On stochastic partial differential equations in Sobolev spaces

1. Foundation

We are dealing with the following general equation

$$du_t = [(1/2)a_t^{ij} D_{ij} u_t + b_t^i D_i u_t + c_t u_t + f_t]\, dt + [\sigma_t^{ik} D_i u_t + \nu_t^k u_t + g_t^k]\, dw_t^k, \quad (1.1)$$

where $a^{ij}, f, \sigma^{ik}, g^k$ are real-valued functions defined for $\omega \in \Omega$, $t \geq 0$, $x \in \mathbb{R}^d$, u is a W_p^2-valued function, $p \geq 2$, $i, j = 1, ..., d$. Apart from boundedness of a, b, c, σ, ν, Lipschitz continuity of σ, ν in x, and appropriate measurability of everything in (1.1), the main assumption is that the matrix

$$(a^{ij} - \alpha^{ij}), \quad \text{where} \quad \alpha^{ij} = \sigma^{ik}\sigma^{jk},$$

is uniformly nondegenerate.

The development of the L_p-theory started about 1994. Before that various aspects of the L_2-theory for general equations were investigated by Pardoux, Krylov, Rozovskii, Gyongy, Flandoli, Brzezniak, Da Prato, and few other researchers.

The interest in L_p-theory is easy to explain, for instance, from computational point of view. If we want to know how fast, say finite-difference approximating schemes will converge to the true solution of an SPDE, we need to know how smooth the true solution is. In the framework of the L_2-theory in order to guarantee that the solution has one continuous derivative in x, we need u to have $[d/2]+2$ generalized derivatives summable to the second power. This requires the coefficients to have $[d/2]$ derivatives in x. In the framework of the L_p-theory the true solution is continuously differentiable in x if the coefficients of the equation are merely continuous (σ, ν Lipschitz continuous) and $p>d$. Another advantage of having an L_p-theory for $p>2$ is that it allows one to get the modulus of continuity of individual trajectories of solutions, whereas for $p=2$ we have only the mean-square continuity.

Passing from the Hilbert space L_2 to L_p with $p\geq 2$ required proving one crucial estimate, which was done in

(Kr,94a) *A generalization of the Littlewood-Paley inequality and some other results related to stochastic partial differential equations*, Ulam Quarterly, Vol 2, No. 4 (1994) 16–26, http://www.ulam.usm.edu/VIEW2.4/krylov.ps

Recall that the space $W^2_p=W^2_p(\mathbb{R}^d)$, $p\in(1,\infty)$, is defined as the closure of $C_0^\infty=C_0^\infty(\mathbb{R}^d)$ with respect to the norm

$$\|u\|_{W^2_p}=\|u\|_{L_p}+\|Du\|_{L_p}+\|D^2u\|_{L_p}.$$

Consider the simplest one-dimensional SPDE

$$du_t(x) = \frac{1}{2}D^2 u_t(x)\,dt + g_t(x)\,dw_t \tag{1.2}$$

given for $t > 0$, $x \in \mathbb{R}^d$, with initial condition $u_0 = 0$, where w_t is a one-dimensional Wiener process. By Itô's formula

$$\|u_t\|_{L_2}^2 + \int_0^t \|Du_s\|_{L_2}^2\,ds = \int_0^t \|g_s\|_{L_2}^2\,ds + m_t,$$

where m_t is a martingale. Hence

$$E\|u_t\|_{L_2}^2 + E\int_0^t \|Du_s\|_{L_2}^2\,ds = E\int_0^t \|g_s\|_{L_2}^2\,ds,$$

and we learn that if g is an L_2-function, we can only control

$$E\int_0^t \|Du_s\|_{L_2}^2\,ds$$

and not, say

$$E\int_0^t \|D^2u_s\|_{L_2}^2\,ds.$$

This is a guideline for the L_p-theory.

The solution of (1.2) is known to be

$$u_t(x) = \int_0^t T_{t-s}g_s(x)\,dw_s, \tag{1.3}$$

where $T_t h(x) = Eh(x + w_t)$ is the heat semi-group. If g is non random, then $u_t(x)$ is a Gaussian random variable with zero mean and its absolute moments are just powers of its second moment. It follows that in this case

$$E\int_0^T ||Du_t||_{L_p}^p\,dt = N(p)\int_0^T\int_{\mathbb{R}^d}[\int_0^t |DT_{t-s}g_s(x)|^2\,ds]^{p/2}\,dx\,dt,$$

and in order to prove that $u \in W_p^1$ we have to estimate the last integral. If g is random, then by the Burkholder-Davis-Gundy inequalities

$$E\int_0^T ||Du_t||_{L_p}^p\,dt \le N(p)E\int_0^T\int_{\mathbb{R}^d}[\int_0^t |DT_{t-s}g_s(x)|^2\,ds]^{p/2}\,dx\,dt,$$

and if we can prove that for each ω

$$\int_0^\infty\int_{\mathbb{R}^d}[\int_0^t |DT_{t-s}g_s(x)|^2\,ds]^{p/2}\,dx\,dt \le N\int_0^\infty \|g_t\|_{L_p}^p\,dt, \tag{1.4}$$

then we would get

$$E\int_0^\infty ||Du_t||_{L_p}^p\,dt \le NE\int_0^\infty \|g_t\|_{L_p}^p\,dt,$$

which is a basic estimate in the L_p-theory of SPDEs.

For functions $f_t(x)$ given for all $t \in (-\infty, \infty)$ and $x \in \mathbb{R}^d$ introduce

$$\mathcal{G}f_t(x) = [\int_{-\infty}^t |DT_{t-s}f_s(x)|^2\,ds]^{1/2}.$$

Our first goal will be to prove the following theorem.

Theorem 1.1. *Let $p \in [2, \infty)$ and $f \in C_0^\infty(\mathbb{R}^{d+1})$. Then*

$$\int_{-\infty}^{\infty} \int_{\mathbb{R}^d} |\mathcal{G}_t f(x)|^p \, dx \, dt \le N \int_{-\infty}^{\infty} \int_{\mathbb{R}^d} |f_t(x)|^p \, dx \, dt. \tag{1.5}$$

where the constant N depends only on d, p.

It turns out that (1.5) is equivalent to (1.4) and for $p = 2$ we have an equality with $N = 1$.

It is also worth noting that from the proof it will be seen that similar estimate is valid in much more general situation, for instance, when T_{t-s} is replaced with $T_{s,t}$, where $T_{s,t}$ is an evolution operator associated with a parabolic equation with coefficients depending only on t in a measurable way. It seems that in this time-inhomogeneous setting it is imposible to prove the result by using semigroup approach. One also can replace DT_{t-s} with the convolution with a kernel depending on s, t and satisfying appropriate conditions, which allows one to consider Lévy processes.

1.1. **Partitions.** The proof of (1.5) is based on the Fefferman-Stein theorem which we discuss in the next subsection. In this section we remind some standard notions from probability theory related to martingales with the underlying measure which is not necessarily a probability measure.

Definition 1.2. Let $\mathbb{Z} = \{n : n = 0, \pm 1, \pm 2, ...\}$ and let $(\mathbb{Q}_n, n \in \mathbb{Z})$ be a sequence of partitions of $\mathbb{R}^d$ each consisting of disjoint bounded Borel subsets $Q \in \mathbb{Q}_n$. We call it *a filtration of partitions* if

(i) the partitions become finer as n increases:

$$\inf_{Q \in \mathbb{Q}_n} |Q| \to \infty \quad \text{as} \quad n \to -\infty, \qquad \sup_{Q \in \mathbb{Q}_n} \operatorname{diam} Q \to 0 \quad \text{as} \quad n \to \infty,$$

(ii) the partitions are nested: for each n and $Q \in \mathbb{Q}_n$ there is a (unique) $Q' \in \mathbb{Q}_{n-1}$ such that $Q \subset Q'$,

(iii) the following regularity property holds: for Q and Q' as in (ii) we have

$$|Q'| \leq N_0 |Q|,$$

where N_0 is a constant independent of n, Q, Q'.

Example 1.3. In the applications in these lectures we will be dealing with the filtration of parabolic dyadic cubes in

$$\mathbb{R}^{d+1} = \{(t, x) : t \in \mathbb{R}, x \in \mathbb{R}^d\},$$

defined by

$$\mathbb{Q}_n = \{Q_n(i_0, i_1, ..., i_d), i_0, i_1, ..., i_d \in \mathbb{Z}\},$$

$$Q_n(i_0, i_1, ..., i_d) = [i_0 4^{-n}, (i_0 + 1)4^{-n}) \times Q_n(i_1, ..., i_d), \tag{1.6}$$

$$Q_n(i_1, ..., i_d) = [i_1 2^{-n}, (i_1 + 1)2^{-n}) \times ... \times [i_d 2^{-n}, (i_d + 1)2^{-n}). \tag{1.7}$$

Definition 1.4. Let $\mathbb{Q}_n$, $n \in \mathbb{Z}$, be a filtration of partitions of $\mathbb{R}^d$.

(i) Let $\tau = \tau(x)$ be a function on $\mathbb{R}^d$ with values in $\{\infty, 0, \pm 1, \pm 2, ...\}$. We call τ *a stopping time* (relative to the filtration) if, for each $n \in \mathbb{Z}$, the set $\{x : \tau(x) = n\}$ is the union of some elements of $\mathbb{Q}_n$.

