Available online at www.sciencedirect.com
ScienceDirect
Journal of Magnetic Resonance 191 (2008) 193–201
JMR Journal of Magnetic Resonance
www.elsevier.com/locate/jmr

Automatic phasing of MR images. Part II: Voxel-wise phase estimation

G. Larry Bretthorst *

Biomedical Magnetic Resonance Laboratory, Department of Radiology, Washington University, Campus Box 8227, 4525 Scott Avenue, Suite 2313, St. Louis, MO 63110, USA

Received 7 September 2007; revised 17 December 2007
Available online 27 December 2007

Abstract

Magnetic resonance images are typically displayed as the absolute value of the discrete Fourier transform of the *k*-space data. However, absorption-mode images, the real part of the discrete Fourier transform of the data after applying an appropriate phase correction, have significant advantages over absolute-value images. In a companion paper, the problem of estimating the phase parameters needed to produce an absorption-mode image when the phase of the complex image varies linearly as a function of position, a situation common in magnetic resonance images, was addressed. However, some magnetic resonance images have phases that can vary in a complicated, nonlinear, positionally dependent fashion. To produce an absorption-mode image from these data, one must first estimate the positionally dependent phase, and then use that phase estimate to produce an absorption-mode image. This paper addresses both of these problems by first using Bayesian probability theory to estimate the constant or zero-order phase *as a function of image position*, and then the calculations are illustrated by using them to generate absorption-mode images from data where the phase of the image is a nonlinear function of position.

Keywords: Absorption-mode images; Zero-order phase estimation; Bayesian probability theory

1. Introduction

Phase estimation is important in magnetic resonance spectroscopy because the spectrum of the data is typically displayed in absorption-mode. An absorption-mode spectrum is the real part of the discrete Fourier transform of the time-domain data after removing the confounding effects of phase. To display an absorption-mode spectrum, both the constant phase (called a zero-order phase) and a phase that varies linearly with frequency (called a first-order phase) must be estimated.

Magnetic resonance images are typically displayed in absolute value mode, rather than in absorption-mode, because estimating the required phase parameters is very difficult. As described in [1], absorption-mode images have significant advantages over absolute value images; these advantages include: eliminating correlations between the signal and the noise, eliminating constant offsets, eliminating absolute values and increasing the signal-to-noise ratio. Part of the problem of producing absorption-mode images was solved by the algorithm presented in [1], an algorithm that can produce an absorption-mode image from any 2D magnetic resonance data sets in which the phase varies linearly as a function of position in both phase-encode and readout directions. Linearly varying, positionally dependent phases are common in spin-echo experiments because the acquisition method tends to remove phase artifacts arising from inhomogeneous magnetic fields. Thus, the only phases arising in these experiments are due to centering the peak signal in the middle of the acquisition window, and these phases vary linearly. However, gradient-echo acquisition methods do not remove these phase artifacts and, consequently, the phase in a gradient-echo experiment can vary in a complicated, positionally dependent, nonlinear fashion.

The effects of these nonlinearly varying phases are illustrated in Fig. 1. This T1-weighted, 2D, gradient-echo image

* Fax: +1 314 362 0526.
E-mail address: gbretthorst@wustl.edu

1090-7807/$ - see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmr.2007.12.011

