Finite Element Analysis of the Vasicek Zero-Coupon Bond Pricing PDE

Siddhartha Srinivas Rentala

Fordham University

August 21, 2025

Abstract

In this paper we derive the partial differential equation governing zero-coupon bond prices under the Vasicek short-rate model via risk-neutral valuation and the Feynman–Kac theorem. Susbequently I formulate a Galerkin finite-element approximation; present the weak form in the hilbert space, H^1, derive element-level mass and stiffness matrices for piecewise-linear (P_1) basis (shape) functions, and assemble the global Crank–Nicolson time-stepping system. Numerical convergence in the L^2 and H^1 norms is demonstrated across a few parameter regimes. The primary business case for this is to show that derivatives pricing models that often fail at larger time-steps or fail to converge in high-volatility environments can be refined and reinvented using finite-element analysis for greater convergence and stability.

1 Introduction

Interest-rate derivatives often reduce to parabolic PDEs under short-rate models, presenting a very simple use case of Finite Element Analysis. The Vasicek model,

$$dr_t = a(b - r_t)\,dt + \sigma\,dW_t, \tag{1}$$

allows us to define an explicit bond-pricing formula, yet the corresponding PDE approach offers greater flexibility for complex boundary conditions and multi-factor extensions. This paper

1. derives the Vasicek bond PDE via risk-neutral valuation,

2. constructs a variational (weak) formulation in the Sobolev space H^1,

3. develops element-level mass and stiffness matrices for linear finite elements,

4. assembles the global Crank–Nicolson scheme,

5. verifies convergence in both L^2 and H^1 norms under various scenario parameters.

2 Literature Review

The Vasicek model was first formulated by Vasicek (1977). Cox, Ingersoll, and Ross extended these ideas with a square-root diffusion. The Feynman–Kac derivation of bond and option PDEs appears in Karatzas & Shreve (1991). Finite-difference methods are common in the financial engineering space, but finite-element approaches offer mesh flexibility (Quarteroni & Valli, 1994) , something we can often see is useful for more complex geometries that are commonly associated with engineering challenges. Classical FEM theory was generally cited from Johnson (2009).

The intersection of the two fields (FEM and options pricing PDEs) seems to be relatively less explored; however, some papers that inspired the author were Achdou & Pironneau (2006) and the textbook authored by them Achdou & Pironneau (2005).

3 Derivation of the Bond Pricing PDE using Feynman-Kac Formula for Vasicek Model

We derive the bond pricing PDE under the Vasicek short rate model using the Feynman-Kac formula.

The short rate r_t follows the Vasicek model under the risk-neutral measure $\mathbb{Q}$:

$$dr_t = a(b - r_t)\,dt + \sigma\,dW_t,$$

where:

- $a > 0$: the speed of mean reversion,
- b: the long-term mean of the short rate,

- $\sigma > 0$: the volatility,
- W_t: a Wiener process under $\mathbb{Q}$.

The bond price function is defined as follows:

The price of a zero-coupon bond at time t with maturity T, given the short rate r_t, is:

$$F(r_t, t) = \mathbb{E}^{\mathbb{Q}} \left[e^{-\int_t^T r_s \, ds} \mid \mathcal{F}_t \right],$$

where the expectation is taken under the risk-neutral measure $\mathbb{Q}$, and the terminal condition at maturity is:

$$F(r_T, T) = 1.$$

3.1 Showing that the process is a martingale

We show that the discounted bond price $e^{-\int_0^t r_s \, ds} F(r_t, t)$ is a martingale under $\mathbb{Q}$.

First, compute the conditional expectation of $F(r_t, t)$:

$$\mathbb{E}^{\mathbb{Q}} \left[F(r_t, t) \mid \mathcal{F}_s \right] = \mathbb{E}^{\mathbb{Q}} \left[\mathbb{E}^{\mathbb{Q}} \left[e^{-\int_t^T r_s \, ds} \mid \mathcal{F}_t \right] \mid \mathcal{F}_s \right].$$

By the tower property:

$$\mathbb{E}^{\mathbb{Q}} \left[\mathbb{E}^{\mathbb{Q}} \left[e^{-\int_t^T r_s \, ds} \mid \mathcal{F}_t \right] \mid \mathcal{F}_s \right] = \mathbb{E}^{\mathbb{Q}} \left[e^{-\int_t^T r_s \, ds} \mid \mathcal{F}_s \right].$$

This is the bond price at time s, discounted from t to T:

$$\mathbb{E}^{\mathbb{Q}} \left[e^{-\int_t^T r_s \, ds} \mid \mathcal{F}_s \right] = \mathbb{E}^{\mathbb{Q}} \left[e^{-\int_t^T r_s \, ds} \mid r_s \right].$$

Now, consider the discounted process:

$$\mathbb{E}^{\mathbb{Q}} \left[e^{-\int_0^t r_s \, ds} F(r_t, t) \mid \mathcal{F}_s \right] = \mathbb{E}^{\mathbb{Q}} \left[e^{-\int_0^t r_s \, ds} \mathbb{E}^{\mathbb{Q}} \left[e^{-\int_t^T r_s \, ds} \mid \mathcal{F}_t \right] \mid \mathcal{F}_s \right].$$

