Ge 101 Problem Set 4
Due: Wednesday, November 28, 2012
Remember your target is 8 hours of work on this; don't overdo it.

1. **The ice-volume and temperature signals in of marine isotopic records**
(a) Given the following information, calculate the mean $\delta^{18}O$ of the ocean at the Last Glacial Maximum:

Last Glacial Maximum (LGM) sea level was 120 meters lower than today.
The mean depth of the modern ocean is 3800 m.
The modern $\delta^{18}O$ of the ocean is 0‰ (SMOW scale).
The mean $\delta^{18}O$ of the ice that has melted since LGM is –30‰.

(b) The measured difference in $\delta^{18}O$ of tropical planktonic forams between LGM and now is 1.8‰. This consists of a contribution from the change in $\delta^{18}O$ of the ocean water added to a change in the temperature-dependent fractionation between water and precipitated carbonate. What temperature change does this imply? Use your answer from part (a) and the Epstein calibration, $T = 16.5 - 4.3(\delta_{carbonate} - \delta_{water}) + 0.14(\delta_{carbonate} - \delta_{water})^2$ (either make an assumption about absolute T, investigate the sensitivity of your answer across a reasonable range in T, or just neglect the quadratic term). What other factors, besides ice volume and surface ocean temperature, could influence this measured change?

2. **Fun with orbital parameters**
It is evident that variations in the Earth's obliquity and the precession of the seasons have no net effect on total annual solar radiation arriving at the Earth, only its seasonal and latitudinal distribution (as an aside, Earth's albedo is seasonally and latitudinally non-uniform so this statement is not true of total annual solar radiation *absorbed* by the Earth). Is the same thing true of the 100 ka variations in the eccentricity of the Earth's orbit? If you consider the variations in Earth-Sun distance, Kepler's second law (i.e., the Earth-Sun line sweeps out equal area in equal time), and the inverse square law for received radiation, then the answer is not at all obvious. Here are the tools you need to determine the answer:

- The radiation received by the earth per unit time can be written W_o/r^2, where r is the Sun-Earth distance at any time and W_o is a constant that depends on things like solar flux, size of the Earth, and choice of units.
- The total radiation received in one year, Ψ, is obtained by integrating $(W_o/r^2)dt$ over 1 year, so if the orbit is circular (i.e., r is constant), the annual total radiation is TW_o/r^2, where T is the length of a year.
- Now, instead of assuming r to be constant, change variables in the integral to the angle θ between the Sun-Earth line and some fixed point in the ecliptic, so now you are integrating over $d\theta$ from 0 to 2π. You will need $dt/d\theta$ to do this change of variables; Kepler's second law tells you $r^2\, d\theta/dt = J$, where J is the angular momentum of the orbit.
- You should also know that $J^2 = GMa(1-\varepsilon^2)$, and $T^2 = (4\pi^2/GM)a^3$, where G is the gravitational constant, M is the total Earth+Sun mass (constant), a is the semi-major axis of the orbit, and ε is the eccentricity. Let us assume that a and T are constants, and that ε is constant over the course of a single year.
- OK, for the non-circular orbit, solve for Ψ in terms of W_o, T, a, and perhaps ε.

Does the answer depend on ε? If so, how strongly, and do you think this will be a significant effect on climate? Your notes from Lecture 14 show the amplitude of the actual eccentricity variations (about 0 to 0.05). Speculate, if you are so inclined, on the radiation received by comets, which have orbits with eccentricities very close to 1.

3. **You'll have to learn discrete Fourier analysis someday, why not now?**
Let's try our own frequency analysis of an ice core record, and see if we get something like Milankovitch forcing to come out of it. You can do this entire exercise in Microsoft Excel, if you have StatPlus and Solver installed. Other software packages such as Matlab, Mathematica, or Maple can do it too. Or, if you're a serious code warrior, build it yourself.

(a) Download the Vostok Ice Core δD record, which you can find either at:
http://www.asimow.com/vostok.1999.temp.dat or
http://cdiac.esd.ornl.gov/ftp/trends/temp/vostok/vostok.1999.temp.dat .

The file gives temperature as well as δD, but it is not an exact function (the temperature must be corrected using some other measurement), so we'll just analyze the δD time series. Take a minute to ponder the amount of work involved in counting 422,700 annual bands and making 3310 isotope measurements!

(b) The data are at 1 meter intervals downcore, but we want data regularly spaced in time for our analysis, so the first thing to do is resample the data to a regular time grid. Use 2048 points, starting at time 0, going to 414000 years, and using an interval of 18*23000/2048 ~ 202.15 years (you can experiment with other gridding schemes, but the total number of points has to be a power of 2 for many FFT routines). Use simple linear interpolation between adjacent points. If you can't figure out a simple way to do this (i.e., without typing 2048 different formulas or numbers), just ask me (hint: if you're working in Excel look at the functions MATCH and INDEX).

