Improving the Law-Eberly Control Scheme
(Optimal control meets d'Alembert)

Institut Henri Poincaré

December 8, 2010

Roger Brockett
Harvard University

What I Hope to Describe

A short description of the Law-Eberly controllability result

A technique for establishing estimates on optimal transfer times

A related problem involving systems with additional symmetries.

The use of modes to study time-optimal paths on a sphere

An interpretation of optimality in terms of wave speed

The possibility of a “turnpike theorem” for time optimal control on the sphere

Outline

1. The Law-Eberly model for a harmonic oscillator with spin
2. Controllability via overlapping invariant subspaces
3. A Liapunov-type lower bound on minimum transfer time
4. Lattice models, waves, wave speed and dispersion
5. Some long chain asymptotic estimates on optimal transfer

Part 1: The work of Law and Eberly

The harmonic oscillator

As is well known, the controlled Schrodinger equation

$$ih\frac{\partial}{\partial t}\psi = \frac{1}{2}\left(-\frac{\partial^2}{\partial x^2} + (x-u)^2\right)\psi(t,x)$$

fails to be controllable because the Lie algebra generated by

$\frac{\partial^2}{\partial x^2} - x^2$ and x is only four dimensional.

$$\mathcal{L} = \alpha\left(\frac{\partial^2}{\partial x^2} - x^2\right) + \beta x + \gamma\frac{\partial}{\partial x} + \delta$$

This is also the Lie algebra of the simplest Kalman filter and in that setting this fact is a blessing.

Two basic points

What is a $\pi/2$ pulse?

$$\frac{d}{dt}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 & u \\ -u & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Then if

$$\int_a^b u(t)\, dt = -\pi/2$$

Then

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} \mapsto \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Two basic points

For distinct frequencies in $e^{-At}Be^{At}$ we have distinct vector fields. For $\dot{x} = (A + uB)x$ we can always introduce $z = e^{-At}x$ so that

$$\dot{z} = u(t)e^{-At}Be^{At}z(t)$$

If e^{At} is a sum of sinusoids then by choosing u to be sinusoidal we can approximate the solution by replacing $e^{-At}Be^{At}$ by

$$B(u) = \tfrac{1}{T}\int_0^T u(t)e^{-At}Be^{At}\,dt$$

This approximation is to be accepted without question.

Harmonic oscillator + spin

From one point of view, the lack of controllability of the harmonic oscillator is a consequence of the lack of distinct frequencies in $e^{-At}Be^{At}$. If the Hamiltonian is changed, for example by a field that couples to spin, then $ad_A(\cdot)$ can gain additional eigenvalues, and $e^{-At}Be^{At}$ additional frequencies.

For the problem of steering the system

From Bloch, B. & Rangan

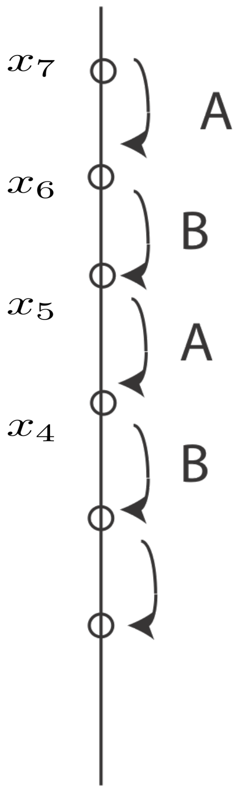
$$\dot{Y} = (u(t)B_r + v(t)B_b)Y,$$

where $u(t)$ and $v(t)$ are defined as before. Then

$$B_r \quad = \quad \left(\begin{array}{c|cc|cc|cc}
0 & 0 & 0 & 0 & 0 & 0 & \dots \\
\hline
0 & 0 & L_0^{(1)}(\eta^2) & 0 & 0 & 0 & \dots \\
0 & -L_0^{(1)}(\eta^2) & 0 & 0 & 0 & 0 & \dots \\
\hline
0 & 0 & 0 & 0 & L_1^{(1)}(\eta^2) & 0 & \dots \\
0 & 0 & 0 & -L_1^{(1)}(\eta^2) & 0 & 0 & \dots \\
\hline
0 & 0 & 0 & 0 & 0 & 0 & \dots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots
\end{array}\right). \quad (1)$$