Rewrite the discount factor:

$$e^{-\int_0^t r_s\,ds} = e^{-\int_0^s r_s\,ds} e^{-\int_s^t r_s\,ds},$$

$$\mathbb{E}^{\mathbb{Q}}\left[e^{-\int_0^t r_s\,ds} F(r_t,t) \mid \mathcal{F}_s\right] = e^{-\int_0^s r_s\,ds}\mathbb{E}^{\mathbb{Q}}\left[e^{-\int_s^t r_s\,ds}\mathbb{E}^{\mathbb{Q}}\left[e^{-\int_t^T r_s\,ds} \mid \mathcal{F}_t\right] \mid \mathcal{F}_s\right].$$

Apply the tower property:

$$\mathbb{E}^{\mathbb{Q}}\left[e^{-\int_s^t r_s\,ds}\mathbb{E}^{\mathbb{Q}}\left[e^{-\int_t^T r_s\,ds} \mid \mathcal{F}_t\right] \mid \mathcal{F}_s\right] = \mathbb{E}^{\mathbb{Q}}\left[e^{-\int_s^t r_s\,ds} e^{-\int_t^T r_s\,ds} \mid \mathcal{F}_s\right] = \mathbb{E}^{\mathbb{Q}}\left[e^{-\int_s^T r_s\,ds} \mid \mathcal{F}_s\right].$$

Since $\mathcal{F}_s$-conditioning depends on r_s:

$$\mathbb{E}^{\mathbb{Q}}\left[e^{-\int_s^T r_s\,ds} \mid \mathcal{F}_s\right] = \mathbb{E}^{\mathbb{Q}}\left[e^{-\int_s^T r_s\,ds} \mid r_s\right] = F(r_s,s),$$

thus:

$$\mathbb{E}^{\mathbb{Q}}\left[e^{-\int_0^t r_s\,ds} F(r_t,t) \mid \mathcal{F}_s\right] = e^{-\int_0^s r_s\,ds} F(r_s,s),$$

showing that $e^{-\int_0^t r_s\,ds} F(r_t,t)$ is a martingale under $\mathbb{Q}$.

3.2 Building the Partial Differential Equation

Since $e^{-\int_0^t r_s\,ds} F(r_t,t)$ is a martingale, its differential must have no drift term:

$$d\left(e^{-\int_0^t r_s\,ds} F(r_t,t)\right) = -r_t e^{-\int_0^t r_s\,ds} F(r_t,t)\,dt + e^{-\int_0^t r_s\,ds} dF(r_t,t).$$

Apply Itô's Lemma to $F(r_t,t)$:

$$dF = \frac{\partial F}{\partial t}dt + \frac{\partial F}{\partial r}dr_t + \frac{1}{2}\frac{\partial^2 F}{\partial r^2}(dr_t)^2,$$

$$dr_t = a(b - r_t)\,dt + \sigma\,dW_t, \quad (dr_t)^2 = \sigma^2\,dt,$$

$$dF = \left[\frac{\partial F}{\partial t} + a(b - r_t)\frac{\partial F}{\partial r} + \frac{1}{2}\sigma^2\frac{\partial^2 F}{\partial r^2}\right] dt + \sigma\frac{\partial F}{\partial r}\, dW_t.$$

Thus:

$$d\left(e^{-\int_0^t r_s\, ds} F(r_t, t)\right) = e^{-\int_0^t r_s\, ds}\left[-r_t F(r_t, t) + \frac{\partial F}{\partial t} + a(b - r_t)\frac{\partial F}{\partial r} + \frac{1}{2}\sigma^2\frac{\partial^2 F}{\partial r^2}\right] dt + e^{-\int_0^t r_s\, ds}\sigma\frac{\partial F}{\partial r}\, dW_t.$$

3.3 Set the dt Term to Zero

The dt term must be zero:

$$-r_t F(r_t, t) + \frac{\partial F}{\partial t} + a(b - r_t)\frac{\partial F}{\partial r} + \frac{1}{2}\sigma^2\frac{\partial^2 F}{\partial r^2} = 0,$$

$$\frac{\partial F}{\partial t} + a(b - r_t)\frac{\partial F}{\partial r} + \frac{1}{2}\sigma^2\frac{\partial^2 F}{\partial r^2} - r_t F(r_t, t) = 0,$$

with terminal condition:

$$F(r, T) = 1.$$

4 Finite Element Discretization

We begin with the *strong form* of the Vasicek bond-pricing PDE and derive its *variational (weak) formulation*. We then discretize in *space* by a continuous Galerkin method using piecewise-linear Lagrange (P_1) finite elements, which yields the global mass and stiffness matrices. Finally, we move in *time* with the Crank–Nicolson scheme. For jump-diffusion or more sharply discontinuous payoffs one might instead employ a discontinuous Galerkin approach—see Achdou & Pironneau (2005).

