

---

# SumPS Analysis: CTF Estimation

---

Notes by James Conway, 26-Mar-2001.

Contact me:

Laboratoire de Microscopie Electronique Structurale  
Institut de Biologie Structurale  
41 rue Jules Horowitz  
38027 Grenoble Cedex 1  
France  
Telephone: +33-(0)476-884568  
Fax: +33-(0)476-885494  
Email: James.Conway@ibs.fr

Contents:

1. Overview .....	1
2. Sample Data .....	1
3. Through the analysis step-by-step .....	3
4. CtfZeros .....	8
5. References.....	9

---

## 1. Overview

The averaged power spectra data produced by the program SumPS (Sum Power Spectra) may be used to determine characteristics of the Contrast Transfer Function (CTF) of an electron micrograph as outlined in [1]. These include the Gaussian decay of signal with spatial frequency and the defocus setting used for imaging the micrograph. (Note that the nominal setting of the defocus on the microscope is insufficiently accurate). A step-by-step instruction manual to the SumPS analysis is given in this document. Use is made of Kaleidagraph (Synergy Software, [www.synergy.com](http://www.synergy.com)), a commercially available graphing program that runs on Macintosh and Windows systems. I have used version 3.08d of Kaleidagraph on Macintosh computers – a recent update to 3.5.1 has become available but I have not yet seen it.

The average spectrum calculated by SumPS is saved as an image in both a 4-byte-real format and as log-values of the averaged spectrum in a 1-byte unsigned integer file. A 1D profile of the whole spectrum, and optionally of angular segments of the spectral plane, are also saved in a text file. This text file is the main source of data for the analysis described here.

The procedure is to plot the profile of the spectrum, select the minima and tail of the plot, fit a Gaussian decay to these selected regions, divide the Gaussian fit from the average and measure the locations of the minima from this deconvolution plot.

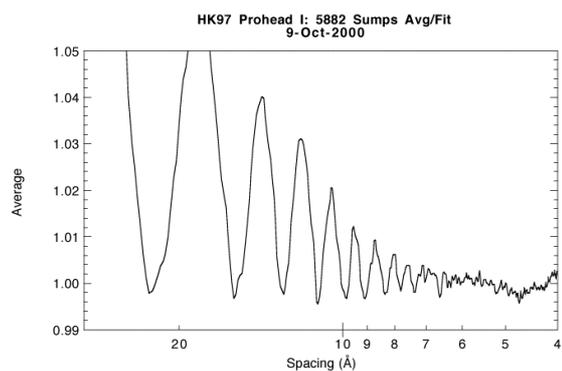
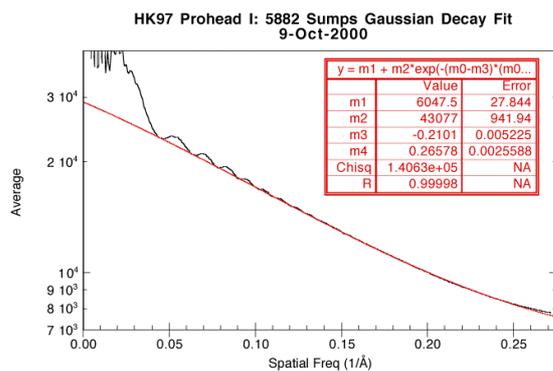
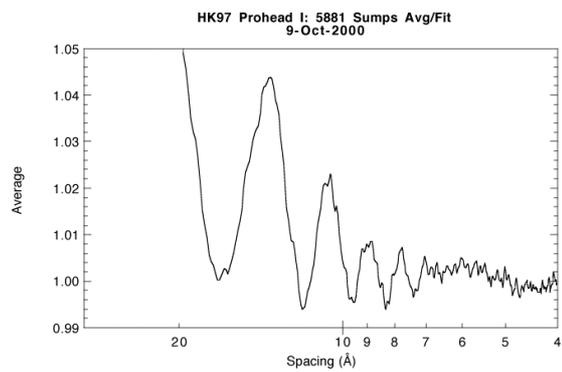
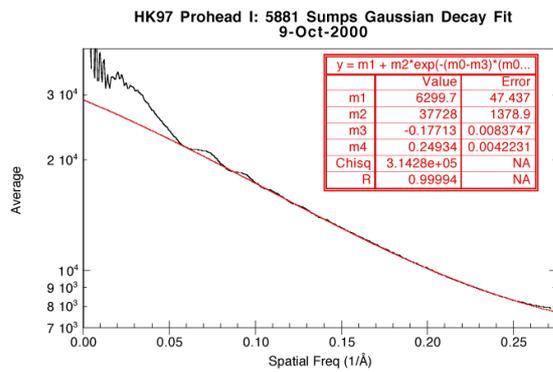
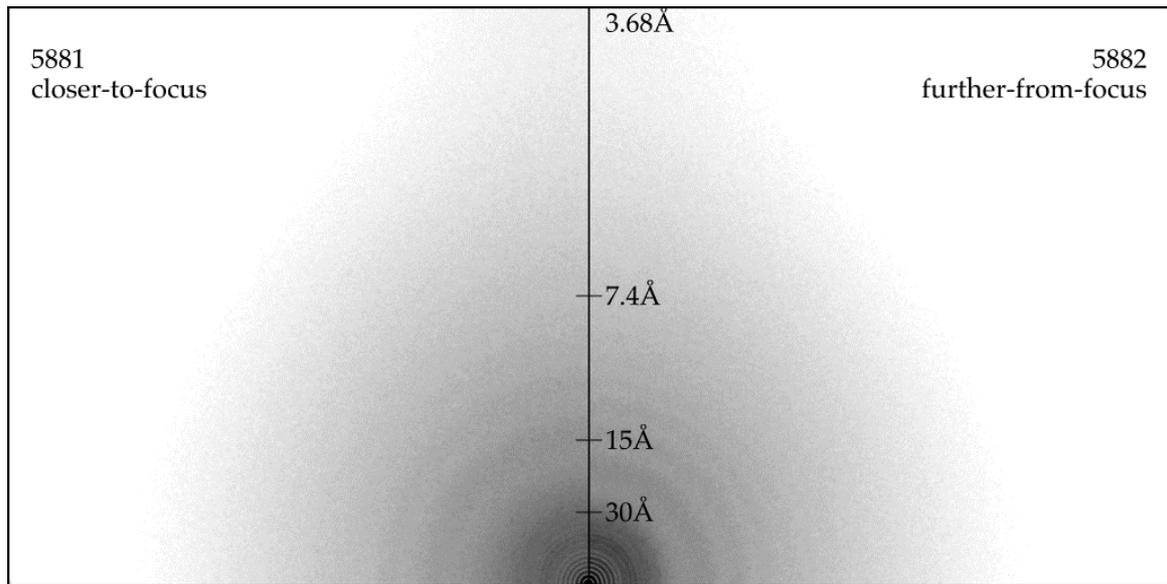
---

