Contents

CHAPTER 5

An analytic approach to SPDEs

N.V. KRYLOV
127 Vincent Hall
University of Minnesota
Minneapolis, MN, 55455
krylov@math.umn.edu

1. Introduction

Evolutional stochastic partial differential equations (SPDEs) arise in many applications of probability theory and have been treated since long ago (see [**30**]). An example of a linear second-order SPDE is given by the following equation in $\mathbb{R}^d$:

$$du = (a^{ij}u_{x^ix^j} + b^iu_{x^i} + cu + f)\,dt + (\sigma^{ik}u_{x^i} + \nu^k u + g^k)\,dw^k_t, \ \ t > 0, \tag{1.1}$$

where summation with respect to the repeated indices i, j, k is assumed as usual, i and j go from 1 to d, and k may run through $1, 2, \ldots$.

The main purpose of this publication is to present a theory of solvability of the Cauchy problem for linear and some quasi-linear equations like (1.1) in spaces of summable functions with exponent of summability $p \geq 2$. If $p = 2$, so that we are concerned with solutions belonging to the Sobolev spaces $W^n_2(\mathbb{R}^d)$, such a theory does exist and is rather complete and satisfactory (see, for instance, [**30**]). Recall that $W^n_2(\mathbb{R}^d)$ is the set of all generalized functions on $\mathbb{R}^d$ whose derivatives up to and including the nth order belong to $L_2(\mathbb{R}^d)$. Some results concerning the solvability of the first boundary-value problem in spaces like $W^n_2(D)$, where D is a smooth domain, can be found in [**2**], [**6**], [**20**], and [**32**]. Roughly speaking, the main tool in W^n_2-theory is integration by parts. There are also approaches based on semigroup methods [**5**], [**6**], which work well for the equations with nonrandom leading coefficients a^{ij} and again in the Hilbert-space framework.

The necessity of the L_p–theory arises, for instance, when one wants to find the solutions numerically. The convergence rate and the way the finite difference should be chosen depend on smoothness properties of solutions. Here by the smoothness properties we mean continuity, Hölder continuity, differentiability, continuity of derivatives, and so on.

One of inconveniences of W^n_2-theory is that $W^n_2(\mathbb{R}^d) \subset \mathcal{C}^{n-d/2}(\mathbb{R}^d)$ only if $2n > d$, and one can prove that the solutions belong to $W^n_2(\mathbb{R}^d)$ only if the coefficients are $n-2$ times continuously differentiable. Therefore, if we want to get the solutions m times continuously differentiable with respect to $x \in \mathbb{R}^d$, we have to suppose that the coefficients of the equation are more than $m + d/2 - 2$ times continuously differentiable even if the free terms are of class $C^\infty_0(\mathbb{R}^d)$. At the same time, $W^n_p(\mathbb{R}^d) \subset \mathcal{C}^{n-d/p}(\mathbb{R}^d)$ if $pn > d$, and, by taking p sufficiently large, we see that the solutions have almost as many usual derivatives as generalized ones. Actually, exactly for this purpose the spaces $W^n_p(\mathbb{R}^d)$ with $p \geq 2$ have already been used in SPDE theory (see, for instance, [**30**]), but the corresponding results, obtained again by integration by parts, were not sharp. It is worth mentioning that sharp results concerning $\mathcal{C}^{2+\alpha}(\mathbb{R}^d)$–theory are recently obtained in [**27**] for equations like (1.1) but with $\sigma \equiv 0$.

Another advantage of the W^n_p setting with $p \geq 2$ can be seen in the case of very popular equations with so-called cylindrical white noise (see, for instance, [**8**], [**26**], [**28**], [**35**], and references therein). Although these equations are covered by the general W^n_p-theory for any $p \geq 2$ (see Section 8.3), for $p = 2$ we get only the solutions summable to any degree, and the solutions become continuous only for $p > 2$. By the way, as in [**30**], we consider n positive and negative, but in contrast with [**30**] we allow n to have non-integer values. For general n we are working in the spaces of Bessel potentials $H^n_p(\mathbb{R}^d)$, and, in the case of equations with cylindrical white noise, we take n slightly less than $(-3/2)$.

Our main tool is the theory of spaces $H^n_p(\mathbb{R}^d)$, borrowed from [**33**], together with a result from [**17**] or [**18**] which is an analog of the so-called maximal regularity property of stochastic convolutions in Hilbert spaces obtained by Da Prato (cf. [**4**]). We also use some results from the theory of parabolic equations and follow a general scheme of proving the solvability of PDEs adopted in this theory. We discuss this general scheme in Sec. 2.

The main source of these notes is author's article [**20**]. However, it underwent a major revision and restructuring here, so that, in author's opinion, the whole subject looks more natural now. Also, the proofs are given with much more details, and several mistakes have been corrected. In this regard the contribution of S. Lototsky, H. Yoo, and A. Zatezalo is greatly appreciated. The author has delivered four lectures on the subject at Workshop/School on SPDEs, Theory and Applications, Los Angeles, January, 3–7, 1996. This was a very good opportunity to think all over again. It is a pleasure to thank the organizers of the Workshop for the invitation and hospitality.

2. Generalities

We want to explain here some basic ideas guiding the investigation we are going to present further.

Rewrite equation (1.1) in the following form

$$du = (Lu + f)\,dt + (\Lambda^k u + g^k)\,dw^k_t, \quad t > 0, \tag{2.1}$$

where

$$Lu = a^{ij}u_{x^ix^j} + b^i u_{x^i} + cu, \quad \Lambda^k u = \sigma^{ik}u_{x^i} + \nu^k u.$$

Recall that we are using the summation convention. Of course, w^k_t are independent Wiener processes and du is Itô's stochastic differential with respect to t. The equation is considered for all $x \in \mathbb{R}^d$ and $t > 0$. The coefficients σ, ν, and the "free" term g may vanish, and then (2.1) becomes a usual parabolic partial differential equation. Therefore, it is natural to recall some ideas and results from the L_p-theory of parabolic partial differential equations.

Let $\mathbb{R}^d$ be a d-dimensional Euclidean space of points $x = (x^1, ..., x^d)$. By a distribution or a generalized function on $\mathbb{R}^d$ we mean an element of the space $\mathcal{D}$ of real-valued Schwarz distributions defined on C^∞_0, where $C^\infty_0 = C^\infty_0(\mathbb{R}^d)$ is the set of all infinitely differentiable functions with compact support.

For given $p \in (1, \infty)$ and $n \in (-\infty, \infty)$, define the space $H^n_p = H^n_p(\mathbb{R}^d)$ (called the space of Bessel potentials or the Sobolev space with fractional derivatives) as the space of all generalized functions u such that $(1 - \Delta)^{n/2}u \in L_p = L_p(\mathbb{R}^d)$. To explain the meaning of this, let us first introduce $(1 - \Delta)^{n/2}$ in the following way (see [**14**]). If $\alpha \in (0, 1)$, then, for a constant $c(\alpha)$ and all $z \leq 0$,

$$(1-z)^\alpha = c(\alpha) \int_0^\infty \frac{e^{-t}e^{zt}-1}{t^\alpha}\,\frac{dt}{t}.$$

If we formally substitute here Δ instead of z, then we get the following definition of $(1-\Delta)^\alpha$

$$(1-\Delta)^\alpha u = c(\alpha) \int_0^\infty \frac{e^{-t}T_t u - u}{t^\alpha}\,\frac{dt}{t}, \tag{2.2}$$

where T_t is the semigroup associated with Δ, that is, a solution of $T_t' = \Delta T_t$ similarly to e^{zt} which is a solution of $f' = zf$. The following is an explicit formula for T_t:

$$T_t u(x) := \frac{1}{(4\pi t)^{d/2}} \int_{\mathbb{R}^d} u(y) e^{-\frac{1}{4t}|x-y|^2}\,dy. \tag{2.3}$$

In the same way for any $\alpha > 0$ we define

$$(1-\Delta)^{-\alpha} u = d(\alpha) \int_0^\infty t^\alpha e^{-t} T_t u\,\frac{dt}{t}, \tag{2.4}$$

where $d(\alpha)$ is an appropriate constant. It turns out (see [**14**]) that formulas (2.2) and (2.4) are sufficient to consistently define $(1-\Delta)^{n/2}$ for any $n \in (-\infty, \infty)$. The result of application of $(1-\Delta)^{-n/2}$ to an $f \in L_p$ is defined as a limit of truncated integrals in (2.2) or (2.4). For a distribution u, we say that $u \in H_p^n$ if there is $f \in L_p$ such that $u = (1-\Delta)^{-n/2} f$ in the sense of distributions, that is, if u and $(1-\Delta)^{-n/2} f$ coincide as elements in the space $\mathcal{D}$. In this case we also write $(1-\Delta)^{n/2} u = f$.

For $u \in H_p^n$ one introduces the norm

$$||u||_{n,p} := ||(1-\Delta)^{n/2} u||_p,$$

where $||\cdot||_p$ is the norm in L_p. It is known (see, for instance, [**33**], [**34**]) that H_p^n is a Banach space with norm $||\cdot||_{n,p}$ and the set C_0^∞ is dense in H_p^n.

Next, for fixed T one introduces the space $H_p^{1,2}(T) = H_p^{1,2}((0,T)\times\mathbb{R}^d)$ as

$$\{u = u(t,x) : ||u||_{1,2,p}^p := \int_0^T ||\frac{\partial u}{\partial t}(t,\cdot)||_p^p\,dt + \int_0^T ||u(t,\cdot)||_{2,p}^p\,dt < \infty\}.$$

The norm $||\cdot||_{1,2,p}$ makes $H_p^{1,2}(T)$ a Banach space.

The investigation of the deterministic counterpart of equation (2.1)

$$\frac{\partial u}{\partial t} = Lu + f \tag{2.5}$$

with zero initial condition goes in the following way (see [**31**]). First, for the simplest equation

$$\frac{\partial u}{\partial t} = \Delta u + f, \tag{2.6}$$

its solvability in $H_p^{1,2}(T)$ is proved by means of explicit formulas and some estimates of heat potentials, provided that $f \in L_p((0,T)\times\mathbb{R}^d)$. In Subsec. 4.1 we use the following theorem which is proved in [**31**] (also see Remark 2.3.2 in [**34**]).

THEOREM 2.1. *For any $f \in L_p((0,T)\times\mathbb{R}^d)$ and $u_0 \in H_p^{2-2/p}$ there exists a unique solution $u \in H_p^{1,2}(T)$ of the heat equation (2.6) with initial data $u(0) = u_0$. In addition,*

$$||u_{xx}||_{L_p((0,T)\times\mathbb{R}^d)} + ||\frac{\partial u}{\partial t}||_{L_p((0,T)\times\mathbb{R}^d)} \leq N(d,p)(||f||_{L_p((0,T)\times\mathbb{R}^d)} + ||u_0||_{2-2/p,p}), \tag{2.7}$$
$$||u||_{1,2,p} \leq N(d,p,T)(||f||_{L_p((0,T)\times\mathbb{R}^d)} + ||u_0||_{2-2/p,p}),$$

where u_{xx} is the matrix of second–order derivatives of u with respect to x.

This theorem yields a bounded operator $\mathcal{R}_1$ which maps any $f \in L_p((0,T) \times \mathbb{R}^d)$ into the solution $u \in H_p^{1,2}(T)$ of the heat equation (2.6) with zero initial data.

Then, the so–called *a priori estimate* is obtained for (2.5). One assumes that there is a solution $u \in H_p^{1,2}(T)$ of (2.5) with zero initial condition and inequality (2.7) is proved, where N is a constant probably depending on T and some characteristics of L (below, N, usually without indices or arguments, denotes various constants, and writing $N(d,p)$ is just to indicate that N depends only on d and p). At this point, one of the most fruitful general ideas in the theory of partial differential equations, linear or not, is used. The idea says that obtaining a priori estimates for solutions in a class of functions implies that the equation is solvable in the same class. We will see how this works below.

Of course, this idea only works "as usual". For instance, there is no trouble to estimate the sup norm of the second order derivative for possible solutions of the equation $(u'')^2 = f$ on $[0,1]$ satisfying $u(0) = 0$ and $u(1) = 1$, where f is a given bounded function. So, one has an a priori estimate, but nevertheless the set of solutions is empty if $f < 0$. Also, the statement "if (2.5) has a solution $u \in H_p^{1,2}(T)$, then (2.7) holds" looks bad from the point of view of formal logic (in which any statement with a false assumption is true). In fact, instead of (2.7), one proves

$$||u||_{1,2,p} \leq N||Lu - \frac{\partial u}{\partial t}||_{L_p((0,T)\times\mathbb{R}^d)} \tag{2.8}$$

for *any* $u \in H_p^{1,2}(T)$ such that in a certain sense $u(0,\cdot) = 0$. A usual way to prove (2.8) consists of observing that (2.8) is true for $L = \Delta$, then applying perturbation methods to get (2.8) for L that are close enough to Δ, and finally, replacing Δ with other operators with constant coefficients and using partitions of unity in order to get small regions in which L is close to an operator with constant coefficients (we will see this in details in the proof of Theorem 5.1).

The last step is to use *the method of continuity*. Instead of (2.5) one considers the following family of equations

$$\frac{\partial u}{\partial t} = L_\lambda u + f \tag{2.9}$$

with $\lambda \in [0,1]$, where

$$L_\lambda u = \lambda\Delta + (1-\lambda)L.$$

Assume that the a priori estimate (2.8) holds with the same constant N for all L_λ in place of L. Also assume that, for a $\lambda = \lambda_0 \in [0,1]$, equation (2.9) with zero initial data has a unique solution $u \in H_p^{1,2}(T)$ for any $f \in L_p((0,T) \times \mathbb{R}^d)$. Then we have an operator $\mathcal{R}_{\lambda_0}$ such that $\mathcal{R}_{\lambda_0} f = u$. By the way, this assumption is satisfied for $\lambda_0 = 1$ by Theorem 2.1. From (2.8) we get that

$$||\mathcal{R}_{\lambda_0} f||_{1,2,p} \leq N||f||_{L_p((0,T)\times\mathbb{R}^d)}. \tag{2.10}$$

For other $\lambda \in [0,1]$ we rewrite (2.9) as

$$\frac{\partial u}{\partial t} = L_{\lambda_0} u + \{(\lambda - \lambda_0)(\Delta - L)u + f\}, \quad u = \mathcal{R}_{\lambda_0}\{(\lambda - \lambda_0)(\Delta - L)u + f\}$$

and we solve the last equation by iterations. Define $u_0 = 0$ and

$$u_{n+1} = \mathcal{R}_{\lambda_0}\{(\lambda - \lambda_0)(\Delta - L)u_n + f\}.$$

Then by (2.10)

$$||u_{n+1} - u_n||_{1,2,p} \leq N|\lambda - \lambda_0|\,||(\Delta - L)(u_n - u_{n-1})||_{L_p((0,T)\times\mathbb{R}^d)}$$

$$\le N_1|\lambda-\lambda_0|\,||u_n-u_{n-1}||_{1,2,p},$$

and we have a strongly convergent sequence u_n in $H^{1,2}_p(T)$ if $N_1|\lambda-\lambda_0|\le 1/2$. Consequently, one can solve (2.9) for such λ. Starting from $\lambda=1$, one reaches $\lambda=0$ in finitely many steps, and this finishes the proof of solvability of (2.5).

There are two central objects in the above argument. These are the Banach space $H^{1,2}_p(T)$ and the operator $L-\partial/\partial t : H^{1,2}_p(T)\to L_p((0,T)\times\mathbb{R}^d)$. Since we want to implement the same kind of argument for equations like (2.1), the first thing to do is to find an appropriate counterpart of $H^{1,2}_p(T)$. This was a problem for some time, since one cannot expect any differentiability property with respect to t for solutions u of (2.1). Then an observation appeared that $H^{1,2}_p(T)$ can also be defined without using $\partial u/\partial t$. Indeed, almost obviously,

$$H^{1,2}_p(T) = \{u : u(t,x) = u(0,x)+\int_0^t f(s,x)\,ds, u,u_x,u_{xx},f \in L_p((0,T)\times\mathbb{R}^d)\}.$$

Somewhat unusual about this definition is that in the formula

$$u(t,x)=u(0,x)+\int_0^t f(s,x)\,ds$$

the function f only belongs to $L_p((0,T)\times\mathbb{R}^d)$ but u,u_x,u_{xx} also belong to $L_p((0,T)\times\mathbb{R}^d)$ (by definition).

Now the guess is natural that a stochastic counterpart $\mathcal{H}^2_p(T)$ of the spaces $H^{1,2}_p(T)$ could be the space of functions $u=u(\omega,t,x)$ such that

$$u(t,x)=u(0,x)+\int_0^t f(s,x)\,ds+\int_0^t g^k(s,x)\,dw^k_s, \tag{2.11}$$

$$E\int_0^T\int_{\mathbb{R}^d}\{|u|+|u_x|+|u_{xx}|+|f|\}^p\,dxdt<\infty,$$

and something of the same type is satisfied for $g=\{g^k\}$. It may look a little bit surprising that one needs $p\ge 2$ and

$$E\int_0^T\int_{\mathbb{R}^d}\{|g|+|g_x|\}^p\,dxdt<\infty,$$

which involves both g *and* g_x, where

$$|g|^2:=\sum_{k=1}^\infty |g^k|^2,\quad |g_x|^2:=\sum_{k=1}^\infty |g^k_x|^2.$$

We will explain later why one needs $p\ge 2$ (see Remark 4.3). The need for some conditions on g_x may be explained by the fact that in parabolic equations one derivative in t is worth two derivatives in x. The stochastic integral in (2.11) has, so to speak, one half derivative in time, so one can expect that it also has one derivative in x. The second derivative should be provided by differentiability of g itself. This is a kind of phenomenological explanation. The real explanation, however, is that with such definition of stochastic $\mathcal{H}^2_p(T)$ spaces one can construct a theory completely analogous to the theory discussed above.

For $p=2$ we can also see the necessity of imposing conditions on g_x from the model one–dimensional equation

$$du=u_{xx}\,dt+g\,dw_t,$$

where $x \in \mathbb{R}$ and w_t is a one–dimensional Wiener process. Forgetting for a while about rigor, by Itô's formula we get

$$d(u^2) = (2uu_{xx} + g^2)\,dt + 2ug\,dw_t, \quad \int_{\mathbb{R}} u^2(t,x)\,dx = \int_{\mathbb{R}} u^2(0,x)\,dx$$

$$+2\int_0^t \int_{\mathbb{R}} uu_{xx}(s,x)\,dx\,ds + \int_0^t \int_{\mathbb{R}} g^2(s,x)\,dx\,ds + \int_0^t \int_{\mathbb{R}} ug(s,x)\,dx\,dw_s.$$

Integrating by parts we see that

$$E||u(t,\cdot)||_2^2 + 2E\int_0^t ||u_x(s,\cdot)||_2^2\,ds = E||u(0,\cdot)||_2^2 + E\int_0^t ||g(s,\cdot)||_2^2\,ds.$$

This shows that without assumptions on g_x we *only* get an estimate of the first derivative of u.

With the spaces $\mathcal{H}_p^2(T)$ at hand, we write (2.1) in an operator form by introducing the operator (L,Λ) which can be applied to *any* element $u \in \mathcal{H}_p^2(T)$. Namely for a $u \in \mathcal{H}_p^2(T)$ we write $(L,\Lambda)u = -(f,g)$ if and only if

$$u(t) = u(0) + \int_0^t [Lu + f](s)\,ds + \int_0^t [\Lambda^k u + g^k](s)\,dw_s^k.$$

The expression $(L,\Lambda)u$ does make sense for any $u \in \mathcal{H}_p^2(T)$ since, if we have (2.11), then $(L,\Lambda)u = (Lu - f, \Lambda u - g)$. Now, instead of equation (2.1) we have the equation $(L,\Lambda)u = -(f,g)$ in the Banach space $\mathcal{H}_p^2(T)$, which makes the situation completely analogous to the one in the theory of parabolic equations.

This finishes our explanation of the basics of the theory, and now we are ready to present the details.

3. The Stochastic Banach Spaces

Let $(\Omega, \mathcal{F}, P)$ be a complete probability space, $(\mathcal{F}_t, t \geq 0)$ be an increasing filtration of σ-fields $\mathcal{F}_t \subset \mathcal{F}$ containing all P-null subsets of Ω, and $\mathcal{P}$ be the predictable σ-field generated by $(\mathcal{F}_t, t \geq 0)$. Let $\{w_t^k; k = 1, 2, ...\}$ be a family of independent one–dimensional $\mathcal{F}_t$–adapted Wiener processes defined on $(\Omega, \mathcal{F}, P)$.

We fix a $p \geq 2$ and an integer $d \geq 1$ and use the notation C_0^∞, $\mathcal{D}$, L_p, H_p^n, $||\cdot||_p$, and $||\cdot||_{n,p}$ introduced in Sec. 2. We are dealing with distributions, and, for a distribution u and $\phi \in C_0^\infty$, by (u,ϕ) or (ϕ,u) we mean the result of application of u to ϕ. Observe that, for $u \in H_p^n$ and $\phi \in C_0^\infty$, by definition

$$(u,\phi) = ((1-\Delta)^{n/2}u, (1-\Delta)^{-n/2}\phi) = \int_{\mathbb{R}^d} [(1-\Delta)^{n/2}u](x)(1-\Delta)^{-n/2}\phi(x)\,dx, \tag{3.1}$$

where the last integral is a usual Lebesgue integral. Since $(1-\Delta)^{n/2}u \in L_p$, one can define (u,ϕ) by (3.1) for any ϕ whose derivatives vanish sufficiently fast at infinity, say exponentially fast.

Sometimes it is useful to notice that $||\cdot||_{n,p} \leq ||\cdot||_{m,p}$ for $m \geq n$. Recall that the set C_0^∞ is dense in H_p^n and the latter is a subset of $\mathcal{D}$ (see, for instance, Theorem 2.3 (ii) of [**33**]). Also recall (see, for instance [**33**]) that, for integers $n \geq 0$, the space H_p^n coincides with the Sobolev space $W_p^n = W_p^n(\mathbb{R}^d)$.

We apply the same definitions to l_2-valued functions g, where l_2 is the set of all real-valued sequences $g = \{g^k; k = 1, 2, ...\}$ with the norm defined by $|g|_{l_2}^2 := \sum_k |g^k|^2$. Specifically,

$$||g||_p := ||\,|g|_{l_2}||_p, \quad ||g||_{n,p} := ||\,|(1-\Delta)^{n/2}g|_{l_2}||_p.$$

Finally, for stopping times τ, we denote $(0,\tau]\!] = \{(\omega,t) : 0 < t \le \tau(\omega)\}$

$$\mathbb{H}^n_p(\tau) = L_p((0,\tau]\!], \mathcal{P}, H^n_p), \quad \mathbb{H}^n_p = \mathbb{H}^n_p(\infty),$$

$$\mathbb{H}^n_p(\tau, l_2) = L_p((0,\tau]\!], \mathcal{P}, H^n_p(\mathbb{R}^d, l_2)), \quad \mathbb{L}_{\dots}\dots = \mathbb{H}^0_{\dots}\dots .$$

The norms in these spaces are defined in an obvious way. By convention, elements of spaces like $\mathbb{H}^n_p$ are treated as functions rather than distributions or classes of equivalent functions, and if we know that a function of this class has a modification with better properties, then we always consider this modification. For instance, if we take $u \in H^n_p$ and $n - d/p > 0$, then u has a bounded continuous modification, but we talk about $\sup_x u(x)$ instead of sup of this modification. Also, elements of spaces $\mathbb{H}^n_p(\tau, l_2)$ need not be defined or belong to H^n_p for all $(\omega,t) \in (0,\tau]\!]$. As usual, these properties are needed only for almost all (ω,t).

For $n \in \mathbb{R}$ and

$$(f,g) \in \mathcal{F}^n_p(\tau) := \mathbb{H}^n_p(\tau) \times \mathbb{H}^{n+1}_p(\tau, l_2),$$

set

$$||(f,g)||_{\mathcal{F}^n_p(\tau)} := ||f||_{\mathbb{H}^n_p(\tau)} + ||g||_{\mathbb{H}^{n+1}_p(\tau,l_2)}.$$

DEFINITION 3.1. For a $\mathcal{D}$-valued function $u \in \bigcap_{T>0} \mathbb{H}^n_p(\tau \wedge T)$, we write $u \in \mathcal{H}^n_p(\tau)$ if $u_{xx} \in \mathbb{H}^{n-2}_p(\tau)$, $u(0,\cdot) \in L_p(\Omega, \mathcal{F}_0, H^{n-2/p}_p)$, and there exists $(f,g) \in \mathcal{F}^{n-2}_p(\tau)$ such that, for any $\phi \in C^\infty_0$, the equality

$$(u(t,\cdot),\phi) = (u(0,\cdot),\phi) + \int_0^t (f(s,\cdot),\phi)\,ds + \sum_{k=1}^\infty \int_0^t (g^k(s,\cdot),\phi)\,dw^k_s \tag{3.2}$$

holds for all $t \le \tau$ with probability 1. We also define $\mathcal{H}^n_{p,0}(\tau) = \mathcal{H}^n_p(\tau) \cap \{u : u(0,\cdot) = 0\}$,

$$||u||_{\mathcal{H}^n_p(\tau)} = ||u_{xx}||_{\mathbb{H}^{n-2}_p(\tau)} + ||(f,g)||_{\mathcal{F}^{n-2}_p(\tau)} + (E||u(0,\cdot)||^p_{n-2/p,p})^{1/p}. \tag{3.3}$$

As always, we drop τ in $\mathcal{H}^n_p(\tau)$ and $\mathcal{F}^n_p(\tau)$ if $\tau = \infty$.

REMARK 3.2. It is worth noting that the elements of $\mathcal{H}^n_p(\tau)$ are assumed to be defined for all (ω,t) and take values in $\mathcal{D}$. Obviously, $\mathcal{H}^n_p(\tau)$ is a linear space. As usual, we identify two elements u_1 and u_2 of $\mathcal{H}^n_p(\tau)$ if $||u_1 - u_2||_{\mathcal{H}^n_p(\tau)} = 0$. Actually, as we will see later, in this way we only identify indistinguishable functions. Also, observe that the series of stochastic integrals in (3.2) converges uniformly in t in probability on $[0, \tau \wedge T]$ for any finite T, since the quadratic variations of these stochastic integrals satisfy (cf. (3.1))

$$\sum_{k=1}^\infty \int_0^{\tau\wedge T} (g^k(s,\cdot),\phi)^2\,ds = \sum_{k=1}^\infty \int_0^{\tau\wedge T} ((1-\Delta)^{(n-1)/2} g^k(s,\cdot), (1-\Delta)^{(1-n)/2}\phi)^2_{L_2}\,ds$$

$$\le ||(1-\Delta)^{(1-n)/2}\phi||_1 \int_0^{\tau\wedge T} \sum_{k=1}^\infty (|(1-\Delta)^{(n-1)/2} g^k(s,\cdot)|^2, |(1-\Delta)^{(1-n)/2}\phi|)_{L_2}\,ds$$

$$\le N \int_0^{\tau\wedge T} ||(\sum_{k=1}^\infty |(1-\Delta)^{(n-1)/2} g^k(s,\cdot)|^2)^{1/2}||^2_p\,ds < \infty \quad \text{(a.s.)},$$

where $N = ||(1-\Delta)^{(1-n)/2}\phi||_1\,||(1-\Delta)^{(1-n)/2}\phi||_q$, $q = p/(p-2)$ and we have used that $p \ge 2$.

As a consequence of the uniform convergence, $(u(t,\cdot),\phi)$ is continuous in t on $[0, \tau \wedge T]$ for any finite T (a.s.).

REMARK 3.3. There can exist only one couple (f,g) for which (3.2) holds. Indeed, if there are two, then one can represent zero as a sum of a continuous process of bounded variation and a continuous local martingale. One knows that this is only possible if both processes vanish. Therefore, the couple (f,g) is uniquely determined by u, and notation $||u||_{\mathcal{H}_p^n(\tau)}$ in (3.3) makes sense.

REMARK 3.4. It is known that the operator $(1-\Delta)^{m/2}$ maps isometrically H_p^n onto H_p^{n-m} for any n,m. Also, the inequalities from Remark 3.2 can be used to show that given $u\in\mathcal{H}_p^n(\tau)$, one can in (3.2) take any infinitely differentiable function ϕ whose derivatives vanish sufficiently fast at infinity, say exponentially fast. This allows us to substitute $(1-\Delta)^{m/2}\phi$ in (3.2) instead of ϕ and shows that the operator $(1-\Delta)^{m/2}$ maps isometrically $\mathcal{H}_p^n(\tau)$ onto $\mathcal{H}_p^{n-m}(\tau)$ for any n,m. The same is true for $\mathbb{H}_p^n(\tau)$.

DEFINITION 3.5. For $u\in\mathcal{H}_p^n(\tau)$, if (3.2) holds, then we write $f=\mathbb{D}u$, $g=\mathbb{S}u$ (for "deterministic" and "stochastic" parts of u) and we also write

$$u(t)=u(0)+\int_0^t \mathbb{D}u(s)\,ds+\int_0^t \mathbb{S}^k u(s)\,dw_s^k,\quad du=f\,dt+g^k\,dw_t^k\quad t\le\tau.$$

REMARK 3.6. It follows from Definitions 3.1 and 3.5 that the operators $\mathbb{D}$ and $\mathbb{S}$ are continuous operators from $\mathcal{H}_p^n(\tau)$ to $\mathbb{H}_p^{n-2}(\tau)$ and $\mathbb{H}_p^{n-1}(\tau,l_2)$ respectively. From Theorem 4.2 and Remark 3.4 it follows that $\mathbb{S}$ maps $\mathcal{H}_p^n(\tau)$ onto $\mathbb{H}_p^{n-1}(\tau,l_2)$. However, at this point we do not know how rich $\mathcal{H}_p^n(\tau)$ is. Nevertheless obviously $H_p^{1,2}(T)\subset\mathcal{H}_p^2(T)$.

