The semigroup approach to stochastic PDEs and their finite element approximation

Mihály Kovács

Department of Mathematics and Statistics
University of Otago
Dunedin, New Zealand

Bielefeld IGK, November 2009

Outline

- Gaussian measures in Hilbert spaces
- Q-Wiener process
- Wiener-integral
- Semigroup framework for linear SPDEs with additive noise
  - Weak solution
- Semigroup framework for semilinear SPDEs with additive noise
  - Mild solution
- Examples: heat and wave equations
- Spatial approximation
- Strong error
  - Heat equation
  - Wave equation
- Weak error
  - Representation of the weak error
  - Heat equation
  - Wave equation

Gaussian measures in Hilbert spaces

- $(\Omega, \mathcal{F}, P)$ probablility space
- U separable Hilbert space
- $\mathcal{B}(U)$ the Borel σ-algebra of U

DEFINITION. A random real random variable is measurable function $X : (U, \mathcal{B}(U)) \to (\mathbf{R}, \mathcal{B}(\mathbf{R}))$, where $\mathcal{B}(\mathbf{R})$ is the real Borel σ-algebra. The law of X is the probability measure $\mu \circ X^{-1}$.

For $v \in U$ let $v' \in U^*$ denote the functional given by $v'(u) = \langle v, u \rangle_U, \; u \in U$.

DEFINITION. A probability measure μ on $(U, \mathcal{B}(U))$ is *Gaussian* if for all $v \in U$, v' has a Gaussian law as a real-valued random variable on the probability space $(U, \mathcal{B}(U), \mu)$. That is, for all $v \in U$ there are $m_v \in \mathbf{R}$ and $\sigma_v \in \mathbf{R}_+$, such that, if $\sigma_v > 0$,

$$\left(\mu \circ (v')^{-1}\right)(A) = \mu\left(\{u \in U : v'(u) \in A\}\right) = \frac{1}{\sqrt{2\pi\sigma_v^2}} \int_A \mathrm{e}^{-\frac{(s-m_v)^2}{2\sigma_v^2}} \, \mathrm{d}s,$$

for all $A \in \mathcal{B}(\mathbf{R})$. If $\sigma_v = 0$, then we require that $\mu \circ (v')^{-1} = \delta_{m_v}$, the Dirac measure concentrated at m_v.

Nuclear operators and trace

Let $\mathcal{L}_1(U)$ denote the set of *nuclear operators* from U to U; that is, $T \in \mathcal{L}_1(U)$ if $T \in L(U)$ (bounded, linear on U) and there are sequences $\{a_j\}, \{b_j\} \subset U$ with $\sum_{j=1}^{\infty} \|a_j\| \|b_j\| < \infty$ and such that

$$Tx = \sum_{j=1}^{\infty} \langle x, b_j \rangle a_j \quad \forall x \in U.$$

It is a Banach space under the norm

$$\|T\|_{\mathrm{Tr}} = \inf \Big\{ \sum_{j=1}^{\infty} \|a_j\| \|b_j\| : Tx = \sum_{j=1}^{\infty} \langle x, b_j \rangle a_j \Big\}.$$

For $T \in \mathcal{L}_1(U)$ the trace of T, $\mathrm{Tr}(T)$, is well defined and is given by

$$\mathrm{Tr}(T) = \sum_{k=1}^{\infty} \langle Te_k, e_k \rangle$$

with $\{e_k\}_{k=1}^{\infty}$ an ONB of U.

Characterization of Gaussian measures

THEOREM. A finite measure μ on $(U, \mathcal{B}(U))$ is Gaussian if and only if

$$\hat{\mu}(u) := \int_U e^{i\langle u,v\rangle_U} \, d\mu(v) = e^{i\langle m,u\rangle_U - \frac{1}{2}\langle Qu,u\rangle_U},$$

where $m \in U$ and $Q \in L(U), Q \geq 0$, with $\mathrm{Tr}(Q) < \infty$. In this case we write $\mu = N(m, Q)$, and m and Q are called the mean and the covariance operator of μ. The measure μ is uniquely determined by m and Q.

COROLLARY. Let μ be a Gaussian measure on U with mean m and covariance operator Q. Then, for all $u, v \in U$,

$$\int_U \langle x, u\rangle_U \, d\mu(x) = \langle m, u\rangle_U,$$

$$\int_U \langle x - m, u\rangle_U \langle x - m, v\rangle_U \, d\mu(x) = \langle Qu, v\rangle_U,$$

$$\int_U \|x - m\|_U^2 \, d\mu(x) = \mathrm{Tr}(Q).$$

Gaussian random variables and the existence of Gaussian measures

DEFINITION. A U-valued random variable X on a probability space $(\Omega, \mathcal{F}, P)$, that is, a measurable mapping $X : (\Omega, \mathcal{F}, P) \to (U, \mathcal{B}(U))$, is Gaussian if the law $\mu = P \circ X^{-1}$ of X is a Gaussian measure on $(U, \mathcal{B}(U))$, that is, $P \circ X^{-1} = N(m, Q)$ for some $m \in U$ and $Q \in L(U)$. We call m the mean and Q the covariance operator of X.

PROPOSITION. If X is a U-valued Gaussian random variable with mean m and covariance operator Q, then for all $u, v \in U$,

$$\begin{aligned}
&\mathbf{E}(\langle X, u\rangle_U) = \langle m, u\rangle_U,\\
&\mathbf{E}(\langle X - m, u\rangle_U \langle X - m, v\rangle_U) = \langle Qu, v\rangle_U,\\
&\mathbf{E}(\|X - m\|_U^2) = \mathrm{Tr}(Q).
\end{aligned}$$

PROPOSITION. If $Q \in L(U)$, $Q \geq 0$, and $\mathrm{Tr}(Q) < \infty$, then there is an orthonormal basis $\{e_k\}_{k\in\mathbf{N}}$ of U such that $Qe_k = \lambda_k e_k$, $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_k \geq \lambda_{k+1} \geq \cdots \geq 0$, $\lambda_k \to 0$ as $k \to \infty$, and 0 is the only accumulation point of $\{\lambda_k\}_{k\in\mathbf{N}}$. Moreover,

$$Qx = \sum_{k=1}^{\infty} \lambda_k \langle x, e_k\rangle_U e_k, \quad x \in U, \text{ and } \mathrm{Tr}(Q) = \sum_{k=1}^{\infty} \lambda_k.$$

THEOREM. Let $m \in U$ and $Q \in L(U)$, $Q \geq 0$, with $\mathrm{Tr}(Q) < \infty$. A U-valued random variable X on $(\Omega, \mathcal{F}, P)$ is Gaussian with $P \circ X^{-1} = N(m, Q)$ if and only if

$$X = m + \sum_{k=1}^{\infty} \sqrt{\lambda_k} \beta_k e_k,$$

where (λ_k, e_k) are the eigenpairs of Q and β_k are independent real random variables with $P \circ \beta_k^{-1} = N(0,1)$ if $\lambda_k > 0$ and $\beta_k = 0$ otherwise. The series converges in $L_2(\Omega, \mathcal{F}, P; U)$.

COROLLARY.(**Existence of Gaussian measures.**) For each $m \in U$ and $Q \in L(U)$, $Q \geq 0$, with $\mathrm{Tr}(Q) < \infty$, there exists $\mu = N(m, Q)$.

Nuclear Q-Wiener processes

DEFINITION. A U-valued stochastic process $\{W(t)\}_{t\geq 0}$ is called a (nuclear) Q-Wiener process if

1. $W(0) = 0$;
2. $\{W(t)\}_{t\geq 0}$ has continuous paths almost surely, that is, the mapping $t \mapsto W(t,\omega)$ is continuous for almost every $\omega \in \Omega$;
3. $\{W(t)\}_{t\geq 0}$ has independent increments, that is, for any finite partition $0 = t_0 \leq t_1 \leq \cdots \leq t_{m-1} \leq t_m < \infty$ the random variables $W(t_1), W(t_2) - W(t_1), \cdots, W(t_m) - W(t_{m-1})$, are independent;
4. the increments have Gaussian laws, more precisely,

$$P \circ (W(t) - W(s))^{-1} = N(0, (t-s)Q), \quad 0 \leq s \leq t.$$

Existence and representation of nuclear Q-Wiener processes

THEOREM. Let $Q \in L(U)$, $Q \geq 0$, with $\mathrm{Tr}(Q) < \infty$. A U-valued process $\{W(t)\}_{t\geq 0}$ is a U-valued Q-Wiener process if and only if

$$W(t) = \sum_{k=1}^{\infty} \sqrt{\lambda_k}\beta_k(t)e_k,$$

where (λ_k, e_k) are the eigenpairs of Q and $\{\beta_k(t)\}_{t\geq 0}$ are independent real-valued standard Brownian motions on $(\Omega, \mathcal{F}, P)$. For each $T > 0$, the series converges in $L_2\big(\Omega, \mathcal{F}, P; C([0, T], U)\big)$. In particular, for every $Q \in L(U)$ with $Q \geq 0$ and $\mathrm{Tr}(Q) < \infty$, there exists a Q-Wiener process.

Cylindrical processes

In practice one may would like to consider a Wiener process with more general covariance operator Q, such as $Q = I$. If $\mathrm{Tr}(Q = \infty)$, then the sum

$$W(t) = \sum_{k=1}^{\infty} \sqrt{\lambda_k}\beta_k(t)e_k$$

does not even converge in $L_2(\Omega, U)$, since

$$\mathbf{E}\Big\| \sum_{j=1}^{\infty} \lambda_j^{1/2}\beta_j(t)e_j \Big\|^2 = \sum_{j=1}^{\infty} \lambda_j \mathbf{E}\big(\beta_j(t)^2\big) = t\sum_{j=1}^{\infty} \lambda_j = t\,\mathrm{Tr}(Q) = \infty.$$

In this case we call the formal sum a cylindrical Q-Wiener process. The important point is that we can still define an integral w.r.t. cylindrical processes.

REMARK.

- The sum converges in a suitable larger Hilbert space where one obtains a nuclear Wiener process. However, this is not unique.
- The real processes $W_x(t) = \sum_{k=1}^{\infty} \sqrt{\lambda_k}\beta_k(t)\langle e_k, x\rangle_U$ are well-defined real valued Brownian motions with covariance $\mathbf{E}(W_x(t)^2) = t\|Q^{1/2}x\|^2$.

The Wiener(-Itô) integral

For stochastic equations with additive noise the complete theory of the Itô integral is not needed since the integrand is deterministic.

Let U, H be separable Hilbert spaces and let $F : [0, \infty) \to L(U, H)$ ($L(U, H)$ is the space of bounded linear operators from U to H) be strongly continuous; that is, $t \mapsto F(t)x$ is continuous for each $x \in U$. Let

$$W(t) = \sum_{k=1}^{\infty} \sqrt{\lambda_k}\beta_k(t)e_k; \;\; Qe_k = \lambda_k e_k, \tag{1}$$

where the sum is formal if $\mathrm{Tr}(Q) = \infty$. Define, first formally,

$$\int_0^t F(s)\,\mathrm{d}W(s) := \sum_{k=1}^{\infty} \sqrt{\lambda_k} \int_0^t F(s)e_k\,\mathrm{d}\beta_k(s). \tag{2}$$

Each term in the expansion is defined in terms of real-valued Itô integrals as

$$\int_0^t F(s)e_k\,\mathrm{d}\beta_k(s) = \sum_{j=1}^{\infty} \int_0^t \langle F(s)e_k, \phi_j\rangle\,\mathrm{d}\beta_k(s)\,\phi_j,$$

where $\{\phi_j\}_{j=1}^{\infty}$ is an orthonormal basis for H.

