

# Derivations and Applications for SimpleDifferentialOperators.jl

Chiyoungh Ahn and Jesse Perla

April 4, 2019

## 1 Overview

This package is intended to be a stop-gap while more complete and higher-performance solutions are implemented (in particular, the evolution of <https://github.com/JuliaDiffEq/DiffEqOperators.jl/>). In the meantime, the package hopefully provides a solution for discretizing operators and solving ODEs/PDEs.

The focus is on discretizing linear (and, with some careful checks, affine) operators. In the case of models where the process is non-linear, it is most appropriate for algorithms that iteratively solve linear ODEs/PDEs. Some notation used throughout the document

- Let  $x \in R$  is the general “spatial” state variable
- Derivatives are denoted by the operator  $\partial_x$  and univariate derivatives such as  $\partial_x \tilde{v}(x) \equiv \tilde{v}'(x)$ .
- Use the vertical bar to denote operator evaluation at a particular point. That is if  $\tilde{B} \equiv \partial_x|_{x=x_0}$  then  $\tilde{B}\tilde{v}(x) = \partial_x \tilde{v}(x_0)$ , and if  $\tilde{B} \equiv 1|_{x=x_0}$  then  $\tilde{B}\tilde{v}(x) = \tilde{v}(x_0)$ . For the notation, the 1 is simply the identity operator on the function instead of applying a derivative (i.e.  $\tilde{B} \equiv 1|_{x=x_0}$  then  $\tilde{B}\tilde{v}(\cdot) = 1 \times \tilde{v}(x_0)$ )
- Let  $W_t$  be the Wiener process with the integral defined by the Ito interpretation

### 1.1 Linear Differential Equations

**ODEs (e.g. Steady State)** To understand the class of models that this can support, first look at the sort of ODE that comes out of solving a stationary model. The general pattern is a linear differential operator  $\tilde{L}$ , a boundary condition operator  $\tilde{B}$ , the function of interest  $\tilde{v}(x)$ , and the affine terms  $\tilde{f}(\cdot)$  and  $b$ . The general problem to solve is to find the  $\tilde{v}(x)$  such that.

$$0 = \tilde{L}\tilde{v}(x) - \tilde{f}(x) \tag{1}$$

$$0 = \tilde{B}\tilde{v}(x) - b \tag{2}$$

**Motivating Example** As a simple example, let

- $x \in [x_{\min}, x_{\max}]$  be a state variable on a domain following the SDE

$$dx_t = \mu dt + \sigma dW_t$$

where the variable  $x_t$  is reflected at  $x_{\min}$  and  $x_{\max}$

- The payoffs for state  $x$  are a function  $\tilde{f}(x)$  defines on the domain

- $\tilde{v}(x)$  as the value of the the stream of payoffs discounted at rate  $r > 0$

Then, through standard arguments, the stationary Bellman equation along with boundary conditions is

$$r\tilde{v}(x) = \tilde{f}(x) + \mu\partial_x\tilde{v}(x) + \frac{\sigma^2}{2}\partial_{xx}\tilde{v}(x) \quad (3)$$

$$\partial_x\tilde{v}(x_{\min}) = 0 \quad (4)$$

$$\partial_x\tilde{v}(x_{\max}) = 0 \quad (5)$$

Mapping to the notation of (1) and (2)

$$\tilde{L} \equiv r - \mu\partial_x - \frac{\sigma^2}{2}\partial_{xx} \quad (6)$$

$$\tilde{B} \equiv \begin{bmatrix} \partial_x|_{x=x_{\min}} \\ \partial_x|_{x=x_{\max}} \end{bmatrix} \quad (7)$$

$$b \equiv \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad (8)$$

This package will allow you to define the  $\tilde{L}$ ,  $\tilde{B}$ , and  $b$  to solve for a discretization of the  $\tilde{v}(x)$  function.

**PDEs (i.e. Time-Varying)** The motivating example above has no time-variation in any of the parameters, payoffs, or boundary conditions. Consider that the operators, payoffs, and boundary conditions could change over time – which we denote with a  $t$  subscript. As a variation on (1) and (2)

$$\partial_t\tilde{v}(t, x) = \tilde{L}(t)\tilde{v}(t, x) - \tilde{f}(t, x) \quad (9)$$

$$0 = \tilde{B}(t)\tilde{v}(t, x) - b(t) \quad (10)$$

Subject to an initial condition,  $\tilde{v}(0, x)$  given or potentially a boundary value,  $\tilde{v}(T, x)$  for some  $T$ .

This is a linear PDE where the operators, boundary conditions, and payoffs all may change over time.

**Motivating Example for Dynamics** Going back to the motivating example, consider an extension where We will make the following assumptions

- The discount rate, drift, and payoffs could be time varying. i.e.  $r(t)$ ,  $\mu(t)$  and  $\tilde{f}(t, x)$ .
- After some  $T$  the system is stationary because  $r(t) = r(T)$ ,  $\mu(t) = \mu(T)$  and  $\tilde{f}(t, x) = \tilde{f}(T, x)$  for all  $t \geq T$

Through standard arguments, the Bellman equation is

$$r\tilde{v}(t, x) = \tilde{f}(t, x) + \mu(t)\partial_x\tilde{v}(x) + \frac{\sigma^2}{2}\partial_{xx}\tilde{v}(x) + \partial_t\tilde{v}(t, x) \quad (11)$$

$$\partial_x\tilde{v}(t, x_{\min}) = 0 \quad (12)$$

$$\partial_x\tilde{v}(t, x_{\max}) = 0 \quad (13)$$

Mapping to the notation of the PDE in (9) and (10)

$$\tilde{L}(t) \equiv r(t) - \mu(t)\partial_x - \frac{\sigma^2}{2}\partial_{xx} \quad (14)$$

$$\tilde{B} \equiv \begin{bmatrix} \partial_x|_{x=x_{\min}} \\ \partial_x|_{x=x_{\max}} \end{bmatrix} \quad (15)$$

$$b \equiv \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad (16)$$

In order to find the steady state, we can solve the stationary equation with  $\partial_t \tilde{v}(T, x) = 0$  (i.e. with  $\tilde{L}(T)$  fixed to find  $\tilde{v}(T, x)$ ) and then use the  $\tilde{v}(T, x)$  as a boundary value to solve for the  $\tilde{v}(t, x)$  by solving the PDE in

**Boundary Conditions** The package supports some key boundary conditions used for stochastic processes and ODE/PDEs.

As will become clear in the discretization, whether the boundary condition is homogenous or not (i.e.  $b > 0$  or  $b = 0$ ) is important for the numerical methods. If the boundary conditions are inhomogeneous, then the setup is affine. To detail a few of the one-dimensional versions of the supported boundary conditions

