**PDE in Finance, Spring 2008,**
http://www.math.nyu.edu/faculty/goodman/teaching/PDEfin/index.html
Written by Jonathan Goodman. Please report errors, sources of confusion, and typos on the class Blackboard site.
Last revised: February 8, 2008.

**Section 2: The Fourier transform and the Heat equation**

# 1 The Fourier transform

The Fourier transform is a powerful tool in studying functions in general and solution of PDEs in particular. In this class we explain the Fourier transform and show how it gives a different way to represent a general solution of the initial value problem for the free space heat equation. This explicit solution confirms things about the solution we learned from the Green's function representation. It also explains why the heat equation does not go backwards. In future classes, we will use the Fourier transform to give explicit solutions to more complicated PDEs that arise in finance.

## 1.1 DFT and FFT

The *discrete Fourier transform*, or *DFT*, is the computational side of the Fourier transform. The *fast Fourier transform*, or *FFT*, is a fast algorithm that computes the DFT. We explain the DFT here both because it is a useful computational tool and because it allows us to derive the Fourier transform through a limiting procedure.

We review some linear algebra to establish notation. The *inner product* between two $N$ component complex vectors is

$$\langle v, w \rangle \;=\; \sum_{k=1}^{N} \overline{v}_k w_k \;.$$

The *Euclidian length*, or $l^2$ *norm* is

$$|v| \;=\; \sqrt{\langle v, v\rangle} \;=\; \sqrt{\sum_{k=1}^{N} |v_k|^2} \;.$$

Vectors $v$ and $w$ are *orthogonal* if $\langle v, w\rangle = 0$. Vectors $v_\alpha$, for $\alpha = 1, \ldots, N$ are an *orthonormal basis* if $\langle v_\alpha, v_\beta\rangle = 0$ if $\alpha \neq \beta$, and $|v_\alpha| = 1$ for all $\alpha$. If the $v_\alpha$ are an orthonormal basis and $f$ is any other $N$ component vector, then

$$f \;=\; \sum_\alpha w_\alpha v_\alpha \;,$$

where the weights $w_\alpha$ are given by

$$w_\alpha \;=\; \langle v_\alpha, f\rangle \;.$$

Furthermore, we have the *Plancharel identity*

$$|f|^2 \;=\; \sum_{k=1}^{N} |f_k|^2 \;=\; \sum_{\alpha=1}^{N} |w_\alpha|^2 \; .$$

The DFT is a particular orthonormal basis. For each integer, $\alpha$, let $v_\alpha$ be an $N$ component complex vector with entries

$$v_{\alpha k} = \frac{1}{\sqrt{N}} e^{2\pi i \alpha k/N} \; . \tag{1}$$

The factor $\sqrt{N}$ makes the vectors normalized:

$$|v_\alpha|^2 \;=\; \frac{1}{N} \sum_{k=1}^{N} \left| e^{2\pi i \alpha k/N} \right|^2 \;=\; 1 \; .$$

To calculate $\langle v_\alpha , v_\beta \rangle$, we note the geometric series sum, for any complex number $z \neq 1$ and integers $a < b$,

$$z^a + z^{a+1} + \cdots + z^b \;=\; \frac{z^{b+1} - z^a}{z - 1} \; .$$

Now, if $\alpha$ and $\beta$ are integers with $\alpha \neq \beta$, $1 \leq \alpha \leq N$ and $1 \leq \beta \leq N$, then

$$\langle v_\alpha , v_\beta \rangle \;=\; \frac{1}{N} \sum_{k=1}^{N} \exp\left( 2\pi i (\beta - \alpha) k / N \right) \;=\; \frac{1}{N} \sum_{k=1}^{N} z^k \; ,$$

where

$$z \;=\; \exp\left( \frac{2\pi i (\beta - \alpha)}{N} \right) \; .$$

If $z \neq 1$, then

$$\frac{1}{N} \sum_{k=1}^{N} z^k \;=\; \frac{z}{N(z-1)} \left( z^N - 1 \right) \; .$$