$$B_b = \imath \left(\begin{array}{cc|cc|cc|c}
0 & 0 & 0 & L_0^{(1)}(\eta^2) & 0 & 0 & \dots \\
0 & 0 & 0 & 0 & 0 & 0 & \dots \\
\hline
0 & 0 & 0 & 0 & 0 & L_1^{(1)}(\eta^2) & \dots \\
-L_0^{(1)}(\eta^2) & 0 & & 0 & 0 & 0 & \dots \\
\hline
0 & 0 & 0 & 0 & 0 & & \dots \\
0 & 0 & -L_1^{(1)}(\eta^2) & 0 & & 0 & \dots \\
\hline
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots
\end{array}\right).$$

Rotate within an invariant subspace until the solution lies in the overlap and then switch to the motion which produces a rotation in the second family of invariant subspaces



Illustrating the effect of successive $\pi/2$ pulses from alternating controls

The time optimal problem

- The scheme just described is suitable for establishing controllability but appears to be slower than necessary in that it uses the controls to follow segments lying on a series of great circles which are perpendicular in the sense of spherical geometry. This gives a speed of $2/\pi$ per step.

  What about the time optimal problem? (2 per step.)

Part 2: The simplified models

First of Three Simplified Models: **Toeplitz**

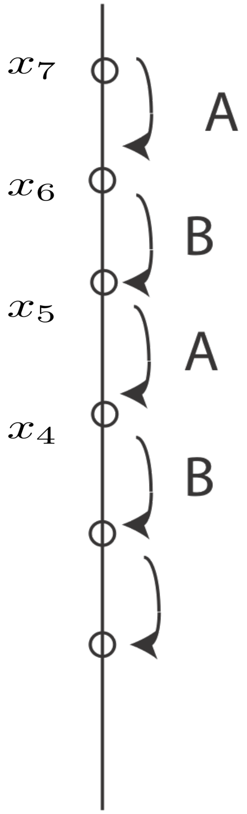
Consider $\dot{x} = (uA + vB)x$ where

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & \dots \\ -1 & 0 & 0 & 0 & 0 & 0 & \dots \\ 0 & 0 & 0 & 1 & 0 & 0 & \dots \\ 0 & 0 & -1 & 0 & 0 & 0 & \dots \\ 0 & 0 & 0 & 0 & 0 & 1 & \dots \\ 0 & 0 & 0 & 0 & -1 & 0 & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \end{bmatrix} \; ; \quad B = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & \dots \\ 0 & 0 & 1 & 0 & 0 & 0 & \dots \\ 0 & -1 & 0 & 0 & 0 & 0 & \dots \\ 0 & 0 & 0 & 0 & 1 & 0 & \dots \\ 0 & 0 & 0 & -1 & 0 & 0 & \dots \\ 0 & 0 & 0 & 0 & 0 & 0 & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \end{bmatrix}$$

The essential point here is that while the operators have many invariant subspaces, for each of the operators these are orthogonal but from one operator to the other, they are linked.

For such systems Law-Eberly's construction shows that any finitely nonzero vector can be transfered to the ground state in k steps if k is the index of the component of x that for which $x_k \neq 0$ and $x_j = 0$ for $j > k$.

