RESEARCH ARTICLE

Phase Correction of Fourier Transform Ion Cyclotron Resonance Mass Spectra Using MatLab

Yulin Qi,[1,3] Christopher J. Thompson,[2] Steve L. Van Orden,[2] Peter B. O'Connor[1]

[1]Department of Chemistry, University of Warwick, Coventry, UK
[2]Bruker Daltonics, 40 Manning Road, Billerica, MA, 01821, USA
[3]Central Laboratory, Shanghai Xuhui Central Hospital, Shanghai, People's Republic of China

Abstract

FT-ICR mass spectrometry has been limited to magnitude mode for almost 40 years due to the data processing methods used. However, it is well known that phase correction of the data can theoretically produce an absorption-mode spectrum with a mass-resolving power that is as much as twice as high as conventional magnitude mode, and that it also improves the quality of the peak shape. Temporally dispersed frequency sweep excitation followed by a time delay before detection results in a steep quadratic variation in the signal phase with frequency. Viewing this, it is possible to find the correct phase function by performing a quadratic least squares fit, modified by iterating through phase cycles until the correct quadratic function is found. Here, we present a robust manual method to rotate these signals mathematically and generate a "phased" absorption-mode spectrum. The method can, in principle, be automated. Baseline correction is also included to eliminate the accompanying baseline drift. The resulting experimental FT-ICR absorption-mode spectra exhibit a resolving power that is at least 50% higher than that of the magnitude mode.

Key words: FTICR, Magnitude-mode, Absorption-mode, Frequency sweep excitation, Phase, Mass resolving power

Introduction

Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR MS) [1, 2] is recognized as being a useful analytical instrument because it provides the highest broadband mass spectral resolving power and mass accuracy currently available [3, 4]. For example, in order to derive the unique elemental composition from the monoisotopic mass for masses >500 Da, one needs accuracy at the sub-ppm level or better [5, 6], which can currently only be achieved by FT-ICR MS. Although it is well known that the mass-resolving power of FT-ICR varies linearly with the applied magnetic field strength and inversely with *m/z* and pressure [7, 8], it is less widely appreciated that the mass-resolving power of FT-ICR can be enhanced by up to a factor of 2 by phasing the magnitude-mode spectrum into an absorption-mode spectrum [9, 10].

Absorption and Dispersion Spectra from Complex Spectra

A digitized time-domain ion cyclotron resonance signal will yield two frequency-domain spectra after Fourier transformation [11, 12], Re(ω) and Im(ω), which correspond to

the real and imaginary parts of the complex spectrum [13]. The general relation is given below:

$$F(\omega) = \int F(t)e^{-i\omega t}dt = A(\omega) + iD(\omega) \qquad (1a)$$

$$\phi(\omega) = \arctan[D(\omega)/A(\omega)] \qquad (1b)$$

$$M(\omega) = [(A(\omega))^2 + (D(\omega))^2]^{\frac{1}{2}} \qquad (1c)$$

Here, $\omega = 2\pi f$ is the angular frequency, $F(\omega)$ represents the frequency domain data, while $A(\omega) = \mathrm{Re}(\omega)$ and $D(\omega) = \mathrm{Im}(\omega)$ are the absorption-mode and dispersion-mode spectra, respectively. $\Phi(\omega)$ and $M(\omega)$ are the phase and magnitude-mode spectra, respectively.

In general, ions are excited to their cyclotron orbits with a phase that is generated by the frequency program of the excitation pulse, and they precess continuously until detection. If this phase function, $\Phi(\omega)$, is known, $F(\omega)$ can be rotated in the complex space to align all signals along the real axis. However, this phase function has generally not been known, so the magnitude-mode spectrum has been used instead.

If the spectrum is misphased by an angle (see Fig. S1, top) $\Phi(\omega)$, which is a function of frequency and time, the corrected spectrum, $G(\omega)$, can be generated by

$$G(\omega) = \int F(t)e^{i\omega t}e^{i\phi(\omega)}dt = \int F(t)e^{i(\omega t + \phi(\omega))}dt. \qquad (2)$$

If $\Phi(\omega)$ is correct, $G(\omega)$ is aligned with the real axis of the complex space, and a clean, properly phased absorption-mode spectrum is available. Thus, the trick is to find the function $\Phi(\omega)$ that encodes the phase for each signal in the entire spectrum. However, it is important to note that any peak can be perfectly phased at some $\Phi(\omega_i)$ between 0 and 2π, but that $\Phi(\omega_i) = \Phi(\omega_i) + 2n_i\pi$ for any integer n_i at any ω_i. In order to generate the useful $\Phi(\omega)$, it is necessary to find the full function, which means determining the individual values of n_i and $\Phi(\omega_i)$.

The Advantages of Absorption Mode Over Magnitude Mode

Even when the very first FT-ICR mass spectra was obtained in 1974 [14, 15], it was readily recognized that an absorption-mode spectrum provides higher resolving power than a magnitude-mode peak [16]. Also, phasing is done routinely in Fourier transform nuclear magnetic resonance (FT-NMR, see Fig. S1 of the Electronic Supplementary Material, ESM) [17, 18].

The absorption-mode spectrum is preferable to the magnitude-mode spectrum for several reasons. First, the peak width at half-maximum height is narrower than its corresponding magnitude mode by a factor of 1.4–2× (Fig. 1) [19–21]. Which means that the mass-resolving power, $m/\Delta m$(FWHM), is enhanced by the same factor. Second, because the height of a perfectly phased absorption-mode spectrum is identical to its corresponding magnitude-mode peak, increasing the resolution will not affect the signal-to-noise ratio of the peak. Other methods that are used to smooth the peak shape, like apodization [22, 23], can be accomplished only at the expense of degradation of the resolution or signal-to-noise ratio. Third, foldover peaks, also known as aliased peaks [24], which are caused by sampling at less than the Nyquist frequency, can easily be detected by their anomalous phase in an absorption-mode spectrum. By contrast, identifying foldover in a magnitude-mode spectrum requires a comparison of the frequencies of the same peak in two spectra acquired at different sampling rates. Fourth, the peaks in absorption-mode spectra have a different shape, and they are very well separated at the baseline, whereas in magnitude mode, overlaps between neighboring peaks can often result in unusual signal shapes at the bases of peaks (see below).

