Optimal Control Solutions to the Magnetic Resonance Selective Excitation Problem

STEVEN CONOLLY, STUDENT MEMBER, IEEE, DWIGHT NISHIMURA, MEMBER, IEEE, AND ALBERT MACOVSKI, FELLOW, IEEE

Abstract—**Most magnetic resonance imaging sequences employ field gradients and amplitude modulated RF pulses to excite only those spins lying in a specific plane. The fidelity of the resulting magnetization distribution is crucial to overall image resolution. Conventional RF-pulse design techniques rely on the small tip-angle approximation to Bloch's equation, which is inadequate for the design of 90° and 180° pulses. This paper demonstrates the existence of a selective pulse, and provides a sound mathematical and computational basis for pulse design. It is shown that the pulses are optimal in the class of piecewise continuous functions of duration T. An optimal pulse is defined as the pulse on the interval that achieves a magnetization profile "closest" to the desired distribution. Optimal control theory provides the mathematical basis for the new pulse design technique. Computer simulations have verified the efficacy of the 90° and the 180° inversion and "pancake-flip" optimal pulses.**

INTRODUCTION

MOST magnetic resonance imaging sequences employ field gradients and radio-frequency (RF) selective pulses to excite only those spins lying in a specific plane. The fidelity of the resulting magnetization distribution is crucial to overall image resolution. Due to the nonlinear relationship between a pulse envelope and its resultant magnetization distribution, there is no simple *inversion* formula to guide the design of selective RF pulses. Hence, conventional pulse design techniques rely on the small tip-angle approximation of Bloch's and its Fourier transform solution of Bloch's equation. These techniques are inadequate for the design of large tip-angle pulses, the 90° and especially the 180° pulse envelopes. The selective excitation problem is to design large tip-angle pulses that achieve high selectivity.

Several researchers have studied the selective excitation problem. Early work was pioneered by Garroway *et al.* [2]. Sutherland and Hutchinson [3] proposed selective excitation sequences for three-dimensional imaging. Hoult's [4] perturbation analysis is useful for computing candidate pulses and for suggesting focusing intervals. Mansfield [5] used rectangular pulses for selective excitation.

Other researchers have used computer simulations to predict the response to various excitation functions [6]–[9]. Caprihan [9] notes that the small tip-angle approximation of the Bloch equation leads to unpredictable errors for the large tip-angle cases. He concludes that a numerical method is required for systematic design of RF pulses.

More recent work has been directed towards numerical inversion of the Bloch equation. Silver *et al.* [10] have used iterative techniques to invert the Bloch–Riccati equation for the 90° case. They also designed a 180° complex (amplitude and phase-modulated) hyperbolic secant pulse [11], which has remarkable robustness to RF inhomogeneity. Lurie [12] has used an optimization algorithm to shape a parameterized sinc function for high selectivity. Nishimura [13] has used sequential pulses to achieve a desired distribution. We presented preliminary results of this investigation in [1]. Lent and Kritzer [14] have independently applied optimal control theory to the selective excitation problem.

The problem still remains, as pointed out in [13], that there is no mathematical foundation for pulse-envelope design. This paper provides both a sound mathematical basis for pulse design and an efficient algorithm to compute the pulses. It will be shown that these pulses are *optimal* in the class of piecewise continuous functions on a specified interval $(0, T)$. Note that we restrict our analysis to amplitude modulated pulses although extension to complex pulses is quite possible.

The first premise made is the definition of an optimal pulse. Most researchers would agree that the best pulse is the one that achieves a magnetization "closest" to the desired distribution. "Closest" can be interpreted with various mathematical formulas. Hence, we formulate selective excitation as a constrained optimization problem. This type of problem lies within the framework of dynamic optimization with nonholonomic constraints [15]. An entire field of research, optimal control theory, is devoted to dynamic optimization problems with differential equations as constraints [16]. Optimal control theory provides necessary and sufficient conditions for determining the optimality of a given pulse. The necessary conditions also structure an efficient algorithm for numerical solution of the optimal pulse [17].

Before we consider algorithms that find an optimal pulse, it is prudent to ask whether such controls exist. The first section of the paper answers the existence question. The second section is a sketch of some fundamental results from optimal control theory. The third section develops and compares various formulations of the selective excitation problem in the optimal control framework. The

Manuscript received December 24, 1985. This work was supported by a National Science Foundation Graduate Fellowship and by General Electric Medical Systems and the National Institutes of Health under Contracts HV38045 and HL34962.

The authors are with the Department of Electrical Engineering, Magnetic Resonance Systems Research Laboratory, Stanford University, Stanford, CA 94305.

IEEE Log Number 8606046.

0278-0062/86/0600-0106$01.00 © 1986 IEEE

last section presents simulation results for both the 90° and the 180° pulses.

I. Controllability of the MR Spin System

Consider the linear system defined by the differential equation

$$\dot{x} = Fx + gu(t) \tag{1}$$

where $u(t)$, the "control," is a scalar input to the system, and $x(t)$ is the $N \times 1$ state vector of the system. $\dot{x}$ denotes the time derivative of the state. F and g are constant $N \times N$ and $N \times 1$ matrices. A *reachable* state is any vector $x(T) \in R^N$ which can be intercepted with some control $u(t)$ in finite time T. This system is *controllable* if and only if the set of reachable states from $x(0)$ is the entire space R^N. The system (1) is controllable if and only if the controllability matrix $C(\mathbf{F}, g)$

$$C(\mathbf{F}, g) = [g, Fg, \cdots, F^{N-1}g]$$

is nonsingular [19]. Note that it is possible to determine the existence of the control without explicitly solving for it.

It is well known that Bloch's equation is nonlinear in the input $\gamma B_1(t) = \omega_1(t)$. It is perhaps not as well known that Bloch's equation lies on the linear fringe of nonlinear systems known as *bilinear* systems [20]. Control theorists have formulated powerful mathematical techniques to analyze bilinear systems. Neglecting relaxation, we can write Bloch's equation in rotating coordinates as

$$\dot{m}(f, t) = [a(f) + b\omega_1(t)]\, m(f, t) \tag{2}$$

where

$$a(f) = \begin{bmatrix} 0 & 2\pi f & 0 \\ -2\pi f & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

and

$$b = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix},$$

and the dot denotes the time derivative. If we employ a slice-select gradient in the z direction, then the spatial frequency variable f is $\gamma Gz/2\pi$. $m(f, t)$ is the 3×1 magnetization vector measured in the rotating coordinates. This system is called bilinear since $\dot{m}(f, t)$ is linear in both $\omega_1(t)$ and $m(f, t)$ when the other is held fixed.

