to always deal with square matrices, $a \in M_N(\mathbb{N})$. However, the parameters p, q are very useful, because they measure the "complexity" of the problem, as shown for instance by the result below.

Let $x!! = (x-1)(x-3)(x-5)\ldots$, with the product ending at 1 or 2. We have:

THEOREM 16.21. *At $p = 1$ we have the formula*

$$I\begin{pmatrix} a_1 & \ldots & a_q \end{pmatrix} = \varepsilon \cdot \frac{(N-1)!!a_1!!\ldots a_q!!}{(N+\Sigma a_i - 1)!!}$$

where $\varepsilon = 1$ if all a_i are even, and $\varepsilon = 0$ otherwise.

PROOF. This follows from the fact that the first slice of O_N is isomorphic to the real sphere $S^{N-1}_{\mathbb{R}}$. Indeed, this gives the following formula:

$$I\begin{pmatrix} a_1 & \ldots & a_q \end{pmatrix} = \int_{S^{N-1}_{\mathbb{R}}} x_1^{a_1}\ldots x_q^{a_q}\, dx$$

But this latter integral can be computed by using polar coordinates, via the various formulae from chapters 5-6, and we obtain the formula in the statement. □

Another instructive computation, as well of trigonometric nature, is the one at $N = 2$. We have here the following result, which completely solves the problem in this case:

THEOREM 16.22. *At $N = 2$ we have the formula*

$$I\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \varepsilon \cdot \frac{(a+d)!!(b+c)!!}{(a+b+c+d+1)!!}$$

where $\varepsilon = 1$ if a, b, c, d are even, $\varepsilon = -1$ is a, b, c, d are odd, and $\varepsilon = 0$ otherwise.

PROOF. When computing the integral over O_2, we can restrict the integration to $SO_2 = \mathbb{T}$, then further restrict the integration to the first quadrant. We get:

$$I\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \varepsilon \cdot \frac{2}{\pi}\int_0^{\pi/2} (\cos t)^{a+d}(\sin t)^{b+c}\,dt$$

This gives the formula in the statement. □

The above computations might tend to suggest that $I(a)$ always decomposes as a product of factorials. However, this is far from being true, but in the 2×2 case it is known that $I(a)$ decomposes as a quite reasonable sum of products of factorials. This is something quite technical, from [11], and we will be back to this, later on.

Let us discuss now the representation theory approach to the computation of $I(a)$. The Weingarten formula reformulates, in "rectangular form", as follows:

THEOREM 16.23. *We have the Weingarten formula*

$$I(a) = \sum_{\pi,\sigma} \delta_\pi(a_l)\delta_\sigma(a_r)W_{kN}(\pi,\sigma)$$

where $k = \Sigma a_{ij}/2$, and where the multi-indices a_l/a_r are defined as follows:

(1) *Start with $a \in M_{p\times q}(\mathbb{N})$, and replace each ij-entry by a_{ij} copies of i/j.*
(2) *Read this matrix in the usual way, as to get the multi-indices a_l/a_r.*

PROOF. This is simply a reformulation of the Weingarten formula. Indeed, according to our definitions, the integral in the statement is given by:

$$I(a) = \int_{O_n} \underbrace{u_{11}\ldots u_{11}}_{a_{11}}\ \underbrace{u_{12}\ldots u_{12}}_{a_{12}}\ \ldots\ \underbrace{u_{pq}\ldots u_{pq}}_{a_{pq}}\,du$$

Thus what we have here is an integral exactly as in the usual Weingarten formula, the multi-indices which are involved being as follows:

$$\begin{aligned} a_l &= (\underbrace{1\ldots 1}_{a_{11}}\ \underbrace{1\ldots 1}_{a_{12}}\ \ldots\ \underbrace{p\ldots p}_{a_{pq}}) \\ a_r &= (\underbrace{1\ldots 1}_{a_{11}}\ \underbrace{2\ldots 2}_{a_{12}}\ \ldots\ \underbrace{q\ldots q}_{a_{pq}}) \end{aligned}$$

The result follows now from the Weingarten formula. □

We are now in position of deriving a first general corollary from our study. This extends the vanishing results appearing before, as follows:

PROPOSITION 16.24. *We have $I(a) = 0$, unless the matrix a is "admissible", in the sense that all $p + q$ sums on its rows and columns are even numbers.*

PROOF. Observe first that the left multi-index associated to a consists of $k_1 = \Sigma a_{1j}$ copies of 1, $k_2 = \Sigma a_{2j}$ copies of 2, and so on, up to $k_p = \Sigma a_{pj}$ copies of p. In the case where one of these numbers is odd we have $\delta_\pi(a) = 0$ for any π, and this gives:

$$I(a) = 0$$

A similar argument with the right multi-index associated to a shows that the sums on the columns of a must be even as well, and we are done. $\square$

A natural question now is whether the converse of Proposition 16.24 holds, and if so, the question of computing the sign of $I(a)$ appears as well. These are both quite subtle questions, and we begin our investigations with a $N \to \infty$ study. We have:

THEOREM 16.25. *The Weingarten matrix is asymptotically diagonal, in the sense that:*

$$W_{kN}(\pi, \sigma) = N^{-k}(\delta_{\pi\sigma} + O(N^{-1}))$$

Moreover, the $O(N^{-1})$ remainder is asymptotically smaller that $(2k/e)^k N^{-1}$.