## 2. Sample Data

Note: This section has been copied from the “SumPS” document.

The following is the analysis of images extracted from a defocus pair of micrographs. Two runs of SumPS are made, one for each dataset of images extracted from a micrograph. The averaged PS images are shown, together with plots of the 1D profiles. The sample is HK97 Prohead I from a Phillips CM20 FEG at 38,000, scanned at  $7\mu\text{m}$  ( $1.842\text{\AA}/\text{pixel}$ ). Micrographs 5881 and 5882 are a focal pair (closer-to- & further-from-focus, respectively).

In this example, data from 607 images have been averaged. In addition to the 1D profile over the whole plane, profiles have been calculated within 10 segments, each corresponding to  $18^\circ$  ( $180^\circ/10$ ). The process was repeated for the further-from-focus micrograph (5882) and results are shown below. The two average power spectra images are shown side-by-side for comparison – minima are light – and plots of the 1D profiles follow. These plots include a Gaussian fit of the minima (left – curves in red) which is subsequently deconvoluted to give the plots on the right where the positions of the

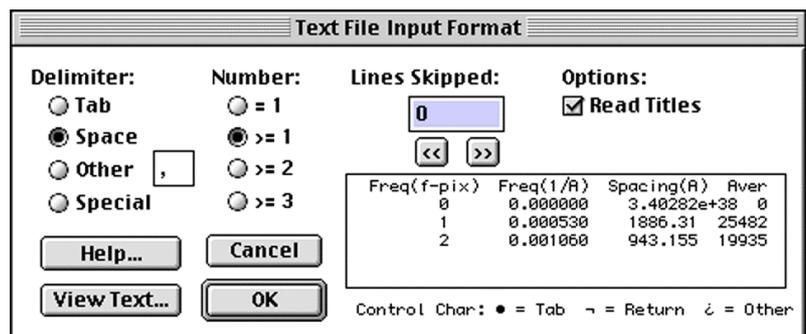


minima are more easily determined. The minima correspond to positions of phase reversal in the phase contrast transfer function of the electron microscope. Their positions can be used to determine the defocus value used for each micrograph. Note that only the full-plane profiles have been plotted. In principle, the plots of segments can be used to estimate and correct for astigmatism, but in practice the curves tend to be noisy, and I have never used them.