(ii) For any $x \in \mathbb{R}^d$ and $n \in \mathbb{Z}$, by $Q_n(x)$ we denote the (unique) $Q \in \mathbb{Q}_n$ containing x.

(iii) For a function $f \in L_{1,loc}$ and $n \in \mathbb{Z}$, we denote

$$f_{|n}(x) = \fint_{Q_n(x)} f(y)\,dy \qquad \Big(\fint_\Gamma f\,dx = \frac{1}{|\Gamma|} \int_\Gamma f\,dx, \quad |\Gamma| = \operatorname{Vol}\Gamma \Big).$$

If we are also given a stopping time τ, we let $f_{|\tau}(x) = f_{|\tau(x)}(x)$ for those x for which $\tau(x) < \infty$ and $f_{|\tau}(x) = f(x)$ otherwise.

We are going to use the following simple and well-known properties of the objects introduced above.

Lemma 1.5. *Let $\mathbb{Q}_n$, $n \in \mathbb{Z}$, be a filtration of partitions of $\mathbb{R}^d$.*

(i) Let $f \in L_{1,loc}$, $f \geq 0$, and let τ be a stopping time. Then

$$\int_{\mathbb{R}^d} f_{|\tau}(x) I_{\tau<\infty} \, dx = \int_{\mathbb{R}^d} f(x) I_{\tau<\infty} \, dx. \tag{1.8}$$

(ii) Let $g \in L_1$, $g \geq 0$, and $\lambda > 0$. Then

$$\tau(x) := \inf\{n : g_{|n}(x) > \lambda\} \quad (\inf \emptyset := \infty) \tag{1.9}$$

is a stopping time. Furthermore, we have

$$0 \leq g_{|\tau}(x) I_{\tau<\infty} \leq N_0 \lambda, \quad |\{x : \tau(x) < \infty\}| \leq \lambda^{-1} \int_{\mathbb{R}^d} g(x) I_{\tau<\infty} \, dx. \tag{1.10}$$

1.2. **Maximal and sharp functions. Fefferman-Stein theorem.** From Lemma 1.5 we derive the following.

Corollary 1.6 (maximal inequality)**.** *Let $f \in L_{1,loc}$. Define the (filtering) maximal function of f by*

$$Mf(x) = \sup_{n<\infty} (|f|)_{|n}(x),$$

so that $Mf = M|f|$. Then, for nonnegative $g \in L_1$, the maximal inequality holds:

$$|\{x : Mg(x) > \lambda\}| \leq \lambda^{-1} \int_{\mathbb{R}^d} g(x) I_{Mg>\lambda} \, dx, \quad \forall \lambda > 0. \tag{1.11}$$

Indeed, for τ as in (1.9), we have $\{x : Mg(x) > \lambda\} = \{x : \tau(x) < \infty\}$.

Remark 1.7. Our interest in estimating $|\{Mg > \lambda\}|$ as in Corollary 1.6 is based on the following formula valid for any $f \geq 0$

$$\int_{\mathbb{R}^d} f(x)\,dx = \int_0^\infty |\{x : f(x) > t\}|\,dt. \tag{1.12}$$

Corollary 1.8. *Let $p \in (1, \infty)$, $g \in L_1$, $g \geq 0$. Then*

$$\|Mg\|_{L_p} \leq q\|g\|_{L_p},$$

where $q = p/(p-1)$.

Indeed, from (1.12), (1.11), and Fubini's theorem we conclude that

$$\|Mg\|_{L_p}^p = \int_0^\infty |\{x : \ Mg(x) > \lambda^{1/p}\}|\,d\lambda \leq \int_{\mathbb{R}^d} g\big(\int_0^\infty \lambda^{-1/p} I_{Mg > \lambda^{1/p}}\,d\lambda\big)\,dx$$

$$= \int_{\mathbb{R}^d} g\big(\int_0^{(Mg)^p} \lambda^{-1/p}\,d\lambda\big)\,dx = q\int_{\mathbb{R}^d} g(Mg)^{p-1}\,dx.$$

Then upon using Hölder's inequality we get

$$\|Mg\|_{L_p}^p \leq q\|g\|_{L_p}\|Mg\|_{L_p}^{p-1}, \quad \|Mg\|_{L_p} \leq q\|g\|_{L_p}.$$

Theorem 1.9. *For any $p \in (1,\infty)$ and $g \in L_p$*

$$\|Mg\|_{L_p} \le q\|g\|_{L_p}.$$

Let $f \in L_{1,loc}$. Define the sharp function of f by

$$f^{\#}(x) = \sup_{n<\infty} ⨍_{Q_n(x)} |f(y) - f_{|n}(x)|\, dy.$$

If we think of $f_{|n}$ as $E(f|\mathbb{Q}_n)$, then

$$f^{\#} = \sup_{n<\infty} E(|f - E(f|\mathbb{Q}_n)|\,|\mathbb{Q}_n).$$

Observe that

$$|f(y) - f_{|n}(x)| = |⨍_{Q_n(x)} (f(y) - f(z))\, dz| \le ⨍_{Q_n(x)} |f(y) - f(z)|\, dz,$$

so that

$$f^{\#}(x) \le \sup_{n<\infty} ⨍_{Q_n(x)} ⨍_{Q_n(x)} |f(y) - f(z)|\, dz dy.$$

On the other hand,

$$|f(y) - f(z)| \le |f(y) - f_{|n}(x)| + |f(z) - f_{|n}(x)|$$

implying that

$$\sup_{n<\infty} \fint_{Q_n(x)} \fint_{Q_n(x)} |f(y) - f(z)| \, dz dy \leq 2 f^{\#}(x).$$

Since

$$|\, |f(y)| - |f(z)| \,| \leq |f(y) - f(z)|,$$

we also conclude that

$$(|f|)^{\#} \leq 2 f^{\#}.$$

Obviously $f^{\#}(x) \leq 2Mf(x)$. It turns out that f and hence Mf are also controlled by $f^{\#}$.

Lemma 1.10. *For $\alpha = (2N_0)^{-1} < 1$, any constant $c > 0$, and $f \in L_1$, we have*

$$|\{x : |f(x)| \geq c\}| \leq \frac{4}{c} \int_{\mathbb{R}^d} I_{Mf(x) > \alpha c} f^{\#}(x) \, dx.$$

Proof. First assume that $f \geq 0$ and set

$$\tau(x) = \inf\{n : f_{|n}(x) > c\alpha\}.$$

Use Lemma 1.5 (ii) and the fact that $f_{|n} \to f$ (a.e.) as $n \to \infty$. Then we find that (a.e.)

$$\{x : f(x) \geq c\} = \{x : f(x) \geq c, \tau(x) < \infty\}$$

$$= \{x : f(x) \geq c, f_{|\tau}(x) \leq c/2\} \subset \{x : |f(x) - f_{|\tau}(x)| \geq c/2\}.$$

By Chebyshev's inequality and Lemma 1.5

$$|\{x : |f(x)| \geq c\}| \leq (2/c) \int_{\mathbb{R}^d} |f(x) - f_{|\tau}(x)| \, dx$$

$$= (2/c) \int_{\mathbb{R}^d} |f - f_{|\tau}|_{|\tau}(x) \, dx = (2/c) \int_{\mathbb{R}^d} |f - f_{|\tau}|_{|\tau}(x) I_{\tau<\infty} \, dx,$$

where the last equality follows from the fact that $|f - f_{|\tau}|_{|\tau}(x) = |f - f_\tau|(x) = 0$ when $\tau(x) = \infty$.

Finally, if at an x we have $\tau(x) = n$, then

$$|f - f_{|\tau}|_{|\tau}(x) = \fint_{Q_n(x)} |f(y) - f_{|\tau}(y)| \, dy = \fint_{Q_n(x)} |f(y) - f_{|n}(y)| \, dy \leq f^{\#}(x).$$

We also notice that $\{\tau(x) < \infty\} = \{Mf(x) > c\alpha\}$ and conclude that for $f \geq 0$

$$|\{x : |f(x)| \geq c\}| \leq \frac{2}{c} \int_{\mathbb{R}^d} I_{Mf(x)>\alpha c} f^{\#}(x) \, dx.$$

For general f we need only add that $(|f|)^{\#} \leq 2f^{\#}$. The lemma is proved.

Theorem 1.11 (Fefferman-Stein). *Let $p \in (1, \infty)$. Then for any $f \in L_p$ we have*

$$\|f\|_{L_p} \leq N \|f^{\#}\|_{L_p},$$

where $N = (2q)^p N_0^{p-1}$.

Proof. As in the proof of Corollary 1.8 we get from Lemma 1.10 that

$$\|f\|_{L_p}^p \le N \int_{\mathbb{R}^d} f^{\#}(Mf)^{p-1}\,dx \le N\|f^{\#}\|_{L_p}\|Mf\|_{L_p}^{p-1}.$$

After that it only remains to use Theorem 1.9 and check that the resulting constant is right.

