Shall we take 13 Mar 2012 off? Discrete FT (FFT) 18.1

Discrete Fourier Transform (FFT)

Fourier modes e^{ikx}

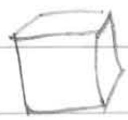
in which of k and x

run over a finite number N of discrete values.

The Fast Fourier Transform (FFT)
evaluates the discrete FT exactly
(to machine precision), and is one of
the most widely used numerical techniques.

Q: Where to use FFT?

A: • In hydro simulations of any kind
where the background is uniform,
and it makes sense to apply
periodic boundary conditions:



Periodic box
of N^3 cells.

- cosmo sims
- hydro sims

• Imaging processing
  - destriping
  - filtering

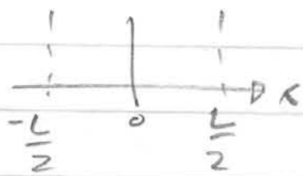
• Signal processing

• Audio processing

Discrete FT

Suppose that $a(x)$ is a continuous, real or complex, periodic function.
Wlog, choose periodic interval to be

$$\left[-\frac{L}{2}, \frac{L}{2}\right]$$



Since $a(x)$ is periodic, its Fourier expansion

$$a(x) = \sum_{m=-\infty}^{\infty} \hat{a}_m e^{2\pi i m x/L}$$

contains only modes $e^{2\pi i m x/L}$ of integral m.

Q: Wavelength λ of Fourier mode $e^{2\pi i m x/L}$?

A: $2\pi i m \lambda / L = 2\pi i \Rightarrow \lambda = \frac{L}{m}$

Q: Frequency ν?

A: $2\pi i m x / L = 2\pi i \nu x \Rightarrow \nu = \frac{m}{L} = \frac{1}{\lambda}$

Q: Angular frequency k?

A: $2\pi i m x / L = i k x \Rightarrow k = \frac{2\pi m}{L}$

Suppose furthermore that $a(x)$ is "smooth" in the specific sense that it contains only modes $|m| \leq \left[\frac{N}{2}\right]$

ie $m = 0, \pm 1, \pm 2, \ldots, \pm \underbrace{\left[\frac{N}{2}\right]}_{\text{denotes largest integer } \leq N/2}$

with the condition that if N is even, then $a_{-\left[\frac{N}{2}\right]} = a_{\left[\frac{N}{2}\right]}$ outermost Fourier coeffs are equal.

$$a(x) = \sum_{m=-[N/2]}^{[N/2]}{}' \; \tilde{a}_m \, e^{2\pi i m x/L}$$

where Σ' means usual sum, except that if N is even, give $\frac{1}{2}$ weight to outermost pair

> Q: Same as counting only one of $a_{-[\frac{N}{2}]}$ or $a_{[\frac{N}{2}]}$
> A: No, as $e^{\cdot}$ is different

ie $\frac{1}{2}\left(a_{-[\frac{N}{2}]} e^{-2\pi i [\frac{N}{2}] x/L} + a_{[\frac{N}{2}]} e^{2\pi i [\frac{N}{2}] x/L}\right)$

which since $a_{-[\frac{N}{2}]} = a_{[\frac{N}{2}]}$ is same as

$$\frac{1}{2} a_{[\frac{N}{2}]} \left(e^{-2\pi i [\frac{N}{2}] x/L} + e^{2\pi i [\frac{N}{2}] x/L}\right)$$

Fourier coeffs of these equal.

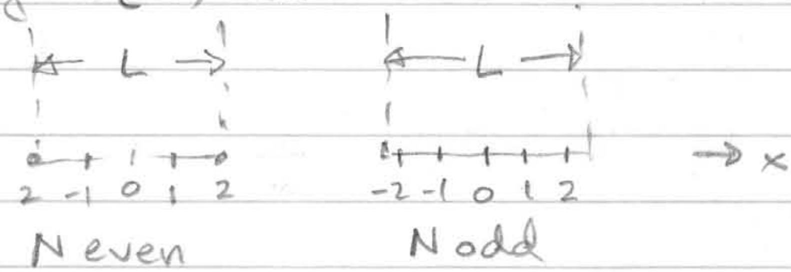
$N = 4$ (even) : $m = 0, \pm 1, \pm 2$
$N = 5$ (odd) : $m = 0, \pm 1, \pm 2$