There is a ground state. A+B is Toeplitz and $\Sigma(A+B)\Sigma = A-B$



Illustrating the effect of successive $\pi/2$ pulses from alternating controls

$$\Sigma = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & \dots \\ 0 & 1 & 0 & 0 & 0 & 0 & \dots \\ 0 & 0 & -1 & 0 & 0 & 0 & \dots \\ 0 & 0 & 0 & -1 & 0 & 0 & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \end{bmatrix}$$

Second Model: Laurent

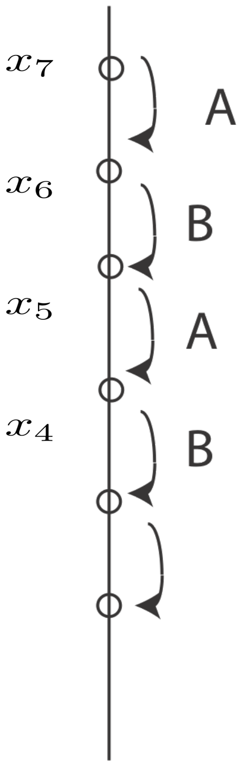
Consider $\dot{x} = (uA + vB)x$ where

$$A = \begin{bmatrix} \dots & \dots & \dots & \dots & \dots & \dots & \dots & \\ \dots & 0 & 1 & 0 & 0 & 0 & 0 & \dots \\ \dots & -1 & 0 & 0 & 0 & 0 & 0 & \dots \\ \dots & 0 & 0 & 0 & 1 & 0 & 0 & \dots \\ \dots & 0 & 0 & -1 & 0 & 0 & 0 & \dots \\ \dots & 0 & 0 & 0 & 0 & 0 & 1 & \dots \\ \dots & 0 & 0 & 0 & 0 & -1 & 0 & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \end{bmatrix} \; ; \; B = \begin{bmatrix} \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & 0 & 0 & 0 & 0 & 0 & 0 & \dots \\ \dots & 0 & 0 & 1 & 0 & 0 & 0 & \dots \\ \dots & 0 & -1 & 0 & 0 & 0 & 0 & \dots \\ \dots & 0 & 0 & 0 & 0 & 1 & 0 & \dots \\ \dots & 0 & 0 & 0 & -1 & 0 & 0 & \dots \\ \dots & 0 & 0 & 0 & 0 & 0 & 0 & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \end{bmatrix}$$

The essential point here is that while the operators have many invariant subspaces, for each of the operators these are orthogonal but from one operator to the other, they are linked.

For such systems Law-Eberly's construction shows that any standard basis vector can be transferred to any other standard basis vector in k steps where the initial and final vectors are e_i and e_{i+k}.

No ground state. A+B is Laurent and $\Sigma(A+B)\Sigma = A-B$.



Law-Eberly scheme no longer establishes complete controllability because there is no ground state. Standard basis elements can be reached from standard basis elements.

$$\Sigma = \begin{bmatrix} \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & 1 & 0 & 0 & 0 & \cdots \\ \cdots & 0 & 1 & 0 & 0 & \cdots \\ \cdots & 0 & 0 & -1 & 0 & \cdots \\ \cdots & 0 & 0 & 0 & -1 & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \end{bmatrix}$$

Third Model: Circulant with n even

Again, $\dot{x} = (uA + vB)x$

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & \ldots & 0 \\ -1 & 0 & 0 & 0 & 0 & \ldots & 0 \\ 0 & 0 & 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & -1 & 0 & 0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & 0 & \ldots & 0 & 1 \\ 0 & 0 & 0 & 0 & \ldots & -1 & 0 \end{bmatrix} \; ; \; B = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & \ldots & -1 \\ 0 & 0 & 1 & 0 & 0 & \ldots & 0 \\ 0 & -1 & 0 & 0 & 0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & 0 & 1 & 0 \\ 0 & 0 & 0 & \ldots & -1 & 0 & 0 \\ 1 & 0 & 0 & \ldots & 0 & 0 & 0 \end{bmatrix}$$

Many invariant subspaces, etc. The Law Eberly construction is only able to show that any basis vector e_i can be transferred to any other basis vector e_j. For large n and $x(0)$ having finite support, the circulant model and the Toeplitz model are more-or-less the same.