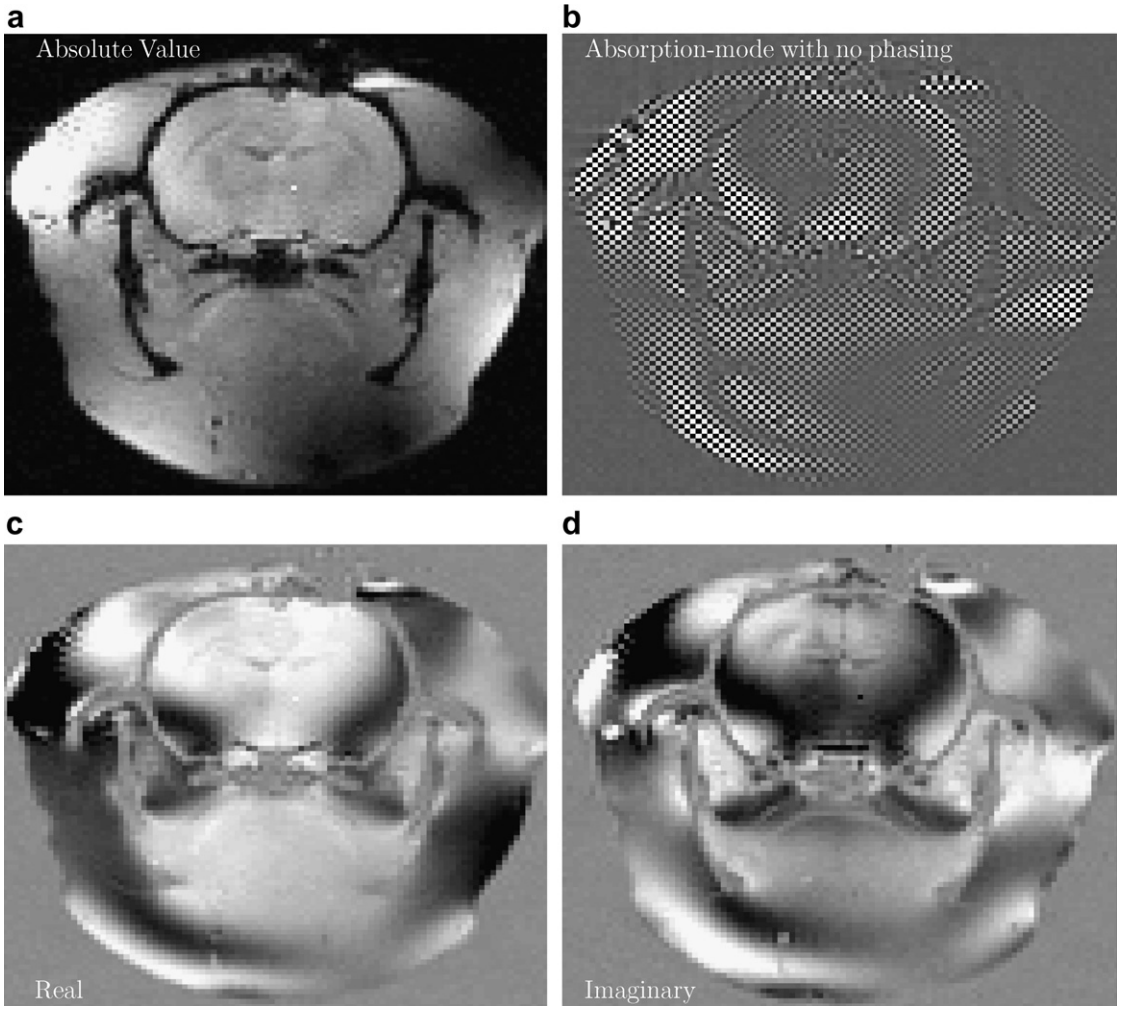

Fig. 1. T1-weighted, 2D, gradient-echo *axial* image of the head of an anesthetized rat. Panel a is an absolute value image. Panel b is an absorption-mode image before any phasing parameters were applied. Panels c and d are the real and imaginary images after applying the linear phasing algorithm [1]. Note that the image intensity oscillates between positive and negative values in both the real and imaginary images. In a properly phased absorption-mode image, panel d should contain only noise.

was acquired at 4.7-T using a 3-cm inner diameter, double-tuned $^{23}Na/^{1}H$ birdcage (quad) coil with TE = 0.01 s. The sample is an anesthetized rat that had undergone lateral ventricular infusion of a sodium shift reagent. The in-plane image resolution is 0.31×0.31 mm^2, the slice thickness is 1 mm, with TE = 10 ms, TR = 500 ms with a 20° flip angle. The matrix size was 128×128 with a field-of-view of 4 cm by 4 cm.

This *axial* image slice was taken through the head. Panel a is the absolute value image. Panel b is the real part of the discrete Fourier transform before applying any phasing parameters. The high frequency oscillations apparent in this panel are largely caused by the origin of time in the discrete Fourier transform. Panels c and d are the real and imaginary images that result from applying the linear phasing algorithm described in [1]. The linear phasing algorithm has removed the high frequency oscillations, but it has left behind a slowly varying, positionally dependent phase variation that causes the signal intensity to oscillate between the real and imaginary parts of the image. In a properly phased absorption-mode image, all of the signal intensity is in the real part of the image; the imaginary part contains only noise.

The calculations presented in [1] were for the first-order phase parameters. A point estimate for the zero-order or constant phase was given in that paper, but it was not derived. In this paper, the posterior probability for the zero-order phase is derived. Interestingly this posterior probability can be obtained in a compact form from which the peak and width of the posterior can then be determined by inspection. Finally, the calculations are illustrated by using them to generate absorption-mode images from data where the phase of the image is a nonlinear, positionally dependent, function of position.

2. The model equation

In Bayesian probability theory, the first step in any inference is to relate the hypothesis of interest to the available data. Here, the hypothesis is about the value of the

phase. This hypothesis is of the form, "the unknown value of the phase is θ". The numerical value of θ indexes a series of hypotheses of this form, and the posterior probability for θ ranks these hypotheses. Low values of the posterior probability indicate that the numerical value of θ is improbable, while high values indicate high probability for the hypothesis.

In the problem being solved, the complex data contain a signal with a positionally varying amplitude with a constant, zero-order phase. No information is available about the functional form of the signal variation other than the signal varies reasonably smoothly. If d_i represents the ith complex data value, then the data are related to the phase through a model given by:

$$d_i = A_i \exp\{-\mathrm{i}\theta\} + n_i \quad (1 \leqslant i \leqslant N), \tag{1}$$

where A_i is the signal intensity or amplitude of the ith signal value and θ is the zero-order phase. The total number of complex data values is N. In this paper we will refer to these data values as voxels as if they are MR volume elements. The number of data values will range from one to the total number of complex voxels in an image. The quantity n_i represents the complex noise. Separating the complex model into its real and imaginary parts, one has

$$d_{\mathrm{R}i} = A_i \cos\theta + n_{\mathrm{R}i} \tag{2}$$

for the real data, and

$$d_{\mathrm{I}i} = -A_i \sin\theta + n_{\mathrm{I}i} \tag{3}$$

for the imaginary data, where $d_{\mathrm{R}i}$ and $d_{\mathrm{I}i}$ represent the real and imaginary parts of the complex data, d_i, and $n_{\mathrm{R}i}$ and $n_{\mathrm{I}i}$ represent the real and imaginary parts of the complex noise. The noise characteristics of the real and imaginary data are assumed to be the same.

