Available online at www.sciencedirect.com

ScienceDirect

Journal of Differential Equations 342 (2023) 282–324

Journal of Differential Equations

www.elsevier.com/locate/jde

Singular perturbations and asymptotic expansions for SPDEs with an application to term structure models

Sergio Albeverio [a], Carlo Marinelli [b,*], Elisa Mastrogiacomo [c]

[a] *Institut für Angewandte Mathematik, Universität Bonn, Endenicher Allee 60, 53115 Bonn, Germany*
[b] *Department of Mathematics, University College London, Gower Street, London WC1E 6BT, UK*
[c] *Dipartimento di Economia, Università degli Studi dell'Insubria, via Montegeneroso, 71, 21100 Varese, Italy*

Received 17 November 2021; revised 7 September 2022; accepted 30 September 2022
Available online 12 October 2022

Abstract

We study the dependence of mild solutions to linear stochastic evolution equations on Hilbert space driven by Wiener noise, with drift having linear part of the type $A + \varepsilon G$, on the parameter ε. In particular, we study the limit and the asymptotic expansions in powers of ε of these solutions, as well as of functionals thereof, as $\varepsilon \to 0$, with good control on the remainder. These convergence and series expansion results are then applied to a parabolic perturbation of the Musiela SPDE of mathematical finance modeling the dynamics of forward rates.

Keywords: Singular perturbations; Asymptotic expansions; Stochastic PDE; Interest rate models

1. Introduction

Consider the family of stochastic evolution equations

$$du_\varepsilon = (A + \varepsilon G)u_\varepsilon \, dt + \alpha \, dt + B \, dW, \qquad u_\varepsilon(0) = u_0, \tag{1}$$

set in a Hilbert space H and indexed by $\varepsilon > 0$, where A and G are linear maximal dissipative operators on H such that $A + \varepsilon G$ is also maximal dissipative, α and B are coefficients satisfying

* Corresponding author.
E-mail address: c.marinelli@ucl.ac.uk (C. Marinelli).

https://doi.org/10.1016/j.jde.2022.09.037

suitable measurability, integrability and regularity conditions, and W is a cylindrical Wiener process. Precise assumptions on the data of the problem are given below.

Our main goal is to obtain an expansion of the difference $u_\varepsilon - u$ as a polynomial in ε plus a remainder term, where u_ε and u are the unique mild solutions to (1) with $\varepsilon > 0$ and $\varepsilon = 0$, respectively. Results in this sense are obtained assuming that the semigroups generated by A and G commute. As a first step, we show that u_ε converges to u as $\varepsilon \to 0$, also in the case where α and B are (random, time-dependent) Lipschitz continuous functions of the unknown, in suitable norms implying the convergence in probability uniformly on compact intervals in time. For such convergence result to hold it is enough that the resolvent of $A + \varepsilon G$ converges to the resolvent of A as $\varepsilon \to 0$ in the strong operator topology, without any commutativity assumption. Sufficient conditions for the convergence of operators in the strong resolvent sense have been largely studied (see, e.g., [9] and references therein) and can be readily applied to obtain convergence results for solutions to stochastic evolution equations. On the other hand, expansions in power series of $u_\varepsilon - u$ are considerably harder to obtain. In fact, it is well known that solutions to singularly perturbed equations, also in the simpler setting of deterministic ODEs, do *not* admit series expansions in the perturbation parameter. This phenomenon appears also in the class of stochastic equations studied here, as it is quite obvious. This is essentially the reason behind the commutativity assumption on the semigroups generated by A and G, as well as on the regularity conditions on the initial datum u_0 and on the coefficients α and B (see §4 below, where asymptotic expansion results are obtained also for functionals of u_ε).

As an application of the abstract results, we consider a singularly perturbed transport equation on $\mathbb{R}$ where, roughly speaking, A and G are the first and second derivative, respectively. This equation can be seen as a singular perturbation of an extension of Musiela's SPDE from a weighted Sobolev space on $\mathbb{R}_+$ to the corresponding one on $\mathbb{R}$. The motivation for considering this problem comes from the interesting article [7], where the author argues that second-order parabolic SPDEs reproduce many stylized empirical properties of forward curves. On the other hand, if forward rates satisfy a Heath-Jarrow-Morton dynamics, the differential operator in the drift of the corresponding SPDE must be of first order. It is then natural to consider singular perturbations of the (first-order) Musiela SPDE by second-order differential operators and to look for conditions implying uniform convergence of the "perturbed" forward rates, as well as of implied bond prices, to the corresponding "unperturbed" forward rates and bond prices, as well as a more precise description of the dependence of the pricing error on the "size" of the perturbation. Results in this regard are obtained in the form of asymptotic expansions in ε of the solution u_ε to a second-order perturbation of a suitable extension of the Musiela SPDE, as well as of functionals thereof.

The rest of the text is organized as follows. In §2 we introduce notation, we recall basic results from semigroup theory, and we establish some inequalities and identities for classes of stochastic convolutions. In §3 we show that a commutativity assumption between the semigroups generated by A and G implies that the closure of $A + \varepsilon G$ converges to A in the strong resolvent sense as $\varepsilon \to 0$. This allows, thanks to a general convergence result for mild solutions to stochastic evolution equations, to deduce the convergence of u_ε to u in a suitable norm. Under further regularity assumptions on u_0, α and B, expansions of the difference $u_\varepsilon - u$ and of functionals thereof as power series in ε are obtained in §4, which is the core of the work. Finally, the applications described above to Musiela's SPDE are developed in §5.

Acknowledgments. The first-named author would like to thank the Economics Department of the Università dell'Insubria, Varese and the Università degli Studi di Milano for warm hospitality.

Large part of the work for this paper has been carried out while the second-named author was visiting the Interdisziplinäres Zentrum für Komplexe Systeme at the University of Bonn. The third-named author would like to thank the Institute of Applied Mathematics and the Hausdorff Center for Mathematics at the University of Bonn for warm hospitality.

2. Preliminaries

Throughout this section we shall use E and F to denote two Banach spaces. The expression $E \hookrightarrow F$ means that E is continuously embedded in F. The domain of a linear operator L with graph in $E \times F$ will be denoted by $\mathsf{D}(L)$. The Banach space of continuous k-linear operators from E^k to F, $k \in \mathbb{N}$, is denoted by $\mathscr{L}_k(E; F)$ (without subscript, as usual, if $k = 1$). Given $h \in E$ and $k \in \mathbb{N}$, we shall set $h^{\otimes k} = (h, \dots, h) \in E^k$. If E and F are Hilbert spaces, $\mathscr{L}^2(E; F)$ will stand for the Hilbert space of Hilbert-Schmidt operators from E to F. An expression of the type $a \lesssim b$ means that there exists a positive constant N such that $a \leq Nb$, and $a \eqsim b$ stands for $a \lesssim b$ and $b \lesssim a$.

We recall the following form of Taylor's formula (see, e.g., [16, p. 349]). Let $U \subseteq E$ be open, $f \in C^m(U; F)$, $x \in U$ and $h \in E$ such that the segment $[x, x+h]$ is contained in U. Then

$$f(x+h) = \sum_{k=0}^{m-1} \frac{1}{k!} D^k f(x) h^{\otimes k} + \int_0^1 \frac{(1-t)^{m-1}}{(m-1)!} D^m f(x+th) h^{\otimes m}\, dt.$$

For the purposes of this section only, we denote a strongly continuous semigroup on a Hilbert space H and its generator by S and A, respectively. As is well known, there exist $M \geq 1$ and $w \in \mathbb{R}$ such that $\|S(t)\| \leq Me^{wt}$ for all $t \geq 0$. Let $m \geq 1$ be an integer. If $\phi \in \mathsf{D}(A^m)$, one has the Taylor-like formula

$$S(t)\phi = \sum_{k=0}^{m-1} \frac{t^k}{k!} A^k \phi + \frac{1}{(m-1)!} \int_0^t (t-u)^{m-1} S(u) A^m \phi\, du$$

(see, e.g., [5, Proposition 1.1.6]). We recall that A^m is a closed operator and that $\mathsf{D}(A^m)$ is a Hilbert space with scalar product

$$\langle \phi, \psi \rangle_{\mathsf{D}(A^m)} = \langle \phi, \psi \rangle + \langle A\phi, A\psi \rangle + \cdots + \langle A^m \phi, A^m \psi \rangle.$$

Let T be a further strongly continuous semigroup on H. We shall say that S and T commute if $S(t)T(t) = T(t)S(t)$ for all $t \in \mathbb{R}_+$. It is immediate that the product semigroup ST is strongly continuous. It also follows that $S(s)T(t) = T(t)S(s)$ for all $t, s \geq 0$: first one proves it for rational s and t, hence the general case follows by density and continuity. For details see, e.g., [9, p. 44]. Moreover T leaves invariant $\mathsf{D}(A)$: in fact, for any $f \in \mathsf{D}(A)$, one has

$$\lim_{h \to 0} \frac{S(h)T(t)f - T(t)f}{h} = T(t)\Big(\lim_{h \to 0} \frac{S(h)f - f}{h}\Big) = T(t)Af.$$

This also implies, by uniqueness of the limit, that $T(t)Af = AT(t)f$. These observations in turn imply that the resolvent R_λ of the generator of T commutes with A, in the sense that, if $f \in \mathsf{D}(A)$, then $R_\lambda f \in \mathsf{D}(A)$ and $R_\lambda Af = AR_\lambda f$ (cf. [13, p. 171]).

All stochastic elements will be defined on a fixed probability space $(\Omega, \mathscr{F}, \mathbb{P})$, endowed with a filtration $(\mathscr{F}_t)_{t\in[0,T]}$, with T a fixed positive number, that is assumed to satisfy the so-called usual assumptions. We shall denote by W a cylindrical Wiener process on a real separable Hilbert space U. We shall denote the closed subspace of $L^p(\Omega; C([0,T]; H))$, $p > 0$, of H-valued adapted continuous processes by C^p, which is hence a quasi-Banach space itself (with the induced quasi-norm). Given a progressively measurable process $C \in L^0(\Omega; L^2(0,T; \mathscr{L}^2(U; H)))$, the stochastic convolution $S \diamond C$ is the H-valued process defined by

$$S \diamond C(t) := \int_0^t S(t-s)C(s)\,dW(s) \qquad \forall t \in [0,T].$$

Similarly, if $f \in L^0(\Omega; L^1(0,T; H))$, we shall define the H-valued process $S * f$ by

$$S * f(t) := \int_0^t S(t-s)f(s)\,ds \qquad \forall t \in [0,T].$$

The stochastic integral of a process F with respect to W will be occasionally denoted by $F \cdot W$ for typographical convenience. Moreover, we recall that, for any $p \in]0,\infty[$ and progressively measurable $\mathscr{L}^2(U; H)$-valued process F, the Burkholder-Davis-Gundy inequality

$$\big\|F \cdot W\big\|_{L^p(\Omega;H)} \leq N_p \big\|F\big\|_{L^p(\Omega;L^2(0,T;\mathscr{L}^2(U;H)))}$$

holds, where N_p is a constant depending on p only (see, e.g., [20]).

Lemma 2.1. *Let $p > 0$, $C \in L^p(\Omega; L^2(0,T; \mathscr{L}^2(U; H)))$ be a progressively measurable process, and $n \geq 0$. One has*

$$\mathbb{E}\bigg\|\int_0^t (t-s)^n S(t-s)C(s)\,dW(s)\bigg\|^p \leq N_p^p M^p\, \mathbb{E}\bigg(\int_0^t (t-s)^{2n} e^{2w(t-s)} \big\|C(s)\big\|^2_{\mathscr{L}^2(U;H)}\,ds\bigg)^{p/2}$$

for every $t \in [0,T]$.