No ground state. A+B is circulant and $\Sigma(A+B)\,\Sigma = A - B$

Third Model: Circulant

Law-Eberly scheme no longer establishes complete controllability because there is no ground state. Standard basis elements can be reached from standard basis elements

Complete controllability must be studied by other means.

Part 3: Probability flow estimates

The starting point

Let $A = -A^T$ be a tridiagonal operator;

$$A(u,v) = \begin{bmatrix} 0 & a_1 & 0 & 0 & 0 & 0 & \dots \\ -a_1 & 0 & a_2 & 0 & 0 & 0 & \dots \\ 0 & -a_2 & 0 & a_3 & 0 & 0 & \dots \\ 0 & 0 & --a_3 & 0 & a_4 & 0 & \dots \\ 0 & 0 & 0 & -a_4 & 0 & a_5 & \dots \\ 0 & 0 & 0 & 0 & -a_5 & 0 & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \end{bmatrix}$$

Then $\phi(x) x^T [N, A] x$ changes as

$$\dot{\phi} = -2(a_1 x_1 x_2 + a_2 x_2 x_3 + a_3 x_3 x_4 + \cdots$$

Getting a bound

Using the fact that x is a unit vector this allows us to put a lower bound on the time required to transfer the state from e_i to e_j assuming that there are magnitude constraints on the a_i.

$$T \geq (e_i^T N e_i - e_j^T N e_j)/\lambda_m$$

where λ_m is the largest eigenvalue of $[A, N]$. Rather than present details for the spin coupled oscillator we consider a prototype with simpler notation.

Part 4: Lattice models and waves

A Lattice model

Toepltz vs. circulant and Laurent systems.
Consider $\dot{x} = A(u,v)x$ with A(u,v) skew-symmetric

$$A(u,v) = \begin{bmatrix} 0 & u & 0 & 0 & 0 & 0 & \ldots \\ -u & 0 & v & 0 & 0 & 0 & \ldots \\ 0 & -v & 0 & u & 0 & 0 & \ldots \\ 0 & 0 & -u & 0 & v & 0 & \ldots \\ 0 & 0 & 0 & -v & 0 & u & \ldots \\ 0 & 0 & 0 & 0 & -u & 0 & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \end{bmatrix}$$

Depending on the support of $x(0)$ may consider A to be part of a Laurent operator or to be completed as a circulant matrix.

A Lattice model

In the circulant case the Law-Eberley argument makes it clear that is clear that any basis vector e_i can be reached from any other basis vector e_j using $A(1,0)$ and $A(0,1)$. E.g., when n =5

$$A(1,1) = \begin{bmatrix} 0 & 1 & 0 & 0 & -1 \\ -1 & 0 & 1 & 0 & 0 \\ 0 & -1 & 0 & 1 & 0 \\ 0 & 0 & -1 & 0 & 1 \\ 1 & 0 & 0 & -1 & 0 \end{bmatrix}$$

The controllability question can be established by other means.

Part 5: Asymptotic analysis

Recall from before

Consider $\dot{x} = (uA + vB)x$ where

$$A = \begin{bmatrix} \dots & \dots & \dots & \dots & \dots & \dots & \dots & \\ \dots & 0 & 1 & 0 & 0 & 0 & 0 & \dots \\ \dots & -1 & 0 & 0 & 0 & 0 & 0 & \dots \\ \dots & 0 & 0 & 0 & 1 & 0 & 0 & \dots \\ \dots & 0 & 0 & -1 & 0 & 0 & 0 & \dots \\ \dots & 0 & 0 & 0 & 0 & 0 & 1 & \dots \\ \dots & 0 & 0 & 0 & 0 & -1 & 0 & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \end{bmatrix} \; ; \; B = \begin{bmatrix} \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & 0 & 0 & 0 & 0 & 0 & 0 & \dots \\ \dots & 0 & 0 & 1 & 0 & 0 & 0 & \dots \\ \dots & 0 & -1 & 0 & 0 & 0 & 0 & \dots \\ \dots & 0 & 0 & 0 & 0 & 1 & 0 & \dots \\ \dots & 0 & 0 & 0 & -1 & 0 & 0 & \dots \\ \dots & 0 & 0 & 0 & 0 & 0 & 0 & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \end{bmatrix}$$

The essential point here is that while the operators have many invariant subspaces, for each of the operators these are orthogonal but from one operator to the other, they are linked.

