RTT TECHNOLOGY TOPIC
September 2005

Matrix Maths in Mobiles
Maths in Mobile Phones Part 1

The role of the Fourier transform in delivering technology and business value

In this month's Hot Topic we study how mathematical transforms deliver technology and business value in our industry in terms of product differentiation (bandwidth, power efficiency, functionality) and intellectual business/patent value.

We use the Fourier Transform as an example.

This is the first of four Hot Topics.

The October Hot Topic ' Maths in Mobile Phones' (2) will cover the Walsh and Hadamard Transform and the combination of Walsh/Hadamard and Fourier Transforms in future radio systems including next generation 1X EV/OFDM and Release 7 HSDPA/OFDM (HSOPA).

The November Hot Topic 'Maths in Mobile Phones'(3) will cover the Discrete Cosine Transform and its application in voice, audio and image processing.

The December Hot Topic 'Maths in Mobile Networks' will cover power control, handover and admission control algorithms using examples from present wide area and local area radio systems.

All the above are examples of transforms used presently in communication systems.

Our starting point is to review the historical body of work upon which transforms are based and areas of present research that promise substantial future performance gain.

We look at how this performance gain can be translated into new revenue and profit opportunity and future competitive advantage.

What are transforms?
Transforms are the process of changing something into something else. Transforms can be unidirectional and irreversible. Setting fire to something is an example of a unidirectional and irreversible transform.

More usefully, transforms can be bi-directional and reversible. AC/DC and DC conversion in electrical engineering is one of the most ubiquitous examples of a day to day use of transforms (hence the use of the word transformer).

Our prime interest though is in the use and application of transforms in communications engineering.

Transforms change the representation of a signal without changing or destroying its information content. They change the way the signal is viewed. The purpose of the transform is to make a signal easier to process. Typically this means making the signal easier to compress and easier to send. 'Easier' implies an ability to process the signal at a lower power and/or to process the signal using less storage or (radio and network) transmission bandwidth.

Transforms in historical context

Transforms can be described and implemented mathematically and exploit techniques which we take for granted but are the result of several thousand years of mathematical progress.

Maths (applied maths) helps us to do things we want and need to do.

The real life signals that need to be transformed in communications all start out and end as analogue waveforms (sound waves, light waves, and radio waves) so an understanding of waveform geometry as a specific part of mathematical theory is particularly important to us.

Information in radio systems is modulated on to radio waves. The radio waves are sinusoidal but carry information that is modulated by changing the phase, frequency and/or amplitude of the (composite) carrier. We express phase, amplitude and frequency in terms of radians (the speed and angle of rotation of the phase vector). These waveforms can be described mathematically and manipulated mathematically. Mathematical manipulation and the efficiency of the mathematical manipulation process are the basis for modern digital radio communication systems.

The history of maths and the contribution that individual mathematicians have made to waveform and transform theory goes something like this:

15000 BC - the start of organised agriculture involves the counting of goats and other mobile objects (arithmetic - the science of numbers). Farmers need to work out the size of their fields (geometry - the science of shapes, from the Greek geo meaning earth and metro to measure) and build structures (trigonometry - the science of angles).

2000 BC - the Egyptians produce a numeral system to support trade and agriculture

1900 -1600 BC - the Babylonians introduce the base 60 system of applied mathematics still used today (360 degrees in a circle, 60 seconds in an hour etc).

640-546 BC - Thales of Miletus predicts the lunar eclipse of 585BC and proves a number of properties of triangles (vertical angles being equal, an angle inscribed in a semicircle is a right angle etc).

572 BC - Pythagoras of Samos produces all those useful equations for triangles and circles and shows how complex shapes can be described in terms of simple straight

lines and circles (an early example of a transform).

427-347 BC - Plato founds his academy in Athens (387 BC) and encourages mathematicians such as Theaetetus and Eudoxus (the exhaustion man) to work on geometry and proportional theory.

300 BC - Euclid produces 'the bible of mathematics', 13 books and 456 propositions on plane and solid geometry and number theory including prime numbers (the infinitude of the primes). Euclid's axioms include the concept of euclidean distance (the straight-line distance between two signal vectors) which still serves today as a starting point for many mathematicians. Note that 'distance' can be related to many performance metrics in present communication systems including system sensitivity and selectivity.