Proof. For any $\delta > 0$, one has

$$\bigg\|\int_0^t (t-s)^n S(t-s)C(s)\,dW(s)\bigg\|_{L^p(\Omega;H)} \leq \sup_{t_0\in[t,t+\delta]}\bigg\|\int_0^t (t_0-s)^n S(t_0-s)C(s)\,dW(s)\bigg\|_{L^p(\Omega;H)}.$$

Since $\big((t_0-\cdot)S(t_0-\cdot)B\big)\cdot W$ is a local martingale, the Burkholder-Davis-Gundy inequality and the ideal property of Hilbert-Schmidt operators yield

$$\sup_{t_0\in[t,t+\delta]}\bigg\|\int_0^t (t_0-s)^n S(t_0-s)C(s)\,dW(s)\bigg\|_{L^p(\Omega;H)}$$

$$\leq N_p M \sup_{t_0 \in [t,t+\delta]} \Big\| (t_0 - \cdot)^n e^{w(t_0-\cdot)} \| C \|_{\mathscr{L}^2(U;H)} \Big\|_{L^p(\Omega; L^2(0,t;H))}.$$

Setting

$$\phi_\delta(t) = \begin{cases} e^{w(t+\delta)}, & \text{if } w \geq 0, \\ e^{wt}, & \text{if } w < 0, \end{cases}$$

one has $e^{w(t_0-s)} \leq \phi_\delta(t-s)$ for all $t_0 \in [t, t+\delta]$ and $s \in [0,t]$, hence

$$\sup_{t_0 \in [t,t+\delta]} \Big\| \int_0^t (t_0 - s)^n S(t_0 - s) C(s)\, dW(s) \Big\|_{L^p(\Omega;H)} \\ \leq N_p M \Big\| (t + \delta - \cdot)^n \phi_\delta(t - \cdot) \| C \|_{\mathscr{L}^2(U;H)} \Big\|_{L^p(\Omega; L^2(0,t;H))},$$

therefore

$$\mathbb{E} \Big\| \int_0^t (t-s)^n S(t-s) C(s)\, dW(s) \Big\|^p \leq N_p^p M^p\, \mathbb{E} \Big(\int_0^t (t+\delta-s)^{2n} \phi_\delta^2(t-s) \| C(s) \|^2_{\mathscr{L}^2(U;H)}\, ds \Big)^{p/2}$$

for all $\delta > 0$. Taking the limit as $\delta \to 0$ proves the claim. □

The following recursive relation for certain stochastic convolutions will be very useful in the sequel.

Lemma 2.2. *Let $C \in L^0(\Omega; L^2(0,T; \mathscr{L}^2(U;H)))$ be a progressively measurable process and define, for every $k \in \mathbb{R}_+$,*

$$\Sigma_k(t) := \int_0^t S(t-s)(t-s)^k C(s)\, dW(s).$$

Then $\Sigma_{k+1} = (k+1) S * \Sigma_k$.

Proof. Let $k \geq 1$. Using the identity

$$(t-s)^k = k \int_s^t (r-s)^{k-1}\, dr,$$

the stochastic Fubini theorem, and the semigroup property, one has, for every $t \in [0,T]$,

$$
\begin{aligned}
\Sigma_k(t) &= \int_0^t S(t-s)(t-s)^k C(s)\,dW(s)\\
&= k\int_0^t S(t-s)\biggl(\int_s^t (r-s)^{k-1}\,dr\biggr)C(s)\,dW(s)\\
&= k\int_0^t\int_0^r S(t-s)(r-s)^{k-1}C(s)\,dW(s)\,dr\\
&= k\int_0^t S(t-r)\int_0^r S(r-s)(r-s)^{k-1}C(s)\,dW(s)\,dr\\
&= kS*\Sigma_{k-1}(t). \quad \square
\end{aligned}
$$

3. Singular perturbations by commuting semigroups

Let us consider the stochastic evolution equation on the Hilbert space H

$$du = Au\,dt + \alpha(u)\,dt + B(u)\,dW, \qquad u(0) = u_0,$$

and the family of stochastic evolution equations on H indexed by a parameter $\varepsilon \geq 0$

$$du_\varepsilon = (A+\varepsilon G)u_\varepsilon\,dt + \alpha(u_\varepsilon)\,dt + B(u_\varepsilon)\,dW, \qquad u_\varepsilon(0) = u_0,$$

where (i) A and G are linear maximal dissipative operators on H such that the closure of $A+\varepsilon G$, denoted by the same symbol, is maximal dissipative as well; (ii) the initial datum u_0 belongs to $L^0(\Omega, \mathscr{F}_0; H)$; (iii) the coefficients

$$\alpha\colon \Omega\times[0,T]\times H \longrightarrow H, \qquad B\colon \Omega\times[0,T]\times H \longrightarrow \mathscr{L}^2(U;H)$$

are Lipschitz continuous in the third variable, uniformly with respect to the other ones, and such that $\alpha(\cdot,\cdot,h)$ and $B(\cdot,\cdot,h)$ are progressively measurable for every $h \in H$. It is well known that under these conditions the above stochastic equations admit unique mild solutions u and u_ε, respectively, with continuous trajectories. Moreover, if $u_0 \in L^p(\Omega, \mathscr{F}_0; H)$ for some $p > 0$, then u and u_ε belong to C^p.

The aim of this section is to provide sufficient conditions ensuring that $u_\varepsilon \to u$ in C^p. We rely on the following convergence result, which is a minor modification of [22, Theorem 2.4] (see also [14]).

Theorem 3.1. *Let $p \in]0,\infty[$, $u_0 \in L^p(\Omega, \mathscr{F}_0; H)$. Assume that $A+\varepsilon G$ converges to A in the strong resolvent sense. Then $u_\varepsilon \to u$ in C^p as $\varepsilon \to 0$.*

We recall that a sequence of maximal dissipative operators (L_n) is said to converge to a maximal dissipative operator L in the strong resolvent sense if $(\lambda - L_n)^{-1}x \to (\lambda - L)^{-1}x$ for all $x \in H$ and all $\lambda > 0$.

The problem of the convergence of u_ε to u is thus reduced to finding sufficient conditions for the convergence of $A + \varepsilon G$ to A in the strong resolvent sense as $\varepsilon \to 0$. In view of the results on asymptotic expansions in the next sections, we limit ourselves to the special case where the semigroups generated by A and G, denoted respectively by S_A and S_G, commute.

Lemma 3.2. *Assume that S_A and S_G commute, i.e. that $S_A(t)S_G(t) = S_G(t)S_A(t)$ for all $t \geq 0$. Then $A + \varepsilon G$ converges to A in the strong resolvent sense as $\varepsilon \to 0$.*

Proof. One has, for any $\lambda > 0$ and $f \in H$,

$$\big(\lambda - (A + \varepsilon G)\big)^{-1} f = \int_0^\infty e^{-\lambda t} S_{A+\varepsilon G}(t) f \, dt = \int_0^\infty e^{-\lambda t} S_A(t) S_{\varepsilon G}(t) f \, dt$$

and $S_{\varepsilon G}(t) f \to f$ as $\varepsilon \to 0$, hence, by dominated convergence,

$$\lim_{\varepsilon \to 0} \big(\lambda - (A + \varepsilon G)\big)^{-1} f = \int_0^\infty e^{-\lambda t} S_A(t) f \, dt = (\lambda - A)^{-1} f. \quad \square$$

Remarks 3.3. (i) Under the assumption that S_A and S_G commute, $\mathsf{D}(A) \cap \mathsf{D}(G)$ is a core for the generator of the product semigroup $S_A S_{\varepsilon G}$, which is contractive and strongly continuous. Its generator is hence equal to the closure of $A + \varepsilon G$. So the hypothesis of maximal dissipativity of (the closure of) $A + \varepsilon G$ is automatically satisfied here.
(ii) It is clear from the proof of the previous lemma that not even the assumption of dissipativity of A and G is needed, but just that the resolvent sets of A and G have non-empty intersection. In particular, the statement of the lemma continues to hold if A and G are maximal quasi-dissipative, i.e. if there exist a and $b \in \mathbb{R}_+$ such that $A - aI$ and $G - bI$ are maximal dissipative. In this respect, as long as one is concerned with applications to the stochastic equation, there is no loss of generality assuming that A and G are dissipative rather than quasi-dissipative, because the latter case reduces to the former by adding a linear term to the drift α.

4. Asymptotic expansion of u_ε

Our next goal is to obtain an expression of the difference $u_\varepsilon - u$ as a polynomial in ε plus a remainder. Once such an expression is obtained, the main issue is to prove estimates on the coefficients of the polynomial and on the remainder. Such estimates will crucially depend on suitable regularity assumptions on the coefficients α and B that will be assumed throughout the section to be random and time-dependent, but not explicitly dependent on u. In particular, let us consider the stochastic evolution equations

$$du = Au \, dt + \alpha \, dt + B \, dW, \qquad u(0) = u_0, \tag{2}$$

and

$$du_\varepsilon = Au_\varepsilon \, dt + \varepsilon G u_\varepsilon \, dt + \alpha \, dt + B \, dW, \qquad u_\varepsilon(0) = u_0, \tag{3}$$

where A and G are maximal dissipative and generate commuting semigroups. As before, we denote the closure of $A + \varepsilon G$, $\varepsilon > 0$, by the same symbol. Moreover, we assume that there exist $p \in [1, \infty[$ and an integer $m \geq 1$ such that

$$\begin{gathered} u_0 \in L^p(\Omega; \mathsf{D}(G^m)), \qquad \alpha \in L^p(\Omega; L^1(0, T; \mathsf{D}(G^m))), \\ B \in L^p(\Omega; L^2(0, T; \mathscr{L}^2(U; \mathsf{D}(G^m)))). \end{gathered} \tag{4}$$

Then equations (2) and (3) admit unique mild solutions u and u_ε in C^p, respectively.[1] Just for convenience, we also assume that $\varepsilon \in [0, 1]$.

All results in this section do not use the assumption that A and G are maximal dissipative, except in an indirect way in Proposition 4.11, namely through Theorem 3.1. In particular, all results except Proposition 4.11 continue to hold under the same assumptions on u_0, α and B, commutativity of S_A and S_G, and the existence of unique solutions u and $u_\varepsilon \in \mathsf{C}^p$ to (2) and (3), respectively.

We begin with a decomposition of u_ε that is essentially of algebraic nature. Namely, recalling assumption (4), let us introduce the adapted H-valued processes $v_1, \dots, v_{m-1}$ defined as

$$\begin{aligned} v_k(t) := t^k S_A(t) G^k u_0 &+ \int_0^t (t-s)^k S_A(t-s) G^k \alpha(s)\, ds \\ &+ \int_0^t (t-s)^k S_A(t-s) G^k B(s)\, dW(s) \end{aligned}$$

for each $k \in \{1, \dots, m-1\}$, and the family of adapted H-valued processes $(R_{m,\varepsilon})_{\varepsilon \in]0,1]}$ defined as

$$\begin{aligned} R_{m,\varepsilon}(t) := \frac{\varepsilon^m}{(m-1)!} S_A(t) \int_0^t (t-r)^{m-1} S_{\varepsilon G}(r) G^m u_0\, dr \\ + \frac{\varepsilon^m}{(m-1)!} \int_0^t S_A(t-s) \int_0^{t-s} (t-s-r)^{m-1} S_{\varepsilon G}(r) G^m \alpha(s)\, dr\, ds \\ + \frac{\varepsilon^m}{(m-1)!} \int_0^t S_A(t-s) \int_0^{t-s} (t-s-r)^{m-1} S_{\varepsilon G}(r) G^m B(s)\, dr\, dW(s). \end{aligned}$$

Then one has the following Taylor-like expansion of u_ε around u, with $R_{m,\varepsilon}$ playing the role of a remainder term.

[1] Since α and B are not functions of the unknown, the only non-trivial issue is the pathwise continuity, which follows by the contractivity of the semigroups generated by A and (the closure of) $A + \varepsilon G$ (see, e.g., [8, §6.2]).

Proposition 4.1. *One has*

$$u_\varepsilon = u + \sum_{k=1}^{m-1} \frac{\varepsilon^k}{k!} v_k + R_{m,\varepsilon} \qquad \forall \varepsilon \in]0,1].$$

Moreover, the processes $v_1, \dots, v_{m-1}$ and $R_{m,\varepsilon}$ belong to C^p.