- Reflecting Barriers (i.e. a homogeneous Neumann Boundary Conditions)

$$\partial_x \tilde{v}(x_{\min}) = 0 \quad (17)$$

$$\partial_x \tilde{v}(x_{\max}) = 0 \quad (18)$$

or in operator form

$$\tilde{B} \equiv \begin{bmatrix} \partial_x|_{x=x_{\min}} \\ \partial_x|_{x=x_{\max}} \end{bmatrix} \quad (19)$$

$$b \equiv \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad (20)$$

- Mixed boundary conditions (i.e a homogeneous Robin Boundary Conditions):

$$\underline{\xi} \tilde{v}(x_{\min}) + \partial_x \tilde{v}(x_{\min}) = 0 \quad (21)$$

$$\bar{\xi} \tilde{v}(x_{\max}) + \partial_x \tilde{v}(x_{\max}) = 0 \quad (22)$$

Note that when  $\underline{\xi} = \bar{\xi} = 0$ , this nests the reflecting barriers. In operator form,

$$\tilde{B} \equiv \begin{bmatrix} \partial_x|_{x=x_{\min}} + \underline{\xi} 1|_{x=x_{\min}} \\ \partial_x|_{x=x_{\max}} + \bar{\xi} 1|_{x=x_{\max}} \end{bmatrix} \quad (23)$$

$$b \equiv \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad (24)$$

- Absorbing Barriers (i.e. a homogenous or inhomogeneous Dirichlet Boundary Conditions)

$$\tilde{v}(x_{\min}) = b_1 \quad (25)$$

$$\tilde{v}(x_{\max}) = b_2 \quad (26)$$

In the case of  $b_1 = b_2 = 0$ , this is homogeneous. In operator form,

$$\tilde{B} \equiv \begin{bmatrix} 1|_{x=x_{\min}} \\ 1|_{x=x_{\max}} \end{bmatrix} \quad (27)$$

$$b \equiv \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \quad (28)$$

Of course, many models would have different boundary conditions on different sides of the domain, which entails mixing and matching rows in the  $B$  and  $b$  matrices.

## 2 Discretization

### 2.1 Notation

This section defines the grids and other notation for the discretization.

- Define an irregular grid  $\{x_i\}_{i=0}^{M+1}$  with **boundary nodes**,  $x_0 = x_{\min}$  and  $x_{M+1} = x_{\max}$ . Denote the **extended grid** as  $\bar{x} \equiv \{x_i\}_{i=0}^{M+1}$  and the **interior grid**, a collection of nodes excluding the boundary nodes, as  $x \equiv \{x_i\}_{i=1}^M$ .
- Recall that the continuous functions, prior to discretization, are denoted like  $\tilde{u}(x)$ . The discretization of  $\tilde{u}(x)$  on the interior  $x \in \mathbb{R}^M$  is denoted  $u \equiv \{\tilde{u}(x_i)\}_{i=1}^M \in \mathbb{R}^M$  and the discretization of  $\tilde{u}(x)$  on  $\bar{x} \in \mathbb{R}^{M+2}$  is  $\bar{u} \equiv \{\tilde{u}(x_i)\}_{i=0}^{M+1} \in \mathbb{R}^{M+2}$ .
- When we discretize a particular operator, e.g.  $\tilde{L}$ , we will drop the tilde to become  $L$ . The typical size of this, before applying boundary conditions, is  $L \in \mathbb{R}^{M \times (M+2)}$ .
- Denote the backward and forward distance between the grid points as

$$\Delta_{i,-} \equiv x_i - x_{i-1}, \text{ for } i = 1, \dots, M+1 \quad (29)$$

$$\Delta_{i,+} \equiv x_{i+1} - x_i, \text{ for } i = 0, \dots, M \quad (30)$$

- Define the vector of backwards and forwards first differences, padding with  $\Delta_{0,-} = \Delta_{M+1,+} = 0$ , as

$$\Delta_- \equiv \begin{bmatrix} 0 \\ \text{diff}(z) \end{bmatrix} \in \mathbb{R}^{M+2} \quad (31)$$

$$\Delta_+ \equiv \begin{bmatrix} \text{diff}(z) \\ 0 \end{bmatrix} \in \mathbb{R}^{M+2} \quad (32)$$

- Some special matrices to help in the composition notation:
  - $\mathbf{I}_N$  is the  $N \times N$  identity matrix. Always drop the subscript when the dimensions are unambiguous, as it would be the same in the code
  - $\mathbf{0}_N$  is the column vector of  $N$  0s, and  $\mathbf{0}_N^\top$  a row vector
  - $\mathbf{0}_{N \times M}$  is the  $N \times M$  matrix of 0s

In order to discretize these operators with finite-differences, we need to choose a stencil. Denote the discretization

- For the first-derivative operator  $\tilde{L}_1 \equiv \partial_x$ , denote the stencils for discretizing with backwards and forward first-differences respectively as  $L_{1-} \in \mathbb{R}^{M \times (M+2)}$  and  $L_{1+} \in \mathbb{R}^{M \times (M+2)}$ .<sup>1</sup>
- For the second-derivative operator  $\tilde{L}_2 \equiv \partial_{xx}$ , always use central differences and denote the discretized operator as  $L_2 \in \mathbb{R}^{M \times (M+2)}$

For first-derivatives, the choice of  $L_{1-}$  vs.  $L_{1+}$  or a combination of them, will use upwind finite differences.

---

<sup>1</sup>Note that the stencil for both of these only really needs to be defined on  $M \times (M+1)$  but we will pad a column with 0s to make composition easier. In the current form, the package composes operators as sparse matrices. Depending on the circumstances, this code will execute slower than a hand-tweaked model creating composed operators directly. In many cases, this wouldn't be a problem, but in some algorithms where operators need to be redefined frequently in tight loops, it might be. In those cases, use the output of this package for test-suites on hand-built discretizations.

### 3 Discretizing Operators with a Regular Grid

In this section, we study discretization schemes under regular grids, i.e., grids such that  $x_{i+1} - x_i = \Delta$  for all  $i = 0, \dots, M$  for some fixed  $\Delta > 0$ .

Throughout, take a function of interest  $\tilde{v}(x)$  defined on the grid, and define  $\bar{v} \equiv \{\tilde{v}(x_i)\}_{i=0}^{M+1}$  and  $v \equiv \{\tilde{v}(x_i)\}_{i=1}^M$ .

#### 3.1 Discretized Operators

In this section we derive the stencils for operators of various orders.