THEOREM 3.7. *The spaces $\mathcal{H}_p^n(\tau)$ and $\mathcal{H}_{p,0}^n(\tau)$ are Banach spaces with norm (3.3). In addition if $\tau\le T$, where T is a finite constant, then for $u\in\mathcal{H}_p^n(\tau)$*

$$||u||_{\mathbb{H}_p^n(\tau)}\le N(d,T)||u||_{\mathcal{H}_p^n(\tau)},\quad E\sup_{t\le\tau}||u(t,\cdot)||_{n-2,p}^p\le N(d,T)||u||_{\mathcal{H}_p^n(\tau)}^p.\tag{3.4}$$

Proof. We first deal with (3.4). Obviously

$$||u||_{\mathbb{H}_p^n(\tau)}=||(1-\Delta)u||_{\mathbb{H}_p^{n-2}(\tau)}\le||u||_{\mathbb{H}_p^{n-2}(\tau)}+||u||_{\mathcal{H}_p^n(\tau)},$$

so that to prove (3.4) we only need to prove that

$$E\sup_{t\le\tau}||u(t,\cdot)||_{n-2,p}^p\le N||u||_{\mathcal{H}_p^n(\tau)}^p.\tag{3.5}$$

Owing to Remark 3.4 we may and will assume that $n=2$. Take a nonnegative function $\zeta\in C_0^\infty$ with unit integral, for $\varepsilon>0$ define $\zeta_\varepsilon(x)=\varepsilon^{-d}\zeta(x/\varepsilon)$, and for generalized functions u let $u^{(\varepsilon)}(x)=u*\zeta_\varepsilon(x)$. Observe that $u^{(\varepsilon)}(x)$ is a continuous (infinitely differentiable) function of x for any distribution u. Plugging in $\zeta_\varepsilon(\cdot-x)$ instead of ϕ in (3.2), we get that for any x the equality

$$u^{(\varepsilon)}(t,x)=u^{(\varepsilon)}(0,x)+\int_0^t f^{(\varepsilon)}(s,x)\,ds+\sum_{k=1}^\infty\int_0^t g^{(\varepsilon)k}(s,x)\,dw_s^k\tag{3.6}$$

holds almost surely for all $t\le\tau$. If necessary, we redefine the stochastic integrals in (3.6) in such a way that (3.6) would hold for all ω,t, and x such that $t\le\tau$. Here

$$E||u^{(\varepsilon)}(0,\cdot)||_p^p\le E||u(0,\cdot)||_p^p\le E||u(0,\cdot)||_{n-2/p,p}^p\le||u||_{\mathcal{H}_p^n(\tau)}^p,$$

where we use that, by Minkowski's inequality, $||h^{(\varepsilon)}||_p\le||\zeta_\varepsilon||_1\,||h||_p=||h||_p$. Similarly,

$$|\int_0^t f^{(\varepsilon)}(s,x)\,ds|^p\le T^{p-1}\int_0^\tau|f^{(\varepsilon)}(s,x)|^p\,ds,$$

$$E\sup_{t\le\tau}||\int_0^t f^{(\varepsilon)}(s,\cdot)\,ds||_p^p \le T^{p-1}E\int_0^\tau ||f(s,\cdot)||_p^p\,ds \le T^{p-1}||u||^p_{\mathcal{H}_p^n(\tau)}.$$

Finally, by Burkholder–Davis–Gundy inequalities

$$E\sup_{t\le\tau}|\sum_{k=1}^\infty\int_0^t g^{(\varepsilon)k}(s,x)\,dw_s^k|^p \le NE|\int_0^\tau\sum_{k=1}^\infty|g^{(\varepsilon)k}(s,x)|^2\,ds|^{p/2}$$
$$= NE|\int_0^\tau |g^{(\varepsilon)}|^2_{l_2}(s,x)\,ds|^{p/2},$$

and as above (the first term below makes sense by virtue of (3.6))

$$E\sup_{t\le\tau}||\sum_{k=1}^\infty\int_0^t g^{(\varepsilon)k}(s,\cdot)\,dw_s^k||_p^p \le \int_{\mathbb{R}^d}E\sup_{t\le\tau}|\sum_{k=1}^\infty\int_0^t g^{(\varepsilon)k}(s,x)\,dw_s^k|^p\,dx$$
$$\le NE\int_{\mathbb{R}^d}|\int_0^\tau|g^{(\varepsilon)}|^2_{l_2}(s,x)\,ds|^{p/2}\,dx \le NE(\int_0^\tau ||\,|g^{(\varepsilon)}|^2_{l_2}(s,\cdot)||_{p/2}\,ds)^{p/2}$$
$$= NE(\int_0^\tau ||\,|g^{(\varepsilon)}|_{l_2}(s,\cdot)||_p^2\,ds)^{p/2} \le NE\int_0^\tau ||\,|g^{(\varepsilon)}|_{l_2}(s,\cdot)||_p^p\,ds$$
$$\le N||g||^p_{\mathbb{L}_p(\tau,l_2)} \le N||u||^p_{\mathcal{H}_p^n(\tau)}.$$

This, along with (3.6), leads to

$$E\sup_{t\le\tau}||u^{(\varepsilon)}(t,\cdot)||_p^p \le N||u||^p_{\mathcal{H}_p^n(\tau)}. \tag{3.7}$$

Furthermore, by using the fact that $||h^{(\varepsilon)}-h^{(\gamma)}||_p\to 0$ whenever $h\in L_p$ and $\varepsilon,\gamma\to 0$ and by considering $u^{(1/m)}-u^{(1/k)}$ instead of $u^{(\varepsilon)}$, we easily see that that $u^{(1/m)}(t\wedge\tau,x)$ is a Cauchy sequence in $L_p(\Omega,B([0,T],L_p))$. Define $\bar u$ as its limit in this space. Then, for a subsequence m', we have $u^{(1/m')}(t,\cdot)\to\bar u(t,\cdot)$ in L_p if $t\le\tau$ with probability 1. On the other, hand $u^{(1/m)}(t,\cdot)\to u(t,\cdot)$ in the sense of distributions for all ω and t such that $t\le\tau(\omega)$. Therefore with probability one we have $u(t,\cdot)\in L_p$ for $t\le\tau$. Now, (3.7) and Fatou's lemma yield (3.5) for $n=2$. As explained above, this proves (3.4).

Next, we derive the first assertion of our theorem from (3.4). As usual, we have only to check the completeness of $\mathcal{H}_p^n(\tau)$. If $\{u_j\}$ is a Cauchy sequence in $\mathcal{H}_p^n(\tau)$, then it is a Cauchy sequence in $\mathbb{H}_p^n(\tau\wedge T)$ for any T, and there is $u\in\bigcap_{T>0}\mathbb{H}_p^n(\tau\wedge T)$ such that $||u-u_j||_{\mathbb{H}_p^n(\tau\wedge T)}\to 0$ as $j\to\infty$. Furthermore, u_{jxx} form a Cauchy sequence and therefore converge in $\mathbb{H}_p^{n-2}(\tau)$. It follows easily that $||u_{xx}-u_{jxx}||_{\mathbb{H}_p^{n-2}(\tau)}\to 0$.

Also, for $u_j(0),f_j,g_j$ corresponding to u_j, there is $u(0)\in L_p(\Omega,\mathcal{F}_0,H_p^{n-2/p})$ and $(f,g)\in\mathcal{F}_p^{n-2}(\tau)$ such that

$$E||u(0)-u_j(0)||^p_{n-2/p,p}\to 0,\quad ||f-f_j||_{\mathbb{H}_p^{n-2}(\tau)}\to 0,\quad ||g-g_j||_{\mathbb{H}_p^{n-1}(\tau,l_2)}\to 0.$$

By using the argument from Remark 3.2, one can show that for any $\phi\in C_0^\infty$ equality (3.2) holds in $(0,\tau]\!]$ almost everywhere.

On the other hand, (3.4) also implies that for u (at least for a modification of u) it holds that

$$E\sup_{t\le\tau\wedge T}||u(t,\cdot)-u_j(t,\cdot)||^p_{n-2,p}\to 0$$

for any constant $T<\infty$. Adding to this that the processes $(u_j(t,\cdot),\phi)$ are continuous (a.s.) (see Remark 3.2), we conclude that $(u(t,\cdot),\phi)$ is also continuous (a.s.). Thus, for any $\phi\in C_0^\infty$, equality (3.2) not only holds in $(0,\tau]\!]$ almost everywhere but also for all $t\le\tau$ almost surely. Hence, $u\in\mathcal{H}_p^n(\tau)$ and $u_j\to u$ in $\mathcal{H}_p^n(\tau)$. The theorem is proved.

REMARK 3.8. We could replace the first term on the right in (3.3) with $||u||_{\mathbb{H}^n_p(\tau)}$ and, for bounded τ, we would get an equivalent norm by virtue of (3.4). The form of (3.3) that we have chosen is convenient in the future when we need certain constants to be independent of T, see, for instance, Theorem 4.10.

REMARK 3.9. In Theorems 7.1 and 7.2 below, we prove much sharper estimates than (3.5).

We also need the following properties of the spaces $\mathcal{H}^n_p(\tau)$ and $\mathbb{H}^n_p(\tau)$.

THEOREM 3.10. *Take $g \in \mathbb{H}^n_p(l_2)$. Then there exists a sequence $g_j \in \mathbb{H}^n_p(l_2)$, $j = 1, 2...$, such that $||g - g_j||_{\mathbb{H}^n_p(l_2)} \to 0$ as $j \to \infty$ and*

$$g^k_j = \begin{cases} \sum_{i=1}^j I_{(\tau^j_{i-1}, \tau^j_i]}(t) g^{ik}_j(x) & \text{if} \quad k \le j, \\ 0 & \text{if} \quad k > j, \end{cases}$$

where τ^j_i are bounded stopping times, $\tau^j_{i-1} \le \tau^j_i$, and $g^{ik}_j \in C^\infty_0$.

Proof. The argument in Remark 3.4 and the fact that C^∞_0 is dense in any H^n_p show that we only need to consider $n = 0$. Further, one can easily understand that the set of $g \in \mathbb{L}_p(l_2)$ for which the statement holds forms a linear *closed* subspace $\mathbb{L}$ of $\mathbb{L}_p(l_2)$. We have to prove that $\mathbb{L} = \mathbb{L}_p(l_2)$. If this is not true, then, by Riesz's theorem, there is a nonzero $h \in \mathbb{L}_q(l_2)$ with $q = p/(p-1)$ such that

$$E\int_0^\infty \int_{\mathbb{R}^d} (h, g)_{l_2} \, dxdt = 0$$

for any $g \in \mathbb{L}$. In particular,

$$E\int_0^\infty I_{(0,\tau]}(\int_{\mathbb{R}^d} h^k g \, dx) dt = 0$$

for any bounded stopping time τ, $k \ge 1$, and $g \in C^\infty_0$. Since $\int_{\mathbb{R}^d} h^k g \, dx$ is equal to a predictable function (a.e.), it follows that $\int_{\mathbb{R}^d} h^k g \, dx = 0$ on $(0, \infty]\!]$ (a.e.). By taking g from a countable subset $\mathcal{G}$ in C^∞_0 that is dense in L_p, we get that, on a subset of $(0, \infty]\!]$ of full measure,

$$\int_{\mathbb{R}^d} h^k g \, dx = 0 \quad \forall g \in \mathcal{G}, k \ge 1.$$

But then $h^k = 0$ on $(0, \infty]\!] \times \mathbb{R}^d$ (a.e.). This contradicts that $h \ne 0$ and proves the theorem.

THEOREM 3.11. *Let $T \in (0, \infty)$. If $u_j \in \mathcal{H}^n_p(T)$, $j = 1, 2, ...,$ and $||u_j||_{\mathcal{H}^n_p(T)} \le K$, where K is a finite constant, then there exists a subsequence j' and a function $u \in \mathcal{H}^n_p(T)$ such that*

(i) $u_{j'}$, $u_{j'}(0, \cdot)$, $\mathbb{D}u_{j'}$, and $\mathbb{S}u_{j'}$ converge weakly to u, $u(0, \cdot)$, $\mathbb{D}u$, and $\mathbb{S}u$ in $\mathbb{H}^n_p(T)$, $L_p(\Omega, H^{n-2/p}_p)$, $\mathbb{H}^{n-2}_p(T)$, and $\mathbb{H}^{n-1}_p(T, l_2)$ respectively;

(ii) $||u||_{\mathcal{H}^n_p(T)} \le K$;

(iii) for any $\phi \in C^\infty_0$ and any $t \in [0, T]$ we have $(\phi, u_{j'}(t, \cdot)) \to (\phi, u(t, \cdot))$ weakly in $L_p(\Omega)$.

Proof. From properties of L_p spaces, it follows that there exists a subsequence j' such that $u_{j'}$, $u_{j'}(0, \cdot)$, $\mathbb{D}u_{j'}$, $\mathbb{S}u_{j'}$ converge weakly to some u, u_0, f, g in $\mathbb{H}^n_p(T)$, $L_p(\Omega, H^{n-2/p}_p)$, $\mathbb{H}^{n-2}_p(T)$, and $\mathbb{H}^{n-1}_p(T, l_2)$ respectively. Then, for any $\phi \in C^\infty_0$, the expressions $(u_{j'}(t, \cdot), \phi)$, $(u_{j'}(0, \cdot), \phi)$, $(\mathbb{D}u_{j'}(s, \cdot), \phi)$, and $(\mathbb{S}^k u_{j'}(s, \cdot), \phi)$ in the formula

$$(u_{j'}(t, \cdot), \phi) = (u_{j'}(0, \cdot), \phi) + \int_0^t (\mathbb{D}u_{j'}(s, \cdot), \phi) \, ds + \sum_{k=1}^\infty \int_0^t (\mathbb{S}^k u_{j'}(s, \cdot), \phi) \, dw^k_s$$

converge weakly in corresponding spaces. Since integration and stochastic integration can be considered as continuous linear operators and any continuous linear operator is also weakly continuous, we have that, for any $\phi \in C_0^\infty$,

$$(u(t,\cdot),\phi) = (u_0,\phi) + \int_0^t (f(s,\cdot),\phi)\,ds + \sum_{k=1}^\infty \int_0^t (g^k(s,\cdot),\phi)\,dw_s^k \tag{3.8}$$

for almost all $(\omega, t) \in \Omega \times [0,T]$.

By Banach–Saks theorem, there is a sequence $(v_{j'}, \mathbb{D}v_{j'}, \mathbb{S}v_{j'})$ of convex combinations of $(u_{j'}, \mathbb{D}u_{j'}, \mathbb{S}u_{j'})$ which converges strongly to (u, f, g) in $\mathbb{H}_p^n(T) \times \mathbb{H}_p^{n-2}(T) \times \mathbb{H}_p^{n-1}(T, l_2)$. From (3.4), it follows that

$$E \sup_{t\le T} ||v_j - v_i||_{n-2,p}^p \to 0$$

as $i, j \to \infty$. Therefore, there is a H_p^{n-2}–valued function v on $\Omega \times [0,T]$ such that

$$E \sup_{t\le T} ||v_j - v||_{n-2,p}^p \to 0.$$

In particular, $(v_j(t,\cdot),\phi) \to (v(t,\cdot),\phi)$ uniformly on $[0,T]$ in probability for any $\phi \in C_0^\infty$. On the other hand, the strong convergence of v_j to u in $\mathbb{H}_p^n(T)$ implies that $(v_j(t,\cdot),\phi) \to (u(t,\cdot),\phi)$ on $\Omega \times [0,T]$ in measure. This shows that $(v(t,\cdot),\phi) = (u(t,\cdot),\phi)$ on $\Omega \times [0,T]$ (a.e.). Because of arbitrariness of ϕ and the fact that C_0^∞ is dense in the separable spaces conjugate to H_p^m, $u = v$ (as generalized functions) on $\Omega \times [0,T]$ (a.e.).

Thus, $v \in \mathbb{H}_p^n(T)$. Also, $(v_j(t),\phi)$ are given by equations similar to (3.8), which implies that $(v_j(t),\phi)$ are continuous in t (a.s.). The uniform convergence of $(v_j(t),\phi)$ to $(v(t),\phi)$ in probability yields the continuity of $(v(t),\phi)$ (a.s.). By the above, (3.8) holds for almost all $(\omega,t) \in \Omega \times [0,T]$ if we replace $(u(t,\cdot),\phi)$ by $(v(t,\cdot),\phi)$. Since the latter is continuous and the right–hand side of (3.8) is continuous, $(v(t,\cdot),\phi)$ equals the right–hand side of (3.8) for all $t \in [0,T]$ (a.s.). Hence, $v \in \mathcal{H}_p^n(T)$ and we have proved assertion (i) for v instead u, which is irrelevant.

Assertion (ii) follows from the equality $u = v$ on $\Omega \times [0,T]$ (a.e.) and from the fact that the norm of a weak limit is less than the liminf of norms.

To prove (iii), take $\phi \in C_0^\infty$ and $\xi \in L_q(\Omega)$ with $q = p/(p-1)$ and write

$$E\xi(u_j(t,\cdot),\phi) = E\xi(u_j(0,\cdot),\phi) + E\xi\int_0^t (\mathbb{D}u_j(s,\cdot),\phi)\,ds + E\xi\int_0^t (\mathbb{S}^k u_j(s,\cdot),\phi)\,dw_s^k.$$

By what has been said about the properties of the operators of integration and by (i),

$$\begin{aligned}\lim_{j'\to\infty} E\xi(u_{j'}(t,\cdot),\phi) = \lim_{j'\to\infty}&[E\xi(u_{j'}(0,\cdot),\phi) + E\xi\int_0^t (\mathbb{D}u_{j'}(s,\cdot),\phi)\,ds \\ &+ E\xi\int_0^t (\mathbb{S}^k u_{j'}(s,\cdot),\phi)\,dw_s^k] = E\xi(u(0,\cdot),\phi) \\ &+ E\xi\int_0^t (\mathbb{D}u(s,\cdot),\phi)\,ds + E\xi\int_0^t (\mathbb{S}^k u(s,\cdot),\phi)\,dw_s^k = E\xi(u(t,\cdot),\phi),\end{aligned}$$

which proves (iii) and the theorem.

4. Model Equations

As explained in Sec. 2, to implement our general scheme we have to show that (at least) simple equations are solvable in $\mathcal{H}_{p,0}^n(\tau)$.

Except for Subsec. 4.2, we will always understand equations like (2.1) in the sense of Definition 3.5, which means that we will be looking for a function $u \in \mathcal{H}^n_{p,0}(\tau)$ such that

$$\mathbb{D}u = Lu + f, \quad \mathbb{S}u = \Lambda u + g.$$

In this section, we consider equation (2.1) when $b^i = c = \nu^k = 0$ and the coefficients a and σ do not depend on x. Throughout the section, we fix real-valued functions $a^{ij}(t)$ and l_2-valued functions $\sigma^i(t) = \{\sigma^{ik}(t), k \geq 1\}$ defined for $i, j = 1, ..., d$ on $\Omega \times (0, \infty)$. Define

$$\alpha^{ij}(t) = \frac{1}{2}(\sigma^i(t), \sigma^j(t))_{l_2}$$

and assume that a and σ are $\mathcal{P}$-measurable functions, and in the matrix sense

$$(a^{ij}) = (a^{ij})^*, \quad K(\delta^{ij}) \geq (a^{ij}) \geq (a^{ij} - \alpha^{ij}) \geq \delta(\delta^{ij}),$$

where K and δ are some fixed strictly positive constants. By the way, the assumption that $a > \alpha$ is necessary even to have L_2–theory for SPDEs with constant coefficients (see [**30**] or Remark 4.8).

Equation (2.1) takes the following form:

$$du(t,x) = (a^{ij}(t)u_{x^ix^j}(t,x) + f(t,x))\,dt + (\sigma^{ik}(t)u_{x^i}(t,x) + g^k(t,x))\,dw^k_t, \;\; t > 0. \qquad (4.1)$$

Our plan is as follows. In Subsec. 4.1 we consider the case of the heat equation with random right–hand side and get basic a priori estimates. The results of Subsec. 4.2 show how to reduce general case of equation (4.1) to the case of the heat equation and get some representation formulas for solutions. At that point it turns out that the assumption that $u \in \mathcal{H}^n_p$ is not very convenient, and we consider larger spaces. In the final Subsec. 4.3, we present the basic a priori estimates and the existence and uniqueness theorem for equation (4.1).

4.1. Particular Case $a^{ij} = \delta^{ij}$, $\sigma = 0$. We start with the equation

$$du(t,x) = (\Delta u(t,x) + f(t,x))\,dt + g^k(t,x)\,dw^k_t, \;\; t > 0. \qquad (4.2)$$

We need a lemma from [**17**] or [**20**]. Remember that the operators T_t are defined by (2.3) and, as always, $p \geq 2$.

LEMMA 4.1. *Let $-\infty \leq a < b \leq \infty$, $g \in L_p((a,b) \times \mathbb{R}^d, l_2)$. Then*

$$\int_{\mathbb{R}^d}\int_a^b [\int_a^t |\nabla T_{t-s}g(s,\cdot)(x)|^2_{l_2}\,ds]^{p/2}\,dt\,dx \leq N(d,p)\int_{\mathbb{R}^d}\int_a^b |g(t,x)|^p_{l_2}\,dt\,dx.$$

THEOREM 4.2. *Take $f \in \mathbb{H}^{-1}_p$, $g \in \mathbb{L}_p(l_2)$. Then*
(i) equation (4.2) with zero initial condition has a unique solution $u \in \mathcal{H}^1_p$;
(ii) for this solution, we have

$$||u_{xx}||_{\mathbb{H}^{-1}_p} \leq N(d,p)(||f||_{\mathbb{H}^{-1}_p} + ||g||_{\mathbb{L}_p(l_2)}); \qquad (4.3)$$

(iii) for this solution, we have $u \in C_{loc}([0,\infty), L_p)$ almost surely, and, for any $\lambda, T > 0$,

$$E\sup_{t\leq T}(e^{-p\lambda t}||u(t,\cdot)||^p_p) + E\int_0^T e^{-p\lambda t}||\,|u|^{(p-2)/p}|u_x|^{2/p}(t,\cdot)||^p_p\,dt \leq$$

$$N(d,p,\lambda)(||e^{-\lambda t}f||^p_{\mathbb{H}^{-1}_p(T)} + ||e^{-\lambda t}g||^p_{\mathbb{L}_p(T,l_2)}). \qquad (4.4)$$

Proof. It is well known that there exists a continuous linear operator

$$P : H_p^{-1} \to (L_p)^{d+1}$$

such that if $h \in H_p^{-1}$ and $Ph = (h_0, \tilde{h}^1, ..., \tilde{h}^d)$, then $h = h_0 + \operatorname{div} \tilde{h}$ and

$$||\tilde{h}||_p + ||h_0||_p \leq N(d,p)||h||_{-1,p}, \quad ||h||_{-1,p} \leq N(d,p)\{||\tilde{h}||_p + ||h_0||_p\}. \tag{4.5}$$

Actually, one can take $\tilde{h} = -\operatorname{grad}(1-\Delta)^{-1}h$ and $h_0 = h - \operatorname{div} \tilde{h} = (1-\Delta)^{-1}h$. Indeed, $||h_0||_p = ||h||_{-2,p} \leq ||h||_{-1,p}$. Also, the fact that $\partial/\partial x^i$ is a bounded operator from H_p^n to H_p^{n+1} for any n (see, for instance, [**33**]) means that $(\partial/\partial x^i)(1-\Delta)^{-1/2}$ is a bounded operator from H_p^n to H_p^n and $(\partial/\partial x^i)(1-\Delta)^{-1}$ is a bounded operator from H_p^n to H_p^{n-1}. This is why $||\tilde{h}||_p \leq N(d,p)||h||_{-1,p}$. This gives the first estimate in (4.5). On the other hand, $(1-\Delta)^{-1/2}h = (1-\Delta)^{-1/2}h_0 + \operatorname{div}(1-\Delta)^{-1/2}\tilde{h}$, and both operators $(1-\Delta)^{-1/2}$ and $\operatorname{div}(1-\Delta)^{-1/2}$ are bounded on L_p.

Define $(f_0, \tilde{f}) = Pf$. Then equation (4.2) takes the form

$$du = (\Delta u + f_0 + \operatorname{div} \tilde{f})\,dt + g^k\,dw_t^k, \tag{4.6}$$

and we supply it with zero initial condition. We will prove that, for arbitrary $f_0, \tilde{f}^i \in \mathbb{L}_p$, our assertions hold for (4.6) in place of (4.2). Of course, in (4.3) and (4.4), by f we mean $f_0 + \operatorname{div} \tilde{f}$.

A particular case. First we consider the case in which

$$\begin{aligned} f_0(t,x) = \sum_{i=1}^m I_{(\tau_{i-1},\tau_i]}(t) f_{0i}(x), \quad \tilde{f}(t,x) = \sum_{i=1}^m I_{(\tau_{i-1},\tau_i]}(t)\tilde{f}_i(x), \\ g(t,x) = \sum_{k=1}^m g^k(t,x)h_k, \quad g^k(t,x) = \sum_{i=1}^m I_{(\tau_{i-1},\tau_i]}(t) g^{ik}(x), \end{aligned} \tag{4.7}$$

where (h_k) is the standard orthonormal basis in l_2, $m < \infty$, τ_i are bounded stopping times, $\tau_{i-1} \leq \tau_i$, and $f_{0i}, \tilde{f}_i, g^{ik} \in C_0^\infty$.

Set

$$v(t,x) = \int_0^t g^k(s,x)\,dw_s^k = \sum_{i,k=1}^m g^{ik}(x)(w_{t\wedge\tau_i}^k - w_{t\wedge\tau_{i-1}}^k),$$

$$u(t,x) = v(t,x) + \int_0^t T_{t-s}[\Delta v + f](s,\cdot)(x)\,ds, \quad \forall t \geq 0. \tag{4.8}$$

As easy to see, the function $u - v$ is infinitely differentiable in (t,x) and satisfies the equation

$$\frac{\partial z}{\partial t} = \Delta z + \Delta v + f.$$

It follows that, for any x, the function $u(t,x)$ satisfies almost surely the following form of (4.6):

$$u(t,x) = \int_0^t (\Delta u(s,x) + f(s,x))\,ds + \sum_{k=1}^m \int_0^t g^k(s,x)\,dw_s^k. \tag{4.9}$$

Next, we want to obtain some bounds on norms of u. Let

$$u_1(t,x) = \int_0^t T_{t-s} f(s,x)\,ds.$$

According to Theorem 2.1 (for any ω),

$$||u_{1xx}||_{L_p(\mathbb{R}_+,H_p^{-1})} \le N||f||_{L_p(\mathbb{R}_+,H_p^{-1})}. \tag{4.10}$$

Furthermore, use again that the operators $(\partial/\partial x^i)(1-\Delta)^{-1/2}$ are bounded in L_p for any $p>1$. Then

$$||u_{xx}-u_{1xx}||^p_{\mathbb{H}_p^{-1}} \le N||u_x-u_{1x}||^p_{\mathbb{L}_p} = N\int_0^\infty\int_{\mathbb{R}^d} E|u_x-u_{1x}|^p(t,x)\,dx\,dt. \tag{4.11}$$

To make some further transformations of this formula, we note that if $z^k = z^k(x)$ are bounded Borel functions, then, by Itô's formula applied to the increment over $[0,t]$ of

$$(\int_r^t T_{t-s}z^k\,ds)(w^k_{r\wedge\tau_2}-w^k_{r\wedge\tau_1})$$

(with no summation in k) as a function of r, we obtain (a.s.)

$$0 = -\int_0^t (w^k_{r\wedge\tau_2}-w^k_{r\wedge\tau_1})T_{t-r}z^k\,dr + \int_0^t I_{(\tau_1,\tau_2]}(r)(\int_r^t T_{t-s}z^k\,ds)\,dw^k_r.$$

By using this for our particular g, or by using the stochastic version of the Fubini theorem and coming back to (4.8), for any $t\ge 0$ and $x\in\mathbb{R}^d$, we get

$$u_x(t,x)-u_{1x}(t,x) = v_x(t,x) + \int_0^t T_{t-s}\sum_{k=1}^m\int_0^s \Delta g^k_x(r,x)\,dw^k_r\,ds$$

$$= v_x(t,x) - \sum_{k=1}^m\int_0^t\int_r^t \frac{d}{ds}T_{t-s}g^k_x(r,x)\,ds\,dw^k_r = \sum_{k=1}^m\int_0^t T_{t-r}g^k_x(r,x)\,dw^k_r \quad \text{(a.s.)}.$$

Hence, by the Burkholder–Davis–Gundy inequality,

$$E|u_x-u_{1x}|^p(t,x) \le NE[\int_0^t\sum_{k=1}^m |T_{t-r}g^k_x(r,x)|^2\,dr]^{p/2}$$

$$= NE[\int_0^t |T_{t-r}g_x(r,x)|^2_{l_2}\,dr]^{p/2}.$$

By plugging this into (4.11) and by applying Lemma 4.1, we obtain

$$||u_x-u_{1x}||^p_{\mathbb{L}_p} \le NE\int_0^\infty\int_{\mathbb{R}^d}[\int_0^t |\nabla T_{t-s}g(s,x)|^2_{l_2}\,ds]^{p/2}\,dx\,dt \le N||g||^p_{\mathbb{L}_p(l_2)}.$$

This along with (4.10) gives us (4.3). However, we do not know yet that $u\in\mathcal{H}^1_p$.

Our next step is to prove (4.4) for sufficiently large λ. We repeat briefly the corresponding arguments from [**25**] or [**30**]. From (4.9) and Itô's formula,

$$|u(t,x)|^p e^{-\lambda t}$$

$$= \int_0^t e^{-\lambda s}(p|u|^{p-2}u\Delta u + p|u|^{p-2}uf + \frac{1}{2}p(p-1)|u|^{p-2}|g|^2_{l_2} - \lambda|u|^p)(s,x)\,ds$$

$$+ p\sum_{k\le m}\int_0^t e^{-\lambda s}|u|^{p-2}ug^k(s,x)\,dw^k_s.$$

We integrate here with respect to x and use the stochastic Fubini theorem and the fact that $u(t,x), g(t,x)$, and their derivatives decrease very fast when $|x| \to \infty$. Then we integrate by parts in $\int |u|^{p-2}u\Delta u\,dx$ and also notice that, for $q = p/(p-1)$,

$$\int_{\mathbb{R}^d} |u|^{p-2}uf(s,x)\,dx = -(p-1)\int_{\mathbb{R}^d} |u|^{p-2}u_x \cdot \tilde{f}(s,x)\,dx + \int_{\mathbb{R}^d} |u|^{p-2}uf_0(s,x)\,dx,$$

$$|\int_{\mathbb{R}^d} |u|^{p-2}u_x \cdot \tilde{f}(s,x)\,dx| \leq \int_{\mathbb{R}^d} (|u|^{(p-2)/2}|u_x|)^q |u|^{q(p-2)/2}\,dx$$
$$+ ||\tilde{f}(s,\cdot)||_p^p \leq N||f(s,\cdot)||_{-1,p}^p + N_1||u(s,\cdot)||_p^p + \frac{1}{2}||\,|u|^{(p-2)/p}|u_x|^{2/p}(s,\cdot)||_p^p,$$

$$\int_{\mathbb{R}^d} |u(s,x)|^{p-2}u(s,x)f_0(s,x)\,dx \leq ||f_0(s,\cdot)||_p^p + ||u(s,\cdot)||_p^p$$
$$\leq N||f(s,\cdot)||_{-1,p}^p + ||u(s,\cdot)||_p^p.$$

In this way, for $\lambda \geq p(p-1)N_1 + p + p(p-1)/2$, we get

$$||u(t,\cdot)||_p^p e^{-\lambda t} + \frac{p(p-1)}{2}\int_0^t ||\,|u|^{(p-2)/p}|u_x|^{2/p}(s,\cdot)||_p^p e^{-\lambda s}\,ds$$
$$\leq N\int_0^t [||f(s,\cdot)||_{-1,p}^p + ||g(s,\cdot)||_p^p]e^{-\lambda s}\,ds$$
$$+ p\sum_{k\leq m}\int_0^t e^{-\lambda s}\{\int_{\mathbb{R}^d} |u|^{p-2}ug^k(s,x)\,dx\}\,dw_s^k,$$

where $N = N(p)$. After this, basically, one takes expectations and applies certain standard transformations based on the Burkholder–Davis–Gundy inequalities. For more detail we refer the reader to [**25**] or [**30**].

