SAXS CORRELATION FUNCTIONS : NEW SOFTWARE AT DARESBURY

Anthony J Ryan
Manchester Materials Science Centre, UMIST, Grosvenor Street, Manchester M1 7HS
and DRAL Daresbury Laboratory, Warrington, WA4 4AD.
(email : tony.ryan@umist.ac.uk)

1. Introduction

A set of programs for performing correlation function analysis of one dimensional SAXS patterns have recently been written by Tom Nye (an undergraduate in the Applied Mathematics Department at Cambridge) who worked for me for a month in the summer. This report is based on the instruction manual he wrote to accompany the programs (which must be the first test programs I have ever received with decent documentation!). The programs are based on the SUN network at Daresbury, but have also been copied across and are working on a SUN at UMIST. From a given SAXS image they calculate the one and three dimensional correlation functions, and analyse the one dimensional function in terms of an ideal lamellar morphology[1]. The two correlation functions (called Γ_1 and Γ_3 respectively) are essentially Fourier transforms of the given one dimensional SAXS curve[2]. They are often interpreted in terms of an imaginary rod moving through the structure of the material from which the SAXS curve was obtained. $\Gamma_1(r)$ is the probability that a rod of length r moving through the material has equal electron densities at either end. Hence a frequently occurring spacing within a structure shows up as a peak in the one dimensional correlation function. The difference between Γ_1 and Γ_3 lies in the assumptions made about the experimental material. The interpretation of Γ_1 assumes that, within the SAXS length scale, spacings occur along one fixed axis, but that the axis assumes all possible directions throughout the material. Similarly, Γ_3 assumes that spacings can occur in all three dimensions within the SAXS length scale.

The task of calculating and interpreting the correlation functions can be broken down into three logical parts:

• Extrapolation of the experimental SAXS curve to $q=\infty$ and $q=0$.
This is a mathematical requirement for the Fourier transform to be performed. Any SAXS experiment gives a finite number of intensity values at finite values of q necessitating this extrapolation and a numerical integration to calculate the transforms. It should be stressed that extrapolation to $q=\infty$ (tail fitting) is the most problematic task of correlation function analysis, and can greatly influence results obtained; the programmes are called `tailfit` and `tailjoin`..

• Fourier transformation of the extrapolated data.
Γ_1 is based on a cosine Fourier transform, while Γ_3 is based on a sine transform. The two functions are related by a simple expression[2]. The program `transform` performs the Fourier transformations, and is relatively simple. It can also re-transform Γ_1 back into a smoothed copy of the extrapolated data.

• Interpretation of Γ_1 based on an ideal lamellar morphology.
A model is required for the interpretation of features of Γ_1 to be possible. The program `extract` performs this analysis. Note that no interpretation of Γ_3 is performed.

Correlation functions are extremely valuable tools in the interpretation of one dimensional SAXS patterns, particularly those for which features are either weak or obscured (for example shoulders). However, as with any involved method of analysis, care should be taken to ensure results are genuine and meaningful.

2. Program structure.

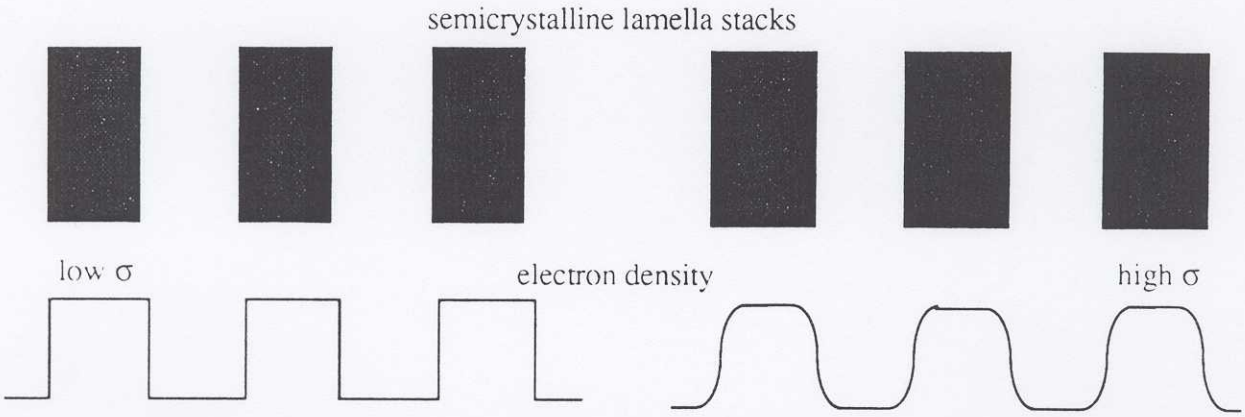
The programs were written in a modular form to enable the code to be incorporated in the GUI version of otoko due over the next few years. The programs also had to be transportable between systems. Each program is run from a UNIX shell script. The shell script runs each of the programs in turn, and also controls graphics output. The programming behind this is complicated, but was written to ease transport of the programs. To plot a graph, a program creates otoko format files containing the data, and a set of shell scripts. The shell scripts run otoko and force it to plot the data contained in the files just created. This may seem an unnecessarily complicated way of producing plots, but it can be used on any system that supports the Xwindows version of otoko, and in general works seamlessly. The programs require an intensity file corrected for the transmission, background and detector response and a q-axis from otoko