Controllability results analogous to the linear system condition have been derived for the bilinear system (2). To apply the results to our spatial controllability problem we must consider an extended system of equations. Let us sample the magnetization vector at K locations along the gradient field. Define the extended $3K \times 1$ state vector

$$M(t) = [m(f_1, t), \cdots, m(f_K, t)].$$

The extended Bloch equation has no explicit dependence on the spatial frequency variable f.

$$\dot{M}(t) = [A + B\omega_1(t)]\, M(t) \tag{3}$$

where A is a block diagonal $3K \times 3K$ matrix with $a(f_i)$ on the diagonal. $\mathbf{B}$ is also block diagonal with $\mathbf{b}$ on each entry

$$A = \begin{bmatrix} \mathbf{a}(\mathbf{f}_1) & & \boldsymbol{O} \\ & \ddots & \\ \boldsymbol{O} & & \mathbf{a}(\mathbf{f}_K) \end{bmatrix}$$

and

$$B = \begin{bmatrix} \mathbf{b} & & \boldsymbol{O} \\ & \ddots & \\ \boldsymbol{O} & & \mathbf{b} \end{bmatrix}.$$

We assume a homogeneous initial distribution with all the spins initially aligned with B_0

$$M(0) = [0, 0, 1, \quad 0, 0, 1, \cdots, 0, 0, 1]^T.$$

Now we define a discrete, but arbitrarily accurate, selective excitation problem. We must find a control $\omega_1(t)$ that will drive $M(0)$ to

$$M(T) = D$$

where D is a $3K \times 1$ vector representing the desired magnetization spatial distribution sampled at the f_i. Here we are concerned only with existence of such a control.

In the Appendix, we show that each isochromat is controllable on its own sphere, independent of other isochromats. In fact, we find that there exists a finite time $T > 0$ such that, given any two profiles, there exists a "bang-bang" control $\omega_1(t)$ that obtains the second profile from the first in less than T units of time, as long as the length of the two profiles is identical at each f_i. The above result does not imply controllability for an *arbitrary* specified time T. Indeed, we know from perturbation analysis [4] that T should be inversely proportional to the spatial bandwidth of the desired profile.

The finite time result provides an interesting contrast with linear systems. For controllable linear systems, any state is (theoretically) reachable instantaneously using impulsive controls. Picture a single spin vector at a specific spatial frequency. If the B_1 pulse is very strong, then we can assume that the magnetic vector rotates about the x-axis. Starting from the "north pole" the vector nutates quickly to any desired latitude. Turning off the control, the vector precesses freely (at fixed radial frequency γGz) to the desired longitude. In some sense, it is the free precession portion of the motion of the entire distribution of spins that requires finite time.

In summary, the controllability result says that any Hermitian profile is attainable as long as the magnetization length is conserved at each spatial frequency and we have a sufficiently long interval T. Of course, these results

also apply to the less restricted control problem using complex $B_1(t)$ modulation.

II. Optimal Control Theory

Although we have established existence of a control $\omega_1(t)$ we have no formula or even an algorithm to find it. We formulate selective excitation as a constrained minimization problem and then apply the results of optimal control theory to obtain necessary conditions on $\omega_1(t)$. Classical applications of the theory have included satellite guidance and robot manipulator control. The necessary conditions provide an efficient algorithm for computing optimal pulses, much as the necessary conditions from calculus allow one to compute extrema of functions.

Optimal control theory addresses the general problem of minimizing some cost functional J over all possible controls $u(t)$

$$\min_{u(t)} J[u(t)] = \phi(\boldsymbol{x}(T)) + \int_0^T L(u, \boldsymbol{x}, t)\, dt \tag{4}$$

where

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{f}[\boldsymbol{x}(t), u(t), t], \qquad \boldsymbol{x}(0) \text{ given.} \tag{5}$$

ϕ usually represents some error function associated with the final state, and L represents some penalty function associated with, say, an expensive control or a slow control. Following the development in [17] we adjoin the differential equations (5) to the system (4) with unspecified Lagrange multiplier $\lambda(t)$

$$\min_{u(t)} \bar{J}\,[u(t)] = \phi(\boldsymbol{x}(T)) + \int_0^T [L(\boldsymbol{x}, u, t) + \boldsymbol{\lambda}^T(t)\,\{\boldsymbol{f}(\boldsymbol{x}, u, t) - \dot{\boldsymbol{x}}\}]\, dt.$$

It is clear that near the minimum of the quadratic function $q(u) = (u - a)^2$, small changes in u effect no (first-order) changes in $q(u)$. Similarly, if a control $u(t)$ is optimal, then a small perturbation $\delta u(t)$ will cause no change in the cost functional $\bar{J}$. This stationarity condition is the first-order necessary condition for an optimal $u(t)$.

The variation in $\bar{J}$ due to a variation $\delta u(t)$ in the control has two terms, one due to explicit dependence on $u(t)$ and one due to implicit dependence through $\boldsymbol{x}(t)$. The implicit variation cannot be computed in general since it obeys a nonlinear differential equation driven by $\delta u(t)$. The unspecified Lagrange multipliers allow us to separate the problem into two distinct differential equations. If we impose the following differential equation on the $\lambda(t)$[1]

$$\dot{\boldsymbol{\lambda}} = -\left[\frac{\partial L}{\partial \boldsymbol{x}} + \boldsymbol{\lambda}^T \frac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}}\right]^T, \qquad \boldsymbol{\lambda}^T(T) = \frac{\partial \phi(\boldsymbol{x}(T))}{\partial \boldsymbol{x}(T)} \tag{6}$$

it is possible to mask all the implicit variations through $\delta \boldsymbol{x}(t)$. We can show that the variation in $\bar{J}$ due to a perturbation in $u(t)$ is

$$\delta \bar{J} = \int_0^T \left[\frac{\partial L}{\partial u} + \boldsymbol{\lambda}^T \frac{\partial \boldsymbol{f}}{\partial u}\right] \delta u(t)\, dt. \tag{7}$$

[1]Gradient vectors are row vectors in the sequel.

For an extremum, $\delta \bar{J}$ must be zero for arbitrary $\delta u(t)$; this can happen only if

$$\boldsymbol{\lambda}^T \frac{\partial \boldsymbol{f}}{\partial u} + \frac{\partial L}{\partial u} = 0. \tag{8}$$

Hence, the necessary conditions comprise two differential equations, (5) and (6), which are coupled by the stationarity condition (8). Note that a computational difficulty arises because the $\boldsymbol{\lambda}$ boundary condition is specified at the right side of the interval, while the $\boldsymbol{x}$ boundary condition is specified at the left side of the time interval. This type of differential equation is called a *two-point boundary-value problem.*

Heuristically, the $\boldsymbol{\lambda}(t)$ represents feedback from the desired end condition, a small $\phi(\boldsymbol{x}(T))$ and small $L(\boldsymbol{x}, u, t)$. This feedback information really must flow backwards in time.

To solve the two-point boundary-value problem, the user makes an educated guess at the control. The differential equation (5) is integrated forward and (6) is integrated backwards from T using the initial control. Subsequent controls are modified to minimize (8)

$$u^{i+1}(t) = u^i(t) - \mu \left[\boldsymbol{\lambda}^T \frac{\partial \boldsymbol{f}}{\partial u^i} + \frac{\partial L}{\partial u^i}\right]. \tag{9}$$

If μ is small, then the perturbation $\delta u^i(t)$ given by (9) produces a variation in the cost functional (7)

$$\delta \bar{J}_i = -\mu \int_0^T \left[\frac{\partial L}{\partial u^i} + \boldsymbol{\lambda}^T \frac{\partial \boldsymbol{f}}{\partial u^i}\right]^2 dt \le 0.$$

So that $\bar{J}$ will always decrease until the stationarity condition (8) is met. Hence, the algorithm will iteratively compute better pulses. When $\delta \bar{J}$ becomes very small, the pulse is close to optimal, barring local minima. The steepest descent parameter μ is chosen by the experimenter. Large values give fast convergence, but can lead to instability. Other minimization algorithms such as the conjugate gradient algorithm or the Davidon–Fletcher–Powell algorithm [18] could substitute for the steepest descent algorithm described here.