For the proof we will need several lemmas, the main point of which is to obtain suitable estimates of the remainder term $R_{m,\varepsilon}$. More precisely, we are going to estimate $R_{m,\varepsilon}$ in $L^p(\Omega; H)$ pointwise with respect to the time variable as well as in C^p. As mentioned above, such estimates do not use the dissipativity of A and G. For this reason, we shall prove them under the sole assumption that A and G are generators of strongly continuous semigroups S_A and S_G, respectively, with $\|S_A(t)\| \leq M_A e^{w_A t}$ and $\|S_G(t)\| \leq M_G e^{w_G t}$ for all $t \in \mathbb{R}_+$, where $M_A, M_G \geq 1$ and $w_A, w_G \in \mathbb{R}$.

Let us set

$$\begin{aligned}
R^1_{m,\varepsilon}(t) &:= S_A(t) \int_0^t (t-r)^{m-1} S_{\varepsilon G}(r) G^m u_0 \, dr, \\
R^2_{m,\varepsilon}(t) &:= \int_0^t S_A(t-s) \int_0^{t-s} (t-s-r)^{m-1} S_{\varepsilon G}(r) G^m \alpha(s) \, dr \, ds, \\
R^3_{m,\varepsilon}(t) &:= \int_0^t S_A(t-s) \int_0^{t-s} (t-s-r)^{m-1} S_{\varepsilon G}(r) G^m B(s) \, dr \, dW(s),
\end{aligned}$$

so that

$$R_{m,\varepsilon} = \frac{\varepsilon^m}{(m-1)!} \big(R^1_{m,\varepsilon} + R^2_{m,\varepsilon} + R^3_{m,\varepsilon} \big), \tag{5}$$

and introduce the function $f_\varepsilon \colon \mathbb{R}_+ \to \mathbb{R}_+$ defined as

$$f_\varepsilon(t) := e^{w_A t} \int_0^t (t-r)^{m-1} e^{\varepsilon w_G r} \, dr. \tag{6}$$

Lemma 4.2. *One has, for every $t \in [0,T]$ and $\varepsilon \in [0,1]$,*

$$\|R^1_{m,\varepsilon}(t)\| \leq M_A M_G f_\varepsilon(t) \big\|u_0\big\|_{\mathsf{D}(G^m)}$$

and

$$\|R^2_{m,\varepsilon}(t)\| \leq M_A M_G \int_0^t f_\varepsilon(t-s) \big\|\alpha(s)\big\|_{\mathsf{D}(G^m)} \, ds.$$

Proof. Both estimates are immediate consequences of Minkowski's inequality and the definition of f. For instance, the second one is given by

$$\begin{aligned}\|R^2_{m,\varepsilon}(t)\| &\leq M_A M_G \int_0^t e^{w_A(t-s)} \big\|\alpha(s)\big\|_{\mathsf{D}(G^m)} \int_0^{t-s} (t-s-r)^{m-1} e^{\varepsilon w_G r}\,dr\,ds \\ &= M_A M_G \int_0^t f_\varepsilon(t-s) \big\|\alpha(s)\big\|_{\mathsf{D}(G^m)}\,ds. \quad \square\end{aligned}$$

The running maximum of the function f_ε defined in (6) will be denoted by f^*_ε, i.e. $f^*_\varepsilon(t) := \max_{s\in[0,t]} f_\varepsilon(s)$.

Lemma 4.3. *One has, for every $t \in [0, T]$ and $\varepsilon \in [0, 1]$,*

$$\big\|R^1_{m,\varepsilon}(t)\big\|_{L^p(\Omega;H)} \leq M_A M_G f_\varepsilon(t) \big\|u_0\big\|_{L^p(\Omega;\mathsf{D}(G^m))}$$

and

$$\big\|R^2_{m,\varepsilon}(t)\big\|_{L^p(\Omega;H)} \leq M_A M_G f^*_\varepsilon(t) \big\|\alpha\big\|_{L^p(\Omega;L^1(0,t;\mathsf{D}(G^m)))}.$$

Proof. The first estimate is evident. The second one follows by

$$\begin{aligned}\int_0^t f_\varepsilon(t-s) \big\|\alpha(s)\big\|_{\mathsf{D}(G^m)}\,ds &\leq \big\|f_\varepsilon(t-\cdot)\big\|_{L^\infty(0,t)} \big\|\alpha\big\|_{L^1(0,t;\mathsf{D}(G^m))} \\ &= \big\|f_\varepsilon\big\|_{L^\infty(0,t)} \big\|\alpha\big\|_{L^1(0,t;\mathsf{D}(G^m))}. \quad \square\end{aligned}$$

Lemma 4.4. *One has, for every $\varepsilon \in [0, 1]$,*

$$\big\|R^1_{m,\varepsilon}\big\|_{\mathsf{C}^p} \leq M_A M_G f^*_\varepsilon(T) \big\|u_0\big\|_{L^p(\Omega;\mathsf{D}(G^m))}$$

and

$$\big\|R^2_{m,\varepsilon}\big\|_{\mathsf{C}^p} \leq M_A M_G f^*_\varepsilon(T) \big\|\alpha\big\|_{L^p(\Omega;L^1(0,T;\mathsf{D}(G^m)))}.$$

Proof. The first estimate is again evident, thanks to Lemma 4.2. The second one follows by

$$\int_0^t f_\varepsilon(t-s) \big\|\alpha(s)\big\|_{\mathsf{D}(G^m)}\,ds \leq f^*_\varepsilon(t) \big\|\alpha\big\|_{L^1(0,t;\mathsf{D}(G^m))} \leq f^*_\varepsilon(T) \big\|\alpha\big\|_{L^1(0,T;\mathsf{D}(G^m))}. \quad \square$$

The estimates of $R^3_{m,\varepsilon}$ are more delicate. The reason is that the double integral in its definition, i.e.

$$R^3_{m,\varepsilon}(t) := \int_0^t S_A(t-s) \int_0^{t-s} (t-s-r)^{m-1} S_{\varepsilon G}(r) G^m B(s)\, dr\, dW(s),$$

is not a stochastic convolution. In fact, while it can be written as

$$\int_0^t R(t-s)B(s)\, dW(s), \qquad R(t) := S_A(t) \int_0^t (t-r)^{m-1} S_{\varepsilon G}(r) G^m\, dr,$$

the family of operators $(R(t))_{t\in\mathbb{R}_+}$ is not a semigroup. Unfortunately we are not aware of any maximal inequalities for such "nonlinear" stochastic convolutions. We shall nonetheless obtain estimates on the remainder term $R^3_{m,\varepsilon}$ by different arguments.

Lemma 4.5. *One has, for every $t \in [0, T]$ and $\varepsilon \in [0, 1]$,*

$$\big\| R^3_{m,\varepsilon}(t) \big\|_{L^p(\Omega;H)} \leq N_p M_A M_G f^*_\varepsilon(t) \big\| B \big\|_{L^p(\Omega;L^2(0,t;\mathscr{L}^2(U;\mathsf{D}(G^m))))}.$$

Proof. We shall use an argument analogous to the one used in the proof of Lemma 2.1. Let us set

$$C(t,s) := S_A(t-s) \int_0^{t-s} (t-s-r)^{m-1} S_{\varepsilon G}(r) G^m B(s)\, dr,$$

so that $R^3_{m,\varepsilon}(t) = \big(C(t,\cdot)\cdot W\big)_t$. Then

$$\big\| R^3_{m,\varepsilon}(t) \big\|_{L^p(\Omega;H)} \leq \sup_{t_0\in[t,t+\delta]} \big\| \big(C(t_0,\cdot)\cdot W\big)_t \big\|_{L^p(\Omega;H)},$$

with

$$\big\| \big(C(t_0,\cdot)\cdot W\big)_t \big\|_{L^p(\Omega;H)} \leq N_p \big\| C(t_0,\cdot) \big\|_{L^p(\Omega;L^2(0,t;\mathscr{L}^2(U;H)))}$$

and

$$\big\| C(t_0,s) \big\|_{\mathscr{L}^2(U;H)} \leq M_A M_G e^{w_A(t_0-s)} \int_0^{t_0-s} (t_0-s-r)^{m-1} e^{\varepsilon w_G r} \big\| B(s) \big\|_{\mathscr{L}^2(U;\mathsf{D}(G^m))}\, dr,$$

where

$$\exp(w_A(t_0-s)) \leq \exp\big(w_A(t+\delta 1_{\{w_A\geq 0\}}-s)\big),$$
$$\int_0^{t_0-s} (t_0-s-r)^{m-1} e^{\varepsilon w_G r}\, dr \leq \int_0^{t+\delta-s} (t+\delta-s-r)^{m-1} e^{\varepsilon w_G r}\, dr$$

hence

$$\big\|C(t_0,s)\big\|_{\mathscr{L}^2(U;H)} \leq M_A M_G \big\|B(s)\big\|_{\mathscr{L}^2(U;\mathsf{D}(G^m))} \cdot \\ \cdot \exp\big(w_A(t+\delta 1_{\{w_A\geq 0\}}-s)\big) \int_0^{t+\delta-s} (t+\delta-s-r)^{m-1} e^{\varepsilon w_G r}\,dr$$

for all $t_0 \in [t, t+\delta]$. In particular,

$$\sup_{t_0\in[t,t+\delta]} \big\|C(t_0,s)\big\|_{\mathscr{L}^2(U;H)} \leq M_A M_G \big\|B(s)\big\|_{\mathscr{L}^2(U;\mathsf{D}(G^m))} \cdot \\ \cdot \exp\big(w_A(t+\delta 1_{\{w_A\geq 0\}}-s)\big) \int_0^{t+\delta-s} (t+\delta-s-r)^{m-1} e^{\varepsilon w_G r}\,dr.$$

Moreover, setting

$$f_{\varepsilon,\delta}(t) := \exp\big(w_A(t+\delta 1_{\{w_A\geq 0\}})\big) \int_0^{t+\delta} (t+\delta-r)^{m-1} e^{\varepsilon w_G r}\,dr,$$

we can write

$$\begin{aligned}\big\|R^3_{m,\varepsilon}(t)\big\|_{L^p(\Omega;H)} &\leq \sup_{t_0\in[t,t+\delta]} \big\|\big(C(t_0,\cdot)\cdot W\big)_t\big\|_{L^p(\Omega;H)} \\ &\leq N_p \sup_{t_0\in[t,t+\delta]} \big\|C(t_0,\cdot)\big\|_{L^p(\Omega;L^2(0,t;\mathscr{L}^2(U;H)))} \\ &\leq N_p \Big\|\sup_{t_0\in[t,t+\delta]} \big\|C(t_0,\cdot)\big\|_{\mathscr{L}^2(U;\mathsf{D}(G^m))}\Big\|_{L^p(\Omega;L^2(0,t))} \\ &\leq N_p M_A M_G \bigg\|\bigg(\int_0^t f^2_{\varepsilon,\delta}(t-s)\big\|B(s)\big\|^2_{\mathscr{L}^2(U;\mathsf{D}(G^m))}\,ds\bigg)^{1/2}\bigg\|_{L^p(\Omega)}.\end{aligned}$$

Since $\delta > 0$ is arbitrary, taking the limit as $\delta \to 0$ yields

$$\begin{aligned}\big\|R^3_{m,\varepsilon}(t)\big\|_{L^p(\Omega;H)} &\leq N_p M_A M_G \big\|f_\varepsilon(t-\cdot)B\big\|_{L^p(\Omega;L^2(0,t;\mathscr{L}^2(U;\mathsf{D}(G^m))))} \\ &\leq N_p M_A M_G f^*_\varepsilon(t) \big\|B\big\|_{L^p(\Omega;L^2(0,t;\mathscr{L}^2(U;\mathsf{D}(G^m))))}. \qquad \square\end{aligned}$$

Lemma 4.6. *One has, for every $\varepsilon \in [0,1]$,*

$$\big\|R^3_{m,\varepsilon}\big\|_{\mathsf{C}^p} \leq \frac{T^m}{m} N_p M_A^2 M_G (e^{w_A T} \vee 1)\big(e^{(w_A+\varepsilon w_G)T} \vee 1\big)\big\|B\big\|_{L^p(\Omega;L^2(0,T;\mathscr{L}^2(U;\mathsf{D}(G^m))))}.$$