### 3. Through the analysis step-by-step

The 1D profiles were plotted in a Macintosh/Windows program called Kaleidagraph (Synergy Software, [www.synergy.com](http://www.synergy.com)). The Gaussian fit, deconvolution, and estimation of the positions of minima were also done with Kaleidagraph. Instructions for performing these steps follows, and the positions can be used with another program CTFZEROS for estimating the defocus value of the micrograph. (Note that I have used version 3.08d of Kaleidagraph on Macintosh computers – a recent update to 3.5.1 has become available but I have not yet seen it).

1. Transfer the SumPS output files to a Macintosh (or Windows). For example, use “Fetch” on a Mac specifying “Ascii” (or “Text”) for the 1D profile file, and “Binary” for the others.
2. Import the SumPS profile into Kaleidagraph. Use the File->Open menu item to see the dialog box at right. Notice the settings selected: delimiter is “Space” and the number is “>=1”; 0 lines are skipped, and the option to read titles is selected.



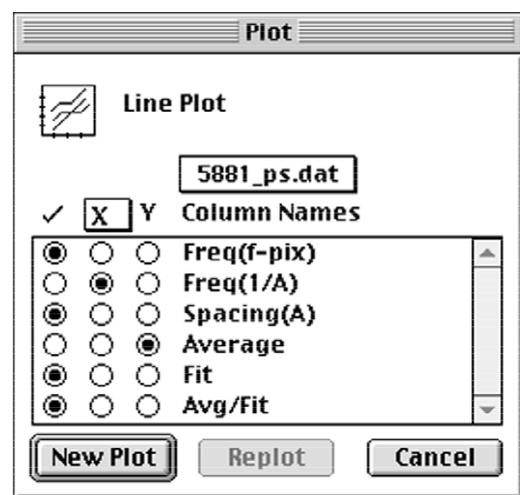
Click “OK” to import the file.

3. The Data window will look as shown at right. Note that the left-most column, C0, is empty due to the leading spaces in the input text file. This column can be removed by using the Data->Delect Column menu item. Since column arithmetic is used in the procedure, you should be consistent about keeping or deleting the empty column. I usually delete it, and subsequent screen shots of the Data window have this done. The columns following are spatial frequencies (in Fourier pixels and in inverse Ångstroms), spacing (Å), and the radial average. If segment averages have been calculated, they will follow. Two additional columns need to be added – one for values of the Gaussian fit, the other for the results of the Average/Fit. If there are no segment average columns, then use the Data->Append Column menu item. If there are segment average columns, select the first and use the Data->Insert Column menu item. The Data window should now look like this:

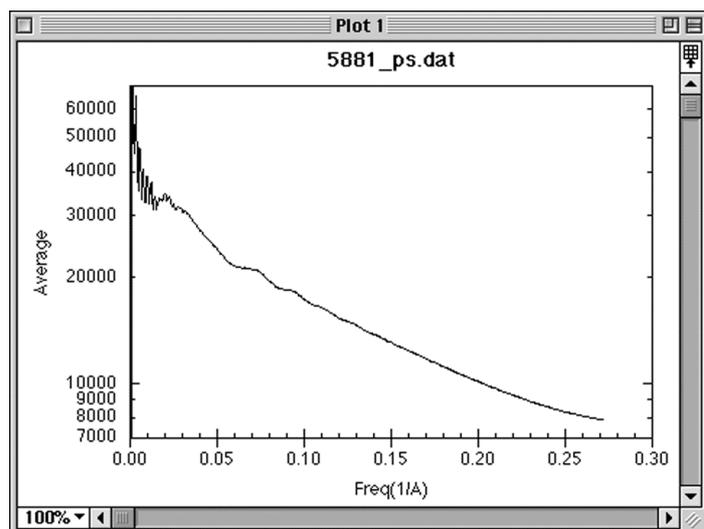
	C0	C1 Freq(f-pix)	C2 Freq(1/Å)
0		0.0000	0.0000
1		1.0000	0.00053000
2		2.0000	0.00106000
3		3.0000	0.00159000
4		4.0000	0.00212100
5		5.0000	0.00265100
6		6.0000	0.00318100
7		7.0000	0.00371100
8		8.0000	0.00424100
9		9.0000	0.00477100
10		10.0000	0.00530100
11		11.0000	0.00583100