As long as $\alpha$ and $\beta$ are integers, $z^N = 1$ because

$$z^N \;=\; \exp\left( 2\pi i (\beta - \alpha) N/N \right) \; .$$

If $\alpha \neq \beta$ then $z \neq 1$ because $\beta - \alpha \neq 0$ and

$$-1 < \frac{\beta - \alpha}{N} < 1 \; .$$

The DFT of a vector $f$ is its representation in terms of the $v_\alpha$. The expansion coefficients are called discrete *Fourier coefficients* and are written $\widehat{f}_\alpha$. Written out in components, we have

$$f_k \;=\; \frac{1}{\sqrt{N}} \sum_{\alpha=1}^{N} \widehat{f}_\alpha e^{2\pi i \alpha k/N} \; , \tag{2}$$

where

$$\widehat{f}_\alpha = \frac{1}{\sqrt{N}} \sum_{k=1}^{N} e^{-2\pi i \alpha k/N} f_k \; . \tag{3}$$

The Plancharel identity is

$$\sum_\alpha \left| \widehat{f}_\alpha \right|^2 = \sum_\alpha |f_k|^2 \tag{4}$$

There are other possible conventions for the factors of $N$. For example, we may omit the $\sqrt{N}$ from (2) if we replace $\sqrt{N}$ by $N$ in (3) and put a factor of $N$ left side of (4).

If the $\widehat{f}_\alpha$ are given, then (2) defines $f_k$ for any integer $k$, not just those between 1 and $N$. You should verify that this function satisfies $f_{k+N} = f_k$, which is to say that $f_k$ is a periodic function of $k$ with period $N$. Similarly, the discrete Fourier coefficients $\widehat{f}_\alpha$ are defined for any integer $\alpha$ and are a periodic function of $\alpha$. The fact that $\widehat{f}_\alpha = \widehat{f}_{\alpha+N}$ is called *aliasing*, because $\alpha$ and $\alpha+N$ are different names for the same number. Therefore, we may use different conventions for the starting and stopping value of $\alpha$ and $k$. For $\alpha$, in stead of $\{1, 2, \ldots, N\}$, it is convenient to use (if $N$ is even) $\left\{\frac{-N}{2}+1, \ldots, 0, \ldots, \frac{-N}{2}\right\}$. For $k$, it may be simpler to use $\{0, 1, \ldots, N-1\}$. All that matters is that in (2) we sum over a full set of $\alpha$ and in (3) we sum over a full set of $k$.

The FFT is a clever algorithm computes the DFT quickly. The direct formulas (3) and (2) require $N$ operations for each coefficient, which gives $N^2$ operations in all. For large $N$ this is a lot. The FFT algorithm rearranges (3) in a clever way to compute all the $\widehat{f}_\alpha$ from all the $f_k$ in $2N\log_2(N)$. For large $N$ this is much smaller. The FFT leads to fast (faster) algorithms for computing anything that can be expressed in terms of the DFT.

One example is the autocorrelation function.

## 2 DFT to Fourier transform

The Fourier modes (1) are complex exponentials. As a function of $k$, they oscillate at a rate that depends on the wave number, $\alpha$. The representation (2) gives $f_k$ as a sum of oscillating exponentials with various wave numbers. Both the space variable, $k$, and the wave number variable, $\alpha$ are discrete, taking integer values only. In our application to the heat equation, we want to express a function of the continuous variable $x$ as a sum (or integral) of complex exponentials $e^{ipx}$ for a continuous range of $p$ values. We can arrive at these continuous representations from the discrete ones by a limiting process.

We explain this in the one dimensional case for simplicity. Suppose $f(x)$ is a smooth function that decays to zero rapidly as $x \to \pm\infty$. We want a representation

$$f(x) = \int \widehat{f}(p) e^{ipx} dp \; , \tag{5}$$

together with a formula for $\widehat{f}(p)$. To do this, we derive a representation formula that works for closely spaced but still discrete points, then take the limit that removes the discreteness. We will see that the discreteness in $p$ disappears in at the same time.

For a large integer $N > 0$, define (this makes things simple later)

$$\epsilon = \sqrt{2\pi/N} \; ,$$

the lattice points $x_k = k\epsilon$, and the discretely sampled function $F_k = f(x_k)$. We will take $k$ in the range $\frac{-N}{2} + 1 \leq k \leq \frac{N}{2}$ (assuming $N$ is even). Two things happen as $N$ increases, the sampling points $x_k$ get closer together, and the range of the sampling points expands. In the limit the sampling points become a continuous that stretches from $-\infty$ to $\infty$ in $x$.

