PDE Lattice for Hull-White model

Lecture notes

Vladimir Piterbarg
Bank of America
1999-2000

1 Hull-White Model

- Formulas using *short rate state* $x(t)$. Here a is mean reversion parameter, σ is volatility, $F(0,t)$ is instantaneous forward rate (known at time 0)

$$\begin{aligned} r(t) &= F(0,t) + x(t), \\ dx(t) &= (\theta(t) - ax(t))\, dt + \sigma\, dW(t). \end{aligned}$$

- The process $x(t)$

$$\begin{aligned} x(T) &= x(t)\, e^{-a(T-t)} + I(T) - I(t)\, e^{-a(T-t)} + \sigma \int_t^T e^{-a(T-s)}\, dW_s, \\ I(\tau) &= \frac{\sigma^2}{2} b^2(\tau), \quad b(\tau) = \frac{1-e^{-a\tau}}{a} \end{aligned} \tag{1}$$

- Bond prices are functions of short rate:

$$P(t,T) = p(x(t), t, T),$$

where the (deterministic!) function $p(x,t,T)$ is given by

$$\begin{aligned} p(x,t,T) &= \frac{P(0,T)}{P(0,t)} \exp\{-b(T-t)\cdot x + a(t,T)\}, \\ a(t,T) &= -\frac{\sigma^2}{2} \frac{1-e^{-2at}}{2a} b^2(T-t). \end{aligned}$$

- HW model is Markovian in one factor:
  - Factor is short-rate state $x(t)$ which is a Markov process (see formula (1)).
  - The whole term structure at time t is a deterministic function of factor $x(t)$.

2 Rollback

- Rollback = backward induction.
- Given the value of an instrument at time T, compute its value at some previous time t, $t < T$.
- Values of *all* instruments (non-path-dependent) are deterministic functions of the short rate state x!
- Simple example: rollback a bond $P(\cdot, T)$
  - At maturity time T, the value of a bond is a function of $x = x(T)$ that is equal to 1.0 for all x;
  - At some previous time t, (see formulas before)
  $$P(t, T) = p(x(t), t, T).$$
- Same applies to all instruments. "Rollback" is a procedure where
  $$\text{Value at } T \text{ as a function of } x(T) \Longrightarrow \text{Value at } t \text{ as a function of } x(t).$$
- Suppose we have a payoff X at time T that is a (deterministic) function of $x(T)$, such that
  $$X = V(x(T)).$$
  Denote the result of rollback (deterministic function of $x(t)$) by
  $$\rho\{V, T \to t\}(x),$$
  so that we have (compare with the diagram above) for the symbolic representation of rollback
  $$V(x(T)) \Longrightarrow \rho\{V, T \to t\}(x(t)).$$
- See Figure 1.

3 Forward measure in rollback

- From general risk-neutral valuation result,

$$\begin{aligned}\rho\{V, T \to t\}(x(t)) &= \pi_t(X) \\ &= \mathbf{E}\left(e^{-\int_t^T r(s)\,ds} X \,\middle|\, x(t) \right) \\ &= \mathbf{E}\left(e^{-\int_t^T r(s)\,ds} V(x(T)) \,\middle|\, x(t) \right).\end{aligned}$$

  Apply T-forward measure

$$\rho\{V, T \to t\}(x(t)) = P(t, T)\,\mathbf{E}^T\left(V(x(T)) \,|\, x(t) \right).$$

- Of course $P(t, T)$ is a deterministic function of $x(t)$.
- We would like to write "$\mathbf{E}^T$ part" as a deterministic function of $x(t)$ as well.
- The quantity $\mathbf{E}^T\left(V(x(T)) \,|\, x(t) \right)$ can be computed if we know the distribution of $x(T)$ given $x(t) = x$. Need the distribution under T-forward measure.

