Fourier Transform: the discrete, infinite case (discrete, nonperiodic).

Our concern here is with functions/signals $f\colon \mathbb{Z} \to \mathbb{C}$. It is not obvious that the norm of such signal is finite, so we have to make some restriction. We also write $f = (f(n))_{n=-\infty}^{+\infty}$, mimicking the way we write a vector in $\mathbb{C}^N$.

Definition 1. *The ℓ^2-**norm** $\|f\|_{\ell^2(\mathbb{Z})} = \|f\|_{\ell^2}$ of a function $f\colon \mathbb{Z} \to \mathbb{C}$ is*

$$\|f\|_{\ell^2} := \sqrt{\sum_{n=-\infty}^{+\infty} |f(n)|^2} \in [0, +\infty].$$

The space $\ell^2(\mathbb{Z})$ is the set of those $f\colon \mathbb{Z} \to \mathbb{C}$ for which $\|f\|_{\ell^2} < +\infty$.

Obvious examples of function in ℓ^2 are those which vanish for all, but a finite number of arguments in $\mathbb{Z}$: if there is $M > 0$, that is, such that $f(n) = 0$ whenever $|n| > M$. An obvius example of function which is not in ℓ^2 is the constant function $g(n) = 1$, for all n in $\mathbb{Z}$.

There are more interesting examples.

- Let $p > 0$. $f(n) = \frac{1}{n^p}$ for $n \geqslant 1$, and $f(n) = 0$ otherwise, defines a function in ℓ^2 if, and only if, $p > \frac{1}{2}$.

- Let $r \in \mathbb{R}$. $g(n) = r^{|n|}$ defines a function in ℓ^2 if, and only if, $|r| < 1$ and in that case $\|g\|_{\ell^2}^2 = \frac{1}{1 - r^4}$.

After the restriction to functions in $\ell^2(\mathbb{Z})$ is done, we can develop (with care) some linear algebra.

Linear algebra in $\ell^2(\mathbb{Z})$

The Cauchy-Schwarz inequality is the basis upon which rigorous arguments in ℓ^2 can be based.

Theorem 2. *[Cauchy-Schwarz inequality] Let $f, g \in \ell^2(\mathbb{Z})$. Then, there exists*

$$\sum_{j=-\infty}^{+\infty} \overline{g(j)} f(j) := \lim_{M \to \infty} \sum_{|j| \leqslant M} \overline{g(j)} f(j) \in \mathbb{C}.$$

Let $<g, f>_{\ell^2} := \sum_{j=-\infty}^{+\infty} \overline{g(j)} f(j)$. Then, the Cauchy-Schwarz inequality holds:

$$|<g, f>_{\ell^2}| \leqslant \|g\|_{\ell^2} \|f\|_{\ell^2}.$$

Proof. The inequality holds for functions vanishing off a finite set of arguments and the general case can be established passing in the limit. Here are the details of the limiting procedure. Let $S_M = \sum_{|j| \leqslant M} \overline{g(j)} f(j)$. Then, if $M < N$,

$$\begin{aligned}
|S_N - S_M| &= \left| \sum_{M<|j|\leqslant N} \overline{g(j)} f(j) \right| \\
&\leqslant \sqrt{\sum_{M<|j|\leqslant N} |g(j)|^2} \cdot \sqrt{\sum_{M<|j|\leqslant N} |f(j)|^2} \\
&\quad \text{by the Cauchy} - \text{Schwarz inequality in the finite case} \\
&\to 0 \text{ as } M \to \infty \text{ since the series } \sum_{n=-\infty}^{+\infty} |g(n)|^2 \text{ and } \sum_{n=-\infty}^{+\infty} |f(n)|^2 \text{ converge.}
\end{aligned}$$

Hence, by Cauchy test for sequences, S_M has a limit in $\mathbb{C}$, which is -by definition- $\sum_{|j| \leqslant M} \overline{g(j)} f(j)$. Also,

$$\begin{aligned}
|<g, f>_{\ell^2}| &= \lim_{M\to\infty} |S_M| \\
&\leqslant \lim_{M\to\infty} \left(\sqrt{\sum_{|j|\leqslant M} |g(j)|^2} \cdot \sqrt{\sum_{|j|\leqslant M} |f(j)|^2} \right) \\
&= \|g\|_{\ell^2} \|f\|_{\ell^2},
\end{aligned}$$

hence inequality also holds for infinite sums.