Finally, the assertion about the arbitrariness of λ in (4.4) can be easily justified by rescaling arguments when instead of f, g, and w one takes $(c^2 f, cg)(c^2 t, cx)$ and $c^{-1}w_{c^2 t}$ and gets $u(c^2 t, cx)$ instead of $u(t,x)$.

From our explicit formulas and from the particular choice of f and g, it also follows that $u \in C_{loc}([0,\infty), H_p^n)$ for any n (and for any ω). This proves (iii).

From estimates (4.3) and (4.4), we conclude that $u \in \bigcap_{T>0}\mathbb{H}_p^1(T)$. Furthermore, from the pointwise equation (4.9) by the stochastic Fubini theorem it follows easily that u solves (4.2) in the sense of Definition 3.1. Hence $u \in \mathcal{H}_p^1$, which proves a part of assertion (i). The uniqueness in (i) follows from the fact that for $f \equiv 0, g \equiv 0$ we have the heat equation and the uniqueness of its solution in our class of functions is a standard fact (we say more about this in the proof of Lemma 4.9). This completes the proof in the case of step functions f, g.

General case. In the case of general f, g, we observe that the uniqueness in (i) is proved as above. As far as other assertions are concerned we are going to use Theorem 3.10 and Remark 3.6.

If we consider all functions $f_0, \tilde{f}^i, g^k$ as one sequence, then, by Theorem 3.10, we can approximate them by functions $f_{0j}, \tilde{f}_j^i, g_j^k$ of type (4.7). Let u_j be the corresponding solutions of (4.6). By the result for the particular case, $\{u_j\}$ is a Cauchy sequence in $\mathcal{H}_p^1$, and, by Theorem 3.7, there is a $u \in \mathcal{H}_p^1$ to which u_j converges in $\mathcal{H}_p^1$. Remark 3.6 and the convergence $||u_{xx} - u_{jxx}||_{\mathbb{H}_p^{-1}} \to 0$ prove that $\mathbb{D}u = \Delta u + f$ and $\mathbb{S}u = g$. In particular, this proves assertion (i).

Assertion (ii) follows from the construction of u. From assertion (iii) available in the particular case, we get that u_j is a Cauchy sequence in $L_p(\Omega, C([0,T], L_p))$ for any T. Therefore, it converges in this space to a function $\bar{u}$. It follows easily that, for any $\phi \in C_0^\infty$,

$$(\bar{u}(t,\cdot),\phi) = \int_0^t \{(\bar{u}(s,\cdot),\Delta\phi) + (f(s,\cdot),\phi)\}\,ds + \int_0^t (g^k(s,\cdot),\phi)\,dw_s^k$$

for all t (a.s.). Therefore, $u - \bar{u}$ is a generalized solution of the heat equation with zero initial condition and with bounded L_p–norm (a.s.). This implies that $||(u-\bar{u})(t,\cdot)||_p = 0$ for all t (a.s.), so that $u \in C([0,T], L_p)$ for all T (a.s.). Finally, we get (4.4) by Fatou's lemma taking into account that

$$\int_0^T \int_{\mathbb{R}^d} |\nabla(u-u_j)|^p\,dxdt = \int_0^T \int_{\mathbb{R}^d} |\nabla(1-\Delta)^{-1/2}(1-\Delta)^{1/2}(u-u_j)|^p\,dxdt$$
$$\le N \int_0^T \int_{\mathbb{R}^d} |(1-\Delta)^{1/2}(u-u_j)|^p\,dxdt \to 0$$

in probability for any T. The theorem is proved.

REMARK 4.3. Although (2.7) holds for all $p \in (1,\infty)$, it follows from [**17**] that Lemma 4.1 is false if $p < 2$.

4.2. Relation of the Solutions of (4.1) to the Solutions of the Heat Equation. It turns out that the investigation of general equation (4.1) with coefficients independent of x can be quite formally reduced to the particular case of the heat equation. First, we explain how to do this without caring about rigorousness, and then proceed with formal proofs.

The first observation consists of the following. Assume that we have

$$du(t,x) = f(t,x)\,dt + g^k(t,x)\,dw_t^k, \tag{4.12}$$

and we define a process x_t and a function v by

$$x_t^i = \int_0^t \sigma^{ik}(s)\,dw_s^k, \quad i = 1,...,d, \quad v(t,x) = u(t, x - x_t). \tag{4.13}$$

Then, by applying formally the Itô–Wentzell formula, we get

$$\begin{aligned} dv(t,x) =& [f(t,x-x_t) + \alpha^{ij}(t) v_{x^i x^j}(t,x) - (g_{x^i}(t,x-x_t), \sigma^i(t))_{l_2}]\,dt \\ &+ [g^k(t,x-x_t) - v_{x^i}(t,x)\sigma^{ik}(t)]\,dw_t^k. \end{aligned} \tag{4.14}$$

This shows how to introduce the terms $v_{x^i}\sigma^{ik}$ in equation (4.12) and also shows again a kind of necessity for g to have the first derivatives in x.

This device alone is not sufficient, since, if we had $\Delta u + \bar{f}$ instead of f in (4.12), then, in (4.14), we would get the second order differential operator $(\delta^{ij} + \alpha^{ij})\partial^2/\partial x^i \partial x^j$ with coefficients strongly related to the coefficients of $v_{x^i}\sigma^{ik}(t)$. We could get around this difficulty if we manage to start with equations with more general operators L instead Δ. Here the second observation comes that if, instead of (4.12), we consider

$$du(t,x) = (\Delta u + \bar{f})\,dt + g^k(t,x)\,dw_t^k,$$

and take expectations in the counterpart of (4.14) corresponding to this equation, then, assuming that σ is nonrandom, we get indeed an equation for $Ev(t,x)$ with operator L different from Δ. By the way, this method of studying parabolic equations with coefficients independent of x was applied in [**18**] in order to show that "whatever" estimate is true for the heat equation, it is also true for any parabolic equation with coefficients independent of

x. Of course, taking expectations "kills" all randomness in the equation, and therefore we use a conditional expectation.

DEFINITION 4.4. Denote by $\mathfrak{D}$ the set of all $\mathcal{D}$-valued functions u (written as $u(t,x)$ in a common abuse of notation) on $\Omega \times [0,\infty)$ such that, for any $\phi \in C_0^\infty$,

(i) the function (u,ϕ) is $\mathcal{P}$–measurable,

(ii) for any $\omega \in \Omega$ and $T \in (0,\infty)$, we have

$$\int_0^T \sup_{x\in\mathbb{R}^d} |(u(t,\cdot),\phi(\cdot - x))|^2\,dt < \infty. \tag{4.15}$$

In the same way, we define $\mathfrak{D}(l_2)$ by considering l_2-valued linear functionals on C_0^∞ and replacing $|\cdot|$ in (4.15) by $|\cdot|_{l_2}$.

REMARK 4.5. Notice that $(u(t,\cdot),\phi(\cdot - x))$ is continuous in x and Borel in t so that (4.15) makes sense. Also, for $p \geq 2$, $q = p/(p-1)$, and any n,

$$\int_0^T \sup_{x\in\mathbb{R}^d} |(u(t,\cdot),\phi(\cdot - x))|^2\,dt \leq \int_0^T ||(u(t,\cdot)||_{n,p}^2 ||\phi||_{-n,q}^2\,dt$$

$$\leq ||\phi||_{-n,q}^2 T^{(p-2)/p} \left(\int_0^T ||(u(t,\cdot)||_{n,p}^p\,dt\right)^{2/p}. \tag{4.16}$$

This shows that if $u \in \mathcal{H}_p^n$, then condition (4.15) is satisfied at least for almost all ω. Also, if $u \in \mathcal{H}_p^n$, then (3.2) holds, which shows that $(u(t,\cdot),\phi)$ is indistinguishable from a predictable process. This is true for any $\phi \in C_0^\infty$. From separability of H_q^{-n}, it follows that we can modify u on a set of probability zero and after this we get a function belonging to $\mathfrak{D}$. This is the sense in which we write

$$\mathcal{H}_p^n \subset \mathfrak{D}. \tag{4.17}$$

DEFINITION 4.6. Let $f,u \in \mathfrak{D}$, $g \in \mathfrak{D}(l_2)$. We say that the equality

$$du(t,x) = f(t,x)\,dt + g(t,x)\,dw_t, \quad t > 0, \tag{4.18}$$

holds *in the sense of distributions* if for any $\phi \in C_0^\infty$, with probability 1 for all $t \geq 0$ we have

$$(u(t,\cdot),\phi) = (u(0,\cdot),\phi) + \int_0^t (f(s,\cdot),\phi)\,ds + \int_0^t (g^k(s,\cdot),\phi)\,dw_s^k. \tag{4.19}$$

Observe that, since $|(g,\phi)(t)|_{l_2}^2$ is locally summable in t, the last series in (4.19) converges uniformly in t on every finite interval of time in probability.

In this subsection, we always understand equation (4.1) in the sense of distributions. Notice that if $u \in \mathcal{H}_p^n$ and u satisfies (4.19), then, by (4.17), $u \in \mathfrak{D}$ and (4.18) holds in the sense of distributions. An advantage of Definition 4.6 is that one need not check summability of any derivative.

LEMMA 4.7. *Let $f,u \in \mathfrak{D}$, $g \in \mathfrak{D}(l_2)$. Assume the definitions in (4.13). Then (4.12) holds (in the sense of distributions) if and only if (4.14) holds (in the sense of distributions).*

Proof. First remember that, for a distribution $\alpha(x)$ and $y \in \mathbb{R}^d$, by $\alpha(x-y)$ we mean the distribution defined by $(\alpha,\phi(\cdot + y))$. Also from relations like (cf. (4.16))

$$\int_0^T \sup_{y\in\mathbb{R}^d} |(v_{xx}(t,\cdot),\phi(\cdot - y))|^2\,dt = \int_0^T \sup_{y\in\mathbb{R}^d} |(v(t,\cdot),\phi_{xx}(\cdot - y))|^2\,dt$$

$$= \int_0^T \sup_{y\in\mathbb{R}^d} |(u(t,\cdot),\phi_{xx}(\cdot+x_t-y))|^2\,dt = \int_0^T \sup_{y\in\mathbb{R}^d} |(u(t,\cdot),\phi_{xx}(\cdot-y))|^2\,dt < \infty,$$

$$\int_0^T \sup_{y\in\mathbb{R}^d} |((g_{x^i}(t,\cdot-x_t),\sigma^i(t))_{l_2},\phi(\cdot-y))|^2\,dt$$

$$= \int_0^T \sup_{y\in\mathbb{R}^d} |((g_{x^i}(t,\cdot-x_t),\phi(\cdot-y)),\sigma^i(t))_{l_2}|^2\,dt$$

$$\le \int_0^T |\sigma^i(t)|_{l_2}^2\,dt \int_0^T \sup_{y\in\mathbb{R}^d} |(g_{x^i}(t,\cdot-x_t),\phi(\cdot-y))|_{l_2}^2\,dt < \infty$$

it follows that $v(t,x)$, $f(t,x-x_t)$ and $(g_{x^i}(t,x-x_t),\sigma^i(t))_{l_2}$ belong to $\mathfrak{D}$ and $g(t,x-x_t)$ and $v_{x^i}(t,x)\sigma^i(t)$ belong to $\mathfrak{D}(l_2)$. Furthermore, for any $\phi\in C_0^\infty$, the function $F(t,x) := (u(t,\cdot-x),\phi)$ has a stochastic differential in t for any x and is infinitely differentiable with respect to x. Now our assertion immediately follows from the Itô–Wentzell formula for $F(t,x_t)$. By the way, one can easily memorize (and perhaps prove) this formula by considering the following computations:

$$d\int_{\mathbb{R}^d} u(t,x-x_t)\phi(x)\,dx = d\int_{\mathbb{R}^d} u(t,x)\phi(x+x_t)\,dx = \int_{\mathbb{R}^d}\phi(x+x_t)du(t,x)\,dx$$

$$+\int_{\mathbb{R}^d} u(t,x)d\phi(x+x_t)\,dx + \int_{\mathbb{R}^d}(du(t,x))d\phi(x+x_t)\,dx$$

$$= \left\{\int_{\mathbb{R}^d}\phi(x+x_t)f(t,x)\,dx + \int_{\mathbb{R}^d} u(t,x)\alpha^{ij}(t)\phi_{x^ix^j}(x+x_t)\,dx\right.$$

$$\left. + \int_{\mathbb{R}^d} g^k(t,x)\phi_{x^i}(x+x_t)\sigma^{ik}(t)\,dx\right\}dt$$

$$+\left\{\int_{\mathbb{R}^d}\phi(x+x_t)g^k(t,x)\,dx + \int_{\mathbb{R}^d} u(t,x)\phi_{x^i}(x+x_t)\sigma^{ik}(t)\,dx\right\}dw_t^k$$

$$= \int_{\mathbb{R}^d}\left\{f(t,x-x_t)+\alpha^{ij}(t)u_{x^ix^j}(t,x-x_t) - g^k_{x^i}(t,x-x_t)\sigma^{ik}(t)\right\}\phi(x)\,dx\,dt$$

$$+\int_{\mathbb{R}^d}\left\{g^k(t,x-x_t) - u_{x^i}(t,x-x_t)\sigma^{ik}(t)\right\}\phi(x)\,dx\,dw_t^k.$$

The lemma is proved.

REMARK 4.8. If, instead of (4.12), u satisfies the equation

$$du(t,x) = (a^{ij}(t)u_{x^ix^j}(t,x)+h(t,x))\,dt + \sigma^{ik}(t)u_{x^i}(t,x)\,dw_t^k,$$

then (4.14) takes the form

$$\frac{\partial}{\partial t}v(t,x) = ((a^{ij}(t)-\alpha^{ij}(t))v_{x^ix^j}(t,x) + h(t,x-x_t) \quad t>0, \tag{4.20}$$

and can be considered on each ω separately. Observe that if $a(t)<\alpha(t)$, then the initial–value problem $v(0)=v_0$ for equation (4.20) is *ill posed*.

This shows that operators appearing in the stochastic term should be subordinated in a certain sense to the operator in the deterministic part of equation. This is needed in order to construct an L_p–theory. On the other hand, take $d=1$ and a one-dimensional Wiener process w_t. Consider the following equation

$$du(t,x) = iu'_x(t,x)\,dw_t.$$

Surprisingly enough and somewhat in spite of what is said above, this equation has a very nice solution for each initial data $u_0\in L_2$. One gets the solution after passing to Fourier transforms. It turns out that $\tilde{u}(t,\xi) = u_0(\xi)\exp\{\xi w_t - (1/2)|\xi|^2 t\}$. The function $\tilde{u}(t,\xi)$ decays vary fast when $|\xi| \to \infty$, which shows that $u(t,x)$ is infinitely differentiable in x. Also notice that, taking expectations, we see that $Eu(t,x) = u_0(x)$ if u_0 is nonrandom, and in this case we get a representation of any L_2 function as an integral over Ω of functions $u(\omega,1,x)$ which are infinitely differentiable in x. However, the major drawback of such equations is that $E|u(t,0)|^p = \infty$ for any $p > 1$ if, for example, $\tilde{u}_0(\xi) \geq \exp\{-\lambda|\xi|\}$, where λ is a constant.

LEMMA 4.9. *Let $f \in \mathfrak{D}$, $g \in \mathfrak{D}(l_2)$, u_0 be a $\mathcal{D}$-valued function on Ω. Then the following assertions hold:*

(i) In $\mathfrak{D}$ there can exist only one (up to evanescence) solution of equation (4.1) with the initial condition $u(0,\cdot) = u_0$.

(ii) Let $\mathcal{F}_t = \mathcal{W}_t \vee \mathcal{B}_t$ for $t \geq 0$, and assume that σ–fields $\mathcal{W}_t$ and $\mathcal{B}_t$ form independent increasing filtrations. Let W and B be sets such that $W \cup B = \{1,2...\}$. Assume that $(w_t^k, \mathcal{W}_t)$ and $(w_t^r, \mathcal{B}_t)$ are Wiener processes for $k \in W$, $r \in B$. Let $u \in \mathfrak{D}$ satisfy equation (4.1) (in the sense of distributions), and let a, f, σ, g be $\mathcal{W}_t$–adapted. Finally, assume that there exists an $n \in (-\infty,\infty)$ such that $f \in \mathbb{H}_2^n(T)$, $g \in \mathbb{H}_2^n(T,l_2)$ for any $T \in (0,\infty)$ and $u(0,\cdot)$ is W_0–measurable and

$$E||u(0,\cdot)||_{n,2}^2 < \infty.$$

Then in $\mathfrak{D}$ there exists a unique solution $\tilde{u}$ of the equation

$$d\tilde{u} = (a^{ij}\tilde{u}_{x^ix^j} + f)\,dt + \sum_{k\in W}(\sigma^{ik}\tilde{u}_{x^i} + g^k)\,dw_t^k, \quad t > 0. \tag{4.21}$$

In addition, for any $\phi \in C_0^\infty$ and $t \geq 0$,

$$(\tilde{u}(t,\cdot),\phi) = E\{(u(t,\cdot),\phi)|\mathcal{W}_t\} \qquad (a.s.). \tag{4.22}$$

Proof. (i) As always, we can take $f \equiv 0, g \equiv 0$, and $u_0 \equiv 0$, and, by Lemma 4.7, it suffices to consider only the case $\sigma \equiv 0$. In this case, for any given $\phi \in C_0^\infty$ we have

$$(u(t,\cdot),\phi) = \int_0^t (u(s,\cdot), L(s)\phi)\,ds, \quad t \geq 0,$$

almost surely. Putting here $\phi(\cdot - x)$ instead of ϕ and observing that both sides are continuous and bounded in (t,x) on $[0,T] \times \mathbb{R}^d$ for any $T < \infty$ (cf. (4.15)), we get that the function $F(t,x) := (u(t,\cdot),\phi(\cdot - x))$ is bounded in (t,x) on $[0,T] \times \mathbb{R}^d$ for any $T < \infty$, infinitely differentiable in x, and satisfies the equation

$$F(t,x) = \int_0^t L(s)F(s,x)\,ds \qquad \forall t,x$$

(a.s.). From the theory of parabolic equations, it follows that $F(t,x) = 0$ for all t,x (a.s.). This means that $(u(t,\cdot),\phi) = 0$ for all t (a.s.). Now take ϕ with unit integral. Then for any x and n with probability 1, $(u(t,\cdot), n^d\phi(n(\cdot - x))) = 0$ for all t. Since this function is continuous in x, we have $(u(t,\cdot), n^d\phi(n(\cdot - x))) = 0$ for all t and x with probability 1. Finally, $(u(t,\cdot), n^d\phi(n(\cdot - x))) \to u(t,x)$ as $n \to \infty$ for all (ω,t,x) in the sense of distributions, which implies that, with probability 1, we have $u(t,\cdot) = 0$ for all t as stated.

(ii) First, notice that, according to [**30**], equation (4.1) has a unique solution v in the space $\mathbb{H}_2^{n+1}(T)$ for any T. The definition of solutions in $\mathbb{H}_2^{n+1}(T)$ from [**30**] is slightly different, but v is continuous (a.s.) as an H_2^n-valued process and

$$E\sup_{t\leq T}||v(t,\cdot)||_{n,2}^2 < \infty \qquad \forall T < \infty, \tag{4.23}$$

so that v is a $\mathfrak{D}$-solution of (4.1). It follows from (i) that our function u coincides with v and therefore belongs to $\mathbb{H}_2^{n+1}(T)$ for any T, and (4.23) holds for u. Furthermore, with probability 1 for all t at once,

$$u(t) = u(0) + \int_0^t [a^{ij}(s)u_{x^ix^j}(s) + f(s)]\,ds + \int_0^t [\sigma^{ik}(s)u_{x^i}(s) + g(s)]\,dw_s^k,$$

where all integrals are taken in the sense of the Hilbert space H_2^{n-1} (see [**30**]). By Theorem 1.4.7 of [**30**], or rather by its Hilbert-space counterpart, there exists an H_2^{n+1}-valued , $\mathcal{W}_t$-predictable function $\bar{u}(t)$ such that, for almost any t, we have

$$\bar{u}(t) = E\{u(t)|\mathcal{W}_t\}, \quad \bar{u}_x(t) = E\{u_x(t)|\mathcal{W}_t\}, \quad \bar{u}_{xx}(t) = E\{u_{xx}(t)|\mathcal{W}_t\} \quad \text{(a.s.)}$$

(conditional expectations of Hilbert–space valued random elements) and

$$\bar{u}(t) = u(0) + \int_0^t [a^{ij}(s)\bar{u}_{x^ix^j}(s) + f(s)]\,ds + \sum_{k\in W}\int_0^t [\sigma^{ik}(s)\bar{u}_{x^i}(s) + g^k(s)]\,dw_s^k \tag{4.24}$$

for almost all t and ω. The right-hand side here is a continuous H_2^{n-1}-valued process which we denote by $\tilde{u}$ and we show that $\tilde{u}$ is the function we need.

By definition and by the equality $\bar{u} = \tilde{u}$ (a.e.), $\tilde{u}$ satisfies (4.24) for all t with probability 1 and also is a continuous process in H_2^{n-1}. This implies that $\tilde{u} \in \mathfrak{D}$ and $\tilde{u}$ is a solution of (4.21). To prove (4.22) for any t, it remains only to observe that, again by Theorem 1.4.7 of [**30**], the conditional expectation $E\{u(t)|\mathcal{W}_t\}$ for any t is equal to the right-hand side of (4.24) almost surely. The lemma is proved.

4.3. General Equation (4.1) with Coefficients Independent of x.

THEOREM 4.10. *Take $n \in \mathbb{R}$ and let $f \in \mathbb{H}_p^{n-1}$, $g \in \mathbb{H}_p^n(l_2)$. Then*
(i) equation (4.1) with zero initial condition has a unique solution $u \in \mathcal{H}_p^{n+1}$;
(ii) for this solution, we have

$$||u_{xx}||_{\mathbb{H}_p^{n-1}} \le N(||f||_{\mathbb{H}_p^{n-1}} + ||g||_{\mathbb{H}_p^n(l_2)}), \quad ||u||_{\mathcal{H}_p^{n+1}} \le N||(f,g)||_{\mathcal{F}_p^{n-1}}, \tag{4.25}$$

where $N = N(d,p,\delta,K)$;
(iii) we have $u \in C_{loc}([0,\infty), H_p^n)$ almost surely and for any $\lambda, T > 0$,

$$E\sup_{t\le T}(e^{-p\lambda t}||u(t,\cdot)||_{n,p}^p) \le N(||e^{-\lambda t}f||_{\mathbb{H}_p^{n-1}(T)}^p + ||e^{-\lambda t}g||_{\mathbb{H}_p^n(T,l_2)}^p), \tag{4.26}$$

where $N = N(d,p,\delta,K,\lambda)$.

Proof. Since one can apply the operator $(I-\Delta)^{n/2}$ to both sides of (4.1), it suffices to prove the theorem only for $n = 0$. Furthermore, as we have already noticed in (4.17), any function $u \in \mathcal{H}_p^1$ also belongs to $\mathfrak{D}$. This and Lemma 4.9 prove the uniqueness in (i). Also, the fact that our norms are translation invariant, combined with Lemma 4.7, shows that, to prove the existence in (i) and all other assertions of the theorem, we only need to consider the case $\sigma \equiv 0$. As in the proof of Theorem 4.2, we can assume that f and g are as in (4.7). In this case, as we know from [**25**] or [**30**], equation (4.1) has a unique $\mathfrak{D}$–solution u that belongs to $C_b([0,T]\times\mathbb{R}^d)$ and $C([0,T],L_2)$ almost surely for any $T < \infty$. It follows that $u \in C([0,T],L_p)$ almost surely for any $T < \infty$. Estimate (4.26) also follows from [**25**] or [**30**] as in the proof of Theorem 4.2. It remains only to prove that $u \in \mathcal{H}_p^1$ and that (4.25) holds. Since u is a $\mathfrak{D}$–solution, to prove that $u \in \mathcal{H}_p^1$, it suffices to prove that $u \in \mathbb{H}_p^1(T)$ for any $T < \infty$.

Since the matrix a is uniformly non-degenerate, by making a nonrandom time change, we can reduce the general case to the case $a \ge I$. In this case, define the matrix-valued function $\bar\sigma(t) = \bar\sigma^*(t) \geq 0$ as a solution of the equation $\bar\sigma^2(t) + 2I = 2a(t)$. Furthermore, without loss of generality, we assume that on our probability space we are also given a d-dimensional Wiener process B_t independent of $\mathcal{F}_t$.

Now, consider the equation

$$dv(t,x) = [\Delta v(t,x) + f(t, x - \int_0^t \bar\sigma(s)\,dB_s)]\,dt + g^k(t, x - \int_0^t \bar\sigma(s)\,dB_s)\,dw_t^k \tag{4.27}$$

with zero initial condition. Replace the predictable σ–field $\mathcal{P}$ with predictable σ–field generated by $\mathcal{F}_t \vee \sigma(B_s; s \leq t)$. Then the spaces $\mathcal{H}_p^n$ become larger. By Theorem 4.2 there is a solution v of (4.27) possessing properties (i) through (iii) listed in Theorem 4.2 (with new $\mathcal{H}_p^1$). Use again that, after changing, if necessary, v on a set of probability zero, the function v becomes a $\mathfrak{D}$–solution of (4.27). Then, by Lemma 4.7, the function $z(t,x) := v(t, x + \int_0^t \bar\sigma(s)\,dB_s)$ is a $\mathfrak{D}$–solution of

$$dz(t,x) = (a^{ij}(t) z_{x^ix^j}(t,x) + f(t,x))\,dt + g^k(t,x)\,dw_t^k + z_{x^i}(t,x)\bar\sigma^{ij}(t)\,dB_t^j,$$

and by Lemma 4.9 there is a solution $\tilde u \in \mathfrak{D}$ of

$$d\tilde u(t,x) = (a^{ij}(t)\tilde u_{x^ix^j}(t,x) + f(t,x))\,dt + g^k(t,x)\,dw_t^k,$$

which is (4.1) in our case. In addition, for any $\phi \in C_0^\infty$ and $t \geq 0$,

$$(\tilde u(t,\cdot),\phi) = E\{(z(t,\cdot),\phi)|\mathcal{F}_t\} = E\{(v(t,\cdot + \int_0^t \bar\sigma(s)\,dB_s),\phi)|\mathcal{F}_t\} \quad \text{(a.s.)}.$$

In particular, it follows from this equality that $\tilde u$ is a $\mathfrak{D}$–solution with respect to the initial predictable σ–field $\mathcal{P}$, and from uniqueness we get $\tilde u = u$. Therefore,

$$(u(t,\cdot),\phi) = E\{(v(t,\cdot + \int_0^t \bar\sigma(s)\,dB_s),\phi)|\mathcal{F}_t\} \quad \text{(a.s.)}. \tag{4.28}$$

It follows that

$$|(u(t,\cdot),\phi)|^p \leq E\{||v(t,\cdot)||_{1,p}^p|\mathcal{F}_t\}||\phi||_{-1,q}^p \tag{4.29}$$

(a.s.) for any $\phi \in C_0^\infty$ and $t \geq 0$, where $q = p/(p-1)$.

Next, take a countable family $\Phi \subset C_0^\infty$ dense in C_0^∞. Observe that, given a distribution ψ, we have $\psi \in H_p^1$ if and only if for any $\phi \in \Phi$ we have $|(\psi,\phi)| \leq N||\phi||_{-1,q}$ with a constant N independent of ϕ. Indeed, in this case there exists a bounded linear functional ℓ on H_q^{-1} such that $\ell(\phi) = (\psi,\phi)$ for any $\phi \in \Phi$. Since $\ell(\phi) = ((1-\Delta)^{-1/2}h,\phi)$ with an $h \in L_p$ and Φ is dense in C_0^∞, we have $\psi = (1-\Delta)^{-1/2}h \in H_p^1$ indeed. In particular, this fact implies that the set $\{(\omega,t) : w(\omega,t,\cdot) \in H_p^1\}$ is measurable (even predictable) for any $w \in \mathfrak{D}$, say for $w = u$.

We also know that $v \in \mathcal{H}_p^1$, which implies that $E||v(t,\cdot)||_{1,p}^p < \infty$ for almost all t. Fix such a t. Then there exists a set Ω' of probability 1 such that $E\{||v(t,\cdot)||_{1,p}^p|\mathcal{F}_t\} < \infty$ on Ω' and (4.29) holds for all $\omega \in \Omega'$ and $\phi \in \Phi$. Hence, $u(t,\cdot) \in H_p^1$ for the chosen t and all $\omega \in \Omega'$. In particular, $u(t,\cdot) \in H_p^1$ for almost all (ω,t), and from (4.29) it follows that

$$||u(t,\cdot)||_{1,p}^p \leq E\{||v(t,\cdot)||_{1,p}^p|\mathcal{F}_t\} \quad \text{(a.s.)}, \quad ||u||_{\mathbb{H}_p^1(T)} \leq ||v||_{\mathbb{H}_p^1(T)} < \infty.$$

Thus $u \in \mathbb{H}_p^1(T)$ for any $T < \infty$ indeed and $u \in \mathcal{H}_p^1$.

Similarly, from the equality

$$(u_{xx}(t,\cdot),\phi) = E\{(v_{xx}(t,\cdot + \int_0^t \bar\sigma(s)\,dB_s),\phi)|\mathcal{F}_t\} \quad \text{(a.s.)},$$

one gets

$$||u_{xx}(t,\cdot)||^p_{-1,p} \le E\{||v_{xx}(t,\cdot)||^p_{-1,p}|\mathcal{F}_t\} \quad \text{(a.s.)}.$$

This and the properties of v immediately yield (4.25). The theorem is proved.

