

To Prof. P. R. Pisharoty

With regards

Suhasini

PRL

TECHNICAL NOTE

PRL-TN-85-47

Decomposition of data into a
sum of non-linear functions -
Application to Astronomical
Observations.

By

B.G. Anandarao and
Suhasini R. Rao

August 1985

PHYSICAL RESEARCH LABORATORY
AHMEDABAD - 9

Decomposition of data into a sum of non-linear
functions- Application to Astronomical observations.

B.G.Anandarao and Suhasini R.Rao
Physical Research Laboratory
Ahmedabad-380 009, India.

Abstract:

In astronomical observations we often encounter spectra which are a result of superposition of individual contributions from phenomena of similar kind. We describe here a method together with its algorithm by which one can decompose a given set of data into a number of non-linear functions and apply the same to a couple of astronomical data-sets. We discuss the modifications to be made in the algorithm, appropriate to the particular problem in order to arrive at a solution. This method can be used in high-resolution spectroscopic and photometric data on astronomical objects to extract valuable information which otherwise is not apparent.

1. Introduction:

In many scientific problems, the data obtained represent the resultant of superposition of a number of functions of a similar type, but of different parametric values. The type or form of these functions depends upon the physical nature of the source. Therefore, the problem that we address ourselves to in this paper, is to resolve or decompose into individual functions (with different parameters) from a given set of data-assuming the functional form.

One often encounters spectra in astronomical observations, which are a result of superposition of several individual profiles which are not resolved unambiguously either due to a poor signal-noise ratio or to closeness of the profiles making the resolving power of the instrument insufficient. For instance, in the case of Seyfert galaxies, spectroscopic observations (of Wilson as referred to by Burbidge et al, 1959) showed partially resolved structured profile for several forbidden emission lines and the hydrogen - recombination line $H\alpha$. These structures are thought to represent the turbulent motions of several individual gas clouds. By resolving these structures into individual gaussian profiles, one can obtain the density in each gas cloud from the peak intensity of each of

these profiles and the turbulent velocities from the width of each of these profiles. In the case of a planetary nebula (Wilson, 1950; Osterbrock et al, 1966) one expects the emission lines to give a double-peaked profile indicating an expanding shell with respect to the central star in the nebula. However, in certain cases, even the high-resolution Spectroscopic technique may not be able to resolve the double-peaked structure due to the reasons mentioned earlier. Furthermore, in the case of near - and far-infrared photometric data sometimes, one needs to fit a two - or multiple - temperature model (i.e. Planck's function with different temperatures and emissivities) representing different physical conditions of the emitting regions.

We describe here a numerical method using which we can resolve or decompose a set of data into a number of non-linear (or linear) functions of assumed form. We consider two specific astronomical data-sets and apply the method to resolve the unresolved features.

2. Method:

Mathematically, the problem under investigation is to fit the data with a model function which is a sum of a number of non-linear functions of similar nature

having physically meaningful parameters of different magnitudes. The individual functions may be Gaussian, Lorentzian, exponential, Fourier or Planckian.

The model function $F(X)$ which should represent the given data-set $P(X_i)$ is written as

$$F(X_i) = \sum_{k=1}^M f_k(X_i, \theta_j), \quad i = 1, 2, 3, \dots, N. \\ j = 1, 2, \dots, l \quad \dots\dots(1)$$

where $f_k(X_i, \theta_j)$ represent the individual functions that sum up to give $F(X_i)$; θ_j are the physical parameters in each function that need be fixed so as to obtain a good fit; N is the number of data points and, M is the number of functions required to fit the data; and l is the number of parameters (θ_j).

The deviation of the model from the data at each data point is defined as

$$D(X_i) = F(X_i) - P(X_i) \quad \dots\dots(2)$$

then the residual sum of squares ϕ can be written as

$$\phi(\theta_1, \theta_2, \dots, \theta_L) = \frac{1}{N} \sum_{i=1}^N \left[D(X_i)^2 \right]; \quad N \geq L \quad \dots(3)$$

where $L = 1 \times M$, represents the total number of parameters that need be evaluated.