3. Programs

`Tailfit` fits either a Porod[3] or sigmoid[4] profile to the tail of the intensity data. It then passes the fit parameters to `tailjoin` for extrapolation of the data to q=∞. The sigmoid profile should generally be used. It uses an intensity profile of the form:

$$I(q) = B + Ke^{\sigma^2 q^2} / q^4$$

where B is a Bonart thermal background, K is the Porod constant, and σ describes the electron density profile at the interface between crystalline and amorphous sections



The parameter σ is in Å, and is a measure of the width of the distribution that gives rise to the sigmoid electron density profile. Fitting with a sigmoid profile is a non-linear problem, and requires a robust fitting algorithm: a Levenburg-Marquart method[5] is used.

The Porod tail profile has the form

$$I(q) = B + K / q^4$$

and usually gives poor quality correlation functions, it is included for completeness. However, if a Porod profile is used, a greater number of structural parameters can be extracted, including interface surface to volume ratios and the Porod characteristic chord length. Fitting becomes a linear problem, and so can be less problematic than fitting to a sigmoid profile.

The most important parameters controlled by the user are those concerning which data are used for the tail fitting. If time-resolved data is being analysed, the user can input different channel numbers for each frame. Choosing limits can be difficult, and some experimentation may be required. As a general rule, the start channel should have a q value about twice that at which any peak occurred in the intensity data. The end channel should be as large as possible, but should avoid any detector noise.

Since the experimental data is extrapolated to very high q, the majority of the data used in the Fourier transform comes from the tail fit. It's therefore very important to check that the tail fit is good. The tail affects points on the correlation function at low values of r (real space coordinate) to the greatest extent, but these points are the most important in the extraction of ideal lamellar morphology parameters. Hence the results from an analysis session depend greatly on the tail fit. Choosing channel limits is a payoff between using as many points as possible to ensure a good fit, but wanting to keep as many points as possible from your experimental data in the extrapolated data that is passed to the Fourier transform. A noisy tail on the experimental data will result in poor tail fits, possibly making correlation function analysis impossible.

`Tailjoin` extrapolates the experimental data back to q=0 using a Guinier model[6]. It then creates otoko files containing intensities from q=0 to beyond q=0.6 using the parameters found by `tailfit`. The intensity profile used by the Guinier model has the form:

$$I(q) = Ae^{Bq^2}$$

where B is negative. The Guinier model assumes the small angle scattering arises from particles and the parameter B is related to the radius of gyration of those particles. This obviously has dubious applicability to polymer systems. However, the correlation function is affected by the Guinier back-extrapolation to the greatest extent at large values of r, and so the back-extrapolation only has a small effect on the analysis. The Guinier profile is fitted to the first few genuine scattering points after the beamstop. If your experimental data does not contain an upturn in intensity at low q, back extrapolation may fail. As an alternative to the Guinier profile, a Vonk profile can also be used

Due to the nature of the tail fitting, the join between experimental data and data in the calculated tail usually involves a step that could cause ripples in the correlation function. Hence this join is smoothed using a Savitzky-Golay[5] smoothing algorithm that smoothes the joins without greatly altering higher moments of the data. The point in q to which extrapolation is performed affects the correlation functions, particularly if it is too small. The value of q=0.6 used by the programs was decided on after experimentation, and gives smooth correlation functions without loss of speed. The lower the truncation point is, the rougher the correlation functions, while the higher the truncation point is, the slower the transform. A truncation point of q=0.6 corresponds to fluctuations in the correlation functions of about 10 Å. These are usually not observable.