Corollary 3. *If $f, g \in \ell^2$ and $\alpha \in \mathbb{C}$, then $f + g \in \ell^2$ and $\alpha f \in \ell^2$. Moreover,*

$$\|f + g\|_{\ell^2} \leqslant \|f\|_{\ell^2} + \|g\|_{\ell^2} \text{ and } \|\alpha f\|_{\ell^2} \leqslant |\alpha| \|f\|_{\ell^2}.$$

- If $f, g \in \ell^2(\mathbb{Z})$ and $\alpha \in \mathbb{C}$, then $(f + g)(j) := f(j) + g(j)$ and $(\alpha f)(j) := \alpha f(j)$ define the sum $f + g$ and the product αf. These operations make $\ell^2(\mathbb{Z})$ into a vector space. The sum $f + g$ might be seen as the superposition of the two signals and the product αf might be seen as the amplification of a signal f by a facor α.

- We want to measure the size of a signal. There are several ways to do that. The simplest one is using the *standard inner product*: $<g, f> = \sum_{j=-\infty}^{+\infty} \overline{g(j)} f(j)$. To measure the size, then, we use the *standard norm*: $\|f\| = \sqrt{<f, f>} = \sqrt{\sum_{j=-\infty}^{+\infty} |f(j)|^2}$.

Properties of the inner product on the vector space $(\mathbb{C}^N, +, \cdot)$.

- $\forall f, g \in \mathbb{C}^N \Rightarrow <g, f> = \overline{<f, g>}$

- $\forall f,g,h\in\mathbb{C}^N \forall a,b\in\mathbb{C}\Rightarrow <h,af+bg> = a<h,f> + b<h,g>$

- $\forall f,g,h\in\mathbb{C}^N \forall a,b\in\mathbb{C}\Rightarrow <af+bg,h> = \bar{a}<f,h> + \bar{b}<g,h>$

- $\forall f\in\mathbb{C}^N \Rightarrow <f,f> \geqslant 0$ and $<f,f \geqslant 0$ if and only if $f=0$

- $\forall f,g\in\mathbb{C}^N \Rightarrow |<g,f>| \leqslant \|f\|\cdot\|g\|$ (Cauchy-Schwarz inequality)

We can use the norm to define a *distance* between $f,g\in\mathbb{C}^N$. We set it to be $\|g-f\|$.

The only property whose verification os not trivial is the Cauchy-Schwarz inequality. We start with an obvious inequality and do some algebra:

$$
\begin{aligned}
0 &\leqslant \sum_{j,k=0}^{N-1} |f(j)g(k)-g(j)f(k)|^2 \\
&= \sum_{j,k=0}^{N-1} (f(j)g(k)-g(j)f(k))\overline{(f(j)g(k)-g(j)f(k))} \\
&= \sum_{j,k=0}^{N-1} |f(j)|^2|g(k)|^2 - \sum_{j,k=0}^{N-1} f(j)\overline{g(j)}\,\overline{f(k)}\,g(k) - \sum_{j,k=0}^{N-1} g(j)\overline{f(j)}\,\overline{g(k)}\,f(k) + \\
&\quad \sum_{j,k=0}^{N-1} |g(j)|^2|f(k)| \\
&= 2\sum_{j=0}^{N-1} |f(j)|^2\cdot\sum_{k=0}^{N-1} |g(k)|^2 - 2\sum_{j=0}^{N-1} f(j)\overline{g(j)}\cdot\overline{\sum_{k=0}^{N-1} g(k)\overline{f(k)}} \\
&= 2\left[\sum_{j=0}^{N-1} |f(j)|^2\sum_{k=0}^{N-1} |g(k)|^2 - \left|\sum_{j=0}^{N-1} f(j)\overline{g(j)}\right|^2\right] \\
&= 2[\|f\|^2\|g\|^2 - |<g,f>|^2].
\end{aligned}
$$

We have then $|<g,f>|^2 \leqslant |<g,f>|^2 + \frac{1}{2}\sum_{j,k=0}^{N-1} |f(j)g(k)-g(j)f(k)|^2 \leqslant \|f\|^2\|g\|^2$, as wished.

We can also deduce the cases of equality.