**First Derivative Operators** To discretize the  $\tilde{L}_1$  operator, we can use a backward difference approximation

$$\tilde{L}_1 \tilde{v}(x_i) \equiv \partial_x \tilde{v}(x_i) \approx \frac{\bar{v}_i - \bar{v}_{i-1}}{\Delta}, \text{ for } i = 1, \dots, M \quad (33)$$

And with forward differences

$$\tilde{L}_1 \tilde{v}(x_i) \equiv \partial_x \tilde{v}(x_i) \approx \frac{\bar{v}_{i+1} - \bar{v}_i}{\Delta}, \text{ for } i = 1, \dots, M \quad (34)$$

In order to calculate the derivatives for all  $i = 1, \dots, M$  (i.e. in the interior) we can stack these up and apply to the extension  $\bar{v}$ .

$$\{\partial_x \tilde{v}(x_i)\}_{i=1}^M \approx L_{1-} \cdot \bar{v} \quad (35)$$

or,

$$\{\partial_x \tilde{v}(x_i)\}_{i=1}^M \approx L_{1+} \cdot \bar{v} \quad (36)$$

Where we define  $L_{1-}$  from applying (33) to the  $\bar{v}$  vector for all  $i = 1, \dots, M$

$$L_{1-} \equiv \frac{1}{\Delta} \begin{bmatrix} -1 & 1 & 0 & \dots & 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & \ddots & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & -1 & 1 & 0 \end{bmatrix}_{M \times (M+2)} \quad (37)$$

And similarly define  $L_{1+}$  from applying (34) to the  $\bar{v}$  vector for all  $i = 1, \dots, M$

$$L_{1+} \equiv \frac{1}{\Delta} \begin{bmatrix} 0 & -1 & 1 & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 & -1 & 1 \end{bmatrix}_{M \times (M+2)} \quad (38)$$

$$(39)$$

It is important to note that while these operators map the  $\bar{v}$  (i.e. including the boundary points), the operator only maps to points on the interior i.e.  $i = 1, \dots, M$ .

**Second Derivative Operators** To discretize the  $\tilde{L}_2$  second order operator, we can use central differences

$$\tilde{L}_2 \tilde{v}(x_i) \equiv \partial_{xx} \tilde{v}(x_i) \approx \frac{\bar{v}_{i+1} - 2\bar{v}_i + \bar{v}_{i-1}}{\Delta^2}, \text{ for } i = 1, \dots, M \quad (40)$$

In order to calculate the derivatives for all  $i = 1, \dots, M$  (i.e. in the interior) we can stack these up and apply to the extension  $\bar{v}$ .

$$\{\partial_{xx} \tilde{v}(x_i)\}_{i=1}^M \approx L_2 \cdot \bar{v} \quad (41)$$

Where we define  $L_2$  from applying (40) to the  $\bar{v}$  vector for all  $i = 1, \dots, M$

$$L_2 \equiv \frac{1}{\Delta^2} \begin{bmatrix} 1 & -2 & 1 & 0 & \dots & 0 & 0 & 0 & 0 \\ 0 & 1 & -2 & 1 & \dots & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 1 & -2 & 1 & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 & 1 & -2 & 1 \end{bmatrix}_{M \times (M+2)} \quad (42)$$

**Identity Operators** For simplicity in composition, also consider the discretization of the identity operator (i.e. not applying any derivatives or stencils). For simplicity, define the identity operator as the 0-th order operator  $\tilde{L}_0 \equiv I$  so that  $\tilde{L}_0 \tilde{v}(x) = \tilde{v}(x)$ .

With this, the operator applied to the  $\bar{v}$  vector is trivial

$$\tilde{L}_0 \tilde{v}(x_i) \equiv \tilde{v}(x_i) = \bar{v}_i, \text{ for } i = 1, \dots, M \quad (43)$$

And stacking it up for all  $i = 1, \dots, M$ ,

$$\{\tilde{v}(x_i)\}_{i=1}^M \approx L_0 \cdot \bar{v} \quad (44)$$

Where

$$L_0 \equiv \begin{bmatrix} 0 & 1 & 0 & 0 & \dots & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 & 1 & 0 \end{bmatrix}_{M \times (M+2)} \quad (45)$$

This operator trivially maps from the  $\bar{v}$  to extract the interior  $v$  and ignoring the boundaries  $\bar{v}_0$  and  $\bar{v}_{P+1}$ . Its primary role will be when composing operators rather than being used directly.

### 3.2 Representation with Special Matrices

See Appendix A for more details on the notation of special matrices (e.g. Toeplitz, Tridiagonal, and Banded)

These can be defined in more compact forms as banded matrices

$$L_{1-} \equiv \frac{1}{\Delta} [\text{tridiag}(\mathbf{0}_{M-1}, -\mathbf{1}_M, \mathbf{1}_M) \quad \mathbf{0}_M] \quad (46)$$

$$L_{1+} \equiv \frac{1}{\Delta} [\mathbf{0}_M \quad \text{tridiag}(\mathbf{0}_{M-1}, -\mathbf{1}_M, \mathbf{1}_M)] \quad (47)$$

$$L_2 \equiv \frac{1}{\Delta^2} \text{tridiag}^+(\mathbf{1}_M, -2\mathbf{1}_M, \mathbf{1}_M) \quad (48)$$

$$L_0 \equiv [\mathbf{0}_M \quad \mathbf{I}_M \quad \mathbf{0}_M] \quad (49)$$

where  $\text{tridiag}(x, y, z)$  is a matrix whose lower, main, upper diagonal vectors are  $x$ ,  $y$ , and  $z$ , respectively, and  $\text{tridiag}^+(x, y, z)$  is a matrix whose main, upper, and second upper diagonal vectors are  $x$ ,  $y$ , and  $z$ , respectively.

### 3.3 Applying Boundary Conditions

Boundary conditions can be applied manually by using operators on extended grids,  $\bar{x}$ , to find solutions on extended grids. Suppose that we want to solve a system  $\tilde{L}\tilde{v}(x) = \tilde{f}(x)$  for  $\tilde{v}(x)$  where  $L$  is a linear combination of discretized differential operators for some  $f$  that represents the values of a function  $\tilde{f}(\cdot)$  on discretized interior  $x$ . To solve the system under boundary conditions on  $v$ , one can construct and solve the following extended system:

$$\begin{bmatrix} L \\ B \end{bmatrix} \bar{v} = \begin{bmatrix} f \\ b \end{bmatrix} \quad (50)$$

with  $M_E$  by  $(M+2)$  matrix  $B$  and  $M_E$ -length vector  $b$  that represent the current boundary conditions, where  $M_E$  is the number of boundary conditions to be applied. The solution of (50),  $\tilde{v}(\bar{x})$  can be decomposed into

$$\bar{v} = \begin{bmatrix} \tilde{v}(x_0) \\ v \\ \tilde{v}(x_{M+1}) \end{bmatrix} \quad (51)$$

which also gives the solution for  $v$ .