4 Forward measure in rollback (cont)

- Recall that under risk-neutral measure

$$x(T) = x(t)\,e^{-a(T-t)} + I(T) - I(t)\,e^{-a(T-t)} + \sigma \int_t^T e^{-a(T-s)}\,dW_s.$$

Change to T-forward measure. Now $W^T(\cdot)$ is a (driftless) Brownian motion,

$$dW_s = dW_s^T - \sigma b(T-s)\,ds.$$

Substitute

$$\begin{aligned} x(T) &= x(t)\,e^{-a(T-t)} + I(T) - I(t)\,e^{-a(T-t)} \\ &\quad + \sigma \int_t^T e^{-a(T-s)}\left(dW_s^T - \sigma b(T-s)\,ds\right) \\ &= x(t)\,e^{-a(T-t)} \\ &\quad + I(T) - I(t)\,e^{-a(T-t)} - I(T-t) \\ &\quad + \sigma \int_t^T e^{-a(T-s)}\,dW_s^T. \end{aligned}$$

- Denote

$$\begin{aligned} d(t,T) &= I(T) - I(t)\,e^{-a(T-t)} - I(T-t), \\ N(t,T) &= \sigma \int_t^T e^{-a(T-s)}\,dW_s^T. \end{aligned}$$

- Then (under T-forward measure)

$$x(T) = x(t)\,e^{-a(T-t)} + d(t,T) + N(t,T).$$

5 Forward measure in rollback (cont)

- From previous slide

$$x(T) = x(t)\,e^{-a(T-t)} + d(t,T) + N(t,T),$$

  where

  - $d(t,T)$ is a deterministic function;
  - $N(t,T)$ is a Gaussian random variable with **zero mean** and standard deviation

$$\begin{aligned}
\nu(t,T) &\triangleq \sqrt{\operatorname{Var} N(t,T)} \\
&= \sqrt{\operatorname{Var}\left\{\sigma \int_t^T e^{-a(T-s)}\,dW_s^T\right\}} \\
&= \sqrt{\sigma^2 \int_t^T \left(e^{-a(T-s)}\right)^2\,ds} \\
&= \sigma\sqrt{\frac{1-e^{-2a(T-t)}}{2a}}.
\end{aligned}$$

6 Forward measure in rollback (cont)

- Plugging the expression for $x(T)$ in terms of $x(t)$ into forward-measure valuation formula we get,

$$
\begin{aligned}
& \mathbf{E}^T \left(V(x(T)) | \, x(t) = x \right) \\
= & \mathbf{E}^T \left(V \left(x \cdot e^{-a(T-t)} + d(t,T) + N(t,T) \right) \right) \\
= & \int_{-\infty}^{\infty} V \left(x \cdot e^{-a(T-t)} + d(t,T) + \nu(t,T) \cdot z \right) n(z) \, dz,
\end{aligned}
$$

  where $n(z)$ is the **standard** Gaussian density

$$
n(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}.
$$

- So how do we calculate $\mathbf{E}^T \left(V(x(T)) | \, x(t) = x \right)$ as a function of x?

  1. Fix x;
  2. Perform numerical integration of $V \left(x \cdot e^{-a(T-t)} + d(t,T) + \nu(t,T) \cdot z \right) n(z)$ (as a function of z with x fixed);

- Usually we need to know $\mathbf{E}^T \left(V(x(T)) | \, x(t) = x \right)$ **for all** x in some range; computing integrals for each x is very slow.

- We will discuss more efficient methods later.

- For now we just assume that if we have a function $V(x)$, we can "rollback" it to some previous time.

- See Figure 1 again.