Corollary 4. *$\forall f,g\in\mathbb{C}^N$: $|<g,f>| = \|f\|\cdot\|g\|$ if and only if $0 = \sum_{j,k=0}^{N-1} |f(j)g(k) - g(j)f(k)|^2$, and the latter holds if and only if there are $a,b\in\mathbb{C}$, not both vanishing, such that $af+bg=0$.*

That is, equality holds if and only if f and g are linearly dependent. (**Exercise:** prove the last "if and only if" in the corollary).

We consider vectors $f = \begin{pmatrix} f(1) \\ \ldots \\ f(N) \end{pmatrix} \in \mathbb{C}^N$, which we can consider as functions f: $\{0, 1, \ldots, N-1\} \to \mathbb{C}$, and matrices $A = (a_{jk})_{j=0,\ldots,N-1} = \begin{pmatrix} a_{00}\, a_{01} \ldots\ldots\ldots\ldots a_{0,N-1} \\ a_{10}\, a_{11} \ldots\ldots\ldots\ldots a_{1,N-1} \\ \ldots \\ a_{N-1,0} a_{N-1,1} \ldots a_{N-1,N-1} \end{pmatrix} = \begin{pmatrix} \boldsymbol{a}_0^t \\ \boldsymbol{a}_{t_1} \\ \ldots \\ \boldsymbol{a}_{t_N} \end{pmatrix}$, where $\boldsymbol{a}_0^t$, ..., $\boldsymbol{a}_N^t$ are the rows of A and the symbol ^{t} stands for *transpose*, $\boldsymbol{a}_j^t = (a_{j0}, \ldots, a_{j,N-1})$, which is the transpose of the column vector $\boldsymbol{a}_j = \begin{pmatrix} a_{j0} \\ \ldots \\ a_{j,N-1} \end{pmatrix}$.

The product follows the usual row-times-column rule:

$$Af = \begin{pmatrix} \sum_{j=0}^{N-1} a_{0,k} f_k \\ \sum_{j=0}^{N-1} a_{1,k} f_k \\ \ldots \\ \sum_{j=0}^{N-1} a_{N-1,k} f_k \end{pmatrix}.$$

Convention about summation. When convenient, we think of vectors $f \in \mathbb{C}^N$ as of N-*periodic* functions $f: \mathbb{Z} \to \mathbb{C}$; $f(j + kN) = f(j)$ whenever $j, k \in \mathbb{Z}$. Clearly, it suffices to know the value of $f(j)$ for just N consecutive values of j, in order to know $f(j)$ for all $j \in \mathbb{Z}$.

This way we have that for $f \in \mathbb{C}^N$,

$$\sum_{j=0}^{N-1} f(j+l) = \sum_{j=0}^{N-1} f(j) \text{ for } l \in \mathbb{Z}.$$

We also extend by periodicity $N \times N$ matrices:

$$a_{j+lN,k+mN} = a_{j,k} \text{ for } j, k, l, m \in \mathbb{Z}.$$

Adjoints, self-adjoint matrices and unitary matrices.

Let $A = (a_{j,k})_{j,k=0,\ldots,N-1}$ be a matrix with entries $a_{j,k} \in \mathbb{C}$. The adjoint of A is the matrix $A^* = (\overline{a_{k,j}})_{j,k=0,\ldots,N-1} = \bar{A}^t$. A matrix A is *selfadjoint* if

$$A = A^*.$$

Theorem 5. *The adjoint A^* of a matrix A satisfies $< A^* f, g > = < f, Ag >$ whenever f, $g \in \mathbb{C}^N$.*

Proof.

$$\begin{aligned}
<A^* f, g> &= \sum_{j=0}^{N-1} \overline{(A^* f)(j)}\, g(j) \\
&= \sum_{j=0}^{N-1} \sum_{k=1}^{N-1} \overline{(A^*)_{j,k} f(k)}\, g(j) \\
&= \sum_{j=0}^{N-1} \sum_{k=1}^{N-1} \overline{\overline{a_{k,j}} f(k)}\, g(j) \\
&= \sum_{j=0}^{N-1} \sum_{k=1}^{N-1} \overline{f(k)}\; a_{k,j} g(j) \\
&= \sum_{j=0}^{N-1} \sum_{k=1}^{N-1} \overline{f(k)}\, (A g)(k) \\
&= <f, A g>.
\end{aligned}$$

A matrix U is unitary if $U^* U = U U^* = \mathrm{Id}$, where Id is the identity matrix $\mathrm{Id} = \begin{pmatrix} 10\ldots0 \\ 01\ldots0 \\ \ldots \\ 00\ldots1 \end{pmatrix}$.