Remark 1.12. By Hölder's inequality, for any $p \in [1,\infty]$

$$f^{\#}(x) \le \sup_{n<\infty} \Big(\fint_{Q_n(x)} |f(y) - f_{|n}(y)|^p\,dy\Big)^{1/p}.$$

The maximal function introduced in Corollary 1.6 is related to the underlying filtration of partitions. Below we are also using the following more traditional maximal function:

$$\mathbb{M}g(x) = \sup_{r>0} \fint_{B_r(x)} |g(y)|\,dy, \tag{1.13}$$

where $B_r(x)$ is the open ball of radius r centered at x. Let Mg be the maximal function associated with the filtration of dyadic cubes Q_n introduced in (1.7). It turns out that, in a sense, Mg and $\mathbb{M}g$ are comparable.

First, since $Q_n(x) \subset B_{r_n}(x)$ with $r_n = 2^{-n}\sqrt{d}$, we have $|B_{r_n}(x)| = N(d)|Q_n(x)|$,

$$\fint_{Q_n(x)} |g|\,dy \le \frac{|B_{r_n}(x)|}{|Q_n(x)|} \fint_{B_{r_n}(x)} |g|\,dy \le N(d)\mathbb{M}g(x),$$

and $Mg \le N\mathbb{M}g$.

On the other hand, we have the following.

Lemma 1.13. *There is a constant* $N = N(d)$ *such that if* $g \in L_1$, *then for any* $\lambda > 0$

$$|\{x : \mathbb{M}g(x) > N\lambda\}| \le N|\{x : Mg(x) > \lambda\}|. \tag{1.14}$$

Proof. Without losing generality we may assume that $g \ge 0$. If x_0 and $\lambda > 0$ are such that $\mathbb{M}g(x_0) > \lambda$, then for an $r > 0$ we have

$$\int_{B_r(x_0)} g(y)\,dy > \lambda|B_r(x_0)|. \tag{1.15}$$

Set

$$n = -[\log_2 r], \quad i_k = [x^k 2^n], \quad \bar{x}_0 = (i_1 2^{-n}, ..., i_d 2^{-n}).$$

Then $2^{-n} \le r < 2^{-n+1}$, $|x_0^k - \bar{x}_0^k| < 2^{-n} \le r$ and $B_r(x_0)$ is covered by the union of 2^d dyadic cubes each of which has $\bar{x}_0$ as one of its vertices and they are taken from the family $\mathbb{Q}_{n-2}$. Owing to (1.15) the integral of g over at least one of these cubes Q is greater than

$$\lambda 2^{-d}|B_r(x_0)| = N\lambda r^d \ge N_*\lambda|Q|.$$

Furthermore, it is not hard to see that $x_0 \in 2Q$, where by $2Q$ we mean the twice dilated Q with the center of dilation being that of Q.

Now define $\tau(x) = \inf\{n : g_{|n}(x) > N_*\lambda\}$. Then $\tau \le n-2$ on $Q \in \mathbb{Q}_{n-2}$. Actually, it may happen that $\tau = m < n-2$ on Q. In that case $Q \subset Q' \in \mathbb{Q}_m$ and $\tau = m$ on Q'. Since $x_0 \in 2Q'$ we conclude that x_0 is in the union over j and m of twice dilated dyadic cubes Q_{jm} from the family $\mathbb{Q}_m$ composing $\{x : \tau(x) = m\}$. Hence,

$$|\{x : \mathbb{M}g(x) > \lambda\}| \le 2^d \sum_m \sum_j |Q_{jm}|$$

$$= 2^d |\{x : \tau(x) < \infty\}| = 2^d |\{x : Mg(x) > N_*\lambda\}|.$$

This proves the lemma.

Here is the classical maximal function estimate.

Theorem 1.14. *Let $p \in (1,\infty)$ and $g \in L_p$. Then $\mathbb{M}g \in L_p$ and*

$$\|\mathbb{M}g\|_{L_p} \le N \|g\|_{L_p}, \tag{1.16}$$

where N is independent of g.

Proof. Without losing generality we assume that $g \geq 0$. If $g \in L_1$, then (1.16) is obtained by replacing λ with $\lambda^{1/p}$ in (1.14), integrating with respect to λ, remembering (1.12), and using Corollary 1.8.

1.3. **Proof of Theorem 1.1.** First I remind what Theorem 1.1 is about. We define

$$T_t g(x) = \int_{\mathbb{R}^d} t^{-d/2} p\big(\frac{|x-y|}{t^{1/2}}\big) g(y)\, dy, \quad p(z) = (2\pi)^{-d/2} e^{-z^2/2},$$

$$\mathcal{G} f_t(x) = [\int_{-\infty}^{t} |DT_{t-s} f_s(x)|^2\, ds]^{p/2}.$$

Our assertion is that if $p \in [2, \infty)$ and $f \in C_0^\infty(\mathbb{R}^{d+1})$, then

$$\int_{-\infty}^{\infty} \int_{\mathbb{R}^d} |\mathcal{G}_t f(x)|^p\, dx\, dt \leq N \int_{-\infty}^{\infty} \int_{\mathbb{R}^d} |f_t(x)|^p\, dx\, dt,$$

where N is independent of f. In other words, for $f \in C_0^\infty(\mathbb{R}^{d+1})$ and

$$u = \mathcal{G} f$$

we have to prove that

$$\|u\|_{L_p(\mathbb{R}^{d+1})} \leq N(d,p) \|f\|_{L_p(\mathbb{R}^{d+1})}. \tag{1.17}$$

First we state three auxiliary results which we prove in Subsection 1.4.

Lemma 1.15. *We have*

$$\|u\|_{L_2(\mathbb{R}^{d+1})} = N(d)\|f\|_{L_2(\mathbb{R}^{d+1})}.$$

Next, according to (1.13) introduce the maximal function of a real-valued function h given on $\mathbb{R}^d$ relative to balls. We denote this function $\mathbb{M}_x h$ to emphasize that this maximal function is taken with respect to x. Similarly, for functions h on $\mathbb{R}$ we introduce $\mathbb{M}_t h$ as the maximal function of h relative to symmetric intervals:

$$\mathbb{M}_t h(t) = \sup_{r>0} \frac{1}{2r} \int_{-r}^{r} |h(t+r)| \, dr.$$

For a function $h(t,x)$ set

$$\mathbb{M}_x h(t,x) = \mathbb{M}_x(h(t,\cdot))(x), \quad \mathbb{M}_t h(t,x) = \mathbb{M}_t(h(\cdot,x))(t).$$

Lemma 1.16. *Set*

$$Q_0 = [-4,0] \times [-1,1]^d \tag{1.18}$$

and assume that $f_t = 0$ *for* $t \notin (-12,12)$. *Then for any* $(t,x) \in Q_0$

$$\int_{Q_0} |u_s(y)|^2 \, dsdy \le N \mathbb{M}_t \mathbb{M}_x |f|^2(t,x), \tag{1.19}$$

where N *depends only on* d.

Lemma 1.17. *Assume that $f_t(x) = 0$ for $t \geq -8$. Then for any $(t,x) \in Q_0$*

$$\int_{Q_0} |u_s(y) - u_t(x)|^2 \, dsdy \leq N \mathbb{M}_t \mathbb{M}_x |f|^2(t,x), \tag{1.20}$$

where the constant N depends only on d.

Now we start proving Theorem 1.1. Equation (1.17) follows from Lemma 1.15 if $p = 2$. Hence we may concentrate on $p > 2$. We start considering this case by claiming that at each point in $\mathbb{R}^{d+1}$

$$(\mathcal{G}f)^{\#} \leq N(d)(\mathbb{M}_t \mathbb{M}_x |f|^2)^{1/2}, \tag{1.21}$$

where the sharp function $(\mathcal{G}f)^{\#}$ is defined relative to the parabolic dyadic cubes of type (1.6).

Remark 1.12 shows that to prove (1.21) it suffices to prove that for each $Q = Q_n(i_0, ..., i_d)$ (see (1.6)) and $(t,x) \in Q$

$$⨍_Q |\mathcal{G}f - (\mathcal{G}f)_Q|^2 \, dyds \leq N(d) \mathbb{M}_t \mathbb{M}_x |f|^2(t,x), \tag{1.22}$$

where

$$(\mathcal{G}f)_Q = ⨍_Q \mathcal{G}f \, dyds.$$

To prove (1.22), observe that if a constant $c \neq 0$, then

$$\mathcal{G}[f_{c^2\cdot}(c\,\cdot)](t,x) = \mathcal{G}f(c^2t, cx).$$

This and the fact that dilations do not affect averages show that it suffices to prove (1.22) for $Q = Q_{-1}(i_0, ..., i_d)$. In that case Q is just a shift of Q_0 from (1.18). Furthermore, the shift is harmless since $\mathbb{M}_x$ and $\mathbb{M}_t$ are defined in terms of balls rather than dyadic cubes.

Thus let $Q = Q_0$ and take a function $\zeta \in C_0^\infty(\mathbb{R})$ such that $\zeta(t) = 1$ on $[-8, 8]$, $\zeta(t) = 0$ outside of $[-12, 12]$, and $1 \geq \zeta \geq 0$. Set

$$\alpha = f\zeta, \quad \beta = f - \alpha.$$

Observe that

$$\mathcal{G}f \leq \mathcal{G}\alpha + \mathcal{G}\beta, \quad \mathcal{G}\beta \leq \mathcal{G}f.$$

It follows that for any constant c

$$|\mathcal{G}f - c| \leq |\mathcal{G}\alpha| + |\mathcal{G}\beta - c|$$

and the left-hand side of (1.22) is less than

$$2 \fint_{Q} |\mathcal{G}\alpha|^2 \, dyds + 2 \fint_{Q} |\mathcal{G}\beta - c|^2 \, dyds.$$

We finally take $c = (\mathcal{G}\beta)_Q$ and obtain (1.22) from Lemmas 1.20 and 1.21.