7 Forward-measure rollback – summary

- Our main goal so far: express

$$\rho\left\{V,T\to t\right\}\left(x\left(t\right)\right)=P\left(t,T\right)\mathbf{E}^{T}\left(V\left(x\left(T\right)\right)\middle|\,x\left(t\right)\right)$$

as a deterministic function of $x\left(t\right)$. Steps:

1. Express bond as a function of $x\left(t\right)$;

$$P\left(t,T\right)\big|_{x(t)=x}=p\left(x,t,T\right).$$

2. Use some numerical scheme with payoff $V\left(\cdot\right)$ and standard deviation $\nu\left(t,T\right)$ to obtain function $U\left(y\right)$,

$$U\left(y\right)=\int_{-\infty}^{\infty}V\left(y+\nu\left(t,T\right)\cdot z\right)n\left(z\right)\,dz;$$

3. Account for the drift. Since

$$\begin{aligned}&\mathbf{E}^{T}\left(V\left(x\left(T\right)\right)\middle|\,x\left(t\right)=x\right)\\=&\int_{-\infty}^{\infty}V\left(x\cdot e^{-a(T-t)}+d\left(t,T\right)+\nu\left(t,T\right)\cdot z\right)n\left(z\right)\,dz\end{aligned}$$

we have

$$\mathbf{E}^{T}\left(V\left(x\left(T\right)\right)\middle|\,x\left(t\right)=x\right)=U\left(x\cdot e^{-a(T-t)}+d\left(t,T\right)\right);$$

4. Looking ahead, the function $U\left(y\right)$ will be computed (on step 2) for some knots y_n, $n=1,\ldots,N$. To adjust for the drift (Step 3) we have to interpolate U over the knots so we can compute it for all y.

8 Twice-exercisable bond option

- A twice-exercisable Bermuda-style bond option is a right to buy a bond on one of the two dates.
- Consider the dates

$$0 = t_0 < t_1 < t_2 < t_3,$$

  and two strikes, K_1 (for date t_1) and K_2 (for date t_2).

- The bond that the holder has the right to buy is the discount bond $P(\cdot, t_3)$ (paying on t_3).
- A holder of twice-exercisable bond option has the right (not the obligation) to exercise on dates t_1 and t_2,
  - If he exercises on date t_1 he pays the strike K_1 and gets the discount bond;
  - If he exercises on date t_2 he pays the strike K_2 and gets the discount bond;
  - Once exercised, the option goes away.

9 Backward induction for Bermuda-style options

- American and Bermudan-style options are valued by looking at each exercise date "backwards".
- Denote E_i the "exercise" value of the option on date t_i, $i = 1, 2$;
- Denote H_i the "hold" value of the option on date t_i, $i = 1, 2$;
- Denote B_i the Bermuda-option value of the option on date t_i, $i = 0, 1, 2$;
- The holder will exercise on t_i ($i = 1, 2$) if and only if $E_i > H_i$.
- We have

$$
\begin{aligned}
H_2 &= 0, \\
E_2 &= P(t_2, t_3) - K_2, \\
B_2 &= \max\{E_2, H_2\}, \\
H_1 &= \pi_{t_1}(B_2), \\
E_1 &= P(t_1, t_3) - K_1, \\
B_1 &= \max\{E_1, H_1\}, \\
B_0 &= \pi_{t_0}(B_1).
\end{aligned}
$$

10 Backward induction for Bermuda-style options (cont)

- Everything can be expressed in terms of short rate(s). Define deterministic functions

$$e_i(x), \quad h_i(x), \quad b_i(x)$$

by

$$\begin{aligned} E_i &= e_i(x(t_i)), \\ H_i &= h_i(x(t_i)), \\ B_i &= b_i(x(t_i)). \end{aligned}$$

- Rewrite in deterministic-function form

$$\begin{aligned} h_2(x) &= 0, \\ e_2(x) &= p(x, t_2, t_3) - K_2, \\ b_2(x) &= \max\{e_2(x), h_2(x)\}, \\ h_1(x) &= \rho\{b_2, t_2 \to t_1\}(x), \\ e_1(x) &= p(x, t_1, t_3) - K_2, \\ b_1(x) &= \max\{e_1(x), h_1(x)\}, \\ b_0(x) &= \rho\{b_1, t_1 \to t_0\}(x). \end{aligned}$$

- Final value of the Bermuda option (recall that $x(0) = 0$)

$$\text{value}(t = t_0) = b_0(0).$$

- Two "rollbacks" need to be performed:
  - One from t_2 to t_1 with payoff $b_2(x)$;
  - The other from t_1 to t_0 with payoff $b_1(x)$.