Spectral Phase and Frequency-Sweep Excitation

In Fourier transform nuclear magnetic resonance (FT-NMR) [25, 26] spectroscopy, signals at all frequencies are excited near-simultaneously by a short RF pulse. This can be done because an entire FT-NMR spectrum typically covers a bandwidth of a few kilohertz on a several hundred
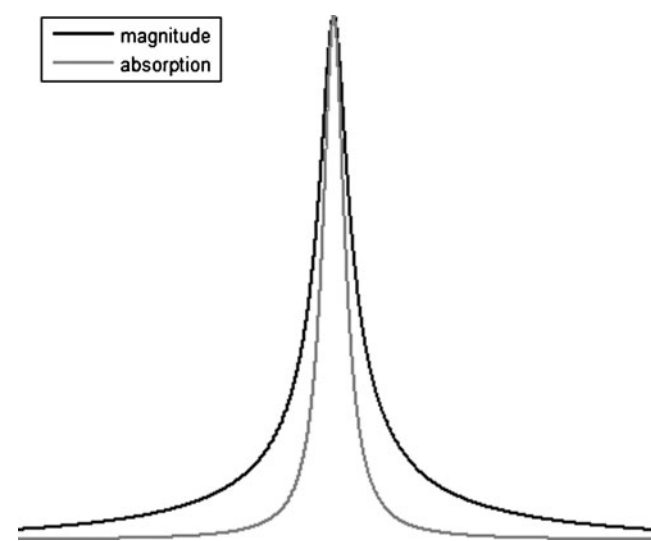

Figure 1. Simulated magnitude-mode and absorption-mode Lorentzian peakshapes, corresponding to the Fourier transformation of a "high-pressure" limited ICR signal that damped nearly to zero during the data acquisition period

megahertz signal [27]. Thus, in NMR, this short pulse essentially simultaneously excites all of the spins, and the phase angle can usually be represented by a linear equation that is similar to the last two terms in Eq. 3 (below), so that for NMR, $A=0$ in Eq. 3 [28, 29]. Furthermore, the phase variation in FT-NMR is usually less than one cycle ($<2\pi$) across the spectrum [30, 31], so phasing becomes relatively simple for NMR: choose a peak at one end of the spectrum, vary C until the absorption-mode spectrum is equal in height to the magnitude-mode spectrum, and then vary B using the same equation for a peak at the other end of the spectrum. The peaks between the two points will then all exhibit the absorption-mode shape, as shown in Fig. S1 of the ESM [19].

$$\Phi(\omega) = A\omega^2 + B\omega + C. \tag{3}$$

For FT-ICR, it was recently demonstrated that, because of linear frequency sweep excitation, $\Phi(\omega)$ is a quadratic function [32, 33]. Recovery of the absorption-mode spectrum from the complex spectrum requires the correct phase angle function, $\Phi(\omega)$; see Eq. 3.

This phase, $\Phi(\omega)$, has several contributions arising from [34]: (a) a zero-order phase shift ($C \neq 0$) due to phase accumulation at the beginning of excitation; (b) a linear shift ($B \neq 0$) due to the time delay between excitation and detection; and (c) a quadratic phase shift ($A \neq 0$) for the frequency-sweep excitation [35] or stored-waveform inverse Fourier transform (SWIFT) [36].

In FT-ICR, the frequency bandwidth is much larger than it is in FT-NMR. Ions are excited by frequency-sweep excitation [35], so the phase $\Phi(\omega)$ increases quadratically with frequency (see Fig. 2), and various electronic circuit components (e.g., filters) can also introduce additional nonlinear phase variation. Moreover, $\Phi(\omega)$ can vary by 2π or more over even a single nominal m/z because $\sin[\Phi + 2n\pi] = \sin(\Phi)$, which is known as "phase wrapping" [13]. This effect makes it relatively difficult (compared to FT-NMR) to determine the variation of $\Phi(\omega)$ with frequency across a broadband spectrum. So, strictly speaking, the phase $\Phi(\omega)$ of a spectral peak can be determined only within $0–2\pi$, so Eq. 3 should be replaced by

$$\Phi(\omega_i) = A\omega_i^2 + B\omega_i + C + 2\pi n_i. \tag{4}$$

Here, $0 \leq C \leq 2\pi$, and n_i is the number of phase cycles for a peak at ω_i. Because of this complication, FT-ICR mass spectra have generally been limited to magnitude mode. However, because any peak can be "phased" iteratively, a set of phases, Φ_i, can be generated for a matched set of peaks in the frequency domain. Thus, finding the correct $\Phi(\omega)$ for all values of ω_i becomes a quadratic least squares fitting problem where the values of n_i can take on a range of specific integer values. The range of these n_i values can be
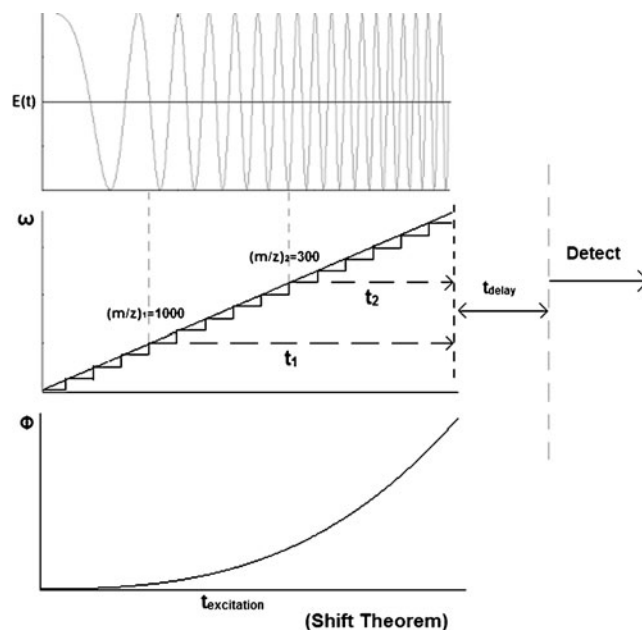

Figure 2. *Top*: time-domain profiles of the excitation for frequency-sweep excitation. *Middle*: frequency for a continuous (*upper line*) or stepwise (*lower line*) frequency sweep. *Bottom*: quadratic phase accumulation for the stepwise frequency sweep

readily approximated from the duration of excitation and the excite–detect delay time. It is then straightforward to iterate through the expected range of n_i values, phase correct $G(\omega)$ at each iteration (as in Eq. 2), and check the absorption-mode spectrum for symmetry around peaks and/or maximize the positive signal amplitude.