In computing the DFT of $F$, we use $k = \sqrt{N/2\pi}\, x_k$ to re-express the quantity in the exponent as :

$$\frac{2\pi\alpha k}{N} \;=\; \left(\sqrt{\frac{2\pi}{N}}\alpha\right) x_k \;= p_\alpha x_k \; ,$$

where $p_\alpha = \epsilon\alpha$ with the same $\epsilon$ as before. Note that the $1/\sqrt{N}$ factor in the DFT formulas (3) and (2) may be written $\epsilon/\sqrt{2\pi}$. Therefore the DFT formulas become

$$\widehat{F}(p_\alpha) \;=\; \frac{1}{\sqrt{2\pi}}\,\epsilon \sum_k e^{-ip_\alpha x_k} f(x_k) \; , \tag{6}$$

and

$$f(x_k) \;=\; \frac{1}{\sqrt{2\pi}}\,\epsilon \sum_\alpha e^{ip_\alpha x_k} \widehat{F}(p_\alpha) \; . \tag{7}$$

Be careful to remember that (7) applies only for those $k$ values that were used in (6). As $N \to \infty$, the range of those valid $x$ values also becomes infinite.

We recognize the sum in (6) as a Riemann sum approximation to an integral. Therefore, as $N \to \infty$, the numbers $\widehat{F}(p)$ will converge to

$$\widehat{f}(p) \;=\; \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-ipx} f(x)\, dx \; . \tag{8}$$

Using this in (7) gives, in the limit,

$$f(x) \;=\; \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{ipx} \widehat{f}(p)\, dp \; . \tag{9}$$

The Plancharel formula becomes

$$\int_{-\infty}^{\infty} \left|\widehat{f}(p)\right|^2 \, dp \;=\; \int_{-\infty}^{\infty} |f(x)|^2 \, dx \; . \tag{10}$$

These are the formulas we will use. The second expresses the function $f$ as a superposition of plane waves with all wave numbers. The first is a formula for the appropriate weight function. It takes some mathematical analysis to give a complete proof that the limits are as they seem, but they are not very interesting, or very difficult for someone with the right training.

As with the DFT, there are variants of (8) and (9) that put the $2\pi$ factors in different places. If $f(x)$ is a probability density, it is common to define

$$\widehat{f}(p) \;=\; \int_{-\infty}^{\infty} e^{-ipx} f(x)\, dx \;=\; E\left[e^{-ipX}\right] \;.$$

This gives

$$f(x) \;=\; \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ipx} \widehat{f}(p)\, dp \;.$$

Another possibility (often used by electrical engineers) is

$$\widehat{f}(p) \;=\; \int_{-\infty}^{\infty} e^{-2\pi ipx} f(x)\, dx \;,$$

and

$$f(x) \;=\; \int_{-\infty}^{\infty} e^{2\pi ipx} \widehat{f}(p)\, dp \;.$$

The Plancharel formula also implies the important *uniqueness* property. If $f(x) = \frac{1}{\sqrt{2\pi}} \int e^{ipx} g(p) dp$, then $g(p) = \widehat{f}(p)$ as given by (8). There is only one Fourier representation of $f$. The point is that the Fourier functions $e^{ipx}$ are complete but not *overcomplete*. In an $n$ dimensional space, $v_1$, …, $v_m$ form a basis if $n = m$ and they are linearly independent. In that case, any vector, $f$, may be written in the form

$$f \;=\; \sum_{k=1}^{m} w_k v_k \;. \tag{11}$$

If $m > n$ and the $v_k$ span the space, we say they are overcomplete. It is possible to represent any $f$, but the coefficients $w_k$ are not unique. The formula (11) does not determine the $w_k$ uniquely. Overcomplete sets are very useful, for example, in signal processing. The Plancharel formula rules this out for the Fourier representation, because if $f(x) = \frac{1}{\sqrt{2\pi}} \int e^{ipx} g(p) dp$, then subtracting the two representations gives $0 = \frac{1}{\sqrt{2\pi}} \int e^{ipx} \left(g(p) - h(p)\right) dp$. But (10) them implies that $\int \left(g(p) - h(p)\right)^2 dp = 0$, which implies that $g(p) - h(p) = 0$ for all $p$ (almost all, if you know what that means).

