Extended DISPersion-Absorption (eDISPA) approach to automatic phasing of HR-NMR Spectra

Carlos Cobas [1], Stanislav Sýkora [2]

[1] Mestrelab Research, Xosé Pasín 6, Santiago de Compostela, 15706 Spain; carlos@mestrec.com
[2] Extra Byte, Via Raffaello Sanzio 22/C, Castano Primo (Mi), Italy I-20022; sykora@ebyte.it

INTRODUCTION

We present an extension of DISPA plots in which the in-phase component of a spectrum is plotted against its out-of-phase component, disregarding the frequency coordinate. Starting from such plots, it is possible to derive real-valued functionals $Q(\varphi_0,\varphi_1)\{S\}$ of the spectra S which, when evaluated as functions of the two phase correction angles, exhibit a maximum at the correct phase-correction values. Standard optimisation of $Q(\varphi_0,\varphi_1)\{S\}$ amounts to a novel automatic phase correction algorithm which is completely objective and void of any a-priori discrimination between experimental data points. Being of a particular integral transform type, it is also surprisingly insensitive to the experimental noise (S/N ratio) and to baseline imperfections, allowing reliable and objective automatic phasing of spectra which would defy manual procedures.

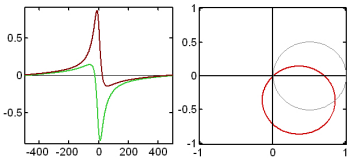
STANDARD PHASING PROCEDURES

The majority of phasing procedures, both manual or automatic, require the selection of two sufficiently distant spectral peaks/multiplets centered in relatively narrow spectral windows Despite the use of a variety of algorithms, all such procedures in essence estimate the complex phase of these two selected signals or signal groups and extrapolate them linearly over the whole spectrum. Because they use a limited subset of available data points, these methods are overly sensitive to limited digital resolution, experimental noise, and spectral artifacts. They are best suited for manual phasing since an experienced operator is often capable to partially compensate their drawbacks. What we are looking for, however, is a stable procedure based on noise-insensitive spectral integrals and suitable for automated black-box implementation.

Classical DISPA

DISPA (Dispersion-Absorption) plots were introduced in 1978 and employed to analyze the shapes and phases of NMR and ESR spectral lines. They show each complex data point of a spectrum in the Cartesian plane of complex numbers, disregarding its frequency offset.

Figure 1. Classical DISPA plots



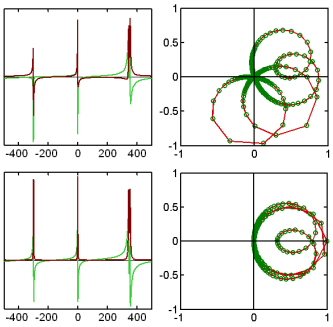
Given a complex Lorentzian line (left; real part is red, imaginary part green), its DISPA plot (right, red trace) is always a circle, whose orientation makes it possible to read out the line's phase angle (in this case, -45º). If the line were properly shaped, its DISPA plot would coincide with the thin gray circle.

EXTENSION to HR-NMR SPECTRA

A HR-NMR spectrum hardly ever contains just one line, but that is no obstacle to plotting the data points in the Cartesian complex numbers diagram. Fig.2 shows simulated examples of such extended DISPA (eDISPA) plots. What is interesting is that while variations in the receiver phase (better known as φ_0) simply rotate the pattern, any linear phase variation proportional to frequency offset (φ_1) results in a proportional azimuthal spread of the 'lobes' corresponding to individual spectral lines. A phased spectrum with $\varphi_0 = \varphi_1 = 0$ exhibits the most closely packed eDISPA plot, with all the lobes pointing to the right. This holds even when the spectral lines are not exactly Lorentzian; in this case the lobes are not perfectly circular, but the statement is still valid.