#### 3.3.1 Mixed Boundary Conditions

Recall mixed boundary conditions from (21) and (22). Note that reflecting barrier conditions are special cases with  $\bar{\xi} = \underline{\xi} = 0$ . Using forward difference and backward difference discretization scheme for the lower bound and upper bound respectively, we have

$$\frac{\bar{v}_1 - \bar{v}_0}{\Delta} - \underline{\xi}\bar{v}_0 = 0 \quad (52)$$

$$\frac{\bar{v}_{M+1} - \bar{v}_M}{\Delta} - \bar{\xi}\bar{v}_{M+1} = 0 \quad (53)$$

Thus, the corresponding boundary condition matrix  $B$  is

$$B = \begin{bmatrix} -\frac{1}{\Delta} + \underline{\xi} & \frac{1}{\Delta} & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & -\frac{1}{\Delta} & \frac{1}{\Delta} + \bar{\xi} \end{bmatrix}_{2 \times (M+2)} \quad b = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad (54)$$

which provides the identical system as

$$B = \begin{bmatrix} -1 + \underline{\xi}\Delta & 1 & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & -1 & 1 + \bar{\xi}\Delta \end{bmatrix}_{2 \times (M+2)} \quad b = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad (55)$$

Subtracting the first row of  $\bar{L}_{1-}$  by the first row of  $B$  in (84) multiplied by  $(-1 + \underline{\xi}\Delta)^{-1}\Delta^{-1}$  gives, with the corresponding row operation matrix  $R$ ,

$$RL_{1-} = \frac{1}{\Delta} \begin{bmatrix} 0 & 1 + (-1 + \underline{\xi}\Delta)^{-1} & 0 & 0 & \dots & 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & \ddots & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 & -1 & 1 & 0 \end{bmatrix}_{M \times (M+2)} \quad (56)$$

note that there is no zero element in the first and last column for nodes on boundaries. Hence, solving the corresponding extended system,

$$\begin{bmatrix} L \\ B \end{bmatrix} = \begin{bmatrix} f \\ b \end{bmatrix} \quad (57)$$

is identical as solving the following system

$$R \begin{bmatrix} L \\ B \end{bmatrix} = R \begin{bmatrix} f \\ b \end{bmatrix} \quad (58)$$

$$= \begin{bmatrix} f \\ b \end{bmatrix} \quad (59)$$

as  $b$  is a zero vector so that the row operations  $R$  do not change anything on the RHS. Furthermore, limited to the interior, solving  $v$  in the system above is identical as solving the following system with an operator  $L^B$  on interior nodes:

$$L^B v = f \quad (60)$$

where we have  $L = L_{1-}$  and  $L^B = L_{1-}^B$  with

$$L_{1-}^B \equiv \frac{1}{\Delta} \begin{bmatrix} 1 + (-1 + \underline{\xi}\Delta)^{-1} & 0 & 0 & \dots & 0 & 0 & 0 \\ -1 & 1 & 0 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & -1 & 1 & 0 \\ 0 & 0 & 0 & \dots & 0 & -1 & 1 \end{bmatrix}_{M \times M} \quad (61)$$

instead of solving the full system with boundary conditions. Similarly, one can define differential operators on the interior as follows:

$$L_{1-}^B \equiv \frac{1}{\Delta} \begin{bmatrix} 1 + (-1 + \underline{\xi}\Delta)^{-1} & 0 & 0 & \dots & 0 & 0 & 0 \\ -1 & 1 & 0 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & -1 & 1 & 0 \\ 0 & 0 & 0 & \dots & 0 & -1 & 1 \end{bmatrix}_{M \times M} \quad (62)$$

$$L_{1+}^B \equiv \frac{1}{\Delta} \begin{bmatrix} -1 & 1 & 0 & \dots & 0 & 0 & 0 \\ 0 & -1 & 1 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & -1 & 1 \\ 0 & 0 & 0 & \dots & 0 & 0 & -1 + (1 + \bar{\xi}\Delta)^{-1} \end{bmatrix}_{M \times M} \quad (63)$$

$$L_2^B \equiv \frac{1}{\Delta^2} \begin{bmatrix} -2 - (-1 + \underline{\xi}\Delta)^{-1} & 1 & 0 & \dots & 0 & 0 & 0 \\ 1 & -2 & 1 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & -2 & 1 \\ 0 & 0 & 0 & \dots & 0 & 1 & -2 + (1 + \bar{\xi}\Delta)^{-1} \end{bmatrix}_{M \times M} \quad (64)$$

which can be defined in more compact forms:

$$L_{1-}^B \equiv \text{tridiag}(-\mathbf{1}_{M-1}, [1 + (-1 + \underline{\xi}\Delta)^{-1} \quad \mathbf{1}_{M-1}^T]^T, \mathbf{0}_{M-1}) \quad (65)$$

$$L_{1+}^B \equiv \text{tridiag}(\mathbf{0}_{M-1}, [-\mathbf{1}_{M-1}^T \quad -1 + (1 + \bar{\xi}\Delta)^{-1}]^T, \mathbf{1}_{M-1}) \quad (66)$$

$$L_2^B \equiv \text{tridiag}(\mathbf{1}_{M-1}, [-2 - (-1 + \underline{\xi}\Delta)^{-1} \quad -2\mathbf{1}_{M-2}^T \quad -2 + (1 + \bar{\xi}\Delta)^{-1}]^T, \mathbf{1}_{M-1}) \quad (67)$$



### 3.3.2 Absorbing Boundary Conditions

To apply an absorbing barrier condition  $\tilde{v}(x_{\min}) = S$  for some  $S \in \mathbb{R}$ , with one reflecting barrier condition on the upper bound  $v'(x_{\max}) = 0$ , one can use

$$B = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & -1 & 1 \end{bmatrix}_{2 \times (M+2)} \quad b = \begin{bmatrix} S \\ 0 \end{bmatrix} \quad (68)$$

Similarly, one can apply an absorbing condition on the upper bound  $\tilde{v}(x_{\max}) = S$  for some  $S \in \mathbb{R}$  and the

$$B = \begin{bmatrix} -1 & 1 & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & 1 \end{bmatrix}_{2 \times (M+2)} \quad b = \begin{bmatrix} 0 \\ S \end{bmatrix} \quad (69)$$

## 4 Discretizing Operators with a Irregular Grid

### 4.1 Applying Boundary Conditions

Instead of solving (50) for a value function  $\tilde{v}(\bar{x})$  on the extended grid, one can perform Gaussian elimination to reduce the system and solve  $\tilde{v}(x)$ , which gives the identical solution as the interior of  $\tilde{v}(\bar{x})$ .

### 4.2 Irregular grids

Define the vectors of backward and forward distance for interior nodes as follows:

$$\Delta_{-}^{\circ} = \{\Delta_{i,-}\}_{i=1}^M \quad (70)$$

$$\Delta_{+}^{\circ} = \{\Delta_{i,+}\}_{i=1}^M \quad (71)$$

We can then define the following operators on  $\bar{x}$ :

$$L_{1-} \equiv \begin{bmatrix} -\Delta_{1,-}^{-1} & \Delta_{1,-}^{-1} & 0 & \dots & 0 & 0 & 0 \\ 0 & -\Delta_{2,-}^{-1} & \Delta_{2,-}^{-1} & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \Delta_{M-1,-}^{-1} & 0 & 0 \\ 0 & 0 & 0 & \dots & -\Delta_{M,-}^{-1} & \Delta_{M,-}^{-1} & 0 \end{bmatrix}_{M \times (M+2)} \quad (72)$$