The finite Fourier transform.

Let $f \in \mathbb{C}^N$. Its *finite Fourier transform* $\hat{f} \in \mathbb{C}^N$ is defined by

$$\hat{f}(k) = \sum_{j=0}^{N-1} f(j) \exp\left(-2\pi i\, \frac{j k}{N}\right).$$

We might see $f \mapsto \hat{f} = \mathcal{F} f$ as the linear map induced by the matrix

$$\mathcal{F} := (\mathcal{F}_{j,k})_{j,k=0,\ldots,N-1} = \left(\exp\left(-2\pi i\, \frac{j k}{N}\right)\right)_{j,k=0,\ldots,N-1}.$$

Observe that the matrix $\mathcal{F} = \mathcal{F}^t$ is symmetric: $\mathcal{F}_{j,k} = \mathcal{F}_{k,j}$.

The adjoint matrix is

$$\mathcal{F}^* := (\mathcal{F}^*_{j,k})_{j,k=0,\ldots,N-1} = \left(\overline{\exp\left(2\pi i\, \frac{j k}{N}\right)}\right)_{j,k=0,\ldots,N-1} = \left(\exp\left(2\pi i\, \frac{j k}{N}\right)\right)_{j,k=0,\ldots,N-1}.$$

Theorem 6. *The matrix* $\frac{1}{\sqrt{N}} \mathcal{F}$ *is unitary.*

Proof. Let $z = \exp\left(2\pi i \frac{j}{N}\right)$, with $j = 0, \ldots, N-1$. Then,

$$0 = z^N - 1 = (z-1)(1 + z + \ldots + z^{N-1}) = (z-1) \sum_{l=0}^{N-1} z^l = (z-1) \sum_{l=0}^{N-1} \exp\left(2\pi i \frac{j l}{N}\right),$$

so that either $j=0$ (i.e. $z=1$), or $j=1,\dots,N-1$, and $\sum_{l=0}^{N-1}\exp\left(2\pi i\frac{jl}{N}\right)=0$.

We have then that:

$$\begin{aligned}(\mathcal{F}^*\mathcal{F})_{m,n} &= \sum_{l=0}^{N-1}\mathcal{F}^*_{m,l}\mathcal{F}_{l,n}\\ &= \sum_{l=0}^{N-1}\exp\left(2\pi i\frac{ml}{N}\right)\exp\left(-2\pi i\frac{nl}{N}\right)\\ &= \sum_{l=0}^{N-1}\exp\left(2\pi i\frac{(m-n)\,l}{N}\right)\\ &= \begin{cases} N \text{ if } m-n=0\\ 0 \text{ if } m-n=0\end{cases}\\ &= N\,\mathrm{Id};\end{aligned}$$

hence, $\left(\frac{1}{\sqrt{N}}\mathcal{F}\right)^*\left(\frac{1}{\sqrt{N}}\mathcal{F}\right)=\mathrm{Id}$, as wished.

We immediately deduce two of the most important properties of the Fourier transfrm.

Theorem 7. ***(Fourier inversion formula)****. Let $f\in\mathbb{C}^N$ and let $\hat{f}\in\mathbb{C}^N$ be its Fourier transform. Then, f can be reconstructed as*

$$f(j)=\frac{1}{N}\mathcal{F}^*\hat{f}(j)=\frac{1}{N}\sum_{k=0}^{N-1}\hat{f}(k)\exp\left(2\pi i\frac{kj}{N}\right).$$

Proof. By definition of $\hat{f}$, $\frac{1}{N}\mathcal{F}^*\hat{f}=\frac{1}{N}\mathcal{F}^*\mathcal{F}f=\frac{1}{N}N\,\mathrm{Id}\,f=f$, by the Theorem above.