In this paper, a broadband phase-correction method based on MATLAB (MathWorks, Natick, MA, USA) that can phase a spectrum into the absorption mode is demonstrated. The algorithm calculates the optimum value of $\Phi(\omega)$ for each individual data point in the spectrum so that the entire spectrum can be simply phased into the absorption mode. This method has been tested for petroleum mixtures, proteins, and top-down MS/MS spectra from ECD (electron capture dissociation), and yields a much-improved peak shape compared with a magnitude-mode spectrum, particularly at the bases of peaks. Furthermore, many low-intensity peaks that overlap at the baseline can be resolved using this method.

Experimental Methods

Sample Description and Preparation

The petroleum sample is a distillate fraction from crude oil, diluted initially in dichloromethane and then diluted to 500 μg/mL in a 85:15 methanol:toluene mixture and used without additional purification. Ubiquitin (purchased from Sigma-Aldrich Chemical Co., Gillingham, UK) was diluted

to 100 fmol/μL and 200 fmol/μL in a 50:50 water:methanol mixture for the molecular ion charge state distribution and 11^+ ECD spectra, respectively. All solvents were HPLC grade, and obtained from Sigma-Aldrich.

Instrumentation

All spectra were recorded on a SolariX 12T FT-ICR mass spectrometer (Bruker Daltonics, Bremen, Germany) using either electrospray ionization (ESI) or atmospheric pressure photoionization (APPI). Broadband frequency sweep (chirp) excitation (92–938 kHz (m/z 2000–200) for petroleum and 92–1250 kHz (m/z 2000–150) for ubiquitin, both at a sweep rate of 125 Hz/μs) was followed by direct-mode image current detection to yield 4 M data-point (32 bit) time-domain data sets.

Theory

Figure 2 illustrates how the FT-ICR spectral phase varies quadratically with excitation frequency. To a first approximation, each ion with a particular m/z value will be excited at the instant that the excitation frequency matches its own cyclotron resonance frequency during the frequency sweep. Furthermore, it is normal to delay the onset of detection by several milliseconds, because the magnitude of the RF signal excitation V_{p-p} is >100 V, whereas the RF signal for detection is usually at the microvolt level, and the residual "ring down" of the RF signal usually saturates the pre-amplifier for a few milliseconds. In most recent versions of the FT-ICR MS software, t_{delay} is approximately 3 ms, and the definition of t_{delay} has a jitter of nanoseconds at most. While stable on a scan-to-scan basis at the nanosecond scale, this value of t_{delay} varies with different pulse sequences and different instruments (particularly from different manufacturers), and therefore must be accounted for, as it can cause difficulties when calculating the phase. The "shift theory" of Fourier analysis [37] points out that the accumulated phase between the end of excitation and the onset of detection is ωt_{delay}, which usually exceeds 2π. The phase accumulation during the delay period will vary linearly with frequency, as represented by the last two terms in Eq. 3. The first term of Eq. 3 is related to the fact that a linear frequency sweep excitation of the form $\sin(\omega t + at^2)$ (where a is the sweep rate) is also a quadratic phase program. Thus, one can model the accumulation of phase during the frequency sweep as uniformly accelerated linear motion, where the phase increases quadratically with time (similar to the linear distance traveled by an object with constant acceleration). In reality, however, for the Bruker data system, the real excitation is generated in the instrument control software as a series of single-frequency excitation events of identical length. For the examples used here, there were approximately 1500–2000 frequency steps, each of 5 μs length. Note that this algorithm instead assumes a true linear frequency shift, which is a possible source of error; although this is apparently not significant (see below).

Furthermore, it is approximated that an ion with a given m/z is instantaneously excited to its coherent cyclotron orbit when the excitation reaches the resonance frequency [38]. The phase of the ion accumulates quadratically before that point, but after that instant, the ion cyclotron phase will accumulate linearly (e.g., for an ion with an m/z value of 1000 in Fig. 2, the linear accumulation will be $w_1(t_1+t_{delay})$). In conclusion, the total cyclotron phase will be the sum of the quadratic phase accumulation during the frequency sweep and the linear increase between the instant of excitation and the onset of detection, as shown in Eq. 3.

Computational Algorithm

The entire process of phasing the FT-ICR mass spectrum is performed within MatLab, a detailed explanation of the algorithm for the process is given in "Appendix I," and all of the MatLab code used in the process is listed in "Appendix II." The process can be broken down into the following steps:

Step 1: Read the data
Step 2: Generate the frequency basis set and calibrate the m/z domain
Step 3: Calculate the optimal phase for a set of peaks in a narrow range of the spectrum (perhaps an isotopic distribution)
Step 4: Phase that region of the spectrum
Step 5: Iteratively extend the range of the phase function by progressively adding more peaks and recalculating the phase function until the function can phase correct the entire spectrum
Step 6: Baseline correction