Let

$$a_n \equiv a(x_n), \quad n = 0, \pm 1, \ldots, \pm [\tfrac{N}{2}]$$

denote values of $a(x)$ at

$$x_n = \frac{L}{N} n$$



The <u>sampling theorem</u> asserts that <u>Fourier coeffs</u> $\tilde{a}_m$ are related to a_n by

$$\boxed{\begin{aligned} \tilde{a}_m &= \frac{1}{N} \sum_{n=-[\frac{N}{2}]}^{[N/2]}{}' \; a_n e^{-2\pi i m n/N} \\ a_n &= \sum_{m=-[\frac{N}{2}]}^{[N/2]}{}' \; \tilde{a}_m e^{2\pi i m n/N} \end{aligned}}$$

constitute a discrete FT pair.

Proof: DIY

By definition

$$a_n = \sum_m{}' \tilde{a}_m e^{2\pi i m n/N}$$

Take

$$\sum_n{}' a_n e^{-2\pi i m n/N}$$

$$= \sum_n{}' \left(\sum_{m'}{}' \tilde{a}_{m'} e^{2\pi i m' n/N} \right) e^{-2\pi i m n/N}$$

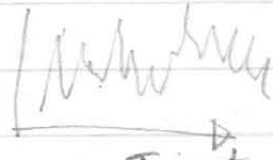
$$= \sum_{m'}{}' \tilde{a}_{m'} \underbrace{\sum_n{}' e^{2\pi i (m'-m) n/N}}_{\text{can show this } = N\delta_{m'm}}$$

$$= N \tilde{a}_m \quad \checkmark$$

$$\boxed{\begin{array}{l}\text{The discrete FT effectively assumes} \\ \text{(a) } x \text{ is periodic} \quad (\Rightarrow k \text{ discrete}) \\ \text{(b) } k \text{ is periodic} \quad (\Rightarrow x \text{ discrete}).\end{array}}$$

In literature "FT" is often treated as synonymous with "FFT", but it ain't so.

Ex PS8/ Have time series, eg the brightness of a variable AGN.
You think the series contains periodicities, so you want to FT.



You only observe for a finite time.
So you decide to FFT. In doing so, you
(a) truncate time series
(b) make it periodic

Truncation ⇒ ringing

Trunc in real space ⇒ ringing in Fourier space.
" Fourier " ⇒ ——— real space.

Maybe easier to understand latter.

Trunc in Fourier space
⇒ remove freqs above some cutoff.
⇒ typically introduces visible
oscillations near cutoff freq.

fourier.nb take $N_{seq} = 2^4, 2^8$.

Make periodic ⇒ aliasing

Make periodic in real space ⇒ aliasing in real space
——— Fourier ⇒ ——— fourier "

E.g. in FFT, freq is periodic, so
frequency $m + iN \equiv$ frequency m
$\uparrow$ any integer

If signal contains freqs higher
than your FFT admits, then
those freqs will be aliased to
lower freq.

Nyquist frequency

The highest frequency that an FFT
of N points admits is:

$m = \left[\frac{N}{2}\right]$ ← Nyquist frequency

or in real units

$\nu = \frac{m}{L} = \frac{[N/2]}{L}$ ← box size.

Q: Why factor of $1/2$?
A: Essentially because N points encode Re & Im parts of $[\frac{N}{2}]$ complex numbers.

Organization of FFT in codes

FFT maps

$$a_n \underset{\text{backward FFT}}{\overset{\text{forward FFT}}{\rightleftarrows}} \hat{a}_m$$

Typically a_n is a real sequence.
If so, then

$$\hat{a}_{-m} = \tilde{a}^*_m$$

ie $\hat{a}_m$ is complex, with $\hat{a}_m$ and $\hat{a}_m$ effectively encoding the Re and Im parts of $\hat{a}_m$:

$$\mathrm{Re}\,(\tilde{a}_m) = \frac{1}{2}(\hat{a}_m + \hat{a}_{-m})$$
$$\mathrm{Im}\,(\hat{a}_m) = \frac{1}{2i}(\hat{a}_m - \hat{a}_{-m})$$