The latter series converges in $L_2(\Omega, H)$, because, by the isometry of the real-valued Itô integral; that is,

$$\mathbf{E}\Big(\Big|\int_0^t f(s)\,\mathrm{d}\beta_k(s)\Big|^2\Big) = \int_0^t |f(s)|^2\,\mathrm{d}s, \tag{3}$$

and Parseval's identity, we have for fixed $t > 0$,

$$\begin{aligned}
\mathbf{E}\Big(\Big\|\int_0^t F(s)e_k\,\mathrm{d}\beta_k(s)\Big\|^2\Big) &= \mathbf{E}\Big(\sum_{j=1}^{\infty}\Big|\int_0^t \langle F(s)e_k, \phi_j\rangle\,\mathrm{d}\beta_k(s)\Big|^2\Big) \\
&= \sum_{j=1}^{\infty}\int_0^t |\langle F(s)e_k, \phi_j\rangle|^2\,\mathrm{d}s \\
&= \int_0^t \|F(s)e_k\|^2\,\mathrm{d}s < \infty,
\end{aligned}$$

since F is strongly continuous.

The Itô isometry

THEOREM. Let $F : [0, \infty) \to L(U, H)$ be strongly continuous and let $\{W(t)\}_{t \geq 0}$ be a Q-Wiener process given by (1). Assume that the operator $Q_F(t) \in L(H)$, which is defined by

$$Q_F(t)x = \int_0^t F(s) Q F^*(s) x \, \mathrm{d}s, \quad x \in H,$$

has finite trace for all $t \geq 0$. Then the series in (2) converges in $L_2(\Omega, H)$ and defines an H-valued Gaussian random variable $\int_0^t F(s) \, \mathrm{d}W(s)$ with zero mean and covariance operator $Q_F(t)$. Moreover, we have the isometry

$$\mathbf{E}\Big(\Big\| \int_0^t F(s) \, \mathrm{d}W(s) \Big\|^2\Big) = \mathrm{Tr}(Q_F(t)) = \int_0^t \|F(s) Q^{1/2}\|_{\mathrm{HS}}^2 \, \mathrm{d}s. \qquad (4)$$

Hilbert-Schmidt operator $B : U \to H$:

$$\|B\|_{\mathrm{HS}}^2 = \sum_{l=1}^{\infty} \|B\varphi_l\|_H^2 < \infty, \qquad \{\varphi_l\} \text{ arbitrary ON basis in } U.$$

Note: $\|B\|_{\mathrm{HS}}^2 = \mathrm{Tr}(B^* B) = \mathrm{Tr}(B B^*) = \|B^*\|_{\mathrm{HS}}^2$.

PROOF. Since F is strongly continuous it follows that $Q_F(t)$ is well defined as a Bochner integral. Its trace is

$$\begin{aligned}\mathrm{Tr}(Q_F(t)) &= \sum_{j=1}^{\infty} \int_0^t \langle F(s)QF^*(s)\phi_j, \phi_j\rangle \,\mathrm{d}s = \int_0^t \mathrm{Tr}(F(s)QF^*(s))\,\mathrm{d}s \\ &= \int_0^t \mathrm{Tr}(F(s)Q^{1/2}(F(s)Q^{1/2})^*)\,\mathrm{d}s = \int_0^t \|F(s)Q^{1/2}\|_{\mathrm{HS}}^2\,\mathrm{d}s,\end{aligned}$$

where $\{\phi_j\}_{j=1}^{\infty}$ is an orthonormal basis of H. This is the last equality in (4).

Next we show that the series in (2) converges in $L_2(\Omega, H)$.

$$\mathbf{E}\Big(\Big\|\int_0^t F(s)\,\mathrm{d}W(s)\Big\|^2\Big) = \mathbf{E}\Big(\Big\|\sum_{k=1}^{\infty}\lambda_k^{\frac12}\int_0^t F(s)e_k\,\mathrm{d}\beta_k(s)\Big\|^2\Big)$$

$$= \mathbf{E}\Big(\sum_{j=1}^{\infty}\Big|\sum_{k=1}^{\infty}\lambda_k^{\frac12}\int_0^t \langle F(s)e_k,\phi_j\rangle\,\mathrm{d}\beta_k(s)\Big|^2\Big)$$

$$= \sum_{j=1}^{\infty}\mathbf{E}\Big(\Big|\sum_{k=1}^{\infty}\lambda_k^{\frac12}\int_0^t \langle F(s)e_k,\phi_j\rangle\,\mathrm{d}\beta_k(s)\Big|^2\Big)$$

$$= \sum_{j=1}^{\infty}\sum_{k,l=1}^{\infty}\lambda_k^{\frac12}\lambda_j^{\frac12}\,\mathbf{E}\Big(\int_0^t \langle F(s)e_k,\phi_j\rangle\,\mathrm{d}\beta_k(s)\int_0^t \langle F(s)e_l,\phi_j\rangle\,\mathrm{d}\beta_l(s)\Big)$$

{Independence of β_k, β_l and real Itô isometry}

$$= \sum_{j=1}^{\infty}\sum_{k=1}^{\infty}\lambda_k\int_0^t |\langle F(s)e_k,\phi_j\rangle|^2\,\mathrm{d}s = \sum_{j=1}^{\infty}\sum_{k=1}^{\infty}\int_0^t |\langle F(s)Q^{\frac12}e_k,\phi_j\rangle|^2\,\mathrm{d}s$$

$$= \int_0^t\sum_{k=1}^{\infty}\sum_{j=1}^{\infty}|\langle F(s)Q^{\frac12}e_k,\phi_j\rangle|^2\,\mathrm{d}s = \int_0^t\sum_{k=1}^{\infty}\|F(s)Q^{\frac12}e_k\|^2\,\mathrm{d}s = \int_0^t \|F(s)Q^{\frac12}\|_{\mathrm{HS}}^2\mathrm{d}s.$$

Thus the series in (2) converges in $L_2(\Omega, H)$ to a random variable, which is Gaussian, because it is the limit of Gaussian random variables. Finally, a similar calculation shows that

$$\mathbf{E}\Big(\Big\langle \int_0^t F(s)\,\mathrm{d}W(s), x\Big\rangle\Big\langle \int_0^t F(s)\,\mathrm{d}W(s), y\Big\rangle\Big) = \langle Q_F(t)x, y\rangle, \quad x, y \in H,$$

so that the covariance operator of $\int_0^t F(s)\,\mathrm{d}W(s)$ is indeed $Q_F(t)$. $\square$

REMARK. The Fourier expansion (1) of $W(t)$ might not converge in $L_2(\Omega, U)$ ($\mathrm{Tr}(Q) = \infty$) but the expansion (2) of the stochastic integral does converge in $L_2(\Omega, H)$ and gives an H-valued random variable provided that $\mathrm{Tr}(Q_F(t)) < \infty$.

Stochastic convolution

An important special case, as we will see, is the stochastic convolution; that is, when

$$F(s) = \mathrm{e}^{-(t-s)A}B.$$

COROLLARY. Let $-A$ generate a C_0-semigroup e^{-tA} on H, let $B \in L(U,H)$, and let $W(t)$ be a Q-Wiener process in U. Assume that the operator $Q_A(t)$, which is defined by

$$Q_A(t)x = \int_0^t \mathrm{e}^{-sA} B Q B^* \mathrm{e}^{-sA^*} x \, \mathrm{d}s, \tag{5}$$

has finite trace for all $t \geq 0$. Then the stochastic convolution

$$W_A(t) = \int_0^t \mathrm{e}^{-(t-s)A} B \, \mathrm{d}W(s) := \sum_{k=1}^{\infty} \sqrt{\lambda_k} \int_0^t \mathrm{e}^{-(t-s)A} B e_k \, \mathrm{d}\beta_k(s)$$

is a well-defined Gaussian random variable with zero mean and covariance operator $Q_A(t)$. Moreover, we have the isometry

$$\mathbf{E}\Big(\|W_A(t)\|^2\Big) = \mathrm{Tr}(Q_A(t)) = \int_0^t \|\mathrm{e}^{-sA} B Q^{1/2}\|_{\mathrm{HS}}^2 \, \mathrm{d}s.$$

Semigroup approach to SPDEs: the linear case

Linear SPDEs with additive noise:

$$\begin{cases} \mathrm{d}X(t) + AX(t)\,\mathrm{d}t = B\,\mathrm{d}W(t), \ t > 0 \\ X(0) = X_0 \end{cases}$$

- H, U Hilbert spaces
- $W(t) = \sum_{k=1}^{\infty} \sqrt{\lambda_k}\beta_k(t)e_k$, Q-Wiener process on U, $Qe_k = \lambda_k e_k$
- Filtration $\mathcal{F}_s := \bigcap_{r>s} \tilde{\mathcal{F}}_r^0$, where

  $$\mathcal{N} := \{C \in \mathcal{F} : P(C) = 0\}, \ \tilde{\mathcal{F}}_s := \sigma(\beta_k(r) : r \leq s, k \in \mathbf{N}), \ \tilde{\mathcal{F}}_s^0 := \sigma(\mathcal{N} \cup \tilde{\mathcal{F}}_s)$$

- $B \in L(U, H)$
- $\{X(t)\}_{t\geq 0}$, H-valued stochastic process
- X_0 is $\mathcal{F}_0$-measurable

- $-A : \mathcal{D}(A) \subset H \to H$ is linear operator, generating a strongly continuous semigroup (C_0-semigroup) of bounded linear operators $\{S(t)\}_{t\geq 0} \subset L(H)$; that is,
  - $S(0) = I$;
  - $S(t+s) = S(t)S(s)$ for all $s, t \geq 0$;
  - $\{S(t)\}_{t\geq 0}$ is strongly continuous on $[0, \infty)$, that is, $t \mapsto S(t)x$ is continuous on $[0, \infty)$ for all $x \in H$;
  - $\lim_{h\to 0^+} \frac{S(t+h)x - S(t)x}{h} = -Ax$ for all $x \in \mathcal{D}(A)$;

In this case $u(t) = S(t)x$ is the unique (mild) solution of the deterministic equation

$$u(t) + A\int_0^t u(s)ds = x,\ x \in H,\ t \geq 0,$$

and if $x \in \mathcal{D}(A)$, then u is the unique (strong) solution of

$$\dot{u}(t) + Au(t) = 0,\ t > 0;\ u(0) = x.$$

Sometimes the semigroup $S(t)$ generated by A is also written as $S(t) = e^{-tA}$ in analogy with matrix exponentials. In several cases this can be made rigorous using a functional calculus.

Weak solution

DEFINITION. (**Weak Solution.**) An H-valued process $\{X(t)\}_{t\in[0,T]}$ is a weak solution of the linear SPDE if $X(t)$ is $\mathcal{F}_t$-measurable ($t \in [0,T]$), $\{X(t)\}_{t\in[0,T]}$ has Bochner integrable trajectories P-almost surely and

$$\langle X(t), \eta\rangle + \int_0^t \langle X(s), A^*\eta\rangle \, \mathrm{d}s = \langle \xi, \eta\rangle + W_{B^*\eta}(t)$$

$$P\text{-a.s.}, \ \forall \eta \in \mathcal{D}(A), \ t \in [0,T].$$

Recall that

$$W_{B^*\eta}(t) = \sum_{k=1}^{\infty} \sqrt{\lambda_k}\beta_k(t)\langle e_k, B^*\eta\rangle_U = \sum_{k=1}^{\infty} \sqrt{\lambda_k}\beta_k(t)\langle Be_k, \eta\rangle_H .$$

Note, that $W_{B^*\eta}(t) = \int_0^t l_\eta B \, \mathrm{d}W(s)$, where

$$l_\eta : H \to \mathbf{R}, \quad l_\eta(h) := \langle h, \eta\rangle, \ h \in H.$$

The obvious candidate for the solution is given by the variation of constants formula

$$X(t) = S(t)\xi + \int_0^t S(t-s)B \, \mathrm{d}W(s).$$

THEOREM. (**Existence and uniqueness of the weak solution.**) If

$$\int_0^T \|S(r)BQ^{1/2}\|_{\mathrm{HS}}^2 \, \mathrm{d}r < \infty,$$

then

$$X(t) = S(t)\xi + \int_0^t S(t-s)B \, \mathrm{d}W(s)$$

is a weak solution of the linear SPDE and it is unique up to modification. That is, if $Y(t)$ is another weak solution then $X(t) = Y(t)$, P-a.s.