$$L_{1+} \equiv \begin{bmatrix} 0 & -\Delta_{1,+}^{-1} & \Delta_{1,+}^{-1} & \dots & 0 & 0 & 0 \\ 0 & 0 & -\Delta_{2,+}^{-1} & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & -\Delta_{M-1,+}^{-1} & \Delta_{M-1,+}^{-1} & 0 \\ 0 & 0 & 0 & \dots & 0 & -\Delta_{M,+}^{-1} & \Delta_{M,+}^{-1} \end{bmatrix}_{M \times (M+2)} \quad (73)$$

$$L_2 \equiv 2 \begin{pmatrix} (\Delta_{1,+} + \Delta_{1,-})^{-1} \Delta_{1,-}^{-1} & -\Delta_{1,-}^{-1} \Delta_{1,+}^{-1} & (\Delta_{1,+} + \Delta_{1,-})^{-1} \Delta_{1,+}^{-1} & \dots & 0 & 0 \\ 0 & \ddots & \ddots & \ddots & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & (\Delta_{M,+} + \Delta_{M,-})^{-1} \Delta_{M,-}^{-1} & -\Delta_{M,-}^{-1} \Delta_{M,+}^{-1} & (\Delta_{M,+} + \Delta_{M,-})^{-1} \Delta_{M,+}^{-1} \end{pmatrix}_{M \times (M+2)} \quad (74)$$

Note that we use the following discretization scheme from ?:

$$v''(x_i) \approx \frac{\Delta_{i,-}\tilde{v}(x_i + \Delta_{i,+}) - (\Delta_{i,+} + \Delta_{i,-})\tilde{v}(x_i) + \Delta_{i,+}\tilde{v}(x_i - \Delta_{i,-})}{\frac{1}{2}(\Delta_{i,+} + \Delta_{i,-})\Delta_{i,+}\Delta_{i,-}}, \text{ for } i = 1, \dots, M \quad (75)$$

for second-order derivatives.

And one for identity matrix:

$$I \equiv \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 & 0 \end{bmatrix}_{M \times (M+2)} \quad (76)$$

Alternatively, in a more compact form, using vectors distances for interior nodes:

$$L_{1-} \equiv [\text{tridiag}(\mathbf{0}_{M-1}, -(\Delta_-^\circ)^{-1}, (\Delta_-^\circ)^{-1}) \quad \mathbf{0}_M] \quad (77)$$

$$L_{1+} \equiv [\mathbf{0}_M \quad \text{tridiag}(\mathbf{0}_{M-1}, -(\Delta_-^\circ)^{-1}, (\Delta_-^\circ)^{-1})] \quad (78)$$

$$L_2 \equiv 2 \odot \text{tridiag}^+ [(\Delta_-^\circ + \Delta_+^\circ)^{-1} \odot (\Delta_-^\circ)^{-1}, -(\Delta_-^\circ \odot \Delta_+^\circ)^{-1}, (\Delta_-^\circ + \Delta_+^\circ)^{-1} \odot (\Delta_+^\circ)^{-1}] \quad (79)$$

$$I \equiv [\mathbf{0}_M \quad \text{diag}(\mathbf{1}_M) \quad \mathbf{0}_M] \quad (80)$$

#### 4.2.1 Mixed boundary conditions

Recall mixed boundary conditions from (21) and (22). Note that reflecting barrier conditions are special cases with  $\bar{\xi} = \underline{\xi} = 0$ . Using forward difference and backward difference discretization scheme for the lower bound and upper bound respectively, we have

$$\frac{\bar{v}_1 - \bar{v}_0}{\Delta_{0,+}} - \underline{\xi}\bar{v}_0 = 0 \quad (81)$$

$$\frac{\bar{v}_{M+1} - \bar{v}_M}{\Delta_{M+1,-}} - \bar{\xi}\bar{v}_{M+1} = 0 \quad (82)$$

Thus, the corresponding boundary condition matrix  $B$  is

$$B = \begin{bmatrix} -\frac{1}{\Delta_{0,+}} + \underline{\xi} & \frac{1}{\Delta_{0,+}} & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & -\frac{1}{\Delta_{M+1,-}} & \frac{1}{\Delta_{M+1,-}} + \bar{\xi} \end{bmatrix}_{2 \times (M+2)} \quad b = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad (83)$$

which provides the identical system as

$$B = \begin{bmatrix} -1 + \underline{\xi}\Delta_{1,-} & 1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & -1 & 1 + \bar{\xi}\Delta_{M,+} \end{bmatrix}_{2 \times (M+2)} \quad b = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad (84)$$

since  $\Delta_{0,+} = \Delta_{1,-}$  and  $\Delta_{M+1,-} = \Delta_{M,+}$

The first columns of all the extension operators above,  $\bar{L}_{1-}, \bar{L}_{1+}, \bar{L}_2, \bar{I}$ , have non-zero element only in the first rows. Thus, a single Gaussian elimination for the first extension grid will suffice

to remove the extended. Likewise, in the last columns of all the extension operators have non-zero element only in the last row.

Subtracting the first row of  $\bar{L}_{1-}$  by the first row of  $B$  in (84) multiplied by  $(-1 + \xi \Delta_{1,-})^{-1} \Delta_{1,-}^{-1}$  gives, with the corresponding row operation matrix  $R$  for Gaussian elimination,

$$RL_{1-} = \begin{bmatrix} 0 & \Delta_{1,-}^{-1} [1 + (-1 + \xi \Delta_{1,-})^{-1}] & 0 & \dots & 0 & 0 & 0 \\ 0 & -\Delta_{2,-}^{-1} & \Delta_{2,-}^{-1} & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \Delta_{M-1,-}^{-1} & 0 & 0 \\ 0 & 0 & 0 & \dots & -\Delta_{M,-}^{-1} & \Delta_{M,-}^{-1} & 0 \end{bmatrix}_{M \times (M+2)} \quad (85)$$

note that there is no zero element in the first and last column for nodes on boundaries. Hence, solving the corresponding extended system,

$$\begin{bmatrix} L \\ B \end{bmatrix} = \begin{bmatrix} f \\ b \end{bmatrix} \quad (86)$$

is identical as solving the following system

$$R \begin{bmatrix} L \\ B \end{bmatrix} = R \begin{bmatrix} f \\ b \end{bmatrix} \quad (87)$$

$$= \begin{bmatrix} f \\ b \end{bmatrix} \quad (88)$$

as  $b$  is a zero vector so that the row operations  $R$  do not change anything on the RHS. Furthermore, limited to the interior, solving  $v$  in the system above is identical as solving the following system with an operator  $L^B$  on interior nodes:

$$L^B v = f \quad (89)$$

where we have  $L = L_{1-}$  and  $L^B = L_{1-}^B$  with

$$L_{1-}^B \equiv \begin{bmatrix} \Delta_{1,-}^{-1} [1 + (-1 + \xi \Delta_{1,-})^{-1}] & 0 & 0 & \dots & 0 & 0 & 0 \\ -\Delta_{2,-}^{-1} & \Delta_{2,-}^{-1} & 0 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & -\Delta_{M-1,-}^{-1} & \Delta_{M-1,-}^{-1} & 0 \\ 0 & 0 & 0 & \dots & 0 & -\Delta_{M,-}^{-1} & \Delta_{M,-}^{-1} \end{bmatrix}_{M \times M} \quad (90)$$