Results and Discussion

Absorption-Mode Versus Magnitude-Mode Display

As recently discussed by the Marshall group [39], petroleum samples provide a particularly good test of FT-ICR mass spectrometric performance because they have extraordinarily complex sample compositions over a wide dynamic range. Thus, broadband fast Fourier transform real and phase-corrected absorption-mode ESI FT-ICR mass spectra of crude oil at 12T are shown in Fig. 3. In this case, the mass-resolving power is increased by at least 50% for the absorption mode $(m/\Delta m_{50\%} \sim 850,000)$ compared to its corresponding magnitude mode $(m/\Delta m_{50\%} \sim 500,000)$ (Fig. 3, bottom inset). Several small, negative direction peaks in the absorption-mode spectrum (Fig. 3, bottom) are the errors that arise from the assumption of a true linear frequency shift (see Fig. 2), because the excitation is in fact generated as a series of single-frequency pulses, as mentioned
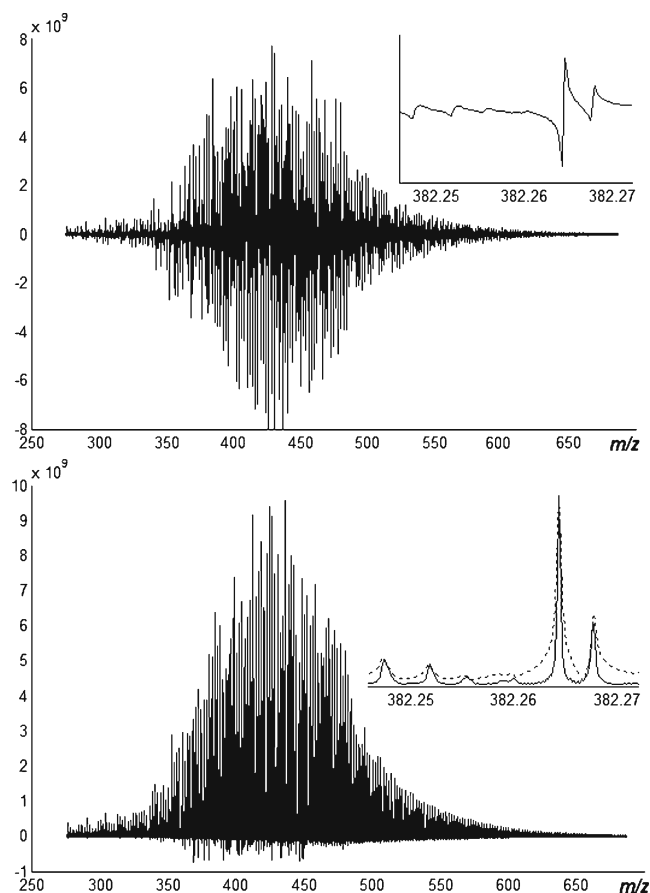
Figure 3. ESI FT-ICR mass spectra of petroleum. *Top*: unphased real data spectrum following the fast Fourier transform. *Bottom*: phase-corrected absorption-mode spectrum; *inset*: magnitude-mode (*dotted line*, $m/\Delta m_{50\%} \sim 500$ k) versus absorption-mode (*solid line*, $m/\Delta m_{50\%} \sim 850$ k) spectra for some peaks extracted from the entire spectrum

before. The use of a stored-waveform inverse Fourier transform (SWIFT) [40] or a more finely grained excitation pulse series should eliminate these.

If we consider Fig. 3, it is particularly interesting to note (Fig. 3, bottom inset) that the magnitude-mode spectrum (dotted line) and the absorption-mode spectrum (solid line) have significantly different peak shapes, which lead to some unusual baseline effects that are commonly observed in FT-ICR mass spectra. Note that the distortion of the magnitude mode depends on the degree of peak overlap, the relative magnitudes of the overlapping peaks, and the signal damping rate [41], as well as the relative phase of the overlapping peaks [42]. In particular, the "tails" of the magnitude mode's Lorentzian function tend to distort the positions of neighboring smaller peaks. Because the "tails" of the dispersion-mode spectrum extend much farther than the "tails" of the pure absorption-mode spectrum, and the magnitude-mode peak is a vector summation of both, neighboring peaks must be farther apart to achieve the same degree of overlap as absorption-mode peaks. Furthermore, the deviation from this summation (see Fig. 3, inset) suggests that the absorption-

mode spectrum should provide better mass accuracy than the magnitude-mode spectrum, over and above the increase in mass accuracy associated with the improved resolution.

Resolving Power

Figure 4 shows the comparison of magnitude-mode and phase-corrected absorption mode tandem mass spectra for the 11^+ charge stage of ubiquitin. Again, the mass-resolving power for the absorption mode was observed to be >50% higher than that of the magnitude mode at half height, and many overlapping isotopic peaks were better resolved in the absorption mode. Also notice that the peaks are separated far
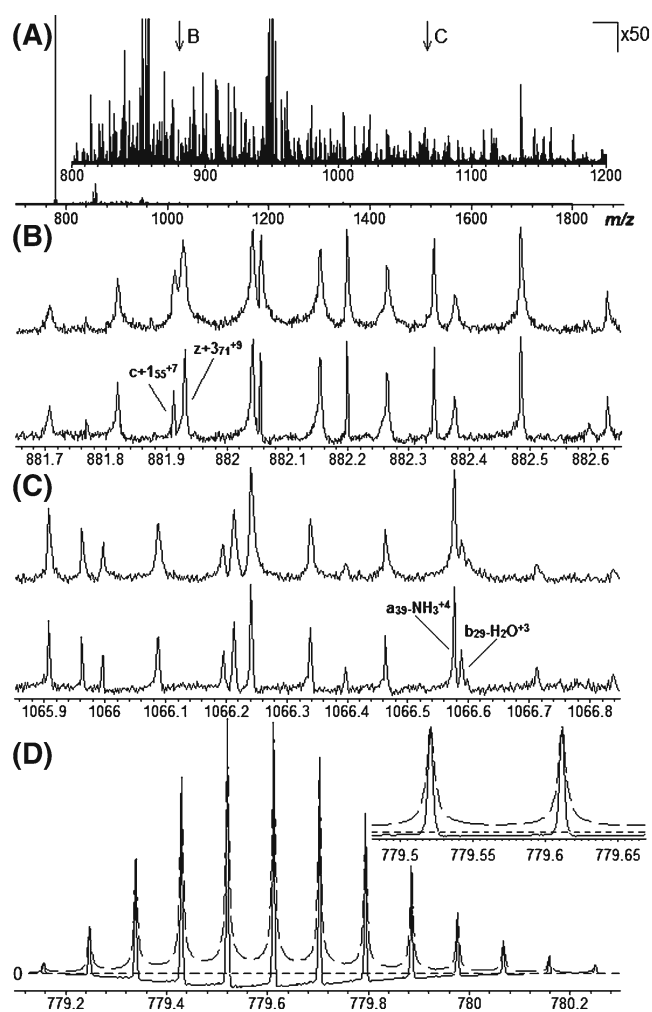

Figure 4. (**A**) Magnitude-mode ECD spectrum for the 11^+ charge stage of ubiquitin, along with a close-up that shows the peak density of the raw data. (**B**) The absorption- (*bottom*) and magnitude- (*top*) mode spectra of the region of (**A**) around *m/z* 882. (**C**) The absorption- (*bottom*) and magnitude- (*top*) mode spectra of the region of (**A**) around *m/z* 1066. In both (**B**) and (**C**), the absorption-mode spectra clearly improve peak shapes and resolution without distorting the intensities. (**D**) The absorption- (*solid*) and magnitude- (*dotted*) mode spectra of the precursor ion isotopic distribution at *m/z* 779–780

better at their bases (Fig. 4D) than in magnitude mode, as in Fig. 3. As is commonly observed in unapodized FT-ICR spectra, the baseline of the magnitude-mode isotopic distribution (Fig. 4, bottom) rises and falls through the peak cluster due to the summation of the Lorentzian "tails" of the peak shape distributions. This effect is not observed in the absorption-mode spectrum, but a slight distortion of the peak shape and a negative baseline drift (due to imperfect phase correction) is observed, which can be corrected (see below). In spite of these peak shape distortions, the absorption mode still generates far sharper spectra.