REMARK. The concept of weak solution is necessary for two reasons.

- The relation $X(t) \in \mathcal{D}(A)$ is seldom true
- For the integral $\int_0^t B \, \mathrm{d}W(t)$ to exist one needs $\|BQ^{1/2}\|_{\mathrm{HS}}^2 < \infty$.

Semigroup approach to SPDEs: the semilinear case

Here we consider equations written formally as

$$\begin{aligned} &\mathrm{d}X(t) + AX(t)\,\mathrm{d}t = f(X(t))\,\mathrm{d}t + B\,\mathrm{d}W(t), \quad 0 < t < T, \\ &X(0) = \xi. \end{aligned} \tag{6}$$

- H, U separable Hilbert spaces
- $W(t) = \sum_{k=1}^{\infty} \sqrt{\lambda_k}\beta_k(t)e_k$, Q-Wiener process on U, $Qe_k = \lambda_k e_k$
- Filtration $\mathcal{F}_s := \bigcap_{r>s} \tilde{\mathcal{F}}_r^0$, where

$$\mathcal{N} := \{C \in \mathcal{F} : P(C) = 0\},\ \tilde{\mathcal{F}}_s := \sigma(\beta_k(r) : r \leq s, k \in \mathbf{N}),\ \tilde{\mathcal{F}}_s^0 := \sigma(\mathcal{N} \cup \tilde{\mathcal{F}}_s)$$

- $-A$ generates a C_0-semigroup $\{S(t)\}_{t\geq 0}$
- $B \in L(U, H)$
- $f : H \to H$
- $\{X(t)\}_{t\geq 0}$, H-valued stochastic process
- X_0 is $\mathcal{F}_0$-measurable

The main difference when dealing with this kind of equations compared to the one before is that, in general, there is no explicit representation of the solution of (6). Another solution concept is more convenient in this case.

Mild solution

An H-valued process $\{X(t)\}_{t\in[0,T]}$ is a mild solution of (6) if $X(t)$ is $\mathcal{F}_t$-measurable ($t \in [0,T]$),

$$X \in C\big([0,T]; L_2(\Omega, \mathcal{F}, P; H)\big)$$

and, for all $t \in [0,T]$,

$$X(t) = S(t)\xi + \int_0^t S(t-s)f(X(s))\,\mathrm{d}s + \int_0^t S(t-s)B\,\mathrm{d}W(s) \quad P\text{-a.s.}$$

THEOREM. (**Existence and uniqueness of the mild solution.**) If $\xi \in L_2(\Omega, \mathcal{F}_0, P; H)$,

$$\int_0^T \|S(s)BQ^{1/2}\|_{\mathrm{HS}}^2\,\mathrm{d}s < \infty$$

and $f : H \to H$ satisfies the global Lipschitz condition

$$\|f(x) - f(y)\|_H \leq K\|x - y\|_H, \quad \forall x, y \in H,$$

for some $K > 0$, then there is a unique mild solution of (6).

PROOF. (**Sketch.**) The proof is a fixed point argument.

First, it is not difficult to show that

$$Z_{[a,b]} := \Big\{ X \in C\big([a,b]; L_2(\Omega, \mathcal{F}, P; H)\big) : X(t) \text{ is } \mathcal{F}_t\text{-measurable } (t \in [a,b]) \Big\}$$

with norm $\|Y\|_{Z_{[a,b]}} = \sup_{t\in[a,b]}(\mathbf{E}\|Y(t)\|_H^2)^{1/2}$ is a Banach space.

Then, define

$$F(Y)(t) := S(t)\xi + \int_0^t S(t-s)f(Y(s))\mathrm{d}s + \int_0^t S(t-s)B\,\mathrm{d}W(s)$$

and show that $F : Z_{[0,\tau]} \to Z_{[0,\tau]}$ for some $\tau > 0$ and that it is a contraction; that is,

$$\|F(Y_1) - F(Y_2)\|_{Z_{[0,\tau]}} \le L\|Y_1 - Y_2\|_{Z_{[0,\tau]}},\ L < 1.$$

This yields a unique fixed point of F and hence a unique mild solution on $[0,\tau]$. Finally, repeat the argument on $[\tau, 2\tau]$, $[2\tau, 3\tau]$ and so on, to get a unique solution on $[0, T]$.

Examples: heat equation

The stochastic heat equation is formally

$$\begin{cases} \frac{\partial u(x,t)}{\partial t} - \Delta u(x,t) = \dot{W}(x,t), & x \in D,\ t > 0 \\ u(x,t) = 0, & x \in \partial D,\ t > 0 \\ u(x,0) = u_0(x), & x \in D \end{cases}$$

where $D \subset \mathbf{R}^d$ is a bounded domain and $\Delta = \sum_{k=1}^d \partial/\partial \xi_k^2$ denotes the Laplace operator. In order to put the equation into the semigroup framework define $H = U = L_2(\mathcal{D})$ and recall the Sobolev spaces

$$H^k = H^k(\mathcal{D}) = \Big\{ v \in L_2(\mathcal{D}) : D^\alpha v \in L_2(\mathcal{D}),\ |\alpha| \le k \Big\},$$
$$H_0^1 = H_0^1(\mathcal{D}) = \Big\{ v \in H^1(\mathcal{D}) : v|_{\partial \mathcal{D}} = 0 \Big\}.$$

We consider $A = -\Delta$ as an unbounded linear operator on H with domain of definition $\mathcal{D}(A) = H^2 \cap H_0^1$.

It is well known that A is self-adjoint positive definite and that the eigenvalue problem

$$A\phi_j = \mu_j \phi_j$$

provides an orthonormal basis $\{\phi_j\}_{j=1}^{\infty}$ for H and an increasing sequence of eigenvalues

$$0 < \mu_1 < \mu_2 \leq \cdots \leq \mu_j \leq \cdots, \quad \mu_j \approx j^{2/d} \to \infty \text{ as } j \to \infty. \tag{7}$$

The operator $-A$ is the infinitesimal generator of the semigroup $S(t) = \mathrm{e}^{-tA} \in L(H)$ defined by

$$S(t)v = \mathrm{e}^{-tA}v = \sum_{j=1}^{\infty} \mathrm{e}^{-t\mu_j} \langle v, \phi_j \rangle \phi_j.$$

The semigroup is analytic and, in particular, by a simple calculation using Parseval's identity we have

$$\int_0^T \|A^{1/2}\mathrm{e}^{-tA}v\|^2 \,\mathrm{d}t = \int_0^T \sum_j \mu_j \mathrm{e}^{-2t\mu_j} \langle v, \phi_j \rangle^2 \,\mathrm{d}t \leq \frac{1}{2}\|v\|^2. \tag{8}$$

The stochastic heat equation can now be written

$$\begin{aligned} &\mathrm{d}X + AX\,\mathrm{d}t = \mathrm{d}W, \quad t > 0, \\ &X(0) = 0, \end{aligned}$$

where, for simplicity, we have set $X_0 = 0$. It is of the form of a linear SPDE with $B = I$, and its unique weak solution is given by the stochastic convolution

$$X(t) = W_A(t) = \int_0^t S(t-s)\,\mathrm{d}W(s)$$

provided that

$$\int_0^T \|S(t)Q^{1/2}\|_{\mathrm{HS}}^2\,\mathrm{d}t < \infty. \tag{9}$$

Taking an orthonormal basis $\{f_k\}$ of H we compute, using (8),

$$\int_0^T \|S(t)Q^{1/2}\|_{\mathrm{HS}}^2\,\mathrm{d}t = \int_0^T \|\mathrm{e}^{-tA}Q^{1/2}\|_{\mathrm{HS}}^2\,\mathrm{d}t = \int_0^T \sum_k \|\mathrm{e}^{-tA}Q^{1/2}f_k\|^2\,\mathrm{d}t$$

$$= \sum_k \int_0^T \|A^{1/2}\mathrm{e}^{-tA}A^{-1/2}Q^{1/2}f_k\|^2 \mathrm{d}t \le \frac{1}{2}\sum_k \|A^{-1/2}Q^{1/2}f_k\|^2 = \frac{1}{2}\|A^{-1/2}Q^{1/2}\|_{\mathrm{HS}}^2$$

Thus, (9) holds if

$$\|A^{-1/2}Q^{1/2}\|_{\mathrm{HS}}^2 < \infty.$$

- If $\mathrm{Tr}(Q) < \infty$, then

$$\|A^{-1/2}Q^{1/2}\|_{\mathrm{HS}}^2 \leq \|A^{-1/2}\|_{L(H)}^2 \|Q^{1/2}\|_{\mathrm{HS}}^2 = \|A^{-1/2}\|_{L(H)}^2 \,\mathrm{Tr}(Q) < \infty$$

  and hence there is a weak solution in any spatial dimension.
- If $Q = I$, then, using (7),

$$\|A^{-1/2}Q^{1/2}\|_{\mathrm{HS}}^2 = \|A^{-1/2}\|_{\mathrm{HS}}^2 = \sum_k \mu_k^{-1} \sim \sum_k k^{-2/d}.$$

  This is finite if and only if $d = 1$. Thus white noise is too irregular in higher spatial dimensions.