	0	1	2	3	4	5
	Freq(f-pix)	Freq(1/Å)	Spacing(Å)	Average	Fit	Avg/Fit
0	0	0.000000	3.4028e+38	0.0000		
1	1	0.0005300000	1886.3	2.5482e+05		
2	2	0.0010600000	943.16	1.9935e+05		
3	3	0.0015900000	628.77	1.2229e+05		
4	4	0.0021210000	471.58	51474		
5	5	0.0026510000	377.26	45064		
6	6	0.0031810000	314.39	65765		
7	7	0.0037110000	269.47	60586		
8	8	0.0042410000	235.79	39286		
9	9	0.0047710000	209.59	35295		
10	10	0.0053010000	188.63	46907		
11	11	0.0058310000	171.48	46389		

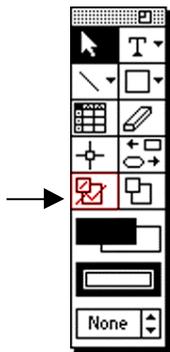
4. To plot the average, a pre-configured template can be used (such as a previously created plot) via the Gallery->Template menu item. This is by far the quickest method, but make sure that any curve fits have been disabled before using a Plot window as a template, otherwise the fits are automatically reproduced. If no template is available, then use the Gallery->Linear->Line menu item, and select the columns as shown to the right. Note that the Frequency column must be used rather than the easier-to-interpret Spacing for the Gaussian fit. The Plot window will need adjustments made to the X and Y axes, which can be done by double-clicking on one of the axes, or choosing the Plot->Axis Options menu item. The X-axis should be linear (not logarithmic) and start at 0. The Y-axis should be logarithmic and of course the minimum and maximum values must be positive.



The Plot Window (right) shows suitable choices that have been made for the limits of the axes. Cosmetic issues, such as the representation of ticks, gridlines and numbers, can be found in the Kaleidagraph manual and are not described here. Structure information is visible on the left, between spatial frequencies of 0 to 0.033 (1/30Å) followed by the tail with several dips at 0.058, 0.085 and 0.105 Å<sup>-1</sup>, which correspond to 1/12Å, 1/17Å and 1/9.5Å respectively. These minima are the positions of phase reversal in the CTF. In order to locate them precisely, we need to divide out the decay of the curve by fitting a Gaussian to the regions of the curve in the vicinities of the minima where there is no signal, just noise.

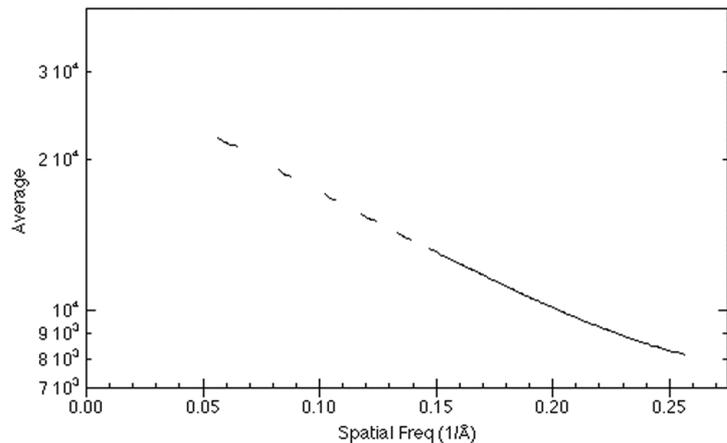
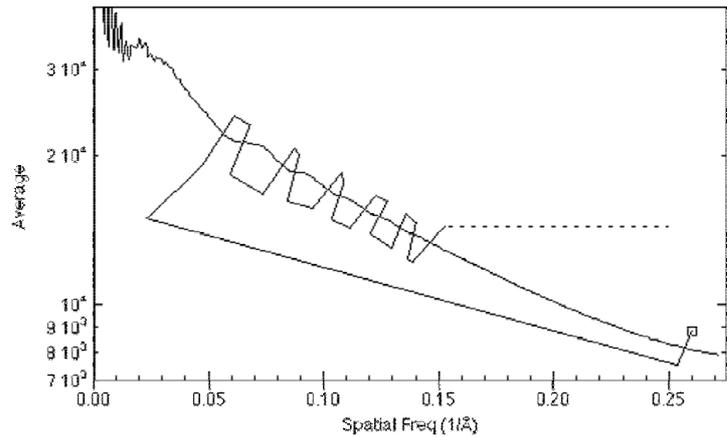