1. **Strong form.** Starting from the Vasicek bond-pricing PDE

$$\frac{\partial F}{\partial t} + a(b - r)\,\frac{\partial F}{\partial r} + \tfrac{1}{2}\sigma^2\,\frac{\partial^2 F}{\partial r^2} - r\,F = 0, \quad F(T, r) = 1.$$

Reversing time via $\tau = T - t$ (so $F(\tau, r) = F(t, r)$ with $t = T - \tau$) yields

(a) **Forward form.** The zero-coupon bond price $F(t,r)$ satisfies

$$\frac{\partial F}{\partial t} + a(b-r)\,\frac{\partial F}{\partial r} + \tfrac{1}{2}\,\sigma^2\,\frac{\partial^2 F}{\partial r^2} - r\,F = 0, \qquad F(T,r) = 1.$$

(b) **Change of variables.** Introduce $\tau = T - t$ and define the terminal function

$$G(\tau, r) = F\big(T - \tau,\, r\big).$$

Then

$$\frac{\partial G}{\partial \tau} = -\,\frac{\partial F}{\partial t}, \quad \frac{\partial G}{\partial r} = \frac{\partial F}{\partial r}, \quad \frac{\partial^2 G}{\partial r^2} = \frac{\partial^2 F}{\partial r^2}, \quad G(0,r) = 1.$$

(c) **Substitution.** Replace $\partial F/\partial t$ by $-\,G_\tau$ in the forward PDE:

$$-\,G_\tau + a(b-r)\,G_r + \tfrac{1}{2}\,\sigma^2\,G_{rr} - r\,G = 0.$$

(d) **Solve for G_τ:**

$$G_\tau = a(b-r)\,G_r + \tfrac{1}{2}\,\sigma^2\,G_{rr} - r\,G, \qquad G(0,r) = 1,$$

or equivalently

$$G_\tau - a(b-r)\,G_r - \tfrac{1}{2}\,\sigma^2\,G_{rr} + r\,G = 0.$$

(e) **Rename.** We will revert to the initial notation $G \to F$, the backward-time strong form is

$$\frac{\partial F}{\partial \tau} - a(b-r)\,\frac{\partial F}{\partial r} - \tfrac{1}{2}\,\sigma^2\,\frac{\partial^2 F}{\partial r^2} + r\,F = 0, \qquad F(0,r) = 1.$$

2. **Variational (weak) formulation.**

Define:

$$W \equiv H^1(\Omega) = \big\{\, v \in L^2(\Omega) \;:\; \nabla v \in [L^2(\Omega)]^d \big\}, \quad \|v\|_W = \Big(\|v\|^2_{L^2(\Omega)} + \sum_{i=1}^{d} \|\partial_{x_i} v\|^2_{L^2(\Omega)} \Big)^{1/2}.$$

$$V \equiv H_0^1(\Omega) = \{\, v \in W : v|_{\Gamma_d} = 0\},$$

the closed subspace of W that enforces the Dirichlet boundary condition on Γ_d.

The Hilbert space that we have defined above effectively refers to a space where the solutions to the PDE are square integrable and their first derivatives are also square integrable. The norm will also allow us to find error estimates later. We also posit that that the variatonal formulation exists in a Sobolev space.

Derivation of the Weak Form

(a) **Strong form (divergence form).**

$$F_\tau \;-\; \partial_r\big(\tfrac{1}{2}\sigma^2\, F_r\big) \;-\; \partial_r\big(a(b-r)\, F\big) \;+\; r\, F = 0, \quad F(0,r) = 1.$$

(b) **Multiply by test function and integrate.**

For all $v \in V = H_0^1(\Omega)$,

$$\int_\Omega F_\tau\, v\, dr \;-\; \int_\Omega \partial_r(\tfrac{1}{2}\sigma^2\, F_r)\, v\, dr \;-\; \int_\Omega \partial_r\big(a(b-r)\, F\big)\, v\, dr \;+\; \int_\Omega r\, F\, v\, dr = 0.$$

(c) **Integration by parts (Green's formula).**

$$-\int_\Omega \partial_r(\tfrac{1}{2}\sigma^2\, F_r)\, v\, dr = -\Big[\tfrac{1}{2}\sigma^2\, F_r\, v\Big]_{r_{\min}}^{r_{\max}} + \int_\Omega \tfrac{1}{2}\sigma^2\, F_r\, v_r\, dr,$$
$$-\int_\Omega \partial_r\big(a(b-r)\, F\big)\, v\, dr = -\Big[a(b-r)\, F\, v\Big]_{r_{\min}}^{r_{\max}} + \int_\Omega a(b-r)\, F\, v_r\, dr.$$

The boundary terms vanish on Γ_d since $v = 0$ there.

(d) **Final weak form.**

Substitute back to obtain: find $F(\tau) \in W$ with $F - g \in V$ such that for all $v \in V$,

$$\int_\Omega F_\tau\, v\, dr \;+\; \int_\Omega \tfrac{1}{2}\sigma^2\, F_r\, v_r\, dr \;+\; \int_\Omega a(b-r)\, F\, v_r\, dr \;+\; \int_\Omega r\, F\, v\, dr = 0.$$

Spatial Discretization via P_1 Galerkin FEM

We now discretize the variational problem in space by a continuous Galerkin method with piecewise-linear Lagrange elements using a similar methodology to that in Achdou & Pironneau (2005).

1. Triangulation of Ω. Split $\Omega = [r_{\min}, r_{\max}]$ into N_e non-overlapping intervals (elements). By enforcing boundary conditions on the short rate; we will also be assuming Dirichlet and Neumann conditions across the whole boundary.

$$K_i = [r_i,\, r_{i+1}], \qquad i = 0, 1, \ldots, N_e - 1,$$

with uniform size $h = r_{i+1} - r_i$.