3. The bayesian calculations

Initially, this problem will be solved assuming that the standard deviation of the noise prior probability, σ, is known. Later this constraint will be relaxed and, if σ is not known, it will be removed using the sum and product rules. All of the information in the data relevant to phase estimation is summarized in the posterior probability for the phase. This posterior probability will be written as $P(\theta|\sigma DI)$, which is read as the posterior probability for the phase, θ, given the standard deviation of the noise prior probability σ, *all* of the data, D, and the background information I. The posterior probability for the phase is computed by applying Bayes' theorem [2]:

$$P(\theta|\sigma DI) = \frac{P(\theta|\sigma I)P(D|\sigma\theta I)}{P(D|\sigma I)}, \tag{4}$$

where the prior probability for the phase, $P(\theta|\sigma I)$, represents what is known about the phase before acquiring the data; the direct probability for the data, $P(D|\sigma\theta I)$, is a marginal probability from which the dependence on the amplitudes has been removed, and $P(D|\sigma I)$ is a normalization constant. If this probability density function is normalized at the end of the calculation, one obtains

$$P(\theta|\sigma DI) \propto P(\theta|I)P(D|\theta\sigma I), \tag{5}$$

where σ was dropped from the prior probability for the phase, $P(\theta|I)$.

The direct probability for the data, $P(D|\theta\sigma I)$, is a marginal probability. Marginal probabilities are those from which one or more parameters have been removed using the sum rule of probability theory. In this case, the A_i have been removed. To proceed, these amplitudes must be reintroduced into the direct probability and then removed using the sum rule. Reintroducing the amplitudes, one obtains

$$P(\theta|\sigma DI) \propto P(\theta|I) \int \mathrm{d}AP(AD|\theta\sigma I) \tag{6}$$

as the posterior probability for the phase, where an "A" without a subscript is being used to mean the collection of *all* of the amplitudes. Applying the product rule, the right-hand side of this equation may be factored:

$$P(\theta|\sigma DI) \propto P(\theta|I) \int dAP(A|\sigma I)P(D|A\theta\sigma I), \tag{7}$$

where $P(A|\sigma I)$ is the joint prior probability for the amplitudes. The probability for the data given the parameters is represented by $P(D|A\theta\sigma I)$ and is essentially the likelihood.

The prior probability for θ will be assigned as a uniform bounded prior:

$$P(\theta|I) = \begin{cases} \frac{1}{2\pi} & \text{if } 0 \leqslant \theta \leqslant 2\pi \\ 0 & \text{otherwise.} \end{cases} \tag{8}$$

When the prior probability for the amplitudes is assigned, a correlated prior will be used to impose constraints on the amplitudes. For example, in a typical gradient-echo image, the amplitude is a slowly varying function of position. Consequently, one might want to impose a condition on the amplitudes that indicates that adjacent amplitudes are approximately equal. This would be equivalent to assigning a constraint on the first or second derivative of the amplitude. Such constraints are easily imposed using Gaussian prior probabilities, and these priors have maximum entropy for a given value of the constraints. The amplitude prior will be assigned as a generalized Gaussian of the form:

$$P(A|\beta\sigma I) \propto \sigma^{-N} \exp\left\{-\frac{\beta^2}{2\sigma^2}\sum_{k=1}^{N}\sum_{l=1}^{N} A_k U_{kl} A_l\right\}, \tag{9}$$

where the hyperparameter, β, part of I, has been made explicit. A more general version of this prior that includes mean-value estimates of the amplitudes could be assigned. However, for the purposes of this paper, this more general prior is unnecessary; smoothness constraints, etc., may be expressed using this simpler form. Also, factors involving σ, but not β, are being tracked, because eventually σ will

be marginalized from the posterior probability for the phase. The matrix, U_{kl}, specifies how the amplitudes are related to each other, and the parameter, β, expresses how strongly this prior information is believed. Note that the functional form of this prior, writing β/σ rather than just β, is for notational convenience only.

As noted earlier, the noise in the real and imaginary data have the same noise characteristics and thus can be assigned using the same noise prior probability. If each data value is assumed logically independent, then the direct probability for the data may be written as

$$P(D|A\theta\sigma I) \propto \prod_{i=1}^{N} \sigma^{-2} \exp\left\{-\frac{(d_{\mathrm{R}i} - A_i\cos\theta)^2}{2\sigma^2}\right\} \times \exp\left\{-\frac{(d_{\mathrm{I}i} + A_i\sin\theta)^2}{2\sigma^2}\right\}. \tag{10}$$

Using Eqs. (8)–(10), the posterior probability for the phase may be written as:

$$P(\theta|\sigma DI) \propto \int \mathrm{d}A\sigma^{-3N} \exp\left\{-\frac{Q}{2\sigma^2}\right\}, \tag{11}$$

where several constants that cancel when this distribution is normalized have been dropped. The quantity, Q, is given by

$$Q \equiv \sum_{k=1}^{N}\sum_{l=1}^{N} \beta^2 A_l U_{kl} A_k + \sum_{i=1}^{N}(d_{\mathrm{R}i} - A_i\cos\theta)^2 + \sum_{i=1}^{N}(d_{\mathrm{I}i} + A_i\sin\theta)^2. \tag{12}$$