Note: from now on, we will not use Φ_1 but the closely related, and physically more justified, dimensionless ratio $\tau = t_0/D_w$, where D_w is the *dwell time* between consecutive FID data points and t_0 is the *equivalent* delay between the true start of the FID and the first data point (including filter effects). The φ_1 variation over the the whole spectral width is then equal to $2\pi\tau$. Furthermore, we will define Φ_0 as the complex phase of any data points located at the carrier offset (normally the center of the spectral window).

Figure 2. eDISPA plots of multi-line spectra



The normalized eDISPA plots on the right correspond to the composite-line spectra S on the left. The unphased top spectrum has $\varphi_0 = 45^o$ and $\tau = 0.5$. The bottom spectrum contains exactly the same spectral lines but is perfectly phased, with $\varphi_0 = 0$ and $\tau = 0$.

Notice the DISPA curve of the partially resolved doublet (two 5 Hz wide lines set 10 Hz apart) and the scarce digitization of the two lobes corresponding to the two sharper (2 Hz) singlets (the digital resolution of these simulated spectra is 0.5 Hz).

When the two phasing parameters φ_0 and τ vary, each point in the eDISPA plot undergoes a rotation around the origin. While φ_0 rotates the whole figure, increasing τ causes a progressive spread of the individual 'lobes'. For large values of τ, the spread can exceed 360º, giving rise to a false, partial 'refocusing' of the individual lobes.

The Q-FACTOR FUNCTIONAL $q(\varphi,\tau)\{S\}$

Exploiting the insight gained with eDISPA, we have designed a real-valued functional, the quality factor $\mathbf{q}(\varphi,\tau)$, which acts on a spectrum S corrected by the given phasing parameters φ and τ, and attains a maximum in close vicinity of the correct phasing values. To develop the functional, we have tested parameterized families of functionals on a large number of randomly simulated spectra. Here we will avoid the details of this process and concentrate on the final formula, comment its parameters and illustrate its performance.

Suppose that the spectrum S is an array of N complex values y_k, k = 1,2,…,N and denote as $Y_k = R_k + jI_k$ the corresponding values after a normalization and phase correction with parameters (φ,τ). Explicitly,

$$Y_k = (y_k/y_{max}) * \exp[-j(\varphi+2\pi\tau(k-N)/N)],$$

with φ expressed in radians and y_{max} the maximum of all $|y_k|$ (notice that phase correction does not change the modulus of any of the data points). The application of the functional $\mathbf{q}(\varphi,\tau)$ to S is then caried out according to the **integral-type formula**

$$\mathbf{q} \equiv \mathbf{q}(\varphi,\tau)\{S\} = \sum_{k'} |Y_{k'}|^a * R_{k'}^b * \exp(-w(2k-N)/N),$$

where the summation index k' extends only over those points for which $|y_{k'}|/y_{max} \geq c$.

The parameters a, b, c and w specify the type of the q-factor functional. When b = 1 and a = c = w = 0, the formula reduces to the plain sum (integral) of the real parts of all data points – a functional whose optimization sometimes gives reasonable values for φ and τ, but fails just as often.

A major step towards better reliability consists in increasing the value of a. This amounts **to weighing each point with a power of its own modulus** and thus reducing sensitivity to baseline distortions and, more importantly, giving the EDISPA lobes a more elongated, ellipsoidal shape which sharpens the maximum of **q**. After empirical testing, we have settled for **a = 2**.

The sensitivity to baseline distortions (due in part to known FFT artifacts associated with t_0) is further suppressed by setting the **relative intensity threshold** c to a value such as **c = 0.1** (lower if the spectrum contains only one intense and narrow multiplet, plus some very weak lines). This has also a beneficial effect on reliability (though not as marked as that of parameter a), and it dramatically improves computing efficiency since only a fraction of data points needs to be treated.