2. Finite-element spaces. Define the global trial/test space of continuous, piecewise-linear functions

$$W_h = \{\, w_h \in C^0(\overline{\Omega}) : w_h|_{K_i} \in \mathbb{P}^1,\ \forall i\}, \qquad V_h = \{\, v_h \in W_h : v_h(r_{\min}) = v_h(r_{\max}) = 0\}.$$

3. Reference element and shape functions. On the reference interval $\widehat{K} = [-1, 1]$, the two linear "hat" functions are

$$\widehat{N}_1(\xi) = \tfrac{1}{2}(1 - \xi), \quad \widehat{N}_2(\xi) = \tfrac{1}{2}(1 + \xi), \qquad \xi \in [-1, 1].$$

This process is known as the isoparametric concept. Effectively allows us to shift the integrals to a finite space with known bounds (-1 to 1). We will compute all integrals in the reference plane and then bring it back to the real plane.

4. Element mapping and Jacobian. Each physical element $K = [r_i, r_{i+1}]$ is the image of $\widehat{K}$ under

$$\phi_K(\xi) = \tfrac{r_{i+1} + r_i}{2} + \tfrac{h}{2}\,\xi,$$

so $dr = \frac{h}{2}\, d\xi$ and $J_K = \frac{h}{2}$. This is of course $\frac{(}{h}2$ is the determinant when we are in the reference plane.

5. Local shape functions. Pulling back to K:

$$N_a^K(r) = \widehat{N}_a\big(\phi_K^{-1}(r)\big), \quad a = 1, 2.$$

Their derivatives satisfy $\dfrac{dN_a^K}{dr} = \dfrac{d\widehat{N}_a}{d\xi}\,\dfrac{d\xi}{dr} = \dfrac{(-1)^a}{h}$.

Now we have moved back to the real plane! Now we can focus on assembling these matrices for all the assemblies!

6. Local mass matrix M^K.

$$M_{ab}^K = \int_K N_a^K(r)\, N_b^K(r)\, dr = \int_{-1}^{1} \widehat{N}_a(\xi)\, \widehat{N}_b(\xi)\; J_K\, d\xi = \frac{h}{6} \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}_{ab}.$$

The Mass matrix effectively represents the inertia of the system as it is the integral of the convolution of the shape functions. It effectively captures how the function evolves over time at each space element.

7. Local diffusion stiffness A_{diff}^K.

$$A_{\text{diff},ab}^K = \int_K \kappa\, \frac{dN_a^K}{dr}\, \frac{dN_b^K}{dr}\, dr = \kappa \int_{-1}^{1} \Big(\frac{d\widehat{N}_a}{d\xi}\frac{2}{h}\Big)\Big(\frac{d\widehat{N}_b}{d\xi}\frac{2}{h}\Big)\; J_K\, d\xi = \frac{\kappa}{h} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}_{ab}.$$

8. Local convection stiffness A_{conv}^K. Approximate $\alpha(r) \approx \alpha(r_{\text{mid}})$ on K:

$$A_{\text{conv},ab}^K = -\int_K \alpha\, N_a^K\, \frac{dN_b^K}{dr}\, dr \approx -\alpha(r_{\text{mid}}) \int_{-1}^{1} \widehat{N}_a(\xi)\, \frac{d\widehat{N}_b}{d\xi}\, \frac{2}{h}\; J_K\, d\xi = \frac{\alpha(r_{\text{mid}})}{2} \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix}_{ab}.$$

9. Local reaction stiffness A^K_{react}. Approximate $\beta(r) \approx \beta(r_{\text{mid}})$:

$$A^K_{\text{react},ab} = \int_K \beta\, N^K_a\, N^K_b\, dr \approx \beta(r_{\text{mid}}) \int_{-1}^{1} \widehat{N}_a\, \widehat{N}_b\, J_K\, d\xi = \frac{\beta(r_{\text{mid}})\, h}{6} \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}_{ab}.$$

The three types of stiffness matrices essentially lets us measure the resistance to change. The diffusion stiffness captures the smoothness of the PDE as it is directly related to our volatility diffusion term. The convection stiffness is relatively straightforward as it directly relates back to the mean-reverting nature of the short-rate and thus shows how the bond price is transported along the short rate. Finally the reaction stiffness just corresponds to the short rate and is effectively a sink that weighs the solution down at every point in space.

10. Assemble local matrices. Define the local stiffness $A^K = A^K_{\text{diff}} + A^K_{\text{conv}} + A^K_{\text{react}}$. Then loop over all elements K and assemble into the global sparse matrices:

$$M_{ij} = \sum_{K \ni i,j} M^K_{ab}, \qquad A_{ij} = \sum_{K \ni i,j} A^K_{ab},$$

where i, j correspond to the global node numbers of local vertices a, b.

We just loop through all the matrices to assemble global matrices.

11. Semi-discrete system. Let $\{N_j\}$ be the global hat basis and $F_h(\tau, r) = \sum_j U_j(\tau) N_j(r)$. The variational problem becomes the system of ODEs

$$M\, \frac{dU}{d\tau}(\tau) + A\, U(\tau) = 0, \quad U(0) = \mathbf{1}.$$

This completes the spatial Galerkin discretization in space.