11 Feynman-Kac for rollback

- The scheme above works well except actual mechanics of a "rollback" are not clear.
- We need to be able to compute

$$g(y) = \int_{-\infty}^{\infty} G(y + cz)\, n(z)\, dz \qquad (2)$$

where $n(z)$ is a standard Gaussian density. This has to be done for all y (or at least all y in some range)!

- Direct integration is highly impractical.
- Recall your Stochastic Calculus (1999) lectures, Lecture 5 Theorem 2.

Theorem 11.1 (Feynman-Kac special case) *The function $g(y)$ as defined in (2) is equal to*

$$g(y) = f(y, 0)$$

*where function $f(y,t)$ satisfies the **heat equation***

$$\frac{\partial f(y,t)}{\partial t} = -\frac{c^2}{2}\frac{\partial^2 f(y,t)}{\partial y^2}$$

for

$$(y,t) \in \mathbb{R} \times [0,1]$$

with terminal condition

$$f(y,1) = G(y).$$

- This is a very standard problem from numerical analysis and there are tons of methods for solving this particular equation.

12 Integrating heat equation

- We need to solve the heat equation is presented in Theorem: find $f(y,t)$ for
$$(y,t) \in \mathbb{R} \times [0,1]$$
such that inside the domain
$$\frac{\partial f(y,t)}{\partial t} = -\frac{c^2}{2}\frac{\partial^2 f(y,t)}{\partial y^2} \tag{3}$$
and for $t = 1$
$$f(y,1) = G(y). \tag{4}$$

- In our equation $y \in (-\infty, \infty)$. Numerical methods work best in bounded domain. Choose some
$$-\infty < y_{\min} < 0 < y_{\max} < \infty$$
such that
$$|y_{\min}|, |y_{\max}| \text{ large} \tag{5}$$
(will choose later). Solve (3) in the domain
$$(y,t) \in [y_{\min}, y_{\max}] \times [0,1].$$
However, need boundary conditions for $y = y_{\min}$ and $y = y_{\max}$.

- Boundary conditions don't really matter as long as (5) is satisfied. Use constant boundary conditions for example:
$$\begin{aligned} f(y_{\min}, t) &= G(y_{\min}), \quad t \in [0,1], \\ f(y_{\max}, t) &= G(y_{\max}), \quad t \in [0,1]. \end{aligned} \tag{6}$$

- Goal: solve numerically (3), (4) and (6).

13 Finite-difference scheme

- Recall the problem from previous slide: solve

$$\begin{aligned}
\frac{\partial f(y,t)}{\partial t} &= -\frac{c^2}{2}\frac{\partial^2 f(y,t)}{\partial y^2}, \\
f(y,1) &= G(y), \quad y \in [y_{\min}, y_{\max}], \\
f(y_{\min}, t) &= G(y_{\min}), \quad t \in [0,1], \\
f(y_{\max}, t) &= G(y_{\max}), \quad t \in [0,1].
\end{aligned}$$

- Discretize the domain. Choose integers N, M and define

$$\begin{aligned}
h &= (y_{\max} - y_{\min})/N, \\
y_n &= y_{\min} + h \cdot n, \quad n = 0, \ldots, N,
\end{aligned}$$

and

$$\begin{aligned}
k &= 1/M, \\
t_m &= k \cdot m, \quad m = 0, \ldots, M,
\end{aligned}$$

and

$$f_{nm} = f(y_n, t_m).$$

- Lattice terminal conditions:

$$f_{nM} = G(y_n), \quad n = 0, \ldots, N.$$

- Lattice boundary conditions:

$$f_{0m} = G(y_0), \quad f_{Nm} = G(y_N), \quad m = 0, \ldots, M.$$

- See Figure 2.