After having proved (1.21), by combining the Fefferman-Stein theorem with the L_q, $q > 1$, boundedness of the maximal operators we conclude (recall that $p > 2$)

$$\|u\|^p_{L_p(\mathbb{R}^{d+1})} \leq N \|(\mathbb{M}_t \mathbb{M}_x |f|^2)^{1/2}\|^p_{L_p(\mathbb{R}^{d+1})}$$

$$= N \int_{\mathbb{R}^d} \big(\int_{\mathbb{R}} (\mathbb{M}_t \mathbb{M}_x |f|^2)^{p/2} \, dt \big) \, dx \leq N \int_{\mathbb{R}^d} \big(\int_{\mathbb{R}} (\mathbb{M}_x |f|^2)^{p/2} \, dt \big) \, dx$$

$$= N \int_{\mathbb{R}} \big(\int_{\mathbb{R}^d} (\mathbb{M}_x |f|^2)^{p/2} \, dx \big) \, dt \leq N \|f\|^p_{L_p(\mathbb{R}^{d+1})}.$$

This proves the theorem.

$1.4.$ **Auxiliary estimates on $\mathcal{G}$.** Throughout the subsection f is a fixed element of $C_0^\infty(\mathbb{R}^{d+1})$ and

$$u_t(x) = \mathcal{G} f_t(x) = \Big[\int_{-\infty}^{t} |DT_{t-s} f_s(x)|^2 \, ds\Big]^{1/2}.$$

First we improve Lemma 1.15

Lemma 1.18. *For any* $T \in (-\infty, \infty]$

$$\|u\|_{L_2(\mathbb{R}^{d+1}\cap\{t\le T\})} \le N(d,K)\|f\|_{L_2(\mathbb{R}^{d+1}\cap\{t\le T\})}.$$

To proceed further we need some notation. According to (1.13) introduce the maximal function of a real-valued function h given on $\mathbb{R}^d$ relative to balls. We denote this function $\mathbb{M}_x h$ to emphasize that this maximal function is taken with respect to x. Similarly, for functions h on $\mathbb{R}$ we introduce $\mathbb{M}_t h$ as the maximal function of h relative to symmetric intervals:

$$\mathbb{M}_t h(t) = \sup_{r>0} \frac{1}{2r} \int_{-r}^{r} |h(t+r)|\, dr.$$

For a function $h(t,x)$ set

$$\mathbb{M}_x h(t,x) = \mathbb{M}_x(h(t,\cdot))(x), \quad \mathbb{M}_t h(t,x) = \mathbb{M}_t(h(\cdot,x))(t).$$

Notice the following consequence of Lemma 1.15, in which and below we denote by $B_r(x)$ the open ball of radius r centered at x and $B_r = B_r(0)$.

Corollary 1.19. *Set*

$$Q_0 = [-4,0] \times [-1,1]^d \tag{1.23}$$

and assume that $f=0$ outside of $[-12,12]\times B_{3d}$. Then for any $(t,x)\in Q_0$

$$\int_{Q_0}|u_s(y)|^2\,dsdy\le N\mathbb{M}_t\mathbb{M}_x|f|^2(t,x),\tag{1.24}$$

where N depend only on d.

Indeed, for $g:=|f|^2$ the left-hand side is less than

$$N\int_{\mathbb{R}^{d+1},s\le 0}g\,dyds\le N\int_{-12}^{0}\int_{|y|\le 3d}g\,dyds$$

$$\le N\int_{-12}^{0}\int_{|x-y|\le 4d}g\,dyds\le N\int_{-12}^{0}\mathbb{M}_xg(s,x)\,ds\le N\mathbb{M}_t\mathbb{M}_xg(t,x).$$

Here is a generalization of Corollary 1.19.

Lemma 1.20. *Assume that $f(t,x)=0$ for $t\notin(-12,12)$. Then (1.24) holds again for any $(t,x)\in Q_0$.*

Proof. We take a $\zeta\in C_0^\infty(\mathbb{R}^d)$ such that $\zeta=1$ in B_{2d} and $\zeta=0$ outside of B_{3d}. Set $\alpha=\zeta f$ and $\beta=(1-\zeta)f$. Since $\mathcal{G}f\le\mathcal{G}\alpha+\mathcal{G}\beta$ and $\mathcal{G}\alpha$ admits the stated estimate, it suffices to concentrate on $\mathcal{G}\beta$. In other words, in the rest of the proof we may assume that $f_t(x)=0$ for $x\in B_{2d}$.

It turns out that in this case for $0 > s > r > -12$ and $x, y \in [-1,1]^d$ we have

$$|DT_{s-r}f_r(y)|^2 \leq N\mathbb{M}_x f_r^2(x).$$

To show this observe that

$$T_t g(y) = \int_{\mathbb{R}^d} t^{-d/2} p\big(\frac{|y-z|}{t^{1/2}}\big) g(z)\,dz, \quad p(z) = (2\pi)^{-d/2} e^{-z^2/2},$$

$$|Dp\big(\frac{|y-z|}{t^{1/2}}\big)| \leq N t^{-1/2} q\big(\frac{|y-z|}{t^{1/2}}\big), \quad q = p^{1/2},$$

And for $|z| \geq 2d$ ($x, y \in [-1,1]^d$) we have

$$|y-z| \geq |z| - d \geq (1/2)|z|, \quad |z| \geq |z-x| - d \geq |z-x| - (1/2)|z|,$$

$$|y-z| \geq (1/3)|x-z|, \quad q\big(\frac{|y-z|}{t^{1/2}}\big) \leq \kappa\big(\frac{|x-z|}{t^{1/2}}\big), \quad \kappa = p^{1/6}.$$

Hence,

$$|DT_{s-r}f_r(y)| \leq N(s-r)^{-(d+1)/2} \int_{\mathbb{R}^d} \kappa\big(\frac{|x-z|}{(s-r)^{1/2}}\big)|f_r(z)|\,dy$$
$$= N(s-r)^{-(d+1)/2} \int_{\mathbb{R}^d} \kappa\big(\frac{|z|}{(s-r)^{1/2}}\big)|f_r(x-z)|\,dy.$$

We transform the last integral by using the formula

$$\int_{\mathbb{R}^d} F(z) Q(|z|)\,dz = -\int_\varepsilon^\infty Q'(\rho)\Big(\int_{|z|\leq\rho} F(z)\,dz\Big)\,d\rho \tag{1.25}$$

where F and G satisfy appropriate conditions. This yields

$$|DT_{s-r}f_r(y)|$$

$$\leq (s-r)^{-(d+2)/2}\int_0^\infty |\kappa'|(\rho/\sqrt{s-r})|\Big(\int_{|z|\leq\rho}|f_r(x-z)|\,dz\Big)\,d\rho.$$

$$\leq N\mathbb{M}_x f_r(x)(s-r)^{-(d+2)/2}\int_1^\infty |\kappa'|(\rho/\sqrt{s-r})\rho^d\,d\rho$$

$$= N\mathbb{M}_x f_r(x)(s-r)^{-1/2}\int_{(s-r)^{-1/2}}^\infty |\kappa'|(\rho)\rho^d\,d\rho \leq N\mathbb{M}_x f_r(x).$$

Also observe that by Hölder's inequality $(\mathbb{M}_x f_r)^2 \leq \mathbb{M}_x f_r^2$. Then for $(s,y)\in Q_0$ we obtain

$$|u_s(y)|^2 = \int_{-12}^s |DT_{s-r}f_r(y)|^2\,dr \leq N\int_{-12}^0 \mathbb{M}_x|f_r|^2(x)\,dr,$$

where the last expression is certainly less than the right-hand side of (1.24). The lemma is proved.

Lemma 1.21. *Assume that $f(t,x)=0$ for $t\geq -8$. Then for any $(t,x)\in Q_0$*

$$\int_{Q_0}|u_s(y)-u_t(x)|^2\,dsdy \leq N\mathbb{M}_t\mathbb{M}_x|f|^2(t,x), \tag{1.26}$$

where the constant N depends only on K and d.