Theorem 8. ***(Plancherel's Formula)*** *Let $f,g\in\mathbb{C}^N$. Then, $\frac{1}{N}\|\hat{f}\|^2=\|f\|^2$ and $\frac{1}{N}<\hat{g},\hat{f}>=<g,f>$.*

Proof. The first conclusion follows from the second after setting $f=g$. We have:

$$\begin{aligned}<\hat{g},\hat{f}> &= <\hat{g},\hat{f}>\\ &= <\mathcal{F}g,\mathcal{F}f>\\ &= <\mathcal{F}^*\mathcal{F}g,f>\\ &= <N\,\mathrm{Id}\,g,f>\\ &= N<g,f>.\end{aligned}$$

Linearity of the Fourier transform. Let $f,g\in\mathbb{C}^N$ and $\lambda,\mu\in\mathbb{C}$. Then,

$$\mathcal{F}(\lambda f+\mu g)=\lambda\mathcal{F}f+\mu\mathcal{F}g.$$

Convolution and its relation with the Fourier tranform. Let $f,\ g\in\mathbb{C}^N$. Their convolution is $f{*}g\in\mathbb{C}^N$,

$$f{*}g(k)=\sum_{j=0}^{N-1} f(k-j)g(j).$$

Here our convention on periodization becomes useful. For instance, if $N=4$,

$$\begin{aligned}
f{*}g(3) &= f(3)g(0)+f(2)g(1)+f(1)g(0)+f(0)g(-1)\\
&= f(3)g(0)+f(2)g(1)+f(1)g(0)+f(0)g(-1+4)\\
&= f(3)g(0)+f(2)g(1)+f(1)g(0)+f(0)g(3).
\end{aligned}$$

For $f,g\in\mathbb{C}^N$, let their pointwise product be $f\cdot g\in\mathbb{C}^N$, defined by $(f\cdot g)(j)=f(j)g(j)$. We will often drop the multiplication symbol $\cdot$, as usual.

Theorem 9. *Let $f,g\in\mathbb{C}^N$. Then, $\mathcal{F}(f{*}g)=\mathcal{F}f\cdot\mathcal{F}g$,*

$$\widehat{f{*}g}(j)=\hat{f}(j)\hat{g}(j)\text{ for } j=0,\dots,N-1.$$

Proof. We change variable $k-l=m$ and use periodicity from third to fourth line:

$$\begin{aligned}
\widehat{f{*}g}(j) &= \sum_{k=0}^{N-1}(f{*}g)(k)\exp\left(-2\pi i\frac{kj}{N}\right)\\
&= \sum_{k=0}^{N-1}\sum_{l=0}^{N-1} f(k-l)g(l)\exp\left(-2\pi i\frac{kj}{N}\right)\\
&= \sum_{l=0}^{N-1} g(l)\sum_{k=0}^{N-1} f(k-l)\exp\left(-2\pi i\frac{kj}{N}\right)\\
&= \sum_{l=0}^{N-1} g(l)\sum_{m=0}^{N-1} f(m)\exp\left(-2\pi i\frac{(l+m)\,j}{N}\right)\\
&= \sum_{l=0}^{N-1} g(l)\sum_{m=0}^{N-1} f(m)\exp\left(-2\pi i\frac{lj}{N}\right)\exp\left(-2\pi i\frac{mj}{N}\right)\\
&= \sum_{l=0}^{N-1} g(l)\exp\left(-2\pi i\frac{lj}{N}\right)\sum_{m=0}^{N-1} f(m)\exp\left(-2\pi i\frac{mj}{N}\right)\\
&= \hat{g}(j)\hat{f}(j),
\end{aligned}$$

as wished.

From the usual properties of the multiplication between functions, we obtain the basic properties of convolution.

Proposition 10. *Let $f, g, h \in \mathbb{C}^N$ and $\lambda, \mu \in \mathbb{C}$. Then,*

1. $f * g = g * f$ *(commutativity).*

2. $(f * g) * h = f * (g * h)$ *(associativity).*

3. $(\lambda f) * g = \lambda (f * g)$ *(mixed associativity).*

4. $(\lambda f + \mu g) * h = \lambda (f * h) + \mu (g * h)$ *(distributive property).*

Let's prove the first to see how the Fourier transform can help.

$$\begin{aligned} \widehat{g * f} &= \hat{g} \cdot \hat{f} \\ &= \hat{f} \cdot \hat{g} \\ &= \widehat{f * g}, \end{aligned}$$

hence, by Fourier inversion,

$$g * f = \frac{1}{N} \mathcal{F}^* \mathcal{F}(g * f) = \frac{1}{N} \mathcal{F}^* \mathcal{F}(f * g) = f * g.$$

The other proofs are similar.