14 Finite-difference scheme (cont)

- Lattice equation at point $y = y_n, t = t_m$,
  - Time derivative

$$\frac{\partial f(y,t)}{\partial t} = \frac{f(y_n, t_{m+1}) - f(y_n, t_m)}{k} = \frac{f_{n,m+1} - f_{nm}}{k}.$$

  - Space derivative. Fix α,

$$\begin{aligned} 0 &\le \alpha \le 1, \\ \beta &= 1 - \alpha. \end{aligned}$$

The space derivative is a mixture of second-order discretizations for the derivative on time slices t_m and t_{m+1}:

$$\frac{\partial^2 f(y,t)}{\partial y^2} = \alpha \frac{f_{n+1,m} - 2f_{nm} + f_{n-1,m}}{h^2} + \beta \frac{f_{n+1,m+1} - 2f_{n,m+1} + f_{n-1,m+1}}{h^2}.$$

- Equation on time slice t_m,

$$\frac{f_{n,m+1} - f_{nm}}{k} = -\frac{c^2}{2}\left(\alpha \frac{f_{n+1,m} - 2f_{nm} + f_{n-1,m}}{h^2} + \beta \frac{f_{n+1,m+1} - 2f_{n,m+1} + f_{n-1,m+1}}{h^2}\right).$$

- See Figure 3.

15 Finite-difference scheme (cont)

- Introduce dimensionless parameter τ,

$$\tau = \frac{k}{h^2}\frac{c^2}{2}.$$

Then

$$\begin{aligned} f_{n,m+1} - f_{nm} = & -\tau\alpha\left(f_{n+1,m} - 2f_{nm} + f_{n-1,m}\right) \\ & -\tau\beta\left(f_{n+1,m+1} - 2f_{n,m+1} + f_{n-1,m+1}\right), \end{aligned}$$

so

$$\begin{aligned} & \left(-\tau\alpha f_{n+1,m} + (1+2\tau\alpha)\, f_{nm} - \tau\alpha f_{n-1,m}\right) \\ = & \ \tau\beta f_{n+1,m+1} + (1-2\tau\beta)\, f_{n,m+1} + \tau\beta f_{n-1,m+1}. \end{aligned} \qquad (7)$$

- The equation connects the value of the unknown function on time slices t_{m+1} and t_m. We know the value of the function for time slice $t_M = 1.0$. In backward induction, we use the equation (7) to solve for the value of f on time slice t_{M-1}, then t_{M-2} and so on until $t_0 = 0$.

- See Figure 2 again.

16 Explicit and Crank-Nicholson

- Main equation to solve

$$\begin{aligned} &\left(-\tau\alpha f_{n+1,m} + (1+2\tau\alpha)\, f_{nm} - \tau\alpha f_{n-1,m}\right) \\ &= \tau\beta f_{n+1,m+1} + (1-2\tau\beta)\, f_{n,m+1} + \tau\beta f_{n-1,m+1}. \end{aligned}$$

- Simplest case (explicit discretization) is when $\alpha = 0$, $\beta = 1$:
  - Equation

$$f_{nm} = \tau f_{n+1,m+1} + (1-2\tau)\, f_{n,m+1} + \tau f_{n-1,m+1};$$

  - Same as trinomial tree: $\{\tau, 1-2\tau, \tau\}$ as rollback probabilities;
  - Simple but only conditionally stable; if "probabilities" are negative, the method blows up.

- Best numerical properties is with Crank-Nicholson scheme, $\alpha = \beta = 0.5$:
  - Unconditionally stable for all discretization steps h and k;
  - Second-order accurate in time.