All these Fourier transform identities may be summarized in the informal statement of Dirac (his book on quantum mechanics):

$$\int_{-\infty}^{\infty} e^{ipx}\, dx \;=\; 2\pi \delta(p) \;. \tag{12}$$

For example, if we seek a formula for $\widehat{f}$ so that (9) is satisfied, we multiply (9) by $e^{-iqx}$ and integrate over $x$. On the left we get $\int e^{-iqx} f(x) dx$. On the right we get a double integral. Changing the order of integration makes this

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \left( \int_{-\infty}^{\infty} e^{i(p-q)x} \, dq \right) \widehat{f}(p) \, dp \;=\; \sqrt{2\pi} \int_{-\infty}^{\infty} \delta(p-q) \widehat{f}(p) \, dp$$
$$= \sqrt{2\pi} \widehat{f}(q) \; .$$

This is the formula (8).

## 3 Properties of the Fourier transform

The Fourier transform is a powerful tool for studying functions partly because of the way it treats derivatives. For example, differentiating (8) with respect to $p$ gives

$$\partial_p \widehat{f}(p) \;=\; \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-ipx} \left( -ixf(x) \right) dx \; . \tag{13}$$

By the uniqueness theorem, this shows that the Fourier transform of $g(x) = xf(x)$ is $\widehat{g}(p) = i\partial_p \widehat{f}(p)$. Similarly, differentiating (9) gives

$$\partial_x f(x) \;=\; \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{ipx} \left( ip\widehat{f}(p) \right) dp \; . \tag{14}$$

This implies that the Fourier transform of $g(x) = f'(x)$ is $\widehat{g}(p) = -ip\widehat{f}(p)$. In other words, differentiation *in the space domain* (going from $f(x)$ to $f'(x)$) is equivalent to multiplication in the *frequency domain* (going from $\widehat{f}(p)$ to $ip\widehat{f}(p)$), and differentiation in the frequency domain is equivalent to multiplication in the space domain.

One use of these relations is to derive the Fourier representation of the solution of the heat equation. If

$$u(x,t) \;=\; \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{ipx} \, \widehat{u}(p,t) \, dt \; ,$$

then

$$\partial_t u(x,t) \;=\; \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{ipx} \, \partial_t \widehat{u}(p,t) \, dt \; ,$$

and applying (14) twice,

$$\frac{1}{2} \partial_x^2 u(x,t) \;=\; \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{ipx} \left( -p^2 \widehat{u}(p,t) \right) \, dt \; .$$

Therefore, given the uniqueness theorem (think this through), $u$ satisfies the heat equation if $\widehat{u}$ satisfies

$$\partial_t \widehat{u}(p,t) \;=\; -p^2 \, \widehat{u}(p,t) \; ,$$

which gives

$$\widehat{u}(p,t) \;=\; e^{-p^2 t/2} \, \widehat{u}(p,0) \; . \tag{15}$$

A partial differential equation in the space domain is equivalent to a family of ordinary differential equations (one for each $p$) in the frequency domain. (Note: we found the $e^{-p^2t/2}$ last class. We only were missing the Fourier transform.) Since the heat equation is so easy to solve in the frequency domain, we must ask what information we can get about the solution this way.

The FFT is a way to use these formulas to give a fast and accurate solution to the initial value problem for the heat equation. Just take the FFT of the initial data, multiply by $e^{-p^2t/2}$, then take the inverse FFT. When this works, it is probably the fastest method (multigrid supporters can please shut up). The disadvantage is that Fourier transform solution formulas and FFT based algorithms are available for a very limited range of problems.

The differentiation formulas imply a relationship between rate at which $\widehat{f}(p) \to 0$ as $|p| \to \infty$ and the smoothness of $f(x)$. Start with the observation that if

$$M_0 \;=\; \int_{-\infty}^{\infty} \left|\widehat{f}(p)\right| \, dp \;<\; \infty \;,$$

then $|f(x)| \leq M_0' = M_0/\sqrt{2\pi}$ for all $x$ (We will write $M$ for $M'$ from now on). That is, if the Fourier transform decays to zero fast enough so that the integral is finite, then $f$ is bounded. Similarly, the formula for $f'$ (14) implies that if

$$M_1 \;=\; \int_{-\infty}^{\infty} \left|p\, \widehat{f}(p)\right| \, dp \;<\; \infty \;,$$

then $\partial_x f(x)$ is bounded too. Continuing in this way, if

$$M_k \;=\; \int_{-\infty}^{\infty} \left|p^k\, \widehat{f}(p)\right| \, dp \;<\; \infty \;,$$

then $\partial_x^k f(x)$ is bounded. Thus, if the Fourier transform decays for large $p$ fast enough, then all the derivatives of $f$ exist and are bounded. A function that has many derivatives is called *smooth* The homework will have some calculations that illustrate this.