5. The regions to be used in fitting the Gaussian are selected with the Crop tool from the toolbox (arrowed below, at left). With this tool, a polygon is drawn around the regions

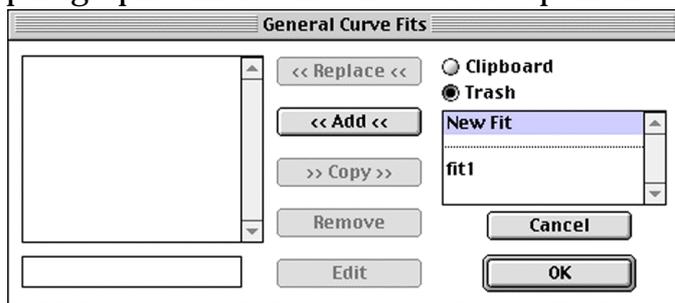


of the curve that are to be included in the fitting. An example of selection in progress, using the 5881 PS curve displayed above, is shown at upper-right. The dotted line is the current line segment, which will be placed on the starting point (small black square) to complete the selection.

Note that the visible minima have been chosen along with the tail so as to constrain the Gaussian fit to the shape of the decay, even though the very high frequencies ( $8\text{\AA}$  and higher) will not be used in the 3D analysis. Once the selection is complete, the selected regions of the curve will remain visible, and the rest will be erased, as shown in the plot at lower-right. The data corresponding to the cropped (or masked) regions are indicated in the Data Window with red backgrounds. If the masking is not optimal, the mask can be deleted by double-clicking on the Crop tool in the toolbox to restore the full curve. Note that the segments are not connected in the plot at right. Missing data breaks are enabled and disabled with the Format->Plot Extras menu item which opens a dialog box.

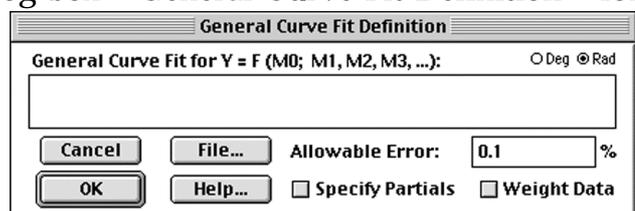


6. The Gaussian fit must be set up if it is not already in place. The Curve Fit->General menu item will show a sub-menu where the user-specified fits will be listed. This paragraph will describe how to set up the fit – if the Gaussian fit is already listed, go on to paragraph 7. Select the Curve Fit->General->Edit General menu item which brings up a dialog box that will look similar to the one shown at left (General Curve Fits).



Select “New Fit” on the right hand side, and click on the “<<Add<<” button to copy it to the left hand side. Select the “New Fit” item on

the left, and rename it to “Gaussian” using the data entry field underneath – this name will appear in the Curve Fit->General sub-menu. Select the new “Gaussian” item and click the Edit button to open a new dialog box – General Curve Fit Definition – for describing the fit. Kaleidagraph comes with a folder of predefined fits, and we can most easily set up the Gaussian fit by reading in the appropriate file. Click on the “File” button which opens a text window, go to the File->Open menu



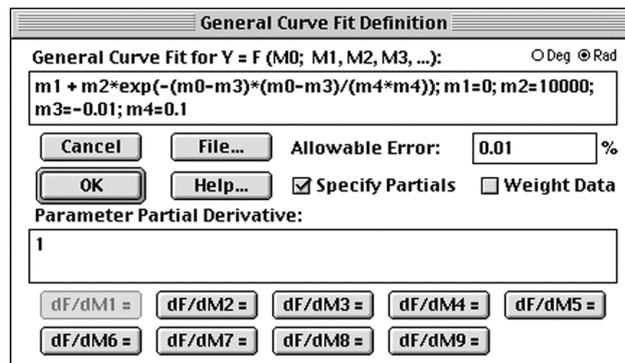
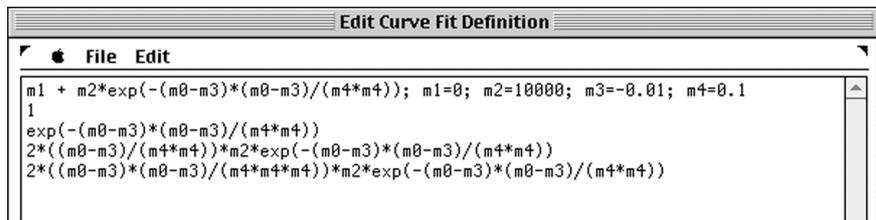
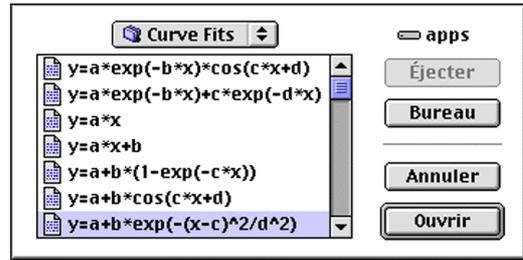
item, and navigate into the folder where the Kaleidagraph application is located. Open the “Examples Folder” and then the “Curve Fits” folder, and look for a file called:

$$y = a + b * \exp(-(x - c)^2 / d^2)$$

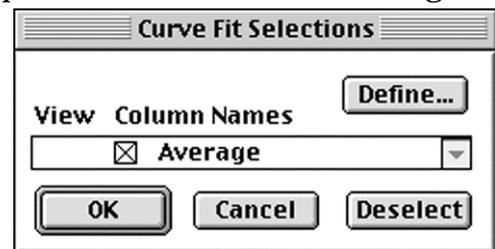
as shown at right. If these folders are not present, then copy them from the Kaleidagraph installation media. In the “Edit Curve Fit Definition” window,

change the starting values for m1-m4 as shown, then select the File->Exit menu item. In the “General Curve Fit Definition” dialog

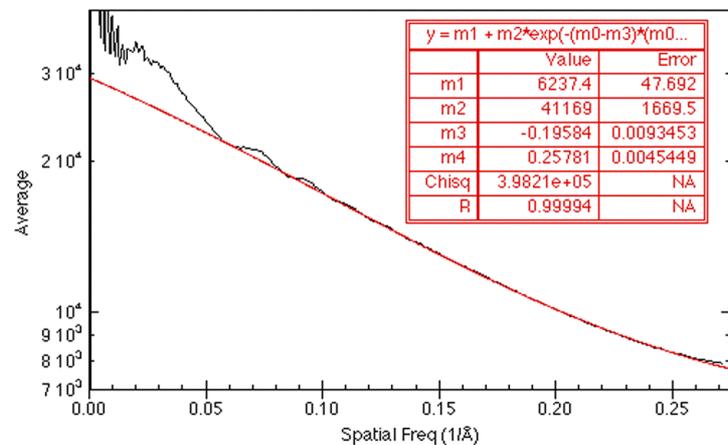
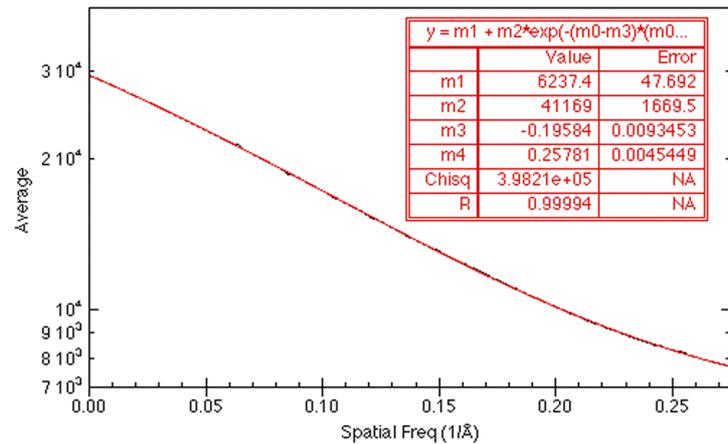
box select the ‘Specify Partials’ option, and enter an Allowable Error of (say) 0.01%. The final form of the dialog box is shown below. Click the “OK” button to exit. When you quit from Kaleidagraph, remember to allow the saving of macro’s so that this fit will be available the next time you run the program.



- With the minima and tail of the plot selected as in (5), select the Gaussian fit from the Curve Fit->General submenu. A dialog box will open – select the curve (Average) as shown, and also check the settings by pressing on the “Define...” button. This will show a dialog box from (6) – ensure that the “Specify Partials” check-box is enabled and that the starting values for m1-m4 are reasonable. Note also the “Allowable Error” field. If this is too small, or the data too noisy, then the fit may be unsuccessful, or maybe unreasonable and will likely take a long time to settle (it will oscillate between two sets of values for m1-m4). If the fit takes more than a second or two, the Allowable Error is probably too small. Try values from 0.01% (noisy data) to 0.001% (good data). The point is not to get an exact fit – that is not possible in any case because the selections will include the minima and neighbouring values – but to get a useful fit, A final point – if the fit returns an error about a singular matrix, then there are two possibilities – you used the Spacing instead of Frequency column for the X-axis (step 4), or the starting values are not good. In the latter case, ensure that m1 is small (eg, 0), m2 is about the magnitude of the curve at Freq=0 (about 30000 in this example), m3 is negative (try values between -0.1 and -0.3), and m4 is positive (0.1 to 0.3). When successful, the fit will display on top of the selected curve, as shown below. If the information about the fit does not display, use the Plot->Display Equation menu item to make it visible. The values for m1-m4