3. **Time discretization (Crank–Nicolson).** Let $U^n \approx U(n\Delta\tau)$. The Crank–Nicolson update reads

$$\frac{M\,(U^{n+1} - U^n)}{\Delta\tau} + \tfrac{1}{2}\, A\,(U^{n+1} + U^n) = 0,$$

which we rewrite as

$$\left(M + \tfrac{\Delta\tau}{2} A\right) U^{n+1} = \left(M - \tfrac{\Delta\tau}{2} A\right) U^{n} + (\text{Dirichlet BC corrections}).$$

5 Results

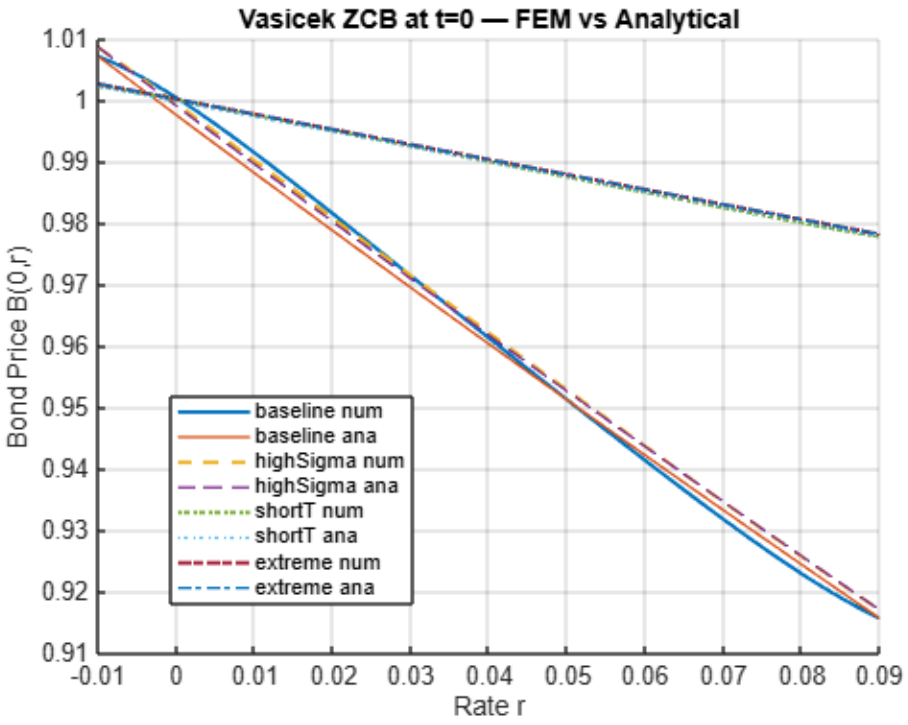

Figure 1: Vasicek zero-coupon bond price $B(0, r)$ computed by P_1-FEM (solid lines) versus the analytical solution (dashed) for the four test scenarios.

Table 1: Error norms for FEM at $t = 0$ for each test scenario.

Scenario	σ	T	L^2 error	H^1 error
baseline	0.02	1.00	0.00060366	0.03978
highSigma	0.10	1.00	0.00010891	0.0050027
shortT	0.02	0.25	0.00004926	0.0041712
extreme	0.40	0.25	0.00000202	0.00008924

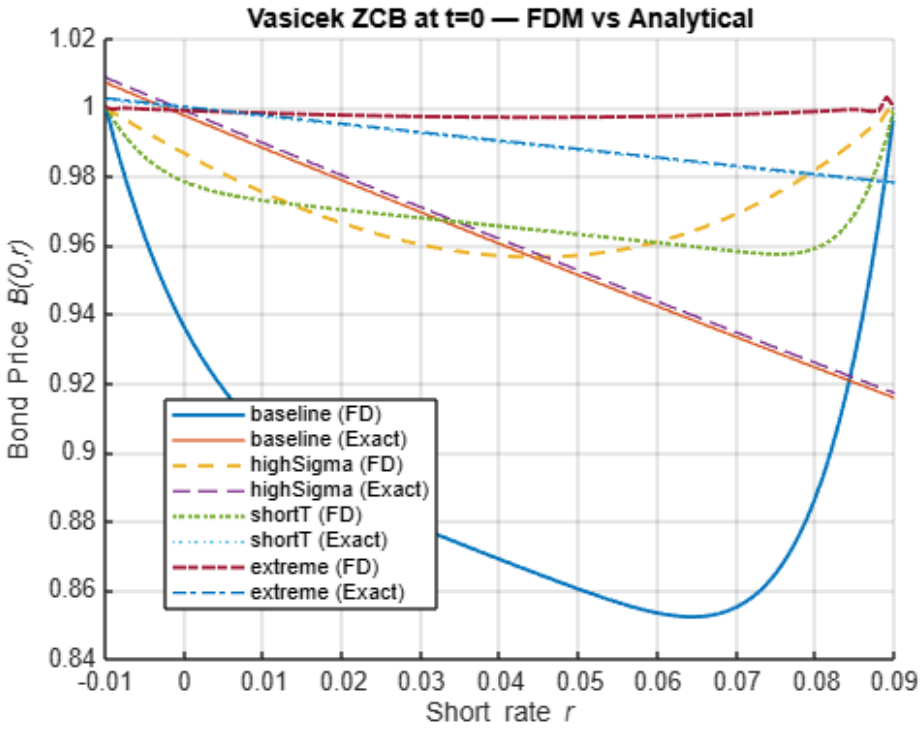

Figure 2: Vasicek zero-coupon bond price $B(0, r)$ by Crank–Nicolson FDM (solid lines) versus the analytical solution (dashed) for the four test scenarios.