287-212 BC - Archimedes equates the area of a circle with a triangle and works out the value of Pi using inscribed and circumscribed polygons. (as well as that business with the bath).

A bit of a gap due to the Romans being more interested in building roads than maths.

430-501 AD - Tsu Ch'ung-chih - the famous Chinese mathematician defines Pi more accurately.

529 AD - the closing of the Alexandrian library marks a shift towards Arabia in terms of mathematical innovation including the evolution of algebra (the art of calculating with unknown quantities represented by letters after the arabic al gebr/al jabr meaning to equalise).

900 AD - Al Khumar Rizmi works on algorithmic approaches to problem solving (algorithm - the step by step iterative process for solving problems using a pre determined set of rules within a finite number of steps). The word algorithm is named after him.

1452-1519 - Leonardo Da Vinci and the arrival of renaissance man including;

1501-1576 - Gerolamo Cardamo of Milan works on the mathematics of probability and the application of negative numbers (used in IQ quadrature modulation).

1550-1617 - John Napier and Henry Briggs (1561 to 1631) work on logarithms.

1596-1650 - Rene Descartes - works on merging algebra and Euclidean geometry, gives us Cartesian co-ordinates still used as the basis for modern constellation diagrams.

1601-1665 - Fermat and his theorem. Pierre De Fermat of Toulouse works on analytic geometry and number theory and differential calculus (determining the rate of change of a quantity). Produces irritating problems for other mathematicians (such as why a cube cannot be divided into two cubes).

1642-1726 - Isaac Newton develops a geometric approach for calculating Pi (as well as that business with the apple).

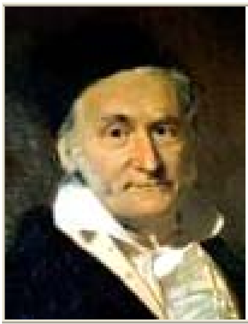
1707-1783 -.Leonard Euler proves that complex numbers (the extension of a real number by the inclusion of an imaginary number) can be related directly to the real trigonometric functions of sines and cosines - the basis of the fourier base and fourier transform used in all present OFDM radio systems.

1768-1830 - Jean Baptiste Joseph Fourier narrowly avoids being guillotined in the French revolution and produces a paper in 1807 on the use of sinusoids to represent temperature distribution. Uses the technique in a study on heat propagation which finally gets published in 1822 (Theorie Analytique De La Chaleur.). In the meantime, claims that any continuous periodic signal can be represented as the sum of properly chosen sinusoidal waves. Challenged by Lagrange (1736-1813) and Laplace (1749-1827). Lagrange asserts that sinusoids cannot be used to describe signals with corners eg square waves. Lagrange is (sort of) right but in practice over time it is shown that sinusoids can be used to the point where the difference between the two has zero energy (the Gibbs effect). This becomes the basis for translating signals from the time domain to the frequency domain and back again, taking simple elements in the frequency domain and combining them to create complex signals in the time domain (the inverse discrete fourier transform used in OFDM signal synthesis) and decomposing complex time domain signals into simple frequency components (the discrete fourier transform used in OFDM signal analysis). Unfortunately the process involves a lot of calculations and proves to be quite laborious.

Jean Baptiste Joseph Fourier 1768-1830


1777-1855 - Carl Friedrich Gauss develops the theory of complex numbers following on from Euler and works out simpler and faster ways of doing Fourier Transforms. Forgets to tell anyone how it's done. His work on the Fast Fourier Transform is published after his death in neo Latin and remains unread by anyone who could possibly understand it. Also produces a body of work on non euclidean geometry (triangles with more than 180 degrees in their angles) and work on statistics (Gaussian distribution) which was to prove fundamental to our later understanding of white noise and filter design.