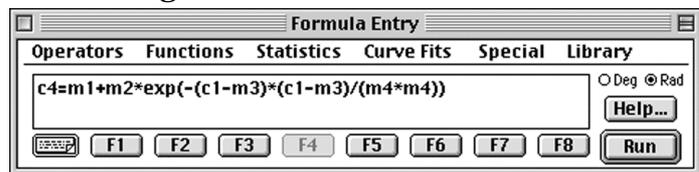


should be recorded – they will be used for deconvolution of the Gaussian decay from the images with the program CtfMix, prior to 3D reconstruction. If the fit is not extrapolated to the limits of the X-axis, then use the Format->Curve Fit Options menu item to open a dialog box that will allow this to be enabled. Finally, the complete data curve may be redrawn without triggering a re-fit by the following method (which is poorly documented in the Kaleidagraph manual): go to the Data Window corresponding to this plot (5881.dat), select the “Freq(1/Å)” and “Average” columns (or select all columns to save time); chose the Functions->Unmask menu item to remove the masking; reselect the Plot Window and chose the Plot->Plot Style menu item; dismiss the resulting dialog box by hitting the OK button (ie, don't change anything). This is sufficient to redraw the curve, unmasked, without changing the fit. The resulting Plot Window is shown at right. Note that keyboard short-cuts save a lot of time: click on the Data Window to bring it to the front; command-



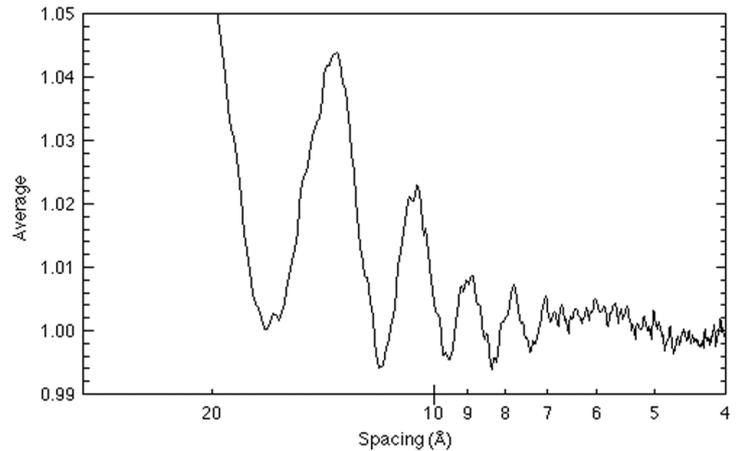
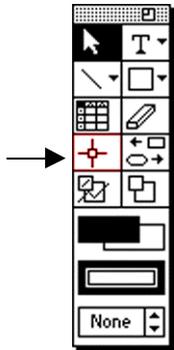
A to select all, command-] to unmask, click on the Plot Window to bring it to the front; command-M to open the Plot Style dialog box, and Enter to close it.

- Values corresponding to the fit curve can be generated into the Data Window's 'Fit' column using the Windows->Formula Entry menu item to open the “Formula Entry” window. Eight different formulas maybe stored, and I keep two reserved for generating the “Fit” column of values, and the “Average/Fit” column. In one formula, enter the equation as shown (this is the Gaussian fit equation, of course). Note the columns involved – C1 is the “Freq(1/Å)” column, C4 is the empty “Fit” column. If there is an empty C0, as mentioned above in step 3, you may find that the columns do not match. You can either adjust the names in the Formula Entry window, delete the empty C0, or select the “Freq(f-pix)” so that it is now C0. The m1-m4 coefficients hold values from the most recent fit, so don't forget to run this formula before doing any other fitting. Make sure that the correct Data Window is at the front, then select the “formula Entry window and click the “run” button. Check that the 'Fit' column now holds the newly generated values. The second formula is simply “C5=C3/C4” which writes the values of the “Average” column divided by the “Fit” column into the “Average/Fit” column.