Examples: wave equation

We consider the stochastic wave equation

$$
\begin{aligned}
&\mathrm{d}\dot{u} - \Delta u \,\mathrm{d}t = \mathrm{d}W && \text{in } \mathcal{D} \times \mathbf{R}_+,\\
&u = 0 && \text{on } \partial\mathcal{D} \times \mathbf{R}_+,\\
&u(\cdot,0) = u_0,\ \dot{u}(\cdot,0) = u_1 && \text{in } \mathcal{D}.
\end{aligned}
$$

Let $\dot{H}^{-1} = (H_0^1(\mathcal{D}))^*$. We let $\Lambda = -\Delta$ with $\mathcal{D}(\Lambda) = H_0^1$ and we regard Λ as an operator $H_0^1 \subset \dot{H}^{-1} \to \dot{H}^{-1}$ by

$$
(\Lambda u)(v) = \langle \nabla u, \nabla v \rangle_{L_2(\mathcal{D})}.
$$

Let $U = L_2(\mathcal{D})$ and W be a Q-Wiener process on U as before. We put

$$
X = \begin{bmatrix} u \\ \dot{u} \end{bmatrix}, \quad \xi = \begin{bmatrix} u_0 \\ u_1 \end{bmatrix}, \quad H = L_2(\mathcal{D}) \times \dot{H}^{-1}.
$$

Now we can write

$$\begin{aligned}
\mathrm{d}X = \begin{bmatrix} \mathrm{d}u \\ \mathrm{d}\dot{u} \end{bmatrix} &= \begin{bmatrix} \dot{u}\,\mathrm{d}t \\ \Delta u\,\mathrm{d}t + \mathrm{d}W \end{bmatrix} \\
&= \begin{bmatrix} X_2 \\ -\Lambda X_1 \end{bmatrix} \mathrm{d}t + \begin{bmatrix} 0 \\ I \end{bmatrix} \mathrm{d}W \\
&= \begin{bmatrix} 0 & I \\ -\Lambda & 0 \end{bmatrix} X\,\mathrm{d}t + \begin{bmatrix} 0 \\ I \end{bmatrix} \mathrm{d}W \\
&= -AX\,\mathrm{d}t + B\,\mathrm{d}W,
\end{aligned}$$

where

$$A = \begin{bmatrix} 0 & -I \\ \Lambda & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ I \end{bmatrix}.$$

So we have

$$\begin{aligned}
&\mathrm{d}X + AX\,\mathrm{d}t = B\,\mathrm{d}W, \quad t > 0, \\
&X(0) = \xi,
\end{aligned} \tag{10}$$

where

$$\mathcal{D}(A) = \left\{ x \in H : Ax = \begin{bmatrix} -x_2 \\ \Lambda x_1 \end{bmatrix} \in H = L_2(\mathcal{D}) \times \dot{H}^{-1} \right\} = H_0^1(\mathcal{D}) \times L_2(\mathcal{D}).$$

Hence, in this case, $U \neq H$ and $B \neq I$. In order to see what $S(t) = e^{-tA}$ is, we note that $y(t) = S(t)x$ is the solution of

$$\dot{y} + Ay = 0; \quad y(0) = x,$$

that is,

$$\ddot{y}_1 + \Lambda y_1 = 0; \quad y_1(0) = x_1,\ \dot{y}_1(0) = x_2.$$

We solve it using an eigenfunction expansion:

$$\begin{aligned} y_1(t) &= \sum_{j=1}^{\infty} \cos(\sqrt{\mu_j}t)\langle x_1, \phi_j\rangle \phi_j + \frac{1}{\sqrt{\mu_j}} \sin(\sqrt{\mu_j}t)\langle x_2, \phi_j\rangle \phi_j \\ &= \cos(t\Lambda^{1/2})x_1 + \Lambda^{-1/2}\sin(t\Lambda^{1/2})x_2, \end{aligned}$$

and

$$y_2 = \dot{y}_1(t) = -\Lambda^{1/2}\sin(t\Lambda^{1/2})x_1 + \cos(t\Lambda^{1/2})x_2.$$

Now we can write the semigroup as

$$S(t) = \mathrm{e}^{-tA} = \begin{bmatrix} \cos(t\Lambda^{1/2}) & \Lambda^{-1/2}\sin(t\Lambda^{1/2}) \\ -\Lambda^{1/2}\sin(t\Lambda^{1/2}) & \cos(t\Lambda^{1/2}) \end{bmatrix}.$$

With $\xi = 0$ the evolution problem (10) has the unique weak solution

$$\begin{aligned} X(t) &= \int_0^t S(t-s)B\,\mathrm{d}W(s) \\ &= \begin{bmatrix} \displaystyle\int_0^t \Lambda^{-1/2}\sin((t-s)\Lambda^{1/2})\,\mathrm{d}W(s) \\ \displaystyle\int_0^t \cos((t-s)\Lambda^{1/2})\,\mathrm{d}W(s) \end{bmatrix}. \end{aligned}$$

For the existence and uniqueness of mild solutions one needs

$$\int_0^T \|S(t)BQ^{1/2}\|_{\mathrm{HS}}^2\,\mathrm{d}t < \infty.$$

We have,

$$\int_0^T \|S(t)BQ^{\frac{1}{2}}\|_{\mathrm{HS}}^2 \,\mathrm{d}t = \int_0^T \sum_k \|S(t)BQ^{\frac{1}{2}} f_k\|_H^2 \,\mathrm{d}t$$

$$= \int_0^T \sum_k \Big(\|\Lambda^{-\frac{1}{2}} \sin(t\Lambda^{\frac{1}{2}}) Q^{\frac{1}{2}} f_k\|_{L_2(\mathcal{D})}^2 + \|\cos(t\Lambda^{\frac{1}{2}}) Q^{\frac{1}{2}} f_k\|_{\dot{H}^{-1}}^2 \Big) \,\mathrm{d}t$$

$$= \int_0^T \Big(\|\Lambda^{-\frac{1}{2}} \sin(t\Lambda^{\frac{1}{2}}) Q^{\frac{1}{2}}\|_{\mathrm{HS}}^2 + \|\Lambda^{-\frac{1}{2}} \cos(t\Lambda^{\frac{1}{2}}) Q^{\frac{1}{2}}\|_{\mathrm{HS}}^2 \Big) \,\mathrm{d}t.$$

This must be finite.

- If $\mathrm{Tr}(Q) < \infty$:

$$\|\Lambda^{-\frac{1}{2}} \sin(t\Lambda^{\frac{1}{2}}) Q^{\frac{1}{2}}\|_{\mathrm{HS}}^2 \le \|\Lambda^{-\frac{1}{2}}\|_{L(L_2(\mathcal{D}))}^2 \|\sin(t\Lambda^{\frac{1}{2}})\|_{L(L_2(\mathcal{D}))}^2 \,\mathrm{Tr}(Q) < \infty,$$

  and similarly for cosine, so the condition holds in any spatial dimension.

- For $Q = I$ we have

$$\|\Lambda^{-\frac{1}{2}} \sin(t\Lambda^{\frac{1}{2}})Q^{\frac{1}{2}}\|_{\mathrm{HS}}^2 = \|\Lambda^{-\frac{1}{2}} \sin(t\Lambda^{\frac{1}{2}})\|_{\mathrm{HS}}^2$$
$$\leq \|\Lambda^{-\frac{1}{2}}\|_{\mathrm{HS}}^2 \|\sin(t\Lambda^{\frac{1}{2}})\|_{L(L_2(\mathcal{D}))}^2 \leq \|\Lambda^{-\frac{1}{2}}\|_{\mathrm{HS}}^2.$$

Similarly for the cosine operator. Here $\|\Lambda^{-\frac{1}{2}}\|_{\mathrm{HS}} < \infty$ if and only if $d = 1$ as we saw for the heat equation.

Note: this is where one needs the choice $H = L_2(\mathcal{D}) \times \dot{H}^{-1}$. Otherwise, if one takes $H = H_0^1(\mathcal{D}) \times L_2(\mathcal{D})$, then in case $Q = I$ one would need, for example,

$$\int_0^T \|\cos(t\Lambda^{\frac{1}{2}})\|_{\mathrm{HS}}^2 \, \mathrm{d}t < \infty$$

which does not hold in any spatial dimension.

Spatial approximation

Consider

$$\begin{cases} \mathrm{d}X_h(t) + A_h X_h(t)\,\mathrm{d}t = B_h\,\mathrm{d}W(t), \ t > 0, \\ X_h(0) = X_{0,h} \end{cases}$$

- $V_h \subset H$ finite dimensional
- $B_h : U \to V_h$, "approximation of B"
- $-A_h$ generates a C_0-semigroup $E_h(t) = e^{-tA_h}$ on V_h
- $X_{0,h}$ approximates X_0 in V_h

The weak solution is given by

$$X_h(t) = E_h(t)X_{0,h} + \int_0^t E_h(t-s)B_h\,\mathrm{d}W(s)$$

REMARK 1. In the rest of the lecture the semigroups e^{-tA} and e^{-tA_h} are denoted by $E(t)$ and $E_h(t)$ instead of $S(t)$ and $S_h(t)$, respectively, to avoid confusion with the finite element spaces S_h.
REMARK 2. In what follows (as before) we do not assume that A and Q commute unless it is explicitly stated.

Strong and weak error

- Strong error:

$$\|X_h(t) - X(t)\|_{L_2(\Omega,H)} = (\mathbf{E}\|X_h(t) - X(t)\|^2)^{1/2}$$

- Weak error:

$$\mathbf{E}G(X_h(T)) - \mathbf{E}G(X(T))$$

for $G : H \to \mathbb{R}$.

Strong error of the FEM for the stochastic heat equation

Recall, that the stochastic heat equation with additive noise can be written as

$$\begin{cases} \mathrm{d}X(t) + AX(t)\,\mathrm{d}t = B\,\mathrm{d}W(t), \ t > 0, \\ X(0) = X_0 \end{cases}$$

- $U = H = L_2(\mathcal{D})$, $\mathcal{D} \subset \mathbf{R}^d$
- $A = -\Delta$, $D(A) = H^2(\mathcal{D}) \cap H_0^1(\mathcal{D})$
- $W(t) = \sum_{k=1}^{\infty} \sqrt{\lambda_k}\beta_k(t)e_k$, Q-Wiener process on U, $Qe_k = \lambda_k e_k$

Analytic semigroup:

$$\mathrm{e}^{-tA}v = \sum_{j=1}^{\infty} \mathrm{e}^{-t\mu_j}(v, \phi_j)\phi_j$$

can be extended as a holomorphic function of t.

Smoothing property:

$$\|A^{\beta}\mathrm{e}^{-tA}v\| \le Ct^{-\beta}\|v\|, \quad \beta \ge 0$$

$$\int_0^t \|A^{1/2}\mathrm{e}^{-sA}v\|^2\,\mathrm{d}s \le C\|v\|^2$$

Regularity of the solution

In order to describe regularity of functions on a fine scale we define norms

$$|v|_s = \Big(\sum_j \mu_j^s |\langle v, \phi_j \rangle|^2 \Big)^{1/2} = \|A^{s/2} v\|, \quad s \in \mathbf{R}.$$

For $s \geq 0$ we define the corresponding spaces:

$$\dot{H}^s = \Big\{ v \in H : |v|_s < \infty \Big\},$$

$\dot{H}^{-s}$ is the closure of H with respect to the $\dot{H}^s$-norm.

The negative order space $\dot{H}^{-s}$ can be identified with the dual space $(\dot{H}^s)^*$. Then we have $\dot{H}^s \subset H = \dot{H}^0 \subset \dot{H}^{-s}$. It is known that $\dot{H}^1 = H_0^1$, $\dot{H}^2 = H^2 \cap H_0^1 = \mathcal{D}(A)$.

Finally, set

$$\|v\|^2_{L_2(\Omega, \dot{H}^\beta)} = \mathbf{E}(|v|^2_\beta) = \int_\Omega \int_{\mathcal{D}} |A^{\beta/2} v|^2 \, \mathrm{d}\xi \, \mathrm{d}\mathbf{P}(\omega), \quad \beta \in \mathbf{R}.$$

Regularity of the solution

THEOREM [Yan'04]. If $\|A^{(\beta-1)/2}Q^{1/2}\|_{\mathrm{HS}} < \infty$ for some $\beta \geq 0$, then
$\|X(t)\|_{L_2(\Omega,\dot{H}^\beta)} \leq C\Big(\|X_0\|_{L_2(\Omega,\dot{H}^\beta)} + \|A^{(\beta-1)/2}Q^{1/2}\|_{\mathrm{HS}}\Big)$

Two cases:

- If $\|Q^{1/2}\|_{\mathrm{HS}}^2 = \sum_{j=1}^\infty \|Q^{1/2}e_j\|^2 = \sum_{j=1}^\infty \lambda_j = \mathrm{Tr}(Q) < \infty$, then $\beta = 1$.
- If $Q = I,\ d = 1,\ A = -\frac{\partial^2}{\partial\xi^2}$, then $\|A^{(\beta-1)/2}\|_{\mathrm{HS}} < \infty$ for $\beta < 1/2$.