The first term on the right-hand side of this equation comes from the prior probability for the amplitudes; the second and third terms come from the direct probability or likelihood. Expanding, this quadratic, one obtains

$$Q \equiv 2N\overline{d^2} - 2\sum_{i=1}^{N} A_i T_i + \sum_{l=1}^{N}\sum_{k=1}^{N} A_l V_{kl} A_k, \tag{13}$$

where the T_i are given by

$$T_i \equiv d_{\mathrm{R}i}\cos\theta - d_{\mathrm{I}i}\sin\theta. \tag{14}$$

The interaction matrix, V_{kl}, is given by

$$V_{kl} \equiv \delta_{kl} + \beta^2 U_{kl}, \tag{15}$$

where δ_{kl} is a Kronecker delta function. The mean-square data value is defined as

$$\overline{d^2} \equiv \frac{1}{2N}\sum_{i=1}^{N}(d_{\mathrm{R}i}^2 + d_{\mathrm{I}i}^2). \tag{16}$$

The functional form of Q in Eq. (13) is a quadratic in A_i, so the integrals over A_i in Eq. (11) are Gaussian quadrature integrals. Such integrals are easily evaluated, and only the results are given:

$$P(\theta|\sigma DI) \propto \sigma^{-2N} \exp\left\{-\frac{2N\overline{d^2} - N\overline{h^2}(\theta)}{2\sigma^2}\right\}, \tag{17}$$

where

$$\overline{h^2}(\theta) = \frac{1}{N}\sum_{i=1}^{N} \hat{A}_i T_i \tag{18}$$

is the mean-square projection of the data onto the model for a given θ, and is a sufficient statistic for estimating the phase, θ. The $\hat{A}_i$ are the estimated amplitude of the image voxels and are given by the solution to

$$\sum_{l=1}^{N} V_{kl}\hat{A}_l = d_{\mathrm{R}k}\cos\theta + d_{\mathrm{I}k}\sin\theta. \tag{19}$$

For convenience, this is written as

$$\hat{A}_l \equiv \hat{a}_l\cos\theta + \hat{b}_l\sin\theta, \tag{20}$$

where $\hat{a}_l$ and $\hat{b}_l$ are given by the inverse of the V_{kl} matrix dotted into the column vectors represented by the data, d_{Rl} and d_{Il}, respectively.

The sufficient statistic,

$$\overline{h^2}(\theta) = \frac{1}{N}\sum_{i=1}^{N}[d_{\mathrm{R}i}\hat{a}_i\cos^2\theta - (d_{\mathrm{R}i}\hat{b}_i + d_{\mathrm{I}i}\hat{a}_i)\cos\theta\sin\theta + d_{\mathrm{I}i}\hat{b}_i\sin^2\theta], \tag{21}$$

can be simplified by using

$$\cos^2\theta = \frac{1}{2}(1 + \cos 2\theta), \tag{22}$$

$$\sin^2\theta = \frac{1}{2}(1 - \cos 2\theta) \tag{23}$$

and

$$\sin\theta\cos\theta = \frac{1}{2}\sin 2\theta \tag{24}$$

to replace the sines and cosines appearing in Eq. (21) by a single cosine:

$$\overline{h^2}(\theta) = \frac{Y + \sqrt{W^2 + X^2}\cos(2\theta + \psi)}{2N}, \tag{25}$$

where the constants, W, X, Y and ψ, are given by

$$W \equiv \sum_{i=1}^{N}(\hat{a}_i d_{\mathrm{R}i} - \hat{b}_i d_{\mathrm{I}i}), \tag{26}$$

$$X \equiv -\sum_{i=1}^{N}(\hat{a}_i d_{\mathrm{I}i} + \hat{b}_i d_{\mathrm{R}i}), \tag{27}$$

$$Y \equiv \sum_{i=1}^{N}(\hat{a}_i d_{\mathrm{R}i} + \hat{b}_i d_{\mathrm{I}i}) \tag{28}$$

and

$$\psi \equiv \tan^{-1}\left(\frac{X}{W}\right). \tag{29}$$

Substituting the sufficient statistic, Eq. (25), back into the posterior probability for the phase, Eq. (17), one obtains

$$P(\theta|\sigma DI) \propto \sigma^{-2N} \times \exp\left\{-\frac{4N\overline{d^2} - Y - \sqrt{W^2+X^2}\cos(2\theta+\psi)}{4\sigma^2}\right\}, \tag{30}$$

where a number of constants that cancel when this distribution is normalized have been dropped.

If the standard deviation of the noise is known, as it often is in images, then this equation may be further simplified:

$$P(\theta|\sigma DI) \propto \exp\left\{\frac{\sqrt{W^2+X^2}\cos(2\theta+\psi)}{4\sigma^2}\right\}. \tag{31}$$

In this form, it is obvious that the peak of this distribution occurs when the argument of the cosine is zero,

$$\theta_{\mathrm{Max}} = -\frac{\psi}{2}. \tag{32}$$

Noting this maximum, one can Taylor expand about this maximum to second order and use the width of this Gaussian approximation as an estimate of the width of the posterior probability to obtain

$$\theta_{\mathrm{est}} = -\frac{\psi}{2} \pm \sqrt{\frac{8\sigma^2}{\sqrt{W^2+X^2}}} \tag{33}$$

as the peak ± standard deviation estimate of the phase. This estimate can be calculated directly without ever having to evaluate the posterior probability. Experience with this approximation indicates that even at the modest signal-to-noise ratio of 3:1 *and one* complex data *value*, this approximation differs from the exact calculation by less than 10 percent. When the number of complex data values is greater than one, this estimate of the phase is very different from $\tan^{-1}(\mathrm{I}/\mathrm{R})$, where R and I are the real and imaginary parts of the complex image. It is very different because this formula uses multiple data values to estimate the phase. However, when only a single data value is analyzed, Eq. (33) reduces to $\tan^{-1}(\mathrm{I}/\mathrm{R})$, even when β is not zero. Eq. (33) looks very different from $\tan^{-1}(\mathrm{I}/\mathrm{R})$, but that difference is superficial and is caused by the introduction of the half-angle formulas used to obtain a single cosine term.