Proof. The left-hand side of (1.26) is certainly less than a constant times

$$\sup_{(s,y)\in Q_0}[|D_s u(s,y)|^2+|Du(s,y)|^2]. \tag{1.27}$$

Fix $(s,y)\in Q_0$ and note that $s\geq -4$ and by Minkowski's inequality (the derivative of a norm is less then the norm of the derivative)

$$|Du_s(y)|^2\leq \int_{-\infty}^{-8} I^2(r,s,y)\,dr.$$

where

$$I(r,s,y):=|D^2T_{s-r}f_r(y)|$$

$$=(s-r)^{-(d+2)/2}|\int_{\mathbb{R}^d}|(D^2p)(z/\sqrt{s-r})f_r(y-z)|\,dz$$

$$\leq (s-r)^{-(d+2)/2}\int_{\mathbb{R}^d}q(|z|/\sqrt{s-r})|f_r(y-z)|\,dz.$$

Also use again (1.25). Then we see that for $s>r$

$$I(r,s,y)\leq (s-r)^{-(d+3)/2}\int_0^\infty |q'|(\rho/\sqrt{s-r})\big(\int_{B_\rho(y)}|f_r(z)|\,dz\big)\,d\rho.$$

Since for a constant $\nu>0$ we have $B_\rho(y)\subset B_{\nu+\rho}(x)$ ($|x|,|y|\leq d$), we obtain

$$I(r,s,y)\leq N\mathbb{M}_xf_r(x)(s-r)^{-(d+3)/2}\int_0^\infty |q'|(\rho/\sqrt{s-r})(\nu+\rho)^d\,d\rho$$

$$=N\mathbb{M}_xf_r(x)(s-r)^{-1}\int_0^\infty |q'(\rho)|(\nu/\sqrt{s-r}+\rho)^d\,d\rho.$$

For $r\leq -8$ we have $s-r\geq 4$ and we conclude

$$\int_0^\infty |q'(\rho)|(\nu/\sqrt{s-r}+\rho)^d\,d\rho\leq N,\quad I(r,s,y)\leq N(s-r)^{-1}\mathbb{M}_x f_r(x),$$

$$|Du(s,y)|^2\leq N\int_{-\infty}^{-8}\mathbb{M}_x f_r^2(x)\frac{dr}{(4-r)^2}.$$

We transform the last integral integrating by parts or using (1.25) to find

$$|Du(s,y)|^2\leq N\int_{-\infty}^{-8}\frac{1}{(4-r)^3}\Big(\int_r^0\mathbb{M}_x f_p^2(x)\,dp\Big)\,dr$$

$$\leq N\mathbb{M}_t\mathbb{M}_x f^2(t,x)\int_{-\infty}^{-8}\frac{|r|}{(4-r)^3}\,dr=N\mathbb{M}_t\mathbb{M}_x f^2(t,x).$$

We thus have estimated part of (1.27).

The estimate of $D_s u$ is obtained similarly and the lemma is proved.

2. Theory in the whole space

Now we describe the function spaces in which (1.1):

$$du_t=[(1/2)a_t^{ij}D_{ij}u_t+b_t^iD_iu_t+c_tu_t+f_t]\,dt+[\sigma_t^{ik}D_iu_t+\nu_t^ku_t+g_t^k]\,dw_t^k,\qquad(2.1)$$

is treated.

For a stopping time τ and integers $\gamma=0,1,...$ we denote

$$\mathbb{H}_p^\gamma(\tau)=L_p((0,\tau]],\mathcal{P},W_p^\gamma),\quad\mathbb{H}_p^\gamma=\mathbb{H}_p^\gamma(\infty),\quad\mathbb{L}_{...}...=\mathbb{H}^0_{...}....$$

For distributions u_t, f_t, g^k_t, depending on t and ω and measurable in a natural sense, we write

$$du_t = f_t\,dt + g^k_t\,dw^k_t, \quad t > 0 \tag{2.2}$$

if for any $\phi \in C^\infty_0$ with probability 1 we have

$$(u_t,\phi) = (u_0,\phi) + \int_0^t (f_s,\phi)\,ds + \int_0^t (g^k_s,\phi)\,dw^k_s$$

for all $t \geq 0$ assuming that all integrals and the sum here make sense.

For $\gamma \geq 2$ and a function $u \in \mathbb{H}^\gamma_p(\tau)$ we write $u \in \mathcal{H}^\gamma_{p,0}(\tau)$ if there exists $f \in \mathbb{H}^{\gamma-2}_p(\tau)$ $g \in \mathbb{H}^{\gamma-1}_p(\tau)$ such that equality (2.2) holds in the sense of distributions. In this case we define

$$\|u\|_{\mathcal{H}^\gamma_p(\tau)} = \|u_{xx}\|_{\mathbb{H}^{\gamma-2}_p(\tau)} + \|f\|_{\mathbb{H}^{\gamma-2}_p(\tau)} + \|g\|_{\mathbb{H}^{\gamma-1}_p(\tau)}.$$

These spaces are convenient for treating equations with zero initial condition. The general case reduces to this one in a standard way.

Theorem 2.1. *Take a*

$$p \geq 2$$

and a bounded stopping time τ. Under the above condition of uniform nondegeneracy:

the matrix

$$(a^{ij}-\alpha^{ij}), \quad \textit{where} \quad \alpha^{ij}=\sigma^{ik}\sigma^{jk},$$

is uniformly nondegenerate,

also suppose that (a^{ij}) is uniformly continuous in x uniformly with respect to (ω,t) and the coefficients b,c,σ,ν are bounded and σ,ν satisfy the Lipschitz condition in x uniformly with respect to t,ω. Finally, assume that in (2.1):

$$du_t=[(1/2)a_t^{ij}D_{ij}u_t+b_t^iD_iu_t+c_tu_t+f_t]\,dt+[\sigma_t^{ik}D_iu_t+\nu_t^ku_t+g_t^k]\,dw_t^k,$$

the summation in k is restricted to a finite set. Then for any $f\in\mathbb{L}_p(\tau)$ and $g\in\mathbb{H}_p^1(\tau)$ there exists a unique $u\in\mathcal{H}_{p,0}^2(\tau)$ satisfying (2.1). Furthermore,

$$\|u\|_{\mathbb{H}_p^2(\tau)}\le N(\|f\|_{\mathbb{L}_p(\tau)}+\|g\|_{\mathbb{H}_p^1(\tau)}),$$

where N is independent of f and g.

There are a few extensions of this result when the right-hand sides f,g belong to spaces $\mathbb{H}$ and are nonlinear functions of u.

By the way, neither Theorem 2.1 nor its natural extension to nonzero initial-value problem are applicable to

$$du_t = D^2 u_t \, dt + u_+^\lambda \phi^k \, dw_t^k, \tag{2.3}$$

since the range of the summation in k is $1, 2, ...$ Actually, the solutions of this equation never belong to $\mathcal{H}_p^2(\tau)$ but rather to $\mathcal{H}_p^\gamma(\tau)$, where γ is any number $< 1/2$. In particular, this shows the necessity to treat the equations in the spaces $\mathcal{H}_p^\gamma(\tau)$ with arbitrary $\gamma \in \mathbb{R}$. It is also worth noting that quite often in the literature after work by J. Walsh in 1986 the second term on the right in (2.3) is written by using the so-called space-time white noise. The form (2.3) was probably first introduced by T. Funaki in 1983 and turns out more convenoent in many respects.

The proof of this theorem goes the following way. Usual perturbation argument allows us to assume that

$$b = c = \nu = 0.$$

Partitions of unity, another standard and almost trivial technique, reduces further the situation to the one that a and σ depend only on (ω, t). To explain better two more ideas we assume that

$$d = 1$$

and there is only one Wiener process involved, so that we are dealing with the equation

$$du_t = ((1/2)a_t D^2 u_t + f_t)\,dt + (\sigma_t D u_t + g_t)\,dw_t. \tag{2.4}$$

Introduce

$$v_t(x) = u_t(x - \xi_t), \quad \xi_t := \int_0^t \sigma_s\,dw_s.$$

Then by Itô-Wentzell formula equation (2.4) becomes

$$dv_t = ((1/2)\bar{a}_t D^2 v_t + \bar{f}_t)\,dt + \bar{g}_t\,dw_t, \tag{2.5}$$

where

$$\boxed{\bar{a}_t = a_t - \sigma_t^2}, \quad \bar{f}_t(x) = f_t(x - \xi_t) - \sigma_t D g_t(x - \xi_t), \quad \bar{g}_t = g_t(x - \xi_t)$$

(observe that the first derivative of g now enters $\bar{f}$ and that we need $\bar{a}_t > 0$ even if $\bar{g} = 0$).

Our nondegeneracy assumption is that $\bar{a}_t \geq \delta > 0$, where δ is a constant. Now imagine that

$$\bar{a} = \text{const}.$$

Then changing time shows that we may assume that $\bar{a} = 1$, in which case we first find a solution $y \in \mathbb{H}_p^1$ of the equation

$$dy_t = (1/2)\bar{a}D^2 y_t \, dt + \bar{g}_t \, dw_t \tag{2.6}$$

with zero initial condition. Recall that $\bar{a} = 1$ so that by the first part of the lectures

$$y_t = \int_0^t T_{t-s}\bar{g}_s \, dw_s$$

and

$$\|Dy\|_{\mathbb{L}_p} \leq N(d,p)\|\bar{g}\|_{\mathbb{L}_p} = N(d,p)\|g\|_{\mathbb{L}_p}. \tag{2.7}$$

One can differentiate through (2.6) which leads to

$$dDy_t = (1/2)\bar{a}D^2(Dy_t) \, dt + D\bar{g}_t \, dw_t$$

and similarly to (2.7)

$$\|D^2 y\|_{\mathbb{L}_p} \leq N(d,p)\|D\bar{g}\|_{\mathbb{L}_p} = N(d,p)\|Dg\|_{\mathbb{L}_p}.$$

Also by classical results (or by proceeding very similarly to what we have done in the first part) there is a $z \in \mathbb{H}_p^2$ satisfying

$$dz_t = ((1/2)\bar{a}D^2 z_t + \bar{f}_t) \, dt$$

with zero initial condition and

$$\|D^2 z\|_{\mathbb{L}_p} \leq N(d,p)\|\bar{f}\|_{\mathbb{L}_p} \leq N(d,p)\|g\|_{\mathbb{L}_p} + N(d,p)\|Dg\|_{\mathbb{L}_p}.$$

Since obviously $v = w + z$ we obtain the desired result:

$$\|u\|_{\mathbb{H}^2_p(\tau)} \leq N(\|f\|_{\mathbb{L}_p(\tau)} + \|g\|_{\mathbb{H}^1_p(\tau)}),$$

if $\bar{a}$ is a constant.