Proof. Thanks to the stochastic Fubini theorem, $R^3_{m,\varepsilon}(t)$ can be written as

$$\begin{aligned}&\int_0^t S_{\varepsilon G}(r) \int_0^{t-r} (t-r-s)^{m-1} S_A(t-s) G^m B(s)\,dW(s)\,dr \\ &\qquad = \int_0^t S_{A+\varepsilon G}(r) \int_0^{t-r} (t-r-s)^{m-1} S_A(t-r-s) G^m B(s)\,dW(s)\,dr,\end{aligned}$$

thus also, setting

$$\Phi(t) := \int_0^t (t-s)^{m-1} S_A(t-s) G^m B(s)\,dW(s) \qquad \forall t \in [0,T],$$

as $S_{A+\varepsilon G} * \Phi(t)$, with

$$\begin{aligned}\big\|S_{A+\varepsilon G} * \Phi(t)\big\| &\leq M_A M_G \int_0^t e^{w_A(t-s)} e^{\varepsilon w_G(t-s)} \big\|\Phi(s)\big\|\,ds \\ &\leq M_A M_G \big(e^{(w_A+\varepsilon w_G)T} \vee 1\big) \int_0^T \big\|\Phi(t)\big\|\,dt.\end{aligned}$$

Minkowski's inequality yields

$$\big\|S_{A+\varepsilon G} * \Phi\big\|_{\mathsf{C}^p} \leq M_A M_G \big(e^{(w_A+\varepsilon w_G)T} \vee 1\big) \int_0^T \big\|\Phi(t)\big\|_{L^p(\Omega;H)}\,dt,$$

where, by Lemma 2.1,

$$\begin{aligned}\big\|\Phi(t)\big\|_{L^p(\Omega;H)} &\leq N_p M_A \big\|(t-\cdot)^{m-1} e^{w_A(t-\cdot)} \|B\|_{\mathscr{L}^2(U;\mathsf{D}(G^m))}\big\|_{L^p(\Omega;L^2(0,t))} \\ &\leq N_p M_A t^{m-1} (e^{w_A t} \vee 1) \big\|B\big\|_{L^p(\Omega;L^2(0,t;\mathscr{L}^2(U;\mathsf{D}(G^m))))},\end{aligned}$$

hence

$$\int_0^T \big\|\Phi(t)\big\|_{L^p(\Omega;H)}\,dt \leq N_p M_A \big\|B\big\|_{L^p(\Omega;L^2(0,T;\mathscr{L}^2(U;\mathsf{D}(G^m))))} (e^{w_A T} \vee 1) \int_0^T t^{m-1}\,dt.$$

From this it follows that

$$\big\|R^3_{m,\varepsilon}\big\|_{\mathsf{C}^p} \leq \frac{T^m}{m} N_p M_A^2 M_G (e^{w_A T} \vee 1) \big(e^{(w_A+\varepsilon w_G)T} \vee 1\big) \big\|B\big\|_{L^p(\Omega;L^2(0,T;\mathscr{L}^2(U;\mathsf{D}(G^m))))}. \qquad \square$$

Remark 4.7. Lemmas 4.3, 4.4, and 4.5 continue to hold also if $p \in]0,1[$, while Lemma 4.6 does not, as its proof uses the Minkowski inequality, which reverses when $p < 1$.

We now have all the necessary tools to prove Proposition 4.1.

Proof of Proposition 4.1. It follows by commutativity of S_A and S_G that

$$u_\varepsilon = S_A S_{\varepsilon G} u_0 + S_A S_{\varepsilon G} * \alpha + S_A S_{\varepsilon G} \diamond B,$$

where, by the Taylor-like formula for strongly continuous semigroups of §2,

$$S_{\varepsilon G}(t) = \sum_{k=0}^{m-1} \frac{t^k}{k!} \varepsilon^k G^k + \frac{\varepsilon^m}{(m-1)!} \int_0^t (t-r)^{m-1} S_{\varepsilon G}(r) G^m,$$

as an identity of linear operators on $\mathsf{D}(G^m)$. Since $u_0 \in L^p(\Omega; \mathsf{D}(G^m))$, one has

$$S_{A+\varepsilon G}(t) u_0 = \sum_{k=0}^{m-1} \frac{\varepsilon^k t^k}{k!} S_A(t) G^k u_0 + \frac{\varepsilon^m}{(m-1)!} S_A(t) \int_0^t (t-r)^{m-1} S_{\varepsilon G}(r) G^m u_0 \, dr.$$

Similarly, since $\alpha \in L^p(\Omega; L^1(0,T; \mathsf{D}(G^m)))$ and $B \in L^p(\Omega; L^2(0,T; \mathscr{L}^2(U; \mathsf{D}(G^m))))$,

$$\begin{aligned} S_{A+\varepsilon G} * \alpha(t) &- \sum_{k=0}^{m-1} \int_0^t \frac{\varepsilon^k (t-s)^k}{k!} S_A(t-s) G^k \alpha(s) \, ds \\ &= \frac{\varepsilon^m}{(m-1)!} \int_0^t S_A(t-s) \int_0^{t-s} (t-s-r)^{m-1} S_{\varepsilon G}(r) G^m \alpha(s) \, dr \, ds \end{aligned}$$

as well as

$$\begin{aligned} S_{A+\varepsilon G} \diamond B(t) &- \sum_{k=0}^{m-1} \int_0^t \frac{\varepsilon^k (t-s)^k}{k!} S_A(t-s) G^k B(s) \, dW(s) \\ &= \frac{\varepsilon^m}{(m-1)!} \int_0^t S_A(t-s) \int_0^{t-s} (t-s-r)^{m-1} S_{\varepsilon G}(r) G^m B(s) \, dr \, dW(s). \end{aligned} \tag{7}$$

Therefore, recalling the definition of the processes $v_1, \dots, v_{m-1}$ and $(R_{m,\varepsilon})_{\varepsilon \in]0,1]}$, the desired decomposition follows. It remains to check that v_k belongs to C^p for every $k = 1, \dots, m-1$. For $m = 2$ one has

$$v_1 = \frac{1}{\varepsilon}\big(u - u_\varepsilon - R_{2,\varepsilon}\big),$$

where u and u_ε belong to C^p thanks to assumption (4), and $R_{2,\varepsilon}$ belongs to C^p by Lemmas 4.4 and 4.6. Therefore, choosing $\varepsilon \in]0,1]$ arbitrarily, the claim for $m = 2$ follows. By induction on m, the proof is completed. □

Remark 4.8. Estimates for the C^p-norm of v_k can be obtained in a more direct (and precise) way exploiting the dissipativity of A. In fact, one has

$$\big\|(\cdot)^k S_A G^k u_0\big\|_{\mathsf{C}^p} \leq M_A T^k e^{w_A T} \big\|u_0\big\|_{L^p(\Omega;\mathsf{D}(G^k))},$$

and

$$\bigg\|\int_0^{\cdot} (\cdot - s)^k S_A(\cdot - s) G^k \alpha(s)\,ds\bigg\|_{\mathsf{C}^p} \leq M_A T^k e^{w_A T} \big\|\alpha\big\|_{L^p(\Omega;L^1(0,T;\mathsf{D}(G^k)))},$$

(in fact just assuming that A is the generator of a strongly continuous semigroup), as well as, by maximal estimates for stochastic convolutions,

$$\bigg\|\int_0^{\cdot} (\cdot - s)^k S_A(\cdot - s) G^k B(s)\,dW(s)\bigg\|_{\mathsf{C}^p} \lesssim T^k e^{w_A T} \big\|B\big\|_{L^p(\Omega;L^2(0,T;\mathscr{L}^2(U;\mathsf{D}(G^k))))}.$$

Alternative assumptions on A yield similar estimates for the stochastic convolution, for instance if A generates an analytic semigroup. We shall not pursue this issue here.

The main result of the section now follows easily.

Theorem 4.9. *One has, for every $t \in [0,T]$,*

$$\begin{aligned}\big\|R_{m,\varepsilon}(t)\big\|_{L^p(\Omega;H)} \leq \frac{\varepsilon^m}{(m-1)!} M_A M_G \Big(& f_\varepsilon(t) \big\|u_0\big\|_{L^p(\Omega;\mathsf{D}(G^m))} \\ & + f_\varepsilon^*(t) \big\|\alpha\big\|_{L^p(\Omega;L^1(0,t;\mathsf{D}(G^m)))} + N_p f_\varepsilon^*(t) \big\|B\big\|_{L^p(\Omega;L^2(0,t;\mathscr{L}^2(U;\mathsf{D}(G^m))))}\Big).\end{aligned}$$

Moreover,

$$\begin{aligned}\big\|R_{m,\varepsilon}\big\|_{\mathsf{C}^p} \leq \frac{\varepsilon^m}{(m-1)!} M_A M_G \Big(& f_\varepsilon^*(T) \big\|u_0\big\|_{L^p(\Omega;\mathsf{D}(G^m))} \\ & + f_\varepsilon^*(T) \big\|\alpha\big\|_{L^p(\Omega;L^1(0,T;\mathsf{D}(G^m)))} \\ & + N_p \frac{T^m}{m} M_A (e^{w_A T} \vee 1)\big(e^{(w_A + \varepsilon w_G)T} \vee 1\big) \big\|B\big\|_{L^p(\Omega;L^2(0,T;\mathscr{L}^2(U;\mathsf{D}(G^m))))}\Big).\end{aligned}$$

In particular,

$$\lim_{\varepsilon \to 0} \frac{\big\|R_{m,\varepsilon}\big\|_{\mathsf{C}^p}}{\varepsilon^{m-1}} = 0.$$

Proof. This is an immediate consequence of the previous lemmas in this section, upon observing that f_ε and f_ε^* converge pointwise to a finite limit as $\varepsilon \to 0$. □

Remarks 4.10. (i) In view of Remark 4.7, the estimate in $L^p(\Omega; H)$ of $R_{m,\varepsilon}(t)$ in the previous theorem can be extended to the range of exponents $p \in]0, 1[$ thanks to the inequality

$$\|R^1_{m,\varepsilon} + R^2_{m,\varepsilon} + R^3_{m,\varepsilon}\|_{L^p(\Omega;H)} \le 3^{1/p-1}\Big(\|R^1_{m,\varepsilon}\|_{L^p(\Omega;H)} + \|R^2_{m,\varepsilon}\|_{L^p(\Omega;H)} + \|R^3_{m,\varepsilon}\|_{L^p(\Omega;H)}\Big).$$

(ii) It seems interesting to remark that, without any dissipativity assumption on A and G, the previous theorem implies that, as soon as $m \ge 1$, one has $u_\varepsilon \to u$ in C^p as $\varepsilon \to 0$ without the need to appeal to Theorem 3.1.

We are now going to identify the process v_k as the k-th derivative at zero of $\varepsilon \mapsto u_\varepsilon$. We shall actually prove more than this, namely that u_ε is m times continuously differentiable with respect to ε.