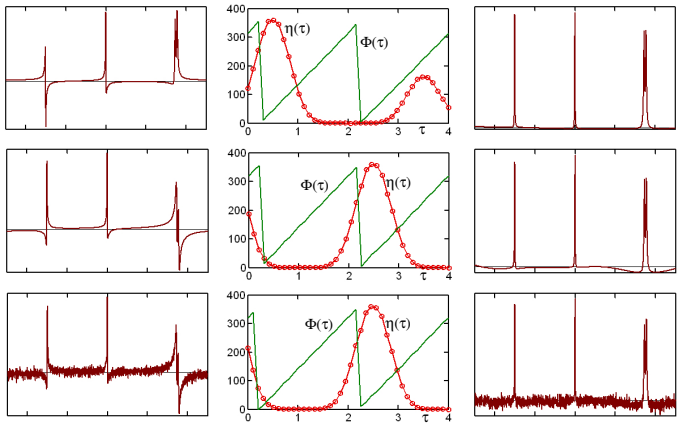
The **offset-weighing coefficient** w suppresses false maxima of **q** in cases of large τ values. By giving a larger weight to the central portions of the spectrum, it makes it possible to resolve situations when the lines on both extremes of the spectrum are nearly in phase, while those in the center are mis-phased. The fact that w attenuates the spectral wings is not deleterious, since noise is attenuated as well. We have settled for **w = 2**.

The eDISPA ALGORITHM

To better explain the proposed eDISPA q-factor method, consider again the unphased spectrum of Figure 2. In Figure 3 we plot two functions $\eta(\tau)$ and $\Phi(\tau)$ obtained in this way: for each τ, we compute $\mathbf{q}(\varphi,\tau)S$ for an array of values φ (since the variations are quite smooth, 5º or 10º steps are quite sufficient). From these, we estimate the location $\Phi(\tau)$ of the maximum of $\mathbf{q}(\varphi,\tau)\{S\}$ with respect to φ and denote its value as $Q(\tau)$. This defines two new functions, and $Q(\tau)$ and $\Phi(\tau)$. To facilitate human inspection, the shown function $\eta(\tau)$ is a normalized 4*th*-power of the variation index of $Q(\tau)$, i.e., $\eta(\tau) = \kappa([Q(\tau)-\min(Q(\tau))]/[\max(Q(\tau))-\min(Q(\tau))])^4$, with the coefficient κ chosen to be 360, in order to make the vertical range of $\eta(\tau)$ compatible with that of $\Phi(\tau)$.

Given an experimental spectrum S, once the functions $\eta(\tau)$ and $\Phi(\tau)$ have been computed, the optimal phasing parameter τ_{opt} corresponds to the absolute maximum of $\eta(\tau)$ and $\varphi_{opt} = \Phi(\tau_{opt})$.

Figure 3. Automatic phasing via the eDISPA q-factor



On the left are simulated unphased spectra, in the center are the graphs of $\eta(\tau)$ and $\Phi(\tau)$ described above, and on the right are the spectra phased with the computed optimal parameters $(\varphi_{opt},\tau_{opt})$. The **top row** spectrum is the same as in Fig.2 (top), generated with $\tau = 0.5$ and $\varphi = 45^o$. Notice the false local maximum of $\eta(\tau)$, due to the fact that the spectrum has so few lines. The **middle row** shows the same spectrum simulated with $\tau = 2.5$ and $\varphi = 50^o$. The final result is again correct, even though even an experienced operator might get tempted to search the solution at much lower τ values (particularly considering the presence of the t_0-related baseline artifact). Finally, the **bottom row** shows the same situation, but with the addition of a considerable amount of experimental noise to which the method is clearly quite resistant, undoubtedly due to the fact that it is based on spectral integrals rather than any local features.