Now the next step is to minimise the function with respect to θ_p ($p = 1, 2, \dots, L$) for a good fit. Assuming that ϕ is differentiable, one can write these conditions as,

$$\frac{\partial \phi(\theta)}{\partial \theta_p} = 0, \quad p = 1, 2, \dots, L \quad \dots (4)$$

(1) These equations may be linear or nonlinear in θ depending upon the functional form. So, the Problem reduces to determining the parameter set θ_p by solving these equations.

In solving such problems (nonlinear least squares and optimisation) the generalised Newton-Raphson method (Powell 1964, 1965) can be used very efficiently provided the initial approximations are chosen properly. The method essentially is an iterative procedure based on a Taylor Series expansion about the current approximation to the required Solution. In case the function to be minimised is a sum of squares and if each term in the sum of squares is relatively small at the minimum, further approximation can be made and the procedure gains substantially because there is no need to evaluate the second derivatives explicitly.

Applying Taylor's series expansion in the neighbourhood of θ_p^0 (Present approximation) and assuming that

$\theta_p = \theta_p^0 + \delta_p$, is a later approximation, one can write Eq. 4 as

$$\frac{\partial \phi}{\partial \theta_p} = \frac{\partial \phi}{\partial \theta_p^0} + \delta_p \sum_{q=1}^L \frac{\partial^2 \phi}{\partial \theta_p \partial \theta_q} + O(\delta_p^2) = 0$$

for $p=1, 2, \dots, L$ (5)

Substituting equation (3) in equation (5), and neglecting $O(\delta_p^2)$ terms,

$$\sum_{i=1}^N \frac{\partial(D_i^2)}{\partial \theta_p^0} + \sum_{q=1}^L \left[\sum_{i=1}^N \frac{\partial^2(D_i^2)}{\partial \theta_p \partial \theta_q} \right]_{\theta=\theta^0} \delta_p = 0,$$

$p=1, 2, \dots, L$

which is

$$\sum_{i=1}^N \left[\frac{\partial D_i}{\partial \theta_p} D_i \right]_{\theta=\theta^0} + \sum_{q=1}^L \left[\sum_{i=1}^N \frac{\partial^2 D_i}{\partial \theta_p \partial \theta_q} D_i + \frac{\partial D_i}{\partial \theta_p} \frac{\partial D_i}{\partial \theta_q} \right]_{\theta=\theta^0} \delta_p = 0$$

.....(6)

Near the minimum, if $D_i(\theta_p^\circ + \delta_p) = 0$, then $D_i(\theta_p)$ is of the order of δ_p ; thus the second derivative in the above equation turns out to be the order of δ_p^2 and can therefore be neglected. The eqn. (6). then reduces to,

$$\sum_{i=1}^N \left[\frac{\partial D_i}{\partial \theta_p} D_i \right]_{\theta=\theta^\circ} + \sum_{q=1}^L \left[\sum_{i=1}^N \left\{ \frac{\partial D_i}{\partial \theta_p} \frac{\partial D_i}{\partial \theta_q} \right\} \right]_{\theta=\theta^\circ} \delta_p = 0 \quad \dots(7)$$

Defining, the matrices,

$$g = \sum_{i=1}^N \left[\frac{\partial D_i}{\partial \theta_p} D_i \right]_{\theta=\theta^\circ} ; p = 1, 2, \dots, L$$

$$G = \sum_{i=1}^N \left[\frac{\partial D_i}{\partial \theta_p} \frac{\partial D_i}{\partial \theta_q} \right]_{\theta=\theta^\circ} \quad \dots(8)$$

one can write (7) as

$$\delta = - G^{-1} \cdot g \quad \dots(9)$$

which gives the solution for δ_p . and the next

approximation $\theta_p = \theta_p^0 + \delta_p$ can be calculated.

3. Algorithm:

The following are the steps we follow in solving the equation (9)

(i) Depending upon the nature of the source from which the data are obtained, we define the functional form of f_k in equation (1) and the number M of such functions which together will represent the data-set.

(ii) Then we choose a set of initial values for θ_p^0 based upon a priori knowledge of the source of the data. We shall comment on this aspect later in this paper.

(iii) We compute the deviations,

$$D_i(\theta_p^0) = P_i - \sum_{k=1}^N f_{ki}(\theta_1^0, \theta_2^0, \dots, \theta_p^0, \dots, \theta_L^0)$$

(iv) then the column matrix elements

$$g_p = \sum_{i=1}^N \frac{\partial D_i}{\partial \theta_p}(\theta^0) D_i(\theta^0) \quad p=1,2,\dots,L$$

are determined.