III. Magnetic Resonance Selective Excitation Formulation

There are several formulations of the selective excitation problem that will fit the general structure above. We will describe three here: the minimum-distance, minimum-time, and minimum-energy. We utilize the discrete selective excitation problem (3) exclusively.

Minimum-Distance

Recall the discrete selective excitation problem (3). To get a very close fit to the desired magnetization, we might use

$$\phi(\boldsymbol{M}(T)) = \tfrac{1}{2}[\boldsymbol{M}(T) - \boldsymbol{D}]^T\,[\boldsymbol{M}(T) - \boldsymbol{D}],$$

$$= \tfrac{1}{2}[\|\boldsymbol{M}\|^2 + \|\boldsymbol{D}\|^2] - \mathbf{D}^T\boldsymbol{M}(T).$$

This represents the Euclidean squared-distance between the sampled desired magnetization and the distribution actually achieved by a control $\omega_1(t)$. Since the first two terms are constants of the control, we could instead set

$$\phi_d(\boldsymbol{M}(T)) = -\boldsymbol{D}^T\boldsymbol{M}(T). \tag{10}$$

This choice has numerical advantages in the computation of the adjoint equations. To ensure reasonable power consumption we could either limit the magnitude of the pulse, or we could trade off some slice definition for weaker pulses. The former is known as a hard constraint, the latter a soft constraint. Here we adopt the soft constraint, since it leads to simpler necessary conditions. We employ hard constraints in the minimum-time formulation. Hence, $L_d(\boldsymbol{M}, \omega_1, t)$ is

$$L_d(\boldsymbol{M}, \omega_1, t) = \tfrac{1}{2}\alpha\omega_1^2(t). \tag{11}$$

α is a relative weighting factor to penalize high energy pulses. In simulation, the user can start with α close to zero, and increase it until the pulse is sufficiently weak. The boundary conditions are

$$\boldsymbol{M}(0) = \boldsymbol{M}_0 \triangleq [0, 0, 1, \cdots, 0, 0, 1]^T,$$

$$\lambda^T(T) = \frac{\partial\phi_d}{\partial\boldsymbol{M}(T)} = -\boldsymbol{D}^T. \tag{12}$$

Note that the Lagrange multipliers will become the reflection of the magnetization vector when the algorithm converges. Substituting (3) and (11) into (6) we find that the Lagrange multipliers obey

$$\dot{\lambda}(t) = [A + B\omega_1(t)]\,\lambda(t). \tag{13}$$

This system is called self-adjoint since the Lagrange multipliers obey the same equation as the $\boldsymbol{M}(t)$. Finally, the stationarity condition (8) obtains

$$\lambda^T(t)\,\boldsymbol{BM}(t) + \alpha\omega_1(t) = 0. \tag{14}$$

Note that the optimal control is a filtered version of the state $\boldsymbol{M}(t)$.

Minimum-Energy

In this formulation, we ensure that the final state is achieved precisely,

$$\boldsymbol{M}(T) = \boldsymbol{D}. \tag{15}$$

We constrain the time interval to $(0, T)$ and minimize the energy expended to achieve (15). Hence,

$$\phi_e = \nu^T[\boldsymbol{M}(T) - \boldsymbol{D}],\ L_e(\boldsymbol{M}, \omega_1, t) = \tfrac{1}{2}\omega_1^2(t). \tag{16}$$

The unspecified vector ν is picked to satisfy (15). The necessary conditions for the minimum-energy formulation follow from (6)–(8).

$$\dot{\boldsymbol{M}} = [A + B\omega_1(t)]\,\boldsymbol{M}(t),$$

$$\dot{\lambda}(t) = [A + B\omega_1(t)]\,\lambda(t),$$

$$\boldsymbol{M}(0) = \boldsymbol{M}_0,\ \lambda(T) = \nu,$$

$$\lambda^T\boldsymbol{BM} + \omega_1(t) = 0. \tag{17}$$

Due to the skew symmetry of $[A, B]$ we again have a self-adjoint system.

Minimum-Time Formulation

In the minimum-time formulation we desire a pulse that achieves

$$\boldsymbol{M}(t_p) = \boldsymbol{D}$$

in the minimum time t_p. To constrain the pulse energy, we restrict our search to bounded controls. Hence, $L_t = 1$, $\phi_t = 0$, and we minimize

$$\min_{|\omega_1(t)| \le U} J[\omega_1(t)] = \int_0^{t_p} 1\,dt = t_p.$$

The stationarity equation (8) gives ambiguous results for this linear problem. Recall that a linear function $l(x) = ax + b$ does not satisfy the necessary condition for a minimum $dl/dx = 0$, for any x. Just as the extrema of a linear function lie at the endpoints of the allowed interval, the extremals of a linear functional are always found at the boundary $|\omega_1(t)| = U$. A formal theorem of Pontryagin [16] states that the optimal control must minimize

$$L + \lambda^T[A + B\omega_1(t)]\,\boldsymbol{M}(t).$$

It should be clear that the optimal bounded control is

$$\omega_1(t) = -U\ \mathrm{sgn}\ \{\lambda^T\boldsymbol{BM}\}$$

where $\mathrm{sgn}(x)$ is the sign function

$$\mathrm{sgn}(x) = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{if } x = 0. \\ -1 & \text{if } x < 0 \end{cases}$$

This type of control is appropriately referred to as "bang-bang," since the control jumps instantaneously from $+U$ to $-U$ when the switching function, $\lambda^T\boldsymbol{BM}$, changes sign. Because our initial development did not allow for variations in the end time, we require one additional equation. In addition to implicit variations through $\boldsymbol{M}(t)$, $\bar{J}[\omega_1(t)]$ experiences implicit variations through changes in t_p. It can be shown [17] that if the so-called *transversality* condition,

$$\lambda^T(t_p)\,[A + B\omega_1(t_p)]\,\boldsymbol{M}(t_p) = -1$$

is satisfied, then the equation (7) for the variation in $\bar{J}$ is still valid. The differential equations in $\boldsymbol{M}$ and λ are again given by (17).