Carl Friedrich Gauss 1777-1855



1845-1918 - Georg Cantor works on exotic number theory and transcendental and irrational numbers (A number is transcendental if it is not the solution of any polynomial equation with integer co-efficients, a number is irrational if there is no limit to the number of its decimal places. Pi and Euler's constant are examples of irrational numbers though bizarrely when added together they become rational).

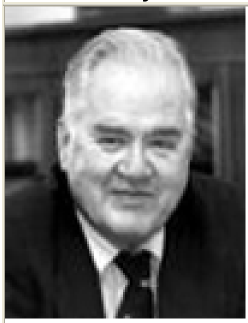
1887-1920 Srinivasa Ramanujan produces ever more accurate approximations of Pi

1862-1943 - David Hilbert works on algebraic number theory and geometry including the functional analysis of space and distance, the concept of vectors as linear sequences of complex numbers and functions and related work on harmonic and functional analysis. Hilbert is rather over shadowed by Einstein (1879-1955) but his contributions to modern communications, signal processing and vector theory remain very significant. Has a transform named after him (Hilbert transform used in wideband 90 degree phase shifts).

The above represents a selective but fundamental body of work that modern living mathematicians can draw on to produce transforms optimised to perform specific tasks.

In 1965, 110 years after the death of Carl Friedrich Gauss, JW Cooley at the IBM laboratories working with John W Tukey of Princetown University published a paper which showed how the Fourier series could be reduced recursively by decimating in time or frequency, an approach similar to Gauss whose work at the time was unknown to them. This became the basis for the Fast Fourier Transform and is often now cited as the genesis of modern digital signal processing.

John Tukey 1915-2000



The Cooley-Tukey algorithm inspired a whole generation of mathematicians to against the amount of memory space needed against the number of clock cycles needed against the processing delay and power limitations of the host device. Examples include the Guo and Burrus wavelet based FFT in 1996, the Edelman fast multipole method (based on the compressibility of the Fourier matrix rather than the compressibility of the data), and other options like the Good Thomas algorithm, the Rader Brenner algorithm, the Bruun algorithm and the Winograd algorithm.

Some of the techniques (for example Guo and Burrus) draw on work done by Benoit Mandelbrot. Mandelbrot's work on fractal geometry has been translated into Basque, Brazilian, Bulgarian, Chinese, Czech, Italian, Portuguese, Rumanian and Spanish so is unlikely to share the fate of Heinrich Gauss's work on the FFT.

Fractals have particularly compelling advantages in image processing, particularly where natural objects (trees and leaves) need to be encoded. We revisit this topic in November when we look at fractal compression schemes as an extension of existing DCT based image compression (Mandelbrot maths in mobile phones).

Benoit Mandelbrot's most recent published work in 2004 with RL Hudson applies the concepts of fractal geometry and fractal transforms to the evaluation of risk in financial markets - the science of fiscal transforms. A profit and loss statement, balance sheets and share price movements are mathematical expressions and are therefore amenable to fiscal mathematical trend and pattern analysis. Fractal transforms are particularly good at exposing patterns and coupling effects in fiscal systems.

Back in the engineering rather than fiscal domain, fractal transforms are increasingly relevant to the modelling of offered traffic behaviour in multi media networks. (See RTT's March 2003 Hot Topic on Turbulent Networks for a flavour of this aspect of transform value). We revisit this topic again in December when we study 'Maths in Mobile Networks'.

Like Cooley and Tukey, Mandelbrot's work continues to inspire mathematicians to produce faster and more effective transforms both for signal processing and image compression applications.

Note in passing how pure maths (maths for its own sake) and applied maths requires certain infrastructural conditions in order to flourish and deliver political, social and economic value. Plato's academy and the Alexandrian Library (387 BC to AD529), the growth of Islam after 650 AD, the Gutenberg Press in 1450, the postal service that allowed Newton to correspond with his peers and more recently the internet are examples of changes that have facilitated mathematical progress.

So how do Fourier transforms deliver performance differentiation ?
Well for a start present radio systems including WiFi, WiMax, UWB , DAB/DVB/DRM and (fairly soon) HSOPA all depend on the Fourier transform in order to support relatively high data rates.