This relation between smoothness and decay goes the other way as well. If $f(x)$ decays rapidly as $|x| \to \infty$, then $\widehat{f}$ is a smooth function of $p$.

This can be applied to the heat equation through the solution formula (15). Suppose only that $|\widehat{u}(p,0)| \leq S < \infty$ for all $p$. Then, for $u(x,t)$, we have

$$M_k \;=\; \int_{-\infty}^{\infty} \left|p^k\, \widehat{u}(p,t)\right| \, dp \;\leq\; S \int_{-\infty}^{\infty} \left|p^k\, e^{-p^2t/2}\right| \, dp \;<\; \infty \;.$$

Thus, even when the initial data $u(x,0)$ is discontinuous, the solution at any positive time is differentiable to any order (is *infinitely differentiable*).

Some intuition for all this comes from $\partial_x e^{ipx} = ip \cdot e^{ipx}$. That is, if $v_p(x) = e^{ipx}$, then the derivative of $v_p$ is on the order of $p$ times $v_p$ itself. This means that $v_p$ varies at a rate $p$, or on a length scale $1/p$. A general functions has some rapidly varying and some slowly varying components. The function $f$ is smooth if its rapidly varying components have very small amplitude, i.e. if $\widehat{f}(p)$ is small for large $p$.

The smoothness and decay relations have important implications for approximating functions using the FFT. If $f(x)$ is very smooth, then $\widehat{f}$ is rapidly decaying. This, in turn, implies that we get a reasonable approximation of $f$ if we *truncate* the Fourier integral (9):

$$f_R(x) \;=\; \frac{1}{\sqrt{2\pi}} \int_{-R}^{R} e^{ipx}\, \widehat{f}(p)\, dp \;.$$

If $M_k < \infty$, then the error, $g_R(x) = f(x) - f_R(x)$ satisfies (use $\left|p^{-k}\right| \leq R^{-k}$ if $|p| \geq R$)

$$\begin{aligned} |g_R(x)| \;&=\; \frac{1}{\sqrt{2\pi}} \left| \int_{|p|\geq R} e^{ipx}\, \widehat{f}(p)\, dp \;\right| \\ &\leq\; \frac{1}{\sqrt{2\pi}} \int_{|p|\geq R} \left| p^k\, \widehat{f}(p) \right| R^{-k}\, dp \\ &\leq\; \frac{M_k}{R^k} \;. \end{aligned}$$

For example, if $M_2 < \infty$, then the error arising from truncating the Fourier transform at wave-number $R$ decays at least as fast as $1/R^2$. A more sophisticated calculation along the same lines as this one shows that if $f$ is an analytic function of $x$, then the error from Fourier truncation decays exponentially as $R \to \infty$. The bottom line is that very smooth functions may be approximated easily with a small range of Fourier modes. This explains why the FFT is so effective in solving the heat equation. At time $t$ the Fourier transform decays extremely quickly. There a very small range of wave-numbers suffices to represent $u(x,t)$ well.

The Fourier transform tells us what goes wrong when we try to run the heat equation backwards. Running it forward shrinks high frequency modes, so running it backwards must blow them up by an exponential factor. In particular, the solution formula (15) gives

$$\widehat{u}(p,0) \;=\; e^{p^2 t/2}\, \widehat{u}(p,t) \;.$$

This makes it nearly impossible to specify $u(\cdot,t)$ and find initial data $u(\cdot,0)$ that evolves into it. It is unlikely that the inversion formula (9) will make sense (converge) for functions that grow exponentially for large $p$. We might try to get around this by specifying $u(\cdot,t)$ so that $e^{p^2 t/2}\, \widehat{u}(p,t)$ makes sense, for example, by making $\widehat{u}(p,t) = 0$ for $|p| > R$. Not may functions are represented this way.