## DOING A SPECTRAL CLEAN ON IRREGULARLY-SAMPLED DATA

We provide a step-by-step procedure for doing spectral clean including cautionary and helpful remarks. The fundamental reference is Roberts, Lehár, & Dreher 1987, AJ, 93, 968, particularly Appendix B and it's certainly nice to look at their figures and examples.

1. You have $N$ irregularly-spaced samples in time. $N$ need not need be even, as we required in the regularly-sampled case.

2. Rezero the *time* values by subtracting the mean. This forces the window function to have zero slope at the origin $\nu = 0$. This, in turn, allows you to use a nice symmetric real Gaussian as the CLEAN beam.

3. Subtract the mean from the *signal* to get rid of the DC component. This is not necessary, but why bother to clean away the DC component, whose value is accidental anyway?

4. You need to select a value of $f_{max}$ to use in the cleaning process, using the distribution of sample intervals as a guide. By "sample intervals" we mean $\Delta t_n = t_{n+1} - t_n$. For the regularly-sampled case, $f_{max} = f_{Nyquist} = \frac{1}{2\Delta t}$. For the irregularly-sampled case, the Nyquist criterion doesn't apply, so perhaps a reasonable choice might be $f_{max} = \frac{1}{\langle \Delta t_n \rangle}$; or $f_{max} \frac{1}{\min(\Delta t_n)}$. Alternatively, such a choice might be physically unrealistic, so you might use considerations such as a theoretical limit.

5. Pick $\Delta\nu$, the frequency interval between points on your computed DIRTY spectrum. A reasonable choice is $\Delta\nu = \frac{1}{T_{max}}$, where $T_{max}$ is the total time interval covered by the samples.

6. Generate the DIRTY spectrum $D(f_j)$ (same as $D_j$) and DIRTY beam (also called the window function) $W(j)$ using a snail-type DFT; you can use my `dft.pro`, which uses the same conventions and normalizations as IDL's `FFT`. There would be four parameters. Three inputs: the set of measruement times $t_n$, the (possibly complex, but here we assum real) input signal $S(t_n)$, and the set of output frequencies $f_j$. One output: the "dirty" complex spectrum evaluated at $f_j$, $D(f_j)$ (same as $D_j$).

   You need to derive $D(f_j)$ for frequencies $-f_{max} \rightarrow +f_{max}$. The total number is $2J + 1$, not $2J$; here $J = \frac{f_{max}}{\Delta f}$. Note that $f = 0$ occurs at $j = J$.

   You need to derive $W(f_j)$ for frequencies $-2f_{max} \rightarrow +2f_{max}$. The total number is $4J + 1$, not $4J$; here $J = \frac{f_{max}}{\Delta f}$. You calculate $W(f_j)$ using the DFT, replacing the input signal $S(t_n)$ by 1.0. Note that $f = 0$ occurs at $j = 2J$.

7. Search for the highest point in $D(j)$, using the "modulus" of the complex number; in IDL you do this with `abs`. Restrict your attention to $f > 0$, i.e. $j > J$. The reason: $S(t_n)$ is real so $D(f_j)$ is Hermitian, so the negative frequency components are just the complex conjugates of the positive frequency ones.

   Let this highest point in $D(f_j)$ occur at $j = j_{max}$.

8. We want the complex contribution of this spectral point $D(j_{max})$ to all frequencies in the spectrum. This consists of *two* contribution:

    (a) The contribution from the positive frequency component to all spectral points; call this $CC^+$, where $CC$ stands for "Clean Component" and the + stands for positive frequency contribution. You get this by shifting $W(f)$ so that it's centered on the signal and multiplying it by the complex value $D(j_{max})$. In IDL this would look like

    $$CC^+ = D(j_{max}) * shift(W, 1 + 2J + j_{max}) \tag{1}$$

    This centers the window function on the spectral point at $j_{max}$.

    (b) The contribution from the negative frequency component, $CC^-$. This point is distanced $2f(j_{max})$ away from the positive frequency point, so the window function is centered $2j_{max}$ away, i.e.

    $$CC^- = conj(D(j_{max})) * shift(W, 1 + 2J - j_{max}) \tag{2}$$

    Note that we have to use the complex conjugate of $D(j_{max})$!

9. Correct the dirty map using an appropriately-considered gain factor

    $$D' = D - gain * (CC^+ + CC^-) \tag{3}$$

    You might wish to force the zero frequency component of $D'$ (that's $D'(J)$) to be zero.

    **Important Comment:** When your software development reaches this point, *test it* by inputting a single, regularly sampled cosine wave whose frequency is an integral multiple of $\Delta\nu$ and make sure that its contribution is *completely removed* in one pass when $gain = 1$, i.e. that $D'(j) = 0$. Then do the same for a sine wave. Repeat for a non-integral multiple of $\Delta\nu$; in this case it will require multiple passes[1] and using $gain = 1$ won't work well. If your software passes these tests, then the rest is gravy.

    Another **Important Comment:** What to use for *gain*? If your data are noise-free, you can use a high *gain*. As the noise increases, *gain* needs to decrease. Try some experiments and use your best judgement.

10. Record the clean component, and its position, from the previous step. Cycle back to number 7 and iterate.

11. When you are finished iterating (this is a judgment call) generate the CLEAN spectrum by convolving the clean components with a suitable Gaussian.

[1]Why? Because your window function can be centered only at integral multiples.