PROOF. It is convenient, for the purposes of this proof, to drop the indices k, N. We know that the Gram matrix is given by $G(\pi, \sigma) = N^{loops(\pi,\sigma)}$, so we have:

$$G(\pi, \sigma) = \begin{cases} N^k & \text{for } \pi = \sigma \\ N, N^2, \ldots, N^{k-1} & \text{for } \pi \neq \sigma \end{cases}$$

Thus the Gram matrix is of the following form, with $||H||_\infty \leq N^{-1}$:

$$G = N^k(I + H)$$

Now recall that for any $K \times K$ complex matrix X, we have the following lineup of standard inequalities, which are all elementary:

$$||X||_\infty \leq ||X|| \leq ||X||_2 \leq K||X||_\infty$$

In the case of our matrix H, the size is $K = (2k)!!$, so we have:

$$||H|| \leq KN^{-1}$$

We can use now the following basic inversion formula:

$$(I + H)^{-1} = I - H + H^2 - H^3 + \ldots$$

We conclude from this that we have the following estimate:

$$||I - (I + H)^{-1}|| \leq \frac{||H||}{1 - ||H||}$$

By putting everything together, we obtain the following estimate:

$$\begin{aligned}||I-N^kW||_\infty &= ||I-(1+H)^{-1}||_\infty\\ &\leq ||I-(1+H)^{-1}||\\ &\leq ||H||/(1-||H||)\\ &\leq KN^{-1}/(1-KN^{-1})\\ &= K/(N-K)\end{aligned}$$

Together with the standard estimate $K \approx (2k/e)^k$, this gives the result. $\square$

As a continuation of this, we have the following result:

THEOREM 16.26. *We have the estimate*

$$I(a) = N^{-k}\left(\prod_{i=1}^p\prod_{j=1}^q a_{ij}!! + O(N^{-1})\right)$$

when all a_{ij} are even, and $I(a) = O(N^{-k-1})$ otherwise.

PROOF. By using the above results, we obtain the following estimate:

$$\begin{aligned}I(a) &= \sum_{\pi,\sigma}\delta_\pi(a_l)\delta_\sigma(a_r)W_{kN}(\pi,\sigma)\\ &= n^{-k}\sum_{\pi,\sigma}\delta_\pi(a_l)\delta_\sigma(a_r)(\delta_{\pi\sigma}+O(N^{-1}))\\ &= N^{-k}\left(\#\{\pi|\delta_\pi(a_l)=\delta_\pi(a_r)=1\}+O(N^{-1})\right)\end{aligned}$$

In order to count the partitions appearing in the set on the right, it is convenient to view the multi-indices a_l, a_r in a rectangular way, as follows:

$$a_l = \begin{pmatrix}\underbrace{1\ldots 1}_{a_{11}} & \ldots & \underbrace{1\ldots 1}_{a_{1q}}\\ \ldots & \ldots & \ldots\\ \underbrace{p\ldots p}_{a_{p1}} & \ldots & \underbrace{p\ldots p}_{a_{pq}}\end{pmatrix} \quad , \quad a_r = \begin{pmatrix}\underbrace{1\ldots 1}_{a_{11}} & \ldots & \underbrace{q\ldots q}_{a_{1q}}\\ \ldots & \ldots & \ldots\\ \underbrace{1\ldots 1}_{a_{p1}} & \ldots & \underbrace{p\ldots p}_{a_{pq}}\end{pmatrix}$$

In other words, the multi-indices a_l/a_r are now simply obtained from the matrix a by "dropping" from each entry a_{ij} a sequence of a_{ij} numbers, all equal to i/j.

These two multi-indices, now in matrix form, have total length $2k = \Sigma a_{ij}$. We agree to view as well any pairing of $\{1,\ldots,2k\}$ in matrix form, by following the same convention.

With this picture, the pairings π which contribute are simply those interconnecting sequences of indices "dropped" from the same a_{ij}, and this gives the following results:

(1) In the case where one of the entries a_{ij} is odd, there is no pairing that can contribute to the leading term under consideration, so we have $I(a) = O(N^{-k-1})$, and we are done.