For such systems Law-Eberly's construction shows that any standard basis vector can be transferred to any other standard basis vector in k steps where the initial and final vectors are e_i and e_{i+k}.

No ground state. A+B is Toplitz

Formulating a regular lattice problem

$$A(u,v) = \begin{bmatrix} 0 & u & 0 & 0 & 0 & 0 & ... \\ -u & 0 & v & 0 & 0 & 0 & ... \\ 0 & -v & 0 & u & 0 & 0 & ... \\ 0 & 0 & -u & 0 & v & 0 & ... \\ ... & ... & ... & ... & ... & ... & ... \end{bmatrix}$$

When $u = v = 1$ this is "locally Laurent" and when $u = -v = 1$ it is similar to a locally Laurent.

$$A(1,-1) = \Sigma A(1,1) \Sigma.$$

$$\Sigma = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & ... \\ 0 & 1 & 0 & 0 & 0 & 0 & ... \\ 0 & 0 & -1 & 0 & 0 & 0 & ... \\ 0 & 0 & 0 & -1 & 0 & 0 & ... \\ ... & ... & ... & ... & ... & ... & ... \end{bmatrix}$$

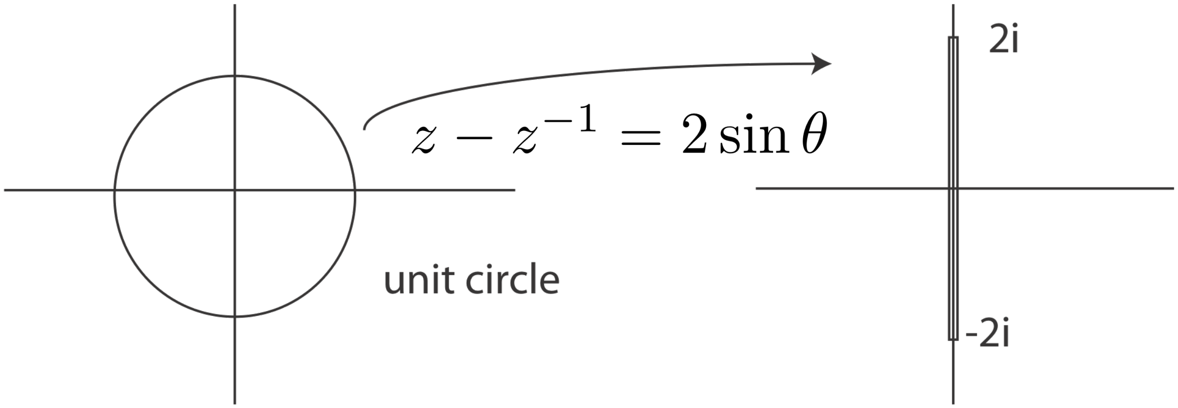
Because u and v enter the appropriate Hamiltonian linearly, if $|u| \leq 1$ and $|v| \leq 1$ then either we have singular arcs or else both u and v take on their extreme values at all time. Thus we would be looking for piecewise exponential solutions

$$x(T) = e^{At_1} e^{Bt_2} \cdots e^{At_k} x(0)$$

We also know that $B = \Sigma A \Sigma$ and so this can also be written in terms of A and $\Sigma = \Sigma^{-1} = \Sigma^T$

$$x(T) = e^{At_1} \Sigma e^{At_2} \Sigma \cdots e^{At_k} x(0)$$

The Laurent operator can be identified with multiplication by $z - z^{-1}$ as in $\mathbb{H}_2(D)$ theory. The spectrum of the Laurent operator is the image of the unit circle under the map $z \to z - z^{-1}$. For circulant matrices the domain is restricted to the nth roots of unity.