`Transform` performs the integration's necessary to calculate the correlation functions and second moment of the data. It also has the capability of re-transforming Γ_1 back into a scattering curve. Γ_1 and Γ_3 are given by:

$$\Gamma_1(r) = \frac{1}{Q}\int_0^\infty I(q)q^2\cos(qr)dq$$

$$\Gamma_3(r) = \frac{1}{Q}\int_0^\infty I(q)q^2(\frac{\sin(qr)}{qr})dq$$

where $I(q)$ is the scattering intensity and Q is the second moment or invariant given by

$$Q = \int_0^\infty I(q)q^2dq$$

Hence $\Gamma_1(0) = \Gamma_3(0) = 1$. Notice that every point in the extrapolated dataset will be used to calculate each point on the correlation functions, leading to a smooth correlation function. Of course, the integration is numerical and is only performed up to q=0.6 as discussed in the last section. Together with the fluctuations introduced by this truncation, fluctuations are also introduced into the correlation functions by the finite gap between points in the extrapolated

dataset - we don't have intensity as a continuous function of q. As a final comment, note that the numerical integration takes the form of a trapezium approximation.

`Extract` interprets Γ_1 in terms of an ideal lamellar morphology, and extracts structural parameters. The program also displays the results of calculating the moments of the experimental SAXS curve, and Porod results. If tail fitting was performed using a Porod instead of a sigmoid profile, a more detailed Porod analysis is performed. The program first decides whether a lamellar interpretation can be applied. It searches for the first local minimum with a negative Γ_1 coordinate and the first local maximum with a positive Γ_1 coordinate. If these cannot be found, extraction of the lamella structural parameters is abandoned (for that particular frame). If these features are found extraction can be performed, and will employ these two features. Note these criteria very carefully: local minima above the abscissa will be ignored and not used in the calculation of structural parameters. Similarly, local maxima below the abscissa will not be used in the calculation of structural parameters. Indeed, the interpretation of any one dimensional correlation function deviating from the ideal lamellar model is not yet properly understood.

A diagram of the one dimensional correlation function from an real lamellar morphology is given later. It is essentially a damped Patterson function and consists of a gradually decaying oscillation, with an initial linear section at low values of r. Structural parameters are derived from the positions of the first local minimum and local maximum, and the position and gradient of the linear section, the complete set of parameters extracted are given below.

Parameter	**Symbol**	**Measurement**
Long period	L_p	As in diagram
Bulk crystallinity	ϕ	$\Gamma min / (\Gamma min + \Gamma^*)$
Shortest block thickness	L_c	As in diagram
Longest block thickness	L_a	$L_p - L_c$
Local crystallinity	V_1	L_c / L_p
Average core thickness	D_0	As in diagram
Average interface thickness	D_{tr}	As in diagram
Polydispersity		$\Gamma min / \Gamma max$
Electron density contrast	$\Delta\rho$	$Q / \phi (1-\phi)$
Specific inner surface		$2\phi / L_c$
Non-ideality		$(L_p - L_p^*)^2 / L_p^2$

Most of these parameters are given in reference[1]. The polydispersity measurement was suggested to Tom by Guy Eeckhaut of ICI. The problem of determining D_{tr} and D_0 has not been straightforward. Algorithms have been developed with consistency as a priority, so that even if the structural interpretation of these parameters is dubious, the line they specify gives consistent values to the bulk crystallinity, hard block thickness and so on. This has not been an easy task, made harder by the fact that poor tail fits radically alter the appearance of Γ_1 at low r.

`Extract` is an exception to the rule that all user input occurs in `tailinput`, should the lamella model fail then the user is requested to input values of the crystallinity, and some control over the initial straight-line portion is also possible. If a Porod tail fit has been used then extract will perform the surface to volume analysis.