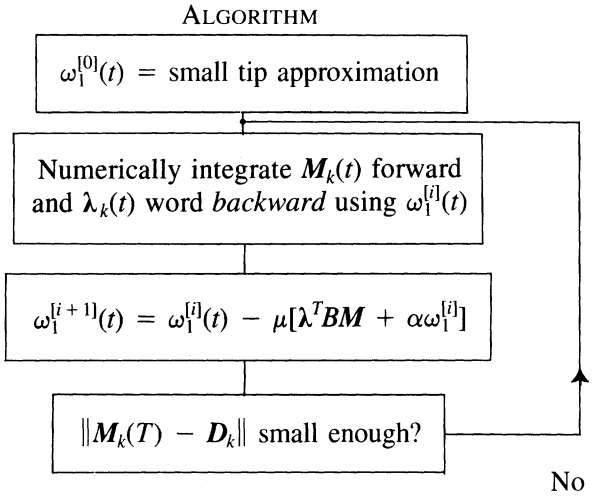
Comparison of the Formulations

Perhaps the most attractive formulation is the minimum-time, since it uses all the available energy and ostensibly affords the briefest pulse. This is also the hardest to implement, since algorithms for computing the optimal switching times are not yet fully developed. Furthermore, it is likely that the researcher is willing to trade off some slice definition for a shorter pulse. The rigid requirement of exactly matching the desired magnetization could lead

to very long "minimum-time" pulses. Bang-bang algorithms are currently being researched [22] and this approach may become realistic. The minimum-energy formulation also suffers from the rigid requirement on the final magnetization. Furthermore, the algorithm is much more computationally intensive than the minimum-distance algorithm. In fact, each ν update [17] requires the inverse of a $3K \times 3K$ matrix. Recall that K is the number of samples in the spatial domain. Since K should be in the hundreds for good resolution, this is a demanding computation. The minimum-distance formulation has produced promising results in simulation. The only drawback of the minimum-distance pulse is that the final profile has some error. This error was shown to be very small in the simulations.

IV. Simulation Results

Below is a symbolic flowchart of the algorithm used to solve the minimum-distance two-point boundary-value problem presented in Section III. All integrations were computed by solving the Bloch equation exactly for a piecewise constant approximation to the input [9]. Since the pulses will ultimately be piecewise constant when implemented with a digital-to-analog converter, this approximation is extremely accurate.



A. 90° Saturation Pulse

The 90° pulse design incorporated a fixed-duration refocusing gradient after the pulse. This alters the basic minimum-distance formulation only slightly;

$$\phi_{90} = -M^T(T)\, R^T D,$$

and

$$\lambda(T) = -R^T D.$$

The refocusing rotation matrix R represents the free precession of the extended magnetization vector under the influence of the negative gradient for a period of $T/2$. R is a block diagonal $3K \times 3K$ matrix with the ith entry

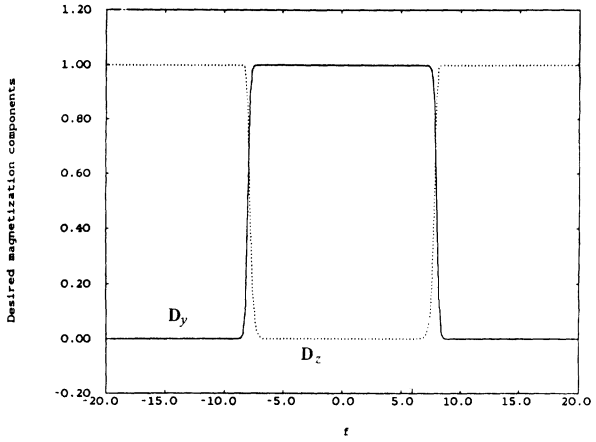


Fig. 1. Desired D_z, D_y distribution for the 90° pulse.

$$R_{i,i} = \begin{bmatrix} \cos(\pi f_i T) & -\sin(\pi f_i T) & 0 \\ \sin(\pi f_i T) & \cos(\pi f_i T) & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

The desired distribution for the 90° was shaped from a Fermi function for its soft shoulders and fast transitions. The desired response D is plotted continuously versus f in Fig. 1. f is slightly modified from that discussed in the paper. $f = \gamma G_z T/2\pi$. Caprihan [9] discusses the simple time scale change $\tau = t/T$ that leads to this variable. Note that all pertinent parameters of the pulse and the slice are embedded in the "sample" slice shown. We design the pulse over a normalized interval. If the user wishes to halve the slice thickness, then the product of the gradient field and the pulse time must be doubled. Essentially all optimal pulses are scaled versions of the pulses shown here. If one cannot extend the gradient or the excitation time further, then he should recompute a pulse with fewer "sinc lobes." To give a rough idea of the selectivity of these pulses, assume that we have $G = 0.5$ G/cm, T = 5 ms, then each unit of spatial frequency f corresponds to about 1 mm. The pulses plotted in this paper, $\omega_1(t)$, $t \in [0, 1]$, can be related to their experimental equivalent $B_1(t)$, $t \in [0, T]$ through the relation

$$B_1(t) = \frac{1}{(\gamma T)}\, \omega_1\left(\frac{t}{T}\right).$$

With the values given in the example above, $B_1(t) = (1/134)\, \omega_1(t/5)$ G, where t is measured in ms. After running the optimization routine for a few hundred iterations with μ about 0.1, we developed the magnetization profile shown in Fig. 2.

Comparison of Figs. 1 and 2 shows that the pulse is indeed quite effective. Fig. 3 presents both the initial (Fourier transform) pulse and the optimal pulse. Note that the Fourier transform pulse is quite accurate even for a 90° degree flip-angle. As one might expect, the optimal pulse resembles the one designed in [8].

B. 180° Inversion

Here we wish to tip the in-slice spins from [0, 0, 1] to the antiparallel position [0, 0, −1]. Since the significant

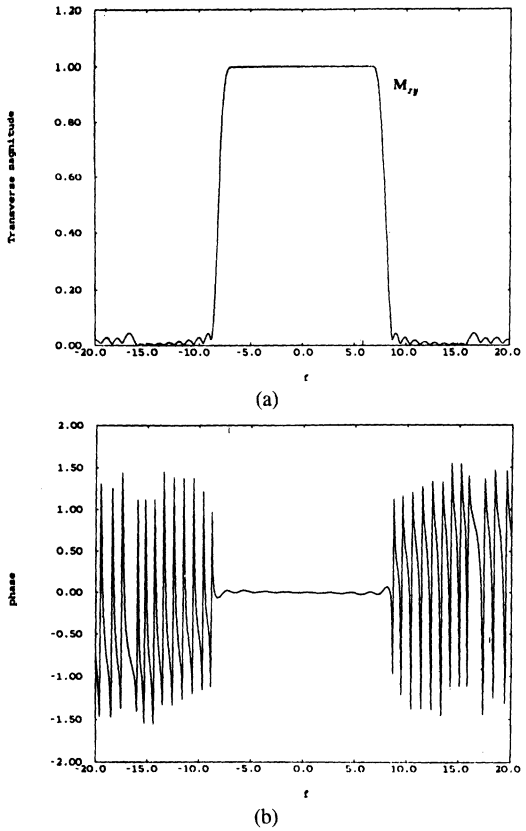


Fig. 2. Achieved 90° magnitude and phase profile.