(2) In the case where all the entries a_{ij} are even, the pairings that contribute to the leading term are those connecting points inside the pq "dropped" sets, i.e. are made out of a pairing of a_{11} points, a pairing of a_{12} points, and so on, up to a pairing of a_{pq} points. Now since an x-point set has $x!!$ pairings, this gives the formula in the statement. □

In order to further advance, let us formulate a key definition, as follows:

DEFINITION 16.27. *The Brauer space D_k is defined as follows:*

(1) *The points are the Brauer diagrams, i.e. the pairings of $\{1, 2, \ldots, 2k\}$.*
(2) *The distance function is given by $d(\pi, \sigma) = k - loops(\pi, \sigma)$.*

It is indeed well-known, and elementary to check, that d satisfies the usual axioms for a distance function. This is something standard, and heavily used in probability theory, and for some comments and examples here, we refer to [10], [23] and related papers. Now the point is that we have a series expansion of the Weingarten function in terms of paths on the Brauer space, originally found by Collins in [18] in the unitary case, then by Collins and Śniady [23] in the orthogonal case. We present here a slightly modified statement, along with a complete proof, by using a somewhat lighter formalism:

THEOREM 16.28. *The Weingarten function W_{kN} has a series expansion in N^{-1},*

$$W_{kN}(\pi, \sigma) = N^{-k-d(\pi,\sigma)} \sum_{g=0}^{\infty} K_g(\pi, \sigma) N^{-g}$$

where the objects on the right are defined as follows:

(1) *A path from π to σ is a sequence $p = [\pi = \tau_0 \neq \tau_1 \neq \ldots \neq \tau_r = \sigma]$.*
(2) *The signature of such a path is $+$ when r is even, and $-$ when r is odd.*
(3) *The geodesicity defect of such a path is $g(p) = \Sigma_{i=1}^r d(\tau_{i-1}, \tau_i) - d(\pi, \sigma)$.*
(4) *K_g counts the signed paths from π to σ, with geodesicity defect g.*

PROOF. Let us go back to the proof of our main estimate so far, established in the above. We can write the Gram matrix in the following way:

$$G_{kn} = N^{-k}(I + H)$$

In terms of the Brauer space distance, the formula of H is simply:

$$H(\pi, \sigma) = \begin{cases} 0 & \text{for } \pi = \sigma \\ N^{-d(\pi,\sigma)} & \text{for } \pi \neq \sigma \end{cases}$$

Consider now the set $P_r(\pi, \sigma)$ of r-paths between π and σ. According to the usual rule of matrix multiplication, the powers of H are given by:

$$H^r(\pi, \sigma) = \sum_{p \in P_r(\pi,\sigma)} H(\tau_0, \tau_1) \ldots H(\tau_{r-1}, \tau_r) = \sum_{p \in P_r(\pi,\sigma)} N^{-d(\pi,\sigma)-g(p)}$$

We can use now the following standard inversion formula:

$$(1+H)^{-1} = 1 - H + H^2 - H^3 + \ldots$$

By using this formula, we obtain:

$$\begin{aligned} W_{kN}(\pi,\sigma) &= N^{-k}\sum_{r=0}^\infty (-1)^r H^r(\pi,\sigma) \\ &= N^{-k-d(\pi,\sigma)}\sum_{r=0}^\infty \sum_{p\in P_r(\pi,\sigma)} (-1)^r N^{-g(p)} \end{aligned}$$

Now by rearranging the various terms of the double sum according to their geodesicity defect $g = g(p)$, this gives the following formula:

$$W_{kN}(\pi,\sigma) = N^{-k-d(\pi,\sigma)}\sum_{g=0}^\infty K_g(\pi,\sigma)N^{-g}$$

Thus, we have obtained the formula in the statement. □

In order to discuss now the $I(a)$ reformulation of the above result, it is convenient to use the total length of a path, defined as follows:

$$d(p) = \sum_{i=1}^r d(\tau_{i-1},\tau_i)$$

Observe that, in terms of this quantity, we have the following formula:

$$d(p) = d(\pi,\sigma) + g(p)$$

With these conventions, we have the following result:

THEOREM 16.29. *The integral $I(a)$ has a series expansion in N^{-1} of the form*

$$I(a) = N^{-k}\sum_{d=0}^\infty H_d(a)N^{-d}$$

where the coefficient on the right can be interpreted as follows:

(1) *Starting from $a \in M_{p\times q}(\mathbb{N})$, construct the multi-indices a_l, a_r as usual.*
(2) *Call a path "a-admissible" if its endpoints satisfy $\delta_\pi(a_l) = 1$ and $\delta_\sigma(a_r) = 1$.*
(3) *Then $H_d(a)$ counts all a-admissible signed paths in D_k, of total length d.*

PROOF. We combine first the above results, in the following way:

$$\begin{aligned} I(a) &= \sum_{\pi,\sigma}\delta_\pi(a_l)\delta_\sigma(a_r)W_{kN}(\pi,\sigma) \\ &= N^{-k}\sum_{\pi,\sigma}\delta_\pi(a_l)\delta_\sigma(a_r)\sum_{g=0}^\infty K_g(\pi,\sigma)N^{-d(\pi,\sigma)-g} \end{aligned}$$