As bit rates have increased over the radio physical layer it has become progressively more attractive to use orthogonal frequency division multiplexing (OFDM) to improve resilience to intersymbol interference. An early example in terms of a European specification is the DAB standard (1991) but OFDM is now a fundamental part of local area WiFi, wide area WiMax and personal area UWB radio systems. The characteristics of these radio systems have been addressed in our August 2004 and August 2005 Hot Topics but let us briefly revisit some of the maths behind OFDM and the differentiation achievable through optimised implementation of the transforms and their associated iterative algorithms.

Characteristics of the Fourier Transform
Fourier's fundamental assertion that has made him famous through time is simply that any complex signal viewed in the frequency domain can be understood as a composite of a number of simple sinusoidal signals that have been added together.

To place this into our specific context of interest (OFDM), his theorem led to the understanding that a composite waveform can be described as a mix of sine and cosine waves (90 degrees apart) that also have specific phase and amplitude characteristics. More specifically signal points in the frequency domain can be combined to create complex signals in the time domain (the Inverse Fourier Transform) and complex time domain signals can be resolved back into specific signal points in the frequency domain (the Fourier Transform).

If the process is applied to a discrete waveform sample (as in an OFDM signal burst of a defined and repetitive length), then it is described as a discrete fourier transform (DFT) or inverse discrete fourier transform (IDFT).

The overall purpose of the Fourier transform in the context of implementing an OFDM transmitter and receiver is to take a complex time varying waveform and express it as a tractable mathematical formula. Tractable implies an ability to either compact the formula itself and/or the data used in the formula - the basis for the algorithms used in Fast Fourier Transforms.

An IDFT synthesis equation on the transmit path multiplies the discrete frequency domain by a sinusoid and sums over the appropriate time domain section (a 3.2 microsecond burst in WiFi).

A DFT analysis equation on the receive path multiplies the time domain signal by a sinusoid and sums over the appropriate time domain section (3.2 microsecond burst).

If it was a continuous signal (which it isn't) the transform would integrate rather than sum.

The IDFT and DFT as used in a WiFi transceiver
Take a WiFi transceiver as an example. In essence, a WiFi transceiver is based on a mix of 17th century (cartesian) and 19th century (fourier) mathematical principles. The modulation can either be BPSK, QPSK, 16 QAM or 64 QAM.

Using a 16 QAM signal as an example, 4 data bits are mapped on to a 16 QAM constellation (based on Cartesian co-ordinates).

The position of the symbol that relates specifically to the 4 bits is described as a complex number with two values, one of which describes the amplitude and the second of which describes the phase. The phase co-ordinate has a numerator j after it which can be +j or -j. The j numerator denotes whether the symbol is mapped to a positive or negative phase shift.

These two values are then put into a bin and the process is repeated 52 times and then another 12 times using zeros as a value to make up 64 complex number bins.

The job of the IDFT (inverse discrete fourier transform) is to do the maths to calculate the sine and cosine waveform values needed to create the composite time domain sine and cosine waveforms that will represent the 64 frequency sub carriers each of which has a specific phase and amplitude. The back end of the waveform is copied to the front in order to provide a guard period and the waveform is clocked out at 20 MHz to form a 3.2 microsecond modulated burst with a .8 microsecond guard period.

In the receiver, a DFT is applied to the composite time domain waveform and (all being well), the waveform across the burst is resolved into 64 frequency sub carriers each with the phase and amplitude information needed to decide which 4 bits were originally modulated onto the signal.

The process exploits a number of convenient properties of sinusoidal waveforms.