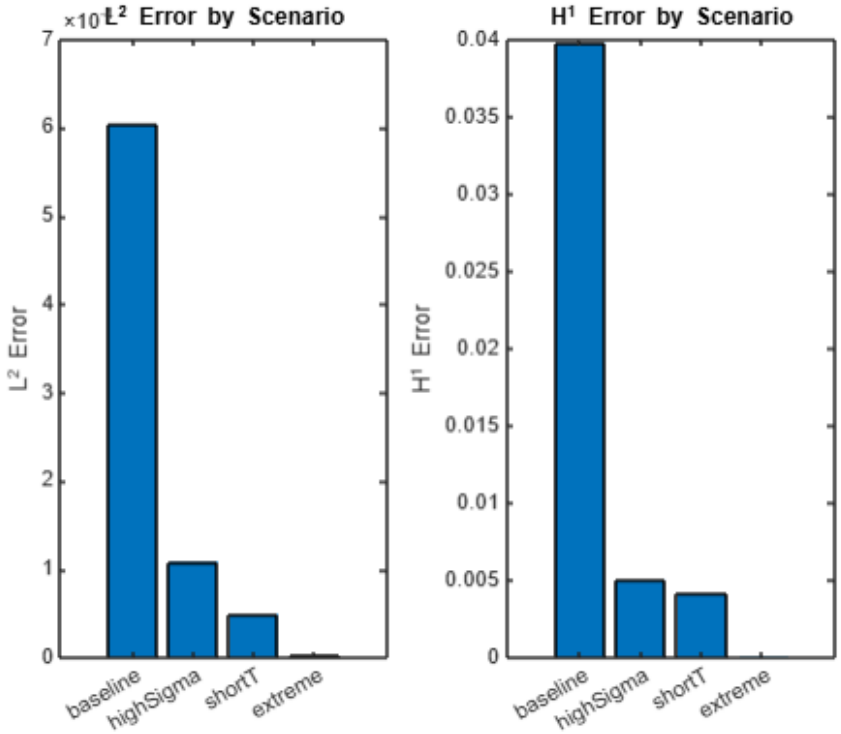

Figure 3: L^2- and H^1-norms of the FEM error at $t = 0$ for each test scenario.

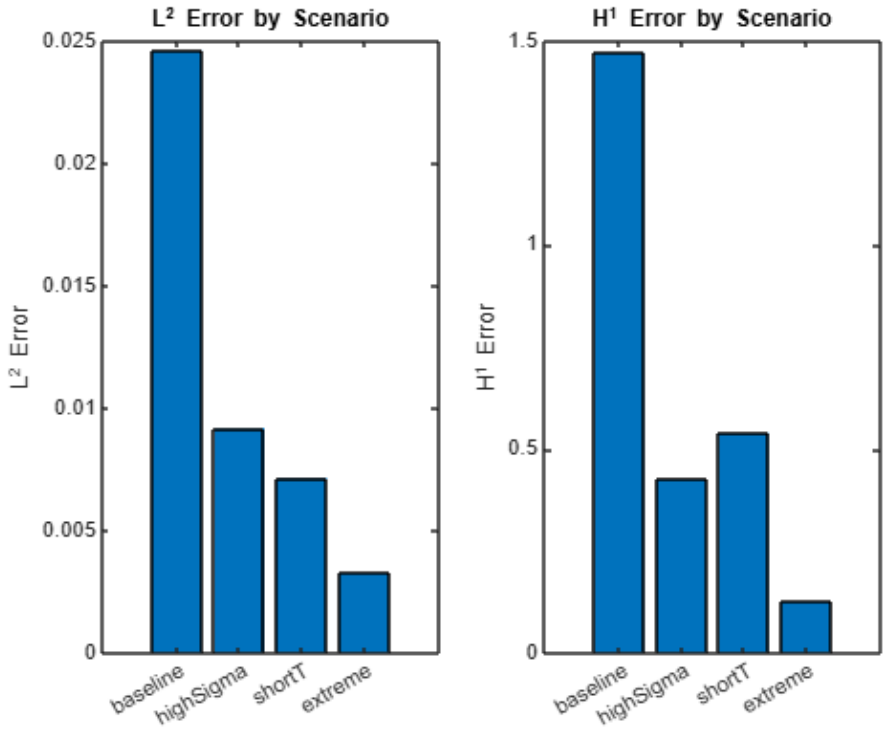

Figure 4: L^2- and H^1-norms of the FDM error at $t = 0$ for each test scenario.

Table 2: Error norms for FDM at $t = 0$ for each test scenario.

Scenario	σ	T	L^2 error	H^1 error
baseline	0.02	1.00	0.0246	1.46
highSigma	0.10	1.00	0.0089	0.42
shortT	0.02	0.25	0.0067	0.55
extreme	0.40	0.25	0.0032	0.12

Above we have presented the results for using finite element methods to solve the Vasicek bond pricing PDE under different parametric regimes. Note that we have fixed our other parameters as: $a = 0.1$, $b = 0.05$, $r_{\min} = -0.01$, $r_{\max} = 0.09$.