It turns out that the general case of nonconstant $\bar{a}$ reduces to the particular one. To show how to do this, istead of (2.4) consider the equation

$$d\hat{u}_t = ((1/2)a_t D^2 \hat{u}_t + f_t)\,dt + (\sigma_t D\hat{u}_t + g_t)\,dw_t + \hat{\sigma}_t D\hat{u}_t\,dB_t, \tag{2.8}$$

where B_t is a Wiener process independent of w_t and $\hat{\sigma}_t$ is such that

$$\bar{a}_t := a_t - \sigma_t^2 - \hat{\sigma}_t^2 = \text{const} > 0.$$

Then by introducing v as above: $v_t(x) = u_t(x - \xi_t)$, with

$$\xi_t = \int_0^t \sigma_s\,dw_s + \int_0^t \hat{\sigma}_s\,dB_s$$

we again come to (2.5):

$$dv_t = ((1/2)\bar{a}_t D^2 v_t + \bar{f}_t)\,dt + \bar{g}_t\,dw_t,$$

with constant $\bar{a}$ the same $\bar{g}_t(x) = g_t(x - \xi_t)$ and

$$\bar{f}_t(x) = f_t(x - \xi_t) - (\sigma_t + \hat{\sigma}_t)Dg_t(x - \xi_t).$$

By the above

$$\|D^2\hat{u}\|_{\mathbb{L}_p} = \|D^2 v\|_{\mathbb{L}_p} \leq N(\|f\|_{\mathbb{L}_p} + \|Dg\|_{\mathbb{L}_p}).$$

Now it only remains to observe that almost obviously

$$u_t = E(\hat{u}_t | w_s, s \leq t)$$

implying that

$$\|D^2 u\|_{\mathbb{L}_p} \leq \|D^2\hat{u}\|_{\mathbb{L}_p} \leq N(\|f\|_{\mathbb{L}_p} + \|Dg\|_{\mathbb{L}_p}),$$

which is the desired result.

Remark 2.2. The last part of the above argument allows one to get a rather unusual information about ordinary parabolic equations with coefficients depending only on t. By the way, the information is stated in terms of PDE but I do not know of any proof of this result, which would not involve probability theory.

For simplicity again assume that $d = 1$ and concider the following heat equation

$$\frac{\partial}{\partial t} u_t(x) = (1/2)D^2 u_t(x) + f_t(x),$$

with zero initial condition. Assume that for infinitely differentiable solutions, say with support bounded in x, of this equation we have an estimate

$$\|u\|_U \leq N_0 \|f\|_F, \tag{2.9}$$

where U and F are some Banach spaces of functions on $(0,T) \times \mathbb{R}$. The crucial assumption on U and F is that we suppose that the norms in these spaces are translation invariant with respect to shifts of the x coordinate by any bounded continuous function of t, that is if, say $u \in U$ and ξ_t is a (deterministic) continuous function of $t \in (0,T)$, then $v_t(x) := u_t(x - \xi_t)$ satisfies

$$v \in U, \quad \|v\|_U = \|u\|_U.$$

Next, take a bounded measurable function a_t depending only on t such that

$$a_t \geq 1$$

and consider the equation

$$\frac{\partial}{\partial t} u_t(x) = (1/2) a_t D^2 u_t(x) + f_t(x) \tag{2.10}$$

with zero initial condition. The claim is that for infinitely differentiable solutions with support bounded in x, of this equation estimate (2.9) holds with the same N_0, which, in particular, does not depend on the bound of a_t from above.

To prove this claim introduce $\sigma_t = (a_t - 1)^{1/2}$ and consider the SPDE

$$dv_t = (1/2)a_t D^2 v_t + f_t)\, dt + \sigma_t D v_t\, dw_t$$

with zero initial condition. "Obviously" for the solution u_t of (2.10) we have

$$u_t(x) = E v_t(x),$$

which by Minkovski's inequality implies that

$$\|u\|_U \le E\|v\|_U. \tag{2.11}$$

On the other hand, by Itô-Wentzell formula

$$z_t(x) := v_t(x - \xi_t), \quad \xi_t := \int_0^t \sigma_s\, dw_s$$

satisfies

$$\frac{\partial}{\partial t} z_t(x) = (1/2)D^2 z_t(x) + g_t(x),$$

where $g_t(x) = f_t(x - \xi_t)$. By the assumption, for each realization of ξ.

$$\|v\|_U = \|z\|_U \le N_0 \|g\|_F = N_0 \|f\|_F.$$

Hence,

$$E\|v\|_U \le N_0 \|f\|_F$$

and by combining this with (2.11) we get the result.

3. SPDEs in half-spaces

It turns out that the theory of SPDEs in domains is much harder than in the whole space. For quite some time the only available results were obtained in function spaces with low regularity or under some compatibility conditions on the data. We refer to works by Pardoux, Flandoli, Brzezniak, Da Prato, Zabczyk, and others.

Let us come back to the simplest equation (1.2):

$$du_t(x) = \frac{1}{2} D^2 u_t(x)\, dt + g_t(x)\, dw_t$$

again with zero initial condition but considered only in the half-line

$$\mathbb{R}_+ = (0, \infty)$$

and with zero lateral condition. The solution is given by the same formula (1.3) but with T_t defined as the heat semigroup in $\mathbb{R}_+$ with zero lateral condition. In particular, when $g \equiv 1$

$$u_t(x) = \int_0^t T_{t-s}1(x)\,dw_s.$$

It is very easy to prove that $u_t(x)$ is infinitely differentiable with respect to x in $\mathbb{R}_+$. However, its second order derivative cannot be bounded near the origin. To understand that write

$$u_t(x) = (1/2)\int_0^t D^2 u_s(x)\,ds + w_t$$

at $x = 0$. Then the left-hand side vanishes owing to the boundary condition and we see that w_t is equal to a function which is differentiable in t, which is known to be false.

This is the reason why we needed to introduce function spaces where elements are allowed to have derivatives blowing up near the boundary.

In the general theory we obtained solutions of class $\mathfrak{H}^\gamma_{p,\theta,0}(\tau)$, where γ is a number indicating the number of derivatives the functions of this class possess, the power of summability $p \geq 2$, and θ is the parameter controlling the blow up near the boundary.

The spaces $\mathfrak{H}^\gamma_{p,\theta,0}(\tau)$ are introduced for all $\gamma \in \mathbb{R}$ and any domain. However, it is easier to explain what they are when $\gamma = 2$ and the domain the equation is considered in is

$$\mathbb{R}^d_+ = \{x \in \mathbb{R}^d : x^1 > 0\}.$$

We say that $u \in \mathfrak{H}^2_{p,\theta,0}(\tau)$ if (2.2):

$$du_t = f_t\,dt + g^k_t\,dw^k_t, \;\; t > 0$$

holds in the sense of distributions on $\mathbb{R}^d_+$, $u_0 = 0$ and u, f, and g are such that

$$E\int_0^\tau \int_{\mathbb{R}^d_+} (x^1)^{\theta-d}(|(x^1)^{-1}u_t(x)|^p + |Du_t(x)|^p + |x^1 D^2 u_t(x)|^p\,dxdt < \infty,$$

$$E\int_0^\tau \int_{\mathbb{R}^d_+} (x^1)^{\theta-d}(|x^1 f_t(x)|^p + |g_t(x)|^p + |x^1 Dg_t(x)|^p)\,dxdt < \infty.$$

For the same reasons as above we needed better results than those for $p = 2$. One more drawback of the Hilbert space theory is that, in the general case, no matter to which class $\mathfrak{H}^\gamma_{2,\theta}(\tau)$ a function u belongs with as large γ as you wish one cannot conclude that

$$u_t(x) \to 0 \quad \text{as} \quad x \downarrow 0, \tag{3.1}$$

that is that the boundary condition is satisfied pointwise. One needs this kind of global continuity of solutions, for instance, for numerical approximations. Proving (3.1)

became a major challenge for the theory. The current state of the art is that (3.1) is finally established in my paper

(Kr,07) *Maximum principle for SPDEs and its applications*, pp. 311-338 in "Stochastic Differential Equations: Theory and Applications, A Volume in Honor of Professor Boris L. Rozovskii", P.H. Baxendale, S.V. Lototsky eds., Interdisciplinary Mathematical Sciences, Vol. 2, World Scientific, 2007.

for the one-dimensional case assuming that the coefficients are independent of x. In addition the convergence in (3.1) is shown to become extremely slow as the constant of nondegeneracy of the equation becomes small.

This behavior is absolutely different from what is happening with ordinary parabolic equations. Typically their solutions decay linearly as x approaches the boundary regardless of the size of the constant of ellipticity.

