

Outline

Fourier Transforms in Crystallography  
They are EVERYWHERE

L. F. Ten Eyck

2005 IUCr Crystallographic Computing School

- 1 Introduction
  - Physics
  - Mathematics

Outline

- 1 Introduction
  - Physics
  - Mathematics

Diffraction is a Physical Fourier Transform

- Scattering of a wave by an object is the Fourier transform of the scattering function.

$$F(h) = \int_{-\infty}^{+\infty} f(x) \exp(2\pi i h \cdot x) dx$$

- Fourier transform of periodic functions are discrete.

$$F(h) = 0 \text{ if } h \text{ is not an integer.}$$

- Crystals are (**MOSTLY**) periodic.
- The scattering function can thus be reconstructed by summing the Fourier series

$$f(x) = \sum_{h=-\infty}^{\infty} F(h) \exp(-2\pi i h \cdot x)$$

Outline

- 1 Introduction
  - Physics
  - Mathematics

General Properties of Fourier Transforms

The FT of a real function is Hermitian Symmetric, and Vice-Versa

- If  $f(x)$  is real,  $F(h) = F^*(-h)$ .
- If  $F(h) = F^*(-h)$ ,  $f(x)$  is real.

Gaussian Functions

- If  $f(x) = A \exp(-ax^2)$  then

$$F(h) = \frac{A}{\sigma^3} \exp(-\pi h^2 / \sigma^2)$$

where  $\sigma^2 = a/4\pi$ .

- Wide Gaussians in real space are narrow in reciprocal space
- Narrow Gaussians in real space are wide in reciprocal space

## General Properties of Fourier Transforms

Continued...

### Convolution Theorem

- The Fourier transform of a convolution is the product of the Fourier transforms of the functions

$$C(\tau) = \int_{-\infty}^{\infty} f(t)g(\tau - t)dt$$

$$C(h) = F(h)G(h)$$

$$\therefore C(\tau) = \int_{-\infty}^{\infty} F(h)G(h) \exp(2\pi i h \cdot \tau) dh$$

FFTs

## Convolutions

The Second Most Important Concept in Crystallography

### Crystals as Convolutions

- Molecules: Atomic distribution functions and point locations
- Unit cells: Mosaicity distribution
- Material heterogeneity: Distribution of atomic types
- Crystal shape

### Measurements as Convolutions

- Beam Profile
- Aperture and Collimator
- Dispersion

FFTs

## Finite Discrete Fourier Transforms

### Definition

Let  $X(t)$  be a sequence of  $N$  points with  $t = 0, 1, \dots, N-1$ . The Discrete Fourier Transform of  $X(t)$  and the inverse DFT are

$$X(\hat{t}) = \sum_{t=0}^{N-1} X(t) \exp(2\pi i t \hat{t} / N)$$

$$X(t) = \frac{1}{N} \sum_{\hat{t}=0}^{N-1} X(\hat{t}) \exp(-2\pi i t \hat{t} / N)$$

The function  $e(x) = \exp(2\pi i x)$  has the properties that  $e(x) = 1$  if  $x$  is an integer,  $e(-x) = e^*(x)$ ,  $e(x+y) = e(x)e(y)$ , and thus  $e(x+N) = e(N)e(x) = e(x)$ .

FFTs

## Properties of DFT's

DFT's are DIFFERENT

### Continuous

Orthogonality:

$$\int_{-\infty}^{+\infty} e(i\hat{t})e(-tu)dt = \delta(\hat{t}-u)$$

Periodicity:

$$\text{If } f(x) = f(x + Na)$$

$F(h)$  is NOT periodic

### Discrete

Orthogonality:

$$\sum_{t=0}^{N-1} e(i\hat{t})e(-tu) = N\delta(\text{mod}(\hat{t}-u, N))$$

Periodicity: for all integer  $k$ ,

$$X(t) = X(t + kN)$$

$$X(\hat{t}) = X(\hat{t} + kN)$$

FFTs

## Factoring DFT's

Let  $N=AB$ . Then  $\hat{t} = \hat{a} + \hat{b}A$  and  $t = b + aB$  cover all values of  $\hat{t}$  and  $t$  as  $\hat{a}, \hat{b} = 0, \dots, A-1$  and  $b, \hat{a} = 0, \dots, B-1$ . Then