(v) We determine the matrix elements,

$$G_{pq} = \left[\sum_{i=1}^N \frac{\partial D_i}{\partial \theta_p} \frac{\partial D_i}{\partial \theta_q} \right]_{\theta = \theta^0} \quad ; \quad p=1, 2, \dots, L$$

$$q=1, 2, \dots, L$$

(vi) We then solve the equations

$$\delta = -G^{-1} g$$

(vii) Then we check whether $\text{Max}_p \left| \frac{\delta_p}{\theta_p^0} \right| \leq \eta$,

where η is a pre-assigned accuracy limit. If this inequality is satisfied then we go to the following step, or else we start all over again from the step (iii) above.

(viii) We then calculate $\epsilon (= \frac{1}{N} \sum_{i=1}^N D_i^2)$,

the residual sum of squares, with the new $\theta_p = \theta_p^0 + \delta_p$ and check whether $\epsilon \leq \gamma$, a limit given to the standard deviation based on the amplitude of the errors in the data. If this inequality is satisfied then we stop the iteration or else we start from step (iii) above with new θ_p^1 .

The following remarks on the nature of the solution

are in order. In general, the solution to the problem discussed here is not unique because of its non-linearity. Therefore, care should be taken in accepting a particular solution and rejecting several others. The method that we adopted here searches for a solution in the neighbourhood of the initial choice θ_p^0 . Thus, unless the initial choice is properly made, the method either diverges giving no solution, or even if it converges, it does so to a solution which may be numerically correct but has no physical significance or meaning. The first two steps in the algorithm as given above, namely, (i) choosing the functional form and the number of such functions, and (ii) assuming a set of initial values θ_p^0 , require a prior knowledge as to the nature of the source from which the data have been observed.

Even after choosing the functional form and the number of functions, one needs enough prior knowledge regarding the object in order to judge whether or not a solution is meaningful. Sometimes the stability of the parameters also causes problems in convergence of the method. In such a case the iteration need not be done simultaneously for all the parameters. One can estimate a subset of parameters which are stable, and

substituting these, the other parameters can be estimated. The whole iteration process is repeated once or twice to get the full set of parameters. The matrix G which consists of partial derivatives at data points gives an indication of the order in which the parameters to be iterated.

We have used Gauss-Seidel iterative method for solving the matrix equation given in (vi) above. In general, we have found that the method leads to fairly fast convergence, provided that the matrix G is well-conditioned.

4. Examples:

We give in this section two examples of the type of functions that we have used for two data sets obtained from astronomical observations.

(a) Gaussian decomposition of spectroscopic data:

It is fairly reasonable to assume that the emission-lines from astronomical sources have profiles of Gaussian shape. This is because in most of such cases, the line broadening is caused by the motion of particles in thermodynamic equilibrium or by the random turbulent motion of the gas. In such a situation, the resulting spectral line intensity distribution can be best approximated by a Gaussian. The data we have used for the

decomposition are the Hydrogen Balmer-alpha line observations made by Anandarao et al (1985) on the peculiar Mira variable star R Aqr by using a high-resolution Fabry-Perot Spectrometer. The model function F_i in equation (1) is

$$F_i = \sum_{k=1}^M f_{ki}(\theta_j); \quad i = 1, 2, \dots, N$$

$$j = 1, 2, \dots, 1$$

$$f_{ki} = A_k \exp \left\{ - \frac{(\lambda_{ok} - \lambda_i)^2}{2\sigma_k^2} \right\}$$

The parameters A_k , λ_{ok} and σ_k , represent the maximum amplitude, the wavelength of maximum amplitude and the width of the profile number K respectively. We have chosen $M=4$ so that the observed profile represents two expanding shells with respect to the central star - each shell represented by two Gaussian profiles one approaching edge and the other receding edge.