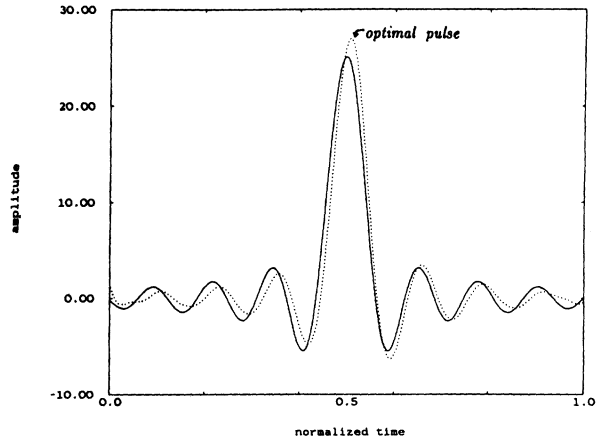


Fig. 3. Initial and optimal 90° pulses.

magnetization component is not transverse, no refocusing period is required. Hence, the minimum-distance equations need no change. The desired distribution is plotted in Fig. 4. The optimization routine required considerable amounts of time to find a pulse that achieved the distribution shown in Fig. 5. Typical μ values were 0.025. The pulse that achieved the distribution is shown in Fig. 6. Note that the pulse represents a significant departure from the sinc pulse.

C. Spin-Echo Generation: The "Pancake Flipper" Pulse

The popular 2D FT technique for generating selective spin-echoes usually employs the sequence 90°–τ–180°–

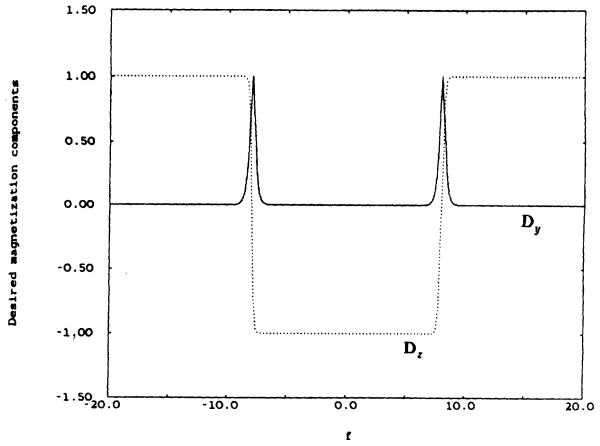


Fig. 4. Desired magnetization components at time T for 180°.

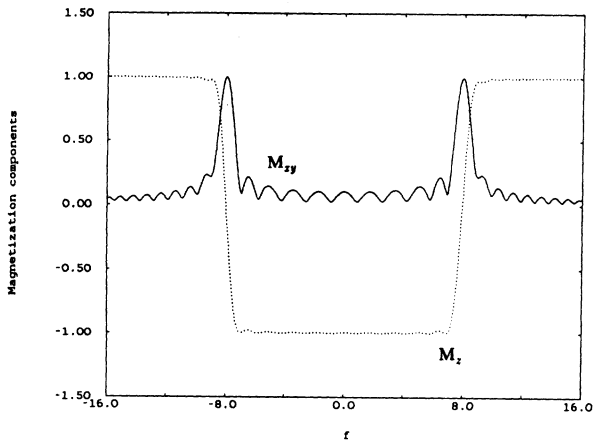


Fig. 5. Achieved 180° transverse and $\mathbf{M}_z$ magnetization profile.

τ. Due to the infidelity of conventional selective 180's, researchers have resorted to nonselective 180's which generate a very large FID from the imperfectly inverted spins in the out-of-slice region.

It is important to realize that an inversion pulse is not necessarily a good rotation about the x-axis, due to the asymmetry introduced by the gradient field. To demonstrate this we simulated the sequence 90°–180° with two inversion pulses. In Fig. 7(a) we simulated the optimal 90° followed by a sinc pulse with the proper area. In Fig. 7(b), we used the optimal inversion pulse designed above (Fig. 6). The results are plotted in Fig. 7. Clearly, one must design pancake pulses with attention to the phase of the resulting profile.

One approach to designing a selective rotation involves simultaneous optimization of the control for two independent profile simulations. Stated simply, one must ensure that the in-slice components initially along the x-axis are returned to the x-axis at the end of the pulse, and that in-slice components initially along the y-axis be flipped to the negative y-axis. Since the Bloch equation is linear in the initial condition, success on the two cases above guarantees a "pancake flip" response to any transverse initial conditions. Hence, we modified the optimal control program to handle a $6K \times 1$ dimensional extended magnetization vector. The two initial distributions correspond to

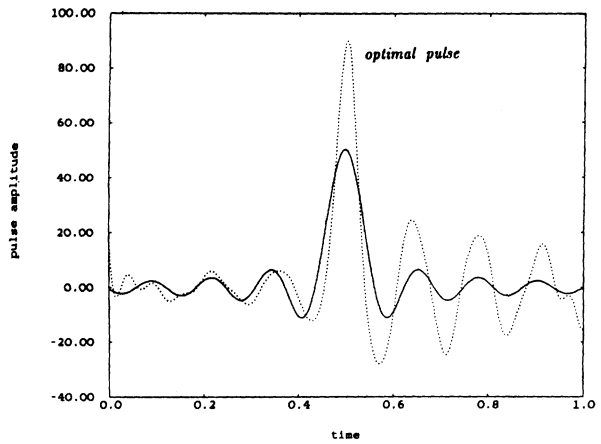


Fig. 6. Initial and optimal pulses for the 180° inversion.

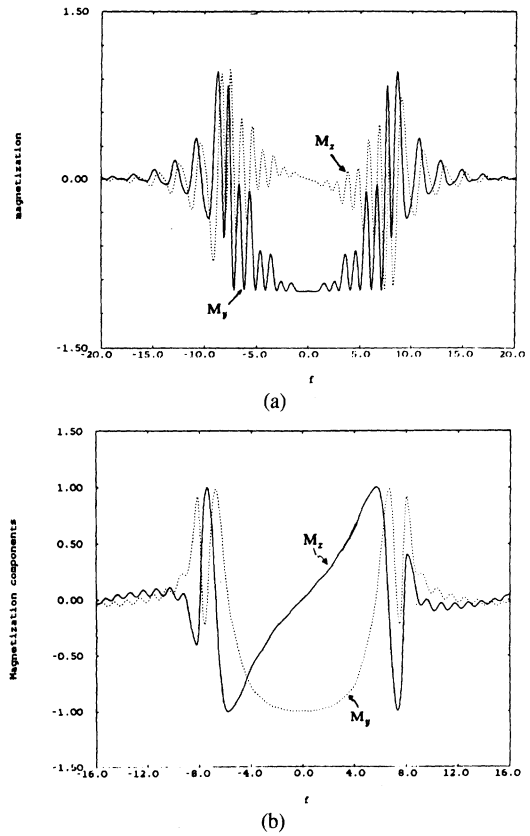


Fig. 7. Results of using inversion pulses as "pancake flipper."