$$\begin{aligned} \hat{X}(\hat{a} + \hat{b}A) &= \sum_{b=0}^{B-1} \sum_{a=0}^{A-1} X(b + aB) e\left(\frac{(b + aB)(\hat{a} + \hat{b}A)}{AB}\right) \\ &= \sum_{b=0}^{B-1} \sum_{a=0}^{A-1} X(b + aB) e\left(\frac{\hat{a}b}{AB} + \frac{\hat{a}a}{A} + \frac{\hat{b}b}{B} + a\hat{b}\right) \\ &= \sum_{b=0}^{B-1} \sum_{a=0}^{A-1} X(b + aB) e\left(\frac{\hat{a}b}{AB}\right) e\left(\frac{\hat{a}a}{A}\right) e\left(\frac{\hat{b}b}{B}\right) \end{aligned}$$

because  $e(a\hat{b}) = 1$ .

FFTs

## Properties of DFT's

### Factoring (cont.)

$$\hat{X}(\hat{a} + \hat{b}A) = \sum_{b=0}^{B-1} e\left(\frac{\hat{b}\hat{b}}{B}\right) \left\{ e\left(\frac{\hat{a}\hat{b}}{AB}\right) \sum_{a=0}^{A-1} X(aB + b) e\left(\frac{a\hat{a}}{A}\right) \right\}$$

This reduces the  $N^2$  problem of computing  $N$  Fourier coefficients, each requiring a sum of  $N$  terms, to  $B$  problems of size  $A^2$  followed by  $A$  problems of size  $B^2$ .

$$BA^2 + AB^2 = AB(A + B) \approx N \log_2 N$$

FFTs

## Properties of DFT's

Factoring can be used to design code that

- Is specialized for real and Hermitian symmetric data
- Incorporates crystallographic symmetry elements
- Factors of two and four do not require multiplication
- Details in Ten Eyck (1973), *Acta Cryst.* **A29**, 183-191.

## Properties of DFT's

### DFT's are Convolutions

- A sequence of samples is a convolution of the sampled function with a  $\delta$  function. If  $f(u)$  is periodic with period  $a$ ,

$$\begin{aligned} X(t) &= f\left(\frac{ta}{N}\right) \\ &= \int_0^a f(u) \delta\left(t - u\frac{N}{a}\right) du \end{aligned}$$

- $\hat{X}(\hat{t})$  is the sum of Fourier coefficients spaced  $N$  points apart.

$$\hat{X}(\hat{t}) = \sum_{k=-\infty}^{+\infty} \mathbf{F}(\hat{t} + kN)$$

- This places a lower limit on useful values of  $N$ .

## Properties of DFT's

### Multiple conventions are in use

- Normalization:
  - None
  - $1/\sqrt{N}$  in each direction
  - $1/N$  in one direction, 1 in the other
- Sign of exponent
  - $\exp(+2\pi i \frac{t\hat{t}}{N})$  forward,  $\exp(-2\pi i \frac{t\hat{t}}{N})$  reverse
  - $\exp(-2\pi i \frac{t\hat{t}}{N})$  forward,  $\exp(+2\pi i \frac{t\hat{t}}{N})$  reverse

## Calculation of Fourier Summations

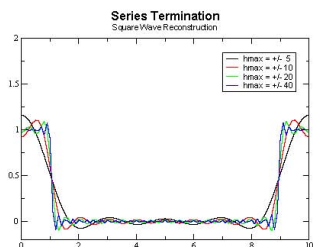
### Summations

The FFT is a very fast way of calculating Fourier summations.

- Direct cost is  $MN$  where  $M$  is the number of Fourier coefficients and  $N$  is the number of grid points.
- Typical values are  $M = 10,000$  and  $N = 128^3 = 2097152$ .
- FFT cost is  $N \log_2 N$ .
- Cost ratio for this case is  $\frac{M}{\log_2 N} = 476$  in favor of FFT.