Let us denote by $H_d(\pi,\sigma)$ the number of signed paths between π and σ, of total length d. In terms of the new variable $d=d(\pi,\sigma)+g$, the above expression becomes:

$$\begin{aligned}
I(a) &= N^{-k}\sum_{\pi,\sigma}\delta_\pi(a_l)\delta_\sigma(a_r)\sum_{d=0}^\infty H_d(\pi,\sigma)N^{-d}\\
&= N^{-k}\sum_{d=0}^\infty\left(\sum_{\pi,\sigma}\delta_\pi(a_l)\delta_\sigma(a_r)H_d(\pi,\sigma)\right)N^{-d}
\end{aligned}$$

We recognize in the middle the quantity $H_d(a)$, and this gives the result. □

We derive now some concrete consequences from the abstract results in the previous section. First, let us recall the following result, due to Collins and Śniady [23]:

THEOREM 16.30. *We have the estimate*

$$W_{kN}(\pi,\sigma)=N^{-k-d(\pi,\sigma)}(\mu(\pi,\sigma)+O(N^{-1}))$$

where μ is the Möbius function.

PROOF. We know from the above that we have the following estimate:

$$W_{kN}(\pi,\sigma)=N^{-k-d(\pi,\sigma)}(K_0(\pi,\sigma)+O(N^{-1}))$$

Now since one of the possible definitions of the Möbius function is that this counts the signed geodesic paths, we have $K_0=\mu$, and we are done. □

Let us go back now to our integrals $I(a)$. We have the following result:

THEOREM 16.31. *We have the estimate*

$$I(a)=N^{-k-e(a)}(\mu(a)+O(N^{-1}))$$

where the objects on the right are as follows:

(1) $e(a)=\min\{d(\pi,\sigma)|\pi,\sigma\in D_k,\delta_\pi(a_l)=\delta_\sigma(a_r)=1\}$.
(2) $\mu(a)$ *counts all a-admissible signed paths in D_k, of total length $e(a)$.*

PROOF. We know that we have an estimate of the following type:

$$I(a)=N^{-k-e}(H_e(a)+O(N^{-1}))$$

Here, according to the various notations above, $e\in\mathbb N$ is the smallest total length of an a-admissible path, and $H_e(a)$ counts all signed a-admissible paths of total length e. Now since the smallest total length of such a path is of course attained when the path is just a segment, we have $e=e(a)$ and $H_e(a)=\mu(a)$, and we are done. □

At a more advanced level now, and still on the same topic, integration over O_N, we have the following result, due to Collins-Matsumoto [19] and Zinn-Justin [99]:

THEOREM 16.32. *We have the formula*

$$W_{kn}(\pi,\sigma)=\frac{\sum_{\lambda\vdash k,\,l(\lambda)\leq k}\chi^{2\lambda}(1_k)w^\lambda(\pi^{-1}\sigma)}{(2k)!!\prod_{(i,j)\in\lambda}(n+2j-i-1)}$$

where the various objects on the right are as follows:

(1) *The sum is over all partitions of $\{1,\ldots,2k\}$ of length $l(\lambda)\leq k$.*
(2) *w^λ is the corresponding zonal spherical function of (S_{2k},H_k).*
(3) *$\chi^{2\lambda}$ is the character of S_{2k} associated to $2\lambda=(2\lambda_1,2\lambda_2,\ldots)$.*
(4) *The product is over all squares of the Young diagram of λ.*

PROOF. This is something quite technical, that we will not attempt to explain here, and for details on all this, we refer to the papers [19], [99]. □

It is of course possible to deduce from this a new a formula for the integrals $I(a)$, just by putting together the various formulae that we have. Let us just record here:

THEOREM 16.33. *The possible poles of $I(a)$ can be at the numbers*

$$-(k-1),-(k-2),\ldots,2k-1,2k$$

where the number $k\in\mathbb N$ associated to the admissible matrix $a\in M_{p\times q}(\mathbb N)$ is given by:

$$k=\sum a_{ij}/2$$

PROOF. We know from the above that the possible poles of $I(a)$ can only come from those of the Weingarten function. On the other hand, Theorem 16.32 tells us that these latter poles are located at the numbers of the form $-2j+i+1$, with (i,j) ranging over all possible squares of all possible Young diagrams, and this gives the result. □

As a last topic, let us discuss Gram determinants. In what regards the symmetric group S_N, we have the following result, that we already know, from the above:

THEOREM 16.34. *The determinant of the Gram matrix of S_N is given by*

$$\det(G_{kN})=\prod_{\pi\in P(k)}\frac{N!}{(N-|\pi|)!}$$

with the convention that in the case $N<k$ we obtain 0.