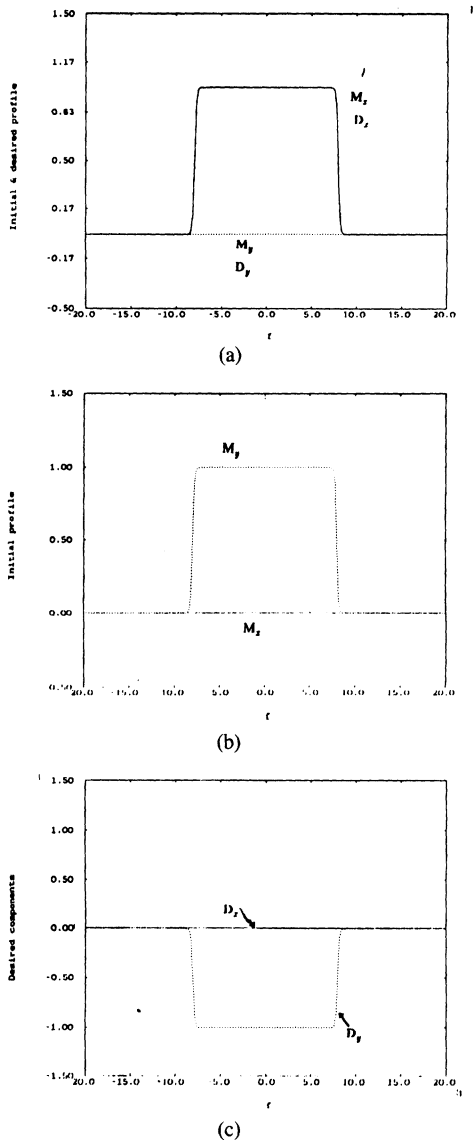


Fig. 8. Initial and desired profiles for 180° rotation. (a) Initial and desired profile along x. (b) Initial conditions along y. (c) Desired final profile for (b).

the final magnetization profile following a selective 90° and refocusing. The initial distributions are plotted in Fig. 8(a) and (b). The desired distributions, representing a perfect rotation about the x-axis, are plotted in Fig. 8(a) and (c). No refocusing was used. A sinc pulse with sufficient area to rotate on-resonance components by 180° was used as the initial control. The resulting magnetization profiles are plotted in Fig. 9. Fig. 10 compares the optimal control with the sinc pulse. Note the remarkable symmetry of the pulse, and its similarity in shape to the optimal 90° pulse shown in Fig. 3.

Offset Slices

As a final example, we tried to generate a thin-slice (2 mm if $G = 0.5$ G/cm, $T = 5$ ms) transverse 180° flip using offset slices [26]. The basic idea is to shift either the 90° profile or the 180° profile relative to the other. The effective slicewidth is approximately the intersection of the two regions. This is a control regime where an algorithm is far superior to Fourier transform design since the second pulse can be optimized to work on the initial distribution left by the first pulse. Also, the "shoulders" of the slices become of central importance rather than the center, and such tradeoffs could be of some advantage. Here we simply demonstrate the use of the optimal control pulses designed above in an offset sequence. Fig. 11 presents the results of the sequence. The first slice was offset by $f = 14$. In a sequence with focusing intervals, the 14 mm region to the right of the slice will never refocus. To the left of the slice, there will be a 14 mm inversion slice. This should not cause a large FID, since it generates transverse magnetization only at the fringe. Hence, the effective slice is the intersection of the two pulses the area between $f = 6$ and $f = 8$.

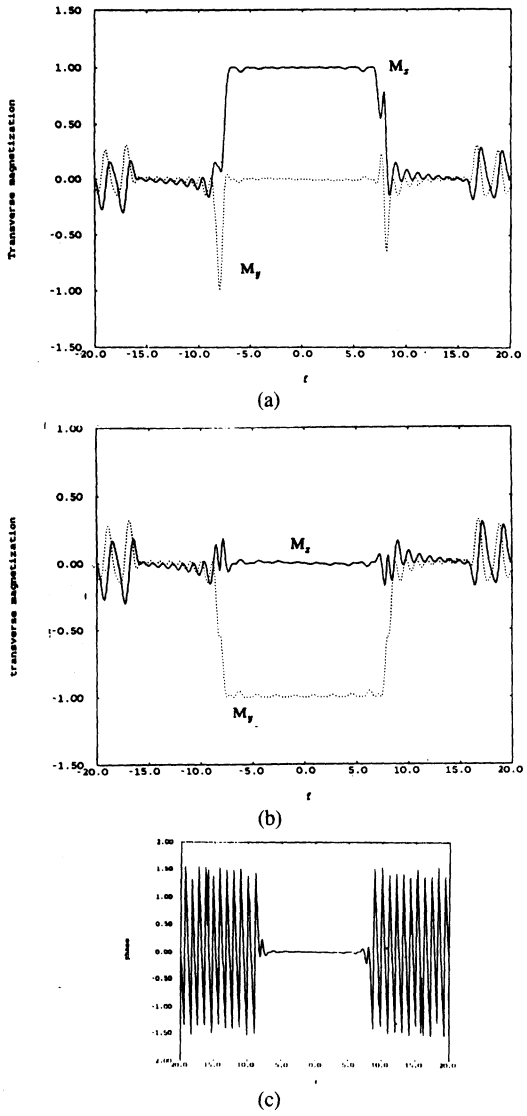


Fig. 9. The achieved magnetization profiles. (a) For initial conditions along x. (b) For initial conditions along y. (c) Phase response.

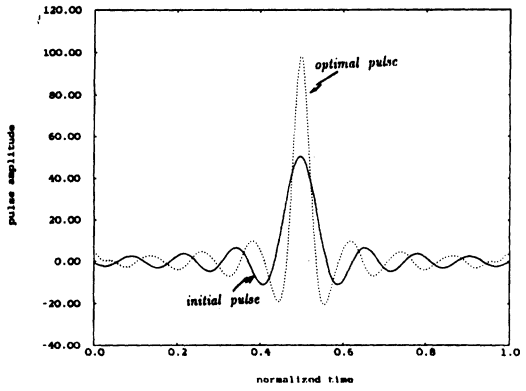


Fig. 10. Optimal and initial pancake flipper controls.

Discussion

After establishing that the magnetic resonance spin system is spatially controllable, we structured the selective excitation problem as a dynamic optimization problem. A mathematical basis for selective pulse design was introduced. The techniques of optimal control theory lend themselves immediately to the selective excitation problem. We argued that the most promising formulation is the minimum-distance. Simulation results demonstrate that the pulses are indeed quite selective. We intend to implement the pulses on a General Electric 1.5 T system in the near future. Although the pulses designed with optimal control theory require substantial computation time, the pulses need only be computed once. Furthermore, to halve the slicewidth, one simply doubles the duration-gradient product. Hence, all optimal 90° pulses have the same shape and need only be recomputed if one cannot push one's gradients or excitation time any more.

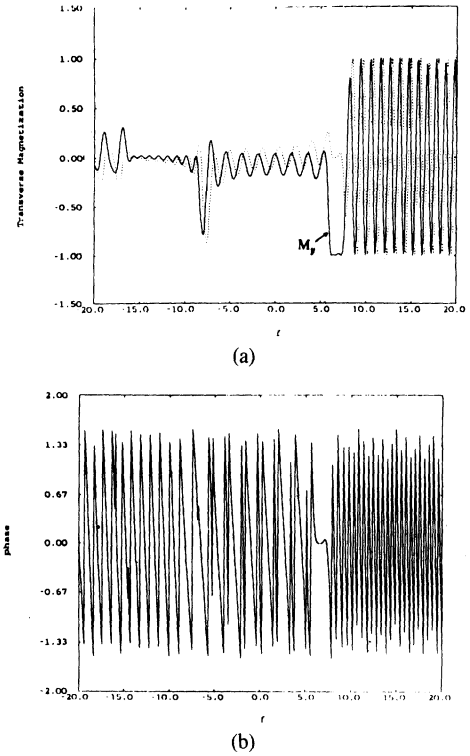


Fig. 11. Thin slice "offset" sequence response.