REMARK 4.11. By using the self-similarity of equation (4.1), it is possible to obtain further estimates from estimates like (4.26) . For instance, remembering that $H^1_p = W^1_p$, one sees that, for $n = 1$ and $\lambda = 1/p$, estimate (4.26) implies that

$$E\sup_{t\le T}\{||u_x(t,\cdot)||^p_p + ||u(t,\cdot)||^p_p\} \le N(d,p,\delta,K)e^T(||f||^p_{\mathbb{L}_p(T)} + ||g_x||^p_{\mathbb{L}_p(T,l_2)} + ||g||^p_{\mathbb{L}_p(T,l_2)}).$$

Let us take a constant $c > 0$ and consider $(c^2 f, cg)(c^2 t, cx), c^{-1}w_{c^2 t}$, and $u(c^2 t, cx)$ instead of f, g, w, and u. Then, from the last estimate, we get

$$\begin{aligned}
&E\sup_{t\le T}\{c^{p-d}||u_x(c^2t,\cdot)||^p_p + c^{-d}||u(c^2t,\cdot)||^p_p\}\\
&\qquad\le Ne^T(c^{2p-(d+2)}||f||^p_{\mathbb{L}_p(c^2T)} + c^{2p-(d+2)}||g_x||^p_{\mathbb{L}_p(c^2T,l_2)} + c^{p-(d+2)}||g||^p_{\mathbb{L}_p(c^2T,l_2)}),
\end{aligned}$$

the constant N being the same as above. It follows that, for $c, T \ge 0$,

$$\begin{aligned}
&E\sup_{t\le T}\{||u_x(t,\cdot)||^p_p + c^{-p}||u(t,\cdot)||^p_p\}\\
&\le Ne^{T/c^2}c^{p-2}(||f||^p_{\mathbb{L}_p(T)} + ||g_x||^p_{\mathbb{L}_p(T,l_2)} + c^{-p}||g||^p_{\mathbb{L}_p(T,l_2)}).
\end{aligned}$$

Upon setting $c^2 = T$ and considering $(1-\Delta)^{(n-1)/2}u$ instead of u, we conclude that, under the conditions of Theorem 4.2, for any $T > 0$,

$$\begin{aligned}
&E\{\sup_{t\le T}||u_x(t,\cdot)||^p_{n-1,p} + T^{-p/2}\sup_{t\le T}||u(s,\cdot)||^p_{n-1,p}\} \le\\
&N(d,p,\delta,K)T^{(p-2)/2}(||f||^p_{\mathbb{H}^{n-1}_p(T)} + ||g_x||^p_{\mathbb{H}^{n-1}_p(T,l_2)} + T^{-p/2}||g||^p_{\mathbb{H}^{n-1}_p(T,l_2)}).
\end{aligned}$$

We will later prove a much deeper estimate than (4.26).

5. Equations with Variable Coefficients

Here we are going to state main results concerning L_p–theory of SPDEs in the whole space, the most important of which, Theorem 5.1, will be proved in Sec. 6. Take a stopping time $\tau \le T$ with T being a finite constant. Fix $n \in (-\infty, \infty)$ and fix a number $\gamma \in [0, 1)$ such that $\gamma = 0$ if $n = 0, \pm1, \pm2, ...$; otherwise $\gamma > 0$ and is such that $|n| + \gamma$ is not an integer. Define

$$B^{|n|+\gamma} = \begin{cases} B(\mathbb{R}^d) & \text{if} \quad n = 0,\\ C^{|n|-1,1}(\mathbb{R}^d) & \text{if} \quad n = \pm1, \pm2, ...,\\ C^{|n|+\gamma}(\mathbb{R}^d) & \text{otherwise;}\end{cases}$$

$$B^{|n|+\gamma}(l_2) = \begin{cases} B(\mathbb{R}^d, l_2) & \text{if} \quad n = 0,\\ C^{|n|-1,1}(\mathbb{R}^d, l_2) & \text{if} \quad n = \pm1, \pm2, ...,\\ C^{|n|+\gamma}(\mathbb{R}^d, l_2) & \text{otherwise,}\end{cases}$$

where $B(\mathbb{R}^d)$ is the Banach space of bounded functions on $\mathbb{R}^d$, $C^{|n|-1,1}(\mathbb{R}^d)$ is the Banach space of $|n| - 1$ times continuously differentiable functions whose derivatives of $(|n| - 1)$st order satisfy the Lipschitz condition on $\mathbb{R}^d$, $C^{|n|+\gamma}(\mathbb{R}^d)$ is the usual Hölder space, and l_2 means that instead of real–valued functions we consider l_2–valued ones.

Consider the following nonlinear equation on $[0, \tau]$:

$$\begin{aligned}
du(t,x) =&[a^{ij}(t,x)u_{x^ix^j}(t,x) + f(u,t,x)]\,dt\\
&+ [\sigma^{ik}(t,x)u_{x^i}(t,x) + g^k(u,t,x)]\,dw^k_t,
\end{aligned} \tag{5.1}$$

where a^{ij} and f are real-valued, and σ^i and g are l_2-valued functions defined for $\omega \in \Omega$, $t \geq 0$, $x \in \mathbb{R}^d$, $u \in H_p^{n+2}$, $i,j = 1,...,d$. We consider this equation in the sense of Definition 3.5 (where we take $n+2$ instead of n).

We make the following assumptions, where, as in Section 4, we define

$$\alpha^{ij}(t,x) = \frac{1}{2}(\sigma^i(t,x), \sigma^j(t,x))_{l_2}. \tag{5.2}$$

ASSUMPTION 5.1 (coercivity). For any $\omega \in \Omega$, $t \geq 0$, $x, \lambda \in \mathbb{R}^d$, we have

$$K|\lambda|^2 \geq [a^{ij}(t,x) - \alpha^{ij}(t,x)]\lambda^i\lambda^j \geq \delta|\lambda|^2,$$

where K, δ are fixed strictly positive constants.

ASSUMPTION 5.2 (uniform continuity of a and σ). For any $\varepsilon > 0, i, j$, there exists a $\kappa_\varepsilon > 0$ such that

$$|a^{ij}(t,x) - a^{ij}(t,y)| + |\sigma^i(t,x) - \sigma^i(t,y)|_{l_2}^2 \leq \varepsilon \tag{5.3}$$

whenever $|x-y| \leq \kappa_\varepsilon, t \geq 0, \omega \in \Omega$.

This assumption is actually used only if $n = 0$, and even then we need a stronger condition on σ. For other values of n we impose stronger conditions on a and σ.

ASSUMPTION 5.3. For any i,j,k, the functions $a^{ij}(t,x)$ and $\sigma^{ik}(t,x)$ are real-valued $\mathcal{P} \times \mathcal{B}(\mathbb{R}^d)$–measurable functions, and for any $\omega \in \Omega$ and $t \geq 0$, we have

$$a^{ij}(t,\cdot) \in B^{|n|+\gamma}, \quad \sigma^i(t,\cdot) \in B^{|n+1|+\gamma}(l_2).$$

ASSUMPTION 5.4. For any $u \in H_p^{n+2}$, the functions $f(u,t,x)$ and $g(u,t,x)$ are predictable as functions taking values in H_p^n and $H_p^{n+1}(\mathbb{R}^d, l_2)$, respectively.

ASSUMPTION 5.5. For any $t \geq 0, \omega, i, j$,

$$||a^{ij}(t,\cdot)||_{B^{|n|+\gamma}} + ||\sigma^i(t,\cdot)||_{B^{|n+1|+\gamma}(l_2)} \leq K, \quad (f(0,\cdot,\cdot), g(0,\cdot,\cdot)) \in \mathcal{F}_p^n(\tau).$$

ASSUMPTION 5.6. The functions f, g are continuous in u. Moreover, for any $\varepsilon > 0$, there exists a constant K_ε such that, for any $u, v \in H_p^{n+2}$, t, ω, we have

$$||f(u,t,\cdot) - f(v,t,\cdot)||_{n,p} + ||g(u,t,\cdot) - g(v,t,\cdot)||_{n+1,p} \leq \varepsilon||u-v||_{n+2,p} + K_\varepsilon||u-v||_{n,p}. \tag{5.4}$$

THEOREM 5.1. *Let Assumptions 5.1–5.6 be satisfied and let*

$$u_0 \in L_p(\Omega, \mathcal{F}_0, H_p^{n+2-2/p}).$$

Then the Cauchy problem for equation (5.1) on $[0,\tau]$ with the initial condition $u(0,\cdot) = u_0$ has a unique solution $u \in \mathcal{H}_p^{n+2}(\tau)$. For this solution, we have

$$||u||_{\mathcal{H}_p^{n+2}(\tau)} \leq N\{||f(0,\cdot,\cdot)||_{\mathbb{H}_p^n(\tau)} + ||g(0,\cdot,\cdot)||_{\mathbb{H}_p^{n+1}(\tau,l_2)} + (E||u_0||_{n+2-2/p,p}^p)^{1/p}\},$$

where the constant N depends only on $d, n, \gamma, p, \delta, K, T$, and the functions κ_ε and K_ε.

To discuss the theorem, we need the following lemma.

LEMMA 5.2. *Let $\zeta \in C_0^\infty$ be a nonnegative function such that $\int \zeta(x)\,dx = 1$ and define $\zeta_k(x) = k^d\zeta(kx)$, $k = 1,2,3,...$. We assert that, for any $u \in H_p^n$, we have the following:*

(i) $||au||_{n,p} \leq N\bar{a}||u||_{n,p}$, where $\bar{a} = ||a||_{B^{|n|+\gamma}}$ and the constant N depends only on d, p, n, and γ;

(ii) $||u * \zeta_k||_{n,p} \leq ||u||_{n,p}$, $||u - u * \zeta_k||_{n,p} \to 0$.

The same assertions hold true for Banach–space valued a with natural definition of $\bar{a}$.

Proof. If $n \neq 0, \pm 1, \pm 2, \ldots$ (and $\gamma > 0$), then one gets (i) by Corollary 2.8.2 (ii) of [**33**]. If n is a nonnegative integer, then (i) follows from the Leibnitz rule (and the fact that $H_p^n = W_p^n$). For negative integers (and generally for negative n) (i) follows easily by duality, that is, by using the fact that if $u = (1-\Delta)^{-n/2} f$, then

$$(au, \phi) = (f, (1-\Delta)^{-n/2}(a\phi)) \leq ||f||_p \, ||a\phi||_{-n,q} \qquad q = p/(p-1).$$

As for (ii), the first inequality follows from Minkowski's inequality and the second one is derived as usual owing to denseness of C_0^∞ in H_p^n. The lemma is proved.

REMARK 5.3. As we have said above, by solution to the Cauchy problem for equation (5.1) on $[0, \tau]$ with the given initial condition u_0, we understand a function $u \in \mathcal{H}_p^{n+2}(\tau)$ such that for any test function $\phi \in C_0^\infty$, one has almost surely

$$\begin{aligned}(u(t,\cdot), \phi) = (u_0, \phi) + \int_0^t (a^{ij}(s,\cdot) u_{x^i x^j}(s,\cdot) + f(u,s,\cdot), \phi)\, ds \\ + \int_0^t (\sigma^{ik}(s,\cdot) u_{x^i}(s,\cdot) + g^k(u,s,\cdot), \phi)\, dw_s^k, \quad \forall t \in [0,\tau].\end{aligned}$$

It is important that, under our assumptions, the equation makes sense for $u \in \mathcal{H}_p^{n+2}(\tau)$, since by Lemma 5.2 we have $a^{ij} u_{x^i x^j} \in H_p^n$, $\sigma^i u_{x^i} \in H_p^{n+1}(\mathbb{R}^d, l_2)$ whenever $u \in H_p^{n+2}$.

REMARK 5.4. Two main ideas in the proof of this theorem are quite standard. The first one, reduction to equations with constant coefficients, will be seen very clearly. The second one, which is somewhat hidden, consists of introducing the new unknown function $v = (1-\Delta)^{n/2} u$, which reduces the case of general n to the case $n = 0$. The function v satisfies

$$\begin{aligned}dv = \Big\{(1-\Delta)^{n/2}(a^{ij}(1-\Delta)^{-n/2} v_{x^i x^j}) + (1-\Delta)^{n/2} f\Big\}\, dt \\ + \Big\{(1-\Delta)^{n/2}(\sigma^{ik}(1-\Delta)^{-n/2} v_{x^i}) + (1-\Delta)^{n/2} g^k\Big\}\, dw_t^k.\end{aligned}$$

This is a pseudo–differential equation, and we note that more general pseudo–differential equations can be considered too. Also, this equations shows a need to have smoothness assumptions on a, σ in x if we are interested in $n \neq 0$ both positive or negative.

REMARK 5.5. By Theorem 14.2 of [**14**], for any $u \in H_p^{n+2}$ and $m \in [n, n+2]$, we have

$$||u||_{m,p} \leq N ||u||_{n+2,p}^\theta ||u||_{n,p}^{1-\theta} \leq N\theta\varepsilon ||u||_{n+2,p} + N(1-\theta)\varepsilon^{-\theta/(1-\theta)} ||u||_{n,p},$$

where $\theta = (m-n)/2$ and N depends only on d, n, m, and p. This shows that the right–hand side in (5.4) can be replaced by $||u - v||_{n+\varepsilon+1,p}$ once $|\varepsilon| < 1$. As an example, one can take $f = f_0(x) \sup_x |u_x|$ if $(n+1)p > d$ and $f_0 \in H_p^n$. Indeed, by Sobolev's embedding theorems, $H_p^{n+1+\varepsilon} \subset C^1$ if $(n+\varepsilon)p > d$. Therefore,

$$||f(u,t,\cdot) - f(v,t,\cdot)||_{n,p} \leq ||f_0||_{n,p} \sup_x |(u-v)_x| \leq N ||u-v||_{n+1+\varepsilon,p}.$$

REMARK 5.6. A typical application of Theorem 5.1 occurs when $f(u,t,x) = b^i(t,x) u_{x^i} + c(t,x) u + f(t,x)$ and $g(u,t,x) = \nu(t,x) u + g(t,x)$, so that (5.1) becomes

$$du = (a^{ij} u_{x^i x^j} + b^i u_{x^i} + cu + f)\, dt + (\sigma^{ik} u_{x^i} + \nu^k u + g^k)\, dw_t^k. \tag{5.5}$$

To describe the appropriate assumptions, we take $\varepsilon \in (0,1)$ and denote

$$\begin{array}{llll} n_b = n + \gamma & \text{if} \quad n \geq 0, & n_b = 0 & \text{if} \quad n \in (-1, 0], \\ n_\nu = n + 1 + \gamma & \text{if} \quad n \geq -1, & n_\nu = 0 & \text{if} \quad n \in (-2, -1], \\ n_c = n + \gamma & \text{if} \quad n \geq 0, & n_c = 0 & \text{if} \quad n \in (-2, 0], \end{array}$$

$$n_b = -n - 1 + \varepsilon \quad \text{if} \quad n \leq -1,$$
$$n_\nu = -n - 2 + \varepsilon \quad \text{if} \quad n \leq -2,$$
$$n_c = -n - 2 + \varepsilon \quad \text{if} \quad n \leq -2.$$

Assume that b, c, and ν are appropriately measurable and

$$b^i(t, \cdot) \in B^{n_b}, \quad c(t, \cdot) \in B^{n_c}, \quad \nu(t, \cdot) \in B^{n_\nu}(\mathbb{R}^d, l_2),$$
$$f(t, \cdot) \in H_p^n, \quad g(t, \cdot) \in H_p^{n+1}(\mathbb{R}^d, l_2).$$
$$||b^i(t, \cdot)||_{B^{n_b}} + ||c(t, \cdot)||_{B^{n_c}} + ||\nu(t, \cdot)||_{B^{n_\nu}(\mathbb{R}^d, l_2)} \leq K, \quad (f(\cdot, \cdot), g(\cdot, \cdot)) \in \mathcal{F}_p^n(\tau).$$

It turns out then that the assumptions of Theorem 5.1 about $f(u, t, x)$ and $g(u, t, x)$ are satisfied. To show this, it suffices to apply Remark 5.5 and to notice that, for instance, if $n \geq -1$, then $||\nu u||_{n+1,p} \leq N||u||_{n+1,p}$ by Lemma 5.2; if $n \in (-2, -1]$, then obviously $||\nu u||_{n+1,p} \leq ||\nu u||_p \leq N||u||_p = N||u||_{n+1+(-n-1),p}$, and $-n - 1 \in [0, 1)$; if $n \leq -2$, then Lemma 5.2 yields $||\nu u||_{n+1,p} \leq ||\nu u||_{n+2-\varepsilon_1,p} \leq N||u||_{n+2-\varepsilon_1,p}$, where $\varepsilon_1 \in (0, \varepsilon)$. The terms $||b^i u_{x^i}||_{n,p}, ||cu||_{n,p}$ are considered similarly.

Actually, the above conditions on b, c, and ν can be considerably relaxed if one makes use of deeper theorems about multipliers from [**33**].

Conditions (5.6) and (5.7) of the following theorem are discussed in Remarks 5.9 and 5.8.

THEOREM 5.7. *Assume that for $m = 1, 2, 3, \ldots$, we are given a_m^{ij}, σ_m^i, f_m, g_m, and u_{0m} having the same sense as in Theorem 5.1 and verifying the same assumptions as a^{ij}, σ^i, f, g, and u_0 with the same constants $\delta, K, \kappa_\varepsilon$, and K_ε. Let $\zeta(x)$ be a real function of class C_0^∞ such that $\zeta(x) = 1$ if $|x| \leq 1$ and $\zeta(x) = 0$ if $|x| \geq 2$. Define $\zeta_r(x) = \zeta(x/r)$ and assume that, for any $r = 1, 2, 3, \ldots$, $i, j = 1, \ldots, d$, $t \geq 0$, and $\omega \in \Omega$,*

$$||\zeta_r\{a^{ij}(t, \cdot) - a_m^{ij}(t, \cdot)\}||_{n,p} + ||\zeta_r\{\sigma^i(t, \cdot) - \sigma_m^i(t, \cdot)\}||_{n+1,p} \to 0 \tag{5.6}$$

as $m \to \infty$. Finally, let $E||u_{0m} - u_0||_{n+2-2/p,p}^p \to 0$ and

$$||f(u, \cdot, \cdot) - f_m(u, \cdot, \cdot)||_{\mathbb{H}_p^n(\tau)} \to 0, \quad ||g(u, \cdot, \cdot) - g_m(u, \cdot, \cdot)||_{\mathbb{H}_p^{n+1}(\tau, l_2)} \to 0 \tag{5.7}$$

whenever $u \in \mathcal{H}_p^{n+2}(\tau)$. Take the function u from Theorem 5.1 and for any m define $u_m \in \mathcal{H}_p^{n+2}(\tau)$ as the (unique) solution of the Cauchy problem for the equation

$$\begin{aligned} du_m(t, x) = &[a_m^{ij}(t, x) u_{m x^i x^j}(t, x) + f_m(u_m, t, x)]\, dt \\ &+ [\sigma_m^{ik}(t, x) u_{m x^i}(t, x) + g_m^k(u_m, t, x)]\, dw_t^k \end{aligned} \tag{5.8}$$

on $[0, \tau]$ with initial condition $u_m(0, \cdot) = u_{0m}$. Then $||u - u_m||_{\mathcal{H}_p^{n+2}(\tau)} \to 0$ as $m \to \infty$.

Proof. For $v_m = u - u_m$ we have

$$dv_m(t) = [a_m^{ij} v_{m x^i x^j} + F_m(v_m)]\, dt + [\sigma_m^{ik} v_{m x^i} + G_m^k(v_m)]\, dw_t^k,$$

where

$$F_m(v) = (a^{ij} - a_m^{ij}) u_{x^i x^j} + f(u) - f_m(u - v),$$
$$G_m^k(v) = (\sigma^{ik} - \sigma_m^{ik}) u_{x^i} + g^k(u) - g_m^k(u - v).$$

Hence, from our assumptions and by Theorem 5.1, we obtain

$$||u - u_m||_{\mathcal{H}_p^{n+2}(\tau)} \leq N I_m,$$

where N is independent of m and

$$I_m = ||(a^{ij} - a_m^{ij}) u_{x^i x^j}||_{\mathbb{H}_p^n(\tau)} + ||f(u) - f_m(u)||_{\mathbb{H}_p^n(\tau)}$$
$$+ ||(\sigma^i - \sigma_m^i) u_{x^i}||_{\mathbb{H}_p^{n+1}(\tau, l_2)} + ||g(u) - g_m(u)||_{\mathbb{H}_p^{n+1}(\tau, l_2)} + (E||u_0 - u_{0m}||_{n+2-2/p,p}^p)^{1/p}.$$

Next, by our assumptions about convergence of f_m, g_m, u_{0m},

$$\limsup_{m\to\infty} I_m \le \limsup_{m\to\infty}\{||(a^{ij}-a^{ij}_m)u_{x^ix^j}||_{\mathbb{H}^n_p(\tau)} + ||(\sigma^i-\sigma^i_m)u_{x^i}||_{\mathbb{H}^{n+1}_p(\tau,l_2)}\}. \tag{5.9}$$

Here, by Lemma 5.2, for any $v \in C_0^\infty$ and r so large that $v\zeta_r = v$, we have

$$||(a^{ij}-a^{ij}_m)u_{x^ix^j}||_{n,p} \le N||(u-v)_{x^ix^j}||_{n,p} + ||(a^{ij}-a^{ij}_m)v_{x^ix^j}||_{n,p}, \tag{5.10}$$

$$||(a^{ij}-a^{ij}_m)v_{x^ix^j}||_{n,p} = ||\zeta_r(a^{ij}-a^{ij}_m)v_{x^ix^j}||_{n,p} \le N||\zeta_r(a^{ij}-a^{ij}_m)||_{n,p}||v||_{B^{|n|+2+\gamma}},$$

where the constants N do not depend on m and r. Thus,

$$\limsup_{m\to\infty} ||(a^{ij}-a^{ij}_m)u_{x^ix^j}||_{n,p} \le N||(u-v)_{x^ix^j}||_{n,p},$$

and, from the arbitrariness of v, we conclude that the left-hand side is zero for those ω and t for which $u \in H^{n+2}_p$. If we again apply Lemma 5.2, then we see that the pth power of the left-hand side of (5.10) is bounded by an integrable function. This and the dominated convergence theorem imply that

$$\lim_{m\to\infty} ||(a^{ij}-a^{ij}_m)u_{x^ix^j}||_{\mathbb{H}^n_p(\tau)} = 0.$$

Similar arguments take care of the remaining term in (5.9). The theorem is proved.

REMARK 5.8. Condition (5.7) is satisfied for any $u \in \mathcal{H}^{n+2}_p(\tau)$ if and only if it is satisfied for $u(t,x) \equiv \phi(x)$ with any $\phi \in C_0^\infty$.

Indeed, the "only if" part is obvious. In the proof of "if" part notice that, under the "if" assumption, (5.7) is automatically satisfied for u of type

$$v = \sum_{i=1}^{j} I_{(\tau_{i-1},\tau_i]}(t)v_i(x),$$

where τ_i are bounded stopping times and $v_i \in C_0^\infty$. By Theorem 3.10, one can approximate any $u \in \mathcal{H}^{n+2}_p(\tau) \subset \mathbb{H}^{n+2}_p(\tau)$ with functions like v. It remains only to notice that, by the assumptions of the theorem, for instance,

$$||f(u,\cdot,\cdot)-f_m(u,\cdot,\cdot)||_{\mathbb{H}^n_p(\tau)} \le ||f(u,\cdot,\cdot)-f(v,\cdot,\cdot)||_{\mathbb{H}^n_p(\tau)} + ||f(v,\cdot,\cdot)-f_m(v,\cdot,\cdot)||_{\mathbb{H}^n_p(\tau)}$$
$$+||f_m(v,\cdot,\cdot)-f_m(u,\cdot,\cdot)||_{\mathbb{H}^n_p(\tau)} \le N||u-v||_{\mathbb{H}^{n+2}_p(\tau)} + ||f(v,\cdot,\cdot)-f_m(v,\cdot,\cdot)||_{\mathbb{H}^n_p(\tau)}.$$

REMARK 5.9. While checking conditions (5.6) and (5.7), it is useful to bear in mind that, if one defines $\sigma^{ik}_m = \sigma^{ik}$ and $g^k_m = g^k$ for $k \le m$ and $\sigma^{ik}_m = g^k_m = 0$ for $k > m$, then

$$||\zeta_r\{\sigma^i(t,\cdot)-\sigma^i_m(t,\cdot)\}||_{n+1,p} \to 0, \quad ||g(u,\cdot,\cdot)-g_m(u,\cdot,\cdot)||_{\mathbb{H}^{n+1}_p(\tau,l_2)} \to 0.$$

Indeed, for instance,

$$||g(u,\cdot,\cdot)-g_m(u,\cdot,\cdot)||^p_{\mathbb{H}^{n+1}_p(\tau,l_2)} = E\int_0^\tau ||\big(\sum_{k>m}|(1-\Delta)^{(n+1)/2}g^k(u,t,\cdot)|^2\big)^{1/2}||^p_p\,dt,$$

which goes to zero by the dominated convergence theorem.

This fact allows one to approximate solutions of (5.1) by solutions of

$$\begin{aligned} du_m(t,x) =&[a^{ij}(t,x)u_{mx^ix^j}(t,x) + f(u_m,t,x)]\,dt \\ &+ \sum_{k\le m}[\sigma^{ik}(t,x)u_{mx^i}(t,x) + g^k(u_m,t,x)]\,dw^k_t. \end{aligned} \tag{5.11}$$

Before stating the following corollary, take the functions ζ_k from Lemma 5.2 and, for a function $h = h(u,t,x)$ defined for $t > 0$, $x \in \mathbb{R}^d$, and u in a function space, let

$$\hat{h}_m(u,t,x) = \int_{\mathbb{R}^d} h(u(\cdot + y), t, x - y)\zeta_m(y)\, dy. \tag{5.12}$$

To get a better idea about this definition, notice that if, for instance, $h(u,t,x) = c(t,x)u(t,x)$, where $c(t,x)$ is a given function, then $h(u(\cdot + y), t, x - y) = c(t, x - y)u(t,x)$, so that $\hat{h}_m(u,t,x) = c_m(t,x)u(t,x)$ with $c_m(t,x) = c(t,x) * \zeta_m(x)$.

COROLLARY 5.10. Under the assumptions of Theorem 5.1, define

$$(a_m, \sigma_m)(t,x) = (a,\sigma)(t,x) * \zeta_m(x),$$

$$(f_m, g_m) = (\hat{f}_m, \hat{g}_m) \quad \textit{or} \quad (f_m, g_m)(u,t,x) = (f,g)(u,t,x) * \zeta_m(x).$$

Then the assumptions of Theorem 5.7 are satisfied, and if we define $u_m \in \mathcal{H}_p^{n+2}(\tau)$ as a solution of the Cauchy problem (5.8) with initial condition $u_m(0,\cdot) = u_0 * \zeta_m$, then $||u - u_m||_{\mathcal{H}_p^{n+2}(\tau)} \to 0$.

To prove this, notice that

$$(\sigma_m^i, \sigma_m^j)_{l_2}\lambda^i\lambda^j = |(\sigma^i\lambda^i) * \zeta_m|_{l_2}^2 \le \zeta_m * [(\sigma^i, \sigma^j)_{l_2}\lambda^i\lambda^j],$$

which guarantees that the approximating equations satisfy the same coercivity condition. Also, for any $u \in H_p^{n+2}$ and $\phi \in C_0^\infty$, the function $(f(u(\cdot + z), t, \cdot - y), \phi) = (f(u(\cdot + z), t, \cdot), \phi(\cdot + y))$ is continuous in z owing to Assumption 5.6 and infinitely differentiable in y as it has to be for any distribution. Therefore, the definition of $\hat{f}_m$ according to (5.12) makes sense as an integral of distribution with respect to the parameter y, namely

$$(\hat{f}_m(u,t,\cdot), \phi) := \int_{\mathbb{R}^d} \zeta_m(y)(f(u(\cdot + y), t, \cdot - y), \phi)\, dy.$$

Furthermore, $\hat{f}_m = \hat{f}_{0m} + \hat{f}_{1m}$, where $f_0 = f(0,t,x)$ and $f_1 = f(u,t,x) - f(0,t,x)$. By Lemma 5.2, we have $\hat{f}_{0m}(t,\cdot) \in H_p^n$. By Assumption 5.6,

$$|(\hat{f}_{1m}(u,t,\cdot), \phi)| \le N \int_{\mathbb{R}^d} \zeta_m(y)||\phi||_{-n,q}||u(\cdot + y)||_{n+2,p}\, dy = N||\phi||_{-n,q}||u||_{n+2,p},$$

where $q = p/(p-1)$. Hence, $\hat{f}_m$ is a H_p^n–valued function. Also, $f(u,t,x) * \zeta_m(x)$ is well defined because of Assumption 5.4. Similarly, g_m is well defined. Minkowski's inequality implies that all Assumptions 5.1–5.6 are satisfied for any m with the same $K, \delta, \kappa_\varepsilon, K_\varepsilon$. Other assumptions of Theorem 5.7 are satisfied due to continuity in the mean of summable functions and Assumptions 5.1–5.6. For instance, for any $\phi \in C_0^\infty$,

$$||f(\phi,\cdot,\cdot) - \hat{f}_m(\phi,\cdot,\cdot)||_{\mathbb{H}_p^n(\tau)} \le I_{1m} + I_{2m},$$

where

$$I_{1m} := ||f(\phi,\cdot,\cdot) - f_m(\phi,\cdot,\cdot)||_{\mathbb{H}_p^n(\tau)}, \quad f_m(\phi,t,x) = \int_{\mathbb{R}^d} f(\phi,t,x-y)\zeta_m(y)\, dy,$$

$$I_{2m} := ||f_m(\phi,\cdot,\cdot) - \hat{f}_m(\phi,\cdot,\cdot)||_{\mathbb{H}_p^n(\tau)}$$

$$\le \int_{\mathbb{R}^d} ||f(\phi,\cdot,\cdot - y) - f(\phi(\cdot + y),\cdot,\cdot - y)||_{\mathbb{H}_p^n(\tau)}\zeta_m(y)\, dy$$

$$= \int_{\mathbb{R}^d} ||f(\phi,\cdot,\cdot) - f(\phi(\cdot + y),\cdot,\cdot)||_{\mathbb{H}_p^n(\tau)}\zeta_m(y)\, dy \le N \int_{\mathbb{R}^d} ||\phi - \phi(\cdot + y)||_{n+2,p}\zeta_m(y)\, dy.$$

Clearly (cf. Lemma 5.2) $I_{1m} + I_{2m} \to 0$ as $m \to \infty$, and the application of Remark 5.8 ends the argument.