  $\|A^{(\beta-1)/2}\|_{\mathrm{HS}}^2 = \sum_j \mu_j^{-(1-\beta)} \approx \sum_j j^{-(1-\beta)2/d} < \infty$ iff $d = 1$, $\beta < 1/2$

Proof with $X_0 = 0$

$$
\begin{aligned}
\|X(t)\|^2_{L_2(\Omega,\dot H^\beta)} &= \mathbf{E}\Big(\Big\| \int_0^t A^{\beta/2}E(t-s)\,\mathrm{d}W(s)\Big\|^2\Big) \\
&= \int_0^t \|A^{\beta/2}E(s)Q^{1/2}\|^2_{\mathrm{HS}}\,\mathrm{d}s \\
&= \int_0^t \|A^{1/2}E(s)A^{(\beta-1)/2}Q^{1/2}\|^2_{\mathrm{HS}}\,\mathrm{d}s \\
&= \sum_{k=1}^\infty \int_0^t \|A^{1/2}E(s)A^{(\beta-1)/2}Q^{1/2}\phi_k\|^2\,\mathrm{d}s \\
&\le C\sum_{k=1}^\infty \|A^{(\beta-1)/2}Q^{1/2}\phi_k\|^2 \\
&= C\|A^{(\beta-1)/2}Q^{1/2}\|^2_{\mathrm{HS}} \qquad \boxed{\int_0^t \|A^{1/2}E(s)v\|^2\,\mathrm{d}s \le C\|v\|^2}
\end{aligned}
$$

The finite element method

Deterministic heat equation, with $u_t = \dfrac{\partial u}{\partial t}$:

$$\begin{cases} u_t - \Delta u = f, & x \in \mathcal{D},\ t > 0 \\ u = 0, & x \in \partial\mathcal{D},\ t > 0 \end{cases}$$

$$\langle u, v \rangle = \int_{\mathcal{D}} uv \, \mathrm{d}\xi$$

$$\langle u_t, v \rangle \underbrace{-(\Delta u, v)}_{=\langle \nabla u, \nabla v \rangle} = \langle f, v \rangle, \quad v \in H_0^1(\mathcal{D}) = \dot{H}^1$$

weak form:

$$\begin{cases} u(t) \in H_0^1(\mathcal{D}), \quad u(0) = u_0 \\ \langle u_t, v \rangle + \langle \nabla u, \nabla v \rangle = \langle f, v \rangle, \qquad \forall v \in H_0^1(\mathcal{D}) \end{cases}$$

The finite element method

weak form: $\begin{cases} u(t) \in H_0^1(\mathcal{D}), \quad u(0) = u_0 \\ \langle u_t, v\rangle + \langle \nabla u, \nabla v\rangle = \langle f, v\rangle, \qquad \forall v \in H_0^1(\mathcal{D}) \end{cases}$

- triangulations $\{\mathcal{T}_h\}_{0<h<1}$, mesh size h
- function spaces $\{S_h\}_{0<h<1}$, continuous piecewise linear functions
- $S_h \subset H_0^1(\mathcal{D}) = \dot{H}^1$

$$\begin{cases} u_h(t) \in S_h, \quad \langle u_h(0), v\rangle = \langle u_0, v\rangle, \qquad \forall v \in S_h \\ \langle u_{h,t}, v\rangle + \langle \nabla u_h, \nabla v\rangle = \langle f, v\rangle, \qquad \forall v \in S_h \end{cases}$$

$$\langle u_{h,t}, v\rangle + \underbrace{\langle \nabla u_h, \nabla v\rangle}_{=\langle A_h u_h, v\rangle} = \underbrace{\langle f, v\rangle}_{=\langle P_h f, v\rangle}, \quad \forall v \in S_h$$

$$\begin{cases} u_{h,t} + A_h u_h = P_h f, \quad t > 0 \\ u_h(0) = P_h u_0 \end{cases}$$

the same abstract framework

The finite element method

- triangulations $\{\mathcal{T}_h\}_{0<h<1}$, mesh size h
- finite element spaces $\{S_h\}_{0<h<1}$
- $S_h \subset H_0^1(\mathcal{D}) = \dot{H}^1$
- S_h continuous piecewise linear functions
- $A_h : S_h \to S_h,$ discrete Laplacian, $\quad \langle A_h \psi, \chi \rangle = \langle \nabla \psi, \nabla \chi \rangle, \ \forall \chi \in S_h$
- $P_h : L_2 \to S_h,$ orthogonal projection, $\quad \langle P_h f, \chi \rangle = \langle f, \chi \rangle, \ \forall \chi \in S_h$

$$\begin{cases} X_h(t) \in S_h, \quad X_h(0) = P_h X_0 \\ \mathrm{d}X_h + A_h X_h \, \mathrm{d}t = P_h \, \mathrm{d}W \end{cases}$$

Weak solution, with $E_h(t) = e^{-tA_h}$, is the stochastic convolution

$$X_h(t) = E_h(t) P_h X_0 + \int_0^t E_h(t-s) P_h \, \mathrm{d}W(s)$$

Error estimates for the deterministic problem

$$\begin{cases} u_t + Au = 0, & t > 0 \\ u(0) = v \end{cases} \qquad \begin{cases} u_{h,t} + A_h u_h = 0, & t > 0 \\ u_h(0) = P_h v \end{cases}$$

$u(t) = E(t)v$ $\qquad$ $u_h(t) = E_h(t)P_h v$

Our basic assumption on the finite element method is that the Ritz projection $R_h : \dot{H}^1 \to S_h$ defined as

$$\langle \nabla R_h v, \nabla \chi \rangle = \langle \nabla v, \nabla \chi \rangle, \; v \in \dot{H}^1, \; \chi \in S_h,$$

satisfies the error bound

$$\|R_h v - v\| \le C h^\beta |v|_\beta, \; v \in \dot{H}^\beta, \; 1 \le \beta \le 2.$$

Example: D is a convex polygonal domain, with a regular family of triangulations of D with maximum mesh-size h.

Error estimates for the deterministic problem

Denote

$$F_h(t)v = E_h(t)P_h v - E(t)v, \qquad |v|_\beta = \|A^{\beta/2}v\|$$

We have, for $0 \le \beta \le 2$ ([Th'06], see also, [Yan'04]),

- $\|F_h(t)v\| \le Ch^\beta |v|_\beta, \quad t \ge 0$
- $\left(\int_0^t \|F_h(s)v\|^2\,\mathrm{d}s\right)^{1/2} \le Ch^\beta |v|_{\beta-1}, \quad t \ge 0$

For piecewise polynomials of order $r-1$: $0 \le \beta \le r$.

Strong convergence

THEOREM [Yan'04]. If $\|A^{(\beta-1)/2}Q^{1/2}\|_{\mathrm{HS}} < \infty$ for some $\beta \in [0,2]$, then
$\|X_h(t) - X(t)\|_{L_2(\Omega,H)} \leq Ch^{\beta}\Big(\|X_0\|_{L_2(\Omega,\dot{H}^{\beta})} + \|A^{(\beta-1)/2}Q^{1/2}\|_{\mathrm{HS}}\Big)$

For piecewise polynomials of order $r-1$: $\beta \in [0,r]$.

Recall: $\|X_h(t) - X(t)\|_{L_2(\Omega,H)} = \Big(\mathbf{E}(\|X_h(t) - X(t)\|^2)\Big)^{1/2}$

Two cases:

- If $\|Q^{1/2}\|_{\mathrm{HS}}^2 = \mathrm{Tr}(Q) < \infty$, then the convergence rate is $O(h)$.
- If $Q = I,\ d = 1$, then the rate is almost $O(h^{1/2})$.

Strong convergence: proof

$$X(t) = E(t)X_0 + \int_0^t E(t-s)\,\mathrm{d}W(s)$$

$$X_h(t) = E_h(t)P_hX_0 + \int_0^t E_h(t-s)\,P_h\mathrm{d}W(s)$$

$$F_h(t) = E_h(t)P_h - E(t)$$

$$X_h(t) - X(t) = F_h(t)X_0 + \int_0^t F_h(t-s)\,\mathrm{d}W(s) = e_1(t) + e_2(t)$$

$\|F_h(t)X_0\| \le Ch^\beta |X_0|_\beta$ (deterministic error estimate)

$\implies \quad \|e_1(t)\|_{L_2(\Omega,H)} \le Ch^\beta \|X_0\|_{L_2(\Omega,\dot{H}^\beta)}$

Strong convergence: proof

$$\begin{cases} \mathbf{E}\Big\| \int_0^t B(s)\,\mathrm{d}W(s) \Big\|^2 = \int_0^t \|B(s)Q^{1/2}\|_{\mathrm{HS}}^2\,\mathrm{d}s \text{ (isometry)} \\ \Big(\int_0^t \|F_h(s)v\|^2\,\mathrm{d}s \Big)^{1/2} \le Ch^\beta |v|_{\beta-1}, \text{ with } v = Q^{1/2}\varphi_l \text{ (deterministic)} \end{cases}$$

$\Longrightarrow$

$$\begin{aligned} \|e_2(t)\|_{L_2(\Omega,H)}^2 &= \mathbf{E}\Big\| \int_0^t F_h(t-s)\,\mathrm{d}W(s) \Big\|^2 = \int_0^t \|F_h(t-s)Q^{1/2}\|_{\mathrm{HS}}^2\,\mathrm{d}s \\ &= \sum_{l=1}^\infty \int_0^t \|F_h(t-s)Q^{1/2}\varphi_l\|^2\,\mathrm{d}s \le C \sum_{l=1}^\infty h^{2\beta} |Q^{1/2}\varphi_l|_{\beta-1}^2 \\ &= Ch^{2\beta} \sum_{l=1}^\infty \|A^{(\beta-1)/2}Q^{1/2}\varphi_l\|^2 = Ch^{2\beta} \|A^{(\beta-1)/2}Q^{1/2}\|_{\mathrm{HS}}^2 \end{aligned}$$

If $\mathrm{Tr}(Q) < \infty$, we may choose $\beta = 1$, otherwise $\beta < 1$.

Note on the simulation

The process $\{P_h W(t)\}_{t\geq 0}$ is a $P_h Q P_h$-Wiener process. There are two ways of simulating it.

- Take a basis $\{\phi_j^h\}_{j=1}^{N_h}$ for S_h. Calculate the matrix $(Q_h)_{i,j} = \langle Q\phi_i^h, \phi_j^h\rangle$. Take the Cholesky decomposition of $Q_h = LL^*$. Take independent Brownian motions in a vector $\beta(t) = (\beta_1(t), ..., \beta_{N_h}(t))$. Then $L\beta(t)$ is the required process with respect to the basis $\{\phi_j^h\}_{j=1}^{N_h}$.
- Solve the eigenvalue problem $Qu = \lambda u$ on S_h with the standard finite element method. This yields the eigenpairs $(\lambda_{j,h}, e_{j,h})_{j=1}^{N_h}$. Take independent Brownian motions $(\beta_j(t))_{j=1}^{N_h}$. Then

  $$W_h(t) = \sum_{j=1}^{N_h} \lambda_{j,h}^{1/2} \beta_j(t) e_{j,h}$$

  is the required process. This can be very expensive. However if the noise is smooth, then it might be enough to take less then N_h terms and still preserve the order of the FEM. For example, for the Gauss kernel, it is enough to take $\sim (\ln N_h)^d$ terms, see [KLL'09a]! So strong spatial correlation helps.