Proposition 4.11. *The map $\varphi\colon \varepsilon \mapsto u_\varepsilon$ is of class C^m from $[0, 1]$ to C^p, with*

$$D^k\varphi(0) = v_k \qquad \forall k \in \{1, \dots, m-1\}.$$

Proof. Let $\varepsilon \in [0, 1]$ and $h \in \mathbb{R}$ be such that $\varepsilon + h \in [0, 1]$. We begin by establishing first-order continuous differentiability. One has

$$\begin{aligned} u_{\varepsilon+h}(t) - u_\varepsilon(t) &= S_{A+\varepsilon G}(t)\big(S_{hG}(t)u_0 - u_0\big) \\ &\quad + \int_0^t S_{A+\varepsilon G}(t-s)\big(S_{hG}(t-s)\alpha(s) - \alpha(s)\big)\,ds \\ &\quad + \int_0^t S_{A+\varepsilon G}(t-s)\big(S_{hG}(t-s)B(s) - B(s)\big)\,dW(s), \end{aligned}$$

where, recalling that $S_{hG} = S_G(h\,\cdot)$,

$$\lim_{h\to 0} \frac{S_{hG}(t)u_0 - u_0}{h} = t \lim_{h\to 0} \frac{S_G(ht)u_0 - u_0}{ht} = tGu_0$$

for every $t \in [0, T]$, hence

$$\lim_{h\to 0} \frac{S_{A+\varepsilon G}\big(S_{hG}u_0 - u_0\big)}{h} = [t \mapsto tGu_0]$$

in C^p by dominated convergence. Similarly, one has

$$\begin{aligned} \lim_{h\to 0} \frac{S_{hG}(t-s)\alpha(s) - \alpha(s)}{h} &= (t-s) \lim_{h\to 0} \frac{S_G(h(t-s))\alpha(s) - \alpha(s)}{h(t-s)} \\ &= (t-s)G\alpha(s) \end{aligned}$$

for all $s, t \in [0, T]$ with $s \leq t$, hence, again by dominated convergence,

$$\lim_{h\to 0} \frac{1}{h} \int_0^{\cdot} S_{A+\varepsilon G}(\cdot - s)\big(S_{hG}(\cdot - s)\alpha(s) - \alpha(s)\big)\, ds$$
$$= \int_0^{\cdot} S_{A+\varepsilon G}(\cdot - s)(\cdot - s)G\alpha(s)\, ds$$

in C^p. The stochastic convolution term cannot be treated in the same way and requires more work. We shall write, for simplicity of notation, S_ε in place of $S_{A+\varepsilon G}$. Introducing the processes $y_\varepsilon = y_\varepsilon^0$ and y_ε^1 defined by

$$y_\varepsilon(t) := \int_0^t S_\varepsilon(t-s)B(s)\, dW(s),$$
$$y_\varepsilon^1(t) := \int_0^t S_\varepsilon(t-s)(t-s)GB(s)\, dW(s),$$

we need to show that

$$\lim_{h\to 0} \frac{y_{\varepsilon+h} - y_\varepsilon}{h} = y_\varepsilon^1 \qquad \text{in } \mathsf{C}^p. \tag{8}$$

Duhamel's formula yields

$$y_{\varepsilon+h}(t) = h\int_0^t S_\varepsilon(t-s)Gy_{\varepsilon+h}(s)\, ds + \int_0^t S_\varepsilon(t-s)B(s)\, dW(s)$$
$$= h\int_0^t S_\varepsilon(t-s)Gy_{\varepsilon+h}(s)\, ds + y_\varepsilon(t),$$

hence

$$\frac{y_{\varepsilon+h}(t) - y_\varepsilon(t)}{h} = \int_0^t S_\varepsilon(t-s)Gy_{\varepsilon+h}(s)\, ds$$
$$= \int_0^t S_\varepsilon(t-s)\int_0^s S_{\varepsilon+h}(s-r)GB(r)\, dW(r)\, ds.$$

Since $S_{\varepsilon+h} \diamond GB$ converges to $S_\varepsilon \diamond GB$ in C^p as $h \to 0$ by Theorem 3.1, it follows by dominated convergence that

$$\lim_{h\to 0}\frac{y_{\varepsilon+h}-y_\varepsilon}{h}=\int_0^{\cdot} S_\varepsilon(\cdot-s)\int_0^s S_\varepsilon(s-r)GB(r)\,dW(r)\,ds \qquad \text{in } \mathsf{C}^p.$$

Moreover, by Lemma 2.2,

$$\int_0^t S_\varepsilon(t-s)\int_0^s S_\varepsilon(s-r)GB(r)\,dW(r)\,ds=\int_0^t S_\varepsilon(t-s)(t-s)GB(s)\,dW(s)=y^1_\varepsilon(t),$$

thus (8) is proved. Furthermore, it follows by the assumptions on B that the same argument also yields the stronger statement

$$\lim_{h\to 0}\frac{G^j y_{\varepsilon+h}-G^j y_\varepsilon}{h}=G^j y^1_\varepsilon \qquad \text{in } \mathsf{C}^p \qquad \forall 0\le j\le m-1 \tag{9}$$

(with j integer). Let us turn to higher-order derivatives. We shall only consider the term involving the stochastic convolution, as the terms involving the initial datum and the deterministic convolution can be treated in a completely analogous (in fact easier) way. We need to show that the k-th derivative of $\varepsilon\mapsto y_\varepsilon$, denoted by $y^{(k)}$, satisfies

$$y^{(k)}_\varepsilon(t)=\int_0^t S_{A+\varepsilon G}(t-s)(t-s)^k G^k B(s)\,dW(s)=: y^k_\varepsilon(t)$$

for all $k\ge 2$, as the case $k=1$ has just been proved. We begin with some preparations. Lemma 2.2 implies that

$$y^k_\varepsilon = kS_\varepsilon * Gy^{k-1}_\varepsilon = k!\,S^{*k}_\varepsilon G^k y_\varepsilon = k!\,S^{*k}_\varepsilon S_\varepsilon \diamond G^k B \tag{10}$$

for every $k\in\{1,\dots,m\}$, where S^{*k}_ε denotes the operation of convolution with S_ε repeated k times, i.e.

$$S^{*1}_\varepsilon\phi := S_\varepsilon * \phi, \qquad S^{*k}_\varepsilon\phi = S_\varepsilon * (S^{*(k-1)}_\varepsilon\phi).$$

It follows by a repeated application of Theorem 3.1 that $G^j y^k_{\varepsilon+h}\to G^j y^k_\varepsilon$ in C^p as $h\to 0$ for all $j,k\in\mathbb{N}$ with $j+k\le m$. We shall now proceed by induction, i.e. we are going to prove that, for any $\varepsilon\in[0,1]$, $y^{(k)}_\varepsilon=y^k_\varepsilon$ implies $y^{(k+1)}_\varepsilon=y^{k+1}_\varepsilon$. Since $y^k_{\varepsilon+h}=kS_{\varepsilon+h}*Gy^{k-1}_{\varepsilon+h}$, Duhamel's formula yields, setting $z:=y^k_{\varepsilon+h}/k$,

$$z(t)=h\int_0^t S_\varepsilon(t-s)Gz(s)\,ds+\int_0^t S_\varepsilon(t-s)Gy^{k-1}_{\varepsilon+h}(s)\,ds,$$

therefore, by the identity $y^k_\varepsilon = kS_\varepsilon * Gy^{k-1}_\varepsilon$,

$$
\begin{aligned}
y^k_{\varepsilon+h}(t) - y^k_\varepsilon(t) &= k\Big(h\int_0^t S_\varepsilon(t-s)Gz(s)\,ds \\
&\qquad + \int_0^t S_\varepsilon(t-s)\big(Gy^{k-1}_{\varepsilon+h}(s) - Gy^{k-1}_\varepsilon(s)\big)\,ds\Big) \\
&= h\int_0^t S_\varepsilon(t-s)Gy^k_{\varepsilon+h}(s)\,ds \\
&\qquad + k\int_0^t S_\varepsilon(t-s)\big(Gy^{k-1}_{\varepsilon+h}(s) - Gy^{k-1}_\varepsilon(s)\big)\,ds,
\end{aligned}
$$

hence

$$
\begin{aligned}
\frac{y^k_{\varepsilon+h}(t) - y^k_\varepsilon(t)}{h} &= \int_0^t S_\varepsilon(t-s)Gy^k_{\varepsilon+h}(s)\,ds \\
&\qquad + k\int_0^t S_\varepsilon(t-s)G\frac{y^{k-1}_{\varepsilon+h}(s) - y^{k-1}_\varepsilon(s)}{h}\,ds,
\end{aligned}
$$

where, as discussed above, $Gy^k_{\varepsilon+h} \to Gy^k_\varepsilon$ in C^p as $h \to 0$, so that, by dominated convergence,

$$
\lim_{h\to 0}\int_0^\cdot S_\varepsilon(\cdot - s)Gy^k_{\varepsilon+h}(s)\,ds = \int_0^\cdot S_\varepsilon(\cdot - s)Gy^k_\varepsilon(s)\,ds
$$

in C^p. The inductive assumption means that

$$
\lim_{h\to 0}\frac{y^{(k-1)}_{\varepsilon+h} - y^{(k-1)}_\varepsilon}{h} = \lim_{h\to 0}\frac{y^{k-1}_{\varepsilon+h} - y^{k-1}_\varepsilon}{h} = y^k_\varepsilon
$$

in C^p for every $\varepsilon \in [0,1]$. The assumptions on B and (10) imply that the inductive assumption also yields, in complete analogy to the argument leading to (9), that

$$
\lim_{h\to 0}\frac{G^j y^{k-1}_{\varepsilon+h} - G^j y^{k-1}_\varepsilon}{h} = G^j y^k_\varepsilon
$$

for every positive integer j such that $j + k \le m$. Therefore, again by dominated convergence, we have

$$
\lim_{h\to 0}\int_0^\cdot S_\varepsilon(\cdot - s)G\frac{y^{k-1}_{\varepsilon+h}(s) - y^{k-1}_\varepsilon(s)}{h}\,ds = \int_0^\cdot S_\varepsilon(\cdot - s)Gy^k_\varepsilon(s)\,ds
$$

in C^p, hence we infer that

$$\lim_{h\to 0}\frac{y^{(k)}_{\varepsilon+h}-y^{(k)}_{\varepsilon}}{h}=\lim_{h\to 0}\frac{y^{k}_{\varepsilon+h}-y^{k}_{\varepsilon}}{h}=(k+1)S_\varepsilon * Gy^k_\varepsilon = y^{k+1}_\varepsilon,$$

thus concluding the proof of the induction step. □

4.1. *Asymptotic expansion of functionals of u_ε*

We are now going to consider asymptotic expansions of processes of the type $F(u_\varepsilon)$, where F is a functional taking values in a Banach space. All assumptions stated at the beginning of the sections are still in force.

We begin with a simple case.

Proposition 4.12. *Let E be a Banach space and $F\colon \mathsf{C}^p \to E$ be of class C^{m-1}, $m \geq 2$. Then there exist $w_1,\dots,w_{m-2}\in E$ and $R_{m-1,\varepsilon}\in E$ such that, for every $\varepsilon \in]0,1]$,*

$$F(u_\varepsilon)=F(u)+\sum_{n=1}^{m-2}\frac{\varepsilon^n}{n!}w_n + R_{m-1,\varepsilon},$$

where

$$w_n=\sum_{j=1}^{n}\sum_{\substack{k_1+\dots+k_n=j\\ k_1+2k_2+\dots+nk_n=n}}\frac{n!}{k_1!\cdots k_n!}D^jF(u)\Big(\big(v_1/1!\big)^{\otimes k_1},\dots,\big(v_n/n!\big)^{\otimes k_n}\Big) \tag{11}$$

and

$$\lim_{\varepsilon\to 0}\frac{R_{m-1,\varepsilon}}{\varepsilon^{m-2}}=0.$$

Proof. Since $\varepsilon\mapsto u_\varepsilon$ is of class C^m from $[0,1]$ to C^p by Proposition 4.11, it follows that $\varepsilon\mapsto F(u_\varepsilon)$ belongs to $C^{m-1}([0,1];E)$. The expression for $F(u_\varepsilon)$ then follows immediately by Taylor's theorem, and the expression for w_n follows by the formula for higher derivatives of composite functions (sometimes called Faà di Bruno's formula – see, e.g., [2, p. 272]). Furthermore, denoting the map $\varepsilon\mapsto u_\varepsilon$ by φ, one has

$$R_{m-1,\varepsilon}=\int_0^1\frac{(1-t)^{m-2}}{(m-2)!}D^{m-1}(F\circ\varphi)(\varepsilon t)\varepsilon^{m-1}\,dt,$$

where $D^{m-1}(F\circ\varphi)$ is bounded in E on the compact interval $[0,1]$ because it is continuous thereon. Denoting the maximum of the E-norm of this function on $[0,1]$ by M_1, we have

$$\frac{R_{m-1,\varepsilon}}{\varepsilon^{m-2}}\leq \varepsilon\frac{M_1}{(m-1)!},$$

where the right-hand side obviously tends to zero as $\varepsilon\to 0$. □

We shall now assume that $F \in C^m(\mathsf{C}^p; E)$ and derive an expansion of $u_\varepsilon - u$ of order $m-1$. Note that an argument based on Taylor's formula for $\varepsilon \mapsto F(u_\varepsilon)$, as in the previous proposition, does not work because $\varepsilon \mapsto u_\varepsilon$ is only of class C^{m-1}, hence its composition with F is also of class C^{m-1}. We are going to use instead a construction based on composition of power series.