17 Matrix formalism

- Implicit schemes (including C-N) requires solving a tridiagonal linear system for each time slice t_m.
- Define time-slice vectors (column vectors)

$$F_m = [f_{0m}, \dots, f_{nm,} \dots, f_{Nm}]^T .$$

- Define tridiagonal matrix

$$R_\xi = \begin{pmatrix} -2\tau\xi & \tau\xi & 0 & \dots & 0 \\ \tau\xi & -2\tau\xi & \tau\xi & \dots & \dots \\ 0 & \dots & \dots & \dots & 0 \\ \dots & \dots & \tau\xi & -2\tau\xi & \tau\xi \\ 0 & \dots & 0 & \tau\xi & -2\tau\xi \end{pmatrix} .$$

- Let E be a unit matrix of the same dimensions.
- The time-stepping equation becomes (only for zero boundary conditions!)

$$(E - R_\alpha)\, F_m = (E + R_\beta)\, F_{m+1},$$

  here F_m is unknown and F_{m+1} is known;
- For non-zero boundary conditions the matrix has to be modified slightly (how? consult a textbook!)
- Tridiagonal linear systems can be solved very fast!
- If using Matlab, use sparse data structures (`help sparfun`).

18 Choice of a domain

- Unanswered question: how to choose $y_{\min}$ and $y_{\max}$?
- General idea: choose a large enough interval so that the short rate $x(\cdot)$ has small probability of ever going outside it.
- For each t, $x(t)$ is Gaussian with drift $I(t)$ and standard deviation

$$\nu(0,t) = \sigma \left(\frac{1 - e^{-2at}}{2a} \right)^{1/2};$$

- Let T be the last "interesting" time; for option on a bond that is the maturity of the bond;
- Fix small probability γ (e.g. $\gamma = 0.05\%$);
- Find a standard normal quantile,

$$\mathbf{Q}\left(\left| \frac{x(T) - I(T)}{\nu(0,T)} \right| > z_\gamma \right) = \gamma;$$

- Rewrite for $x(T)$,

$$\mathbf{Q}\left(I(T) - z_\gamma \nu(0,T) < x(T) < I(T) + z_\gamma \nu(0,T) \right) = 1 - 2\gamma;$$

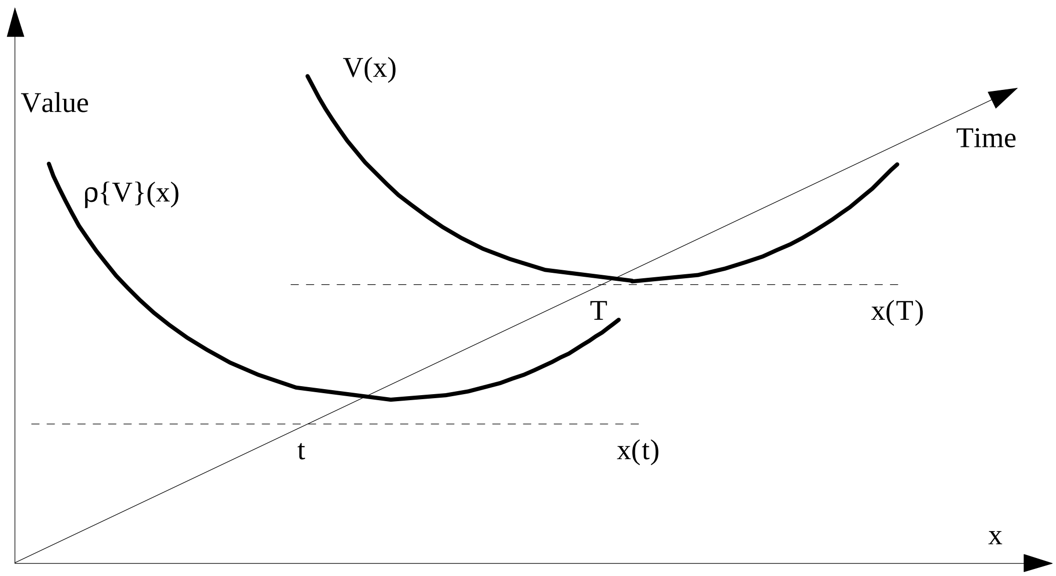
- Set

$$\begin{aligned} y_{\min} &= I(T) - z_\gamma \nu(0,T), \\ y_{\max} &= I(T) + z_\gamma \nu(0,T). \end{aligned}$$