Strong error of the FEM for the stochastic wave equation

Recall that the stochastic wave equation with additive noise (with 0 initial conditions for simplicity) in the abstract form:

$$\begin{cases} \mathrm{d}X(t) + AX(t)\,\mathrm{d}t = B\,\mathrm{d}W(t), \quad t > 0 \\ X(0) = 0 \end{cases}$$

- $X(t)$, $H = \dot{H}^0 \times \dot{H}^{-1}$-valued stochastic process
- $W(t) = \sum_{k=1}^{\infty} \sqrt{\lambda_k}\beta_k(t)e_k$, Q-Wiener process on $U = \dot{H}^0$, $Qe_k = \lambda_k e_k$
- $\Lambda = -\Delta : \; \dot{H}^{-1} \to \dot{H}^{-1}, \quad D(\Lambda) = H_0^1(\mathcal{D}) = \dot{H}^1$
- $\begin{bmatrix} \mathrm{d}u \\ \mathrm{d}u_t \end{bmatrix} = \begin{bmatrix} 0 & I \\ -\Lambda & 0 \end{bmatrix} \begin{bmatrix} u \\ u_t \end{bmatrix} \mathrm{d}t + \begin{bmatrix} 0 \\ I \end{bmatrix} \mathrm{d}W,$

  $X = \begin{bmatrix} u \\ u_t \end{bmatrix}, \; A = -\begin{bmatrix} 0 & I \\ -\Lambda & 0 \end{bmatrix}, \; B = \begin{bmatrix} 0 \\ I \end{bmatrix}$
- $H = \dot{H}^0 \times \dot{H}^{-1}, \quad D(A) = \dot{H}^1 \times \dot{H}^0, \quad U = \dot{H}^0 = L_2(\mathcal{D})$
- $E(t) = e^{-tA} = \begin{bmatrix} \cos(t\Lambda^{1/2}) & \Lambda^{-1/2}\sin(t\Lambda^{1/2}) \\ -\Lambda^{1/2}\sin(t\Lambda^{1/2}) & \cos(t\Lambda^{1/2}) \end{bmatrix}$, C_0-semigroup

Regularity

THEOREM [KLS'09]. If $\|\Lambda^{(\beta-1)/2}Q^{1/2}\|_{\mathrm{HS}} < \infty$ for some $\beta \geq 0$, then there exists a unique weak solution

$$X(t) = \begin{bmatrix} X_1(t) \\ X_2(t) \end{bmatrix} = \int_0^t E(t-s)B\,\mathrm{d}W(s) = \begin{bmatrix} \int_0^t \Lambda^{-1/2}\sin\left((t-s)\Lambda^{1/2}\right)\mathrm{d}W(s) \\ \int_0^t \cos\left((t-s)\Lambda^{1/2}\right)\mathrm{d}W(s) \end{bmatrix}$$

and

$$\|X(t)\|_{L_2(\Omega,\dot{H}^\beta\times\dot{H}^{\beta-1})} \leq C(t)\|\Lambda^{(\beta-1)/2}Q^{1/2}\|_{\mathrm{HS}}.$$

Two cases:

- If $\|Q^{1/2}\|_{\mathrm{HS}}^2 = \mathrm{Tr}(Q) < \infty$, then $\beta = 1$.
- If $Q = I$, then $\|\Lambda^{(\beta-1)/2}\|_{\mathrm{HS}} < \infty$ iff $d = 1,\ \beta < 1/2$.

The finite element method

Spatial discretization

- triangulations $\{\mathcal{T}_h\}_{0<h<1}$, mesh size h
- finite element spaces $\{S_h\}_{0<h<1}$
- $S_h \subset \dot{H}^1 = H_0^1(\mathcal{D})$ continuous piecewise linear functions
- $\Lambda_h : S_h \to S_h,$ discrete Laplacian, $\quad \langle \Lambda_h \psi, \chi \rangle = \langle \nabla \psi, \nabla \chi \rangle, \ \forall \chi \in S_h$
- $P_h : \dot{H}^0 \to S_h,$ orthogonal projection, $\quad \langle P_h f, \chi \rangle = \langle f, \chi \rangle, \ \forall \chi \in S_h$
- $A_h = \begin{bmatrix} 0 & I \\ -\Lambda_h & 0 \end{bmatrix}, \quad B_h = \begin{bmatrix} 0 \\ P_h \end{bmatrix}$
- $\begin{cases} \mathrm{d}X_h(t) + A_h X_h(t)\,\mathrm{d}t = B_h\,\mathrm{d}W(t), \quad t > 0 \\ X_h(0) = 0 \end{cases}$
- $E_h(t) = e^{-tA_h} = \begin{bmatrix} \cos(t\Lambda_h^{1/2}) & \Lambda_h^{-1/2}\sin(t\Lambda_h^{1/2}) \\ -\Lambda_h^{1/2}\sin(t\Lambda_h^{1/2}) & \cos(t\Lambda_h^{1/2}) \end{bmatrix}$

The finite element method (continued)

The weak solution is:

$$X_h(t) = \begin{bmatrix} X_{h,1}(t) \\ X_{h,2}(t) \end{bmatrix}$$

$$= \int_0^t E_h(t-s) B_h \, \mathrm{d}W(s) = \begin{bmatrix} \int_0^t \Lambda_h^{-1/2} \sin\left((t-s)\Lambda_h^{1/2}\right) P_h \, \mathrm{d}W(s) \\ \int_0^t \cos\left((t-s)\Lambda_h^{1/2}\right) P_h \, \mathrm{d}W(s) \end{bmatrix}$$

where

$$\cos(t\Lambda_h^{1/2})v = \sum_{j=1}^{N_h} \cos\left(t\sqrt{\mu_{h,j}}\right) \langle v, \phi_{h,j} \rangle \phi_{h,j}$$

$\mu_{h,j}$, $\phi_{h,j}$ are the eigenpairs of Λ_h

Error estimates for the deterministic problem

$$\begin{cases} v_{tt}(t) + \Lambda v(t) = 0,\ t > 0 \\ v(0) = 0,\ v_t(0) = f \end{cases} \quad \Rightarrow v(t) = \Lambda^{-1/2} \sin\left(t\Lambda^{1/2}\right) f$$

$$\begin{cases} v_{h,tt}(t) + \Lambda_h v_h(t) = 0,\ t > 0 \\ v_h(0) = 0,\ v_{h,t}(0) = P_h f \end{cases} \quad \Rightarrow v_h(t) = \Lambda_h^{-1/2} \sin\left(t\Lambda_h^{1/2}\right) P_h f$$

As for the heat equation, our basic assumption on the finite element method is that the Ritz projection satisfies the error bound

$$\|R_h v - v\| \leq Ch^{\beta} |v|_{\beta},\ v \in \dot{H}^{\beta},\ 1 \leq \beta \leq 2.$$

Then, we have [KLS'09], for $K_h(t) = \Lambda_h^{-1/2} \sin\left(t\Lambda_h^{1/2}\right) P_h - \Lambda^{-1/2} \sin\left(t\Lambda^{1/2}\right)$

$\|K_h(t) f\| \leq C(t) h^2 |f|_2$ "initial regularity of order 3"

$\|K_h(t) f\| \leq 2 |f|_{-1}$ "initial regularity of order 0" (stability)

$\|K_h(t) f\| \leq C(t) h^{\frac{2}{3}\beta} |f|_{\beta - 1}, \quad 0 \leq \beta \leq 3$

$\beta - 1$ can not be replaced by $\beta - 1 - \epsilon$ for $\epsilon > 0$ (J. Rauch 1985)

Strong convergence

THEOREM [KLS'09]. If $\|\Lambda^{(\beta-1)/2}Q^{1/2}\|_{\mathrm{HS}} < \infty$ for some $\beta \in [0, 3]$, then

$$\|X_{h,1}(t) - X_1(t)\|_{L_2(\Omega,\dot{H}^0)} \leq C(t)\, h^{\frac{2}{3}\beta} \big\|\Lambda^{(\beta-1)/2}Q^{1/2}\big\|_{\mathrm{HS}}$$

Higher order FEM: $\quad O(h^{\frac{r}{r+1}\beta}), \quad \beta \in [0, r+1].$

PROOF. $\{f_k\}$ an arbitrary ON basis in $\dot{H}^0$

$$\begin{aligned}
\|X_{h,1}(t) - X_1(t)\|^2_{L_2(\Omega,\dot{H}^0)} &= \mathbf{E}\big(\|X_{h,1}(t) - X_1(t)\|^2\big) \\
&= \mathbf{E}\Big(\Big\|\int_0^t K_h(t-s)\,\mathrm{d}W(s)\Big\|^2\Big) \\
\{\text{Isometry}\} &= \int_0^t \big\|K_h(s)Q^{1/2}\big\|^2_{\mathrm{HS}}\,\mathrm{d}s = \int_0^t \sum_{k=1}^{\infty} \big\|K_h(s)Q^{1/2}f_k\big\|^2\,\mathrm{d}s \\
&\leq C(t)h^{\frac{4}{3}\beta} \sum_{k=1}^{\infty} \Big|Q^{1/2}f_k\Big|^2_{\beta-1} = C(t)h^{\frac{4}{3}\beta}\big\|\Lambda^{(\beta-1)/2}Q^{1/2}\big\|^2_{\mathrm{HS}}
\end{aligned}$$

Strong convergence: special cases and comparison with the heat equation

Two cases:

- If $\|Q^{1/2}\|_{\mathrm{HS}}^2 = \mathrm{Tr}(Q) < \infty$, then $\beta = 1$.
  $\|X_{h,1}(t) - X_1(t)\|_{L_2(\Omega,\dot{H}^0)} \leq C(t)h^{2/3}$
- If $Q = I$, then $\|\Lambda^{(\beta-1)/2}\|_{\mathrm{HS}} < \infty$ iff $d = 1,\ 0 \leq \beta < 1/2$.
  $\|X_{h,1}(t) - X_1(t)\|_{L_2(\Omega,\dot{H}^0)} \leq C(t)h^s, \quad s < 1/3$

Comparison with the heat equation: regularity is the same for both the heat and wave equations:

$$\|X_1(t)\|_{L_2(\Omega,\dot{H}^\beta)} \leq C\|\Lambda^{(\beta-1)/2}Q^{1/2}\|_{\mathrm{HS}}$$

Strong convergence:

$$\|X_{h,1}(t) - X_1(t)\|_{L_2(\Omega,\dot{H}^0)} \leq C\, h^{\frac{2}{3}\beta}\, \|\Lambda^{(\beta-1)/2}Q^{1/2}\|_{\mathrm{HS}} \qquad \text{(wave equation)}$$

$$\|X_h(t) - X(t)\|_{L_2(\Omega,\dot{H}^0)} \leq C\, h^{\beta}\, \|\Lambda^{(\beta-1)/2}Q^{1/2}\|_{\mathrm{HS}} \qquad \text{(heat equation)}$$

Weak error representation: preliminaries

Consider a general linear SPDE with additive noise

$$\begin{cases} \mathrm{d}X(t) + AX(t)\,\mathrm{d}t = B\,\mathrm{d}W(t), \ t > 0, \\ X(0) = X_0 \end{cases}$$

and its spatial approximation

$$\begin{cases} \mathrm{d}X_h(t) + A_hX_h(t)\,\mathrm{d}t = B_h\,\mathrm{d}W(t), \ t > 0, \\ X_h(0) = X_{0,h}. \end{cases}$$

- $V_h \subset H$ finite dimensional
- $B_h : U \to V_h$, "approximation of B"
- $-A_h$ generates a C_0-semigroup $E_h(t) = e^{-tA_h}$ on V_h
- $X_{0,h}$ approximates X_0 in V_h

Weak error representation: preliminaries

Now consider the simple SPDE

$$\mathrm{d}Y(t) = E(T-t)B\,\mathrm{d}W(t),\ \ t \in (0,T];\ \ Y(0) = E(T)X_0,$$

with weak (in this simple case also strong) solution

$$Y(t) = E(T)X_0 + \int_0^t E(T-s)B\,\mathrm{d}W(s).$$

Similarly, consider

$$\mathrm{d}Y_h(t) = E_h(T-t)B\,\mathrm{d}W(t),\ \ t \in (0,T];\ \ Y_h(0) = E_h(T)X_{0,h},$$

with weak solution

$$Y_h(t) = E_h(T)X_{0,h} + \int_0^t E_h(T-s)B_h\,\mathrm{d}W(s).$$

Notice that $X(T) = Y(T)$, $X_h(T) = Y_h(T)$ and there is NO DRIFT term for Y and Y_h.