Theorem 4.13. *Let E be a Banach space and $F\colon \mathsf{C}^p \to E$ be of class C^m, $m \geq 1$. Then there exist $w_1, \dots, w_{m-1} \in E$ and $R_{m,\varepsilon} \in E$ such that, for every $\varepsilon \in]0,1]$,*

$$F(u_\varepsilon) = F(u) + \sum_{n=1}^{m-1} \frac{\varepsilon^n}{n!} w_n + R_{m,\varepsilon},$$

where the (w_n) are defined as in (11) and

$$\lim_{\varepsilon \to 0} \frac{R_{m,\varepsilon}}{\varepsilon^{m-1}} = 0.$$

Proof. Taylor's formula applied to F yields

$$\begin{aligned} F(u_\varepsilon) - F(u) &= \sum_{n=1}^{m-1} D^n F(u)(u_\varepsilon - u)^{\otimes n} \\ &\quad + \int_0^1 \frac{(1-t)^{m-1}}{(m-1)!} D^m F(tu_\varepsilon + (1-t)u)(u_\varepsilon - u)^{\otimes m}\, dt, \end{aligned} \tag{12}$$

where, by Proposition 4.1,

$$u_\varepsilon - u = \sum_{k=1}^{m-1} \frac{v_k}{k!} \varepsilon^k + R_{m,\varepsilon} = \sum_{k=1}^{m-1} \frac{v_k}{k!} \varepsilon^k + \bar{R}_{m,\varepsilon} \varepsilon^m, \tag{13}$$

where $\bar{R}_{m,\varepsilon} \in \mathsf{C}^p$ by (5) and Lemmas 4.4 and 4.6. Multilinearity of the higher-order derivatives of F implies that

$$\sum_{n=1}^{m-1} D^n F(u)(u_\varepsilon - u)^{\otimes n} = \sum_{n=1}^{m-1} \frac{w_n}{n!} \varepsilon^n + \sum_{n=m}^{(m-1)m} a_n \varepsilon^n, \tag{14}$$

where w_n, $n = 1, \dots, m-1$ are defined as in (11), and the a_n are (finite) linear combinations of terms of the type

$$D^j F(u)\Big(v_1^{\otimes k_1}, \dots, v_n^{\otimes k_n}, \bar{R}_{m,\varepsilon}^{\otimes k_{n+1}}\Big),$$

where $j \in \{1, \dots, m-1\}$ and $k_1, \dots, k_{n+1} \in \mathbb{N}$, $k_1 + \dots + k_n + k_{n+1} = j$.

Let us show that $w_n \in E$ for every $n = 1, \dots, m-1$: by (11) it suffices to note that, for any $j = 1, \dots, n$ and $k_1 + \dots + k_n = j$,

$$\left\| D^j F(u)\Big(\big(v_1/1!\big)^{\otimes k_1}, \dots, \big(v_n/n!\big)^{\otimes k_n}\Big)\right\|_E \lesssim \left\| D^j F(u)\right\|_{\mathscr{L}_k(\mathsf{C}^p;E)} \|v_1\|_{\mathsf{C}^p}^{k_1} \cdots \|v_n\|_{\mathsf{C}^p}^{k_n},$$

where the right-hand side is finite because $u \in \mathsf{C}^p$ and $F \in C^m(\mathsf{C}^p; E)$ by assumption, and $v_1, \dots, v_{m-1} \in \mathsf{C}^p$ by Proposition 4.1. The proof that $a_n \in E$ for all $n = m, \dots, (m-1)m$, with norms bounded uniformly for $\varepsilon \in [0, 1]$, is entirely similar, as it immediately follows by Lemmas 4.4 and 4.6.

Finally, by multilinearity of $D^m F$, the integral on the right-hand side of (12) can be written as $\sum_{n=m}^{m^2} b_n \varepsilon^n$, where b_n depends on ε. By a reasoning entirely similar to the previous ones, in order to prove that $b_n \in E$ for all n and that their E-norms are bounded as $\varepsilon \to 0$, we proceed as follows: $D^m F$ is continuous, hence bounded on a neighborhood U of u. Without loss of generality, U can be assumed to be convex. Since $u_\varepsilon \to u$ in C^p as $\varepsilon \to 0$ by assumption, $u_\varepsilon \in U$ for ε sufficiently small, hence also $tu_\varepsilon + (1-t)u \in U$, so that $D^m F(tu_\varepsilon + (1-t)u)$ is bounded in $\mathscr{L}_k(\mathsf{C}^p; E)$ uniformly over ε in a (right) neighborhood of zero and $t \in [0, 1]$. Minkowski's inequality now implies that the E-norm of each b_n can be estimated uniformly with respect to ε. Setting

$$R_{m,\varepsilon} := \sum_{n=m}^{(m-1)m} a_n \varepsilon^n + \sum_{n=m}^{m^2} b_n \varepsilon^n,$$

the proof is completed. □

We now consider the case where F is defined only on $C([0, T]; H)$.

Theorem 4.14. *Let E_0 be a Banach space, $F\colon C([0, T]; H) \to E_0$ of class C^m. Assume that there exists $\beta \geq 0$ such that*

$$\left\| D^j F(x)\right\|_{\mathscr{L}_j(C([0,T];H);E_0)} \lesssim 1 + \|x\|_{C([0,T];H)}^{\beta} \qquad \forall j \leq m$$

and let $q > 0$ be defined by

$$\frac{\beta + m}{p} = \frac{1}{q}.$$

Then the conclusions of Theorem 4.13 hold with $E := L^q(\Omega; E_0)$.

Proof. Taylor's theorem implies that (12), (13) and (14) still hold, as identities of E_0-valued random variables. As in the proof of the previous theorem, the integral on the right-hand side of (12) can be written as the finite sum $\sum_{n=m}^{m^2} b_n \varepsilon^n$, with each b_n possibly depending on ε. We have to show that $w_n, a_n, b_n \in L^q(\Omega; E_0)$ for every n, and that the elements in (a_n) and (b_n) that depends on ε remain bounded in $L^q(\Omega; E_0)$ as $\varepsilon \to 0$. To this purpose, denoting the norms of $C([0, T]; H)$ and $\mathscr{L}_j(C([0, T]; H); E_0)$ by $\|\cdot\|$ and $\|\cdot\|_{\mathscr{L}_j}$, respectively, for simplicity of notation, note that one has, for any $j \leq m$,

$$\left\| D^j F(u)\Big(v_1^{\otimes k_1}, \dots, v_n^{\otimes k_n}, \bar{R}_{m,\varepsilon}^{\otimes k_{n+1}}\Big)\right\|_{E_0} \leq \left\| D^j F(u)\right\|_{\mathscr{L}_j} \|v_1\|^{k_1} \cdots \|v_n\|^{k_n} \left\|\bar{R}_{m,\varepsilon}\right\|^{k_{n+1}},$$

where $\|D^j F(u)\|_{\mathscr{L}_j} \lesssim 1 + \|u\|^\beta$ by assumption and $k_1 + \cdots + k_{n+1} = j \leq m$, hence

$$\frac{\beta}{p} + \frac{k_1}{p} + \cdots + \frac{k_{n+1}}{p} \leq \frac{\beta}{p} + \frac{m}{p} = \frac{1}{q}.$$

Applying Hölder's inequality with the exponents implied by this inequality yields

$$\begin{aligned}\Big\| D^j F(u)\Big(v_1^{\otimes k_1}, \dots, v_n^{\otimes k_n}, \bar{R}_{m,\varepsilon}^{\otimes k_{n+1}}\Big)\Big\|_{L^q(\Omega; E_0)} \\ \lesssim \big(1 + \|u\|_{\mathsf{C}^p}^\beta\big)\|v_1\|_{\mathsf{C}^p} \cdots \|v_n\|_{\mathsf{C}^p}\|\bar{R}_{m,\varepsilon}\|_{\mathsf{C}^p},\end{aligned}$$

where we have used the identity $\|z^\beta\|_{L^{p/\beta}(\Omega)} = \|z\|_{L^p(\Omega)}^\beta$, which holds for any positive random variable z. Recalling that $\bar{R}_{m,\varepsilon}$ is bounded in C^p uniformly over $\varepsilon \in [0,1]$, the claim about (w_n) and (a_n) is proved. In order to show that (b_n) enjoys the same properties of (a_n), it is immediately seen that it suffices to bound the norm of $D^m F(tu_\varepsilon + (1-t)u)$ in $L^{q/\beta}(\Omega; \mathscr{L}_m(C(0,T;H); E_0))$, uniformly with respect to ε in a right neighborhood of zero. But

$$\big\|D^m F(tu_\varepsilon + (1-t)u)\big\|_{\mathscr{L}_m} \lesssim 1 + \big\|u + t(u_\varepsilon - u)\big\|^\beta$$

implies

$$\begin{aligned}\big\|D^m F(tu_\varepsilon + (1-t)u)\big\|_{L^{p/\beta}(\Omega;\mathscr{L}_m)} &\lesssim 1 + \big\|u + t(u_\varepsilon - u)\big\|_{\mathsf{C}^p}^\beta \\ &\lesssim 1 + \big\|u\big\|_{\mathsf{C}^p}^\beta + \big\|u_\varepsilon - u\big\|_{\mathsf{C}^p}^\beta,\end{aligned}$$

where the norm in C^p of $u_\varepsilon - u$ tends to zero as $\varepsilon \to 0$. The proof is thus complete. □

Remark 4.15. One could have also approached the problem in a more abstract way, establishing conditions implying that the function F can be "lifted" to a function of class C^m from C^p to $E = L^q(\Omega; E_0)$, and then applying the Theorem 4.13. We have preferred the above more direct way because it could also be applicable, *mutatis mutandis*, in situations where F admits a series representation not necessarily of Taylor's type.

5. Singular perturbations of a transport equation and the Musiela SPDE

5.1. A transport equation

Let w be a fixed strictly positive real number and set, for notational convenience, $L^2_w := L^2(\mathbb{R}, e^{wx}dx)$. Let H be the Hilbert space of absolutely continuous functions $f \in L^1_{\mathrm{loc}}(\mathbb{R})$ such that $f' \in L^2_w$, equipped with the scalar product

$$\langle f, g\rangle := \lim_{x\to+\infty} f(x)g(x) + \langle f', g'\rangle_{L^2_w}.$$

Here and in the following, for any $\phi \in L^1_{\mathrm{loc}}(\mathbb{R})$, we denote by ϕ' its derivative in the sense of distributions. The definition of the scalar product in H is well posed because $f(+\infty) := \lim_{x\to+\infty} f(x)$ exists and is finite for every $f \in H$. In fact, for any $a, x \in \mathbb{R}$ with $x \geq a$, one has

$$|f(x) - f(a)| \leq \int_a^x |f'(y)|\,dy \leq \Big(\int_a^x |f'(y)|^2 e^{wy}\,dy\Big)^{1/2} \Big(\int_a^{+\infty} e^{-wy}\,dy\Big)^{1/2}. \tag{15}$$

We shall denote the norm in H induced by the above scalar product by $\|\cdot\|$. The following simple consequence of the definition of H will be repeatedly used below. For an open interval $I \subseteq \mathbb{R}$ and a natural number n, let $H^n(I)$ be the Sobolev space of functions in $L^2(I)$ with distributional derivatives of all orders up to n also belonging to $L^2(I)$. We shall write, with a harmless abuse of notation, $H^n(\mathbb{R}_+)$ to denote $H^n(]0, +\infty[)$. If f and f' belong to H, then $x \mapsto f'(x)e^{\frac{w}{2}x} \in H^1 := H^1(\mathbb{R})$, hence

$$\lim_{x\to\pm\infty} f'(x)e^{\frac{w}{2}x} = 0 \tag{16}$$

(see, e.g., [3, p. 214]), in particular $\lim_{x\to+\infty} f'(x) = 0$.