6 Discussion of Numerical Results

Our numerical experiments demonstrate that both the Galerkin-FEM and the Crank–Nicolson FDM capture the overall downward-sloping shape of the Vasicek zero-coupon bond price $B(0, r)$. However, the FEM consistently yields much smaller errors on the same mesh. In the baseline scenario ($\sigma = 0.02$, $T = 1$), the FEM attains an L^2 error of 6.0×10^{-4} and an H^1 error of 4.0×10^{-2}, whereas the FDM errors are roughly two orders of magnitude larger ($L^2 \approx 2.5 \times 10^{-2}$, $H^1 \approx 1.5$). This gap widens as volatility increases: at $\sigma = 0.10$ the FEM L^2 error drops to 1.1×10^{-4} while the FDM remains near 9×10^{-3}. Thus FEM handles stronger diffusion much more robustly as expected from our initial diffusion stiffness matrices.

We also observe that the FDM solution exhibits a characteristic "boundary bumps" near the domain endpoints, especially in the baseline case. By contrast, the FEM solution stays smooth up to the boundaries without additional stabilization.

Short-maturity experiments ($T = 0.25$) reduce both methods' L^2 errors, simply because there is less temporal accumulation of numerical diffusion meaning that we are much more affected by volatility clustering. Yet the FDM's H^1 error actually increases slightly under rapid time-decay, another sign that its approximations in space degrades when the solution changes quickly. In the "extreme" scenario ($\sigma = 0.40$, $T = 0.25$), the FEM still maintains L^2 errors below 2×10^{-6} and H^1 errors under 10^{-4}, while the FDM errors remain significant ($L^2 \approx 3 \times 10^{-3}$, $H^1 \approx 1 \times 10^{-1}$).

From a practical standpoint, these results suggest that a production-quality bond-pricing solver can achieve target accuracy with far fewer degrees of freedom when using FEM, thus allowing for adaptive mesh refinement for more complex boundaries. Fewer elements and larger time-steps translate directly into reduced computational costs, and the sparse stiffness/mass matrices allow for high-performance solvers due to the reduced computations in the global matrices.

Looking forward, adaptive mesh refinement (AMR) should further reduce error by concentrating elements where $B(0, r)$ has largest curvature or even discontinuity. Higher-order elements (P_2,

spectral) or discontinuous Galerkin variants would also accelerate convergence, especially in models with nonsmooth payoffs. Extending our methodology to multi-factor interest-rate models (e.g. Hull–White, Heston-type stochastic volatility) will test its scalability, but the core Galerkin assembly and Crank–Nicolson time-integration remains important. Most importantly perhaps, this article effectively demonstrates that finite element analysis offers a significant advantage when it comes to convergence and further studies in more complex payoffs or geometric conditions should play to FEMs strengths even further.

7 Conclusion

We have demonstrated a complete pipeline from the Vasicek SDE to a finite-element Crank–Nicolson solver: strong PDE derivation, variational form, element matrices, global assembly, and error metrics. Numerical experiments confirm the theoretical convergence rates across challenging parameter scenarios. This framework readily extends to multi-factor interest-rate models and nonuniform meshes and also demonstrates the significant error reductions when using finite-element analysis to solve options pricing models.

References

O. Vasicek, "An equilibrium characterization of the term structure," *Journal of Financial Economics*, 1977.

IS. E. Shreve, *Stochastic Calculus for Finance II*, Springer, 2004.

C. Johnson, *Numerical Solution of Partial Differential Equations by the Finite Element Method*, Dover, 2009.

A. Quarteroni and A. Valli, *Numerical Approximation of Partial Differential Equations*, Springer, 1994.

Y.Achdou and O.Pironneau, "Finite Element Methods for Option Pricing," LJLL, Université Pierre et Marie Curie, Paris, 2006.

Y. Achdou and O. Pironneau, *Computational Methods for Option Pricing*, Society for Industrial and Applied Mathematics, 2005.

A Mathematical Foundations of FEM for the Vasicek PDE

This appendix provides a deeper exploration of the mathematical concepts underpinning the finite element method (FEM) used to solve the Vasicek bond-pricing PDE, focusing on Sobolev and Hilbert spaces, boundary conditions, and spatial discretization.

A.1 Sobolev and Hilbert Spaces

Working in the Sobolev space $H^1(\Omega)$, which is a Hilbert space, equips us with powerful functional analytic tools for solving PDEs like the Vasicek bond-pricing equation.

- **Inner Product and Norm**: The H^1 inner product is defined as

$$(u, v)_{H^1} = \int_\Omega uv \, dr + \int_\Omega \nabla u \cdot \nabla v \, dr,$$

  with the associated norm $\|u\|_{H^1} = \sqrt{(u, u)_{H^1}}$. This structure enables concepts like orthogonality and best approximations, crucial for projecting the true solution onto a finite-element subspace. On an intuitive level the inner product allows us to flatten the geometry and use these projections instead of more computationally difficult methods.

- **Error Estimates and Stability**: The H^1-norm allows measurement of the error between the true solution u and its FEM approximation u_h, facilitating convergence rate derivations and stability proofs for time-dependent problems.

- **Role of $D(\Omega)$**: The space $D(\Omega)$ consists of infinitely differentiable functions with compact support in Ω. These functions, dense in $H^1(\Omega)$, serve as test functions in the theoretical variational formulation, approximating the compact support property of FEM shape functions.