19 PDE vs tree

- Both solve the same problem;
- PDE solver is harder to implement, but is well worth the extra trouble.
- Massive amount of research on numerical methods for PDE's is available.
- In lattice, we achieve decoupling of business logic (rules for exercise, etc.) and numerical methods (solving partial differential equations).
- Stability properties much better for lattice
  - Trees may blow up if large time steps are used; need to be careful;
  - Crank-Nicholson is unconditionally stable.
- In multi-dimensional models, trees have no chance against lattices.
- In conclusion, trees are good for
  - Initial evaluation of the model;
  - Rapid prototyping;
  - Modelling one-off instrument.
- Lattices are good for
  - Models in production;
  - Hi-volume runs where speed, accuracy and robustness are important.

Figure 1. Rollback



.

Figure 2. Finite-difference discretization

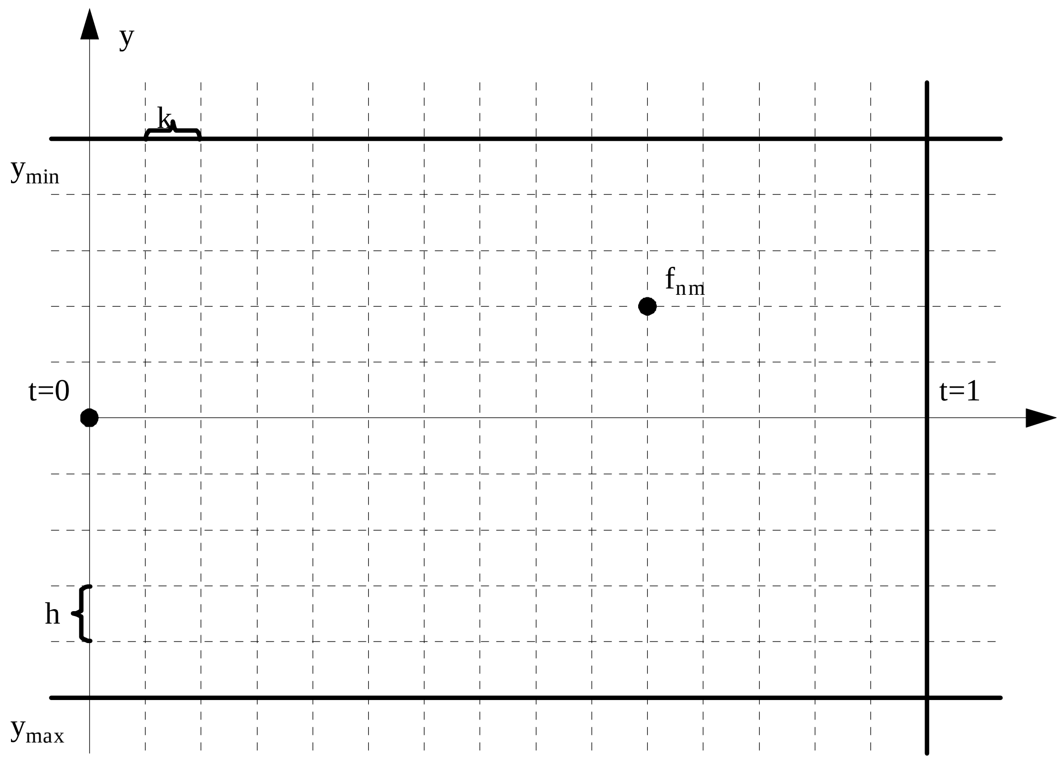


Figure 3. Time stepping in a lattice