As far as we are concerned, a *signal* is function of *time* with values in $\mathbb{C}$. Time can be *discrete* or *continuous* and the signal can be *nonperiodic* or *periodic*:

Continuous Nonperiodic	Discrete Nonperiodic
Continuous Periodic	Discrete Periodicic

In the discrete time case, we assume that the time-interval between successive instants is a fixed time-unit. The case of discrete signals measured at irregularly distributed instants is interesting, but we will not cover it.

- **Continuous nonperiodic.** The signal is a function $f:\mathbb{R}\to\mathbb{C}$.
- **Continuous periodic**, with period $T>0$. The signal is a function $f:\mathbb{R}\to\mathbb{C}$ with the property that $f(t+T)=f(t)\,\forall t\in\mathbb{R}$. Since all that matters is knowledge of f on an interval of length T, we can as well assume that $f:[0,T)\to\mathbb{C}$. (Any other interval having length T would do).
- **Discrete nonperiodic.** The signal is a function $f:\mathbb{Z}\to\mathbb{C}$. We choose the time-unit in such a way the distance between successive instants is one unit.
- **Discrete periodic.** Fix a period $N\geqslant 1$, $N\in\mathbb{N}$. The signal is a function $f:\mathbb{Z}\to\mathbb{C}$ such that $f(n+N)=f(n)\,\forall n\in\mathbb{Z}$. We can identify f with a function $f:\{0,1,\ldots,N\}\to\mathbb{C}$, or with an array (a *vector*) $(f(j))_{j=0}^{N}=\begin{pmatrix} f(0)\\ f(1)\\ \ldots\\ f(N)\end{pmatrix}\in\mathbb{C}^N$.

Discrete periodic

Linear algebra in $\mathbb{C}^N$

The case we consider first is that of discrete, periodic signals, which might be seen as a chapter in linear algebra. We start with a review of basic concepts. We only have to take into account that scalars are complex, rather than real.

- If $f,g\in\mathbb{C}^N$ 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 $\mathbb{C}^N$ 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*: $<g,f>=\sum_{j=0}^{N}\overline{g(j)}f(j)$. To measure the size, then, we use the *standard norm*: $\|f\|=\sqrt{<f,f>}=\sqrt{\sum_{j=0}^{N-1}|f(j)|^2}$.

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:

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

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 1. *$\forall f,g\in\mathbb{C}^N$: $|<g,f>| = \|f\|\cdot\|g\|$ 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).

A *basis* for $\mathbb{C}^N$ is a family $\{f_0,\ldots,f_{N-1}\}$ of elements in $\mathbb{C}^N$ with the property that any element f in $\mathbb{C}^N$ can be written in a unique way as

$$f = \sum_{j=0}^{N-1} \lambda_j f_j,$$

where $\lambda_0,\ldots,\lambda_{N-1}\in\mathbb{C}$ are scalars whose value depend on f and on the basis. The family $\{f_0,\ldots,f_{N-1}\}$ is a basis for $\mathbb{C}^N$ if and only if, for all $\lambda_0,\ldots,\lambda_{N-1}\in\mathbb{C}$,

$$\sum_{j=0}^{N-1} \lambda_j f_j = 0 \text{ if and only if } \lambda_0 = \ldots = \lambda_{N-1} = 0.$$

The basis $\{f_0,\ldots,f_{N-1}\}$ is *orthonormal* if $<f_j,f_k>=0$ whenever $j\neq k$, and $\|f_j\|=1$ for $j=0,\ldots,N-1$.

In this case,

$$f = \sum_{j=0}^{N-1} <f_j,f> f_j.$$

A special ortonormal basis for $\mathbb{C}^N$ consists of the functions δ_j $(j=0,\ldots,N-1)$, defined by

$$\delta_j(k) = \begin{cases} 0 \text{ if } k\neq j\\ 1 \text{ if } k=j \end{cases}$$

The δ_j's are called in different ways in different communities: Dirac's Deltas, Kroenecker's Delta, unit impulses.... They have a privilegde role if we have an interpretation fot the parameter $j \in \{0, 1, ..., N-1\}$. Typically, we see j as "time". Any discrete, periodic signal can be written as linear combination of unit impulses:

$$f = \sum_{j=0}^{N-1} <\delta_j, f> \delta_j = \sum_{j=0}^{N-1} f(j)\delta_j.$$

Linear applications

A *linear application* $A: \mathbb{C}^N \to \mathbb{C}^M$ is a map which satifies

$$A(af + bg) = aAf + bAg$$

whenever $a, b \in \mathbb{C}$ and $f, g \in \mathbb{C}^N$.

A very common case is $M = 1$, $\Lambda: \mathbb{C}^N \to \mathbb{C}$, which we might think as the (scalar) *measurement* Af performed on the signal f. The requirement that Λ be linear much restricts the kind of measurements which are taken into account.

Another common case is $M = N$. We might think of this in terms of a system A performing a linear transformation on the input signals $f \in \mathbb{C}^N$, producing an output signal Af.

Theorem 2. *(Riesz-Fisher: finite version, algebraic part) Let* $\Lambda: \mathbb{C}^N \to \mathbb{C}$ *be linear. Then there exists* $g \in \mathbb{C}$ *such that*

$$\Lambda(f) = <g, f>$$

for $f \in \mathbb{C}^N$. *Moreover,* $g(j) = \overline{\Lambda(\delta_j)}$ $j = 0, ..., N-1$, *provides a formula to recover* g *from* Λ.