Additionally, it is reported that the peaks in a phase-corrected spectrum obtained from Fourier deconvolution will exhibit a characteristic leading effect [43]. Such asymmetry also exists in the magnitude-mode spectrum; however, the broader base of the magnitude mode largely hides this effect. The asymmetry may be caused by frequency drift [44] occurring in the time-domain transients at the early stage, perhaps due to evolution of the ion cloud geometry and concomitant space charge effects immediately following excitation. However, upon viewing the spectra shown here (Fig. 4d), it is clear that the peak asymmetry does exist but that it is not as significant as it is in the absorption-mode spectrum obtained by Fourier deconvolution. This suggests that correction by the phase function, $\Phi(\omega)$, may partially modify these error sources, and investigation of the effect is still underway.

Iteration

Figure 5 illustrates the phase iteration procedure. Figure 5a is simply the magnitude-mode spectrum, but Fig. 5b shows the absorption-mode spectrum correctly phased only in a small region around m/z 450, with the rest of the spectrum periodically coming into and out of phase. The point marked by an arrow is misphased by 4π. When this frequency and phase are added to the fitting list and corrected by 2π, and the quadratic least-squares fit is recomputed, the phase function shifts slightly (see the parameters in the inset) to form Fig. 5c. Note that the periodicity of the absorption-mode spectrum has decreased in Fig. 5c,. This iteration procedure is continued until the best absorption-mode spectrum is obtained (Fig. 5d). It is clear from the small negative blips in Fig. 5d that this phase function is not perfect, which is likely due to the stepped frequency sweep excitation and imperfect centroiding of the peaks in the quadratic least-squares fitting list. However, this phase-corrected absorption-mode spectrum represents a significant improvement over the magnitude-mode spectrum in terms of resolution and peak shape.

Centroiding

One significant source of error that is currently embedded in these calculations is related to the method used to calculate
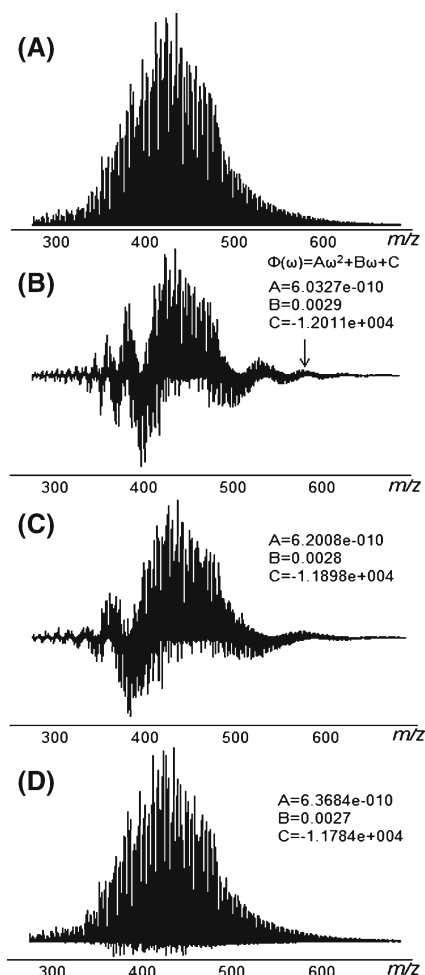

Figure 5. (**A**) The magnitude-mode spectrum of an oil sample. (**B**) The absorption-mode, phase-wrapped spectrum of petroleum (four cycles are apparent in the periodic phasing throughout the entire spectrum, indicating slight errors with the phase function shown in the *inset*). The labeled point (*arrow*) is misphased by 4π, so this frequency and phase pair are added to the fitting list with the phase adjusted by adding 4π. (**C**), (**D**) show further iterations, with more peaks added to the fitting list until the best fit (**D**) is obtained

the peak centroid. The method currently used is the simplest method possible—defining the highest point on the peak (in magnitude mode) as the center of the peak. This method is not particularly accurate, but it improves dramatically with zero filling, so the spectra shown in Figs. 3–5 used two zero fills, thus doubling the length of time-domain data set twice. While not currently tested, it seems likely that improvements can be made to the peak centroiding algorithm, which will improve the accuracy of the Φ/ω pairs that are added to the least squares fit. Possible ways of improving centroiding algorithms include the use of a maximum likelihood peakfitting method [45], a center-of-mass calculation, a least-squares fit to a Lorentzian (or Lorentzian/sinc) function, apodization and fitting to a Gaussian function, or even use of the shifted-basis method [46]. In this work, none of

these centroiding algorithms were tried, but they are an obvious next step.

Baseline Correction

The baseline in the magnitude-mode spectrum is obviously positive because the absolute value is calculated in magnitude mode. In absorption mode, the noise is a Gaussian distribution about the baseline (zero amplitude). Furthermore, the cyclotron excitation and detection range exceeds the sampling rate, which will generate a high-frequency component. Fortunately, some of these peak-shape distortions can be corrected by standard baseline correction procedures. In MatLab, this can be done automatically using a command (see "Appendices I and II") which fits the baseline to a cubic spline and then subtracts it from the signal to flatten the baseline drift. The absorption-mode spectrum after baseline correction shows an improvement in peak shape, which would improve peak selection.

Algorithm Speed

A detailed analysis of the speed of this method is beyond the scope of this paper. Currently, this method is user interactive, requiring manual peak selection and phase assignment, so it is not very fast. However, peak selection and centroiding can be standardized and automated, as can phase determination of individual peaks, so the bulk of the algorithm could be automated. Furthermore, since the excitation pulse does not change from scan to scan, it is likely that an optimal phase function from one scan can be applied to the next scan directly. Inter-scan errors will probably arise from space-charge effects, but these are expected to be minor and can be addressed using the same iterative method as above. Given the speed of modern Fourier transform methods, it is likely that these iterations will be far faster than the data acquisition event itself, so that, in principle, iterative phasing can be done in real time.