(c) Now obtain the discrete Fast Fourier transform (FFT) of the string of 2048 numbers. In old versions of Excel, this was on the Tools menu, under 'Data Analysis..." if you had the Analysis ToolPak installed. In latest versions of Excel you need to get the free StatPlus software, where it's under "Statistics" and then "Time Series/Forecasting". This is a forward (or direct) transform, but note the option for Inverse transform, which we will use in a minute. In the StatPlus dialog box be careful to uncheck "Labels in first row" or it will only take 2047 data points…

The result is a series of 2048 complex numbers that you should number from 0 to 2047. The value at n=0 is the zero frequency term, which is the sum of all the δD values, i.e. 2048 times the mean. The first half of the remaining values ($a_n+b_n i$) represent the amplitude and phase of the set of sinusoidal basis functions with n complete periods in the data range (the second half of the numbers are the negative frequencies, and since the data are real, the negative frequency values are just the complex conjugates of the positive values; they carry no additional information). Hence the complex magnitude (IMABS() in Excel) of the value at n=1 in the FFT is the amplitude of the 414000 year component, and the complex argument (IMARGUMENT() in Excel) is the phase. The n=2 number is the amplitude and phase of the 212000 year component, the n=3 number is the amplitude and phase of the 414000/3 year component and so on down to the n=1024 number which is the amplitude and phase of the 404.3 year oscillation, the highest (or Nyquist) frequency that can be resolved with 202.15-year sampling.

> 1. Plot the complex magnitude of the FFT values from n=1 to n=1024 (on the y-axis) against the period of the oscillation, i.e. against 414000/n years (on the x-axis). Use a log scale for period, but a linear scale for amplitude.

2. What are the periods represented by the three biggest peaks in the frequency spectrum? Can you give an interpretation of each of these peaks in terms of orbital forcing?

(d) Now let us generate synthetic time domain δD records from selected components of the FFT spectrum. You do this by setting to zero all the frequencies you don't want (both in the first half of the FFT and in the second half; i.e. for $0 < n < 1024$, you must zero both the n term and the $2048-n$ term together; if you don't do this the inverse transform will be complex). Then take the inverse Fourier transform of the modified array of 2048 complex frequency domain values, to recover a new array of 2048 real time domain values. You may need to then take the IMREAL() or IMABS() of the result in order to plot it (the numbers are real if you did it right, but Excel does not realize this due to round-off error; note StatPlus returns the real and imaginary parts in separate columns so you can just ignore the nearly-zero column of imaginary parts). Try looking at just the n=0, n=1 (and 2047), and n=4 (and 2044) terms alone (i.e., zero all others), just to get a feel for what they represent.

1. Plot the synthetic time series from the inverse transform of just the n=0 term and the terms at the top of the three biggest peaks in the amplitude spectrum. Compare it to the original resampled time series data. Evaluate the quality of the fit by calculating how much of the variance of the original data is accounted for by this synthetic, i.e. compare the sum of the squares of the residuals between the synthetic and the data to the sum of the squares of the difference of each value in the original time series from the mean δD.

2. Now try adding back in the next few frequencies in decreasing order of amplitude in the original spectrum and calculating new synthetic time series. How many frequencies do you have to include in order to reduce the variance of the residual to 25% of the original variance (that is, ≥75% of the original variance is captured by the synthetic, ≤25% is unexplained)? Plot the simplest synthetic time series that reduces the residual variance to ≤25% of the original data variance on the same graph as the original resampled time series data. Does this synthetic seem to capture most of the essential features of the data?

(e) Without giving away question (c)2 above, let us just imagine that we wanted to find the best fitting model that consists of six sine-wave components with periods of 19, 23, 41, 95, 125, and 406 ka plus a DC (constant) term. There are 13 free parameters in this model...the amplitude and phase at each of the six periods, plus the mean value (which we know, so we don't really have to fit that one). This time you are going to use another of Excel's useful functions, the Solver, which does non-linear minimization. Likewise, Matlab has a number of routines built-in for non-linear multidimensional minimization. Set up a column of 12 places for the *adjustable parameters* (you can put initial guess values in if you want), and build a function of time that is the sum of the DC value plus the six sinusoidal components. Put a column that evaluates this function at each of the times in the resampled δD record. Create a *target cell* that contains the sum of the squared residuals between the data and the model function. Now, under Tools, bring up Solver, and tell it to minimize the target cell with the total squared residual by varying the 12 adjustable parameters (this may take a while, and a lot of memory).

1. Try it a couple times with different initial guess parameters...does it converge to a unique solution, or does it tend to get caught in local minima? How good is the best solution, in terms of reduction in variance of the data? Is it better or worse than the quality of the synthetic time series generated from the six biggest peaks plus DC in the FFT?

2. Those six frequencies are supposed to be the dominant ones in the forcing function, but they seem (at least for me) to give a pretty mediocre description of the actual temperature record. I think this has something to do with the width of the peaks in the FFT spectrum and the non-linearities of the climate system response. What do you think?