PROOF. This is something that we know, the idea being that G_{kN} naturally decomposes as a product of an upper triangular and lower triangular matrix. □

Let us discuss now the case of the orthogonal group O_N. Here the combinatorics is that of the Young diagrams. We denote by $|.|$ the number of boxes, and we use quantity f^λ, which gives the number of standard Young tableaux of shape λ. With these conventions, the result, which is something quite technical, is then as follows:

THEOREM 16.35. *The determinant of the Gram matrix of O_N is given by*

$$\det(G_{kN}) = \prod_{|\lambda|=k/2} f_N(\lambda)^{f^{2\lambda}}$$

where the quantities on the right are $f_N(\lambda) = \prod_{(i,j)\in\lambda}(N + 2j - i - 1)$.

PROOF. This follows from the results of Zinn-Justin in [99]. Indeed, it is known from there that the Gram matrix is diagonalizable, as follows:

$$G_{kN} = \sum_{|\lambda|=k/2} f_N(\lambda)P_{2\lambda}$$

Here $1 = \sum P_{2\lambda}$ is the standard partition of unity associated to the Young diagrams having $k/2$ boxes, and the coefficients $f_N(\lambda)$ are those in the statement. Now since we have $Tr(P_{2\lambda}) = f^{2\lambda}$, this gives the result. See [12], [99]. $\square$

Finally, no book about groups and algebra would be complete without some quantum groups and algebra at the end. Unfortunately, we are here, with our Gram determinants, into quite advanced things, and so we will have to trick a bit, and take some dirty shortcuts. Instead of formulating a theorem, let us start with a definition:

DEFINITION 16.36. *In analogy with the fact that S_N, O_N are easy, coming from P, P_2, let us denote by S_N^+, O_N^+ the formal objects associated to NC, NC_2.*

Observe that S_N^+, O_N^+ cannot be groups, because the sets of noncrossing partitions and noncrossing pairings NC, NC_2 do not contain the basic crossing $\slash\!\!\!\backslash$, and so are not categories of partitions in the sense of chapter 15. This being said, the axiom stating that $\slash\!\!\!\backslash$ must be in the category was coming from the fact that the standard coordinates $u_{ij} : G \to \mathbb{C}$ of a compact Lie group $G \subset_u U_N$ commute, and so in the lack of this axiom, we can only have some kind of "quantum groups", which are beasts a bit like groups, save for the fact that the coordinates $u_{ij} : G \to \mathbb{C}$ do not longer commute.

And this latter fact is true, with S_N^+, O_N^+ being indeed quantum groups, called quantum permutation group, and quantum rotation group. Of course, their construction is something which takes some time, explained for instance in [6], and what is said in Definition 16.36 above is something rather advanced, corresponding to their easiness property. But we won't attempt to explain more, and we will take Definition 16.36 as it is.

Now with this definition in hand, and getting back now to our Gram determinant problematics, we would like to compute the Gram determinants for S_N^+, O_N^+, whatever these beasts exactly are, and in practice, we are led to a very concrete and explicit problem, namely that of computing the Gram determinants for NC, NC_2.

Following Di Francesco [26], let us begin with some examples. We first have:

PROPOSITION 16.37. *At $k=2$ the set of partitions for S_N^+ is $NC(2)=\{||,\sqcap\}$, and the corresponding Gram matrix and its determinant are:*

$$\det\begin{pmatrix}N^2&N\\N&N\end{pmatrix}=N^2(N-1)$$

Also, at $k=4$ the set of partitions for O_N^+ is $NC_2(4)=\{\sqcap\sqcap,\bigcap\!\!\!\!\cap\}$, and the corresponding Gram matrix and its determinant are:

$$\det\begin{pmatrix}N^2&N\\N&N^2\end{pmatrix}=N^2(N^2-1)$$

PROOF. This is something which is clear from definitions. □

With a few tricks, we can work out as well the next computation, as follows:

PROPOSITION 16.38. *At $k=3$ the partition set for S_N^+ is $NC(3)=\{|||,\sqcap|,\sqcap\!\!\!\!|\,,|\sqcap,\sqcap\!\sqcap\}$, and the corresponding Gram matrix and its determinant are:*

$$\det\begin{pmatrix}N^3&N^2&N^2&N^2&N\\N^2&N^2&N&N&N\\N^2&N&N^2&N&N\\N^2&N&N&N^2&N\\N&N&N&N&N\end{pmatrix}=N^5(N-1)^4(N-2)$$

Also, at $k=6$ the set of partitions for O_N^+ is $NC_2(6)\simeq NC(3)$, and the corresponding Gram matrix and its determinant are:

$$\det\begin{pmatrix}N^3&N^2&N^2&N^2&N\\N^2&N^3&N&N&N^2\\N^2&N&N^3&N&N^2\\N^2&N&N&N^3&N^2\\N&N^2&N^2&N^2&N^3\end{pmatrix}=N^5(N^2-1)^4(N^2-2)$$

PROOF. We have two formulae to be proved, the idea being as follows:

(1) In what regards S_N^+, the set of partitions here is $NC(3)=P(3)$, and so the corresponding Gram matrix is the one in the statement, exactly as for S_N. By using the Lindstöm formula, from Theorem 16.12, the determinant of this matrix is, as claimed:

$$\begin{aligned}\det&=\prod_{\pi\in P(3)}\frac{N!}{(N-|\pi|)!}\\&=\frac{N!}{(N-3)!}\left(\frac{N!}{(N-2)!}\right)^3\frac{N!}{(N-1)!}\\&=N(N-1)(N-2)N^3(N-1)^3N\\&=N^5(N-1)^4(N-2)\end{aligned}$$

(2) Regarding now O_N^+, the set of partitions here is $NC_2(6)$, and by using the fattening/shrinking identification $NC_2(6)\simeq NC(3)$, we obtain, by using (1):

$$\begin{aligned}\det &= \frac{1}{N^2\sqrt{N}}\times N^{10}(N^2-1)^4(N^2-2)\times\frac{1}{N^2\sqrt{N}}\\ &= N^5(N^2-1)^4(N^2-2)\end{aligned}$$

Thus, we have obtained the formula in the statement. □

In general now, following [26], we have the following result:

THEOREM 16.39. *The determinant of the Gram matrix for O_N^+ is given by*

$$\det(G_{kN})=\prod_{r=1}^{[k/2]}P_r(N)^{d_{k/2,r}}$$

where P_r are the Chebycheff polynomials, given by

$$P_0=1\quad,\quad P_1=X\quad,\quad P_{r+1}=XP_r-P_{r-1}$$

and $d_{kr}=f_{kr}-f_{k,r+1}$, with f_{kr} being the following numbers, depending on $k,r\in\mathbb Z$,

$$f_{kr}=\binom{2k}{k-r}-\binom{2k}{k-r-1}$$

with the convention $f_{kr}=0$ for $k\notin\mathbb Z$.

PROOF. This is something quite heavy, and we refer here to Di Francesco [26]. □

Also following [26], we have as well the following result:

THEOREM 16.40. *The determinant of the Gram matrix for S_N^+ is given by*

$$\det(G_{kN})=(\sqrt{N})^{a_k}\prod_{r=1}^{k}P_r(\sqrt{N})^{d_{kr}}$$

where $d_{kr}=f_{kr}-f_{k,r+1}$, with f_{kr} being the following numbers, depending on $k,r\in\mathbb Z$,

$$f_{kr}=\binom{2k}{k-r}-\binom{2k}{k-r-1}$$

with the convention $f_{kr}=0$ for $k\notin\mathbb Z$, and where $a_k=\sum_{\pi\in\mathcal P(k)}(2|\pi|-k)$.

PROOF. Again, heavy mathematics, and we refer here to Di Francesco [26]. □

We refer to [12], [26], [40] for a further discussion on these topics.

16e. Exercises

Congratulations for having read this book, and no exercises for this final chapter. But you can try instead to read the various books and articles referenced below.

Bibliography

[1] V.I. Arnold, Ordinary differential equations, Springer (1973).

[2] V.I. Arnold, Mathematical methods of classical mechanics, Springer (1974).

[3] V.I. Arnold, Lectures on partial differential equations, Springer (1997).

[4] V.I. Arnold and B.A. Khesin, Topological methods in hydrodynamics, Springer (1998).

[5] M.F. Atiyah, The geometry and physics of knots, Cambridge Univ. Press (1990).

[6] T. Banica, Introduction to quantum groups, Springer (2023).

[7] T. Banica, Introduction to modern physics (2023).

[8] T. Banica, S.T. Belinschi, M. Capitaine and B. Collins, Free Bessel laws, *Canad. J. Math.* **63** (2011), 3–37.

[9] T. Banica, J. Bichon and B. Collins, The hyperoctahedral quantum group, *J. Ramanujan Math. Soc.* **22** (2007), 345–384.

[10] T. Banica and B. Collins, Integration over quantum permutation groups, *J. Funct. Anal.* **242** (2007), 641–657.

[11] T. Banica, B. Collins and J.-M. Schlenker, On polynomial integrals over the orthogonal group, *J. Combin. Theory Ser. A* **118** (2011), 778–795.

[12] T. Banica and S. Curran, Decomposition results for Gram matrix determinants, *J. Math. Phys.* **51** (2010), 1–14.

[13] T. Banica and R. Speicher, Liberation of orthogonal Lie groups, *Adv. Math.* **222** (2009), 1461–1501.

[14] I. Bengtsson and K. Życzkowski, Geometry of quantum states, Cambridge Univ. Press (2006).

[15] G. Björck, Functions of modulus 1 on Z_n whose Fourier transforms have constant modulus, and cyclic n-roots, *NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci.* **315** (1990), 131–140.

[16] R. Brauer, On algebras which are connected with the semisimple continuous groups, *Ann. of Math.* **38** (1937), 857–872.

[17] V. Chari and A. Pressley, A guide to quantum groups, Cambridge Univ. Press (1994).

[18] B. Collins, Moments and cumulants of polynomial random variables on unitary groups, the Itzykson-Zuber integral, and free probability, *Int. Math. Res. Not.* **17** (2003), 953–982.