instead of solving the full system with boundary conditions. Similarly, one can define differential operators on the interior as follows:

$$L_{1-}^B \equiv \begin{bmatrix} \Delta_{1,-}^{-1}[1 + (-1 + \xi\Delta_{1,-})^{-1}] & 0 & 0 & \dots & 0 & 0 & 0 \\ & -\Delta_{2,-}^{-1} & \Delta_{2,-}^{-1} & 0 & \dots & 0 & 0 & 0 \\ & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ & 0 & 0 & 0 & \dots & -\Delta_{M-1,-}^{-1} & \Delta_{M-1,-}^{-1} & 0 \\ & 0 & 0 & 0 & \dots & 0 & -\Delta_{M,-}^{-1} & \Delta_{M,-}^{-1} \end{bmatrix}_{M \times M} \quad (91)$$

$$L_{1+}^B \equiv \begin{bmatrix} -\Delta_{1,+}^{-1} & \Delta_{1,+}^{-1} & 0 & \dots & 0 & 0 & 0 \\ 0 & -\Delta_{2,+}^{-1} & \Delta_{2,+}^{-1} & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & -\Delta_{M-1,+}^{-1} & \Delta_{M-1,+}^{-1} \\ 0 & 0 & 0 & \dots & 0 & 0 & \Delta_{M,+}^{-1}[-1 + (1 + \xi\Delta_{M,+})^{-1}] \end{bmatrix}_{M \times M} \quad (92)$$

$$L_2^B \equiv 2 \begin{pmatrix} \Xi_1 (\Delta_{1,+} + \Delta_{1,-})^{-1} \Delta_{1,+}^{-1} & 0 & \dots & 0 & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & (\Delta_{i,+} + \Delta_{i,-})^{-1} \Delta_{i,-}^{-1} & -\Delta_{i,-}^{-1} \Delta_{i,+}^{-1} & (\Delta_{i,+} + \Delta_{i,-})^{-1} \Delta_{i,+}^{-1} & \dots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 0 & (\Delta_{M,+} + \Delta_{M,-})^{-1} \Delta_{M,-}^{-1} \Xi_M \end{pmatrix}_{M \times M} \quad (93)$$

$$I^B \equiv \begin{bmatrix} 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & \dots & 0 & 1 \end{bmatrix}_{M \times M} \quad (94)$$

where

$$\Xi_1 = -2 \left[ \Delta_{1,-}^{-1} \Delta_{1,+}^{-1} + (-1 + \xi\Delta_{1,-})^{-1} (\Delta_{1,+} + \Delta_{1,-})^{-1} \Delta_{1,-}^{-1} \right] \quad (95)$$

$$\Xi_M = -2 \left[ \Delta_{M,-}^{-1} \Delta_{M,+}^{-1} - (1 + \bar{\xi}\Delta_{M,+})^{-1} (\Delta_{M,+} + \Delta_{M,-})^{-1} \Delta_{M,+}^{-1} \right] \quad (96)$$

Or, alternatively, with vectorized differences:

$$L_{1-}^B \equiv \text{tridiag} \left[ -(\Delta_-^\circ)^{-1}[2 : M], [\Delta_{1,-}^{-1}[1 + (-1 + \xi\Delta_{1,-})^{-1}]; (\Delta_-^\circ)^{-1}[2 : M]] , \mathbf{0}_{M-1} \right] \quad (97)$$

$$L_{1+}^B \equiv \text{tridiag} \left[ \mathbf{0}_{M-1}, [-(\Delta_-^\circ)^{-1}[1 : M-1]; \Delta_{M,+}^{-1}[-1 + (1 + \bar{\xi}\Delta_{M,+})^{-1}]] , (\Delta_-^\circ)^{-1}[2 : M] \right] \quad (98)$$

$$L_2^B \equiv 2 \odot \text{tridiag} \left[ (\Delta_-^\circ + \Delta_+^\circ)^{-1} \odot (\Delta_-^\circ)^{-1}, \right. \quad (99)$$

$$\left. [\Xi_1; -(\Delta_-^\circ \odot \Delta_+^\circ)^{-1}[2 : M-1]; \Xi_M] , \right. \quad (100)$$

$$\left. (\Delta_-^\circ + \Delta_+^\circ)^{-1} \odot (\Delta_+^\circ)^{-1} \right] \quad (101)$$

$$I^B \equiv \text{diag}(\mathbf{1}_M) \quad (102)$$

### 4.2.2 Absorbing barrier conditions

To apply an absorbing barrier condition  $\tilde{v}(x_{\min}) = S$  for some  $S \in \mathbb{R}$ , with one reflecting barrier condition on the upper bound  $v'(x_{\max}) = 0$ , one can use

$$B = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & -1 & 1 \end{bmatrix}_{2 \times (M+2)} \quad b = \begin{bmatrix} S \\ 0 \end{bmatrix} \quad (103)$$

Similarly, one can apply an absorbing condition on the upper bound  $\tilde{v}(x_{\max}) = S$  for some  $S \in \mathbb{R}$  and the

$$B = \begin{bmatrix} -1 & 1 & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & 1 \end{bmatrix}_{2 \times (M+2)} \quad b = \begin{bmatrix} 0 \\ S \end{bmatrix} \quad (104)$$

Note that the corresponding  $B$  and  $b$  are identical with regular grid cases.

### 4.3 Examples

**Examples 4.1.** Consider  $L = L_2$  to solve  $Lv = f$  with  $M = 3$  under uniform grids  $\bar{x} = \{x_0, x_1, x_2, x_3, x_4\}$  and  $\Delta = 1$ , whose corresponding interior grid is  $x = \{x_1, x_2, x_3\}$ . This gives

$$L^B = \begin{bmatrix} -2 & 1 & 0 \\ 1 & -2 & 1 \\ 0 & 1 & -2 \end{bmatrix} \quad (105)$$

so  $L^B v = f$  on the grid  $x$  results in the following system

$$\begin{bmatrix} -1 & 1 & 0 \\ 1 & -2 & 1 \\ 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} \tilde{v}(x_1) \\ \tilde{v}(x_2) \\ \tilde{v}(x_3) \end{bmatrix} = \begin{bmatrix} \tilde{f}(x_1) \\ \tilde{f}(x_2) \\ \tilde{f}(x_3) \end{bmatrix} \quad (106)$$

For the extended system we have

$$L = \begin{bmatrix} 1 & -2 & 1 & 0 & 0 \\ 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 1 & -2 & 1 \end{bmatrix} \quad (107)$$

$$B = \begin{bmatrix} -1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 1 \end{bmatrix} \quad (108)$$