This approximation, Eq. (33), assumes that σ is known, and this may or may not be the case. Using the posterior probability for the phase given the standard deviation of the noise, Eq. (30), the sum and product rules can be used to remove the the dependence on σ:

$$P(\theta|DI) \propto \int \mathrm{d}\sigma P(\sigma\theta|DI) = \int \mathrm{d}\sigma P(\sigma|I)P(\theta|\sigma DI). \tag{34}$$

If the prior probability for the standard deviation, $P(\sigma|I)$, is assigned using a Jeffreys' prior, $1/\sigma$, then the indicated integral may be evaluated and the posterior probability for the phase may be written as

$$P(\theta|DI) \propto [4N\overline{d^2} - N\overline{h^2}(\theta)]^{-N} = [4N\overline{d^2} - Y - \sqrt{W^2+X^2}\cos(2\theta+\psi)]^{-N}. \tag{35}$$

As shown in [3], around the location of the maximum, this is approximately

$$P(\theta|DI) \approx \exp\left\{\frac{\sqrt{W^2+X^2}\cos(2\theta+\psi)}{4\langle\sigma^2\rangle}\right\} \tag{36}$$

and one would estimate

$$\theta_{\mathrm{est}} = -\frac{\psi}{2} \pm \frac{\sqrt{8\langle\sigma^2\rangle}}{\sqrt{W^2+X^2}} \tag{37}$$

where the standard deviation of the noise prior probability has been replaced by (essentially) the mean-square residual:

$$\langle\sigma^2\rangle = \frac{N}{2N-2}[2\overline{d^2} - \overline{h^2}(\theta_{\mathrm{Max}})] \tag{38}$$

where this estimate assumes N is greater than 2.

4. Results and discussion

These calculations will be illustrated using the complex image shown in Fig. 1. Two examples will be given and both examples use only a single complex data value, so $N = 1$. The first example will illustrate the full posterior probability and the Gaussian approximation. The second example will illustrate how the calculations can be used to generate absorption-mode images. In this first example, two different complex data values will be selected, one from a region where there is signal, and one where there is only noise. When a single voxel is used as the data, only the width of the posterior probability depends on U_{kl}; the peak is independent of both U_{kl} and β. In the following examples $U_{kl} = 1$ and $\beta = 0.01$. Fig. 2a is the posterior probability for the phase computed from a voxel having signal-to-noise ratio of approximately 40:1; here, the posterior probability and the Gaussian approximation are indistinguishable. Panel b is the posterior probability (solid line) and the Gaussian approximation (dotted line) computed from a voxel that contained only noise. For this voxel, the posterior probability has a shallow peak near 60° and the posterior extends over the entire nonaliased parameter range. The Gaussian approximation shares this peak, but has a standard deviation that is too small and consequently falls off too quickly. When signal-to-noise ratios are about 3:1, the Gaussian is narrower than the full distribution but *is* a good approximation to the full distribution. For all practical purposes, it is only when a voxel contains only noise that the full distribution is needed. However, this case is almost irrelevant because there is no phase to estimate.

Next the generation of an absorption-mode image is illustrated using the complex image shown in Fig. 1. To generate an absorption-mode image, each complex voxel

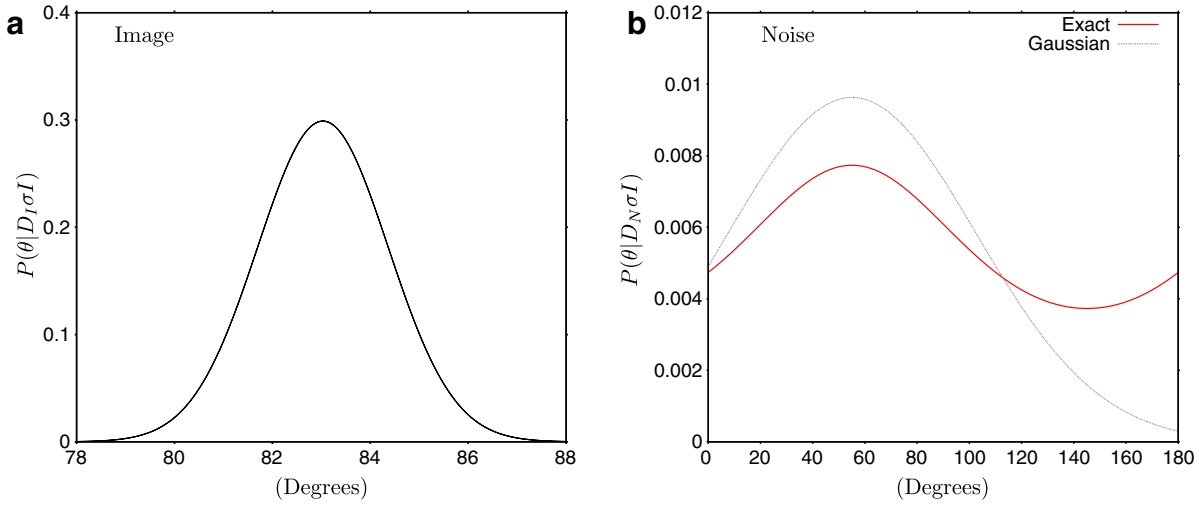

Fig. 2. Panel a is the posterior probability for the phase θ computed from a single voxel taken from the complex image shown in Fig. 1. In this voxel the signal-to-noise ratio is roughly 40:1 and the Gaussian approximation given in Eq. (33) is indistinguishable from the full posterior probability. Panel b is the posterior probability computed from a voxel containing only noise; here the Gaussian approximation (dotted line) is not a good representation of the full posterior (solid line).