Conclusion

A method of producing a broadband absorption-mode spectrum for several spectra from different samples, based on applying an iterative quadratic least squares fitting procedure, has been demonstrated. The improvement in mass resolving power and mass accuracy afforded by this new method are significant when compared with conventional magnitude-mode spectra, particularly at the bases of peaks. Additionally, peak shape is greatly improved at the baseline, except for a general negative drift, which can be corrected. This method can be applied to top-down spectra and other complex spectra in a straightforward approach.

In theory, the largest resolving power enhancement (a factor of 2) would be achieved for on-line LC FT-ICR MS, for which the time-domain acquisition period can be 1 s or less, meaning that the time-domain signal is not damped significantly. Furthermore, displaying the absorption-mode spectrum will have significant advantages when analyzing complex mass spectra, as demonstrated here for petroleum and the ECD spectrum of ubiquitin, so this approach will probably be applied to biological samples in proteomics, lipidomics and metabolomics.

Acknowledgments

Appendix I

Step 1: Read the data. The Bruker data folder (.m or .d extension) is set as the directory in MatLab. The "fopen" command is used to open the file "fid" for read operation. Then the data are zero filled once during the fast Fourier transform (FFT). The negative-frequency half of the data set was discarded because in the fast Fourier transform (without quadrature detection), a time-domain signal is sampled at N equally spaced intervals over a data acquisition period, and only $N/2$ unique complex frequency-domain data are produced, while the FFT cannot distinguish between positive and negative frequencies, so the result will show both positive and negative frequencies that are symmetrically centered at the zero point (0 Hz). This symmetry adds redundant information. Zero filling is used to capture the absorption-mode spectrum information residing in the (non-zero-filled) dispersion data obtained via the Fourier transform [19].

Step 2: Generate the frequency basis set ($N/2$ values that are equally spaced between zero and half the Nyquist frequency), and calibrate the m/z domain. Most FT-ICR data systems do not save the spectral data as (x, y) pairs because the x value is easy to calculate. The x value for any display mode (m/z, Hz, points) can be calculated from the index of the point and an associated transformation/calibration function. The calibration functions used by different instrumental companies vary [47], but they are usually derived from either the Ledford [48] or Francl [49] equations. The calibration of FT-ICR spectra was recently reviewed by Li-Kang Zhang et al. [47].

All of the calibration parameters needed are stored with the spectrum dataset.

Step 3: Calculate an optimal phase, $\Phi(\omega_i) = \Phi(\omega_i) + 2n_i\pi$, for a set of arbitrarily chosen peaks. Several criteria were tested to estimate the best phase values, and dividing the entire spectrum into several segments with small frequency ranges was found to be the best approach. Choose a segment with a certain frequency range and begin by finding the apices of several peaks. Get the real and imaginary values, and calculate their phase values. An initial segment from the middle of the frequency sweep is recommended in order to reduce the first- and second-order correction errors, and zero filling is also suggested here to improve the apex selection accuracy. The optimal phase value for each peak is determined by manually picking the apex, but this could easily be automated.

Step 4: Phase one region of the spectrum by iterating the successive values of n_i. Because the phase accumulation between two adjacent peaks is generally low (within one cycle), it is easy to estimate the value of n_i for a set of consecutive peaks in a narrow frequency range. After a list of peaks has been generated (no more than 20 points), a quadratic least squares fit is applied to the angular frequency and phase value pairs of the above apices in order to generate $\Phi(\omega)$ according to Eq. 3.

Step 5: Extend the range of the phase function. The phase function from step 4 is only a reasonable approximation within the narrow frequency range chosen. When applied to the whole spectrum, it will generate a "phase-wrapping effect" (Fig. 5). This indicates that some of the values of n_i from Eq. 4 are incorrect. Fortunately, it is possible to estimate the number of cycles from the phase-wrapping spectrum. For example, in Fig. 5b, there is a shift of about two cycles for the marked point, so the phase wrapping is within 4π. To correct the phase function, one peak from the low-frequency side is chosen (because a low frequency results in a small phase value, thus minimizing the error), and its corresponding phase value is acquired from the function in step 4. The correct phase value of that point is calculated by iteratively adding $2n\pi$ (n is the phase cycle at that point), as shown in Fig. 5. At each value of n, a new phase function is generated by quadratic least squares fitting to the (ω, Φ) list, and these n values are iterated until the spectrum is fully absorption mode in this region. This procedure is continued by successively adding more peaks until the whole FT-ICR spectrum can be phase corrected simultaneously (Fig. 5d).

Step 6: The baseline of the peaks is corrected automatically using the MatLab command "msbackadj," which estimates the lower-frequency baseline roll within multiple shifted windows using a spline approximation, and then subtracts it.

Appendix II: MatLab Code for Phasing with Full Comments (Start with %)

Step 1: Read the data

```
file = fopen('fid','r');
% open the file 'fid' for read operation.
fid = fread(file,'int32');
% read all data from fid as 32-bit integers
fclose(file);
% close the file, or you can't read any more data from it.


spec = fft(fid,length(fid)*2);
% Fourier transform with one zero fill
spec = spec(1:length(spec)/2);
```
% discard the data in the negative half of the data, because FFT generates even symmetry data around the zero frequency point. This symmetry adds redundant information.

Step 2: Get the frequency, ω, and calibrate the m/z domain

% calibration the m/z domain data follows the calibration equation as approximate. Bruker uses a proprietary calibration function to generate the frequency and m/z domain corresponding to the intensities of the signal for further processing.

Step 3: Calculate an optimal phase, $\Phi(\omega_i) = \Phi(\omega_i) + 2n_i\pi$ for a set of arbitrary chosen peaks

% find the apex of the chosen peaks, then calculate the phase value by Eq.1b, thus the phase is arctangent($Im(\omega)/Re(\omega)$).
```
a = abs(spec(N));
r = real(spec(N));
j = imag(spec(N));
```
% N is the index of assigned peak apex in 'spec' file, a: absolute value of that point, r: real value and j: imaginary value.
% calculate the phase of this point
```
pN = atan(j/r);
```
% then iteratively calculate the phase values for the apex of adjacent peak, and $\Phi(\omega)$ varies within one cycle for consecutive peaks.