Code typically inputs real periodic sequence as

$$a_0, a_1, a_2, \ldots, a_{[\frac{N}{2}]}, \underbrace{a_{-[\frac{N}{2}]+1}}_{=a_{[\frac{N}{2}]+1}}, \ldots, \underbrace{a_{-1}}_{=a_{N-1}}$$

and outputs

$$N\hat{a}_0 = \Sigma a_n$$
$$N\,\mathrm{Re}\,\hat{a}_1 = \mathrm{Re}\,\Sigma a_n e^{-2\pi i mn/N}$$
$$N\,\mathrm{Im}\,\hat{a}_1 = \mathrm{Im} \quad ''$$
$$\vdots$$
$$N\,\mathrm{Re}\,\hat{a}_{[\frac{N}{2}]}$$

Mathematica outputs FFT as complex seq

$$\tilde{a}_0, \hat{a}_1, \ldots, \tilde{a}_{[\frac{N}{2}]}, \underbrace{\hat{a}_{-[\frac{N}{2}]+1}}_{=a_{[\frac{N}{2}]+1}}, \ldots, \underbrace{\tilde{a}_{-1}}_{=\tilde{a}_{N-1}}$$

Mathematica defines FFT

1. with factor $\frac{1}{\sqrt{N}}$ in normalization
2. as complex conjugate of our convention

fourier.nb

With any FFT, check:

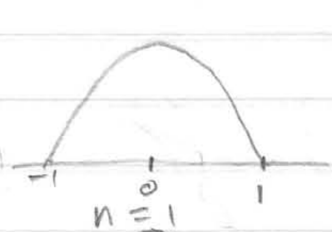
1. Normalization.
2. Phase convention.
3. Ordering of sequences.

Q?
<u>Remedy for ringing caused by truncation</u>
Multiply by a window that goes to zero at edges of data.

Q?
<u>Remedy for aliasing caused by making periodic</u>
Pad data with the mean of the data. Or, pad mean-subtracted or windowed data with zeros. Padding to at least double length eliminates aliasing between start and end of data

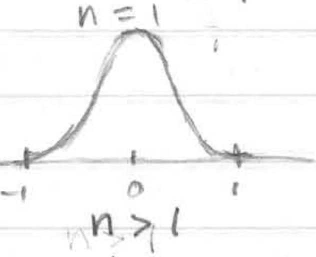
<u>Windowing</u>
Convolution theorem (Proof later):
Multiplication in real space ≡ convolution in Fourier space
——— Fourier ——— ≡ ——— real ———

agn.nb Andrew's favorite windows:
To window over interval [-1, 1],



fourier.nb $w(x) = (1 - x^2)^n$

Why?



(a) non-zero over finite interval;
(b) adjustable parameter n gives useful flexibility.
(c) As n increases, window goes to zero at edges ever more smoothly.
(d) As $n \rightarrow$ large, window $\rightarrow$ Gaussian. Note that FT of a Gaussian is a Gaussian.

Q: To get narrow conv window in F space? A: Broad mult window in real space

Data length

FFT algorithm is fastest when N is a product of small primes (2, 3, 5, ...). Fastest is $N = 2^n$.
Slowest is N = prime number, in which case FFT reverts to SFT.

Correlation function

Q: Examples? Let $a(x)$ be a real random field that is statistically translation invariant. Its lowest order moments are:

1. Mean $\bar{a} = \langle a(x) \rangle$
2. Covariance, or correlation function

$$\xi(\underbrace{|x'-x|}) = \langle \Delta a(x') \, \Delta a(x) \rangle$$

function only of:
- separation $x'-x$ by translation invariance,
- abs value $|x'-x|$ by pair exchange symmetry

where $\Delta a(x) \equiv a(x) - \bar{a}$.

Note: correlation function ξ is often defined to be $1/\bar{a}^2$ of this, so that ξ becomes dimensionless.

Power spectrum P_k

Covariance of Fourier modes a_k is

$\langle \Delta a_{k'}^* \Delta a_k \rangle$ take Fourier series, for simplicity

$$= \left\langle \left(\int_{-\pi}^{\pi} \Delta a(x') \frac{e^{-ik'x'}}{\sqrt{2\pi}} dx' \right)^* \left(\int_{-\pi}^{\pi} \Delta a(x) \frac{e^{-ikx}}{\sqrt{2\pi}} dx \right) \right\rangle$$

$$= \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} \langle \Delta a(x') \Delta a(x) \rangle \frac{e^{ik'x'}}{\sqrt{2\pi}} \frac{e^{-ikx}}{\sqrt{2\pi}} dx' dx$$

$$= \underbrace{\int_{-\pi}^{\pi} \xi(|x'-x|) e^{ik'(x'-x)} d(x'-x)}_{= \sqrt{2\pi}\, P_{k'}} \quad \underbrace{\int_{-\pi}^{\pi} \frac{e^{i(k'-k)x}}{2\pi} dx}_{= \delta_{k'k}}$$

power spectrum = FT of corr fn

$= \sqrt{2\pi}\, \delta_{k'k} P_k$.