7. Using the program on real data

An example of a SAXS pattern from a time resolved experiment on polypropylene is given below. The data were collected in SAXS/WAXS/DSC mode on beamline 8.2 with a time resolution of 6s. Note that the data are still quite noisy but the tail joins are good as indicated by the smooth Γ_1 which turns over at r -> 0. There is good agreement between the long-

spacing from the correlation function and that obtained by applying Braggs Law to the peak in the Lorentz corrected intensity. In this experiment 256 frames were taken during melting and recrystallisation and running the program on the full set took ≈ 45 minutes on a SUN SPARC IPX. The correlation function analysis proved invaluable in analysing the lamella thickness and local crystallinity during melting and recrystallisation and given crucial insights into these processes[7].

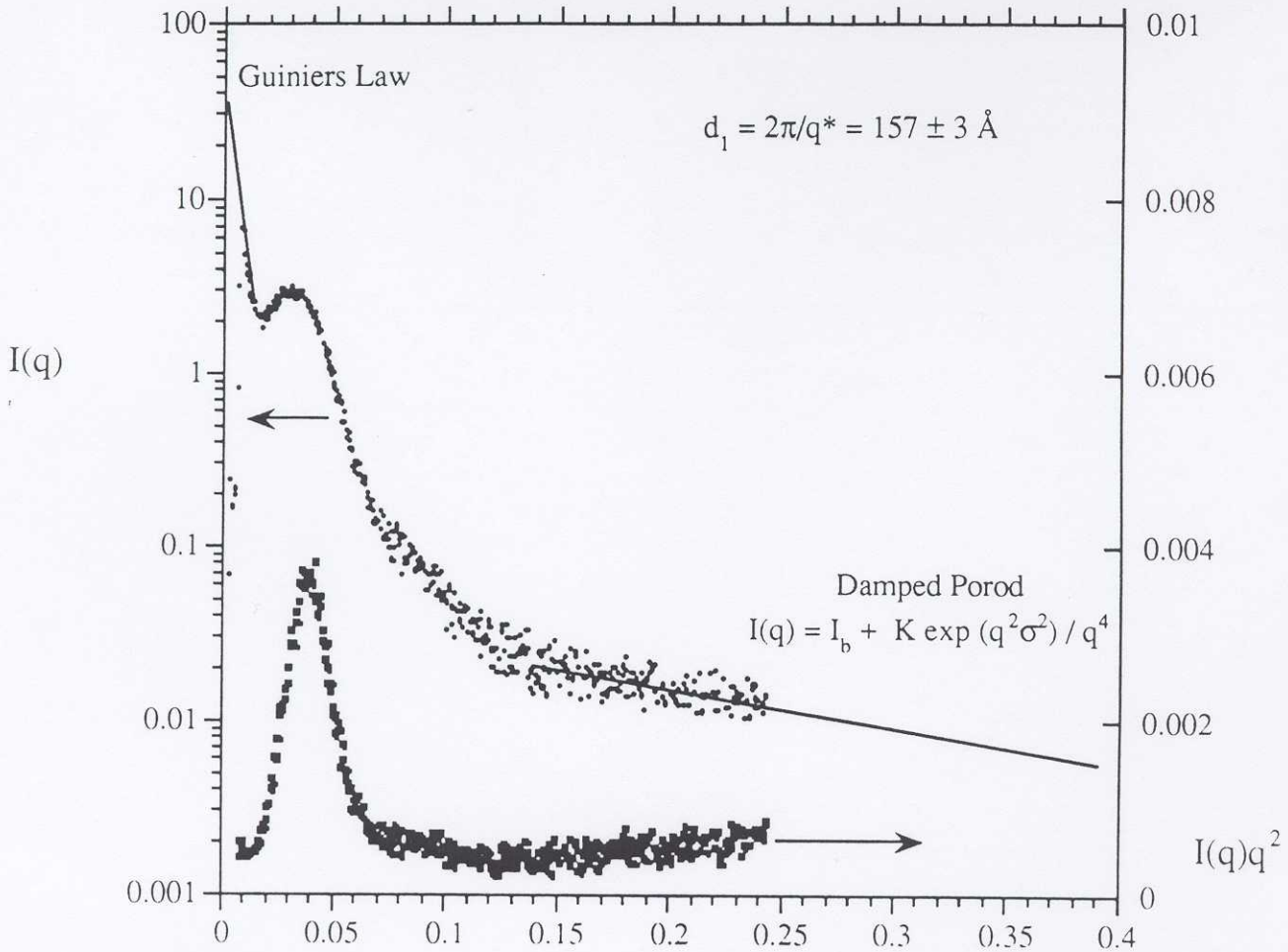
8. Conclusion and recommendations.