Weak error representation: preliminaries

In general, consider the auxiliary problem

$$\mathrm{d}Z(t) = E(T-t)B\,\mathrm{d}W(t),\ t \in (\tau, T];\ Z(\tau) = \xi,$$

where ξ is a $\mathcal{F}_\tau$-measurable random variable. Then the unique weak solution is given by

$$Z(t,\tau,\xi) = \xi + \int_\tau^t E(T-s)B\,\mathrm{d}W(s).$$

Define $u : H \times [0, T] \to \mathbb{R}$ by

$$u(x,t) = \mathbf{E}\left(G(Z(T,t,x))\right).$$

If $G \in C_{\mathrm{b}}^2(H,\mathbb{R})$, then it is well known that u is a solution to Kolmogorov's equation

$$\begin{aligned} &u_t(x,t) + \frac{1}{2}\mathrm{Tr}\left(u_{xx}(x,t)E(T-t)BQ[E(T-t)B]^*\right) = 0,\\ &u(x,T) = G(x), \quad t \in [0,T),\ x \in H. \end{aligned}$$

Weak error representation: preliminaries

$C_{\mathrm{b}}^2(H,\mathbb{R})$ is the set of all real-valued, twice Fréchet differentiable functions G whose first and second derivatives are continuous and bounded. By the Riesz representation theorem, we may identify the first derivative $DG(x)$ at $x \in H$ with an element $G'(x) \in H$ such that

$$DG(x)y = \langle G'(x), y \rangle, \quad y \in H,$$

and the second derivative $D^2G(x)$ with a symmetric linear operator $G''(x) \in \mathcal{B}(H)$ such that

$$D^2G(x)(y,z) = \langle G''(x)y, z \rangle, \quad y, z \in H.$$

We say that $G \in C^2(H,\mathbb{R})$ if G, G', and G'' are continuous, that is, $G \in C(H,\mathbb{R})$, $G' \in C(H,H)$, and $G'' \in C(H,\mathcal{B}(H))$. We define

$$C_{\mathrm{b}}^2(H) := \left\{ G \in C^2(H,\mathbb{R}) : \|G\|_{C_{\mathrm{b}}^2(H)} < \infty \right\},$$

with the seminorm

$$\|G\|_{C_{\mathrm{b}}^2(H)} := \sup_{x\in H} \|G'(x)\|_H + \sup_{x \in H} \|G''(x)\|_{\mathcal{B}(H)}.$$

Weak error representation

THEOREM [KLL'09]. If

$$\mathrm{Tr}\left(\int_0^T E(t)BQ[E(t)B]^*\,\mathrm{d}t\right) < \infty$$

and $G \in C_{\mathrm{b}}^2(H, \mathbb{R})$, then the weak error $e_h(T)$ has the representation

$$\begin{aligned}
e_h(T) &= \mathbf{E}\left(u(Y_h(0),0) - u(Y(0),0)\right) \\
&\quad + \frac{1}{2}\mathbf{E}\int_0^T \mathrm{Tr}\Big(u_{xx}(Y_h(t),t) \\
&\qquad \times [E_h(T-t)B_h + E(T-t)B]Q[E_h(T-t)B_h - E(T-t)B]^*\Big)\,\mathrm{d}t \\
&= \mathbf{E}\left(u(Y_h(0),0) - u(Y(0),0)\right) \\
&\quad + \frac{1}{2}\mathbf{E}\int_0^T \mathrm{Tr}\Big(u_{xx}(Y_h(t),t) \\
&\qquad \times [E_h(T-t)B_h - E(T-t)B]Q[E_h(T-t)B_h + E(T-t)B]^*\Big)\,\mathrm{d}t.
\end{aligned}$$

Weak error representation: proof

If ξ is $\mathcal{F}_t$ measurable, then $u(\xi,t) = \mathbf{E}\left(G(Z(T,t,\xi))\middle|\mathcal{F}_t\right)$. Therefore, by the law of double expectation,

$$\mathbf{E}\Big(u(\xi,t)\Big) = \mathbf{E}\Big(\mathbf{E}\Big(G(Z(T,t,\xi))\Big|\mathcal{F}_t\Big)\Big) = \mathbf{E}\Big(G(Z(T,t,\xi))\Big).$$

Thus,

$$\mathbf{E}\Big(u(Y(0),0)\Big) = \mathbf{E}\Big(G(Z(T,0,Y(0))\Big) = \mathbf{E}\Big(G(Y(T))\Big) = \mathbf{E}\Big(G(X(T))\Big)$$

and with $\xi = Y_h(T)$

$$\mathbf{E}\Big(u(Y_h(T),T)\Big) = \mathbf{E}\Big(G(Z(T,T,Y_h(T)))\Big) = \mathbf{E}\Big(G(Y_h(T))\Big) = \mathbf{E}\Big(G(X_h(T))\Big).$$

Hence,

$$\begin{aligned} e_h(T) &= \mathbf{E}\Big(G(X_h(T)) - G(X(T))\Big) = \mathbf{E}\Big(u(Y_h(T),T) - u(Y(0),0)\Big) \\ &= \mathbf{E}\Big(u(Y_h(0),0) - u(Y(0),0)\Big) + \mathbf{E}\Big(u(Y_h(T),T) - u(Y_h(0),0)\Big). \end{aligned}$$

Weak error representation: proof

Using Itô's formula for $u(Y_h(t), t)$ and Kolmogorov's equation

$$
\begin{aligned}
&\mathbf{E}\Big(u(Y_h(T), T) - u(Y_h(0), 0)\Big) \\
&= \mathbf{E}\int_0^T u_t(Y_h(t), t) \\
&\quad + \frac{1}{2}\operatorname{Tr}\Big(u_{xx}(Y_h(t), t)[E_h(T-t)B_h]Q[E_h(T-t)B_h]^*\Big)\,\mathrm{d}t \\
&= \frac{1}{2}\mathbf{E}\int_0^t \operatorname{Tr}\Big(u_{xx}(Y_h(t), t) \\
&\qquad \times [E_h(T-t)B_h]Q[E_h(T-t)B_h]^* - [E(T-t)B]Q[E(T-t)B]^*\Big)\,\mathrm{d}t.
\end{aligned}
$$

The proof can be finished by algebraic manipulation and playing around with traces.

Applications: heat equation

The stochastic heat equation and its finite element approximation in abstract form are

$$\begin{cases} \mathrm{d}X(t) + AX(t)\,\mathrm{d}t = B\,\mathrm{d}W(t), \ t > 0, \\ X(0) = X_0 \end{cases}$$

and

$$\begin{cases} \mathrm{d}X_h(t) + A_hX_h(t)\,\mathrm{d}t = B_h\,\mathrm{d}W(t), \ t > 0, \\ X_h(0) = X_{0,h}. \end{cases}$$

- $\Lambda := -\Delta$ with $\mathcal{D}(\Lambda) = H^2(D) \cap H_0^1(D)$.
- $U = H := L_2(D)$ with norm $\|\cdot\|$ and inner product $\langle \cdot, \cdot \rangle$
- $V_h = S_h$ (continuous piecewise polynomials of order $r - 1$)
- $A := \Lambda$, $A_h = \Lambda_h$
- $B := I$, $B_h = P_h$
- $X_{0,h} = P_hX_0$
- $W(t) = \sum_{k=1}^{\infty} \sqrt{\lambda_k}\beta_k(t)e_k$, Q-Wiener process on U, $Qe_k = \lambda_k e_k$

Applications: heat equation

THEOREM [DP'09], [GKL'09], [KLL'09]. Let $g \in C_{\mathrm{b}}^2(H, \mathbb{R})$ and assume that $\|A^{\frac{\beta-1}{2}} Q^{\frac{1}{2}}\|_{\mathrm{HS}} < \infty$ for some $\beta \in (0, 1]$. Then, there are $C > 0$, $h_0 > 0$, depending on g, X_0, Q, β, and T but not on h, such that for $h \le h_0$,

$$\Big|\mathbf{E}\big(g(X_h(T)) - g(X(T))\big)\Big| \le Ch^{2\beta}|\ln h|.$$

If, in addition $X_0 \in L_1(\Omega, \dot{H}^{2\beta})$, then C is independent of T as well.

The proof uses the error representation theorem, the basic deterministic finite element estimate

$$\|(E_h(t)P_h - E(t))v\| \le Ch^s t^{-\frac{s-\gamma}{2}}|v|_\gamma,\ 0 \le \gamma \le s \le r,$$

and that

$$\sup_{(x,t)\in H\times[0,T]} \|u_{xx}(x,t)\|_{\mathcal{B}(H)} \le \sup_{x\in H} \|g''(x)\|_{\mathcal{B}(H)}.$$

Applications: heat equation

REMARK. This is twice the rate of the strong convergence under the condition $\|A^{\frac{\beta-1}{2}}Q^{\frac{1}{2}}\|_{\mathrm{HS}} < \infty$.

Special cases:

- $Q = I \Rightarrow d = 1,\ 2\beta < 1$
- $\mathrm{Tr}\, Q < \infty \Rightarrow 2\beta = 2.$

Under a slightly stronger condition on A and Q the result can be extended to the case $\beta > 1$.

THEOREM [KLL'09]. Let $g \in C_{\mathrm{b}}^2(H, \mathbb{R})$ and assume that $\|A^{\beta-1}Q\|_{\mathrm{Tr}} < \infty$ for some $\beta \in [1, \frac{r}{2}]$. Then there are $C > 0$, $h_0 > 0$, depending on g, X_0, Q, β, and T but not on h, such that for $h \le h_0$,

$$\left|\mathbf{E}\big(g(X_h(T)) - g(X(T))\big)\right| \le Ch^{2\beta}|\ln h|.$$

If, in addition $X_0 \in L_1(\Omega, \dot{H}^{2\beta})$, then C is independent of T as well.

REMARK. We have that $\|A^{\frac{\beta-1}{2}}Q^{\frac{1}{2}}\|_{\mathrm{HS}}^2 \le \|A^{\beta-1}Q\|_{\mathrm{Tr}}$.
If A and Q "commute", then $\|A^{\frac{\beta-1}{2}}Q^{\frac{1}{2}}\|_{\mathrm{HS}}^2 = \|A^{\beta-1}Q\|_{\mathrm{Tr}}$.