Although designed specifically for 2D FT experiments, the selective "pancake flipper" pulse might find application in the chemical shift imaging sequence proposed by Joseph [23]. As discussed in this paper, sequential pulses could be optimized to produce very high-resolution "offset" slices. Another application not addressed here is volume-selective excitation [27]. We hope to apply optimal control techniques to design a pulse capable of selectively tipping spins 90° along a constant z-gradient in the presence of time-varying x and y gradients [25]. Finally, optimal control theory generalizes to the multiple control system so that complex excitations could be incorporated into the control.

Perhaps most significant is the flexibility of the optimal control design approach. It allows for inversion of the Bloch equation for any desired distribution.

Appendix
Controllability of the MR Spin System

We adopt the notation of Jurdjevic and Sussman [21]. Throughout, G will refer to a Lie group, and $\mathbf{L}(G)$ will

denote the Lie algebra of G. A Lie algebra is a vector space that is a subset of the set of square matrices with real coefficients. It is defined as a vector space closed on a certain matrix operation. If A and B belong to $\mathbf{L}(G)$, then so does $[A, B] = AB - BA$. The matrix product above is called the Lie bracket. Given a few members of the algebra, one can complete the Lie algebra by recursively computing all the Lie brackets of the original members, and then the Lie brackets of the originals with the first brackets, and so on.

Consider the generic control system (termed "right-invariant") described by an evolution equation in a Lie group G of the form

$$\frac{dx(t)}{dt} = X_0(x(t)) + \sum_{i=1}^{m} u_i(t)\, X_i(x(t)) \qquad \text{(A.1)}$$

where $X_0, X_1, \cdots, X_m$ are right-invariant vector fields on G, and the u_i are piecewise-continuous control functions. We will refer to the system (A.1) as the control system $(\boldsymbol{X}, U)$. We now quote two pertinent theorems from [21].

Theorem A.1: A necessary condition for $(\boldsymbol{X}, U)$ to be controllable is that G be connected and that $\boldsymbol{L} = L(G)$. If G is compact, or if the system is homogeneous, the condition is also sufficient.

In Theorem A.1, $\boldsymbol{L}$ refers to the Lie algebra generated by $X_0, X_1, \cdots, X_m$.

Theorem A.2: Let G be compact, and let $(\boldsymbol{X}, U)$ be controllable. Then there exists $T > 0$ such that, for every $g \in G$, $g' \in G$, there is a control that steers g into g' in less than T units of time.

The central role of the Lie algebra, $\boldsymbol{L}$ and its Lie group is best understood by considering the homogeneous differential equation

$$\frac{dx}{dt} = u(t)\, Ax(t).$$

The solution is

$$x(t) = e^{U(t)A}x(0),$$

where $U(t) = \int_0^t u(\tau)\, d\tau$. This is one description of the space of reachable $x(t)$. An alternative description employs all finite products of exponentials of the form

$$x(t) = e^{At_s}e^{At_{s-1}} \cdots e^{At_1}x(0). \qquad \text{(A.2)}$$

One way to visualize this equivalence is to imagine sampling the control $U(t)$ at very fast intervals. Since the control is arbitrary, the t_i must also be allowed arbitrary value. The set of matrices generated according to (A.2) is exactly the Lie group associated with the trivial Lie algebra generated by A. In the general case of (A.1), the Lie algebra $\boldsymbol{L}$ generated by the X_i spans the space of possible dx/dt, and represents all the differential directions that $x(t)$ can evolve towards. The Lie group associated with $\boldsymbol{L}$ consists of all finite products of exponentials of the elements of $\boldsymbol{L}$. Since this group, which we denote by $\boldsymbol{S}$, is exactly the set of attainable x, the system is controllable if and only if $\boldsymbol{S} = G$. Equivalently, the system is controllable if $\boldsymbol{L}(\boldsymbol{X}) = L(G)$.

For the magnetic resonance selective excitation control system, the appropriate group G is $SO(3)^K$, where $SO(3)$ denotes the group of 3×3 orthogonal matrices with unity determinant, i.e., proper rotation matrices. That is, we wish to send each spatial isochromat to a desired location on its sphere, independent of the final destination of other isochromats. Of course, the ability to do this will depend on the strength of the gradient and the fineness of sampling. Note that G is both compact and connected. A basis for the Lie algebra associated with G is the set of $3K \times 3K$ block diagonal skew-symmetric matrices with only one nonzero block element. A representative of the $3K$ members of the set is

$$\text{DIAGONAL } \{0, 0, \cdots, \boldsymbol{K}_x, 0, \cdots, 0\}$$

where $\boldsymbol{K}_x$ is one of the three basis skew-symmetric matrices.

$$\boldsymbol{K}_x = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix}, \quad \boldsymbol{K}_y = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix},$$

$$\boldsymbol{K}_z = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Recall that there is an exponential map relating the Lie algebra to its group G. Since the exponential of a skew-symmetric matrices is a rotation matrix, the skew-symmetric matrices form a complete basis[2] for $L\{SO(3)^K\}$. Note that there are exactly $3K$ linearly independent skew-symmetric matrices in *any* basis for $\mathbf{L}(G)$.

We now show that the Lie algebra generated by $[A, B]$ of (3) also contains $3K$ linearly independent skew-symmetric matrices. First, we assume that the spatial sampling is uniform with period ϵ so that

$$f_i = i\epsilon.$$

Using the Lie bracket operation, one can construct the Lie algebra, $\mathbf{L}(A, B)$. $A_i = i\epsilon \boldsymbol{K}_z$, $B_i = \boldsymbol{K}_x$. Since $[\boldsymbol{K}_x, \boldsymbol{K}_y] = \boldsymbol{K}_z$, $[\boldsymbol{K}_y, \boldsymbol{K}_z] = \boldsymbol{K}_x$, and $[\boldsymbol{K}_z, \boldsymbol{K}_x] = \boldsymbol{K}_y$, the Lie bracket is isomorphic to the cross product of $\boldsymbol{R}^3$. The ith block of $\boldsymbol{L}(A, B)$ is spanned by

$$\{\boldsymbol{K}_x, (i\epsilon)^2\boldsymbol{K}_x, \cdots, (i\epsilon)^{2K-2}\boldsymbol{K}_x, (i\epsilon)\,\boldsymbol{K}_y, (i\epsilon)^3\boldsymbol{K}_y, \cdots,$$

$$, \cdots, (i\epsilon)^{2K-1}\boldsymbol{K}_y, (i\epsilon)\,\boldsymbol{K}_z, (i\epsilon)^3\boldsymbol{K}_z, \cdots, (i\epsilon)^{2K-1}\boldsymbol{K}_z\}.$$

These block-diagonal matrices are independent if and only if the diagonal "multipliers" are independent. For example, the $\boldsymbol{K}_x$ terms are independent if and only if the $K \times K$ Vandermonde matrix of the diagonal multipliers is nonsingular.

$$V_{\boldsymbol{K}_x} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ \epsilon^2 & (2\epsilon)^2 & \vdots & (K\epsilon)^2 \\ \vdots & \vdots & \vdots & \vdots \\ \epsilon^{2K-2} & (2\epsilon)^{2K-2} & \cdots & (K\epsilon)^{2K-2} \end{bmatrix}$$

[2] $e^{\theta \boldsymbol{K}_x}$ is a rotation about the x-axis by θ.