The wave equation provides intuition

If we pass from $\dot{x} = A(u,v)x$ with A skew-adjoint to

$$\ddot{x} = A^2(u,v)x$$

we are then working with a self adjoint negative semi-definite operator. This equation has the character of a wave equation. Note $(z - z^{-1})^2 = z^2 - 2 + z^{-2}$.

$$A^2(1,1) = \begin{bmatrix} \ldots & -2 & 0 & 1 & 0 & 0 & 0 & \ldots \\ \ldots & 0 & -2 & 0 & 1 & 0 & 0 & \ldots \\ \ldots & 1 & 0 & -2 & 0 & 1 & 0 & \ldots \\ \ldots & 0 & 1 & 0 & -2 & 0 & 1 & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \end{bmatrix}$$

If there are boundary conditions it is dispersive. Note that for $A(1,1)$ and $||x|| = 1$ we have

$$\frac{d}{dt} x^T N x \leq |\lambda_m|(A) = 2$$

The d'Alembert traveling waves

From the point of view of the wave equation

$$\left(\frac{\partial^2}{\partial t^2}-\frac{\partial^2}{\partial z^2}\right)\psi(t,z)=\left(\frac{\partial}{\partial t}+\frac{\partial}{\partial z}\right)\left(\frac{\partial}{\partial t}-\frac{\partial}{\partial z}\right)\psi(t,z)$$

It is natural to try to identify the wave traveling to the left and the wave traveling to the right. The maximum wave speed would be the square root of the largest eigenvalue of A^2. In the case of the Laurent operator the "media" is not dispersive but the other cases correspond to dispersive media–a pulse will not maintain its shape.

For the problem of steering the system

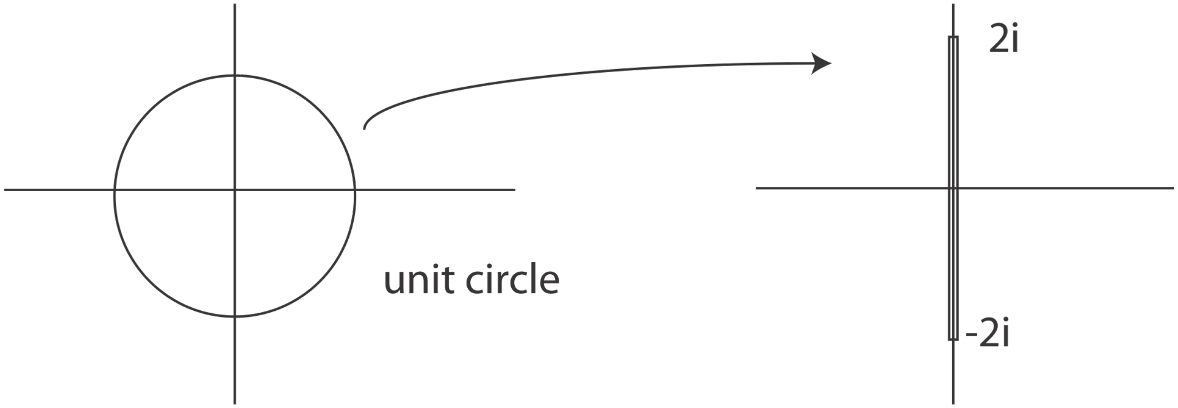
$$\dot{x} = (uA + vB)x \ ; \ |u| \leq 1 \ |v| \leq 1$$

from e_1 to e_k in minimum time, the policy $u = 1$ and $v = 1$ is optimal for certain initial and final states and the average time required corresponds to a wave speed of 2. Compare with $2/\pi$.