When we deal with convolution, we are not merely considering the index set $\{0, 1, \ldots, N-1\}$ as a set, but we are also taking into account the sum operation. In fact, we have the expression $j - k$ with $j, k \in \{0, 1, \ldots, N-1\}$ in the definition of convolution (and when $j < k$ we fix things using periodicity, i.e. replacing it by $j - k + N$).

To make this clear and useful in applications, we introduce the notion of *forward shift* (or *translation operator*) Let $j \in \mathbb{Z}$ and $f \in \mathbb{C}^N$. Then, $\tau_j f \in \mathbb{C}^N$ is defined by $\tau_j f(k) := f(k - j)$, where indeed periodicity is used when it does not hold that $0 \leqslant k - j \leqslant N - 1$.

It is clear that $\tau_j \tau_k = \tau_{j+k}$. Moreover, τ_j is a linear operator:

$$\tau_j(\lambda f + \mu g)(k) = (\lambda f + \mu g)(k - j) = \lambda f(k - j) + \mu g(k - j) = \lambda \tau_j f(k) + \mu \tau_j g(k).$$

Let $A: \mathbb{C}^N \to \mathbb{C}^N$ be a linear operator, which is identified as usual with an $N \times N$ matrix. We say that A is *time invariant* if

$$\tau_j A = A \tau_j$$

for all $j \in \mathbb{Z}$. Clearly, this is the same as requiring the equality to hold for $j = 1$.

The idea is that such A's model devices, or phenomena, whose behavior does not change in time: if a signal f is delayed by a unit of time (becoming $\tau_1 f$), then the output $A f$ of the device changes only insomuch as it is delayed by the same amount of time, $A(\tau_1 f) = \tau_1 (A f)$.

There is an interegting relationship between the shifts and the standard basis $\delta_0, \ldots, \delta_{N-1}$. (We might call δ_k the *unit impulse at time* k, for obvious reasons):

$$\tau_j \delta_k = \delta_{k+j},$$

where we use periodicity in the index of δ: $\delta_{k+N} = \delta_k$ for all $k \in \mathbb{Z}$.

Note that any element f in $\mathbb{C}^N$ can be written:

$$f = \sum_{j=0}^{N-1} f(j)\delta_j = \sum_{j=0}^{N-1} f(j)\tau_j\delta_0 = \left(\sum_{j=0}^{N-1} f(j)\tau_j\right)\delta_0,$$

where in the last expression we have highlighted the fact that f can be seen as a linear combination of translations applied to the unit impulse at 0, which looks a remarkably simple way to look at it.

Theorem 11. *Let A be a time invariant linear operator. Then, we can write A as a convolution operator:*

$$Af = (A\delta_0) * f.$$

Viceversa, if $g \in \mathbb{C}^N$, then the operator $f \mapsto Bf := g * f$ is time invariant and $g = A\delta_0$.

Moreover, the norm of A is

$$|||A||| = \max_{0 \leqslant j \leqslant N-1} |\hat{g}(j)|.$$

We might view $A\delta_0 \in \mathbb{C}^N$ as the *response* of the system A to a unit impulse at time $j = 0$.

Proof. Let A be time invariant. Then,

$$\begin{aligned} Af(k) &= A\left(\sum_{j=0}^{N-1} f(j)\delta_j\right)(k) \\ &= \sum_{j=0}^{N-1} f(j)A(\tau_j\delta_0)(k) \\ &= \sum_{j=0}^{N-1} f(j)\tau_j A(\delta_0)(k) \\ &= \sum_{j=0}^{N-1} f(j)A(\delta_0)(k-j) \\ &= f * A(\delta_0)(k), \end{aligned}$$

as wished.

In the other direction, changing variable of summation $l = m - j$,

$$\begin{aligned}
\tau_j(g{*}f)(k) &= g{*}f(k-j) \\
&= \sum_{l=0}^{N-1} g(k-j-l)f(l) \\
&= \sum_{l=0}^{N-1} g(k-j-l)f(l) \\
&= \sum_{l=0}^{N-1} g(k-m)f(m-j) \\
&= \sum_{l=0}^{N-1} g(k-m)\tau_j f(m) \\
&= g{*}(\tau_j f),
\end{aligned}$$

as wished.