In figure 1, we have given the emission-line profile data alongwith the model-fit with 4 Gaussians. One can see that the fit was remarkably good giving physically meaningful parameters. We have found that a model with

3 Gaussians also fits well enough with the data. However, the case of 3 Gaussians is rather less meaningful because if there were to be 2 expanding shells, then, these should necessarily be represented by 4 profiles. Supposing that there was only one shell then it should be represented by 2 profiles. We tried to fit two profiles to the data. However, the result was not satisfactory and the asymmetry seen on the right-hand-side of the observed profile could only be accounted by assuming that the profile with the maximum amplitude in the data was actually composed of two profiles representing a single shell.

We have solved the equation (9) by taking a set of similar parameters each time. That is, for examples we first keep widths and positions of the Gaussians fixed, and find out the amplitudes; then keep amplitudes and widths fixed and compute the positions; and finally, keep amplitudes and positions fixed and obtain widths. In doing this, it is ofcourse necessary that we first keep the best-known and stable parameters fixed and vary the least known ones. Thus the solutions depend upon the initial values to a considerable extent. We had to do this kind of piece-wise computation, as we can see, in the matrices $[G]_{pq}$ and $[g]_p$, the elements corresponding to different parameters are different

by orders of magnitude. It is felt that this way of solving the problem in a piece-wise manner, reduces the risks involved in wild-goose hunt: because, when we know apriori the values of certain parameters, it is better to fix them first and solve for the unknown parameters; in the end making only slight changes in the parameters assumed apriori.

(b) Multiple Source model for far-infrared data on astronomical objects:

There are a number of sources in the sky which are very strong emitters of far-infrared radiation: molecular clouds in our galaxy, external galaxies, protostars, stars with circumstellar shells and so on. In general, each of these sources can be approximated to be emitting like a blackbody, at a certain temperature with emissivity characteristic to the nature of the source. However, there are a number of sources for which we cannot make this simple approximation. Most often, the far-infrared emission results in as a reradiation from dust grains heated by energetic photons from a hot source embedded in a dusty molecular cloud. The same is the case of a star with a circumstellar envelope. In such cases the far-infrared emission comes from two or more sources, namely, one stellar component and one (or more) dust

component. What we observe is the sum of the radiation from those sources. The problem, therefore, is to decompose the observed far-infrared flux into the fluxes from two or more individual sources.

Let us assume that $P(\lambda_i)$ is the observed data from a source. We use here a simplified version of radiative transfer problem (Jones and Merrill, 1976). We write down the following equation for the model $F(\lambda_i)$

$$F(\lambda_i) = a B(\lambda_i, T_*) + \sum_{k=1}^M b_k B(\lambda_i, T_k) \left\{ 1 - \exp(-2\alpha_k Q_k(\lambda_i)) \right\}$$

where a, b_k are constants which physically represent the source size in steradians; $B(\lambda, T)$ is the well-known Planck's function for blackbody radiation given by

$$B(\lambda, T) = \frac{2 hc}{\lambda^3} \frac{1}{\left\{ e^{hc/KT\lambda} - 1 \right\}}$$

α_k is a constant

$Q_k(\lambda_i)$ are the far-infrared absorption efficiencies of a particular type of grains that we assume; T_* is the effective temperature of the central hot source i.e. a star.

The data used here were obtained by the Infrared Astronomical Satellite (IRAS) on the Mira variable star

U Ori and were made available to us by Professor S.R.Pottasch. In table 1 we have given the fluxes obtained by IRAS detectors in four far-infrared bands centred at wavelengths of 12,25,60 and 100 microns. We know from the type of the star that its effective temperature is ~ 3000 K (typical for M-type stars). We can find readily that the far-IR fluxes given in table 1 cannot be accounted for by the star alone. We also know from the nature of Mira variables that these stars are surrounded by circumstellar dust shells which form due to the shedding of the material from the star during its later stage. With this prior knowledge regarding the object we could explain the observed fluxes very well, in terms of a two temperature model - (i) The star with temperature $T_* = 2900$ K and (ii) a dust shell at temperature $T_d = 80$ K with an assumed silicate grain model. (The details of these results will be published elsewhere). In Table 1 we have given the model fluxes with the above parameters alongwith the source sizes obtained by the model. One can see that the fit is very good.

5. Conclusions:

We have developed a method and its algorithm to decompose observational data into a given number of non-linear (or linear) functions. We have shown that physically

meaningful solutions for the functional parameters can be obtained by giving two examples of data-sets from astronomical observations. In examples where prior knowledge of the source is available, our method yields a better physical insight into the nature of the source as shown in the two examples. Though the mathematical method described here is a standardised one, the details of the usage of its algorithm depends very much on the particulars of the given data. Thus the algorithm is data-dependent.