Step 4: Least squares approximation of the quadratic in a small frequency domain

```
x = [w1 w2 w3...];
y = [p1 p2 p3...];
```

%y is phase values, pN, acquired in step 3, x is their corresponding angular frequency (frequency*2π) which is obtained by calibration from step 2.

para = polyfit(x,y,2);

%quadratic least square fit of (x, y) pairs for a quadratical function

ybest = polyval(para,x);

%ybest is the best fit for the function we get from the regression

plot(x,ybest,x,y,'o'), title('quadratic regression estimate'), grid

%plot the regression curve with the raw (x, y) pairs in spot.

A = para(1),B = para(2),C = para(3)

%output the parameters A, B and C for the phase function

%calculate the initial Φ(ω) function

RightPhase = A.*w.*w + B.*w + C;

% calculate the phase value for the corresponding frequency-domain

PhasedSpec = spec.*exp(-i.* RightPhase);

% correct the phase by rotating the raw data mathematically

plot(mz,abs(spec),mz,real(PhasedSpec))

% plot the magnitude-mode versus phased absorption-mode

Step 5: Extend the range of the phase function

% add additional apex outside the frequency domain in step 4 to the fit, the cycle between two phase values can be estimated as shown in Fig. 5, go to step 4, and more frequency/phase pairs can be added continuously to increase accuracy, thus, phase the whole spectrum

Step 6: Baseline correction

baseline = msbackadj(index,real(PhasedSpec),'STEPSI-ZE',201,'WINDOWSIZE',50,'QUANTILEVALUE',0.01,'-SHOWPLOT',1);

%the parameters 'STEPSIZE', 'WINDOWSIZE' and 'QUAN-TILEVALUE' are depending on the size of the data (see MatLab help file)

figure

plot(m,abs(spec),m,real(PhasedSpec),m,baseline)

%plot the baseline in the phase-corrected spectrum

References

1. Marshall, A.G., Hendrickson, C.L., Jackson, G.S.: Fourier transform ion cyclotron resonance mass spectrometry—a primer. *Mass Spectrom. Rev.* **17**, 1–35 (1998)
2. Holliman, C.L., Rempel, D.L., Gross, M.L.: Detection of high mass-to-charge ions by Fourier transform mass spectrometry. *Mass Spectrom. Rev.* **13**, 105–132 (1994)
3. Amster, I.J.: Fourier transform mass spectrometry. *J. Mass Spectrom.* **1996**(31), 1325–1337 (1996)
4. McLafferty, F.W., Senko, M.W., Little, D.P., Wood, T.D., O'Connor, P. B., Speir, J.P., Chorush, R.A., Kelleher, N.L.: Electrospray ionization and Fourier transfrom mass spectrometry in biomedical research. *Adv. Mass Spectrom.* **1995**(13), 115–128 (1995)
5. Dykes, S., Fancy, S.A., Perkins, G.L., Pullen, F.S.: The automation of a commercial Fourier transform mass spectrometer to provide a quick and robust method for determining exact mass for the synthetic chemist. *Eur J Mass Spectrom* (Chichester, UK) **9**, 73–80 (2003)
6. Zubarev, R.A., Håkansson, P., Sundqvist, B.: Accuracy requirements for peptide characterization by monoisotopic molecular mass measure-ments. *Anal. Chem.* **68**, 4060–4063 (1996)
7. Schaub, T.M., Hendrickson, C.L., Horning, S., Quinn, J.P., Senko, M. W., Marshall, A.G.: High-performance mass spectrometry: Fourier transform ion cyclotron resonance at 14.5 Tesla. *Anal. Chem.* **80**, 3985–90 (2008)
8. Ledford Jr., E.B., Ghaderi, S., White, R.L., Spencer, R.B., Kulkarni, P. S., Wilkins, C.L., Gross, M.L.: Exact mass measurement by Fourier transform mass spectrometry. *Anal. Chem.* **52**, 463–468 (1980)
9. Comisarow, M.B.: Signal modeling for ion cyclotron resonance. *J. Chem. Phys.* **69**, 4097–4104 (1978)
10. Comisarow, M.B.: Comprehensive theory for ICR power absorption: applications to line shapes for reactive and non-reactive ions. *J. Chem. Phys.* **55**, 205–217 (1971)
11. Krot, A.M.: The method of eigentransforms in different fields for computing cyclic convolutions and discrete Fourier transforms. *USSR. Comput Math. Math Phys.* **29**, 23–34 (1989)
12. Dzhyugis, A.S.: The fast Fourier transform in odd cosines. *USSR. Comput Math. Math Phys.* **29**, 88–90 (1989)
13. Craig, E.C., Santos, I., Marshall, A.G.: Dispersion vs. absorption (DISPA) method for automatic phase correction of Fourier transform ion cyclotron resonance mass spectra. *Rapid Commun. Mass Spectrom.* **1**, 33–37 (1987)
14. Comisarow, M.B., Marshall, A.G.: Fourier transform ion cyclotron resonance spectroscopy. *Chem. Phys. Lett.* **25**, 282–283 (1974)
15. Comisarow, M.B., Marshall, A.G.: Frequency-sweep Fourier transform ion cyclotron resonance spectroscopy. *Chem. Phys. Lett.* **26**, 489–490 (1974)
16. Comisarow, M.B., Marshall, A.G.: Selective-phase ion cyclotron resonance spectroscopy. *Can. J. Chem.* **52**, 1997–1999 (1974)
17. Derome, A.E.: Modern NMR Techniques for Chemistry Research. Pergamon, Oxford (1987)
18. Fukushima, E., Roeder, S.B.W.: Experimental Pulse NMR: A Nuts and Bolts Approach. Addison-Wesley Publishing Co, Reading (1981)
19. Marshall, A.G., Verdun, F.R.: Fourier Transforms in NMR, Optical, and Mass Spectrometry: A User's Handbook. Elsevier, Amsterdam (1990)
20. Marshall, A.G., Comisarow, M.B., Parisod, G.: Relaxation and spectral line shape in Fourier transform ion cyclotron resonance spectroscopy. *J. Chem. Phys.* **71**, 4434–4444 (1979)
21. Comisarow, M.B., Marshall, A.G.: Theory of Fourier transform ion cyclotron resonance mass spectroscopy. I. Fundamental equations and low-pressure line shape. *J. Chem. Phys.* **64**, 110–119 (1976)
22. Aarstol, M., Comisarow, M.B.: Apodization of FT-ICR spectra. *Int. J. Mass Spectrom. Ion Process.* **76**, 287–297 (1987)
23. Brenna, J.T., Creasy, W.R.: Experimental evaluation of apodization functions for quantitative Fourier transform mass spectrometry. *Int. J. Mass Spectrom. Ion Process.* **90**, 151–166 (1989)
24. Cody, R.B., Kinsinger, J.A., Ghaderi, S., Amster, I.J., McLafferty, F. W., Brown, C.E.: Developments in analytical Fourier-transform mass spectrometry. *Anal. Chim. Acta.* **178**, 43–66 (1985)
25. Shibata, J.H.: ABCs of FT-NMR. *J. Chem. Educ.* **79**, 1324 (2002)
26. Shaw, D.: Fourier Transform NMR Spectroscopy. Elsevier, Amsterdam (1984)
27. Ernst, R.R.: Numerical Hilbert transform and automatic phase correc-tion in magnetic resonance spectroscopy. *J. Magn. Reson.* **1**, 7–26 (1969)
28. Liu, J., Koenig, J.L.: An automatic phase correction method in nuclear magnetic resonance imaging. *J. Magn. Reson.* **86**, 593–604 (1990)
29. Lindon, J.C., Ferrige, A.G.: Digitisation and data processing in Fourier transform NMR. *Prog. Nucl. Magn. Reson. Spectrosc.* **14**, 27–66 (1980)
30. de Brouwer, H.: Evaluation of algorithms for automated phase correction of NMR spectra. *J. Magn. Reson.* **201**, 230–238 (2009)
31. Sotak, C.H., Dumoulin, C.L., Newsham, M.D.: Automatic phase correction of Fourier transform NMR spectra based on the dispersion versus absorption (DISPA) lineshape analysis. *J. Magn. Reson.* **57**, 453–462 (1984)