Q: P_k real, complex, ...?

A: $P_k \propto \langle \Delta a_k^* \Delta a_k \rangle = \langle |\Delta a_k|^2 \rangle$ is +ve real.

Can think of corr fn as matrix

$$\xi(|x'-x|) \equiv \xi_{x'x} \equiv \langle \Delta a_{x'} \Delta a_x \rangle$$

and power spectrum as diagonal matrix

$$P_{k'k} \equiv P_k \delta_{k'k} \equiv \langle \Delta a_{k'}^* \Delta a_k \rangle$$

Q: Why is covariance of Fourier modes diagonal, whereas covariance of real space modes is not?

A: Statistical translation invariance.
Fact that Fourier modes are uncorrelated means power spectrum is greatly superior to corr fn as a statistical tool.
This is true even though

The power spectrum P_k is the FT of the corr fn $\xi(x)$, and therefore P_k and $\xi(x)$ contain identical information.

<u>Warning</u>
Conventions for normalization of P_k, disposition of 2π's, vary.
Discover the convention in your field, and stick to it.

<u>The mean is encoded by zeroth Fourier mode</u>
The zeroth Fourier eigenmode is

$$\frac{e^{i\vec{k}\vec{x}^0}}{\sqrt{2\pi}} = \frac{1}{\sqrt{2\pi}} \quad \text{is constant.}$$

All other Fourier modes are orthog to this.

Thus

$$\Delta a_k \equiv \text{FT of } \Delta a(x) = \begin{cases} 0 & k = 0 \\ a_k & k \neq 0 \end{cases}.$$

So

$$\langle \Delta a^*_{k'} \Delta a_k \rangle = \begin{cases} 0 & k = 0 \\ \langle a^*_{k'} a_k \rangle & k \neq 0 \end{cases}$$

<u>Estimate power spectrum from FFT of data</u>
Estimate

$$P_k = a_k^* a_k = |a_k|^2$$

for $k = 1, 2, \ldots, \left[\frac{N}{2}\right]$.

Shot noise
= $\boxed{\text{self-pair contribution to any statistic}}$.
This definition of shot noise is all you ever need to remember about it.

Shot noise arises whenever a random field is realized by discrete objects
- photons
- galaxies

fourier.nb
Example / PS8.
Estimate power at wavenumber m:

$$a_m^* a_m = \frac{1}{N^2}\left(\sum_{n'}{}' a_{n'}\, e^{-2\pi i m n'/N}\right)^* \left(\sum_n{}' a_n e^{-2\pi i m n/N}\right)$$

$$= \frac{1}{N^2} \sum_{n'}{}' \sum_n{}' a_{n'} a_n\, e^{2\pi i m (n'-n)/N}$$

contains self-pair contributions from individual x-ray photons.
Necessarily $n' = n$ for self-pairs, but each bin contains many photons.
If a_n = # phots in bin n, then

$$(a_m^* a_m)_{\text{shot}} = \frac{1}{N^2} \sum_n a_n = \text{constant indep}^{\text{t}} \text{ of } m$$

If $a_n = \overbrace{W_n}^{\text{weight}}$ times (# phots in bin n) then

$$(a_m^* a_m)_{\text{shot}} = \frac{1}{N^2} \sum_n W_n^2 a_n$$

Why? Go back to definition of shot noise, and think hard.

Cross-correlation

between 2 random fields $a(x)$, $b(x)$
is covariance

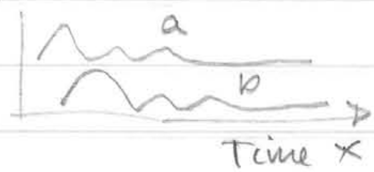
$$\xi^{ab}_{cross}(\underbrace{x'-x}) = \langle \Delta a(x') \, \Delta b(x) \rangle$$

function only of
$x'-x$ by translation invariance

Can be that $\xi^{ab}_{cross}(-(x'-x)) \neq \xi^{ab}_{cross}(x'-x)$.