Aknowledgements:

This work was supported by the Department of Space, The Government of India.

References:

- Anandarao, B.G., Sahu, K.C., and Desai, J.N., 1985
Astrophys. Space Sci., in press.
- Burbidge, E.M., Burbidge, G.R., and Prendergast, K.H., 1959,
Astrophys. J., 130, 26.
- Jones, T.W., and Merrill, K.M., 1976,
Astrophys. J., 209, 509.
- Osterbrock, D.E., Miller, J.S., and Weedman, D.W., 1966,
Astrophys. J., 145, 697.
- Powell, M.J.D., 1964, Comput. J., 7, 155.
- Powell, M.J.D., 1965, Comput. J., 7, 303.
- Wilson, D.C., 1950, Astrophys. J., III, 279.

Table 1.

Results on U Ori from IRAS compared with a two - temperature model.

Photometric band: Effective wavelength in micrometer	Flux in Janskys from IRAS	Model* flux in Janskys
12	635	635
25	273	273
60	42	43.1
100	14	10.2

* The Physical Parameters determined by the Model are as follows:

- (i) $T_* = 2900 \text{ K}$; $a = 1.5 \times 10^{-14}$ steradians.
- (ii) $T_d = 80 \text{ K}$; $b = 8.2 \times 10^{-11}$ steradians.
- (iii) Silicate grains of size 1 micrometer are assumed to constitute the dust shell $\alpha = 2$.

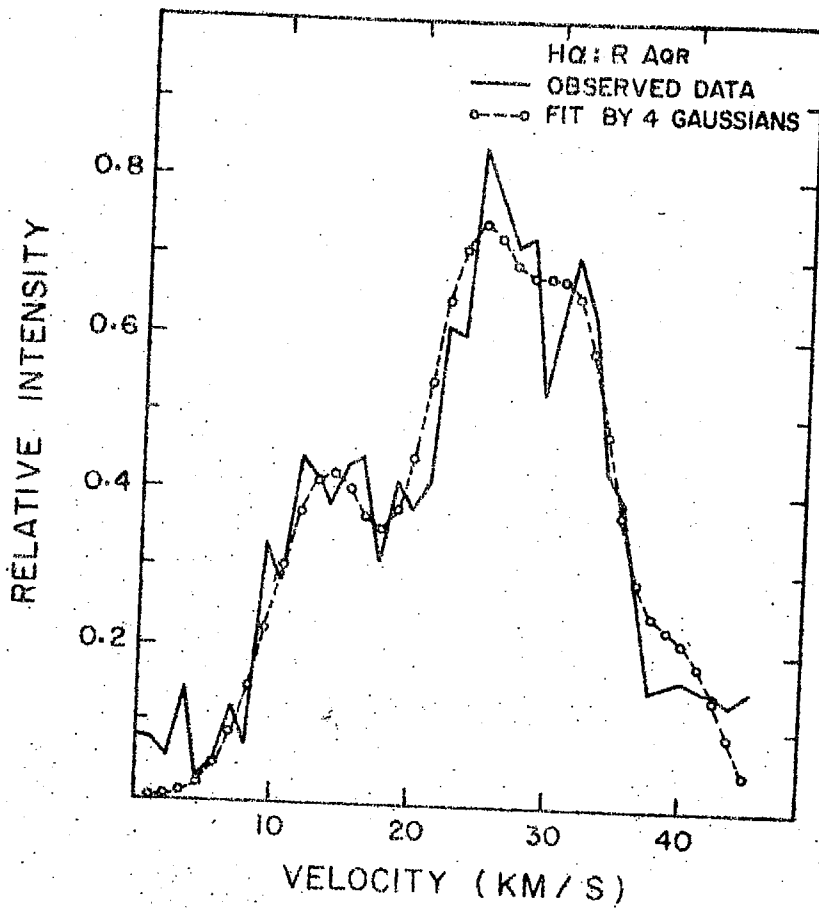


Figure 1 : H α line profile from R Aquarii; observed data fitted with 4 Gaussians using the method described in the text (after, Anandareo et al, 1985).