Traveling waves: circulant case

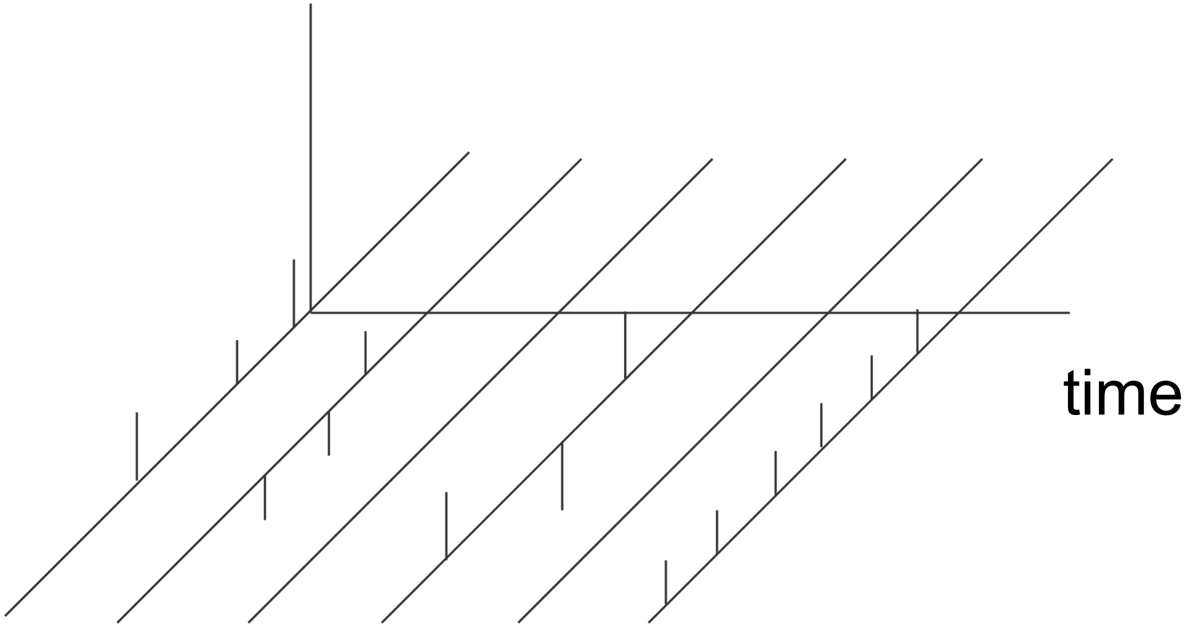
The Laurent operators only have approximate point spectrum whereas circulant operators have eigenvalues. The eigenvectors corresponding to the largest eigenvalues have the largest wave speed. They occur in pairs if $n \in 4\mathbb{Z}$. In terms of the wave picture, they may be forward moving or backwards moving as in d'Alembert's solution of the wave equation.
The eigenvectors corresponding to these eigenvalues have "large support". E.g., for $n = 8$ we have $[1, 0, -1, 0, 1, 0, -1, 0]$ and $[0, 1, 0, -1, 0, 1, 0, -1]$ for $\lambda = -4$.

Traveling waves: circulant case



Now restrict the unit circle to the n roots of unity

- Recall the famous ``turnpike theorem'' in economics which describes the optimal tradeoff between personal consumption and savings in terms of an initial ``establish the right conditions'' phase, followed by a steady policy for the bulk of the time, and a final ``use it all up'' end game.
- Here we can expect something similar. In the first phase we steer to the fastest mode, follow the fastest mode until we are near the desired location and then disassemble the non dispersive fast mode into the desired end state.



Going from the initial circulant coefficients to the highest wave-speed circulant coefficients .

Concluding remarks

- We have presented rather few new “results”, but…
- Lattice problems were introduced and we illustrated how control systems involving them lead to nice questions.
- We have described a wave equation point of view and illustrated its use.
- Unlike traveling waves associated with Toda’s problem, however, here it seems that one can not avoid dispersion.