in the image must be multiplied by $\exp\{-i\hat{\theta}\}$, where $\hat{\theta}$ is an estimate of the phase in a given voxel. This phase is called a point estimate because it replaces the full posterior probability by a single value. One might think that generating an absorption-mode image is as simple as taking $\hat{\theta} = \tan^{-1}(I/R)$ as the point estimate on a voxel by voxel basis, where 'R' and 'I' are the real and imaginary parts of the complex image. This phase, $\hat{\theta}$, is identical to the maximum likelihood estimate of the phase using an image voxel model of the form $A\exp\{i\hat{\theta}\}$ and for a single complex data value is *identical* to the estimate given in Eq. (33) even when an informative Gaussian prior probability is assigned. If one multiplies this complex image voxel by $\exp\{-i\hat{\theta}\}$, then

$$\text{Phased image voxel} = (R + iI) \times \exp\{-i\hat{\theta}\}. \tag{39}$$

Separating this expression into its real and imaginary parts one has

$$\text{Real part} = R\cos\hat{\theta} + I\sin\hat{\theta} \tag{40}$$

and

$$\text{Imaginary part} = -R\sin\hat{\theta} + I\cos\hat{\theta}. \tag{41}$$

But

$$\cos\hat{\theta} = \frac{R}{\sqrt{R^2+I^2}} \quad \text{and} \quad \sin\hat{\theta} = \frac{I}{\sqrt{R^2+I^2}}, \tag{42}$$

so the expressions for the real part of the complex voxel can be simplified:

$$\text{Real part} = \frac{R^2+I^2}{\sqrt{R^2+I^2}} = \sqrt{R^2+I^2}, \tag{43}$$

and similarly the imaginary part is given by

$$\text{Imaginary part} = \frac{-RI+IR}{\sqrt{R^2+I^2}} = 0. \tag{44}$$

Using the phase, $\hat{\theta} = \tan^{-1}(I/R)$, *does not* produce an absorption-model image; rather it is *identical* to computing an absolute value mode image. Consequently, if one wishes to produce an absorption-mode image, at the very minimum, some other point estimate of the phase must be used.

As noted, to remove the effects of the nonlinear phase in any given voxel one must multiply the complex unphased voxel by $\exp\{-i\hat{\theta}\}$, where $\hat{\theta}$ is a point estimate of the phase. The difficulty in generating an absorption-mode images lies in choosing this point estimate. Often, in parameter estimation problems, one chooses this point estimate as the value that maximizes the posterior probability. However, using this maximum for *every* voxel results in an absolute value image rather than an absorption-mode image. The solution is to sample the posterior probability for the phase for each voxel and thus preserve the desired noise properties expected from an absorption-mode image. Additionally, one cannot arbitrarily choose the phase in such a way as to make the amplitude in the real channel positive. This approach introduces a constant offset into the image, and this is most definitely not desired. Indeed, in a perfectly phased image, the noise in both the real and imaginary parts of the images should have zero mean and have the same noise standard deviation. Deviations from these conditions introduce artifacts into the images.

For any given voxel, the posterior probability for the phase summarizes all of the information in the data about the phase. If a sample is drawn from this posterior probability, then this sample can be used to generate an absorption-mode voxel. On average, any given sample from the posterior is as good a candidate point estimate as any other, because on average it reflects everything known about the phase. For a high signal-to-noise ratio voxel,

the posterior probability can be sampled using the Gaussian approximation, Eq. (33), and that sample used to produce an absorption-mode voxel. High signal-to-noise ratio voxels should be phased so that they have positive real amplitudes.

However, high signal-to-noise ratio voxels are not the problem. The problem is what to do about low signal-to-noise ratio voxels. If a sample of the phase were drawn from the posterior probability shown in Fig. 2b, virtually any value of the phase could be generated for that low signal-to-noise ratio voxel, implying that the resulting real phased image could have either positive or negative intensity. Consequently, the positivity condition used on high signal-to-noise ratio voxels cannot be arbitrarily applied to low signal-to-noise ratio voxels; doing so would *necessarily* introduce a constant offset. A criterion must be established that indicates when to apply the positivity condition. From a Bayesian perspective, this is a model selection calculation with a decision.