COROLLARY 5.11. Let the assumptions of Theorem 5.1 be satisfied and also let them be satisfied for a $p = q$, where $q \geq 2$. Then the solution u from Theorem 5.1 belongs also to $\mathcal{H}_q^{n+2}(\tau)$.

To prove this, without loss of generality, assume $p < q$ and let v be the solution of the same initial value problem but belonging to $\mathcal{H}_q^{n+2}(\tau)$ (such a unique v exists by Theorem 5.1). We have only to show that $v = u$. In light of Corollary 5.10, we can approximate both v and u by solutions of equations with smooth coefficients and with $(f_m, g_m)(u,t,x) = (f,g)(u,t,x) * \zeta_m(x)$. We only need to show that the approximating solutions coincide. Observe that, for any $r \geq n$, for instance,

$$||f_m(u,\cdot,\cdot) - f_m(v,\cdot,\cdot)||_{r,p} \leq N||f(u,\cdot,\cdot) - f(v,\cdot,\cdot)||_{n,p} \leq N\varepsilon||u-v||_{n+2,p}$$
$$+NK_\varepsilon||u-v||_{n,p} \leq N\varepsilon||u-v||_{r+2,p} + NK_\varepsilon||u-v||_{r,p},$$

where N's depend only on m, r, n, p. This shows that, for any fixed m, Assumptions 5.1–5.6 are satisfied with any large n both for p and q.

Therefore, we can suppose from the very beginning that Assumptions 5.1–5.6 are satisfied for any large n for p and q. In this case, by Theorem 5.1, $u \in \mathcal{H}_p^r(\tau)$ for any r, and hence, also invoking Theorem 3.7,

$$E\int_0^\tau ||u_{xx}(t,\cdot)||_{r,p}^p\,dt < \infty, \quad E\sup_{t\leq\tau}||u(t,\cdot)||_{C^r}^p < \infty$$

for any $r = 1, 2, 3, \ldots$. Since $||u_{xx}(t,\cdot)||_{r,i}^i \leq N||u_{xx}(t,\cdot)||_{r,p}^p||u(t,\cdot)||_{C^{r+2}}^{i-p}$ for $i \geq p$, it follows that

$$\int_0^\tau ||u_{xx}(t,\cdot)||_{r,i}^i\,dt < \infty \quad \text{(a.s.)}$$

for any $i \geq p$. Take here $r = 0$ and $i = q$ and define

$$\tau_k = \tau \wedge \inf\{t : \int_0^t ||u_{xx}(s,\cdot)||_q^q\,ds \geq k\}.$$

Then obviously $u \in \mathcal{H}_q^2(\tau_k)$. Since v lies in the same class, by uniqueness $u(t,\cdot) = v(t,\cdot)$ for $t \leq \tau_k$ (a.s.). It remains only to observe that $\tau_k \uparrow \tau$ when $k \to \infty$.

Next general result concerns the maximum principle.

THEOREM 5.12 (maximum principle). *Let the assumptions of Theorem 5.1 be satisfied and let u be the function existence of which is asserted in this theorem. Assume that*

$$f(v,t,x) = b^i(t,x)v_{x^i}(t,x) + v(t,x)c(t,x) + f(t,x), \quad g^k(v,t,x) = v(t,x)\nu^k(t,x),$$

where $b^i(t,x), c(t,x), \nu^k(t,x)$ are certain bounded functions on $(0,\tau]\!] \times \mathbb{R}^d$ and $f(t,x) = f(0,t,x) \geq 0$. Also assume that for any ω we have $u_0 \geq 0$. Then $u(t,\cdot) \geq 0$ for all $t \in [0,\tau]$ almost surely.

Proof. By virtue of Remark 5.9, we can only concentrate on equations with finite number of Wiener processes like (5.11), which in our case is the following

$$du(t,x) = [a^{ij}(t,x)u_{x^ix^j}(t,x) + b^i(t,x)u_{x^i}(t,x) + u(t,x)c(t,x) + f(t,x)]\,dt$$
$$+ \sum_{k\leq m}[\sigma^{ik}(t,x)u_{x^i}(t,x) + u(t,x)\nu^k(t,x)]\,dw_t^k. \quad (5.13)$$

Corollary 5.10 and the explanation after (5.12) allow us to assume that $u_0(x)$ and the coefficients and f in (5.13) are infinitely differentiable in x. After this, by multiplying u_0 and f by a cut–off function of x and by using Remark 5.6 and Theorem 5.7, we convince ourselves that we can assume that u_0 and $f(t,x)$ have supports with respect to x in a fixed ball. In this case the assumptions of Theorem 5.1 are satisfied for $p = 2$ again by Remark 5.6 (and Hölder's inequality), which, by Corollary 5.11, yields $u \in \mathcal{H}_2^n(\tau)$ for any n. Now our assertion follows from the maximum principle proved in [**25**] (see Theorem 4.2 there). The only point to mention is that in [**25**] we considered (5.13) on $[0, T]$, but we always can continue our data in an appropriate way after τ, which is assumed to be less than T. The theorem is proved.

By the way the proof of the maximum principle in [**25**] is based on representation formulas like (4.28). A heuristic derivation of many other formulas of that kind can be found in [**15**].

6. Proof of Theorem 5.1

The proof we present here is quite typical for proofs of solvability of equations with variable coefficients on the basis of solvability of equations with constant ones. The same type of arguments is commonly used in the theory of partial differential equations for proving the solvability in Sobolev or Hölder spaces. First we need some auxiliary constructions. Fix a $T \in (0, \infty)$.

DEFINITION 6.1. Assume that, for $\omega \in \Omega$ and $t \geq 0$, we are given operators

$$L(t, \cdot) : H_p^{n+2} \to H_p^n, \;\; \Lambda(t, \cdot) : H_p^{n+2} \to H_p^{n+1}(\mathbb{R}^d, l_2).$$

Assume that

(*i*) for any ω and t, the operators $L(t, u)$ and $\Lambda(t, u)$ are continuous (with respect to u);

(*ii*) for any $u \in H_p^{n+2}$, the processes $L(t, u)$ and $\Lambda(t, u)$ are predictable;

(*iii*) for any $\omega \in \Omega$, $t \geq 0$, and $u \in H_p^{n+2}$, we have

$$||L(t, u)||_{n,p} + ||\Lambda(t, u)||_{n+1,p} \leq N_{L,\Lambda}(1 + ||u||_{n+2,p}),$$

where $N_{L,\Lambda}$ is a constant.

Then for a function $u \in \mathcal{H}_p^{n+2}(T)$, we write

$$(L, \Lambda)u = -(f, g)$$

if $(f, g) \in \mathcal{F}_p^n(T)$, and, in the sense of Definition 3.5, for $t \in [0, T]$, we have that $\mathbb{D}u(t) = L(t, u(t)) + f(t)$ and $\mathbb{S}u(t) = \Lambda(t, u(t)) + g(t)$, or put otherwise

$$u(t) = u(0) + \int_0^t (L(s, u(s)) + f(s))\, ds + \int_0^t (\Lambda^k(s, u(s)) + g^k(s))\, dw_s^k \quad \text{(a.s.)}.$$

REMARK 6.2. By virtue of our conditions on L and Λ, for any $u \in \mathcal{H}_p^{n+2}(T)$, we have $(L(u), \Lambda(u)) \in \mathcal{F}_p^n(T)$. Also, $(L, \Lambda)u = (L(u) - \mathbb{D}u, \Lambda(u) - \mathbb{S}u)$. In particular, the operator (L, Λ) *is well defined* on $\mathcal{H}_p^{n+2}(T)$, and, as follows easily from Definition 6.1 (*iii*),

$$||(L, \Lambda)u||_{\mathcal{F}_p^n(T)} \leq (1 + 2N_{L,\Lambda})||u||_{\mathcal{H}_p^{n+2}(T)} + 2N_{L,\Lambda}T^{1/p}.$$

In terms of Definition 6.1, Theorem 4.10 has the following version.

THEOREM 6.3. *Let a and σ satisfy the assumptions from the beginning of Sec. 4. Define*

$$Lu = a^{ij}u_{x^ix^j}, \;\; \Lambda u = \sigma^i u_{x^i}.$$

Then the operator (L, Λ) is a one-to-one operator from $\mathcal{H}_{p,0}^{n+2}(T)$ onto $\mathcal{F}_p^n(T)$ and the norm of its inverse is less than a constant depending only on d, p, δ, and K (thus independent of T).

Next, we prove a perturbation result. It needs a proof because we do not allow ε to depend on T.

THEOREM 6.4. *Take the operators L and Λ from Theorem 6.3, and let some operators L_1 and Λ_1 satisfy the requirements from Definition 6.1. We assert that there exists a constant $\varepsilon \in (0,1)$ depending only on d, p, δ, and K such that if, for a constant K_1 and any $u, v \in H_p^{n+2}$, $t \geq 0$, $\omega \in \Omega$, we have*

$$||L_1(t,u) - L_1(t,v)||_{n,p} + ||\Lambda_1(t,u) - \Lambda_1(t,v)||_{n+1,p}$$
$$\leq \varepsilon ||u_{xx} - v_{xx}||_{n,p} + K_1 ||u - v||_{n+1,p}, \tag{6.1}$$

then, for any $(f,g) \in \mathcal{F}_p^n(T)$, there exists a unique solution $u \in \mathcal{H}_{p,0}^{n+2}(T)$ of the equation

$$(L + L_1, \Lambda + \Lambda_1)u = -(f,g). \tag{6.2}$$

Furthermore, for this solution u, we have

$$||u||_{\mathcal{H}_p^{n+2}(T)} \leq N||(L_1(\cdot,0) + f, \Lambda_1(\cdot,0) + g)||_{\mathcal{F}_p^n(T)}, \tag{6.3}$$

where N depends only on d, p, δ, K, K_1, and T and N is independent of T if $K_1 = 0$.

Proof. First notice that, by interpolation theorems, $||u||_{n+1,p} \leq \varepsilon ||u_{xx}||_{n,p} + N(\varepsilon, d, p)||u||_{n,p}$. Therefore, without loss of generality we assume that instead of (6.1) we have

$$||L_1(t,u) - L_1(t,v)||_{n,p} + ||\Lambda_1(t,u) - \Lambda_1(t,v)||_{n+1,p}$$
$$\leq \varepsilon ||u_{xx} - v_{xx}||_{n,p} + K_1 ||u - v||_{n,p}.$$

Now fix $(f,g) \in \mathcal{F}_p^n(T)$. Take $u \in \mathcal{H}_{p,0}^{n+2}(T)$, observe that $(L_1(u), \Lambda_1(u)) \in \mathcal{F}_p^n(T)$, and, by using Theorem 6.3, define $v \in \mathcal{H}_{p,0}^{n+2}(T)$ as the unique solution of the equation $(L,\Lambda)v = -(f + L_1(u), g + \Lambda_1(u))$. By denoting $v = Ru$, we define an operator $R : \mathcal{H}_{p,0}^{n+2}(T) \to \mathcal{H}_{p,0}^{n+2}(T)$. Equation (6.2) is equivalent to the equation $u = Ru$. Therefore, to prove the existence and uniqueness of solutions to (6.2), it suffices to show that, for an integer $m > 0$, the operator R^m is a contraction in $\mathcal{H}_{p,0}^{n+2}(T)$.

By Theorem 6.3, for $t \leq T$,

$$||Ru - Rv||^p_{\mathcal{H}_p^{n+2}(t)} \leq N||(L_1(u) - L_1(v), \Lambda_1(u) - \Lambda_1(v))||^p_{\mathcal{F}_p^n(t)}$$
$$\leq N_0 \varepsilon^p ||u - v||^p_{\mathcal{H}_p^{n+2}(t)} + N_0 K_1^p \int_0^t E||u(s) - v(s)||^p_{n,p}\, ds,$$

with a constant N_0 depending only on d, p, δ, and K. This gives the desired result if $K_1 = 0$. Also, in this case estimate (6.3) follows obviously with N independent of T.

In the general case, by Theorem 3.7,

$$E||u(s) - v(s)||^p_{n,p} \leq N_1 ||u - v||^p_{\mathcal{H}_p^{n+2}(s)},$$

where $s \leq T$ and N_1 depends only on d, p, and T. It follows that, for $t \leq T$ and $\theta := N_0 \varepsilon^p$, we have

$$||Ru - Rv||^p_{\mathcal{H}_p^{n+2}(t)} \leq \theta ||u - v||^p_{\mathcal{H}_p^{n+2}(t)} + N_2 \int_0^t ||u - v||^p_{\mathcal{H}_p^{n+2}(s)}\, ds,$$

where N_2 depends only on d, p, δ, K, K_1, and T. Hence, by induction,

$$||R^m u - R^m v||^p_{\mathcal{H}_p^{n+2}(t)} \leq \theta^m ||u - v||^p_{\mathcal{H}_p^{n+2}(t)}$$
$$+ \sum_{k=1}^m \binom{m}{k} \theta^{m-k} N_2^k \int_0^t \frac{(t-s)^{k-1}}{(k-1)!} ||u - v||^p_{\mathcal{H}_p^{n+2}(s)}\, ds, \quad ||R^m u - R^m v||^p_{\mathcal{H}_p^{n+2}(T)}$$
$$\leq \sum_{k=0}^m \binom{m}{k} \theta^{m-k} \frac{1}{k!} (TN_2)^k ||u - v||^p_{\mathcal{H}_p^{n+2}(T)} \leq 2^m \theta^m \max_k \frac{1}{k!} (TN_2/\theta)^k ||u - v||^p_{\mathcal{H}_p^{n+2}(T)}.$$

This allows us to find ε depending only on d, p, δ, and K and m depending on the same things plus K_1 and T, so that R^m is a contraction in $\mathcal{H}^{n+2}_{p,0}(T)$ with coefficient 1/2. Of course, this yields all our assertions. The theorem is proved.

We finish our preparations by showing how Lemma 5.2 will be used.

REMARK 6.5. To some extent, in what follows, the most important consequence of assertion (i) in Lemma 5.2 is that if $\bar{a} = ||a||_{B^{|n|+\gamma}} < \infty$, then there exists a new norm $]|\cdot|[_{n,p}$ in H^n_p such that

$$]|au|[_{n,p} \leq 2N||a||_B]|u|[_{n,p} \quad (||a||_B = \sup_{x\in\mathbb{R}^d}|a(x)|),$$

where N is the same constant as in Lemma 5.2. To show this, it suffices to observe that, for $a_m(x) = a(x/m)$, $u_m(x) = u(x/m)$, and $m \geq 1$, we have

$$||(m^2-\Delta)^{n/2}(au)||_p = m^{n-d/p}||(1-\Delta)^{n/2}(a_m u_m)||_p \leq$$

$$N\bar{a}_m m^{n-d/p}||(1-\Delta)^{n/2}u_m||_p = N\bar{a}_m||(m^2-\Delta)^{n/2}u||_p \leq$$

$$N(||a||_B + \frac{1}{m^{(|n|+\gamma)\wedge 1}}\bar{a})||(m^2-\Delta)^{n/2}u||_p.$$

Alternatively, it would be sufficient for our needs to know that

$$||au||_{n,p} \leq N(||a||_B||u||_{n,p} + ||a||_{B^{|n|+\gamma}}||u||_{n-1,p}).$$

The author could not find the last inequality in the literature, though some very interesting information related to the subject can be found in [**13**].

Now we perform the main step in proving Theorem 5.1.

LEMMA 6.6. *Let Assumptions 5.1, 5.3, 5.4, and 5.5 be satisfied. Then there exists an* $\varepsilon = \varepsilon(d, p, n, \gamma, \delta, K) > 0$ *such that if* $\tau = T$ *and*

(i) inequality (5.3) holds with this ε *for all* x, y, t, *and* ω *and*

(ii) f *and* g *are independent of* u,

then there exists a unique solution $u \in \mathcal{H}^{n+2}_{p,0}(T)$ *of equation (5.1). Furthermore, for this solution* u, *we have*

$$||u||_{\mathcal{H}^{n+2}_p(T)} \leq N||(f,g)||_{\mathcal{F}^n_p(T)},$$

where N *depends only on* d, p, δ, K, *and* T.

Proof. Define $a(t) = a(t,0)$ and $\sigma(t) = \sigma(t,0)$, take operators L and Λ from Theorem 6.3 corresponding to these $a(t)$ and $\sigma(t)$, and let

$$L_1(t,u)(x) = [a^{ij}(t,x) - a^{ij}(t)]u_{x^ix^j}(x), \quad \Lambda_1(t,u)(x) = [\sigma^i(t,x) - \sigma^i(t)]u_{x^i}(x).$$

In view of Theorem 6.4, to prove existence and uniqueness, we have only to check that if ε in (5.3) is sufficiently small, then the operators L_1 and Λ_1 satisfy condition (6.1) with as small ε as we like and with K_1 under control. Observe that by Lemma 5.2,

$$||L_1(t,u)||_{n,p} \leq N||a(t,\cdot) - a(t,0)||_{B^{|n|+\gamma}}||u_{xx}||_{n,p},$$

$$||\Lambda_1(t,u)||_{n+1,p} \leq N||\sigma(t,\cdot) - \sigma(t,0)||_{B^{|n+1|+\gamma}(l_2)}||u_x||_{n+1,p}.$$

Since $||u_x||_{n+1,p} \leq N(||u_{xx}||_{n,p} + ||u||_{n+1,p})$, our lemma holds true if

$$||a(t,\cdot) - a(t,0)||_{B^{|n|+\gamma}} + ||\sigma(t,\cdot) - \sigma(t,0)||_{B^{|n+1|+\gamma}(l_2)} \leq \varepsilon_0 = \varepsilon_0(d,p,n,\gamma,\delta,K) \quad \forall\ t.$$

Next, observe that, for $a_m(t,x) = a(t/m^2, x/m)$ and $m \geq 1$, we have

$$||a_m(t,\cdot) - a_m(t,0)||_{B^{|n|+\gamma}} \leq \varepsilon + m^{-[(|n|+\gamma)\wedge 1]}K$$

(K comes from Assumption 5.5), and for $|n| + \gamma = 0$ we can even drop the second term on the right. An analogous inequality holds for σ. It follows that, for ε sufficiently small and m sufficiently large, the statements of the lemma are true if we replace a, σ, w_t, f, g, and T in equation (5.1) by $a_m, \sigma_m, mw_{t/m^2}$, $m^{-2}f(t/m^2, x/m)$, $m^{-1}g(t/m^2, x/m)$, and m^2T, respectively. After this, it remains only to fix an appropriate m and make an obvious change of the unknown function in the above-mentioned modification of equation (5.1) (and use $1 - \Delta \sim m^2 - \Delta$ so that the norms $||\cdot||_{n,p}$ of $u(t/m^2, x/m)$ and $u(t,x)$ are comparable). The lemma is proved.

Finally, we need the following result from [**17**], which, in a sense, is essentially covered by Theorem 2.4.7 from [**34**].

LEMMA 6.7. *Let $\delta > 0$ and let $\zeta_k \in C^\infty$, $k = 1, 2, 3, \ldots$. Assume that for any multi-index α and $x \in \mathbb{R}^d$,*

$$\sup_{x \in \mathbb{R}^d} \sum_k |D^\alpha \zeta_k(x)| \le M(\alpha),$$

where $M(\alpha)$ are some constants. Then there exists a constant $N = N(d, n, M)$ such that, for any $f \in H_p^n$,

$$\sum_k ||\zeta_k f||^p_{H_p^n} \le N||f||^p_{H_p^n}.$$

If in addition

$$\sum_k |\zeta_k(x)|^p \ge \delta,$$

then for any $f \in H_p^n$,

$$||f||^p_{n,p} \le N(d, n, M, \delta) \sum_k ||\zeta_k f||^p_{n,p}.$$

REMARK 6.8. We will also use a natural extension of this lemma to the case of Banach–space valued f.

Proof of Theorem 5.1. By Theorem 2.1, for any nonrandom $z \in H_p^{n+2-2/p}$ there exists a unique solution $u \in \mathcal{H}_p^{n+2}(\tau)$ of the equation $du = \Delta u\, dt$ with initial condition z. This theorem also provides estimate (2.7) of the norm of u. From this theorem and the estimate, it follows that there exists a unique solution $\bar{u} \in \mathcal{H}_p^{n+2}(\tau)$ of the equation $du = \Delta u\, dt$ with initial condition u_0 and

$$||\bar{u}||^p_{\mathcal{H}_p^{n+2}(\tau)} \le NE||u_0||^p_{n+2-2/p,p}.$$

After replacing the unknown function u with $v + \bar{u}$, it becomes obvious that, to solve (5.1) in $\mathcal{H}_p^{n+2}(\tau)$ with the initial condition u_0, it suffices to solve the equation

$$\begin{aligned} du(t,x) =& [a^{ij}(t,x)u_{x^ix^j}(t,x) + \bar{f}(u,t,x)]\, dt \\ &+ [\sigma^{ik}(t,x)u_{x^i}(t,x) + \bar{g}^k(u,t,x)]\, dw_t^k \end{aligned}$$

in the space $\mathcal{H}_{p,0}^{n+2}(\tau)$, where (remember $d\bar{u} = \Delta\bar{u}\, dt$)

$$\bar{f}(u,t,x) = \{f(u+\bar{u},t,x) + a^{ij}(t,x)\bar{u}_{x^ix^j}(t,x) - \Delta\bar{u}\}I_A(t),$$

$$\bar{g}^k(u,t,x) = \{\sigma^{ik}(t,x)\bar{u}_{x^i}(t,x) + g^k(u+\bar{u},t,x)\}I_A(t),$$

and A is the set of ω, t for which $\bar{u}(t,\cdot) \in H_p^{n+2}$.

From Lemma 5.2, it follows that $\bar{f}$ and $\bar{g}$ satisfy the same conditions as f and g. This allows us to assume that $u_0 = 0$. Furthermore, one can always extend f, g after time τ by equaling them to zero. That is why without loss of generality we assume that $u_0 = 0$ and $\tau = T$.

Now, define

$$Lu = a^{ij}(t,x)u_{x^ix^j}(t,x), \quad \Lambda u = \sigma^i(t,x)u_{x^i}(t,x),$$

and let $\{\zeta_k : k = 1,2,3,...\}$ be a standard partition of unity such that, for any k, the support of ζ_k lies in a ball B_k of radius $(1/4)\kappa_{\varepsilon/2}$, where κ_ε is taken from Assumption 5.2 and ε from Lemma 6.6. Also for any k, we take a function $\eta_k \in C_0^\infty$ such that $\eta_k = 1$ on B_k, $\eta_k = 0$ outside doubled B_k, and $0 \le \eta_k \le 1$. Denote by x_k the center of B_k, define $L_k(t,x) = \eta_k(x)L(t,x) + (1-\eta_k(x))L(t,x_k)$, and similarly define Λ_k.

Observe that, for any k, the operators L_k, Λ_k satisfy condition (i) of Lemma 6.6. Therefore, if we denote $(f_k, g_k) := (L_k, \Lambda_k)(u\zeta_k)$, then by this lemma

$$||u\zeta_k||_{\mathcal{H}_p^{n+2}(T)} \le N||(f_k,g_k)||_{\mathcal{F}_p^n(T)} = N||(L_k,\Lambda_k)(u\zeta_k)||_{\mathcal{F}_p^n(T)}.$$

Furthermore, by using that $\eta_k = 1$ everywhere where $\zeta_k u \neq 0$ we easily check that

$$(L_k,\Lambda_k)(u\zeta_k) = (L,\Lambda)(u\zeta_k) = \zeta_k(L,\Lambda)u + (uL\zeta_k + 2\zeta_{kx}\cdot au_x, u\Lambda\zeta_k),$$

so that

$$||u\zeta_k||_{\mathcal{H}_p^{n+2}(T)} \le N||\zeta_k(L,\Lambda)u||_{\mathcal{F}_p^n(T)} + N||(uL\zeta_k + 2\zeta_{kx}\cdot au_x, u\Lambda\zeta_k)||_{\mathcal{F}_p^n(T)}.$$

We sum up the pth powers of the extreme terms and apply Lemma 6.7 and Lemma 5.2 in the estimates like the following one:

$$\sum_k ||ua^{ij}\zeta_{kx^ix^j}||_{n,p} \le N||ua^{ij}||_{n,p} \le N||u||_{n,p}.$$

Then we conclude that, for any $u \in \mathcal{H}_{p,0}^{n+2}(T)$,

$$||u||_{\mathcal{H}_p^{n+2}(T)} \le N||(L,\Lambda)u||_{\mathcal{F}_p^n(T)} + N||u||_{\mathbb{H}_p^{n+1}(T)}. \tag{6.4}$$

Next, we show that the last term on the right can be dropped. Indeed, $||u||_{n+1,p} \le \varepsilon||u_{xx}||_{n,p} + N(\varepsilon,d,p)||u||_{n,p}$. Therefore, (6.4) can be modified by replacing the last term with $N||u||_{\mathbb{H}_p^n(T)}$. This can be done with any $t \le T$ in place of T. After this, by Theorem 3.7, the inequality

$$E||u(t,\cdot)||_{n,p}^p \le N||(L,\Lambda)u||_{\mathcal{F}_p^n(T)}^p + N\int_0^t E||u(s,\cdot)||_{n,p}^p\,ds$$

holds for any $t \le T$. By Gronwall's inequality, this yields that

$$||u||_{\mathbb{H}_p^n(T)}^p \le N||(L,\Lambda)u||_{\mathcal{F}_p^n(T)}^p$$

which, along with the modified (6.4), proves that

$$||u||_{\mathcal{H}_p^{n+2}(T)} \le N||(L,\Lambda)u||_{\mathcal{F}_p^n(T)} \tag{6.5}$$

for any $u \in \mathcal{H}_{p,0}^{n+2}(T)$.

This is an a priori estimate. Now we use the standard method of continuity. For $\lambda \in [0,1]$ we consider the equation

$$du = (L_\lambda u + f)\,dt + (\Lambda_\lambda^k u + g^k)\,dw_t^k \tag{6.6}$$

with zero initial condition, where

$$L_\lambda u = \lambda\Delta + (1-\lambda)L, \quad \Lambda_\lambda u = (1-\lambda)\Lambda$$

and (f,g) is an arbitrary element in $\mathcal{F}_p^n(T)$.

Observe that the a priori estimate (6.5) holds with the same constant N for all $L_\lambda, \Lambda_\lambda$ in place of L, Λ. Next, take a $\lambda_0 \in [0,1]$ and assume that for $\lambda = \lambda_0$ equation (6.6) with zero initial data has a unique solution $u \in \mathcal{H}^{n+2}_{p,0}(T)$ for any $(f,g) \in \mathcal{F}^n_p(T)$. By the way, this assumption is satisfied for $\lambda_0 = 1$ by Theorem 6.3. Then we have an operator

$$\mathcal{R}_{\lambda_0} : \mathcal{F}^n_p(T) \to \mathcal{H}^{n+2}_{p,0}(T)$$

such that $\mathcal{R}_{\lambda_0}(f,g) = u$. From (6.5) we get that

$$||\mathcal{R}_{\lambda_0}(f,g)||_{\mathcal{H}^{n+2}_p(T)} \leq N||(f,g)||_{\mathcal{F}^n_p(T)}. \tag{6.7}$$

For other $\lambda \in [0,1]$ we rewrite (6.6) as

$$du = (L_{\lambda_0}u + \{(\lambda - \lambda_0)(\Delta - L)u + f\})\,dt + (\Lambda^k_{\lambda_0}u + \{(\lambda_0 - \lambda)\Lambda^k u + g^k\})\,dw^k_t$$

and we solve the last equation by iterations. Define $u_0 = 0$ and

$$u_{j+1} = \mathcal{R}_{\lambda_0}((\lambda - \lambda_0)(\Delta - L)u_j + f, (\lambda_0 - \lambda)\Lambda u_j + g).$$

Then by (6.7)

$$\begin{aligned}||u_{j+1} - u_j||_{\mathcal{H}^{n+2}_p(T)} &\leq N|\lambda - \lambda_0|\,||((\Delta - L)(u_j - u_{j-1}), \Lambda(u_j - u_{j-1}))||_{\mathcal{F}^n_p(T)}\\ &\leq N_1|\lambda - \lambda_0|\,||u_j - u_{j-1}||_{\mathcal{H}^{n+2}_p(T)},\end{aligned}$$

where N_1 is independent of j, λ, and λ_0. If $N_1|\lambda - \lambda_0| \leq 1/2$, then u_j is a Cauchy sequence in $\mathcal{H}^{n+2}_p(T)$, which converges by Theorem 3.7. Its limit satisfies

$$u = \mathcal{R}_{\lambda_0}((\lambda - \lambda_0)(\Delta - L)u + f, (\lambda_0 - \lambda)\Lambda u + g),$$

which is equivalent to (6.6).

In this way we show that if (6.6) is solvable for a λ_0, then it is solvable for λ satisfying $N_1|\lambda - \lambda_0| \leq 1/2$. In finite number of steps starting with $\lambda = 1$, we get to $\lambda = 0$. This proves the theorem if f, g are independent of u.

To consider general f and g, it remains only to repeat the proof of Theorem 6.4 taking $f(u,t,x)$ and $g(u,t,x)$ instead of $f + L_1(u)$ and $g + \Lambda_1(u)$ there. The theorem is proved.

7. Embedding Theorems for $\mathcal{H}^n_p(\tau)$

The above theory of solvability looks satisfactory only until one tries to apply it to concrete problems when one is interested in getting not only solvability but also some qualitative properties of solutions, like continuity, decay at infinity, compactness of support, and so on. To answer such questions, one has to understand what qualitative properties the solutions have. Since solutions are just arbitrary functions from $\mathcal{H}^n_p(\tau)$, we are actually interested in properties of functions from this space. Let us fix a $T \in [0,\infty)$ and a stopping time $\tau \leq T$.

The first two assertions of the following theorem are straightforward corollaries of two Sobolev's theorems. One says that $H^n_p \subset \mathcal{C}^\alpha$ if $\alpha := n - d/p > 0$, where $\mathcal{C}^\alpha = \mathcal{C}^\alpha(\mathbb{R}^d)$ is the Zygmund space (which differs from the usual Hölder space $C^\alpha = C^\alpha(\mathbb{R}^d)$ only if α is an integer, see [**33**]). The second one says that $H^n_p \subset H^m_q$ if $m < n$ and $n - d/p = m - d/q$.