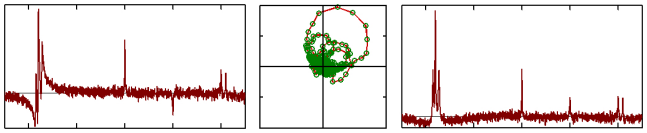
SEARCH FOR MAXIMUM and UNIQUENESS

Considering the smoothness of the function $\mathbf{q}(\varphi,\tau)\{S\}$, the search for its maxima does not represent any problem. It is sufficient to evaluate it on a discrete grid, varying φ by steps of 10º from 0 to 360º values and τ by steps of 0.1 from -1 to 3 (higher values of τ indicate unusual acquisition settings which should be known; in such cases, it is reccomandable to carry out a simple preliminary correction before applying eDISPA). Thereafter, intermediate values of $\mathbf{q}(\varphi,\tau)\{S\}$ and its maxima, can be found quite accurately by quadratic interpolation.

It may happen that $\mathbf{q}(\varphi,\tau)\{S\}$ has two or three maxima of comparable magnitude. This usually indicates the actual presence of multiple acceptable solutions. In such cases, the possible solutions might be subject to a human operator to allow for preferences based on 'external' considerations (such as the employed acquisition delay).

eDISPA PHASING BLACK-BOX

Figure 4. The eDISPA algorithm has been made into a black-box phasing procedure (spectrum-in/spectrum-out). We show here just one result of its application to a rather difficult case (unbalanced line distribution, $\tau = 2.3$, very bad baseline, bad S/N ratio). Note that in this case, the eDISPA plot looks like it might have a τ of almost 0 (wrong) and φ of about 60º (correct). Yet our algorithm is no fool … (thanks, in this case, to the w coefficient).



CONCLUSIONS

The integral eDISPA automatic phasing method is very new and can be certainly still refined. Though it is not foolproof, it already works remarkably well – better, we believe, than most experienced human operators. Its success is undoubtedly due to the combination of physical insight implicit in the extended DISPA plots with an integral-type evaluation procedure which confers it a remarkable insensitivity to experimental noise.

The method is now being incorporated into the **MestreNova** software package and further improvements are expected once a sufficiently large volume of practical experiences gets accumulated.

REFERENCES

This poster is permanently available in *Stan's Library* (www.ebyte.it/library/Library.html), including a complete list of titled references.

A.G.Marshall, D.C.Roe, *Anal.Chem.* **1978**, *50*, 756
D.C.Roe, A.G.Marshall, et al, *Anal.Chem.* **1978**, *50*, 764
A.G.Marshall, *J.Phys.Chem.* **1979**, *83*, 521
A.G.Marshall, D.C.Roe, *J.Magn.Reson.* **1979**, *33*, 551
F.G.Herring, A.G.Marshall, et al, *J.Magn.Reson.* **1980**,*37*,293
A.G.Marshall, R.E.Bruce, *J.Magn.Reson.* **1980**, *39*, 47
T-C.L.Wang, A.G.Marshall, *Anal.Chem.* **1983**, *55*, 2348
P.S.Phillips, F.G.Herring, *J.Magn.Reson.* **1984**, *57*, 43
C.H.Sotak, C.L.Dumoulin, et al, *J. Magn. Reson.* **1984**, *57*, 453
F.G.Herring, P.S.Phillips, *J.Magn.Reson.* **1984**, *59*, 489
E. C.Craig, A.G.Marshall, *J. Magn. Reson.* **1986**, *68*, 283
M.Miglierini, T.Liptaj, R.Palov, *Czech J.Phys.* **1986**, *36*, 729
E. C. Craig, A.G.Marshall, *J. Magn. Reson.* **1988**, *76*, 458
E.A.Wachter, E.Y. Sidky, T.C.Farrar, *J. Magn. Reson.* **1989**, *82*, 352
R.E.Hoffman, F.Delaglio,G.C.Levy, *J. Magn. Reson.* **1992**, *98*, 231
P. Koehl, C.Ling, J.F.Lefèvre, *J. Chim. Phys.*, **1995**, *92*, 1929