If the sampling is done nonuniformly, the ith column of the Vandermonde matrix is replaced by $[1 f_i^2 \cdots f_i^{2K-2}]^T$.

The Vandermonde matrices associated with $\boldsymbol{K}_y$ and $\boldsymbol{K}_z$ are

$$V_{\boldsymbol{K}_y} = V_{\boldsymbol{K}_x} = \begin{bmatrix} \epsilon & \epsilon & \cdots & \epsilon \\ \epsilon^3 & (2\epsilon)^3 & \vdots & (K\epsilon)^3 \\ \vdots & \vdots & \vdots & \vdots \\ \epsilon^{2K-1} & (2\epsilon)^{2K-1} & \cdots & (K\epsilon)^{2K-1} \end{bmatrix}.$$

All the Vandermonde matrices above are nonsingular as long as ϵ is not zero. For the general sampling case, nonsingularity is guaranteed if $f_i > 0$ for all i. This restriction is simply a manifestation of the Hermitian symmetry of the transverse magnetization imposed by amplitude modulation. Hence, there are $3K$ linearly independent skew-symmetric matrices in $\boldsymbol{L}(A, B)$. By Theorem A.1, the MR control system is controllable, and by Theorem A.2, all the states in $SO(3)^K$ can be reached in less than T units of time. Note that we do not know what T is, we simply know it is finite. We know from perturbation analysis that T should be inversely porportional to the bandwidth of the desired profile. In a practical sense though, we have mild spatial dependence since the Vandermonde matrix is notably ill-conditioned [24].

If complex excitation is employed, we should compute the Lie algebra generated by $\{i\epsilon \boldsymbol{K}_z, \boldsymbol{K}_x, \boldsymbol{K}_y\}$. Instead of alternating powers of ϵ, the ith column of all three Vandermonde matrices is $[1, (i\epsilon), (i\epsilon)^2, \cdots, (i\epsilon)^{K-1}]^T$. Note that this matrix is guaranteed nonsingular; there is no restriction to Hermitian profiles. Furthermore, the control problem is better conditioned.

REFERENCES

[1] S. Conolly and A. Macovski, "Selective excitation via optimal control theory," in *Proc. 4th Meet. Soc. Magn. Reson. Med.*, 1985, p. 958.

[2] A. Garroway, P. Grannell, and P. Mansfield, "Image formation in NMR by a selective irradiative process," *J. Phys. C*, vol. 7, pp. L457-L462, 1974.

[3] R. Sutherland and J. Hutchinson, "Three-dimensional nmr imaging using selective excitation," *J. Phys. E*, vol. 11, pp. 79-83, 1978.

[4] D. Hoult, "The solution of the Bloch equations in the presence of a varying B1 field—An approach to selective pulse analysis," *J. Magn. Reson.*, vol. 35, pp. 69-86, 1979.

[5] P. Mansfield, A. Maudsley, P. Morris, and I. Pykett, "Selective pulses in nmr imaging: A reply to criticism," *J. Magn. Reson.*, vol. 33, pp. 261-274, 1979.

[6] P. Locher, "Computer simulation of selective excitation in nmr imaging," *Phil. Trans. R. Soc. Lond. B.*, vol. 289, pp. 537-542, 1980.

[7] P. Joseph, Axel, and M. O'Donnell, "Potential problems with selective excitation pulses in nmr imaging systems," in *Proc. 2nd Meeting Soc. Magn. Reson. Med.*, 1983, p. 182.

[8] W. Loeffler, A. Oppelt, and D. Faul, "Computer simulations of slice selection in nmr imaging," in *Proc. 2nd Meet. Soc. Magn. Reson. Med.*, 1983, pp. 196-197.

[9] A. Caprihan, "Effect of amplitude modulation on selective excitation in nmr imaging," *IEEE Trans. Med. Imaging*, vol. MI-2, pp. 169-175, Dec. 1983.

[10] M. Silver, R. Joseph, and D. Hoult, "Selective pulse creation by inverse solution of the Bloch–Riccati equation," in *Works in Prog., Proc. 2nd Meet. Soc. Magn. Reson. Med.*, 1983, p. 22.

[11] ——, "Highly selective $\pi/2$ and π pulse generation," *J. Magn. Reson.*, vol. 59, pp. 347-351, 1984.

[12] D. Lurie, "A systematic design procedure for selective pulses in nmr imaging," *Magn. Reson. Imaging*, vol. 3, pp. 235-243, 1985.

[13] D. Nishimura, "A multipulse sequence for improved selective excitation in magnetic resonance imaging," *Med. Phys.*, vol. 12, pp. 413-418, 1985.

[14] A. Lent and A. Kritzer, "A new RF pulse shape for narrow-band inversion: The WOW-180," in *Proc. 4th Meet. Soc. Magn. Reson. Med.*, 1985, p. 1015.

[15] I. Gelfand and S. Fomin, *Calculus of Variations*. Englewood Cliffs, NJ: Prentice-Hall, 1963.

[16] L. Pontryagin, B. Boltyanskii, R. Gamkrelidze, and E. Mishchenko, *The Mathematical Theory of Optimal Processes*. New York: Wiley-Interscience, 1962.

[17] A. Bryson, Jr. and Y.-C. Ho, *Applied Optimal Control*, Washington, DC: Hemisphere, 1975.

[18] D. Luenberger, *Introduction to Linear and Nonlinear Programming*. Reading, MA: Addison-Wesley, 1973.

[19] T. Kailath, *Linear Systems*. Englewood Cliffs, NJ: Prentice-Hall, 1980.

[20] C. Bruni, G. DiPillo, and G. Koch, "Bilinear systems: An appealing class of 'nearly linear' systems in theory and applications," *IEEE Trans. Automat. Cont.*, vol. AC-19, pp. 334-348, 1974.

[21] V. Jurdjevic and H. Sussman, "Control systems of lie groups," *J. Diff. Equations*, vol. 12, pp. 313-329, 1972.

[22] A. Weinreb, "Optimal control with multiple bounded inputs," Ph.D. dissertation, Dep. Elec. Eng., Stanford Univ., Stanford, CA, 1984.

[23] P. Joseph, "A spin echo chemical shift MR imaging technique," *J. Comp. Assist. Tomogr.*, vol. 4, pp. 651-658, 1985.

[24] W. Gautschi, "Norm estimates for inverses of Vandermonde matrices," *Numer. Math.*, vol. 23, pp. 337-347, 1975.

[25] A. Macovski, "Volumetric NMR imaging with time-varying gradients," *J. Mag. Res. Med.*, vol. 2, pp. 29-39, 1985.

[26] L. Crooks, *et al.* "Thin definition in magnetic resonance imaging," *Radiology*, vol. 154, pp. 463-467, 1985.

[27] W. Aue, S. Müller, T. Cross, and J. Seelig, "Volume selective excitation. A novel approach to topical NMR," *J. Mag. Reson.*, vol. 56, pp. 350-354, 1984.