Constructing the stacked extended system (50) gives

$$\begin{bmatrix} 1 & -2 & 1 & 0 & 0 \\ 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 1 & -2 & 1 \\ -1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} \tilde{v}(x_0) \\ \tilde{v}(x_1) \\ \tilde{v}(x_2) \\ \tilde{v}(x_3) \\ \tilde{v}(x_4) \end{bmatrix} = \begin{bmatrix} \tilde{f}(x_1) \\ \tilde{f}(x_2) \\ \tilde{f}(x_3) \\ 0 \\ 0 \end{bmatrix} \quad (109)$$

Note that subtracting the first row of  $L$  by  $(-1)$  times the first row of  $B$  returns an identical system as (109). Likewise, subtracting the last row of  $L$  by  $(-1)$  times the last row of  $B$  returns the identical system. Performing the two Gaussian elimination yields the following system:

$$\begin{bmatrix} 0 & -1 & 1 & 0 & 0 \\ 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} \tilde{v}(x_0) \\ \tilde{v}(x_1) \\ \tilde{v}(x_2) \\ \tilde{v}(x_3) \\ \tilde{v}(x_4) \end{bmatrix} = \begin{bmatrix} \tilde{f}(x_2) \\ \tilde{f}(x_3) \\ \tilde{f}(x_4) \\ 0 \\ 0 \end{bmatrix} \quad (110)$$

Note that now we have the first three rows of the coefficient matrix with zero columns on the extended nodes,  $\tilde{v}(x_1 - \Delta)$  and  $\tilde{v}(x_3 + \Delta)$ . Extracting the system corresponding to the first three rows returns the following system, which solves the interior of  $\bar{v}$ , i.e.,  $v$ :

$$\begin{bmatrix} -1 & 1 & 0 \\ 1 & -2 & 1 \\ 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} \tilde{v}(x_1) \\ \tilde{v}(x_2) \\ \tilde{v}(x_3) \end{bmatrix} = \begin{bmatrix} \tilde{f}(x_1) \\ \tilde{f}(x_2) \\ \tilde{f}(x_3) \end{bmatrix} \quad (111)$$

which is identical as (106).

**Examples 4.2.** Consider solving (106), but this time with an absorbing barrier condition on the lower bound,  $\tilde{v}(x_{\min}) = S$  with a boundary condition matrix

$$B = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 1 \end{bmatrix} \quad (112)$$

The corresponding extended system is

$$\begin{bmatrix} 1 & -2 & 1 & 0 & 0 \\ 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 1 & -2 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} \tilde{v}(x_0) \\ \tilde{v}(x_1) \\ \tilde{v}(x_2) \\ \tilde{v}(x_3) \\ \tilde{v}(x_4) \end{bmatrix} = \begin{bmatrix} \tilde{f}(x_1) \\ \tilde{f}(x_2) \\ \tilde{f}(x_3) \\ S \\ 0 \end{bmatrix} \quad (113)$$

Note that subtracting the first row of  $L$  by  $(-1)$  times the first row of  $B$  returns an identical system as (113). Likewise, subtracting the last row of  $L$  by  $(-1)$  times the last row of  $B$  returns the identical system. Performing the two Gaussian elimination yields the following system:

$$\begin{bmatrix} 0 & -2 & 1 & 0 & 0 \\ 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} \tilde{v}(x_0) \\ \tilde{v}(x_1) \\ \tilde{v}(x_2) \\ \tilde{v}(x_3) \\ \tilde{v}(x_4) \end{bmatrix} = \begin{bmatrix} \tilde{f}(x_1) - S \\ \tilde{f}(x_2) \\ \tilde{f}(x_3) \\ 0 \\ 0 \end{bmatrix} \quad (114)$$

Extracting the system corresponding to the first three rows returns the following system, which solves the interior of  $\bar{v}$ , i.e.,  $v$ :

$$\begin{bmatrix} -2 & 1 & 0 \\ 1 & -2 & 1 \\ 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} \tilde{v}(x_1) \\ \tilde{v}(x_2) \\ \tilde{v}(x_3) \end{bmatrix} = \begin{bmatrix} \tilde{f}(x_1) - S \\ \tilde{f}(x_2) \\ \tilde{f}(x_3) \end{bmatrix} \quad (115)$$

## A Special Matrices and Notation

In the code and algebra, a number of the matrices for discretization have special structure.

**Toeplitz Matrices** First, a Toeplitz matrix is one in which every descending diagonal is constant. See [https://en.wikipedia.org/wiki/Toeplitz\\_matrix](https://en.wikipedia.org/wiki/Toeplitz_matrix) for more examples and theory. To create, you need to specify the lower off-diagonals (ordered by rows) and the upper off-diagonals (ordered by columns). They are square only if the number of rows and columns passed in is identical.<sup>2</sup> As an example, to construct

$$vr = [a \ f \ g \ h] \in \mathbb{R}^4 \quad (116)$$

$$vc = [a \ b \ c \ d \ e] \in \mathbb{R}^5 \quad (117)$$

$$\text{toep}(vr, vc) = \begin{bmatrix} a & b & c & d & e \\ f & a & b & c & d \\ g & f & a & b & c \\ h & g & f & a & b \end{bmatrix} \in \mathbb{R}^{4 \times 5} \quad (118)$$

**Banded Matrices** Next, consider a banded matrix ([https://en.wikipedia.org/wiki/Band\\_matrix](https://en.wikipedia.org/wiki/Band_matrix)). This is a sparse matrix where only a “small” number of diagonals below and/or above the main diagonal are non-zero. Unlike a Toeplitz matrix, the diagonals need not be constant.<sup>3</sup>

To denote a banded matrix, one needs to specify (1) how many lower and upper diagonals are in the matrix; and (2) the values of those diagonals. For example, a diagonal matrix has 0 lower and 0 upper diagonals while a tridiagonal matrix has 1 of each.

As an example of the notation, first define vectors of diagonals

$$B_{-1} \equiv [B_{21} \ B_{32} \ B_{43} \ B_{54}] \in \mathbb{R}^4 \quad (119)$$

$$B_0 \equiv [B_{11} \ B_{22} \ B_{33} \ B_{44} \ B_{55}] \in \mathbb{R}^5 \quad (120)$$

$$B_{+1} \equiv [B_{12} \ B_{23} \ B_{34} \ B_{45} \ B_{56}] \in \mathbb{R}^5 \quad (121)$$

$$B_{+2} \equiv [B_{13} \ B_{24} \ B_{35} \ B_{46}] \in \mathbb{R}^4 \quad (122)$$

The  $\text{band}_{\ell,u}(\cdot)$  function is used to define a banded matrix given diagonals where the  $\ell$  is number of bands below the diagonal and  $u$  is the number of bands above the diagonal. The function then takes a list of the diagonals in order (i.e. lower ones, diagonal bands, then upper band)

$$\text{band}_{1,1}(B_{-1}, B_0, B_{+1}) = \begin{bmatrix} B_{11} & B_{12} & 0 & \cdots & \cdots & 0 \\ B_{21} & B_{22} & B_{23} & \ddots & \ddots & \vdots \\ 0 & B_{32} & B_{33} & B_{34} & \ddots & \vdots \\ \vdots & \ddots & B_{43} & B_{44} & B_{45} & 0 \\ \vdots & \ddots & \ddots & B_{54} & B_{55} & B_{56} \end{bmatrix} \in \mathbb{R}^{5 \times 6} \quad (123)$$

<sup>2</sup>This notation is intended to match Matlab’s and Julia’s notation. See <https://www.mathworks.com/help/matlab/ref/toeplitz.html> and <https://github.com/JuliaMatrices/ToeplitzMatrices.jl>. Note that with this interface design, if  $vr_1 \neq vr_1$  there is an error.