Let S_A be the strongly continuous semigroup on H defined by $[S_A(t)f](x) := f(t + x)$. The elementary identity

$$\int_{\mathbb{R}} |f'(t+x)|^2 e^{wx}\,dx = e^{-wt} \int_{\mathbb{R}} |f'(x)|^2 e^{wx}\,dx$$

implies that S_A is a contraction semigroup. Therefore, by the Lumer-Phillips theorem (see, e.g., [29, p. 60]), its generator A is a linear maximal dissipative operator on H. It follows from the definition of S_A that $Af = f'$ on $\mathsf{D}(A) = \big\{f \in H : f' \in H\big\}$.

One has the following formula of integration by parts (that also gives, as a special case, a direct proof of the dissipativity of A).

Lemma 5.1. *If $f, g \in \mathsf{D}(A)$, then*

$$\langle Af, g\rangle = -\langle f, Ag\rangle - w\big\langle f', g'\big\rangle_{L^2_w}.$$

In particular,

$$\langle Af, f\rangle = -\frac{w}{2}\|f'\|^2_{L^2_w}.$$

Proof. By definition one has

$$\langle Af, g\rangle = f'(+\infty)g(+\infty) + \int_{\mathbb{R}} f''(x)g'(x)e^{wx}\,dx,$$

where $g(+\infty)$ is finite and $f'(+\infty) = 0$. Integrating by parts yields

$$\begin{aligned}\langle Af, g\rangle &= \int_{\mathbb{R}} f''(x)g'(x)e^{wx}\,dx \\ &= \lim_{x\to+\infty} f'(x)g'(x)e^{wx} - \lim_{x\to-\infty} f'(x)g'(x)e^{wx}\end{aligned}$$

$$
\begin{aligned}
&- \int_{\mathbb{R}} f'(x)g''(x)e^{wx}\,dx - w \int_{\mathbb{R}} f'(x)g'(x)e^{wx}\,dx \\
&= -\langle f, Ag\rangle - w\big\langle f', g'\big\rangle_{L^2_w},
\end{aligned}
$$

where the two limits are equal to zero thanks to (16). $\square$

Let us now consider the operator A^2, defined on its natural domain $\mathsf{D}(A^2)$ of elements $f \in \mathsf{D}(A)$ such that $Af \in \mathsf{D}(A)$.

Lemma 5.2. *The operator A^2 is quasi-dissipative. More precisely, $A^2 - (w^2/2)I$ is dissipative.*

Proof. Let $f \in \mathsf{D}(A^2)$, and substitute $g = Af$ in the integration by parts formula of the previous lemma. We get

$$
\|Af\|^2 = -\langle A^2 f, f\rangle - w\big\langle f', f''\big\rangle_{L^2_w} = -\langle A^2 f, f\rangle - w\langle Af, f\rangle
$$

i.e.

$$
\langle A^2 f, f\rangle + w\langle Af, f\rangle = -\|Af\|^2,
$$

hence also

$$
\langle A^2 f, f\rangle - \frac{w^2}{2}\|f'\|^2_{L^2_w} = -\|Af\|^2
$$

and

$$
\langle A^2 f, f\rangle - \frac{w^2}{2}\|f\|^2 = -\|Af\|^2 - \frac{w^2}{2}|f(+\infty)|^2 \le 0. \quad \square
$$

Proposition 5.3. *The operator $G := A^2 - (w^2/2)I$ is maximal dissipative.*

Proof. The dissipativity of G has already been proved. Moreover, A^2 is closed, as is every integer positive power of the generator of a strongly continuous semigroup (see, e.g., [5, Proposition 1.1.6]). Hence we only have to show that there exist $\lambda > 0$ such that the image of $\lambda - G$ is H, or, equivalently, that there exists $\lambda > w^2/2$ such that the image of $\lambda - A$ is H. To this purpose, let $f \in H$ and consider the equation $\lambda y - y'' = f$, which yields $\lambda y' - y''' = f'$. Defining (formally, for the moment) z through $y'(x) = z(x)e^{-wx/2}$, one has

$$
y'''(x)e^{wx/2} = z''(x) - wz'(x) + \frac{w^2}{4}z(x), \tag{17}
$$

hence

$$
\lambda y'(x)e^{wx/2} - y'''(x)e^{wx/2} = \big(\lambda - w^2/4\big)z(x) + wz'(x) - z''(x).
$$

We are thus led to consider the equation

$$\big(\lambda - w^2/4\big)z + wz' - z'' = \widetilde{f}, \qquad \widetilde{f}(x) := f'(x)e^{wx/2}.$$

Let us introduce the bounded bilinear form a on H^1 defined as

$$a(\varphi,\psi) := \big(\lambda - w^2/4\big)\int_{\mathbb{R}} \varphi\psi + w\int_{\mathbb{R}} \varphi'\psi + \int_{\mathbb{R}} \varphi'\psi'.$$

One has

$$a(\varphi,\varphi) = \big(\lambda - w^2/4\big)\int_{\mathbb{R}} \varphi^2 + w\int_{\mathbb{R}} \varphi'\varphi + \int_{\mathbb{R}} (\varphi')^2,$$

where $\int_{\mathbb{R}} \varphi'\varphi = 0$, hence, for any $\lambda > w^2/4$, the bilinear form a is coercive on H^1. The Lax-Milgram theorem then yields the existence and uniqueness of a (weak) solution $z \in H^1$. Moreover, the equation satisfied by z implies that, in fact, $z \in H^2$. This immediately yields the existence of a solution y to $\lambda y - y'' = f$. Moreover, by definition of z it is immediate that $y \in H$, the identity $y''(x)e^{wx/2} = z'(x) - wz(x)/2$ implies that $y'' \in L^2_w$, and (17) implies that $y''' \in L^2_w$, i.e. $y \in \mathsf{D}(A^2)$, thus completing the proof. □

Since A is maximal dissipative, the transport equation on H

$$du = Au\,dt + \alpha(u)\,dt + B(u)\,dW, \qquad u(0) = u_0,$$

with α and B satisfying the measurability and Lipschitz continuity assumptions of §3 and $u_0 \in L^p(\Omega; H)$, $p > 0$, admits a unique mild solution $u \in \mathsf{C}^p$ (see, e.g., [8, Chapter 7] for the case $p \geq 2$ and [19] for the general case). Under the same assumptions on α, B, and u_0, the singularly perturbed equation

$$du_\varepsilon = (A + \varepsilon G)u_\varepsilon\,dt + \alpha(u_\varepsilon)\,dt + B(u_\varepsilon)\,dW, \qquad u_\varepsilon(0) = u_0,$$

admits a unique mild solution $u_\varepsilon \in \mathsf{C}^p$, which converges to u in C^p as $\varepsilon \to 0$. Furthermore, if the coefficients α and B do not depend on u and there exists an integer number $m \geq 1$ such that

$$u_0 \in L^p(\Omega; \mathsf{D}(A^{2m})), \quad \alpha \in L^p(\Omega; L^1(0,T; \mathsf{D}(A^{2m}))),$$
$$B \in L^p(\Omega; L^2(0,T; \mathscr{L}^2(U; \mathsf{D}(A^{2m})))),$$

then we can construct a representation of the difference $u_\varepsilon - u$ as a polynomial of degree $m-1$ plus a remainder term of higher order, applying the results of §4.

5.2. *Parabolic approximation of Musiela's SPDE*

Let $u(t,x)$, $t, x \geq 0$, denote the instantaneous forward rate at time t with maturity $t + x$. Musiela observed that the equation for forward rates in the Heath-Jarrow-Morton model can be written as (the mild form of)

$$du(t,x) = \partial_x u(t,x)\,dt + \alpha_0(t,x)\,dt + \sum_{k=1}^{\infty} \sigma_k(t,x)dw^k(t), \tag{18}$$

where $(w^k)_{k\in\mathbb{N}}$ is a sequence of standard real Wiener processes, the volatilities σ_k are possibly random, and α_0 is uniquely determined by (σ_k) if the reference probability measure is such that implied discounted bond prices are local martingales. In particular, in this case it must necessarily hold

$$\alpha_0(t,x) = \sum_{k=1}^{\infty} \sigma_k(t,x) \int_0^x \sigma_k(t,y)\,dy.$$

For details on the financial background we refer to, e.g., [6,10,12,23,24]. There is a large literature on the well-posedness of (18) in the mild sense, also in the (more interesting) case where (σ_k), hence α_0, depend explicitly on the unknown u, with different choices of state space as well as with more general noise (see, e.g., [1,4,10,15,18,28], [25, §20.3]). Here we limit ourselves to the case where (σ_k) are possibly random, but do not depend explicitly on u, and use as state space $H(\mathbb{R}_+)$, which we define as the space of locally integrable functions on $\mathbb{R}_+$ such that $f' \in L^2(\mathbb{R}_+, e^{wx}\,dx)$, endowed with the inner product

$$\langle f, g\rangle = f(+\infty)g(+\infty) + \int_0^{+\infty} f'(x)g'(x)e^{wx}\,dx.$$

This choice of state space, introduced in [10] (cf. also [27]), to which we refer for further details, is standard and enjoys many good properties from the point of view of financial modeling. For instance, forward curves are continuous and can be “flat” at infinity without decaying to zero.

In order to give a precise notion of solution to (18), we recall that the semigroup of left translation on $H(\mathbb{R}_+)$ is strongly continuous and contractive, and that its generator is $A_0\colon \phi \mapsto \phi'$ on the domain $\mathsf{D}(A_0) = \{\phi \in H(\mathbb{R}_+) : \phi' \in H(\mathbb{R}_+)\}$ (see [10]). Moreover, let us assume that there exists $p > 0$ such that

$$\mathbb{E}\biggl(\sum_{k=1}^{\infty} \int_0^T \bigl\|\sigma_k(t,\cdot)\bigr\|^2_{H(\mathbb{R}_+)} dt\biggr)^p < \infty, \qquad \sigma_k(t,+\infty) = 0 \quad \forall k \in \mathbb{N}, \tag{19}$$

so that the random time-dependent linear map

$$\begin{aligned} B_0(\omega,t)\colon \ell^2 &\longrightarrow H(\mathbb{R}_+) \\ (h_k) &\longmapsto \sum_{k=1}^{\infty} \sigma_k(\omega,t,\cdot)h_k \end{aligned}$$

belongs to $L^{2p}(\Omega; L^2(0,T; \mathscr{L}^2(\ell^2; H(\mathbb{R}_+))))$. Setting, for any $\phi \in L^1_{\mathrm{loc}}(\mathbb{R}_+)$,

$$[I\phi](x) := \int_0^x \phi(y)\,dy, \qquad x \geq 0,$$

one has the basic estimate

$$\|\phi\, I\phi\|_{H(\mathbb{R}_+)} \lesssim \|\phi\|^2_{H(\mathbb{R}_+)}$$

for every $\phi \in H(\mathbb{R}_+)$ such that $\phi(+\infty) = 0$ (cf. [10], or see Lemma 5.6 below for a proof in a more general setting). This implies that the assumption on (σ_k) yields $\alpha_0 \in L^p(\Omega; L^1(0,T; H(\mathbb{R}_+)))$. We then have the following well-posedness result for (18), written in its abstract form as

$$du + A_0 u\,dt = \alpha_0\,dt + B_0\,dW, \qquad u(0) = u_0, \tag{20}$$

where W is a cylindrical Wiener process on $U := \ell^2$.

Proposition 5.4. *Let $p > 0$. Assume that $u_0 \in L^p(\Omega, \mathscr{F}_0; H(\mathbb{R}_+))$ and (19) is satisfied. Then (20) has a unique mild solution $u \in \mathsf{C}^p(H(\mathbb{R}_+)) := L^p(\Omega; C([0,T]; H(\mathbb{R}_+)))$, which depends continuously on the initial datum u_0.*

Musiela's equation (20) is closely related to the transport equation studied in §5.1 above, the main difference being the state space. In the following we shall denote the state space of the transport equation by $H(\mathbb{R})$.