Proof.

$$\begin{aligned}\Lambda(f) &= \Lambda\left(\sum_{j=0}^{N-1} f(j)\delta_j\right)\\ &= \sum_{j=0}^{N-1} f(j)\Lambda(\delta_j)\\ &= \sum_{j=0}^{N-1} f(j)\overline{\overline{\Lambda(\delta_j)}}\\ &= <g, f>, \text{with } g = \sum_{j=0}^{N-1} g(j)\delta_j = \sum_{j=0}^{N-1} \overline{\Lambda(\delta_j)}\,\delta_j.\end{aligned}$$

An analogous representation holds in the general case. We only consider $N = M$.

Theorem 3. *(Representation of linear operators on* $\mathbb{C}^N$*: the algebraic part) Let* $A: \mathbb{C}^N \to \mathbb{C}^N$ *be a linear application from* $\mathbb{C}^N$ *to itself. Then, there exists an* $N \times N$ *matrix* $\tilde{A} = [a(k, j)]_{k=0,...,N-1; j=0,...,N-1}$ *such that*

$$A(f) = \tilde{A}f,$$

where on the right we perform row-times-column multiplication, as usual. Moreover, $a(k, j) = <\delta_k, A(\delta_j)>$.

Proof. We can write

$$\begin{aligned} A(f) &= A\left(\sum_{j=0}^{N-1} f(j)\delta_j\right) \\ &= \sum_{j=0}^{N-1} f(j)A(\delta_j) \\ \text{hence,} \\ <\delta_k, A(f)> &= \sum_{j=0}^{N-1} f(j) <\delta_k, A(\delta_j)> \\ &= \sum_{j=0}^{N-1} a(k,j)f(j), \end{aligned}$$

as wished.

From now on we identify $\tilde{A} = A$.

Assuming that $\|f\|$ is a good way to measure the size of the signal $f \in \mathbb{C}^N$, a linear measurement Λ or a linear transformation A will be relatively stable if for $\|f\|$ small, the size $|\Lambda(f)|$ of the measured quantity or the size $\|A(f)\|$ of the transformed signal are small as well. Observe that we have scalar homogeneity w.r.t. f:

$$|\Lambda(a f)| = |a|\,|\Lambda(f)| \text{ and } \|A(a f)\| = |a|\,\|A(f)\|.$$

We can then rescale everything and just consider $\|f\| = 1$, if we wish so.

The *norm* of Λ and A will vbe defined accordingly.

$$\begin{aligned} |||\Lambda||| &= \mathrm{Max}_{\|f\|=1}|\Lambda(f)| = \mathrm{Max}_{f\neq 0}\frac{|\Lambda(f)|}{\|f\|} \\ \text{and} \\ |||A||| &= \mathrm{Max}_{\|f\|=1}\|A(f)\| = \mathrm{Max}_{f\neq 0}\frac{\|A(f)\|}{\|f\|}. \end{aligned}$$

The equality between the expressions on the right in both formulas follows from homogeneity.

Next questions is: is there any reasonable way to compute, or just estimate, the norm of Λ, or of A?

It turns out that the case of Λ $(M = 1)$ is easy, that of A $(M = N)$ much less so.

Theorem 4. *Let* $\Lambda: \mathbb{C}^N \to \mathbb{C}$, $\Lambda(f) = <g, f>$. *Then,* $|||\Lambda||| = \|g\|$.

Proof. By the Cauchy-Schwarz inequality,

$$\begin{aligned} \frac{|\Lambda(f)|}{\|f\|} &= \frac{|<g, f>|}{\|f\|} \\ &\leqslant \frac{\|g\| \cdot \|f\|}{\|f\|} = \|g\|, \end{aligned}$$

then $|||\Lambda||| \leqslant \|g\|$. In the other direction, let $f = g$ if $g \neq 0$ (if $g = 0$ then $\Lambda = 0$ and $|||0||| = 0$ trivially). Then,

$$|||\Lambda||| \geqslant \frac{|\Lambda(g)|}{\|g\|} = \frac{|\langle g, g\rangle|}{\|g\|} = \frac{\|g\|^2}{\|g\|} = \|g\|.$$

This finishes the proof.

When the operator A maps $\mathbb{C}^N$ to $\mathbb{C}^N$, there is no such simple formula, and we have to consider specific classes of operators. The simplest case is that of the *diagonal operators*.

Let $\lambda_0, \lambda_1, \ldots, \lambda_N \in \mathbb{C}$ and form the matrix

$$\mathrm{Diag}(\lambda_0, \ldots, \lambda_N) = \begin{pmatrix} \lambda_0\, 0\, 0 \ldots 0\, 0 \\ 0\, \lambda_1\, 0 \ldots 0\, 0 \\ \ldots \\ 0\, 0\, 0 \ldots 0\, \lambda_{N-1} \end{pmatrix}$$

The effect on a unit impulse at time j is $\mathrm{Diag}(\lambda_0, \ldots, \lambda_N)\delta_j = \lambda_j \delta_j$: the numbers λ_j are the *eigenvalues*, and the unit impulses the *eigenvectors*, of the matrix $\mathrm{Diag}(\lambda_0, \ldots, \lambda_N)$.

Proposition 5. $|||\mathrm{Diag}(\lambda_0, \ldots, \lambda_N)||| = \mathrm{Max}\{|\lambda_0|, \ldots, |\lambda_{N-1}|\}$.

Proof. Exercise.