CONVOLUTION AND CORRELATION

Due Date: 3/26/2014 @ 11:59 pm

You may have noticed from last week's exercise how long it took to perform the discrete Fourier transform of a large array. This greatly limits the applicability of the discrete Fourier transforms, as the time cost scales as $\mathcal{O}(N^2)$. Luckily, the Fast Fourier Transform (FFT) with its $\mathcal{O}(N\log_2 N)$ cost can produce equal results much faster. This enables us to use a range of operations that would otherwise be prohibitively slow. The two most notable examples are convolution,

$$g * h = \int_{-\infty}^{\infty} g(\tau)h(t-\tau)d\tau,$$

and correlation,

$$C(g,h) = \int_{-\infty}^{\infty} g(\tau+t)h(\tau)d\tau.$$

The discrete versions of these two equations are

$$(g*h)_n = \sum_{k=-N/2+1}^{N/2} h_{n-k}g_k, \quad C(g,h)_n = \sum_{k=0}^{N-1} g_{n+k}h_k.$$

By using convolution and correlation theorems,

$$\mathcal{F}(g*h) = \mathcal{F}(g)\mathcal{F}(h), \quad \mathcal{F}(C(g,h)) = \mathcal{F}(g)\mathcal{F}(h)^*,$$

we can perform both operation significantly faster by first computing Fourier transforms of both functions, then multiplying them, and finally computing the inverse transformation of the product.

- Compare the computational cost of the Discrete Fourier Transform from the previous assignment and the FFT. Show that their execution time scales as expected.
- A study of car density as function of time on the turnpike exit to the Dodgers stadium in LA was performed in a period of roughly 25 weeks, with a 5-min sampling. In addition, the start time, end time and the

number of visitors was recorded for the events at the stadium. The data are in files `dodgers.cars.data` and `dodgers.events.data`. Analyze the data and interpret the results.

- Compute the autocorrelation function of the sound of boiling water. If you are so inclined, acquire your own data, otherwise you can use `boiling.data` from the course homepage. The autocorrelation function may drop rapidly and you should use the log scale to visualize it.

- Sunspots are closely correlated with the Sun's magnetic activity. Their number has been recorded since 1700 on a yearly basis and since 1749 on a monthly basis. Using FFT, find any periodicity in the data (`sunspots.yearly.data` and `sunspots.monthly.data`). Compute autocorrelation functions for both data-sets and compare them. Is higher sampling of sunspots warranted?

- Using convolution, predict the shape of a spectral line out of a diffraction-limited spectrograph. Assume that the response function of the spectrograph is a Gaussian.