THEOREM 7.1. *(i) If $\alpha := n - d/p > 0$ and $u \in \mathcal{H}^n_p(\tau)$, then $u \in L_p((0,\tau]], \mathcal{C}^\alpha)$, where $\mathcal{C}^\alpha$ is the Zygmund space. In addition,*

$$E\int_0^\tau ||u(t,\cdot)||^p_{\mathcal{C}^\alpha}\,dt \leq N(d,n,p)||u||^p_{\mathcal{H}^n_p(\tau)}.$$

(ii) If $m < n$, $n - d/p = m - d/q$, and $u \in \mathcal{H}^n_p(\tau)$, then

$$E\int_0^\tau ||u(t,\cdot)||^p_{m,q}\,dt \leq N(d,n,m,p)||u||^p_{\mathcal{H}^n_p(\tau)}.$$

(iii) For any function $u \in \mathcal{H}_2^n(\tau)$, we have $u \in C([0,\tau], H_2^{n-1})$ (a.s.) and

$$E\sup_{t\leq\tau}||u(t,\cdot)||_{n-1,2}^2 \leq N(d,n,T)||u||_{\mathcal{H}_2^n(\tau)}^2.$$

Proof. As we have said before the theorem, we only need to prove the third assertion. By Remark 3.4, we may assume that $n=1$. Denote $u_0 = u(0)$, $v = T_t u_0$. Observe that by Theorem 2.1 we have $v \in \mathcal{H}_2^1(\tau)$. By Minkowski's inequality, $||\zeta * u||_p \leq ||u||_p ||\zeta||_1$, so that

$$||T_t u_0||_2 \leq ||u_0||_2, \quad E\sup_t ||v(t,\cdot)||_2^2 \leq E||u_0||_2^2.$$

In addition, almost obviously, $T_t u_0$ is a continuous (analytic) L_2–valued function in t for $t>0$. Also, we have $(T_t u_0)'_t = T_t \Delta u_0$, which implies that $||T_t u_0 - u_0||_2 \leq t||\Delta u_0||_2 \to 0$ if $u_0 \in H_2^2$. Adding that the set H_2^2 is dense in L_2, we conclude that T_t is a continuous semigroup in L_2. This means that $v \in C([0,\tau], H_2^{n-1})$ and shows that we need only to consider $u - v$. In other words, in the rest of the proof we may and will assume that $u_0 = 0$.

In this case, denote $f = (\mathbb{D}u - \Delta u)I_{t\leq\tau}$, and $g = (\mathbb{S}u)I_{t\leq\tau}$. Solve equation (4.2) on $[0,\infty)$ with zero initial data. By uniqueness, the solution coincides with u on $[0,\tau]$. By Theorem 4.2, assertion (iii) holds. The theorem is proved.

Further results about some basic properties of the spaces $\mathcal{H}_p^n(\tau)$ are collected in the following theorem. As we have seen in Theorem 4.2, assertion (iii) of Theorem 7.1 is true not only for $p=2$ but also for any $p\geq 2$. However, for $p>2$ a much stronger statement (i) of Theorem 7.2 holds.

THEOREM 7.2. *(i) If $p>2$, $1/2>\beta>\alpha>1/p$, then for any function $u \in \mathcal{H}_p^n(\tau)$, we have $u \in C^{\alpha-1/p}([0,\tau], H_p^{n-2\beta})$ (a.s.) and for any stopping time $\eta\leq\tau$,*

$$E||u(t\wedge\eta,\cdot) - u(s\wedge\eta,\cdot)||_{n-2\beta,p}^p \leq N(d,\beta,p,T)|t-s|^{\beta p-1}||u||_{\mathcal{H}_p^n(\tau)}^p \quad \forall t,s\leq T; \tag{7.1}$$

$$E||u(t,\cdot)||_{C^{\alpha-1/p}([0,\tau],H_p^{n-2\beta})}^p \leq N(d,\beta,\alpha,p,T)||u||_{\mathcal{H}_p^n(\tau)}^p. \tag{7.2}$$

(ii) If $q\geq p>2$ and $\theta\in(0,1)$, then for

$$m < n + \frac{d}{q} - \frac{d+2(1-\theta)}{p}, \quad u \in \mathcal{H}_p^n(\tau),$$

we have $u \in L_{p/\theta}((0,\tau), H_q^m)$ (a.s.) and

$$E(\int_0^\tau ||u(t,\cdot)||_{m,q}^{p/\theta}\,dt)^\theta \leq N(d,p,q,n,m,\theta,T)||u||_{\mathcal{H}_p^n(\tau)}^p.$$

In particular (take $\theta = p/q$),

$$E(\int_0^\tau ||u(t,\cdot)||_{m,q}^q\,dt)^{p/q} \leq N(d,p,q,n,m,T)||u||_{\mathcal{H}_p^n(\tau)}^p$$

if

$$q>p>2, \quad m<n-(d+2)(\frac{1}{p}-\frac{1}{q}).$$

To prove the theorem, we need two lemmas. Remember the notation T_t introduced in (2.3).

LEMMA 7.3. *For any $h\in L_p$, $\theta\in[0,1]$, and $t>0$, we have*

$$||e^{-t}T_t h||_p \leq N\frac{1}{t^\theta}||h||_{-2\theta,p}, \quad ||(T_t-1)h||_p \leq Nt^\theta ||h||_{2\theta,p}, \tag{7.3}$$

where $N = N(d,p,\theta)$.

Proof. These inequalities follow from Theorem 14.11 of [**14**]. For the sake of completeness, we prove them. The derivative $(T_t h)'_t$ can be represented as t^{-1} times a convolution of h with a function having finite L_1–norm, the norm being independent of t. By Minkowski's inequality, $||\zeta * h||_p \le ||h||_p ||\zeta||_1$ so that

$$||(T_t h)'_t||_p \le N(d,p)\frac{1}{t}\,||h||_p, \quad ||(e^{-t}T_t h)'_t||_p \le N(d,p)\frac{1}{t}\,||h||_p. \tag{7.4}$$

Hence for $\theta \in (0,1)$, (see (2.2))

$$||(1-\Delta)^\theta (e^{-t}T_t h)||_p = c(\theta)||\int_0^\infty \frac{e^{-(s+t)}T_{s+t}u - e^{-t}T_t h}{s^\theta}\frac{ds}{s}||_p$$

$$\le N\int_0^t \frac{1}{s^\theta}||e^{-(s+t)}T_{s+t}u - e^{-t}T_t h||_p\,\frac{ds}{s} + N\int_t^\infty \frac{ds}{s^{1+\theta}}||h||_p.$$

Here by (7.4)

$$||e^{-(s+t)}T_{s+t}u - e^{-t}T_t h||_p \le N\frac{s}{t}\,||h||_p,$$

$$\int_0^t \frac{1}{s^\theta}||e^{-(s+t)}T_{s+t}u - e^{-t}T_t h||_p\,\frac{ds}{s} \le ||h||_p\int_0^t \frac{1}{ts^\theta}\,ds = N||h||_p\frac{1}{t^\theta}.$$

This proves that

$$||e^{-t}T_t(1-\Delta)^\theta h||_p \le N||h||_p t^{-\theta}, \tag{7.5}$$

which gives the first inequality in (7.3) after replacing $(1-\Delta)^\theta h$ with h. If $\theta = 0$, then one can take $N = 1$ in the first inequality in (7.3), which follows from Minkowski's inequality. If $\theta = 1$, then we use (7.4) and the fact that $(e^{-t}T_t h)'_t = e^{-t}T_t(\Delta - 1)h$.

To prove the second inequality in (7.3) for $t \in [0,1]$ and $\theta > 0$, it suffices to notice that $(T_t h)'_t = \Delta T_t h$ and $\Delta = [\Delta(1-\Delta)^{-1}](1-\Delta)^{1-\theta}(1-\Delta)^\theta$ and use (7.5) in the following estimates

$$||(T_t - 1)h||_p \le \int_0^t ||[\Delta(1-\Delta)^{-1}](1-\Delta)^{1-\theta}T_s[(1-\Delta)^\theta h]||_p\,ds$$

$$\le N\int_0^t ||(1-\Delta)^{1-\theta}T_s[(1-\Delta)^\theta h]||_p\,ds \le N\int_0^t s^{\theta-1}\,ds||h||_{2\theta,p} = Nt^\theta||h||_{2\theta,p}.$$

For $t \ge 1$ or $\theta = 0$, the second inequality in (7.3) is trivial since $||h||_p \le ||h||_{2\theta}$ and $t^0 = 1$. The lemma is proved.

LEMMA 7.4. *Let $\alpha p > 1$ and $p \ge 1$. Then, for any continuous L_p-valued function $h(t)$, and $s \le t$, we have*

$$\begin{aligned} ||h(t)-h(s)||_p^p &\le N(\alpha,p)(t-s)^{\alpha p-1}\int_s^t\int_s^t I_{r_2>r_1}\frac{||h(r_2)-h(r_1)||_p^p}{|r_2-r_1|^{1+\alpha p}}\,dr_1dr_2 \\ &= N(\alpha,p)(t-s)^{\alpha p-1}\int_0^{t-s}\frac{d\gamma}{\gamma^{1+\alpha p}}\int_s^{t-\gamma}||h(r+\gamma)-h(r)||_p^p\,dr \quad (\frac{0}{0} := 0). \end{aligned} \tag{7.6}$$

This is one of embedding theorems for Slobodetskii's spaces (see, for instance, [**33**]). It is to be said that embedding theorems were applied for studying continuity properties of random processes for quite long time; see, for instance, [**1**], [**9**], [**11**], [**19**], [**35**]. The following consequence of (7.6),

$$E\sup_{0\le s<t\le T}\frac{||h(t)-h(s)||_p^p}{(t-s)^{\alpha p-1}} \le N(\alpha,p)\int_0^T\int_0^T I_{r_2>r_1}\frac{E||h(r_2)-h(r_1)||_p^p}{|r_2-r_1|^{1+\alpha p}}\,dr_1dr_2, \tag{7.7}$$

can be used wherever one uses Kolmogorov's continuity criterion (see, [**1**]). By the way, embedding theorems are (for the most part) known for multi–dimensional case as well, and one can use them for studying random fields (cf. [**35**]).

In [**19**], equation (7.7) with $T = \pi$ was used to prove that Fourier's series

$$\frac{1}{\sqrt{\pi}} t\eta_0 + \sqrt{\frac{2}{\pi}} \sum_{n=1}^{\infty} \eta_n \frac{1}{n} \sin nt,$$

where η_n are $N(0,1)$ i.i.d. random variables, converges uniformly on $[0,\pi]$. This is the way the Wiener process is introduced in [**19**]. Estimate (7.7) also gives a very good modulus of continuity for the Wiener process. Of course, an advantage of normal centered variables is that $(E|\xi|^p)^{1/p} = N(p)(E|\xi|^2)^{1/2}$. Instead of this, in our applications here we use Burkholder–Davis–Gundy inequalities. By the way, a very short proof of (7.6) may be found in [**19**]. Finally, notice that the space L_p in (7.6) and (7.7) can be replaced with any Banach space.

Proof of Theorem 7.2. Take $u \in \mathcal{H}_p^n(\tau)$, and define $f = (\mathbb{D}u - \Delta u)I_{t\le\tau}$ and $g = (\mathbb{S}u)I_{t\le\tau}$. Notice that the function u on $[0,\tau]$ satisfies the equation

$$dv = (\Delta v + f)\, dt + g^k\, dw_t^k. \tag{7.8}$$

By Theorem 5.1, equation (7.8) on $[0,T]$ with initial condition $v(0) = u(0)$ has a unique solution $v \in \mathcal{H}_p^n(T)$. The difference $u - v$ satisfies the heat equation on $[0,\tau]$ with zero initial condition. It follows that $u(t,\cdot) = v(t,\cdot)$ on $[0,\tau]$, and by Theorem 5.1,

$$||v||_{\mathcal{H}_p^n(T)} \le N(||f||_{\mathbb{H}_p^{n-2}(\tau)} + ||g||_{\mathbb{H}_p^{n-1}(\tau,l_2)} + (E||u(0)||^p_{n-2/p,p})^{1/p}) \le N||u||_{\mathcal{H}_p^n(\tau)}.$$

The inequality $||v||_{\mathcal{H}_p^n(T)} \le N||u||_{\mathcal{H}_p^n(\tau)}$ and the equality $u(t,\cdot) = v(t,\cdot)$ on $[0,\tau]$ show that we only need to prove the theorem for v, T in place of u, τ. In particular, we may and will assume that $\tau = T$. Also, observe that

$$||f||_{\mathbb{H}_p^{n-2}(T)} + ||g||_{\mathbb{H}_p^{n-1}(T,l_2)} + (E||u(0)||^p_{n-2/p,p})^{1/p} \le N||u||_{\mathcal{H}_p^n(T)}.$$

This allows us to concentrate only on solutions of equations (7.8) on $[0,T]$ and to prove our assertions with $||(f,g)||_{\mathcal{F}_p^{n-2}(T)} + (E||u(0)||^p_{n-2/p,p})^{1/p}$ in place of $||u||_{\mathcal{H}_p^n(T)}$. Therefore, below we take $\tau = T$ and take the function $u \in \mathcal{H}_p^n(T)$ as a solution of (7.8). As in the proof of Theorem 4.2, we may and will assume that f and g are as in (4.7) and $u(0) \in C_0^\infty$.

After these preparations, we are going to prove assertion (i). By Remark 3.4, without loss of generality we assume that $n = 2\beta$. Denote

$$u_1(t) = T_t u_0 + \int_0^t T_{t-s} f(s)\, ds.$$

It is easy to see that

$$u_1(r+\gamma) - u_1(r) = (T_\gamma - 1)u_1(r) + \int_0^\gamma T_{\gamma-\rho} f(r+\rho)\, d\rho.$$

Therefore, $E||u_1(r+\gamma) - u_1(r)||_p^p \le N(A_1(r,\gamma) + B_1(r,\gamma))$, where

$$A_1(r,\gamma) = E||(T_\gamma - 1)u_1(r)||_p^p, \;\; B_1(r,y) = E||\int_0^\gamma T_{\gamma-\rho} f(r+\rho)\, d\rho||_p^p,$$

and, by Lemma 7.4,

$$E||u_1(t) - u_1(s)||_p^p \le N(t-s)^{\alpha p - 1}(I_1(t,s) + J_1(t,s)),$$

with

$$I_1(t,s)=\int_0^{t-s}\frac{d\gamma}{\gamma^{1+\alpha p}}\int_s^{t-\gamma}A_1(r,\gamma)\,dr,\quad J_1(t,s)=\int_0^{t-s}\frac{d\gamma}{\gamma^{1+\alpha p}}\int_s^{t-\gamma}B_1(r,\gamma)\,dr.$$

By using Hölder's inequality and Lemma 7.3 and observing that $(\beta-1)q>-1$ for $q=p/(p-1)$, we get (remember $n=2\beta$)

$$B_1(r,\gamma)=E\int_{\mathbb{R}^d}|\int_0^\gamma \rho^{\beta-1}\rho^{1-\beta}T_\rho f(r+\gamma-\rho)\,d\rho|^p\,dx$$
$$\leq (\int_0^\gamma \rho^{(\beta-1)q}\,d\rho)^{p/q}E\int_0^\gamma \rho^{(1-\beta)p}\int_{\mathbb{R}^d}|T_\rho f(r+\gamma-\rho)|^p\,dxd\rho$$
$$\leq N\gamma^{\beta p-1}E\int_0^\gamma ||f(r+\gamma-\rho)||^p_{n-2,p}\,d\rho=N\gamma^{\beta p-1}E\int_0^\gamma ||f(r+\rho)||^p_{n-2,p}\,d\rho.$$

This and the inequality $\alpha<\beta$ implies that

$$J_1(t,s)\leq N\int_0^{t-s}\frac{d\gamma}{\gamma^{2+(\alpha-\beta)p}}\int_0^\gamma d\rho E\int_s^{t-\gamma}||f(r+\rho)||^p_{n-2,p}\,dr$$
$$\leq N\int_0^{t-s}\frac{d\gamma}{\gamma^{2+(\alpha-\beta)p}}\int_0^\gamma d\rho E\int_0^t ||f(r)||^p_{n-2,p}\,dr$$
$$=N(t-s)^{(\beta-\alpha)p}E\int_0^t ||f(r)||^p_{n-2,p}\,dr.$$

To estimate I_1, we use Lemma 7.3 and Theorem 2.1 to obtain

$$A_1(r,\gamma)\leq N\gamma^{\beta p}E||u_1(r)||^p_{n,p}\leq N\gamma^{\beta p}\{E||u(0)||^p_{n-2/p,p}+E\int_0^r ||f(s)||^p_{n-2,p}\,ds\}$$

$$I_1(t,s)\leq\int_0^{t-s}\frac{d\gamma}{\gamma^{1+\alpha p}}\int_0^t A_1(r,\gamma)\,dr$$
$$\leq N(t-s)^{(\beta-\alpha)p}\{E||u(0)||^p_{n-2/p,p}+E\int_0^t ||f(r)||^p_{n-2,p}\,dr\}.$$

For $u_2:=u-u_1$ we have

$$u_2(r+\gamma)-u_2(r)=(T_\gamma-1)u_2(r)+\int_r^{r+\gamma}T_{r+\gamma-\rho}g^k(\rho)\,dw^k_\rho,$$
$$E||u_2(r+\gamma)-u_2(r)||^p_p\leq N(A_2(r,\gamma)+B_2(r,\gamma)),$$

where

$$A_2(r,\gamma)=E||(T_\gamma-1)u_2(r)||^p_p,\quad B_2(r,\gamma)=E||\int_r^{r+\gamma}T_{r+\gamma-\rho}g^k(\rho)\,dw^k_\rho||^p_p.$$

By Lemma 7.4,

$$E||u_2(t)-u_2(s)||^p_p\leq N(t-s)^{\alpha p-1}(I_2(t,s)+J_2(t,s)),$$

with

$$I_2(t,s)=\int_0^{t-s}\frac{d\gamma}{\gamma^{1+\alpha p}}\int_s^{t-y}A_2(r,\gamma)\,dr,\quad J_2(t,s)=\int_0^{t-s}\frac{d\gamma}{\gamma^{1+\alpha p}}\int_s^{t-\gamma}B_2(r,\gamma)\,dr.$$

Here we use Lemma 7.3, the Burkholder–Davis–Gundy inequalities, and the inequality $(2\beta-1)q>-1$, where $q=p/(p-2)$. Then, similarly to the above calculations, we obtain

$$B_2(r,\gamma)\leq NE\int_{\mathbb{R}^d}[\int_0^\gamma \rho^{2\beta-1}\rho^{1-2\beta}||T_\rho g(r+\gamma-\rho)||^2_{l_2}\,d\rho]^{p/2}\,dx$$

$$\le N\gamma^{\beta p-1} E\int_0^{\gamma} \rho^{(1-2\beta)p/2} ||T_\rho g(r+\gamma-\rho)||_p^p \, d\rho$$

$$\le N\gamma^{\beta p-1} E\int_0^{\gamma} ||(1-\Delta)^{\beta-1/2} g(r+\gamma-\rho)||_p^p \, d\rho,$$

$$J_2(t,s) \le NE\int_0^{t-s} \frac{d\gamma}{\gamma^{2+(\alpha-\beta)p}} \int_s^{t-\gamma} dr \int_0^{\gamma} ||g(r+\gamma-\rho)||_{n-1,p}^p \, d\rho$$

$$\le NE\int_0^{t-s} \frac{d\gamma}{\gamma^{2+(\alpha-\beta)p}} \int_0^{\gamma} d\rho \int_0^t ||g(r)||_{n-1,p}^p \, dr$$

$$= N(t-s)^{(\beta-\alpha)p} E\int_0^t ||g(r)||_{n-1,p}^p \, dr.$$

Finally, again by Lemma 7.3 and Theorem 4.2, we conclude

$$I_2(t,s) \le NE\int_0^{t-s} \frac{d\gamma}{\gamma^{1+(\alpha-\beta)p}} \int_s^{t-\gamma} ||u_2(r)||_{2\beta,p}^p \, dr$$

$$\le N(t-s)^{(\beta-\alpha)p} E\int_0^t ||u_2(r)||_{n,p}^p \, dr \le N(t-s)^{(\beta-\alpha)p} E\int_0^t ||g(r)||_{n-1,p}^p \, dr.$$

Collecting all these estimates, we get (7.1) at least for $\eta = T \,(=\tau)$. In the general case, observe that if $u(t\wedge\eta) - u(s\wedge\eta)$ is not zero, then $s\wedge\eta = s$ and $s \le s\wedge\eta \le t\wedge\eta \le t$. After this it suffices to notice that, instead of points t and s on the left in (7.6), we can obviously take any two points between them.

The proof of (7.2) goes exactly the same way, the only difference being that this time we use (7.7).

To prove assertion (ii) notice that from an interpolation theorem (Theorem 2.4.2 in [**33**]) we have

$$||u||_{m(\theta)-d/p+d/q,q} \le N(d,p,q,m(0),m(1),\theta) ||u||_{m(0),p}^{1-\theta} ||u||_{m(1),p}^{\theta}$$

whenever $1 < p < q < \infty$, $\theta \in (0,1)$, $m(\theta) := (1-\theta)m(0) + \theta m(1) \ne m(0)$, $m(i) \le n$, and $u \in H_p^n$. Theorem 14.2 from [**14**] shows that the case $p = q$ actually need not be excluded. Note also that, under the conditions in (ii), there is a β such that $1/2 > \beta > 1/p$ and $m \le n - 2\beta(1-\theta) - d/p + d/q = m(\theta) - d/p + d/q$, where $m(\theta) := (1-\theta)(n-2\beta) + \theta n$. Therefore,

$$E\Big(\int_0^T ||u(t,\cdot)||_{m,q}^{p/\theta} \, dt\Big)^{\theta} \le E\Big(\int_0^T ||u(t,\cdot)||_{m(\theta)-d/p+d/q,q}^{p/\theta} \, dt\Big)^{\theta}$$

$$\le NE\Big(\int_0^T ||u(t,\cdot)||_{n-2\beta,p}^{(1-\theta)p/\theta} ||u(t,\cdot)||_{n,p}^p \, dt\Big)^{\theta}$$

$$\le NE \sup_{t\le T} ||u(t,\cdot)||_{n-2\beta,p}^{(1-\theta)p} \Big(\int_0^T ||u(t,\cdot)||_{n,p}^p \, dt\Big)^{\theta}.$$

To prove (ii), it remains only to apply Hölder's inequality, (3.4), and (7.2) with an α such that $\beta > \alpha > 1/p$. The theorem is proved.

8. Applications

8.1. Filtering Equation. Take $p \geq 2$ and an integer $d_1 > d$. Let $(\Omega, \mathcal{F}, P)$ be a complete probability space and let w_t be a d_1–dimensional Wiener process on this space. Consider a d_1–dimensional two–component process $z_t = (x_t, y_t)$ with x_t being d–dimensional and y_t $d_1 - d$–dimensional. We assume that z_t is defined as a solution of the system

$$\begin{aligned} dx_t =& b(t, z_t)\, dt + \theta(t, z_t)\, dw_t, \\ dy_t =& B(t, z_t)\, dt + \Theta(t, y_t)\, dw_t \end{aligned} \tag{8.1}$$

with some initial data. The coefficients of (8.1) are assumed to be vectors or matrices of appropriate dimensions satisfying the following assumptions.

ASSUMPTION 8.1. The functions b, θ, B, Θ are Borel measurable and bounded functions of their arguments. Each of them satisfies the Lipschitz condition with respect to z with constant K. Moreover, $\theta(t, x, y)$ is continuously differentiable with respect to x (not z) and its first derivatives with respect to x^is are continuous in y and satisfy the Lipschitz condition, with constant K, with respect to x (not z).

ASSUMPTION 8.2. The symmetric matrix $\Theta\Theta^*$ is invertible and $\Psi := (\Theta\Theta^*)^{-1/2}$ is a bounded function (of (t, y)).

ASSUMPTION 8.3. For any $\xi \in \mathbb{R}^d$ and $z = (x, y) \in \mathbb{R}^{d_1}$ and $t > 0$,

$$|Q(t, y)\theta^*(t, z)\xi|^2 \geq \delta|\xi|^2,$$

where Q is the orthogonal projector on $\ker \Theta$. In other words

$$(\theta(1 - \Theta^*\Psi^2\Theta)\theta^*\xi, \xi) \geq \delta|\xi|^2.$$

One can easily check that this assumption is satisfied if the diffusion matrix of system (8.1) is block upper–triangular:

$$\begin{pmatrix} \theta \\ \Theta \end{pmatrix} = \begin{pmatrix} \theta_1 & \theta_2 \\ 0 & \Theta_2 \end{pmatrix}$$

and the $d \times d$ matrix θ_1 is non-degenerate so that $|\theta_1^*\xi|^2 \geq \delta|\xi|^2$.

We also make the following assumption regarding the initial data for (8.1).

ASSUMPTION 8.4. The random vectors x_0, y_0 are independent of the process w_t. The conditional distribution of x_0 given y_0 has a density, which we denote by $\pi_0(x) = \pi_0(\omega, x)$. We have $\pi_0 \in L_p(\Omega, H_p^{2-2/p})$.

Let

$$a(t, x) = (1/2)\theta\theta^*(t, x, y_t), \quad \sigma(t, x) = \theta\Theta^*\Psi(t, x, y_t),$$
$$\beta(t, x) = \Psi B(t, x, y_t), \quad \beta_t = \beta(t, x_t),$$

and let β^k be the kth coordinate of the vector β where $k = 1, ..., d_1 - d$. Also, for a function u define

$$Lu = \sum_{i,j \leq d} (a^{ij}u)_{x^i x^j} - \sum_{i \leq d} (b^i u)_{x^i}, \quad \Lambda^k u = \beta^k u - \sum_{i \leq d} (\sigma^{ik} u)_{x^i}.$$

Finally, let $\mathcal{F}_t^y = \sigma\{y_s : s \leq t\}$.

THEOREM 8.1. *Under the above assumptions there exists $\pi \in \cap_{T>0}\mathcal{H}_p^2(T)$ such that for any $t \geq 0$, the conditional distribution of x_t given $\mathcal{F}_t^y$ has a density coinciding with $\pi(t, \cdot)$ almost surely. In addition, π satisfies the equation*

$$d\pi = L\pi\, dt + \sum_{k \leq d_1 - d} [\Lambda^k \pi - \bar{\beta}_t^k \pi]\, \{\Psi^k \Theta\, dw_t + [\beta_t^k - \bar{\beta}_t^k]\, dt\}, \tag{8.2}$$

with initial data $\pi(0,\cdot) = \pi_0$*, where the missing arguments are* t, x, y_t*, vector* Ψ^k *is the kth row of* Ψ*, and* $\bar{\beta}_t := (\pi(t,\cdot), \beta(t,\cdot,y_t))$.

To prove the theorem, we need a lemma, but first we make some comments. Notice that a little bit strange way of writing equation (8.2), with dt in two terms, is related to the fact that

$$\bar{w}_t := \int_0^t \{\Psi\Theta\, dw_s + [\beta_s - \bar{\beta}_s]\, ds\}$$

is a $d_1 - d$–dimensional $\mathcal{F}_t^y$–adapted Wiener process (the so–called *innovation process*). Here we do not need this fact.

We will only use that in (8.2) we have $\Psi^k \Theta\, dw_t = d\tilde{w}_t^k$, where $\tilde{w}_t$ is a $d_1 - d$–dimensional $\mathcal{F}_t$–adapted Wiener process. This is easily proved by using Levy's theorem, since it follows from the definition of Ψ that $\langle \Psi\Theta\, dw_t\rangle = Idt$, where I is the $d_1 - d$ unit matrix. In turn, this shows that in notation (5.2) applied to (8.2) we have

$$\alpha^{ij} = \frac{1}{2} \sum_{k \le d_1 - d} \sigma^{ik}\sigma^{jk},$$

which, by Assumption 8.3, guarantees that Assumption 5.1 is satisfied for equation (8.3) below with a K and $\delta/2$ instead of δ. Assumption 8.1 provides certain smoothness of the coefficients of (8.3), which leads us to the conclusion that the function $\bar{\pi}$ considered in the following lemma exists by Theorem 5.1 and Remark 5.6. By the way, the requirement in Assumption 8.1 that first derivatives of $\theta(t,x,y)$ with respect to x^is be continuous in y, which is not needed at this point, will be needed later in some passages to the limit. It is worth noting that this requirement follows automatically from other requirements in Assumption 8.1. Indeed, one knows that if a function $f(x,y)$ is Lipschitz continuous in (x,y) and has bounded generalized second–order derivatives in x, then f_x is Hölder continuous in y with exponent $1/2$.

LEMMA 8.2. *Let* $\bar{\pi}$ *be a unique solution in* $\cap_T \mathcal{H}_p^1(T)$ *of the equation*

$$d\bar{\pi} = L\bar{\pi}\, dt + \sum_{k \le d_1 - d} \Lambda^k \bar{\pi}\, \{\Psi^k \Theta\, dw_t + \beta_t^k\, dt\} \tag{8.3}$$

with initial data $\bar{\pi}(0,\cdot) = \pi_0$*, where we use the same notation as in (8.2). Then* $\bar{\pi}(t,\cdot)$ *is* $\mathcal{F}_t^y$*–measurable.*

Proof. Observe that now we write dt in two terms on the right for the reason that $\Psi^k\Theta\, dw_t + \beta_t^k\, dt = \Psi^k\, dy_t$, which makes the assertion quite natural since everything in (8.3) depends only on y_t. Further (cf. [**24**] or [**30**]), there is an $\mathcal{F}_t^y$–adapted measurable process $\tilde{\beta}_t = (\tilde{\beta}_t^1, ..., \tilde{\beta}_t^{d_1 - d})$ such that the process

$$\tilde{w}_t := \int_0^t \{\Psi\Theta\, dw_s + [\beta_s - \tilde{\beta}_s]\, ds\} \tag{8.4}$$

is a $d_1 - d$–dimensional $\mathcal{F}_t^y$–adapted Wiener process (actually $\tilde{\beta}_t$ is a modification of $E\{\beta_t | \mathcal{F}_t^y\}$). By Itô's formula the function $\tilde{\pi}$ defined by

$$\tilde{\rho}_t \tilde{\pi}(t,x) := \bar{\pi}(t,x), \quad \tilde{\rho}_t := \exp\{\int_0^t \tilde{\beta}_s\, d\tilde{w}_s + \frac{1}{2}\int_0^t |\tilde{\beta}_s|^2\, ds\}$$

satisfies

$$d\tilde{\pi} = L\tilde{\pi}\, dt + \sum_{k \le d_1 - d} [\Lambda^k \tilde{\pi} - \tilde{\beta}_t^k \tilde{\pi}]\, d\tilde{w}_t^k \tag{8.5}$$

with initial data $\bar\pi(0,\cdot)=\pi_0$. One can consider this equation on the probability space $(\Omega,\mathcal{F},P)$ with filtration $\mathcal{F}^y_t$ and Wiener process $\tilde w_t$. Let us denote by $\tilde{\mathcal{H}}^n_p(\tau)$ the $\mathcal{H}^n_p(\tau)$ spaces constructed on the basis of $\mathcal{F}^y_t$ and $\tilde w_t$. Define

$$\tau_n := \inf\{t\ge 0 : \tilde\rho_t \le \frac{1}{n}\}\wedge n.$$

Observe that τ_n are $\mathcal{F}^y_t$–stopping times. By the theory from previous sections, equation (8.5) with initial data u_0 has a unique solution in $\tilde{\mathcal{H}}^1_p(\tau_n)$ (even in $\tilde{\mathcal{H}}^2_p(\tau_n)$). On the other hand, we can replace $\tilde w_t$ in (8.5) with its value from (8.4) and get an equation in terms of the original $\mathcal{H}^n_p(\tau_n)$. This equation has a unique solution in the original $\mathcal{H}^1_p(\tau_n)$. Since $\tilde{\mathcal{H}}^1_p(\tau_n)\subset\mathcal{H}^1_p(\tau_n)$, this implies that any $\mathcal{H}^1_p(\tau_n)$–solution belongs to $\tilde{\mathcal{H}}^1_p(\tau_n)$, and since $\tilde\pi$ is obviously a $\mathcal{H}^1_p(\tau_n)$–solution, we get that $\tilde\pi(t\wedge\tau_n,\cdot)$ is $\mathcal{F}^y_t$–adapted, which, for $n\to\infty$, implies that $\tilde\pi(t,\cdot)$ is $\mathcal{F}^y_t$–adapted. Finally, $\tilde\rho_t$ is $\mathcal{F}^y_t$–adapted, and therefore $\bar\pi(t,\cdot)$ is $\mathcal{F}^y_t$–adapted. The lemma is proved.