[19] B. Collins and S. Matsumoto, On some properties of orthogonal Weingarten functions, *J. Math. Phys.* **50** (2009), 1–18.

[20] B. Collins and I. Nechita, Random quantum channels I: graphical calculus and the Bell state phenomenon, *Comm. Math. Phys.* **297** (2010), 345–370.

[21] B. Collins and I. Nechita, Random quantum channels II: entanglement of random subspaces, Rényi entropy estimates and additivity problems, *Adv. Math.* **226** (2011), 1181–1201.

[22] B. Collins and I. Nechita, Gaussianization and eigenvalue statistics for random quantum channels (III), *Ann. Appl. Probab.* **21** (2011), 1136–1179.

[23] B. Collins and P. Śniady, Integration with respect to the Haar measure on unitary, orthogonal and symplectic groups, *Comm. Math. Phys.* **264** (2006), 773–795.

[24] A. Connes, Noncommutative geometry, Academic Press (1994).

[25] P. Deligne, Catégories tannakiennes, in "Grothendieck Festchrift", Birkhauser (1990), 111–195.

[26] P. Di Francesco, Meander determinants, *Comm. Math. Phys.* **191** (1998), 543–583.

[27] P. Diaconis and M. Shahshahani, On the eigenvalues of random matrices, *J. Applied Probab.* **31** (1994), 49–62.

[28] P.A.M. Dirac, Principles of quantum mechanics, Oxford Univ. Press (1930).

[29] M.P. do Carmo, Differential geometry of curves and surfaces, Dover (1976).

[30] M.P. do Carmo, Riemannian geometry, Birkhäuser (1992).

[31] S. Doplicher and J. Roberts, A new duality theory for compact groups, *Invent. Math.* **98** (1989), 157–218.

[32] V.G. Drinfeld, Quantum groups, Proc. ICM Berkeley (1986), 798–820.

[33] R. Durrett, Probability: theory and examples, Cambridge Univ. Press (1990).

[34] L.C. Evans, Partial differential equations, AMS (1998).

[35] W. Feller, An introduction to probability theory and its applications, Wiley (1950).

[36] E. Fermi, Thermodynamics, Dover (1937).

[37] R.P. Feynman, R.B. Leighton and M. Sands, The Feynman lectures on physics I: mainly mechanics, radiation and heat, Caltech (1963).

[38] R.P. Feynman, R.B. Leighton and M. Sands, The Feynman lectures on physics II: mainly electromagnetism and matter, Caltech (1964).

[39] R.P. Feynman, R.B. Leighton and M. Sands, The Feynman lectures on physics III: quantum mechanics, Caltech (1966).

[40] M. Fukuda and P. Śniady, Partial transpose of random quantum states: exact formulas and meanders, *J. Math. Phys.* **54** (2013), 1–31.

[41] D.J. Griffiths, Introduction to electrodynamics, Cambridge Univ. Press (2017).

[42] D.J. Griffiths and D.F. Schroeter, Introduction to quantum mechanics, Cambridge Univ. Press (2018).

[43] D.J. Griffiths, Introduction to elementary particles, Wiley (2020).

[44] D.J. Griffiths, Revolutions in twentieth-century physics, Cambridge Univ. Press (2012).

[45] U. Haagerup, Orthogonal maximal abelian $*$-subalgebras of the $n \times n$ matrices and cyclic n-roots, in "Operator algebras and quantum field theory", International Press (1997), 296–323.

[46] J. Harris, Algebraic geometry, Springer (1992).

[47] R.A. Horn and C.R. Johnson, Matrix analysis, Cambridge Univ. Press (1985).

[48] K. Huang, Introduction to statistical physics, CRC Press (2001).

[49] K. Huang, Fundamental forces of nature, World Scientific (2007).

[50] J.E. Humphreys, Introduction to Lie algebras and representation theory, Springer (1972).

[51] M. Idel and M.M. Wolf, Sinkhorn normal form for unitary matrices, *Linear Algebra Appl.* **471** (2015), 76–84.

[52] M. Jimbo, A q-difference analog of $U(\mathfrak{g})$ and the Yang-Baxter equation, *Lett. Math. Phys.* **10** (1985), 63–69.

[53] V.F.R. Jones, Index for subfactors, *Invent. Math.* **72** (1983), 1–25.

[54] V.F.R. Jones, On knot invariants related to some statistical mechanical models, *Pacific J. Math.* **137** (1989), 311–334.

[55] V.F.R. Jones, Subfactors and knots, CBMS Lecture Notes (1991).

[56] V.F.R. Jones, Planar algebras I, preprint 1999.

[57] M. Kumar, Quantum: Einstein, Bohr, and the great debate about the nature of reality, Norton (2009).

[58] L.D. Landau and E.M. Lifshitz, Mechanics, Pergamon Press (1960).

[59] L.D. Landau and E.M. Lifshitz, The classical theory of fields, Addison-Wesley (1951).