When the phase estimation calculations given in this paper were first implemented to produce absorption-mode images, determining the voxels to which the positivity condition should be applied was done using a model selection calculation. In that calculation, two models were defined: a ''signal'' model and a ''no signal'' model. The signal model was the model used in this paper, and the no signal model contained only noise. The posterior probability for these models was then computed. For the signal model, this calculation is the same calculation given in this paper with one additional integration over the phase. The calculation of the logarithm of the posterior probability for the no signal model consists of little more than computing the mean-square data value. However, it quickly became apparent that the resulting model selection calculation, when only a single complex data value is used, thresholded voxels based on the signal-to-noise ratio. The posterior probability for the signal model was always 1 for voxels with signal-to-noise ratio larger than about 3:1; consequently, these voxels were always phased with the positivity condition. However, when the signal-to-noise ratio dropped below about 2.5:1, the posterior probability for the no signal model was always 1, and these voxels were always phased without the positivity condition. It was only voxels in the very small signal-to-noise range of 2.5:1–3:1 where a decision had to be made. In test implementations of these calculations, these voxels were phased by using a Markov chain Monte Carlo simulation to draw a sample from the posterior probability for the model. If that sample's model was a signal model, then the voxel was phased with the positivity constraint, and if the sample's model was a no signal model, then the voxel was phased without the positivity constraint. Thus, the distribution of models used to phase the image was the posterior probability distribution for the model.

These additional calculations increased the complexity of the resulting algorithm, slowed it down and added nothing to the resulting images that could not be achieved by simply thresholding the voxels at signal-to-noise ratio 3:1. Thus, in the current algorithm, all of the voxels having signal-to-noise ratios less than 3:1 are phased without the positivity constraint, and those above are phased with the positivity constraint. In both cases, the point estimate of the phase is drawn from an approximation. For voxels having signal-to-noise ratio greater than 3:1 the Gaussian approximation, Eq. (33), is used. The no signal voxels by definition have no signal and are phased using a sample drawn from a uniform distribution.

Fig. 1 illustrates what happens when the linear phasing algorithm, described in [1], is applied to a gradient-echo image. A strong signal persists in the imaginary image, Fig. 1d. Fig. 3a and b are the real and imaginary parts of this same image phased using samples drawn from the posterior probability for the phase, Eq. (31). To generate this image, the noise standard deviation was estimated from a region of the image in which there was no signal. This value, 0.038, was then used in the Gaussian approximation, Eq. (33), to draw samples from the posterior probability of the phase. Note that the real image, Fig. 3a, contains the absorption-mode image plus noise. The mean value of this noise, computed from roughly 4K voxels above the animal's head that contained only noise, was 0.002. This region (not shown) was also used to generate the histogram shown in Fig. 3c. The standard deviation of the noise in this region was 0.037. The imaginary image Fig. 3b contains only noise. Its mean was 5×10^{-5} and its standard deviation was 0.038; these values were computed from all 16K samples in the imaginary image. These same 16K samples were also used to generate the histogram shown in panel Fig. 3d.

The mean and standard deviation of the intensities of the 16K points in the imaginary image, Fig. 3b, are essentially the same as those in the noise region above the mouse's head, Fig. 3a, and in the original unphased image Fig. 1b. Generating an absorption-mode image has moved the signal intensity to the real part of the image, leaving only noise in the imaginary part. The noise level remains unchanged because each complex voxel was multiplied by $\exp\{-i\theta_i\}$, which is a rotation, and rotations do not change the magnitude of a vector, in this case the complex voxel values. However, any two runs of this nonlinear phasing algorithm using a different random number seed will produce slightly different absorption-mode images from the same data set. The reason for this is that this absorption-mode image is the image produced from one Markov chain Monte Carlo sample from the joint posterior probability for all of the phases in the image and different samples from this posterior will produce different images, each equally consistent with what is known about the phase.

Finally, a map of the estimated phases is shown in Fig. 4. This map was unwrapped using the Laplace technique described in [4]. The image phase map is important in magnetic resonance imaging because it is essentially a map of the inhomogeneous magnetic field and can be used to improve the image quality. Additionally, some magnetic