## Series Termination

### Summations



- Square wave of length 100
- Fourier series truncated at 5, 10, 20, and 40
- Note Gibbs' Phenomenon – the overshoot is **constant**

## Calculation of Density Maps

### Tradeoffs

#### Memory Requirements

- Can calculate only an asymmetric unit
  - But sometimes P1 is convenient
- Transform in place
- Hermitian to Real saves factor of 2

#### Speed

- Hermitian to Real saves factor of 2
- Use of symmetry saves factor of  $N_{symops}$

## Structure Factors

## Procedure

- 1 Determine grid spacing and "blur" from resolution limits
- 2 Generate model density from coordinates
  - 1 Use multiple Gaussian scattering factors
  - 2 Transform of Gaussian is Gaussian
  - 3 Mapping of elliptical (anisotropic) density is no more work than spherical density
  - 4 Wrap across boundaries of asymmetric unit and unit cell
- 3 Calculate Real to Hermitian Fourier transform
- 4 Remove "blur"

FFTs

## Structure Factors

Resolution and "Blur"

## Resolution and Accuracy

- 1 Aliasing causes serious contamination by high-order harmonics

$$\mathbf{F}_{FFT}(h) = \sum_{t=-\infty}^{+\infty} \mathbf{F}_c(h + tN)$$

- 2 Structure factors have generally Gaussian distribution – as in Wilson plots.
- 3 Convolution with a Gaussian "blur" causes the error to fall off faster than linearly.

FFTs

## Structure Factors

Resolution and "blur"

- 1 Add "blur"  $B^0$  to every atom, noting that  $(\sin \theta)/\lambda = h/2a$

$$\begin{aligned} \mathbf{F}_B(h) &= \exp(-B^0 h^2/4a^2) \mathbf{F}(h) \\ \mathbf{F}_B(h+N) &= \exp(-B^0 (h+N)^2/4a^2) \mathbf{F}(h+N) \\ \mathbf{F}_B(h-N) &= \exp(-B^0 (h-N)^2/4a^2) \mathbf{F}(h-N) \end{aligned}$$

- 2 The fractional contamination by the two largest contaminating structure factors can be made as small as desired by appropriate choice of  $B^0$  and  $N$ .
- 3 Choice depends on resolution and cell edge.
  - 1  $N \geq 3(a/h_{max})$  and  $B^0 > 15$  is **minimum** for medium resolution
  - 2 At high resolution,  $B^0 \approx 5$  and  $N$  large is required (Tronrud).
- 4 Details in Ten Eyck, *Acta Cryst.* **A33**, 486-492 (1977)

FFTs

## Least Squares Gradients

- 1 If  $\Phi(\mathbf{p}) = \sum_h W(h) [k |\mathbf{F}_o(h)| - |\mathbf{F}_c(h, \mathbf{p})|]^2$ , the derivatives with respect to the parameters  $\mathbf{p}$  are given by

$$\frac{d\Phi(\mathbf{p})}{d\mathbf{p}} = -2 \sum_h W(h) [k |\mathbf{F}_o(h)| - |\mathbf{F}_c(h, \mathbf{p})|] \frac{d|\mathbf{F}_c(h, \mathbf{p})|}{d\mathbf{p}}$$

- 2  $\mathbf{F}_c(h, \mathbf{p}) = \sum_{atoms} f_k(h) \exp(-B_k \sin^2 \theta / \lambda^2) \sum_{pos} \exp(2\pi i \mathbf{h}^T \mathbf{R}_j \mathbf{x}_k)$  so the product above is the product of two Fourier transforms – for the difference map, and for the density due to each atom.
- 3 The sum of products of Fourier transforms can be calculated by convoluting the gradients of the model atomic density with the difference map.

Details are in Tronrud, Ten Eyck and Matthews, *Acta Cryst.* **A43**, 489-501 (1987).

FFTs

## Random Observations

## FFTW

- 1 FFTW (<http://www.fftw.org>) is OUTSTANDING.
- 2 FFTW IS NOT DETERMINISTIC if it is tuning.

## Prototype and Explore

MATLAB (\$\$\$) and Octave (free) both have excellent facilities for playing with Fourier transforms

## Links

- 1 MathWorld: <http://mathworld.wolfram.com> – lots of math; use the search functions.
- 2 Wikipedia: <http://www.wikipedia.org> – excellent material on DFT's and pointers to more.
- 3 FFTW: again, excellent discussion.

FFTs