32. Xian, F., Henderickson, C.L., Blakney, G.T., Marshall, A.: Broadband Phase Correction of Complex FT-ICR Mass Spectra. 56th ASMS Conference on Mass Spectrometry (2008)
33. Xian, F., Hendrickson, C.L., Blakney, G.T., Beu, S.C., Marshall, A.G.: Automated broadband phase correction of Fourier transform ion cyclotron resonance mass spectra. Anal. Chem. (2010) doi:10.1021/ac101091w
34. Craig, E., Marshall, A.G.: Dispersion versus absorption (DISPA) plots as an index of static magnetic field inhomogeneity. Use for adjustment of spinning shims in a superconducting magnet. J. Magn. Reson. **1986** (68), 283–295 (1986)
35. Marshall, A.G., Roe, D.C.: Theory of Fourier transform ion cyclotron resonance mass spectroscopy: response to frequency-sweep excitation. J. Chem. Phys. **73**, 1581–1590 (1980)
36. Guan, S., Marshall, A.G.: Stored waveform inverse Fourier transform (SWIFT) ion excitation in trapped-ion mass spectometry: theory and applications. Int. J. Mass Spectrom. Ion Process. **157–158**, 5–37 (1996)
37. Bracewell, R.N.: The Fourier Transform and its Applications. WCB/McGraw Hill, Boston (2002)
38. Wang, M., Marshall, A.G.: Laboratory-frame and rotating-frame ion trajectories in ion cyclotron resonance mass spectrometry. Int. J. Mass Spectrom. Ion Process. **100**, 323–346 (1990)
39. Kim, S., Rodgers, R.P., Blakney, G.T., Hendrickson, C.L., Marshall, A.G.: Automated electrospray ionization FT-ICR mass spectrometry for petroleum analysis. J. Am. Soc. Mass Spectrom. **20**, 263–268 (2009)
40. Guan, S., Marshall, A.G.: Stored waveform inverse Fourier transform (SWIFT) axial excitation/ejection for quadrupole ion trap mass spectrometry. Anal. Chem. **65**, 1288–1294 (1993)
41. Chow, K.H., Comisarow, M.B.: Frequency (mass) errors and phase dependence in magnitude-mode apodized Fourier transform–ion cyclo-tron resonance spectra. Int. J. Mass Spectrom. Ion Process. **89**, 187–203 (1989)
42. Tolmachev, A.V., Masselon, C.D., Anderson, G.A., Udseth, H.R., Smith, R.D.: Frequency shifts due to the interference of resolved peaks in magnitude-mode Fourier-transform ion cyclotron resonance mass spectra. J. Am. Soc. Mass Spectrom. **13**, 387–401 (2002)
43. Beu, S.C., Blakney, G.T., Quinn, J.P., Hendrickson, C.L., Marshall, A.G.: Broadband phase correction of FT-ICR mass spectra via simultaneous excitation and detection. Anal. Chem. **76**, 5756–5761 (2004)
44. Guan, S., Wahl, M.C., Marshall, A.G.: Elimination of frequency drift from Fourier transform ion cyclotron resonance mass spectra by digital quadrature heterodyning: ultrahigh mass resolving power for laser-desorbed molecules. Anal. Chem. **65**, 3647–3653 (1993)
45. Kaur, P., O'Connor, P.B.: Algorithms for automatic interpretation of high resolution mass spectra. J. Am. Soc. Mass Spectrom. **17**, 459–68 (2006)
46. Savitski, M.M., Ivonin, I.A., Nielsen, M.L., Zubarev, R.A., Tsybin, Y.O., Hakansson, P.: Shifted-basis technique improves accuracy of peak position determination in Fourier transform mass spectrometry. J. Am. Soc. Mass Spectrom. **15**, 457–461 (2004)
47. Zhang, L.K., Rempel, D., Pramanik, B.N., Gross, M.L.: Accurate mass measurements by Fourier transform mass spectrometry. Mass Spectrom. Rev. **24**, 286–309 (2005)
48. Ledford, E.B., Rempel, D.L., Gross, M.L.: Space charge effects in Fourier transform mass spectrometry. II. Mass calibration. Anal. Chem. **56**, 2744–2748 (1984)
49. Francl, T.J., Sherman, M.G., Hunter, R.L., Locke, M.J., Bowers, W.D., McIver Jr., R.T.: Experimental determination of the effects of space charge on ion cyclotron resonance frequencies. Int. J. Mass Spectrom. Ion Process. **54**, 189–199 (1983)