[60] L.D. Landau and E.M. Lifshitz, Quantum mechanics: non-relativistic theory, Pergamon Press (1959).

[61] V.B. Berestetskii, E.M. Lifshitz and L.P. Pitaevskii, Quantum electrodynamics, Butterworth-Heinemann (1982).

[62] S. Lang, Algebra, Addison-Wesley (1993).

[63] P. Lax, Linear algebra and its applications, Wiley (2007).

[64] P. Lax, Functional analysis, Wiley (2002).

[65] P. Lax and M.S. Terrell, Calculus with applications, Springer (2013).

[66] P. Lax and M.S. Terrell, Multivariable calculus with applications, Springer (2018).

[67] B. Lindstöm, Determinants on semilattices, *Proc. Amer. Math. Soc.* **20** (1969), 207–208.

[68] S. Malacarne, Woronowicz's Tannaka-Krein duality and free orthogonal quantum groups, *Math. Scand.* **122** (2018), 151–160.

[69] V.A. Marchenko and L.A. Pastur, Distribution of eigenvalues in certain sets of random matrices, *Mat. Sb.* **72** (1967), 507–536.

[70] M.L. Mehta, Random matrices, Elsevier (2004).

[71] M.A. Nielsen and I.L. Chuang, Quantum computation and quantum information, Cambridge Univ. Press (2000).

[72] S. Raum and M. Weber, The full classification of orthogonal easy quantum groups, *Comm. Math. Phys.* **341** (2016), 751–779.

[73] W. Rudin, Principles of mathematical analysis, McGraw-Hill (1964).

[74] W. Rudin, Real and complex analysis, McGraw-Hill (1966).

[75] W. Rudin, Fourier analysis on groups, Dover (1972).

[76] B.M. Peterson and B. Ryden, Foundations of astrophysics, Cambridge Univ. Press (2010).

[77] B. Ryden, Introduction to cosmology, Cambridge Univ. Press (2002).

[78] D.V. Schroeder, An introduction to thermal physics, Oxford Univ. Press (1999).

[79] J.P. Serre, Linear representations of finite groups, Springer (1977).

[80] G.C. Shephard and J.A. Todd, Finite unitary reflection groups, *Canad. J. Math.* **6** (1954), 274–304.

[81] I.R. Shafarevich, Basic algebraic geometry, Springer (1974).

[82] J.J. Sylvester, Thoughts on inverse orthogonal matrices, simultaneous sign-successions, and tesselated pavements in two or more colours, with applications to Newton's rule, ornamental tile-work, and the theory of numbers, *Phil. Mag.* **34** (1867), 461–475.

[83] P. Tarrago and M. Weber, Unitary easy quantum groups: the free case and the group case, *Int. Math. Res. Not.* **18** (2017), 5710–5750.

[84] N.H. Temperley and E.H. Lieb, Relations between the "percolation" and "colouring" problem and other graph-theoretical problems associated with regular planar lattices: some exact results for the "percolation" problem, *Proc. Roy. Soc. London* **322** (1971), 251–280.

[85] D.V. Voiculescu, K.J. Dykema and A. Nica, Free random variables, AMS (1992).

[86] J. von Neumann, Mathematical foundations of quantum mechanics, Princeton Univ. Press (1955).

[87] S. Weinberg, Foundations of modern physics, Cambridge Univ. Press (2011).

[88] S. Weinberg, Lectures on quantum mechanics, Cambridge Univ. Press (2012).

[89] S. Weinberg, Lectures on astrophysics, Cambridge Univ. Press (2019).

[90] S. Weinberg, Cosmology, Oxford Univ. Press (2008).

[91] D. Weingarten, Asymptotic behavior of group integrals in the limit of infinite rank, *J. Math. Phys.* **19** (1978), 999–1001.

[92] H. Weyl, The theory of groups and quantum mechanics, Princeton Univ. Press (1931).

[93] H. Weyl, The classical groups: their invariants and representations, Princeton Univ. Press (1939).

[94] H. Weyl, Space, time, matter, Princeton Univ. Press (1918).

[95] E. Wigner, Characteristic vectors of bordered matrices with infinite dimensions, *Ann. of Math.* **62** (1955), 548–564.

[96] E. Witten, Quantum field theory and the Jones polynomial, *Comm. Math. Phys.* **121** (1989), 351–399.

[97] S.L. Woronowicz, Compact matrix pseudogroups, *Comm. Math. Phys.* **111** (1987), 613–665.

[98] S.L. Woronowicz, Tannaka-Krein duality for compact matrix pseudogroups. Twisted SU(N) groups, *Invent. Math.* **93** (1988), 35–76.

[99] P. Zinn-Justin, Jucys-Murphy elements and Weingarten matrices, *Lett. Math. Phys.* **91** (2010), 119–127.

[100] B. Zwiebach, A first course in string theory, Cambridge Univ. Press (2004).

Index