We have to prove the estimate for the norm $|||A|||$ of A, wich is of the form $Af = g{*}f$, with $g = A\delta_0$, as we saw before. On the one hand, when $f \neq 0$,

$$\begin{aligned}
\|Af\|^2 &= \|g{*}f\|^2 \\
&= \frac{1}{N}\|\hat{g}\hat{f}\|^2 \\
&= \frac{1}{N}\sum_{j=0}^{N-1} |\hat{g}(j)|^2 |\hat{f}(j)|^2 \\
&\leqslant \max_{0\leqslant j\leqslant N-1} |\hat{g}(j)| \cdot \frac{1}{N}\sum_{j=0}^{N-1} |\hat{f}(j)|^2 \\
&= \Big(\max_{0\leqslant j\leqslant N-1} |\hat{g}(j)|\Big)^2 \|f\|^2.
\end{aligned}$$

Then, $\frac{\|Af\|}{\|f\|} \leqslant \max_{0\leqslant j\leqslant N-1} |\hat{g}(j)|$, hence, $|||A||| \leqslant \max_{0\leqslant j\leqslant N-1} |\hat{g}(j)|$.

In the other direction, we have that $\max_{0\leqslant j\leqslant N-1} |\hat{g}(j)| = \hat{g}(j_0)$ for some $0 \leqslant j_0 \leqslant N-1$. Choose f_0 such that $\hat{f}_0 = \delta_{j_0}$. Then, using Plancherel's formula, we have $\|f_0\|^2 = \frac{1}{N}$ and:

$$\begin{aligned}
\|Af_0\|^2 &= \frac{1}{N}\|\hat{g}\hat{f}_0\|^2 \\
&= \frac{1}{N}\sum_{j=0}^{N-1} |\hat{g}(j)|^2 |\delta_{j_0}(j)|^2 \\
&= \frac{1}{N}|\hat{g}(j_0)|^2 \\
&= \frac{1}{N}\Big(\max_{0\leqslant j\leqslant N-1} |\hat{g}(j)|\Big)^2 \\
&= \Big(\max_{0\leqslant j\leqslant N-1} |\hat{g}(j)|\Big)^2 \frac{1}{N} \\
&= \Big(\max_{0\leqslant j\leqslant N-1} |\hat{g}(j)|\Big)^2 \|f_0\|^2.
\end{aligned}$$

Hence, $|||A||| \geqslant \frac{\|Af_0\|}{\|f_0\|} = \max_{0 \leqslant j \leqslant N-1} |\hat{g}(j)|$. In the end we have $|||A||| = \max_{0 \leqslant j \leqslant N-1} |\hat{g}(j)|$, as we wanted to prove.

Wneh $g \geqslant 0$ the norm $|||A|||$ can be found without before computing the Fourier transform of g.

Remark 12. If $g \geqslant 0$, then $\max_{0 \leqslant j \leqslant N-1} |\hat{g}(j)| = \sum_{k=0}^{N-1} g(k)$.

Let's prove it. In general we have $\max_{0 \leqslant j \leqslant N-1} |\hat{g}(j)| \leqslant \sum_{k=0}^{N-1} |g|(k)$, since for $j = 0, \ldots, N-1$ we have

$$|\hat{g}(j)| = \left| \sum_{k=0}^{N-1} g(k) \exp\left(-2\pi i \frac{jk}{N}\right) \right| \leqslant \sum_{k=0}^{N-1} |g(k)| \cdot \left| \exp\left(-2\pi i \frac{jk}{N}\right) \right| = \sum_{k=0}^{N-1} |g(k)|.$$

In the other direction, $\hat{g}(0) = \sum_{k=0}^{N-1} g(k) = \sum_{k=0}^{N-1} |g(k)|$, since $g \geqslant 0$. Overall,

$$\sum_{k=0}^{N-1} |g(k)| = \sum_{k=0}^{N-1} g(k) = \hat{g}(0) \leqslant \max_{0 \leqslant j \leqslant N-1} |\hat{g}(j)| \leqslant \sum_{k=0}^{N-1} |g|(k),$$

and this proves the Remark.

It can be shown that for general g, such estimate is far from optimal: we can find g's with very small $\max_{0 \leqslant j \leqslant N-1} |\hat{g}(j)|$, yet with very large $\sum_{k=0}^{N-1} |g(k)|$.