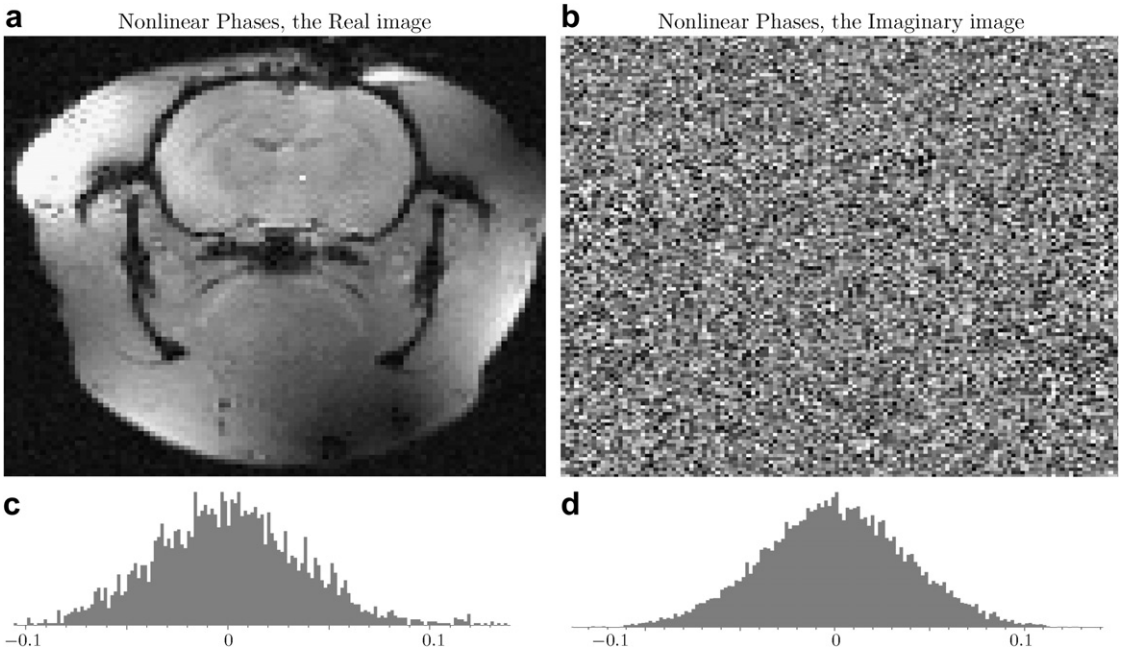

Fig. 3. Panel a is the real image generated by the nonlinear phasing algorithm. The imaginary image b is essentially noise. Panel c is a histogram created from roughly 4K voxel in a region that contained no signal (area above the mouse head, not shown). The mean in this region was 0.002 and the standard deviations was 0.037. Panel d is a histogram created from the entire imaginary image, 16K counts. The mean was 5×10^{-5} while the standard deviation was 0.038.

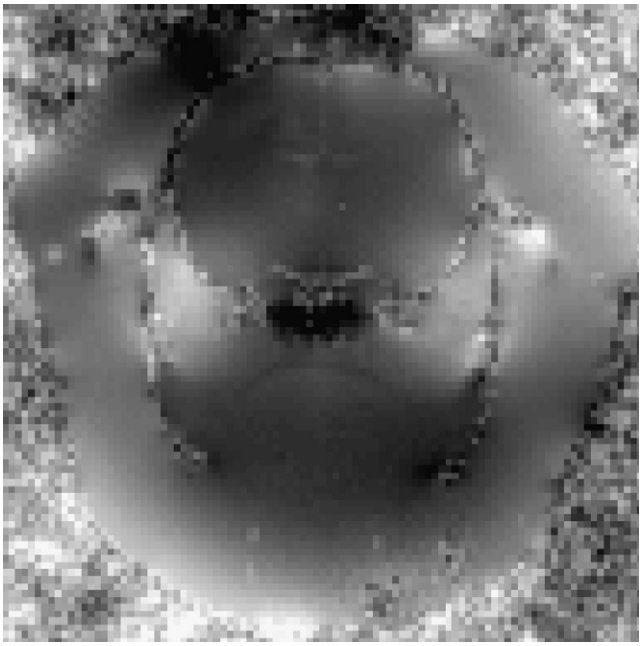

Fig. 4. The nonlinear phasing routine produces an unwrapped map of the phase of the image. Because the linearly varying phase was removed using the algorithms described in [1], this phase map is of deviations from linear. The unwrapping is done using the solution to the Laplace equation described in [4]. This phase map, left, is important in magnetic resonance imaging because, among other things, it is essentially a map of the inhomogeneous magnetic field and can be used to improve the image quality.

resonance imaging experiments encode the information of interest in the phase.

5. Summary and conclusions

The calculation presented in [1] for the first-order phase parameters and the calculation presented here for the zero-order phase parameters enable one to phase most magnetic resonance images. When the phases vary linearly, as in spin-echo experiments, the linear phasing algorithm can be used to phase the images in $n \log n$ time. On a Sun Ultra 60 workstation, the time required to produce an absorption-mode images using the linear phasing algorithm for a typical 128×128 image is less than one second. When the phases vary nonlinearly, as they do in gradient-echo experiments, the calculation presented in this paper can be used to phase the image in n time. Note that the calculations presented in this paper are image domain calculations. Consequently, the algorithm described in this paper operates on the images produced by the linear phasing algorithm. Thus, the slow steps in generating an absorption-mode images are removing the effects of the linear phase and unwrapping the phase maps, both of which require a fast discrete Fourier transform.

Both the calculations presented in [1] and the calculations presented here are more general than they first appear. In [1], the calculations were all described in terms of estimating the first-order phase parameter in a complex data set. However, those calculations could just as easily have been described as estimating the frequency of a complex sinusoid that has a time (or positionally) varying amplitude. In this paper, the calculations were described in a way that naturally leads one to use these calculations to generate absorption-mode images. However, the calcula-

tions themselves are general and apply to any complex data set in which a single constant phase must be estimated.

Acknowledgments

I thank Joseph J. H. Ackerman, Jeffrey J. Neil, Joel R. Garbow, Dmitriy Yablonskiy, Alex Sukstansky and Josh Shimony for encouragement, support, and helpful comments. This work was supported by the Small Animal Imaging Resources Program (SAIRP) of the National Cancer Institute, Grant R24CA83060, and by Grants NS35912, NS41519, NS41519 and HL70037.

References

[1] G.L. Bretthorst, Automatic phasing of MR images. Part I: Linearly varying phase, J. Magn. Reson. 191 (2008) 184–192.
[2] T. Bayes, Rev, An essay toward solving a problem in the doctrine of chances, Philos. Trans. R. Soc. Lond. 53 (1763) 370–418; reprinted in Biometrika 45 (1958) 293–315, and Facsimiles of Two Papers by Bayes, commentary by W. Edwards Deming, New York, Hafner, 1963.
[3] G.L. Bretthorst, Bayesian spectrum analysis and parameter estimation, in: J. Berger, S. Fienberg, J. Gani, K. Krickenberg, B. Singer (Eds.), Lecture Notes in Statistics, vol. 48, Springer-Verlag, New York, 1988.
[4] D.C. Ghiglia, M.D. Pritt, Two-Dimensional Phase Unwrapping, John Wiley and Son, New York, 1998.