To describe the main features of the theory we will concentrate on the case $d = 1$ with one-dimensional Wiener process w_t and the equation

$$du_t(x) = ((1/2)a_t(x)D^2u_t(x) + f_t(x))\,dt + (\sigma_t(x)Du_t(x) + g_t)(x)\,dw_t, \quad x > 0, \quad (3.2)$$

with zero initial and latteral conditions.

For some constants $\delta_0, \delta_1, \in (0,1]$ we assume that

$$\delta_0^{-1} \geq a_t \geq \delta_0, \quad \alpha_t \leq (1-\delta_1)a_t, \quad \alpha = \sigma^2. \tag{3.3}$$

For $R > 0$ set

$$B_R = (-R, R).$$

Everywhere below

$$p \geq 2.$$

Here is a result for equations in the whole space which will be the basis for the theory in $\mathbb{R}_+$.

Theorem 3.1. *Take some $\varepsilon, R \in (0,\infty)$ and a stopping time τ and assume the following:*

(i) There are constants $K \in (0,\infty)$ such that for $|x| < R$ and $y \in \mathbb{R}$, such that $|x-y| \leq \varepsilon R$, we have

$$R|D\sigma_t(y)| \leq K, \quad |\sigma_t(x) - \sigma_t(y)| + |a_t(x) - a_t(y)| \leq \beta_0,$$

where $\beta_0 = \beta_0(d, p, \delta_0, \delta_1) \in (0,1]$ is a (small) constant.

(ii) We are given a function $u \in \mathcal{H}^2_{p,0}(\tau)$ *which is a solution of* (3.2) *with some* $f \in \mathbb{L}_p(\tau)$ *and* $g \in \mathbb{H}^1_p(\tau)$.

Then, for any $r \in (0, R)$, *we have*

$$\|I_{(-r,r)}D^2u\|_{\mathbb{L}_p(\tau)} + R^{-1}\|I_{(-r,r)}Du\|_{\mathbb{L}_p(\tau)} \leq N\big(\|I_{(-R,R)}f\|_{\mathbb{L}_p(\tau)} + \|I_{(-R,R)}Dg\|_{\mathbb{L}_p(\tau)}$$

$$+(R-r)^{-1}\|I_{(-R,R)}g\|_{\mathbb{L}_p(\tau)}\big) + N^*(R-r)^{-2}\|I_{(-R,R)}u\|_{\mathbb{L}_p(\tau)}, \tag{3.4}$$

where $N = N(\delta_0, \delta_1, d, p)$ *and* $N^* = N^*(K, \varepsilon, \delta_0, \delta_1, d, p)$.

Here is the main a priori estimate for SPDEs in domains. By M we denote the operator of multiplying by x.

Theorem 3.2. *Take an* $R \in (0, \infty]$ *and a stopping time* τ *and assume the following.*

(i) For some constants $\varepsilon \in (0, 1]$ *and* $K \in (0, \infty)$ *we have*

$$|x^1 D\sigma_t(x)| \leq K,$$

$$|a_t(x) - a_t(y)| + |\sigma_t(x) - \sigma_t(y)| \leq \beta_0, \tag{3.5}$$

whenever $x, y \in \mathbb{R}_+$, $x, y \leq R$, $|x - y| \leq \varepsilon(x \wedge y)$, $t > 0$, *where* $\beta_0 = \beta_0(d, p, \delta_0, \delta_1) \in (0, 1]$ *is the constant from Theorem 3.1;*

(ii) We have a function u such that $\phi u \in \mathcal{H}^2_{p,0}(\tau)$ for any $\phi \in C_0^\infty((0,R))$ and u satisfies (3.2) in $\mathbb{R}_+$ with some $f \in \mathbb{L}_{p,\theta}(\tau)$, $g \in \mathbb{H}^1_{p,\theta}(\tau)$.

Then for any $r \in (0, R/4)$

$$\|I_{(0,r)}MD^2u\|_{\mathbb{L}_{p,\theta}(\tau)} + \|I_{(0,r)}Du\|_{\mathbb{L}_{p,\theta}(\tau)} \leq N\|I_{(0,R)}Mf\|_{\mathbb{L}_{p,\theta}(\tau)}$$

$$+N\|I_{(0,R)}MDg\|_{\mathbb{L}_{p,\theta}(\tau)} + N\|I_{(0,R)}g\|_{\mathbb{L}_{p,\theta}(\tau)} + N^*\|I_{(0,R)}M^{-1}u\|_{\mathbb{L}_{p,\theta}(\tau)}, \qquad (3.6)$$

where $N = N(d, p, \delta_0, \delta_1)$ and $N^* = N^*(d, p, \delta_0, \delta_1, \varepsilon, K)$.

By letting $r \to \infty$ in (3.6) we get the following.

Corollary 3.3. *If the assumptions of Theorem 3.2 are satisfied with $R = \infty$, then*

$$\|MD^2u\|_{\mathbb{L}_{p,\theta}(\tau)} + \|Du\|_{\mathbb{L}_{p,\theta}(\tau)} \leq N\|Mf\|_{\mathbb{L}_{p,\theta}(\tau)}$$

$$+N\|g\|_{\mathbb{H}^1_{p,\theta}(\tau)} + N^*\|M^{-1}u\|_{\mathbb{L}_{p,\theta}(\tau)}.$$

Remark 3.4. We discuss assumption (3.5) in case that $R = \infty$. It implies that $|a_t(x) - a_t(y)| \leq \beta$ for a small $\beta > 0$ and all $x, y \in \mathbb{R}^d_+$ satisfying $|x - y| \leq x \wedge y$. It turns out that the behavior of $a_t(x)$ near $x = 0$ can be quite irregular.

For instance, introduce the function

$$a(x) = 2 + \cos \ln x.$$

If $x \geq y > 0$ and $|x-y| \leq \varepsilon(x \wedge y)$, then $x \leq (1+\varepsilon)y$, $\ln x \leq \ln y + \ln(1+\varepsilon)$ and, for a $\xi \in (\ln y, \ln x)$, by the mean value theorem we have

$$|a(x) - a(y)| = |\sin \xi|(\ln x - \ln y) \leq \ln(1+\varepsilon),$$

which can be made arbitrarily small by choosing an appropriate ε.

Proof of Theorem 3.2. We are going to apply Theorem 3.1 to shifted B_R. For $n = -1, 0, 1, ...$, set

$$r_n = 2^{-n/3} r.$$

Observe that if $n \geq 0$, then the half width of (r_{n+2}, r_{n-1}) equals $\rho_n := r_{n+2}/2$ and

$$r_{n-1} + \rho_n \leq 2r_{-1} < 4r < R, \quad r_{n+2} - \rho_n = \rho_n.$$

It follows that for $x \in (r_{n+2}, r_{n-1})$ and y, such that $|x-y| \leq \varepsilon \rho_n$, we have

$$R > r_{n-1} \geq x \geq r_{n+2} \geq \rho_n, \quad y \leq x + \varepsilon\rho_n < R, \quad y \geq x - \varepsilon\rho_n \geq \rho_n,$$

$$\rho_n \leq x \wedge y, \quad |x-y| \leq \varepsilon(x \wedge y),$$

so that by our assumptions

$$\rho_n |D\sigma_t(y)| \leq K,$$

$$|\sigma_t(x) - \sigma_t(y)| + |a_t(x) - a_t(y)| \leq \beta_0.$$

Furthermore, if $n \geq 0$, $\zeta \in C_0^\infty((0, R))$ and $\zeta(z) = 1$ for $r_{n+2} \leq z \leq r_{n-1}$, then ζu satisfies (3.2) in $\mathbb{R}^d$ with certain f and g which on (r_{n+2}, r_{n-1}) coincide with the original ones. Finally, if $n \geq 0$, then the distance between the boundaries of (r_{n+1}, r_n) and (r_{n+2}, r_{n-1}) is $(2^{1/3} - 1)r_{n+2}$.

It follows by Theorem 3.1 that for $n \geq 0$

$$\|I_{(r_{n+1},r_n)} D^2 u\|^p_{\mathbb{L}_p(\tau)} + r_{n-1}^{-p} \|I_{(r_{n+1},r_n)} Du\|^p_{\mathbb{L}_p(\tau)} \leq N\big(\|I_{(r_{n+2},r_{n-1})} f\|^p_{\mathbb{L}_p(\tau)}$$

$$+ \|I_{(r_{n+2},r_{n-1})} Dg\|^p_{\mathbb{L}_p(\tau)} + r_{n+2}^{-p} \|I_{(r_{n+2},r_{n-1})} g\|^p_{\mathbb{L}_p(\tau)}\big)$$

$$+ N^* r_{n+2}^{-2p} \|I_{(r_{n+2},r_{n-1})} u\|^p_{\mathbb{L}_p(\tau)}.$$

We multiply both parts by $r_{n+2}^{p+\theta-d}$ and use the facts that $r_{n-1} = 2r_{n+2}$ and on (r_{n+2}, r_{n-1}) the ratio x^1/r_{n+2} satisfies

$$1 \leq x^1/r_{n+2} \leq 2.$$

Then we obtain

$$\|I_{(r_{n+1},r_n)}MD^2u\|^p_{\mathbb{L}_{p,\theta}(\tau)}+\|I_{(r_{n+1},r_n)}Du\|^p_{\mathbb{L}_{p,\theta}(\tau)}$$

$$\leq N\big(\|I_{(r_{n+2},r_{n-1})}Mf\|^p_{\mathbb{L}_{p,\theta}(\tau)}+\|I_{(r_{n+2},r_{n-1})}MDg\|^p_{\mathbb{L}_{p,\theta}(\tau)}$$

$$+\|I_{(r_{n+2},r_{n-1})}g\|^p_{\mathbb{L}_{p,\theta}(\tau)}\big)+N^*\|I_{(r_{n+2},r_{n-1})}M^{-1}u\|^p_{\mathbb{L}_{p,\theta}(\tau)}.$$