Properties of a sinusoid

A cosine /sine function has a period of 360 degrees.
The cosine curve is the same shape as the sine curve but displaced by 90 degrees.
Providing the system is linear, a sinusoidal input results in a sinusoidal output.
The average of the values described by any true sinusoidal waveform is zero if averaged over an even number of cycles.
Multiplying cos (x) by sin(x) produces a third waveform which is smaller in amplitude and at a higher frequency but which will sum to zero when averaged out.
The cosine (real) computation only measures the cosine component in the time series at a particular frequency. The sine (imaginary) computation only measures the sine component in the time series.
The amplitude and phase of the signal can change but the frequency and wave shape remain the same.
The periodic nature of sinusoids provides the basis for implementing orthogonal signals (signals that do not interfere with each other) in either the time or frequency domain.
Being sinusoidal, the sub carriers can be arranged such that the peak signal energy of one sub carrier coincides with the minimum energy in the adjacent sub carrier (a property well known to designers of FFSK modems in the 1980's). This property provides the basis for OFDM signal orthogonality.
If you want to find out whether a time series contains a sine or cosine component, create a cosine function at frequency F and a sine function at frequency F, multiply by the cosine function (which we will call real F (I) and multiply by the sine function which we will call imaginary F (Q). Both should be non-zero values but any modulation applied will show up as sine and cosine off sets or in other words, the IQ modulated waveform.

The IDFT and DFT in wide area OFDM (WiMax and HSOPA)

As discussed in our two previous Hot Topics on OFDM, wide area OFDM implies a specific need to have a relatively slow symbol rate with an extended guard band (to deal with the delay spread on the channel) combined with the ability to support high data rates.

Keeping the symbol rate down requires the implementation of a more complex FFT and/or higher order modulation. For example WiMax has the option of a 256 point FFT and/or 64 QAM. Note that these radio systems in common with WiFi are two way so require both an Inverse Discrete Fourier Transform to achieve the waveform synthesis on the transmit path and the Discrete Fourier Transform to achieve signal analysis on the receive path.

As such, two-way OFDM radio systems are much harder to implement than receive only systems (DAB/DVB broadcasting and DRM). As bit rates increase over time it will be necessary to implement higher order FFT's. This will place a premium on the techniques needed to reduce the complexity/calculation overheads implicit in these schemes. We will need faster more efficient transforms.

Note that these schemes all depend on either the compressibility of the transform itself or the compressibility of the data which in turn depends on the symmetry and essentially repetitive nature of sinusoids and their associated input number series.

Note also how compression techniques become more important as FFT complexity increases. A 32 point FFT is typically ten times faster than a standard FT. A 4096 point FFT (used for example in a DVB H receiver) is typically one thousand times faster.

The impact of transforms on DSP and Microcontroller architectures - maths in the microcontroller.

We said earlier that the efficient implementation of fast Fourier transforms (and transforms in general) imply a trade off between multiplication, division, addition, subtraction, memory space, bus width, bit width and clock cycle bandwidth. These trade offs are in turn dependent on the architecture of the host device.

Devices using parallel processing (super scalar devices capable of operating multiple parallel instructions) have become increasingly useful for all sorts of tasks including the FFT, DCT and related filtering, convolution, correlation, decimation and interpolation routines.

Hardware optimisation includes the integration of hardware repeat loops for filtering and FFT routines and the task specific optimisation of fixed point, floating point and saturation arithmetic - the maths of the microcontroller.

Although FFT algorithms are a commodity item in most DSP libraries there are still plenty of future performance optimisation opportunities particularly as increasingly complex higher order FFT's are introduced into next generation radio systems.

So how do Fourier Transforms add value?

In the context of OFDM, Fourier transforms add value by improving the performance of higher bit rate radio systems both in local area and wide area radio networks. The 'cost' is additional clock cycle/processor bandwidth but this in turn helps sell silicon.

The ownership of the intellectual value that has been built around the FFT, in particular the compression/discard algorithms that speed up the transform will also become increasingly significant as higher order FFT's are deployed over the next three to five years.

The efficiency of these algorithms and their mapping to optimised DSP and microcontroller architectures will translate directly into mobile phone and radio system performance including power budgets and supportable bit rates.

About RTT Technology Topics

RTT Technology Topics reflect areas of research that we are presently working on.

We aim to introduce new terminology and new ideas to clarify present and future technology and business issues.

Do pass these Technology Topics on to your colleagues, encourage them to join our Push List and respond with comments.

Contact RTT

RTT, the Shosteck Group and The Mobile World are presently working on a number of research and forecasting projects in the cellular, two way radio, satellite and broadcasting industry.

If you would like more information on this work then please contact

geoff@rttonline.com

00 44 208 744 3163