Proof of Theorem 8.1. We imitate some proofs from [**24**] or [**30**] omitting some simple arguments which the reader can find there. First, we know (cf. Theorems 2.1 and 2.2 of [**24**]) that there is an $\mathcal{F}^y_t$–adapted measure–valued process π_t such that, for any $\phi\in C^\infty_0$ and $t\ge 0$, we have $E\{\phi(x_t)|\mathcal{F}^y_t\}=(\pi_t,\phi)$ (a.s.). We also know that π_t satisfies the following equation for any $\phi\in C^\infty_0$

$$d(\pi(t,\cdot),\phi)=(\pi(t,\cdot),L^*\phi)\,dt+(\pi(t,\cdot),\Lambda^{k*}\phi-\bar\beta^k_t\phi)\,\{\Psi^k\Theta\,dw_t+[\beta^k_t-\bar\beta^k_t]\,dt\}, \tag{8.6}$$

where $\bar\beta_t=(\pi(t,\cdot),\beta(t,\cdot,y_t))$ and L^*, Λ^* are the operators formally adjoint to L, Λ.

Now, let us assume for a moment that, in addition to our hypotheses, each derivative of b,B,θ with respect to x exists and is bounded. Notice that, by Sobolev's embedding theorem, any measure on $\mathbb{R}^d$ belongs to H^{-n}_p with $n>(p-1)d/p$. This implies that π_t is a $\cap_{T>0}\mathcal{H}^{-n}_p(T)$–solution of (8.6). If we look at $\bar\beta$ as a given process, then by Theorem 5.1 and Remark 5.6 there is a *unique* solution of (8.6) in the class $\cap_{T>0}\mathcal{H}^{-n}_p(T)$ with the initial condition $\pi(0,\cdot)=\pi_0$. On the other hand, since the coefficients are infinitely differentiable (but π_0 belongs only to $L_p(\Omega,H^{2-2/p}_p)$), by the same Theorem 5.1 and Remark 5.6 equation (8.6) has a solution in $\cap_{T>0}\mathcal{H}^2_p(T)$. Owing to the uniqueness, this solution is π. Thus $\pi\in\cap_{T>0}\mathcal{H}^2_p(T)$, which allows us to rewrite (8.6) as (8.2) after integrating by parts, and finishes the proof in the particular case of very smooth coefficients.

Passing to the general case notice that we only need to prove that $\pi\in\cap_{T>0}\mathcal{H}^2_p(T)$, since equation (8.2) is derived from (8.6) as above. Further, take a nonnegative function $\zeta\in C^\infty_0$ with unit integral and for $\varepsilon>0$ define $\zeta_\varepsilon(x)=\varepsilon^{-d}\zeta(x/\varepsilon)$, $h^{(\varepsilon)}(t,x,y)=h(t,x,y)*\zeta_\varepsilon(x)$, where the convolution is taken with respect to x. Consider the process $z^\varepsilon_t=(x^\varepsilon_t,y^\varepsilon_t)$ defined as a solution of

$$\begin{aligned} dx^\varepsilon_t &= b^{(\varepsilon)}(t,z^\varepsilon_t)\,dt+\theta^{(\varepsilon)}(t,z^\varepsilon_t)\,dw_t,\\ dy^\varepsilon_t &= B^{(\varepsilon)}(t,z^\varepsilon_t)\,dt+\Theta(t,y^\varepsilon_t)\,dw_t \end{aligned} \tag{8.7}$$

with initial condition $z^\varepsilon_0=z_0$. If we denote $\mathcal{F}^\varepsilon_t=\sigma\{y^\varepsilon_s : s\le t\}$ and $\pi^\varepsilon(t,x)$ the conditional distribution of x^ε_t given $\mathcal{F}^\varepsilon_t$, then, by the above, $\pi^\varepsilon\in\cap_{T>0}\mathcal{H}^2_p(T)$ and π^ε satisfies the following version of (8.2)

$$d\pi^\varepsilon = L_\varepsilon\pi^\varepsilon\,dt+\sum_{k\le d_1-d}[\Lambda_{\varepsilon k}\pi^\varepsilon-\bar\beta^k_{\varepsilon t}\pi^\varepsilon]\,\{\Psi^k_\varepsilon\Theta_\varepsilon\,dw_t+[\beta^k_{\varepsilon t}-\bar\beta^k_{\varepsilon t}]\,dt\}, \tag{8.8}$$

with initial data $\pi^\varepsilon(0,\cdot)=\pi_0$, where $L_\varepsilon,\Lambda_\varepsilon,\Psi_\varepsilon,\Theta_\varepsilon,\beta_{\varepsilon t}$ are defined in the same way as $L,\Lambda,\Psi,\Theta,\beta_t$ with y^ε_t in place of y_t, and $\bar\beta_{\varepsilon t}:=(\pi^\varepsilon(t,\cdot),\beta_\varepsilon(t,\cdot,y^\varepsilon_t))$.

By Theorem 5.1, we have $||\pi^{\varepsilon}||_{\mathcal{H}^2_p(T)} \leq N$ for any T, where N is independent of ε. By Theorem 3.11 there is a $\pi^0 \in \cap_{T>0}\mathcal{H}^2_p(T)$ and a sequence $\varepsilon_n \to 0$ such that, for any $\xi \in L_q(\Omega)$ with $q = p/(p-1)$ and $\phi \in C_0^\infty$ and $t \geq 0$,

$$\lim_{n\to\infty} E(\pi^{\varepsilon_n}(t,\cdot),\phi)\xi = E(\pi^0(t,\cdot),\phi)\xi. \tag{8.9}$$

This equality and the fact that $(\pi^{\varepsilon_n}(t,\cdot),\phi) = E\{\phi(x_t^{\varepsilon_n})|\mathcal{F}_t^{\varepsilon_n}\}$ (a.s.) can be used instead of equality (3.14) of [**24**]. As there, first for ξ of type $g(y_{t_1},...,y_{t_m})$ and then for arbitrary $\mathcal{F}_t^y$–measurable $\xi \in L_q(\Omega)$, we get that

$$E\xi(\pi^0(t,\cdot),\phi) = E\xi E\{\phi(x_t)|\mathcal{F}_t^y\}.$$

If we knew that $\pi^0(t,\cdot)$ is $\mathcal{F}_t^y$–measurable, then this equality would imply that

$$(\pi^0(t,\cdot),\phi) = E\{\phi(x_t)|\mathcal{F}_t^y\} \quad \text{(a.s.)} \tag{8.10}$$

which is nothing but assertion that the conditional distribution of x_t given $\mathcal{F}_t^y$ has a density coinciding with $\pi^0(t,\cdot)$. Since $\pi^0 \in \cap_{T>0}\mathcal{H}^2_p(T)$, equality (8.10) would imply that π^0 satisfies (8.2) in the same way as at the beginning of the proof. Therefore, in the remaining part of the proof we only need to prove the $\mathcal{F}_t^y$–measurability of $\pi^0(t,\cdot)$.

By Itô's formula and (8.8), the function

$$\bar{\pi}^{\varepsilon}(t,x) := \pi^{\varepsilon}(t,x)\rho_t^{\varepsilon}, \tag{8.11}$$

$$\rho_t^{\varepsilon} := \exp\{\int_0^t (\Theta_\varepsilon^*\Psi_\varepsilon\bar{\beta}_{\varepsilon s}, dw_s) + \int_0^t (\bar{\beta}_{\varepsilon s},\beta_{\varepsilon s})\,ds - \frac{1}{2}\int_0^t |\bar{\beta}_{\varepsilon s}|^2\,ds\}$$

satisfies the linear equation

$$d\bar{\pi}^{\varepsilon} = L_\varepsilon\bar{\pi}^{\varepsilon}\,dt + \sum_{k\leq d_1-d}\Lambda_\varepsilon^k\bar{\pi}^{\varepsilon}\,\{\Psi_\varepsilon^k\Theta_\varepsilon\,dw_t + \beta_{\varepsilon t}^k\,dt\},$$

with initial data $\bar{\pi}^{\varepsilon}(0,\cdot) = \pi_0^{\varepsilon}$. By using that $\sup_{t\leq T}|y_t^{\varepsilon} - y_t| \to 0$ in probability and by applying Theorem 5.7 with $n = -1$ (here we also use Remark 5.8 and the continuity of θ_x with respect to y), one easily proves that $\bar{\pi}^{\varepsilon} \to \bar{\pi}$ in $\mathcal{H}^1_p(T)$ for any T, where $\bar{\pi}$ satisfies (8.3). By Lemma 8.2 we have that $\bar{\pi}(t,\cdot)$ is $\mathcal{F}_t^y$–adapted.

Next, by using the fact that $\pi^{\varepsilon}(t,x)$ is a probability density, one gets from (8.11) that (cf. [**24**])

$$\rho_t^{\varepsilon} \to \int_{\mathbb{R}^d}\bar{\pi}(t,x)\,dx =: \bar{\rho}_t$$

in probability and the last term is again $\mathcal{F}_t^y$–adapted. The process $\bar{\rho}_t$ satisfies the equation obtained by integrating (8.3), which means that

$$d\bar{\rho}_t = \bar{\rho}_t\hat{\beta}_t\,dy_t, \tag{8.12}$$

where

$$\hat{\beta}_t = \int_{\mathbb{R}^d}\beta(t,x)\bar{\pi}(t,x)\,dx(\int_{\mathbb{R}^d}\bar{\pi}(t,x)\,dx)^{-1} \quad (0\cdot 0^{-1} := 0).$$

It follows from (8.12) and the boundedness of β and $\hat{\beta}_t$ that $\bar{\rho}_t \neq 0$ with probability one.

Thus, for any $\phi \in C_0^\infty$,

$$(\pi^{\varepsilon}(t,\cdot)\rho_t^{\varepsilon},\phi) \to (\bar{\pi}(t,\cdot),\phi),$$

$$(\pi^{\varepsilon}(t,\cdot),\phi) = \frac{(\pi^{\varepsilon}(t,\cdot)\rho_t^{\varepsilon},\phi)}{\rho_t^{\varepsilon}} \to (\bar{\pi}(t,\cdot),\phi)(\int_{\mathbb{R}^d}\bar{\pi}(t,x)\,dx)^{-1}$$

in probability. This implies that

$$(\pi^0(t,\cdot),\phi) = (\bar{\pi}(t,\cdot),\phi)(\int_{\mathbb{R}^d}\bar{\pi}(t,x)\,dx)^{-1} \qquad \text{(a.s.)},$$

where the last expression is $\mathcal{F}_t^y$–adapted. The theorem is proved.

REMARK 8.3. This theorem, along with embedding theorems from Sec. 7, extends and generalizes the corresponding results from [**24**] and [**30**] in the way we were talking about in the Introduction.

8.2. On the Notion of Stochastic Integral. To consider applications to equations with infinitely dimensional Wiener processes, we want to discuss the notion of stochastic integral and show that "basically" there is nothing more general than series of usual one–dimensional stochastic integrals. This will show that equations like (2.1), containing series of integrals with respect to Wiener processes, are of a quite general nature.

The first notion was introduced by Paley, Wiener, and Zygmund in [**29**], where the stochastic integral of a *nonrandom smooth* function $f(t)$ against a one-dimensional Wiener process w_t is defined as

$$\int_0^1 f(s)\,dw_s = f(1)w_1 - \int_0^1 w_s f'(s)\,ds. \tag{8.13}$$

Then it is verified that the $L_2(\Omega)$–norm of the stochastic integral is equal to the $L_2(0,1)$–norm of f, which allows one to extend the stochastic integral from smooth functions to all $f \in L_2(0,1)$. One obtains the same integral if, instead of (8.13), one starts with

$$\int_0^1 f(s)\,dw_s = \sum_{i=1}^n a_i(w_{s_i} - w_{s_{i-1}}) \tag{8.14}$$

for functions f such that $f(t) = a_i$ on $(s_{i-1}, s_i]$ and $0 = s_0 \le s_1 \le \dots \le s_n = 1$.

The definition based on (8.14) has an advantage that it can be easily generalized to define an integral of a *nonrandom* function against a random orthogonal measure on a σ–finite measure space $(X, \mathcal{X}, m)$. More precisely, assume that we are given a π–system Π of subsets of X such that $\sigma(\Pi) = \mathcal{X}$, and a random (complex–valued) variable $\mu(\gamma)$ defined for each $\gamma \in \Pi$ (and perhaps not for all $\gamma \in \mathcal{X}$). Assume that $\mu(\gamma) \in L_2(\Omega)$ and $E\mu(\gamma_1)\bar{\mu}(\gamma_2) = m(\gamma_1 \cap \gamma_2)$ for each $\gamma, \gamma_1, \gamma_2 \in \Pi$. Then for functions

$$f(x) = \sum_{i=1}^n a_i I_{\gamma_i}(x),$$

where a_is are constant and $\gamma_i \in \Pi$, one defines the stochastic integral of f against μ as

$$\int_X f(x)\,\mu(dx) = \sum_{i=1}^n a_i \mu(\gamma_i), \tag{8.15}$$

and, again by isometry, one extends the stochastic integral to all $f \in L_2(X, m)$. Such integrals are used in the theory of stationary processes. Surprisingly enough, as we will see, one can also say that this is the most general stochastic integral in Itô's stochastic calculus.

Another advantage of (8.14) is that one can allow f to depend on ω, and if a_i are independent of the process $w_{t+s_{i-1}} - w_{s_{i-1}}$, $t \ge 0$, then (8.14) is again an isometry between a part of $L_2(\Omega \times (0,1])$ and a part of $L_2(\Omega)$. Closing this isometry, K. Itô defines his famous integral.

It turns out that Itô's integral is a particular case of the integral based on (8.15). To be more precise, let w_t be (as usual) a Wiener process with respect to a filtration $\mathcal{F}_t$, $\mathcal{P}$ be the predictable σ–field on $\Omega \times (0,1]$, and Π be the set of all stochastic intervals $(0,\tau]\!]$, where τ are stopping times ≤ 1. Then one gets Itô's integral by taking $X = \Omega \times (0,1]$, $\mathcal{X} = \mathcal{P}$,

$\mu(0,\tau] = w_\tau$ (see more about this in [**19**]). In the same way one defines stochastic integrals with respect to any locally square integrable martingale.

K. Itô [**12**] was also the first to consider integration against measure–valued processes, which is a particular case of integration against martingale measures. Let $p(t,\Gamma)$, $t \geq 0$, be a square integrable process as a function of t with independent increments in time and a random orthogonal measures as a function of Γ for any t. Define $p((s,t],\Gamma) = p(t,\Gamma) - p(s,\Gamma)$. Itô's way of introducing the integral with respect to p is to replace the expression $a_i(w_{s_i} - w_{s_{i-1}})$ in (8.14) with

$$\int_X f_i(x)\, p((s_{i-1},s_i],dx) = \int_X f_i(x)\, p(s_i,dx) - \int_X f_i(x)\, p(s_{i-1},dx), \tag{8.16}$$

where f_i are assumed to be independent of the processes $p((s_{i-1},t],\Gamma)$, $t \geq s_{i-1}$.

More generally, for any $\gamma \in \Pi$ let a process $p(t,\gamma)$ be given, which is a square integrable martingale with respect to a given filtration $\{\mathcal{F}_t\}$. Let $\langle p(\cdot,\gamma_1), p(\cdot,\gamma_2)\rangle_t = q(t,\gamma_1 \cap \gamma_2)$, where $q(t,\cdot)$ is a σ–finite measure on $(X,\mathcal{X})$ for any ω, t, and $q(t,\Gamma)$ increases in t for any $\Gamma \in \mathcal{X}$ and ω. Then there is a measure $q(dt,dx)$ such that

$$q(t,\gamma) = \int_0^1 \int_\gamma q(ds,dx).$$

By following Itô's method based on (8.16), for any $\mathcal{P} \times \mathcal{X}$–measurable $f = f(\omega,t,x)$ such that

$$\int_0^1 \int_X f^2(s,x)\, q(ds,dx) < \infty,$$

one defines the stochastic integral

$$\int_0^1 \int_X f(s,x)\, p(ds,dx). \tag{8.17}$$

This integral is also a particular case of the integral of a *nonrandom* function against a random orthogonal measure. Indeed, define $\bar{X} = \Omega \times (0,1] \times X$, $\bar{\mathcal{X}} = \mathcal{P} \times \mathcal{X}$, and let $m(d\omega dt dx) := P(d\omega) q(dt,dx)$. Also let

$$\bar{\Pi} = \{ (0,\tau]\!] \times \gamma : Eq((0,\tau]\!] \times \gamma) < \infty\}, \quad \mu((0,\tau]\!] \times \gamma) = p(\tau,\gamma).$$

Then on functions

$$f = \sum_{i \leq n} a_i I_{t \leq \tau_i} I_{\gamma_i}(x), \tag{8.18}$$

where $(0,\tau_i]\!] \times \gamma_i \in \bar{\Pi}$ and a_i are some constants, integral (8.17) equals

$$\sum_{i \leq n} a_i p(\tau_i,\gamma_i) = \sum_{i \leq n} a_i \mu((0,\tau_i]\!] \times \gamma_i),$$

which agrees with (8.15). Finally, by functions of type (8.18) one can approximate any $\mathcal{P} \times \mathcal{X}$–measurable function for which (8.17) can be defined.

It is worth mentioning that there are also other notions of martingale measures with respect to which one can define stochastic integration (see [**35**], where the martingale measures discussed above are called orthogonal martingale measures).

Even though the notion of integral of nonrandom functions with respect to random orthogonal measures is very convenient for the purpose of introducing Itô's stochastic integrals (cf. [**19**]), one works almost always with stochastic integrals with variable limits, and a different notation is more appropriate.

In connection with this notice that it is shown in [**10**] how to reduce the stochastic integral with respect to a martingale measure to a series of usual stochastic integrals. This was further used in [**10**] to treat *stochastic equations* containing integrals against martingale measures using *the same notation* as in the case of equations containing just usual stochastic integrals.

To be more precise, it is assumed in [**10**] that $\mathcal{X}$ is countably generated and

$$q(t,\Gamma)=\int_0^t q_s(\Gamma)\,dV_s,$$

where V_s is a predictable increasing process and $q_s(\Gamma)$ is a measure in Γ for any s and predictable in s for any $\Gamma\in\mathcal{X}$. Then it is shown that

$$\int_0^t\int_X f(s,x)\,p(ds,dx)=\sum_{k=1}^{\infty}\int_0^t f_k(s)\,dp_s^k, \tag{8.19}$$

where

$$p_t^k=\int_0^t\int_X \eta_k(s,x)\,p(ds,dx),\quad f_k(s)=\int_X \eta_k(s,x)f(s,x)\,q_s(dx),$$

and for any ω,s the system of functions $\{\eta_k(s,\cdot)\}$ forms an orthonormal basis in $L_2(X,q_s)$.

A particular case of the stochastic integral with respect to a martingale measure is the stochastic integral with respect to the two–dimensional Brownian sheet $W(t,x)$ defined for $t\geq 0, x\in\mathbb{R}$. In this case, one takes $\mathcal{F}_t$ so that the random variables $W(t,x)$ are $\mathcal{F}_t$–measurable, and defines $p((0,\tau]\!]\times(a,b])=W(\tau,b)-W(\tau,a)$. This integral got very popular thanks to the article [**35**]. One can construct $W(s,x)$ by taking independent one–dimensional Wiener processes $w_t^k, k\geq 1$, and an orthonormal basis $\{\eta_k(x),k\geq 1\}$ in $L_2(\mathbb{R})$, and letting

$$W(t,x)=\sum_{k=1}^{\infty}w_t^k\int_0^x \eta_k(y)\,dy\quad t\geq 0, x\in\mathbb{R}.$$

Incidentally, observe that thus defined W is a Gaussian field and

$$EW(s,y)W(t,x)=s\wedge t\sum_{k=1}^{\infty}\int_0^x\eta_k(z)\,dz\int_0^y\eta_k(z)\,dz,$$

where

$$\sum_{k=1}^{\infty}\int_0^x\eta_k(z)\,dz\int_0^y\eta_k(z)\,dz=\begin{cases}\int_{\mathbb{R}}I_{(0,x)}(z)I_{(0,y)}(z)\,dz=x\wedge y & \text{if}\quad x,y\geq 0,\\ \int_{\mathbb{R}}I_{(x,0)}(z)I_{(y,0)}(z)\,dz=|x|\wedge|y| & \text{if}\quad x,y\leq 0,\\ 0 & \text{otherwise.}\end{cases}$$

In this particular case (8.19) becomes

$$\int_0^t\int_{\mathbb{R}}f(s,x)\,W(ds,dx)=\sum_{k=1}^{\infty}\int_0^t\left\{\int_{\mathbb{R}}\eta_k(x)f(s,x)\,dx\right\}dw_s^k. \tag{8.20}$$

By the way, general one–dimensional equations driven by the cylindrical space–time white noise $\dot{B}_t$ were considered in [**7**], where the right–hand side of (8.20) is taken by definition as $\int_0^t\langle f(s,\cdot),dB_t\rangle$. There the series was introduced from the very beginning.

The last integral we want to discuss is the integral against a Hilbert–space valued Wiener processes (see, for instance, [**30**]). Let H be a Hilbert space and w_t be a H–valued Wiener process with covariance operator Q. This operator is known to be nuclear. If h_k are its unit eigenvectors with nonzero eigenvalues, then $w_t^k := (h^k, w_t)(Qh^k, h^k)^{-1/2}$ are independent standard Wiener processes and, for any H–valued process f_t for which the integral $\int_0^t f_t \cdot dw_t$ is defined, the integral can be written as

$$\sum_k \int_0^t f_s^k \, dw_s^k,$$

where $f_s^k = (f_s, h^k)(Qh^k, h^k)^{1/2}$ (see, for instance, [**30**]).

8.3. Equations Driven by Space–Time White Noise. In this subsection, we consider one-dimensional equations with space-time white noise. Thus $d = 1$. Very often (see, for instance [**7**]) one writes these equations in the form

$$du(t,x) = [a(t,x)u''(t,x) + b(t,x)u'(t,x) + f(t,x,u(t,x))]\, dt + h(t,x,u(t,x))\, dB_t, \quad (8.21)$$

where B_t is a cylindrical Wiener process on L_2. There are also very many articles (see, for instance [**8**]) where instead of dB_t one writes $(\partial^2 W/\partial t \partial x)\, dt$. As we explained in Subsec. 8.2 we may as well take $\sum_k \eta_k(x)\, dw_t^k$, where $\{\eta_k(x), k \geq 1\}$ is an orthonormal basis in L_2 and w_t^k are independent $\mathcal{F}_t$–adapted one–dimensional Wiener processes. Thus, instead of (8.21), we will be considering the equation

$$\begin{aligned} du(t,x) =& [a(t,x)u''(t,x) + b(t,x)u'(t,x) + f(t,x,u(t,x))]\, dt \\ &+ g^k(t,x,u(t,x))\, dw_t^k \end{aligned} \quad (8.22)$$

on a time interval $[0, \tau]$, where $g^k := h\eta_k$ and τ is a bounded stopping time.

ASSUMPTION 8.5. The functions $a(t,x) = a(\omega,t,x)$ and $b(t,x) = b(\omega,t,x)$ are real-valued functions defined on $(0,\tau]\!] \times \mathbb{R}$.

(*i*) For any ω and $t \leq \tau(\omega)$, $a(\omega,t,\cdot) \in C^{1,1}$ $(= B^2)$ and $b(\omega,t,\cdot) \in C^{0,1}$ $(= B^1)$ and $||a||_{C^{1,1}} + ||b||_{C^{0,1}} \leq K$. Also $K \geq a \geq \delta$.

(*ii*) For any $x \in \mathbb{R}$, the processes a and b are predictable.

To state the next assumption, take and fix $s \leq \infty$ and finite κ, p, r such that

$$\kappa \in (0, 1/2], \quad p \geq 2r \geq 2, \quad s \leq \infty, \quad \frac{1}{r} + \frac{1}{s} = 1, \quad r < \frac{1}{1 - 2\kappa}. \quad (8.23)$$

ASSUMPTION 8.6. The functions $f(t,x,u)$ and $h(t,x,u)$ are real-valued functions on $(0,\tau]\!] \times \mathbb{R}^2$ such that

(*i*) for any x and u, the processes $f(t,x,u)$ and $h(t,x,u)$ are predictable;

(*ii*) for any ω, t, x, u, and v,

$$|f(t,x,u) - f(t,x,v)| \leq K|u - v|, \quad |h(t,x,u) - h(t,x,v)| \leq \xi(t,x)|u - v|, \quad (8.24)$$

where ξ is certain function of ω, t, x satisfying $||\xi(t,\cdot)||_{2s} \leq K$.

Observe that one of possibilities is $r = 1$, and then $s = \infty$ and (8.24) just means that both f and h satisfy the Lipschitz condition in u with constant K.

To fit equation (8.22) in our general scheme, we need to find an appropriate n such that the assumptions of Theorem 5.1 are satisfied. Define

$$n = -\kappa - 3/2.$$

In the following lemma, we also set

$$R(x) = \chi |x|^{-(1-2\kappa)/2} \int_0^\infty t^{-(5-2\kappa)/4} e^{-tx^2 - 1/(4t)}\, dt,$$

so that, according to (2.3) and (2.4) for right choice of the constant χ, the function $R(x)$ is the kernel of the operator $(1-\Delta)^{(n+1)/2}$. It is easy to prove that R is infinitely differentiable everywhere except the origin, decreases exponentially fast as $|x| \to \infty$, and behaves near the origin like $|x|^{-(1-2\kappa)/2}$ if $\kappa < (1/2)$ and like $-\log|x|$ if $\kappa = 1/2$. Finally, notice that generally speaking, the notation h, ξ, g is used in the lemma for functions different from the ones introduced above.

LEMMA 8.4. *Take some functions $h \in L_p$, $\xi \in L_{2s}$, and set $g^k = \xi h \eta_k$. Then $g = \{g^k\}_{k\geq 1} \in H_p^{n+1}(l_2)$ and*

$$||g||_{n+1,p} = ||\bar{h}||_p \leq (N||\xi||_{2s}||h||_p) \wedge (||\xi h||_2 ||R||_p), \tag{8.25}$$

where $N = ||R||_{2r} < \infty$ and

$$\bar{h}(x) := \left\{ \int_{\mathbb{R}} R^2(x-y)\xi^2(y)h^2(y)\,dy \right\}^{1/2}. \tag{8.26}$$

In addition, if $p(1-2\kappa) > 2$ and $\xi = 1$, then

$$||g||_{n+1,p} \leq N(\kappa)||h||_2^{2\kappa p/(p-2)}||h||_p^{1-2\kappa p/(p-2)}. \tag{8.27}$$

Proof. We know that

$$(1-\Delta)^{(n+1)/2}(\xi h \eta_k)(x) = \int_{\mathbb{R}} R(x-y)\xi(y)h(y)\eta_k(y)\,dy.$$

It follows that by Parseval's theorem

$$|(1-\Delta)^{(n+1)/2}g(x)|_{l_2}^2 = \sum_{k=1}^{\infty} \left(\int_{\mathbb{R}} R(x-y)\xi(y)h(y)\eta_k(y)\,dy \right)^2$$

$$= \int_{\mathbb{R}} R^2(x-y)\xi^2(y)h^2(y)\,dy = \bar{h}^2(x).$$

We thus get the equality in (8.25). To prove the inequality, notice that $\bar{h}^2$ is a convolution and by Minkowski's inequality the L_p–norm of a convolution is less than the L_1–norm of one function times L_p–norm of another. This immediately gives $||\bar{h}||_p \leq ||\xi h||_2 ||R||_p$. Also, by Hölder's inequality,

$$\bar{h}^2(x) \leq ||\xi||_{2s}^2 \left(\int_{\mathbb{R}} R^{2r}(x-y)h^{2r}(y)\,dy \right)^{1/r},$$

which leads to the second inequality in (8.25): $||\bar{h}||_p \leq N||\xi||_{2s}||h||_p$, again by Minkowski's inequality and by the assumption $p \geq 2r$. The finiteness of N follows from the assumption that $r < (1-2\kappa)^{-1}$.

To prove (8.27), we use that $R^2(y) \leq N|y|^{2\kappa-1}$ and we minimize with respect to $\varepsilon > 0$ after the following computations:

$$\Big(\int_{\mathbb{R}} \{\int_{\mathbb{R}} R^2(y)h^2(x-y)\,dy\}^{p/2}\,dx\Big)^{2/p} \leq \Big(\int_{\mathbb{R}} \{\int_{\mathbb{R}} I_{|y|\leq\varepsilon} R^2(y)h^2(x-y)\,dy\}^{p/2}\,dx\Big)^{2/p}$$

$$+\Big(\int_{\mathbb{R}} \{\int_{\mathbb{R}} I_{|y|\geq\varepsilon} R^2(y)h^2(x-y)\,dy\}^{p/2}\,dx\Big)^{2/p} \leq ||I_{|y|\leq\varepsilon}R^2||_1 ||h||_p^2 + ||I_{|y|\geq\varepsilon}R^p||_1^{2/p}||h||_2^2$$

$$\leq N\varepsilon^{2\kappa}||h||_p^2 + N\varepsilon^{-(1-2\kappa-2/p)}||h||_2^2.$$

The lemma is proved.