Upon summing up these inequalities over $n\geq 0$ we conclude

$$\|I_{(0,r)}MD^2u\|^p_{\mathbb{L}_{p,\theta}(\tau)}+\|I_{(0,r)}Du\|^p_{\mathbb{L}_{p,\theta}(\tau)}\leq N\big(\|I_{(0,r_{-1})}Mf\|^p_{\mathbb{L}_{p,\theta}(\tau)}$$

$$+\|I_{(0,r_{-1})}MDg\|^p_{\mathbb{L}_{p,\theta}(\tau)}+\|I_{(0,r_{-1})}g\|^p_{\mathbb{L}_{p,\theta}(\tau)}\big)+N^*\|I_{(0,r_{-1})}M^{-1}u\|^p_{\mathbb{L}_{p,\theta}(\tau)},$$

which is somewhat sharper than (3.6). The theorem is proved.

We end the lectures by showing how to prove Theorem 3.1:

Theorem 3.5. *Take some $\varepsilon,R\in(0,\infty)$ and a stopping time τ and assume the following:*

(i) There are constants $K\in(0,\infty)$ such that for $|x|<R$ and $y\in\mathbb{R}$, such that $|x-y|\leq\varepsilon R$, we have

$$R|D\sigma_t(y)|\leq K,\quad |\sigma_t(x)-\sigma_t(y)|+|a_t(x)-a_t(y)|\leq\beta_0.$$

(ii) We are given a function $u \in \mathcal{H}^2_{p,0}(\tau)$ *which is a solution of* (3.2) *with some* $f \in \mathbb{L}_p(\tau)$ *and* $g \in \mathbb{H}^1_p(\tau)$.

Then, for any $r \in (0, R)$, *we have*

$$\|I_{(-r,r)}D^2u\|_{\mathbb{L}_p(\tau)} + R^{-1}\|I_{(-r,r)}Du\|_{\mathbb{L}_p(\tau)} \le N\big(\|I_{(-R,R)}f\|_{\mathbb{L}_p(\tau)} + \|I_{(-R,R)}Dg\|_{\mathbb{L}_p(\tau)}$$

$$+(R-r)^{-1}\|I_{(-R,R)}g\|_{\mathbb{L}_p(\tau)}\big) + N^*(R-r)^{-2}\|I_{(-R,R)}u\|_{\mathbb{L}_p(\tau)}. \tag{3.7}$$

Proof. We follow a usual procedure taken from the theory of PDEs. Let $\chi(s)$ be an infinitely differentiable function on $\mathbb{R}$ such that $\chi(s) = 1$ for $s \le 1$ and $\chi(s) = 0$ for $s \ge 2$. For $m = 0, 1, 2, ...$ introduce, $(r_0 = r)$

$$r_m = r + (R-r)\sum_{j=1}^{m} 2^{-j}, \quad \xi_m(x) = \chi\big(2^{m+1}(R-r)^{-1}(|x'| - r_m) + 1\big).$$

As is easy to check, for

$$Q(m) = (-r_m, r_m)\,,$$

it holds that

$$\zeta_m = 1 \quad \text{on} \quad Q(m), \quad \zeta_m = 0 \quad \text{outside} \quad Q(m+1).$$

Also (observe that $N2^{m+1}=N_1 2^m$ with $N_1=2N$)

$$|D\zeta_m|\leq N2^m(R-r)^{-1},\quad |D^2\zeta_m|\leq N2^{2m}(R-r)^{-2}.$$

Next, the function $\zeta_m u_t$ is in $\mathcal{H}^2_{p,0}(\tau)$ and satisfies

$$d(\zeta_m u_t)=(L(\zeta_m u_t)+f_{mt})\,dt+(\Lambda^k(\zeta_m u_t)+g^k_{mt})\,dw^k_t,$$

where

$$f_{mt}=\zeta_m f_t+uL_t\zeta_m-a^{ij}_t D_i u_t D_j\zeta_m,\quad g^k_{mt}=\zeta_m g^k_t+u_t\sigma^{ik}_t D_i\zeta_m.$$

Since $\zeta_m u_t(x)=0$ for $|x|\geq R$, it is easyly derived from the result for the whole space that

$$A_m:=\|D^2(\zeta_m u)\|^p_{\mathbb{L}_p(\tau)}+R^{-p}\|D(\zeta_m u)\|^p_{\mathbb{L}_p(\tau)}\leq N(F'+R^{-p}G^0)$$
$$+N^*R^{-2p}U^0+NE\int_0^\tau(B_{mt}+C^1_{mt}+G^1_{mt}+R^{-p}C^0_{mt})\,dt, \tag{3.8}$$

where

$$F'=\|I_{(-R,R)}f\|^p_{\mathbb{L}_p(\tau)},\quad G^0=\|I_{(-R,R)}g\|^p_{\mathbb{L}_p(\tau)},\quad U^0=\|I_{(-R,R)}u\|^p_{\mathbb{L}_p(\tau)},$$

$$B_{mt}=\|uL_t\zeta_m-a^{ij}_t D_i u_t D_j\zeta_m\|^p_{L_p},$$

$$C^1_{mt}=\|D(u_t\sigma^i_t D_i\zeta_m)\|^p_{L_p},\quad G^1_{mt}=\|D(\zeta_m g_t)\|^p_{L_p},$$

$$C^0_{mt} = \|u_t \sigma^i_t D_i \zeta_m\|^p_{L_p}.$$

Observe that by the above mentioned properties of ζ_m and the assumption on b, we have

$$|(L_t - c_t)\zeta_m| \leq N^* 2^{2m}(R-r)^{-2} + N^* R^{-1} 2^m (R-r)^{-1} \leq N^* 2^{2m}(R-r)^{-2}.$$

It follows that

$$B_{mt} \leq N^* 2^{2mp}(R-r)^{-2p} U^0_t + N^* 2^{mp}(R-r)^{-p}\|Du_t\|^p_{L_p(Q(m+1))},$$

where

$$U^0_t = \|u_t\|^p_{L_p((-R,R))}.$$

Furthermore, by interpolation inequalities for any $\gamma > 0$

$$\|Du_t\|^p_{L_p(Q(m+1))} \leq \|D(\zeta_{m+1}u_t)\|^p_{L_p}$$

$$\leq \gamma(R-r)^p 2^{-mp}\|D^2(\zeta_{m+1}u_t)\|^p_{L_p(Q(m+2))} + N\gamma^{-1}2^{mp}(R-r)^{-p}U^0_t,$$

so that for $\gamma \in (0,1)$ (with γ, perhaps, different from the one above)

$$B_{mt} \leq \gamma\|D^2(\zeta_{m+1}u_t)\|^p_{L_p(Q(m+2))} + N^*\gamma^{-1}2^{2mp}(R-r)^{-2p}U^0_t.$$

Similarly, for $\gamma \in (0,1)$

$$C^1_{mt} \leq \gamma \|D^2(\zeta_{m+1}u_t)\|^p_{L_p(Q(m+2))} + N^*\gamma^{-1}2^{2mp}(R-r)^{-2p}U^0_t$$

and almost obviously

$$C^0_{mt} \leq N2^{mp}(R-r)^{-p}U^0_t \leq R^p N\gamma^{-1}2^{2mp}(R-r)^{-2p}U^0_t$$

$$G^1_{mt} \leq N(\|Dg_t\|^p_{L_p((-R,R))} + 2^{mp}(R-r)^{-p}\|g_t\|^p_{L_p((-R,R))}).$$

Hence (3.8) yields

$$A_m \leq \gamma A_{m+1} + NF + N2^{mp}(R-r)^{-p}G^0 + N^*\gamma^{-1}2^{2mp}(R-r)^{-2p}U^0,$$

where

$$F = \|I_{(-R,R)}f\|^p_{\mathbb{L}_p(\tau)} + \|I_{(-R,R)}Dg\|^p_{\mathbb{L}_p(\tau)}.$$

Now we take $\gamma = 8^{-p}$ and get

$$\gamma^m A_m \leq \gamma^{m+1}A_{m+1} + N\gamma^m F + N\gamma^m 2^{mp}(R-r)^{-p}G^0$$

$$+N^*\gamma^m\gamma^{-1}2^{2mp}(R-r)^{-2p}U^0,$$

$$A_0 + \sum_{m=1}^{\infty}\gamma^m A_m \leq \sum_{m=1}^{\infty}\gamma^m A_m$$

$$+NF + N(R-r)^{-p}G^0 + N^*(R-r)^{-2p}U^0. \tag{3.9}$$

We cancel like terms and obtain

$$A_0 \leq NF + N(R-r)^{-p}G^0 + N^*(R-r)^{-2p}U^0.$$

The theorem is proved.

The work was partially supported by NSF grants.