The crux of the problem lies in tail fitting. The Fourier transformation of data seems to be error free, and artefacts introduced by the extrapolation, while certainly present, don't often affect the correlation functions noticeably. But the appearance of the correlation functions, and the structural parameters extracted, do depend greatly on the tail fit. Hence great care should be taken when selecting channel limits for the tail fit. For realtime data it may be worthwhile to go through the dataset frame by frame, using the graphics option to check each fitted tail. It's also clear that the selection of the linear section of Γ_1 in `extract` is yet to be perfected. While this is an annoying problem, it is not a vital aspect of the analysis programs since the user can always intervene. Extraction of structural parameters has concentrated on Γ_1 up to the present. If possible, a program equivalent to `extract` will be written to analyse Γ_3. Future developments might include calculation of the interface distribution function from SAXS data and implication of the various methods for calculating degrees of crystallinity from WAXS data[2].

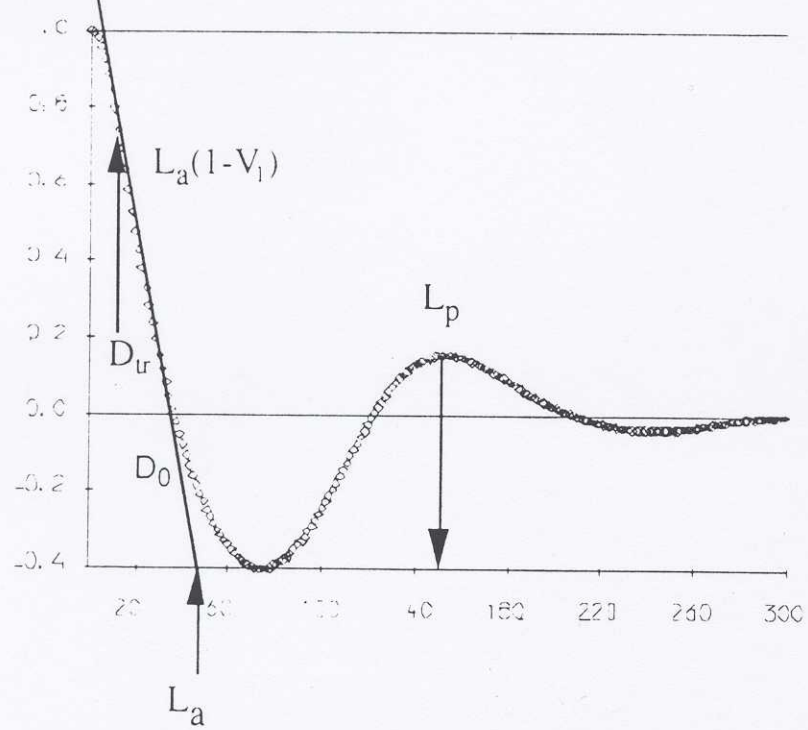
References.

1. Strobl, G. R.; Schneider, M. *J. Polym. Sci.* **1980**, 18, 1343-1359.
2. Baltá Calleja, F. J.; Vonk, C. G. *X-ray Scattering of Synthetic Poylmers*; Elsevier: Amsterdam, 1989; pp.247-251.
3. Baltá Calleja, F. J.; Vonk, C. G. *X-ray Scattering of Synthetic Poylmers*; Elsevier: Amsterdam, 1989; pp.257-261.
4. Koberstein, J.; Stein R. *J. Polym. Sci. Phys. Ed.* **1983**, 21, 2181-2200.
5. Press, W. H. et al. *Numerical Recipes: the Art of Scientific Computing*; Cambridge University Press, 1986.
6. Baltá Calleja, F. J.; Vonk, C. G. *X-ray Scattering of Synthetic Poylmers*; Elsevier: Amsterdam, 1989; pp.260-270.
7. Ryan, A.J.; Stanford, J.L.; Bras, W.; Nye, T.; *J. Polym. Sci . Phys .Ed.*submitted.

EXTRACTING DATA FROM THE CORRELATION FUNCTION



$$\gamma_1(r) = \frac{1}{Q}\int_0^\infty q^2 I(q)\cos(qr)dq$$



frame 15

L_p = 153 Å

L_a = 44 Å

L_c = 109 Å

V_1 = 0.71

D_{tr} = 10 Å