THEOREM 8.5. *Let $\kappa \in (0,1/2)$ and $u_0 \in L_p(\Omega, \mathcal{F}_0, H_p^{(1/2)-\kappa-2/p})$. Assume that*

$$I^p(\tau) := E\int_0^\tau \{||f(t,\cdot,0)||^p_{-3/2-\kappa,p} + ||\bar{h}(t,\cdot,0)||^p_p\}\,dt < \infty, \tag{8.28}$$

where

$$\bar{h}(t,x,0) := \left\{\int_{\mathbb{R}} R^2(x-y)h^2(t,y,0)\,dy\right\}^{1/2}.$$

Then, in the space $\mathcal{H}_p^{(1/2)-\kappa}(\tau)$, equation (8.22) with the initial condition u_0 has a unique solution u. Moreover,

$$||u||_{\mathcal{H}_p^{1/2-\kappa}(\tau)} \le N\{I(\tau) + (E||u_0||^p_{1/2-\kappa-2/p,p})^{1/p}\},$$

where the constant N depends only on κ, p, δ, K, and τ.

Proof. We will apply Theorem 5.1. Its assumptions concerning a and σ ($\equiv 0$) are obviously satisfied. Next, if $u \in H_p^{n+2}$, then $bu' \in H_p^{n+1} \subset H_p^n$, $u \in L_p$, $f(u) - f(0) \in L_p \subset H_p^n$. Also, by (8.28) we have $f(t,\cdot,0) \in H_p^n$ (a.e. on $(0,\tau]\!]$). Furthermore, by Lemma 5.2,

$$||bu'||_{n,p} \le ||bu'||_{-1,p} \le N||u'||_{-1,p} \le N||u||_p = N||u||_{n+2-((1/2)-\kappa),p},$$

$$||f(u) - f(v)||_{n,p} \le ||f(u) - f(v)||_p \le K||u-v||_p.$$

We emphasize that $||\cdot||_p = ||\cdot||_{n+2-((1/2)-\kappa),p}$ and $n+2-((1/2)-\kappa) < n+2$ for $\kappa < 1/2$. Consequently (see Remark 5.5), the assumptions of Theorem 5.1 concerning $bu' + f(u)$ are satisfied. To check the remaining assumptions about $g(u)$, it suffices to notice that, by Lemma 8.4, we have

$$||g(0)(t,\cdot)||_{n+1,p} = ||\bar{h}(t,\cdot,0)||_p, \quad ||g(u)(t,\cdot) - g(v)(t,\cdot)||_{n+1,p}$$
$$\le N||h(t,\cdot,u(t,\cdot)) - h(t,\cdot,v(t,\cdot))||_p \le N||u(t,\cdot) - v(t,\cdot)||_p.$$

The theorem is proved.

REMARK 8.6. We have obtained Theorem 8.5 by checking that all Assumptions 5.1–5.6 are satisfied for $n = -3/2 - \kappa$. After this, of course, all other results from Sec. 5 are available. For example, by the approximation theorem (Theorem 5.7) and Remark 5.9, we have $||u - u_m||_{\mathcal{H}_p^{n+2}(\tau)} \to 0$, where u_m is a unique solution of the following version of (5.11)

$$du(t,x) = [a(t,x)u''(t,x) + b(t,x)u'(t,x) + f(t,x,u(t,x))]\,dt + \sum_{k=1}^m h(t,x,u(t,x))\eta_k\,dw_t^k$$

with initial condition u_0.

REMARK 8.7. Additional information about Hölder continuity properties of the solution is readily obtained from the properties of elements of $\mathcal{H}_p^n(\tau)$ listed in Theorem 7.2.

For example,

$$u \in C^{\alpha-1/p}([0,\tau], H_p^{n+2-2\beta}) \qquad \text{(a.s.)} \tag{8.29}$$

provided $p > 2$ and $1/2 > \beta > \alpha > 1/p$. Here $H_p^{n+2-2\beta} \subset C^\gamma$ if

$$\gamma := n + 2 - 2\beta - 1/p = 1/2 - \kappa - 2\beta - 1/p > 0.$$

Hence, if the inequality $E||u_0||^p_{1/2-\kappa-2/p,p} < \infty$ and (8.28) are satisfied for any $\kappa \in (0,1/2)$ and $p \ge 2$ (say $f(t,x,0) = h(t,x,0) = 0$, $\xi = K$, and u_0 is a deterministic smooth function with compact support), then, after taking p large enough and κ, α, and β small, we see that u satisfies the Hölder condition in x of order $1/2 - \varepsilon$ uniformly with respect to $t \in [0,\tau]$ (a.s.) for any $\varepsilon > 0$.

On the other hand, by taking p large enough, both α and β close to $1/4$, and κ small, we get that u satisfies the Hölder condition in t of order $1/4-\varepsilon$ uniformly with respect to $x \in \mathbb{R}$ (a.s.) for any $\varepsilon > 0$. In terms of parabolic Hölder spaces, this means that

$$u \in C^{1/4-\varepsilon,1/2-\varepsilon}_{t\ \ ,\ \ x}([0,\tau]\times\mathbb{R}) \qquad \text{(a.s.)}.$$

This result agrees well with [**35**], where the same continuity is obtained for equations with constant coefficients. Also, notice that, using standard PDE methods, one can prove interior (with respect to t) estimates which would give similar continuity of u away from $t=0$ under weaker assumptions on u_0. One of results which can be obtained is discussed in the following remark.

REMARK 8.8. There are smoothing properties of equations in the sense that solutions may be much smoother than the initial data. For example, assume that $f(t,x,0) = h(t,x,0) = 0$, $\tau = T$, where T is a constant, $r=1$, and $s=\infty$. Let $u_0 \in L_2(\Omega,\mathcal{F}_0,H_2^{-1+\varepsilon})$ with $\varepsilon \in (0,1/2)$, so that u_0 may be a finite measure or just a delta–function. We claim then that, for any $\varepsilon \in (0,1/4)$,

$$u \in C^{1/4-\varepsilon,1/2-\varepsilon}_{t\ \ ,\ \ x}([\varepsilon,T]\times\mathbb{R}) \qquad \text{(a.s.)}.$$

Indeed, by Theorem 8.5 with $p=2$ and $\kappa = 1/2-\varepsilon$, there is a unique solution $u \in \mathcal{H}_2^{\varepsilon}(T)$ of (8.22) and

$$||u||^2_{\mathcal{H}_2^{\varepsilon}(T)} \le NE||u_0||^2_{-1/2-\kappa,2}.$$

By (3.4) we have

$$\int_0^T E||u(t,\cdot)||^2_{\varepsilon,2}\,dt = ||u||^2_{\mathbb{H}_2^{\varepsilon}(T)} \le NE||u_0||^2_{-1/2-\kappa,2} =: I.$$

By Chebyshev's inequality, for any $\gamma \in (0,T/2)$, Lebesgue measure of the set on $(\gamma,2\gamma)$, where $E||u(t,\cdot)||^2_{\varepsilon,2} > 2I/\gamma$, is less than $\gamma/2$. This implies that for any $\gamma \in (0,T/2)$ there is a point $t_\gamma \in (\gamma,2\gamma)$ such that

$$u(t_\gamma,\cdot) \in L_2(\Omega,\mathcal{F}_{t_\gamma},H_2^{\varepsilon}), \quad E||u(t_\gamma,\cdot)||^2_{\varepsilon,2} \le 2I/\gamma.$$

By considering (8.22) after time t_γ instead of 0 and defining the spaces $\mathcal{H}_2^{1/2-\kappa}(t_\gamma,T)$ in an obvious way, we get by Theorem 8.5 that $u \in \mathcal{H}_2^{1/2-\kappa}(t_\gamma,T)$ for *any* $\kappa \in (0,1/2)$. In addition,

$$\int_{t_\gamma}^T E||u(t,\cdot)||^2_{1/2-\kappa,2}\,dt \le NI/\gamma, \quad E||u(s_\gamma,\cdot)||^2_{1/2-\kappa,2} \le NI/\gamma^2, \tag{8.30}$$

where s_γ is a certain point in $(t_\gamma+\gamma, t_\gamma+2\gamma)$ and $\gamma \le T/4$.

Now we can go to larger powers. Take any $p>4$ and define $A(p,\gamma,R) = \{\omega : ||u(s_\gamma,\cdot)||_p \le R\}$. By Sobolev's embedding theorem, $H_2^{1/2-\kappa} \subset H_p^m$ if $m < 1/2-\kappa$ and $-\kappa = m-1/p$. For $m=0$ we have $\kappa = 1/p$. Since in (8.30) we can take $\kappa = 1/p$, we get

$$P\{A(p,\gamma,R)\} = P\{||u(s_\gamma,\cdot)||_p \le R\} \ge P\{||u(s_\gamma,\cdot)||_{1/2-\kappa,2} \le NR\} \ge 1 - NI\gamma^{-2}R^{-2}.$$

Next, obviously, on the set $A(p,\gamma,R)$ and the time interval (s_γ,T), the assumptions of Theorem 8.5 are satisfied with $\kappa = 1/2-2/p \in (0,1/2)$. Therefore

$$\int_{s_\gamma}^T EI_{A(p,\gamma,R)}||u(t,\cdot)||^p_{2/p,p}\,dt \le NEI_{A(p,\gamma,R)}||u(s_\gamma,\cdot)||^p_p \le NR^p,$$

$$EI_{A(p,\gamma,R)}||u(r^1_\gamma,\cdot)||^p_{2/p,p} \le NR^p/\gamma,$$

for an $r^1_\gamma \in (s_\gamma+\gamma, s_\gamma+2\gamma)$ and $\gamma < T/6$.

For $p > 8$, we can represent $2/p$ as $1/2 - \kappa - 2/p$ with $\kappa = 1/2 - 4/p \in (0, 1/2)$. For those p, by Theorem 8.5,

$$\int_{r_\gamma^1}^T EI_{A(p,\gamma,R)}||u(t,\cdot)||^p_{4/p,p}\,dt \le NEI_{A(p,\gamma,R)}||u(r_\gamma^1,\cdot)||^p_{2/p,p} \le NR^p/\gamma,$$

$$EI_{A(p,\gamma,R)}||u(r_\gamma^2,\cdot)||^p_{4/p,p} \le NR^p/\gamma^2,$$

for some $r_\gamma^2 \in (r_\gamma^1 + \gamma, r_\gamma^1 + 2\gamma) \in (0, T)$. One can keep going this way and, for $p > 4n$, where $n = 1, 2, ...$, find r_γ^n that are close to zero if γ is small, such that $I_{A(p,\gamma,R)}u \in \mathcal{H}_p^{2n/p}(r_\gamma^n, T)$. Here p can be taken arbitrary large, r_γ^n small, and $2n/p$ can be made as close to $1/2$ as we wish. Also, the probability of the set $A(p, \gamma, R)$ can be chosen close to 1. Therefore, we obtain our claim as in Remark 8.7.

REMARK 8.9. If $\varepsilon \ge 0$ and $\kappa + 2\varepsilon < 1/2$, then in (8.21) we could take a noise like $(1 - \Delta)^\varepsilon B_t$, which is even "whiter" than B_t. This would only lead to replacing η_k in (8.22) with $(1 - \Delta)^\varepsilon \eta_k$, and, under natural additional smoothness assumptions on h, would give the assertions of Theorem 8.5 with $\kappa + 2\varepsilon$ in place of κ.

8.4. Non–Explosion for a Nonlinear Equation. Take a, b, h satisfying Assumptions 8.5 and 8.6 from Subsec. 8.3 for $\tau = \infty$, $r = 1$, $s = \infty$, and $\xi \equiv 1$ and take a bounded real-valued $\mathcal{P} \times \mathcal{B}(\mathbb{R})$-measurable function $c(t, x) = c(\omega, t, x)$. Assume that $h(t, x, u) = 0$ for $u \le 0$. Fix a number $\lambda \in [0, 1/2)$ and let

$$g^k(t, x, u) = h(t, x, u)u_+^\lambda \eta_k(x),$$

where, as usual, $\{\eta_k\}$ form an orthonormal basis in L_2. Here it is convenient to assume additionally that each η_k is bounded.

The results from Subsec. 8.3 can be easily applied to prove that if the initial condition u_0 is nonnegative and, say, is nonrandom and belongs to C_0^∞, then the equation

$$\begin{aligned} du(t,x) =& [au''(t,x) + b(t,x)u'(t,x) + c(t,x)u(t,x)]\,dt \\ &+ g^k(t,x,u(t,x))\,dw_t^k \end{aligned} \tag{8.31}$$

has a solution defined for all t in the class of functions such that $\sup_{t \le T, x}|u(t, x)|$ is finite (a.s.) for any $T < \infty$. These facts for equation (8.31) considered on a finite space interval with $a \equiv 1$, $b \equiv c \equiv 0$, and $h(u) = u_+$ and with zero boundary data were discovered in [**26**] with the help of a quite different approach. By using the maximum principle, one can show that our assertion implies the result of [**26**].

First, let us explain why (8.31) is solvable despite the high growth of g in u. It turns out that (8.31) possesses a kind of integral or conservation law. Observe that $u \ge 0$, which follows from the maximum principle if one notices that the solution u of (8.31) also satisfies the equation with $g^k(t, x, u(t, x))$ replaced by $\nu^k(t, x)u(t, x)$, where $\nu^k(t, x) = g^k(t, x, u(t, x))u^{-1}(t, x)$ and $|\nu^k| \le N|u|^\lambda$. Moreover, if, for instance, $a = 1, b = c = 0$, then, by integrating (8.31) formally with respect to x, one obtains that $||u(t, \cdot)||_1$ is a local martingale. It is nonnegative, therefore its trajectories are bounded (a.s.).

This takes care of "almost $u^{1/2}$" in the diffusion term. Indeed, one can rewrite (8.31) with $\xi(t, x)u\eta_k(x)$ in place of $g^k(t, x, u)$, where $\xi(t, x) = h(t, x, u(t, x))u^{\lambda-1}(t, x)$ and $|\xi(t, x)| \le |u^\lambda(t, x)|$. By the above, the latter is summable to power $2s = 1/\lambda$, and $s > 1$ (which is required in Theorem 8.5) for $\lambda < 1/2$. Therefore, u satisfies a linear equation with coefficients under control, and we get that $u(t, x)$ is a bounded continuous function on $[0, T] \times \mathbb{R}$ for any $T < \infty$ as in Remark 8.7 (by letting $p \to \infty$). The rigorous treatment below follows this idea.

By Theorem 8.5, for any $m = 1, 2, 3...$, the equation

$$du_m = (au''_m + bu'_m + cu_m)\,dt + g^k(u_m \wedge m)\,dw^k_t \tag{8.32}$$

with the initial condition $u_m(0,\cdot) = u_0$ has a unique solution $u_m \in \mathcal{H}^{1/2-\kappa}_p(T)$ for any $\kappa \in (0, 1/2)$, $p \geq 2$, and $T < \infty$. By Remark 8.7, the function $u_m(t,x)$ is continuous in (t,x) (a.s.). To proceed with the argument, we need the following lemma to be proved later.

LEMMA 8.10. *For any $T < \infty$,*

$$\lim_{R\to\infty} \sup_m P\{\sup_{t\leq T,x} |u_m(t,x)| \geq R\} = 0. \tag{8.33}$$

Define

$$\tau^R_m = \inf\{t \geq 0 : \sup_x |u_m(t,x)| \geq R\}.$$

Observe that if $m \geq R$ (and R is an integer), then the functions u_m and u_R satisfy the same equation on $[0, \tau^R_m)$ and therefore coincide by uniqueness. In particular, $\tau^m_m \geq \tau^R_m = \tau^R_R$ (a.s.). Therefore, there is no ambiguity in the definition

$$u(t,x) = u_m(t,x) \quad \text{on} \quad [0, \tau^m_m].$$

Of course, u satisfies (8.31) on $(0, \lim \tau^m_m)$ and $|u(t,x)| \leq m$ for $t \leq \tau^m_m$. To finish the proof of our assertions about (8.31), it only remains to notice that $\lim \tau^m_m = \infty$ (a.s.), since, by (8.33),

$$P\{\tau^m_m \leq T\} = P\{\sup_{t\leq T,x} |u_m(t,x)| \geq m\} \leq \sup_n P\{\sup_{t\leq T,x} |u_n(t,x)| \geq m\} \to 0$$

as $m \to \infty$.

Proof of Lemma 8.10. Define

$$\xi_m(t,x) = h(t,x,u_m(t,x) \wedge m)(u_m(t,x) \wedge m)^\lambda_+ u^{-1}_m(t,x) \qquad (0 \cdot 0^{-1} := 0)$$

and notice that ξ_m is a bounded function. It follows from (8.32) that u_m is a solution of the equation

$$dv = (av'' + bv' + cv)\,dt + v\xi_m\eta_k\,dw^k_t. \tag{8.34}$$

Obviously, Assumptions 8.5 and 8.6 from Subsec. 8.3 are satisfied for (8.34). By Remark 8.6 and by virtue of our assumption about boundedness of η_k, Theorem 5.12 is valid for (8.34). Thus, $u_m \geq 0$ (a.s.).

Next, take $\zeta_k(x)$ from Theorem 5.7, multiply (8.32) by $\zeta_k e^{-Kt}$, where $K = \sup(|a''| + |b'| + |c|)$, integrate by parts (that is, use the definition of solutions), and take expectations. Then, for any constant T and stopping time $\tau \leq T$, we obtain

$$e^{-KT}E(\zeta_k, u_m(\tau,\cdot)) \leq E(\zeta_k, u_m(\tau,\cdot))e^{-K\tau} = (\zeta_k, u_0)$$

$$+E\int_0^\tau (a\zeta''_k + (2a' - b)\zeta'_k + (a'' - b' + c - K)\zeta_k, u_m)e^{-Kt}\,dt$$

$$\leq N + NE\int_0^\tau (|\zeta''_k| + |\zeta'_k|, u_m)\,dt \leq N + \frac{N}{k}k^{1-1/p}E\int_0^T ||u_m(t,\cdot)||_p\,dt,$$

$$E(\zeta_k, u_m(\tau,\cdot)) \leq N + Nk^{-1/p}(E\int_0^T ||u_m(t,\cdot)||^p_p\,dt)^{1/p} \leq N + Mk^{-1/p},$$

where the last constant N is independent of m, k, and τ and M is independent of k. Since this inequality is true for any stopping time $\tau \leq T$, with the same N and M for any number $\gamma \in (0,1)$ (see, for instance, Theorem III.6.8 of [**19**]),

$$E\sup_{t\leq T}(\int_{\mathbb{R}} \zeta_k(x)u_m(t,x)\,dx)^\gamma \leq 1+\gamma\frac{N+Mk^{-1/p}}{1-\gamma}, \quad E\sup_{t\leq T}||u_m(t,\cdot)||_1^\gamma \leq \frac{N+1}{1-\gamma},$$

where the latter relation is obtained from the former one by the monotone convergence theorem. It follows that

$$P\{\sup_{t\leq T}||u_m(t,\cdot)||_1 > S\} \leq \frac{N}{\sqrt{S}}, \tag{8.35}$$

where N is independent of m and $S>0$.

Now fix $m, S>0$ and define

$$\tau_m(S) = \inf\{t\geq 0 : ||u_m(t,\cdot)||_1 \geq S\}, \quad h_m(t,x,u) = \xi_m(t,x)u.$$

Observe that $\xi_m \leq u_m^\lambda$, which yields that for $t \leq \tau_m(S)$ we have $||\xi_m(t,\cdot)||_{2s} \leq S^{1/2s}$ with $s = 1/(2\lambda)$. Also, u_m satisfies (8.34). By Theorem 8.5, we obtain that, for any $T<\infty$, $||u_m||_{\mathcal{H}_p^{1/2-\kappa}(T\wedge\tau_m(S))}$ is bounded by a constant independent of m, whenever $p \geq 2r = 2/(1-2\lambda)$ and $1/2 > \kappa > \lambda$. The embedding theorems (cf. Remark 8.7) imply that $E\sup_{t\leq T\wedge\tau_m(S),x}|u_m(t,x)|^p$ is bounded independently of m for all large p (and hence for small p as well).

Thus

$$P\{\sup_{t\leq T,x}|u_m(t,x)| \geq R\} \leq P\{\sup_{t\leq T\wedge\tau_m(S),x}|u_m(t,x)| \geq R\} + P\{\tau_m(S) < T\}$$

$$\leq \frac{N}{R} + P\{\sup_{t\leq T}||u_m(t,\cdot)||_1 > S\} \leq \frac{N}{R} + \frac{N}{\sqrt{S}}$$

with the constants N independent of m, S, R. This leads to (8.33), and the lemma is proved.

REMARK 8.11. In [**24**] one can find a proof that $\sup_m E||u_m||^\gamma_{C([0,T]\times\mathbb{R})} < \infty$ for any $T<\infty$, where $0<\gamma<1-2\lambda$ and γ can be chosen arbitrary close to $1-2\lambda$.

9. Open Problems

- In the filtering problem, we only considered the case of equation (8.1) with noise in observations depending only on t and y_t, and thus independent of the signal x_t. Almost nothing is known about filtering problems in which Θ depends on the signal x_t. Very interesting jump processes may appear.
- The above theory applies to equations in the whole space. It is known that even for zero Dirichlet boundary data the L_2–theory is much more complicated because one needs weighted Sobolev spaces (cf. [**16**]) or compatibility conditions (cf. [**2**], [**6**]). The L_2–theory in [**16**] was developed by considering the stochastic term as a perturbation, which gave a perfect result due to obtaining some sharp estimates which turned out to be usable. Nothing of that kind can happen for general $p>2$, that is why very often, say in [**6**], it is assumed that the stochastic term is sufficiently small and some compatibility conditions hold. A first encouraging step toward developing L_p–theory in weighted Sobolev spaces in bounded domains is made in [**21**] (see also [**23**]).

  By the way, the fact that the derivatives of solutions blow up near the boundary shows that one needs to be careful constructing finite–difference approximations near the boundary.

- The above theory is an L_p–theory. One may ask whether a $C^{2+\alpha}$–theory can be constructed. By a theory we mean not only results that, for f, g^k belonging to a space F, the solution belongs to some kind of stochastic $C^{2+\alpha}$–spaces, but also that *every* element of this stochastic space can be obtained as a solution for certain f, g^k belonging to the same F.

  We only have Theorems 5.1 and 7.2 and Sobolev's embedding $H^n_p \subset \mathcal{C}^{n-d/p}$ if $n - d/p > 0$. Thus, if f, g^k are relatively smooth, the solution belongs to $C^{2+\alpha}$. This is a useful result; for instance, it essentially covers the main result of [**3**] obtained in the whole space for coefficients independent of x. The weakness of such results is that they do not allow natural perturbations and cannot be extended reasonably far. For example, if one knows such results only for equations with constant coefficients, one will not be able to extend them to equations with variable highest order coefficients.

  However, a "right" $C^{2+\alpha}$ theory in the above sense is developed in [**27**], for equations with nonrandom coefficients and with no derivatives of the unknown function in the stochastic term.
- Fully nonlinear equations and uniqueness. In Assumption 5.6 we assume that nonlinear terms in equation (5.1) are strictly subordinated to the main terms. Nothing is known if, instead of linear main operators in (5.1), one has nonlinear ones. Even for the case of linear equations we do not know how strong this subordination has to be. For instance, take the following version of one–dimensional equation (8.22)

$$du(t,x) = u''(t,x)\,dt + h(x)u(t,x)\eta^k(x)\,dw^k_t,$$

  where we only know that $h \in L_2$. Is it true that if $\sup_{t\le T,x}|u(t,x)| < \infty$ (a.s.) and $u(0,\cdot) = 0$, then $u = 0$ on $[0,T]$? The point is that in Assumption 8.6 we have to take $s = 1$, and then $r = \infty$, $p = \infty$ and even $\kappa = 1/2$ is a little bit short of satisfying the last inequality in (8.23) which is needed to provide the right subordination. The author believes that the answer to this particular question is positive, but the general question still remains.
- Superdiffusions. One can derive SPDEs for superdiffusions in multi–dimensional case (see [**22**]). The question arises whether one can get from these equations at least some part of huge and delicate information about superprocesses, available in the literature.

Bibliography

[1] J.M. Bismut, M. Yor: An inequality for processes which satisfy Kolmogorov's continuity criterion. Application to continuous martingales, *J. Funct. Anal.* **51** (1983), 166–173.

[2] Z. Brzezniak: Stochastic partial differential equations in Banach spaces, II, (preprint).

[3] P.-L. Chow, J.-L. Jiang: Stochastic partial differential equations in Hölder spaces, *Probab. Theory Related Fields,* **99** (1994), 1–27.

[4] G. Da Prato: Some results on linear stochastic evolution equations in Hilbert spaces by the semi-group method, *Stochastic Anal. Appl.* **1** (1983), 57–88.

[5] G. Da Prato and J. Zabczyk: *Stochastic equations in infinite dimensions*, Cambridge University Press, New York, NY, 1992.

[6] F. Flandoli: Dirichlet boundary value problem for stochastic parabolic equations : compatibility relations and regularity of solutions. *Stochastics and Stochastics Reports* **29** (1990), 331–357.

[7] T. Funaki: Random motion of strings and related evolution equations, *Nagoya Math. Journal* **89** (1983) 129–193.

[8] I. Gyöngy, D. Nualart and M. Sanz–Solé: Approximation and support theorems in modulus spaces, *Probab. Theory and Related Fields* **101** (1995) 495–509.

[9] A.M. Garsia, E. Rodemich, and H. Rumsey, Jr.: A real variable lemma and the continuity of paths of some Gaussian processes, *Indiana Univ. Math. J.* **20** (1970) 565–578.

[10] I. Gyöngy and N.V. Krylov: On stochastic equations with respect to semi-martingales I, *Stochastics,* **4** (1980) 1–21.

[11] I.A. Ibragimov: Properties of sample functions for stochastic processes and embedding theorems, *Theory Probab. and its Appl.* **18** (1973) 468–480 in Russian; English translation pp. 442–453.

[12] K. Itô: On stochastic differential equations, *Mem. Amer. Math. Soc.* **4** (1951).

[13] C.E. Kenig, G. Ponce and L. Vega: Well-posedness and scattering results for the generalized Korteweg–de Vries equation via the contraction principle, *Comm. Pure Appl. Math.* **46** (1993) 527–620.

[14] M.A. Krasnoselskii, E.I. Pustylnik, P.E. Sobolevski and P.P. Zabrejko: *Integral operators in spaces of summable functions*, Nauka, Moscow, 1966 in Russian; English translation: Noordhoff International Publishing, Leyden, 1976.

[15] N.V. Krylov: On explicit formulas for solutions of evolutionary SPDE's (a kind of introduction to the theory), *Lect. Notes in Control and Inform. Sci.* **176** (1992) 153–164.

[16] N.V. Krylov: A W_2^n-theory of the Dirichlet problem for SPDE in general smooth domains, *Probab. Theory Relat. Fields,* **98** (1994) 389–421.

[17] N.V. Krylov: A generalization of the Littlewood-Paley inequality and some other results related to stochastic partial differential equations, *Ulam Quarterly,* **2** (1994) 16–26, http://www.ulam.usm.edu/VIEW2.4/krylov.ps

[18] N.V. Krylov: A parabolic Littlewood-Paley inequality with applications to parabolic equations, *Topological Methods in Nonlinear Analysis, Journal of the Juliusz Schauder Center,* **4** (1994) 355–364.

[19] N.V. Krylov: *Introduction to the theory of diffusion processes* Amer. Math. Soc., Providence, RI, 1995.

[20] N.V. Krylov: On L_p-theory of stochastic partial differential equations, *SIAM J. Math. Anal.* **27** (1996) 313–340.

[21] N.V. Krylov: One–dimensional SPDEs with constant coefficients in the positive half axis, Submitted to a volume dedicated to R. Liptser's 60th birthday.

[22] N.V. Krylov: On SPDEs and superdiffusions, Annals of Probability, Vol. 25, No. 4 (1997), 1789-1809.

[23] N.V. Krylov, S.V. Lototsky: A Sobolev space theory of SPDEs with constant coefficients on a half line, submitted to SIAM Journal on Mathematical Analysis.

[24] N. V. Krylov, B. L. Rozovskii: On conditional distributions of diffusion processes, *Izvestija Akademii Nauk SSSR, serija matematicheskaja,* **42** (1978), 356–378; English translation: Math. USSR Izvestija, Vol. 12, No. 2 (1978), 336–356.

[25] N. V. Krylov, B. L. Rozovskii: On the characteristics of degenerate second order parabolic Itô equations, *Trudy seminara imeni Petrovskogo,* **8** (1982) 153–168 in Russian; English translation: J. Soviet Math., Vol. 32 (1986), 336–348.

[26] C. Mueller: Long time existence for the heat equation with a noise term, *Probab. Theory Related Fields,* **90** (1991) 505–517.

[27] R. Mikulevicius, B.L. Rozovskii: *Introduction to SPDEs* (forthcoming book).

[28] D. Nualart and E. Pardoux: White noise driven quasilinear SPDEs with reflection. *Probab. Theory Related Fields,* **93** (1992) 77–89.

[29] R.E.A.C. Paley, N. Wiener and A. Zygmund, Note on random functions, *Math. Z.* **37** (1933) 647–668.

[30] B. L. Rozovskii: *Stochastic evolution systems*, Kluwer, Dordrecht, 1990.

[31] O. A. Ladyzhenskaia, V. A. Solonnikov and N. N. Ural'tceva: *Linear and quasi-linear equations of parabolic type*, Nauka, Moscow, 1967 in Russian; English translation: American Math. Soc., Providence, 1968.

[32] S. Lapic: On the first initial-boundary problem for SPDEs on domains with limited smoothness at the boundary, submitted to Potential Anal.

[33] H. Triebel: *Theory of function spaces*, Birkhäuser Verlag, Basel–Boston–Stuttgart, 1983.

[34] H. Triebel: *Theory of function spaces II*, Birkhäuser Verlag, Basel–Boston–Berlin, 1992.

[35] J. B. Walsh: An introduction to stochastic partial differential equations, *Lecture Notes in Math.* **1180** (1986) 266–439.

Index