Q: Example where " $\neq$ " ?

A: a, b two light curves, where
1 lags the other.



Q: Example where ξ^{ab}_{cross} is a function of $|x'-x|$ only?

A: isotropic 2D field (CMB)
or 3D " (galaxies)

Cross-power

$P_k^{ab} = $ FT of $\xi^{ab}(x)$.

Q: P_k^{ab} real, complex, ...?

A: Complex in general,
real if $\xi^{ab}(x)$ is func of $|x|$ only.
May be -ve.
Can look for evidence for lag
$(\xi^{ab}(-x) \neq \xi^{ab}(x))$
by looking at $\mathrm{Im}(P_k^{ab})$.

Shot noise in cross-power

If two fields $a(x)$, $b(x)$ share no members (no common photons, galaxies, ...) then the shot noise is ... (Q?)
zero.
This is how WMAP CMB power spectrum is computed - from cross-powers between maps at distinct wavelength bands - so eliminating uncertainty arising from shot noise subtraction.

If two fields share members, then there will be shot noise from the shared members.

<u>FFT algorithm (sketch)</u> See Numerical Recipes.

Recursively split FFT of length N into 2 FFTs of length $\frac{N}{2}$, if N even.

lowest prime factor, more generally

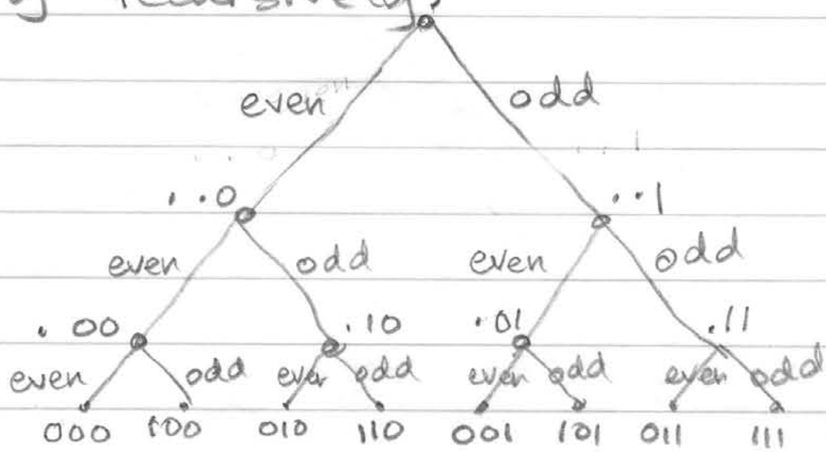
$$\hat{a}_m = \frac{1}{N} \sum{}' a_n e^{-2\pi i m n/N}$$

split sum into even, odd

$$= \frac{1}{N}\left[\sum{}' a_{2n} \underbrace{e^{-2\pi i m 2n/N}}_{= e^{-2\pi i m n/(\frac{N}{2})}} + \sum_n{}' a_{2n+1} \underbrace{e^{-2\pi i m(2n+1)/N}}_{e^{-2\pi i m/N}\, e^{-2\pi i m n/(\frac{N}{2})}} \right]$$

$$= \frac{1}{N}\left[\underbrace{\sum{}' a_{2n} e^{-2\pi i m n/\frac{N}{2}}}_{\text{FFT of length } N/2} + \underbrace{e^{-2\pi i m/N}}_{\text{factor}} \underbrace{\sum{}' a_{2n+1} e^{-2\pi i m n/\frac{N}{2}}}_{\text{FFT of length } N/2} \right]$$

Apply recursively:



← Note bottom layer elements are in bit-reversed order.

Take a_n, write $n = \underbrace{\cdots\cdots}_{0\text{s } \& \text{ 1s}}$ in binary.

Bottom layer is $N = 2$ transform:

$$\hat{a}_m = \frac{1}{2}\left[a_0 + a_1 \underbrace{\frac{1}{2}\left(e^{-2\pi i m/2} + e^{2\pi i m/2}\right)}_{= \begin{cases} 1 & m = 0 \\ -1 & m = \pm 1 \end{cases}} \right]$$