Applications: wave equation

Recall again the stochastic wave equation with additive noise in the abstract form:

$$\begin{cases} \mathrm{d}X(t) + AX(t)\,\mathrm{d}t = B\,\mathrm{d}W(t), \quad t > 0 \\ X(0) = X(0) \end{cases}$$

- $X(t)$, $H = \dot{H}^0 \times \dot{H}^{-1}$-valued stochastic process
- $W(t) = \sum_{k=1}^{\infty} \sqrt{\lambda_k}\beta_k(t)e_k$, Q-Wiener process on $U = \dot{H}^0$, $Qe_k = \lambda_k e_k$
- $\Lambda = -\Delta : \ \dot{H}^{-1} \to \dot{H}^{-1}, \quad D(\Lambda) = H_0^1(\mathcal{D}) = \dot{H}^1$
- $\begin{bmatrix} \mathrm{d}u \\ \mathrm{d}u_t \end{bmatrix} = \begin{bmatrix} 0 & I \\ -\Lambda & 0 \end{bmatrix} \begin{bmatrix} u \\ u_t \end{bmatrix} \mathrm{d}t + \begin{bmatrix} 0 \\ I \end{bmatrix} \mathrm{d}W,$

  $X = \begin{bmatrix} u \\ u_t \end{bmatrix}$, $A = -\begin{bmatrix} 0 & I \\ -\Lambda & 0 \end{bmatrix}$, $B = \begin{bmatrix} 0 \\ I \end{bmatrix}$
- $H = \dot{H}^0 \times \dot{H}^{-1}, \quad D(A) = \dot{H}^1 \times \dot{H}^0, \quad U = \dot{H}^0 = L_2(\mathcal{D})$
- $E(t) = e^{-tA} = \begin{bmatrix} \cos(t\Lambda^{1/2}) & \Lambda^{-1/2}\sin(t\Lambda^{1/2}) \\ -\Lambda^{1/2}\sin(t\Lambda^{1/2}) & \cos(t\Lambda^{1/2}) \end{bmatrix}$, C_0-semigroup

Applications: wave equation

Approximation

$$\begin{cases} \mathrm{d}X_h(t) + A_h X_h(t)\,\mathrm{d}t = B_h\,\mathrm{d}W(t), \quad t > 0 \\ X_h(0) = X_{0,h} \end{cases}$$

- Let $S_h \subset \dot{H}^1$ finite dimensional subspaces and set $V_h := S_h \times S_h$
- Λ_h"discrete Laplacian" defined by

$$\langle \Lambda_h \psi, \chi \rangle = \langle \nabla \psi, \nabla \chi \rangle, \quad \psi, \chi \in S_h$$

  Set

$$A_h := \begin{bmatrix} 0 & -I \\ \Lambda_h & 0 \end{bmatrix}$$

- Let

$$B_h := \begin{bmatrix} 0 \\ P_h \end{bmatrix}$$

  where P_h is the orthogonal projection $\dot{H}^0 \to S_h$

Applications: wave equation

- $$E_h(t) = e^{-tA_h} = \begin{bmatrix} \cos(t\Lambda_h^{1/2}) & \Lambda_h^{-1/2}\sin(t\Lambda_h^{1/2}) \\ -\Lambda_h^{1/2}\sin(t\Lambda_h^{1/2}) & \cos(t\Lambda_h^{1/2}) \end{bmatrix},$$

  C_0-semigroup on V_h

- $$X_{0,h} := \begin{bmatrix} P_h X_1(0) \\ P_h X_2(0) \end{bmatrix}$$

  (Note: P_h can be extended to $\dot{H}^{-1}$)

Applications: Wave equation

THEOREM [KLL'09]. Let $g \in C_{\mathrm{b}}^2(\dot{H}^0, \mathbb{R})$ and assume that $\|\Lambda^{\beta-\frac{1}{2}} Q \Lambda^{-\frac{1}{2}}\|_{\mathrm{Tr}} < \infty$ and that $X_0 \in L_1(\Omega, H^{2\beta})$ for some $\beta \in (0, \frac{r+1}{2}]$. Then, there are $C > 0$, $h_0 > 0$, depending on g, X_0, Q, β, and T but not on h, such that for $h \leq h_0$,

$$\Big|\mathbf{E}\big(g(X_{1,h}(T)) - g(X_1(T))\big)\Big| \leq C h^{\frac{2r}{r+1}\beta}.$$

The proof uses the weak error representation theorem with $G(X) := g(P_1 X)$, where P_1 is the canonical projection $H \to \dot{H}^0$. The relevant deterministic error estimates are

$$\|K_h(t)\| := \|\Lambda_h^{-1/2} \sin(t\Lambda_h^{1/2}) P_h v - \Lambda^{-1/2} \sin(t\Lambda^{1/2}) v\| \leq C(T) h^{\frac{r}{r+1}s} |v|_{s-1}, \quad t \in [0, T],\ s \in [0, r+1].$$

and

$$\|G_h(t)\| := \|\cos(t\Lambda_h^{1/2}) P_h - \cos(t\Lambda^{1/2}) v\| \leq C(T) h^{\frac{r}{r+1}s} |v|_s, \quad t \in [0, T],\ s \in [0, r+1].$$

Wave equation: sketch of proof

Set $G(X) := g(P_1 X)$, where P_1 is the canonical projection $H \to \dot{H}^0$. Then,

$$(u_x(Y(t),t),\Phi) = \mathbf{E}(\langle g'(P_1 Z(Y(t),t,T)), P_1\Phi\rangle|\mathcal{F}_t)$$

and

$$(u_{xx}(Y(t),t)\Phi,\Psi) = \mathbf{E}(\langle g''(P_1 Z(Y(t),t,T))P_1\Phi, P_1\Psi\rangle|\mathcal{F}_t).$$

Recall, the abstract weak error representation:

$$\begin{aligned} e_h(T) &= \mathbf{E}\left(u(Y_h(0),0) - u(Y(0),0)\right) \\ &\quad + \frac{1}{2}\mathbf{E}\int_0^T \mathrm{Tr}\Big(u_{xx}(Y_h(t),t) \\ &\qquad \times [E_h(T-t)B_h - E(T-t)B]Q[E_h(T-t)B_h + E(T-t)B]^*\Big)\,\mathrm{d}t. \end{aligned}$$

Wave equation: sketch of proof

The estimate for the first term is more or less straightforward from the deterministic error estimate and gives

$$|\mathbf{E}\,(u(Y_h(0),0) - u(Y(0),0))\,| \le C \sup_{x\in\dot{H}^0} \|g'(x)\| Ch^{\frac{2r}{r+1}\beta}\mathbf{E}|||X_0|||_{2\beta}.$$

For the second term, one can show, due to the special choice of G,

$$\begin{aligned}
&\mathbf{E}(\mathrm{Tr}(u_{xx}(Y_h(t),t)(E_h(T-t)B_h + E(T-t)B)Q(E_h(T-t)B_h - E(T-t)B)^*))\\
&= \mathbf{E}\,\mathrm{Tr}(K_h(T-t)Q(\Lambda_h^{-\frac{1}{2}}S_h(T-t)P_h + \Lambda^{-\frac{1}{2}}S(T-t))g''(P_1 Z(Y(t),t,T)))
\end{aligned}$$

Wave equation: sketch of proof

Therefore,

$$
\begin{aligned}
&\Big|\mathbf{E}\int_0^T \operatorname{Tr}\big(u_{xx}(Y_h(t),t)\\
&\qquad (E_h(T-t)B_h + E(T-t)B)Q(E_h(T-t)B_h - E(T-t)B)^*\big)\,\mathrm{d}t\Big|\\
&= \Big|\mathbf{E}\int_0^T \operatorname{Tr}(K_h(T-t)Q\\
&\qquad (\Lambda_h^{-\frac{1}{2}}S_h(T-t)P_h + \Lambda^{-\frac{1}{2}}S(T-t))g''(P_1Z(Y(t),t,T)))\,\mathrm{d}t\Big|\\
&= \Big|\mathbf{E}\int_0^T \operatorname{Tr}(K_h(T-t)\Lambda^{\frac{1}{2}-\beta}\Lambda^{\beta-\frac{1}{2}}Q\Lambda^{-\frac{1}{2}}\\
&\qquad \Lambda^{\frac{1}{2}}(\Lambda_h^{-\frac{1}{2}}S_h(T-t)P_h + \Lambda^{-\frac{1}{2}}S(T-t))g''(P_1Z(Y(t),t,T)))\,\mathrm{d}t\Big|\\
&\le CT\sup_{x\in H}\|g''(x)\|_{\mathcal{B}(\dot H^0)}\|\Lambda^{\beta-\frac{1}{2}}Q\Lambda^{-\frac{1}{2}}\|_{\mathrm{Tr}}\|K_h(T-t)\Lambda^{\frac{1}{2}-\beta}\|_{\mathcal{B}(\dot H^0)}\\
&\le CTh^{\frac{2r}{r+1}\beta}\sup_{x\in H}\|g''(x)\|_{\mathcal{B}(\dot H^0)}\|\Lambda^{\beta-\frac{1}{2}}Q\Lambda^{-\frac{1}{2}}\|_{\mathrm{Tr}}.
\end{aligned}
$$

Wave equation: remark

REMARK: The the strong rate $O(\frac{r}{r+1}\beta)$ is obtained under the assumption $\|\Lambda^{\frac{\beta-1}{2}}Q^{\frac{1}{2}}\|_{\mathrm{HS}} < \infty$. It can be shown that

$$\|\Lambda^{\frac{\beta-1}{2}}Q^{\frac{1}{2}}\|_{\mathrm{HS}}^2 \le \|\Lambda^{\beta-\frac{1}{2}}Q\Lambda^{-\frac{1}{2}}\|_{\mathrm{Tr}}$$

with equality when A and Q have a common basis of eigenfunctions ("commute"), in particular, when $Q = I$. If $Q = I$, then $d = 1$ and the weak rate is $O(h^{\frac{r}{r+1}})$.

References: finite elements for SPDEs with Gaussian noise

[DP'09] A. Debussche and J. Printems, *Weak order for the discretization of the stochastic heat equation*, Math. Comp. **78** (2009), 845–863.

[GKL'09] M. Geissert, M. Kovács and S. Larsson, *Rate of weak convergence of the finite element method for the stochastic heat equation with additive noise*, BIT **49** (2009), 343–356.

[KLL'09] M. Kovács, S. Larsson and F. Lindgren, *Weak finite element approximation of linear stochastic evolution equations with additive noise*, preprint.

[KLL'09a] M. Kovács, S. Larsson and F. Lindgren, *Strong convergence of the finite element method with truncated noise for semilinear parabolic stochastic equations with additive noise*, to appear in Numer. Alg.

[KLS'09] M. Kovács, S. Larsson and F. Saedpanah, *Finite element approximation of the linear stochastic wave equation with additive noise*, preprint.

[LM'09] S. Larsson and A. Mesforush, *Finite element approximation of the linear stochastic Cahn-Hilliard equation*, preprint.

[Yan'04] Y. Yan, *Semidiscrete Galerkin approximation for a linear stochastic parabolic partial differential equation driven by an additive noise*, BIT **44** (2004), 829–847.

[Yan'06] Y. Yan, *Galerkin finite element methods for stochastic parabolic partial differential equations*, SIAM J. Numer. Anal. **43** (2005), 1363–1384.

References: finite elements and semigroups and SPDEs

Finite elements for deterministic PDEs with the semigroup approach:

[Th'06] V. Thomée, *Galerkin Finite Element Methods for Parabolic Problems*, Second ed., Springer, 2006.

Operator semigroups:

[ABHN'01] W. Arendt, Ch. Batty, M. Hieber and F. Neubrander, *Vector-valued Laplace Transforms and Cauchy Problems*, Monographs in Mathematics **96**, Birkhäuser, 2001.
[EN'00] K. J. Engel and R. Nagel, *One-Parameter Semigroups for Linear Evolution Equations*, Springer, 2000.
[P'85] A. Pazy, Semigroups of linear operators and applications to partial differential equations. Applied Mathematical Sciences **44**, Springer, 1983.

SPDEs (semigroup approach):

[DZ'92] G. Da Prato and J. Zabczyk, *Stochastic Equations in Infinite Dimensions*, Cambridge University Press, Cambridge (1992).

Infinite dimensional Itô integral (and the variational approach):

[PR'07] C. Prévôt and M. Röckner, *A Concise Course on Stochastic Partial Differential Equations*, Springer, 2007.