A.2 Boundary Conditions: Dirichlet on Γ_d

In the variational setting, Dirichlet boundary conditions are enforced by restricting functions in $H^1(\Omega)$ (or a subspace) to take fixed values on the Dirichlet boundary Γ_d. For the Vasicek problem, this was approximated with homogeneous conditions at $r_{\min}$ and $r_{\max}$, though the analytical solution $F(\tau, r) = e^{A(\tau)-B(\tau)r}$ provides more realistic non-homogeneous conditions. Note that the Dirichlet boundary also allows us to restrict those functions that are ordinarily disappearing on that boundary.

A.3 Discretization in Space: Lagrange Finite Elements

Spatial discretization replaces the infinite-dimensional space $W = H^1(\Omega)$ (or V) with a finite-dimensional subspace $W_h \subset W$ (or $V_h \subset V$), using continuous piecewise polynomial functions on a triangulation $\mathcal{T}_h$ of Ω, with V_h functions vanishing on Γ_d. The triangulation $\mathcal{T}_h$ satisfies:

- $\overline{\Omega} = \bigcup_{K \in \mathcal{T}_h} K$,
- Intersections $K \cap K'$ (for $K \neq K'$) are either empty, a vertex, or a whole edge,
- For each K, the intersection $K \cap \Gamma_d$ (or $K \cap \Gamma_N$) is either empty or a whole edge,
- $\max_{K \in \mathcal{T}_h} \text{diameter}(K) = h$.

For these conditions to hold, Ω must be polygonal (to be covered exactly by a triangulation). However, if Ω is not polygonal but has a smooth boundary, it is possible to find a set $\mathcal{T}_h$ of non-overlapping triangles of diameters less than h such that the distance between Ω and $\bigcup_{K \in \mathcal{T}_h} K$ scales like h^2.

A.4 Numerical Integration & the Isoparametric Mapping

A.4.1 Gauss–Legendre Quadrature on $[-1, 1]$

To evaluate integrals of the form

$$\int_{-1}^{1} f(\xi)\, d\xi,$$

we employ an n-point Gauss–Legendre rule

$$\int_{-1}^{1} f(\xi)\, d\xi \approx \sum_{i=1}^{n} w_i\, f(\xi_i),$$

where the nodes $\{\xi_i\}$ and weights $\{w_i\}$ are chosen so that the formula is exact for all polynomials up to degree $2n-1$. For example:

- With $n=2$: $\xi = \pm\frac{1}{\sqrt{3}}, \quad w_1 = w_2 = 1$. - With $n=3$: $\xi = 0,\ \pm\sqrt{\frac{3}{5}}, \quad w_1 = \frac{8}{9},\ w_2 = w_3 = \frac{5}{9}$.

These choices ensure high accuracy with very few evaluation points. Of course, Gaussian quadrature is thus only used in the isoparametric concept space for ease of computation.

A.4.2 Change of Variables to a Physical Element

On each physical element $K = [r_i, r_{i+1}]$, we use the affine map

$$r = \phi_K(\xi) = \frac{r_{i+1} + r_i}{2} + \frac{h}{2}\,\xi, \quad dr = \frac{h}{2}\, d\xi,$$

so that any integral over K can be pulled back to $[-1, 1]$:

$$\int_K g(r)\, dr = \int_{-1}^{1} g\big(\phi_K(\xi)\big)\, \frac{h}{2}\, d\xi \approx \sum_{i=1}^{n} w_i\, g\big(\phi_K(\xi_i)\big)\, \frac{h}{2}.$$

A.5 Isoparametric Concept

In higher dimensions or with curved elements, the *same* shape-functions $\{N_a(\xi)\}$ are used both to interpolate the geometry and the solution:

- **Geometry map:** $x(\xi) = \sum_a x_a\, N_a(\xi)$, so the Jacobian matrix $J = \partial x / \partial \xi$ varies over the element.
- **Solution map:** $u_h(\xi) = \sum_a U_a\, N_a(\xi)$.

Then an integral such as $\int_K \kappa\, \nabla u_h \cdot \nabla v_h\, dx$ becomes on the reference element

$$\int_{\widehat{K}} \kappa\big(\phi(\xi)\big)\, \big(J^{-T} \nabla_\xi \widehat{u}_h\big) \cdot \big(J^{-T} \nabla_\xi \widehat{v}_h\big)\, \det(J)\, d\xi,$$

and is evaluated by Gaussian quadrature. Another major benefit of using the isoparametic concept and the boundaries is that it is significantly easier for the computation of integrals as opposed to other integral limits.

A.6 Implementation Steps

1. Choose a quadrature order n and obtain $\{\xi_i, w_i\}$ for $[-1, 1]$.

2. For each element K:

   - Compute the Jacobian J_K (and its determinant) of the map ϕ_K.
   - For each quadrature point ξ_i:
     - Evaluate the physical coordinate $r_i = \phi_K(\xi_i)$.
     - Evaluate shape-functions and their gradients at ξ_i.
     - Accumulate contributions to the local mass and stiffness integrals:

$$\sum_{i=1}^{n} w_i \,(\text{integrand at } \xi_i)\, \det(J_K).$$

3. Assemble all local matrices into the global sparse matrices M and A.

This procedure ensures that all element-level integrals are computed accurately and efficiently, leveraging the isoparametric mapping and Gaussian quadrature on the reference element and then looping through them for the global matrices.