<sup>3</sup>In principle there could be banded Toeplitz matrices (i.e., only a certain bandwidth of off diagonals are non-zero), but at this point we are unaware of software packages along those lines.

And other with a variation on the columns with no sub-diagonals and two super-diagonals

$$\text{band}_{0,2}(B_0, B_{+1}, B_{+2}) = \begin{bmatrix} B_{11} & B_{12} & B_{13} & 0 & \cdots & 0 \\ 0 & B_{22} & B_{23} & B_{24} & 0 & \vdots \\ \vdots & 0 & B_{33} & B_{34} & B_{35} & 0 \\ \vdots & \ddots & 0 & B_{44} & B_{45} & B_{46} \\ \vdots & \ddots & \ddots & 0 & B_{55} & B_{56} \end{bmatrix} \in \mathbb{R}^{5 \times 6} \quad (124)$$

Note that a Topelitz matrix where many of the diagonals are 0 can be written as a banded matrix

$$\text{band}_{1,2}(f\mathbf{1}_3, a\mathbf{1}_4, b\mathbf{1}_4, c\mathbf{1}_3) = \text{toep}([a \ f \ 0 \ 0], [a \ b \ c \ 0 \ 0]) \quad (125)$$

$$= \begin{bmatrix} a & b & c & 0 & 0 \\ f & a & b & c & 0 \\ 0 & f & a & b & c \\ 0 & 0 & f & a & b \end{bmatrix} \in \mathbb{R}^{4 \times 5} \quad (126)$$

$$(127)$$

**Tridiagonal Matrices** A final set of matrices are sparse, tridiagonal matrices. This is a particular type of square banded matrix with a single off-diagonal in each direction. As always, the extra structure leads to specialized operations **TODO: START HERE AND CONNECT TO BOTH**

## B Derivation by substitution

One can also derive the formula for  $L_{1-}^B, L_{1+}^B, L_2^B$  in (62), (63), (64) by substitution. For simplicity, here we focus on the case when we have regular grids, i.e.,  $x_{i+1} - x_i = \Delta$  for some  $\Delta > 0$  for all  $i = 0, \dots, M$ .

Using the backwards first-order difference, (21) implies

$$\frac{\tilde{v}(x_1) - \tilde{v}(x_0)}{\Delta} = -\underline{\xi}\tilde{v}(x_0) \quad (128)$$

i.e.,

$$\tilde{v}(x_0) = \frac{1}{1 - \underline{\xi}\Delta} \tilde{v}(x_1) \quad (129)$$

at the lower bound.

Likewise, (22) under the forwards first-order difference yields

$$\frac{\tilde{v}(x_{M+1}) - \tilde{v}(x_M)}{\Delta} = -\bar{\xi}\tilde{v}(x_{M+1}) \quad (130)$$

i.e.,

$$\tilde{v}(x_{M+1}) = \frac{1}{1 + \bar{\xi}\Delta} \tilde{v}(x_M) \quad (131)$$

at the upper bound.



The discretized central difference of second order under (21) at the lower bound is, by substituting (129) in,

$$\frac{\tilde{v}(x_1 + \Delta) - 2\tilde{v}(x_1) + \tilde{v}(x_{\min})}{\Delta^2} = \frac{\tilde{v}(x_1 + \Delta) - \tilde{v}(x_1)}{\Delta^2} - \frac{1}{\Delta} \frac{\tilde{v}(x_1) - \tilde{v}(x_{\min})}{\Delta} \quad (132)$$

$$= \frac{\tilde{v}(x_1 + \Delta) - \tilde{v}(x_1)}{\Delta^2} + \frac{1}{\Delta} \underline{\xi} \tilde{v}(x_1) \quad (133)$$

$$= \frac{1}{\Delta^2} (-1 + \Delta \underline{\xi})^{-1} \tilde{v}(x_1) + \frac{1}{\Delta^2} \tilde{v}(x_1 + \Delta) \quad (134)$$

Similarly, by (22), we have

$$\frac{\tilde{v}(x_{\max}) - 2\tilde{v}(x_M) + \tilde{v}(x_M - \Delta)}{\Delta^2} = \frac{\tilde{v}(x_M - \Delta) - \tilde{v}(x_M)}{\Delta^2} + \frac{1}{\Delta} \frac{\tilde{v}(x_{\max}) - \tilde{v}(x_M)}{\Delta} \quad (135)$$

$$= \frac{\tilde{v}(x_M - \Delta) - \tilde{v}(x_M)}{\Delta^2} - \frac{1}{\Delta} \bar{\xi} \tilde{v}(x_M) \quad (136)$$

$$= \frac{1}{\Delta^2} (-1 - \Delta \bar{\xi})^{-1} \tilde{v}(x_M) + \frac{1}{\Delta^2} \tilde{v}(x_M - \Delta) \quad (137)$$

at the upper bound.

Thus, the corresponding discretized differential operator  $L_{1-}$ ,  $L_{1+}$ , and  $L_2$  are defined as

$$L_{1-}^B \equiv \frac{1}{\Delta} \begin{bmatrix} 1 - (1 - \underline{\xi}\Delta)^{-1} & 0 & 0 & \dots & 0 & 0 & 0 \\ -1 & 1 & 0 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & -1 & 1 & 0 \\ 0 & 0 & 0 & \dots & 0 & -1 & 1 \end{bmatrix}_{M \times M} \quad (138)$$

$$L_{1+}^B \equiv \frac{1}{\Delta} \begin{bmatrix} -1 & 1 & 0 & \dots & 0 & 0 & 0 \\ 0 & -1 & 1 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & -1 & 1 \\ 0 & 0 & 0 & \dots & 0 & 0 & -1 + (1 - \bar{\xi}\Delta)^{-1} \end{bmatrix}_{M \times M} \quad (139)$$

$$L_2^B \equiv \frac{1}{\Delta^2} \begin{bmatrix} -2 - (1 - \underline{\xi}\Delta)^{-1} & 1 & 0 & \dots & 0 & 0 & 0 \\ 1 & -2 & 1 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & -2 & 1 \\ 0 & 0 & 0 & \dots & 0 & 1 & -2 + (1 + \bar{\xi}\Delta)^{-1} \end{bmatrix}_{M \times M} \quad (140)$$