- To plot the Average/Fit against Spacing, another template plot is useful. As before, select the template plot window, and choose the Gallery->Template menu item. Use

the “Spacing(Å)” column for the X-axis, and the “Average/Fit” column for the Y-axis. If you have not used a template, then select the menu item Gallery->Linear->Line, choosing columns as above, and adjust the X-axis to be reversed and plotted as logs. In the example shown at right, the minima are readily visible out to the 5<sup>th</sup> order at ~7.4Å. The positions of the minima can be found with the Identify tool (arrowed, at left) which is above the Crop tool in the Toolbox. Position the Identify cursor at a minimum and click the mouse to see the coordinates displayed in the top left of the Plot Window. Values from this example are: 16.8Å, 11.8Å, 9.6Å, 8.28Å and 7.36Å. These will be used with the program CtfZeros to fit a value for the defocus.



#### 4. CtfZeros

CtfZeros will fit the defocus to the positions of minima using the phase contrast transfer function curve (see [1]). Data are entered as “order,position” so that the example above would look like this:

```

CTF Zeros v1.3
=====
James Conway, compiled on Mar  2 2000
Program to determine the defocus given the measured positions of minima
in the Contrast Transfer Function (CTF). The positions of minima can be
measured by taking the radial average of the power spectrum of the scanned
micrograph.

General parameters
=====
What is the spherical aberration constant (in mm) ? 2
What is the accelerating voltage (in keV) ? 120
What is the percentage of Amplitude Contrast (eg, 7%) ? 7

Summary of input parameters
=====
Spherical aberration constant = 2 mm
Accelerating voltage          = 120 keV
Amplitude Contrast            = 7 %

Summary of derived parameters
=====
Electron wavelength           = 0.0335087 Å

Enter zeros. You will be asked for one zero at a time, along with
the order - an integer from 1 up. Enter 0, 0 to exit the loop.
Input order and position (Å) (eg, 1, 20.2) ? 1: 1,16.8
Input order and position (Å) (eg, 1, 20.2) ? 2: 2,11.8
Input order and position (Å) (eg, 1, 20.2) ? 3: 3, 9.6
Input order and position (Å) (eg, 1, 20.2) ? 4: 4, 8.28
Input order and position (Å) (eg, 1, 20.2) ? 5: 0,0
Order = 1; position = 16.8
Order = 2; position = 11.8
Order = 3; position = 9.6
Order = 4; position = 8.28

```

```

Center of defocus search is 8190 Å
Searching 2000 steps of defocus...
Start: -1000 302 1000
Search finished.
Range of defocus values used: 3190-13190 in 2000 steps of 5
Avg sum-sq-error = 5.44034e-09; defocus = 8295

```

Order	Measured Zero	Calc Zero	Difference	Freq Diff
1	16.8000	16.8350	-0.0350	1.238550e-04
2	11.8000	11.8027	-0.0027	1.963634e-05
3	9.6000	9.5933	0.0067	-7.313158e-05
4	8.2800	8.2784	0.0016	-2.349201e-05
5	-----	7.3806	-----	-----
6	-----	6.7171	-----	-----
7	-----	6.2003	-----	-----
8	-----	5.7828	-----	-----
9	-----	5.4360	-----	-----
10	-----	5.1418	-----	-----

Note the microscope settings: spherical aberration constant (Cs) is 2mm, the accelerating voltage used was 120 kVolts, and the proportion of amplitude contrast assumed is 7%. The positions of minima are entered with their order so that any too difficult to measure can be left out. Often the first order may be difficult because the minimum is not symmetric, and high orders become small and noisy.

The defocus that is fit in this example is 8295Å. The error is  $5.4 \times 10^{-9}$ , which is rather good. Errors above  $1 \times 10^{-6}$  are generally less acceptable.

With this defocus (and micrograph settings used here), and the Gaussian decay from step 7 above, the images can be corrected in the CtfMix program. The procedure is described in another document.

---

## 5. References

1. Conway JF & Steven AC (1999) Methods for reconstructing density maps of "single" particles from cryoelectron micrographs to subnanometer resolution. *J Struct Biol* **128**, 106-118.  
<[http://www.ncbi.nlm.nih.gov:80/entrez/query.fcgi?cmd=Retrieve&db=PubMed&list\\_uids=10600565&dopt=Abstract](http://www.ncbi.nlm.nih.gov:80/entrez/query.fcgi?cmd=Retrieve&db=PubMed&list_uids=10600565&dopt=Abstract)>