As mentioned in the introduction, it has been suggested (see [7] and references therein) that second-order parabolic SPDEs, with respect to the physical probability measure, capture several empirical features of observed forward rates. It seems reasonable to assume that such SPDEs would retain their parabolic character even after changing the reference probability measure to one with respect to which discounted bond prices are (local) martingales, thus excluding arbitrage. It is then natural to consider singular perturbations of the Musiela equation on $H(\mathbb{R}_+)$ adding a singular term εG to the drift A_0 in (20), with $G = A_0^2$, which is, roughly speaking, a second derivative in the time to maturity. On the other hand, if forward rates satisfy the general assumptions of the Heath-Jarrow-Morton model, the HJM drift condition is sufficient and necessary for discounted bond prices to be local martingales. Therefore singular perturbations of the Musiela SPDE introduce arbitrage, in the sense that the implied discounted bond prices may not be local martingales. It is hence interesting to obtain quantitative estimates, loosely speaking, on the arbitrage introduced by a parabolic perturbation of the Musiela SPDE (20). The arguments used in §5.1 for the transport equation, however, give rise to major problems, mainly because boundary terms (at zero) appear that seem difficult to control. To circumvent these issues, we "embed" the abstract Musiela equation (20) into a transport equation on $H(\mathbb{R})$ of the type considered in §5.1, we perturb the equation thus obtained, get asymptotic expansions, and finally "translate" back the results, in a suitable sense, to the Musiela equation.

We need some technical preparations first. Let $H_0(\mathbb{R}_+)$ be the Hilbert space of functions in $H(\mathbb{R}_+)$ that are zero at infinity. The following embeddings and estimates are rather straightforward (see [10] for a proof) and will be repeatedly used below:

(i) $H(\mathbb{R}_+) \hookrightarrow C_b(\mathbb{R}_+)$;

(ii) $H_0(\mathbb{R}_+) \hookrightarrow L^1(\mathbb{R}_+)$;
(iii) $H_0(\mathbb{R}_+) \hookrightarrow L^4_w(\mathbb{R}_+) := L^4(\mathbb{R}_+, e^{wx}\,dx)$.

Let $H^m_w(\mathbb{R}_+)$ be the set of functions in $L^1_{\mathrm{loc}}(\mathbb{R}_+)$ that belong to $L^2_w(\mathbb{R}_+)$ together with all their derivatives up to order m, endowed with the norm defined by

$$\|f\|_{H^m_w(\mathbb{R}_+)} = \sum_{k=0}^m \|f^{(k)}\|_{L^2_w(\mathbb{R}_+)}.$$

Lemma 5.5. *Let $f \in L^1_{\mathrm{loc}}(\mathbb{R}_+)$ and m a positive integer. The following assertions are equivalent:* (a) $f \in \mathsf{D}(A_0^m)$*;* (b) $f' \in H^m_w(\mathbb{R}_+)$*;* (c) $x \mapsto f'(x)e^{wx/2} \in H^m(\mathbb{R}_+)$*. Moreover, for any $f \in \mathsf{D}(A_0^m)$ with $f(+\infty) = 0$,*

$$\|f\|_{\mathsf{D}(A_0^m)} = \|f'\|_{H^m_w(\mathbb{R}_+)} \eqsim \|f' e^{w\cdot/2}\|_{H^m(\mathbb{R}_+)},$$

where the implicit constant depends only on m and w.

Proof. The equivalence of (a) and (b) is immediate by the definition of A_0 and by an inequality completely analogous to (15). In particular, if $f(+\infty) = 0$, the identity $\|f\|_{\mathsf{D}(A_0^m)} = \|f'\|_{H^m_w(\mathbb{R}_+)}$ is a tautology. The other assertions follow by the identity

$$\big(f' e^{w\cdot/2}\big)^{(n)} = \sum_{j=0}^n \binom{n}{j} (w/2)^{n-j} f^{(j+1)} e^{w\cdot/2} \qquad \forall n \in \{1, \dots, m\}. \quad \square$$

Lemma 5.6. *Let $m \geq 1$ be an integer. If $f \in \mathsf{D}(A_0^m)$ with $f(+\infty) = 0$, then*

$$\|f\,If\|_{\mathsf{D}(A_0^m)} \lesssim \|f\|^2_{\mathsf{D}(A_0^m)},$$

where the implicit constant depends only on m and w.

Proof. Let $f \in \mathsf{D}(A_0^m)$ with $f(+\infty) = 0$. In view of the previous lemma, we will bound the $H^m_w(\mathbb{R}_+)$ norm of $(f\,If)' = f'\,If + f^2$ in terms of the $H^m_w(\mathbb{R}_+)$ norm of f'. One has, omitting the indication of $\mathbb{R}_+$ in the notation,

$$\|f'\,If\|_{L^2_w} \leq \|f'\|_{L^2_w} \|If\|_{L^\infty} \leq \|f'\|_{L^2_w} \|f\|_{L^1} \lesssim \|f'\|_{L^2_w} \|f'\|_{L^2_w}$$

and

$$\|f^2\|_{L^2_w} = \|f\|^2_{L^4_w} \lesssim \|f'\|^2_{L^2_w}.$$

Let $1 \leq n \leq m$ be an integer. One has

$$\big(f'\,If\big)^{(n)} = \sum_{j=0}^n \binom{n}{j} f^{(j+1)} (If)^{(n-j)}$$

and

$$\big\| f^{(j+1)} (If)^{(n-j)} \big\|_{L^2_w} \le \big\| f^{(j+1)} \big\|_{L^2_w} \big\| (If)^{(n-j)} \big\|_{L^\infty},$$

where, if $j = n$,

$$\big\| (If)^{(n-j)} \big\|_{L^\infty} = \big\| If \big\|_{L^\infty} \le \big\| f \big\|_{L^1} \lesssim \big\| f' \big\|_{L^2_w},$$

while, if $j \le n-1$,

$$\big\| (If)^{(n-j)} \big\|_{L^\infty} = \big\| f^{(n-1-j)} \big\|_{L^\infty} \lesssim \big\| f^{(n-j)} \big\|_{L^2_w}.$$

Similarly,

$$(f^2)^{(n)} = \sum_{j=0}^{n} \binom{n}{j} f^{(j)} f^{(n-j)} = 2f\, f^{(n)} + \sum_{j=1}^{n-1} \binom{n}{j} f^{(j)} f^{(n-j)}$$

where

$$\big\| f\, f^{(n)} \big\|_{L^2_w} \le \big\| f \big\|_{L^\infty} \big\| f^{(n)} \big\|_{L^2_w} \lesssim \big\| f' \big\|_{L^2_w} \big\| f^{(n)} \big\|_{L^2_w}$$

and, if $1 \le j \le n-1$,

$$\big\| f^{(j)} f^{(n-j)} \big\|_{L^2_w} \le \big\| f^{(j)} \big\|_{L^\infty} \big\| f^{(n-j)} \big\|_{L^2_w} \lesssim \big\| f^{(j+1)} \big\|_{L^2_w} \big\| f^{(n-j)} \big\|_{L^2_w}.$$

The claim is then an immediate consequence of these estimates. $\square$

Proposition 5.7. *Let $p > 0$ and m a positive integer. If*

$$\mathbb{E}\biggl(\sum_{k=1}^{\infty} \int_0^T \big\| \sigma_k(t,\cdot) \big\|^2_{\mathsf{D}(A_0^{2m})} \, dt \biggr)^p < \infty$$

or, equivalently, $B_0 \in L^{2p}(\Omega; L^2(0,T; \mathscr{L}^2(\ell^2; \mathsf{D}(A_0^{2m}))))$, then $\alpha_0 \in L^p(\Omega; L^1(0,T; \mathsf{D}(A_0^{2m})))$.

Proof. One has

$$\big\| B_0(t) \big\|^2_{\mathscr{L}^2(\ell^2; \mathsf{D}(A_0^{2m}))} = \sum_{k=1}^{\infty} \big\| \sigma_k(t,\cdot) \big\|^2_{\mathsf{D}(A_0^{2m})}$$

and, by the previous lemma,

$$\begin{aligned}\big\|\alpha_0(t)\big\|_{\mathsf{D}(A_0^{2m})} &= \Big\|\sum_{k=1}^{\infty}\sigma_k(t,\cdot)\,I\sigma_k(t,\cdot)\Big\|_{\mathsf{D}(A_0^{2m})}\\ &\leq \sum_{k=1}^{\infty}\big\|\sigma_k(t,\cdot)\,I\sigma_k(t,\cdot)\big\|_{\mathsf{D}(A_0^{2m})}\\ &\lesssim \sum_{k=1}^{\infty}\big\|\sigma_k(t,\cdot)\big\|^2_{\mathsf{D}(A_0^{2m})} = \big\|B_0(t)\big\|^2_{\mathscr{L}^2(\ell^2;\mathsf{D}(A_0^{2m}))},\end{aligned}$$

hence

$$\begin{aligned}\big\|\alpha_0\big\|_{L^p(\Omega;L^1(0,T;\mathsf{D}(A_0^{2m})))} &\lesssim \Big\|\big\|B_0\big\|^2_{\mathscr{L}^2(\ell^2;\mathsf{D}(A_0^{2m}))}\Big\|_{L^p(\Omega;L^1(0,T))}\\ &= \big\|B_0\big\|_{L^{2p}(\Omega;L^2(0,T;\mathscr{L}^2(\ell^2;\mathsf{D}(A_0^{2m}))))}. \qquad \square\end{aligned}$$

Recall that the operator $A\colon \phi\mapsto\phi'$ is defined on $\mathsf{D}(A)=\{\phi\in H(\mathbb{R}) : \phi'\in H(\mathbb{R})\}$, and the operator $A_0\colon \phi\mapsto\phi'$ is defined on $\mathsf{D}(A_0)=\{\phi\in H(\mathbb{R}_+) : \phi'\in H(\mathbb{R}_+)\}$.

Lemma 5.8. *There exists a linear continuous extension operator* $L\colon \mathsf{D}(A_0^{2m})\to\mathsf{D}(A^{2m})$ *for every positive integer* m.

Proof. By an extension result due to Stein (see [26, p. 181]), there exists a linear continuous extension operator $L_0\colon H^{2m}(\mathbb{R}_+)\to H^{2m}(\mathbb{R})$. Since a locally integrable function f belongs to $\mathsf{D}(A_0^{2m})$ if and only if $x\mapsto f'(x)e^{wx/2}\in H^{2m}(\mathbb{R}_+)$ by Lemma 5.5, and $f\in H(\mathbb{R}_+)$ implies that $f(+\infty)$ is finite, the map

$$L\colon f\longmapsto \Big[x\mapsto f(+\infty)-\int_x^{+\infty} e^{-wy/2}L_0\big(f'e^{w\cdot/2}\big)(y)\,dy\Big]$$

is well defined on $\mathsf{D}(A_0^{2m})$. Moreover, $L_0\big(f'e^{w\cdot/2}\big)\in H^{2m}(\mathbb{R})$, hence $y\mapsto e^{-wy/2}L_0\big(f'e^{w\cdot/2}\big)(y)\in L^1(x,+\infty)$ by Cauchy's inequality for all $x\in\mathbb{R}$, so that $Lf(x)$ is finite for every $x\in\mathbb{R}$ and $Lf(+\infty):=\lim_{x\to+\infty}Lf(x)=f(+\infty)$. Moreover,

$$x\mapsto e^{wx/2}(Lf)'(x)=L_0\big(f'e^{w\cdot/2}\big)\in H^{2m}(\mathbb{R}),$$

hence $Lf\in\mathsf{D}(A^{2m})$ by an argument completely analogous to the proof of Lemma 5.5. Finally,

$$\begin{aligned}\big\|Lf\big\|_{\mathsf{D}(A^{2m})} &\lesssim |f(+\infty)| + \big\|(Lf)'e^{w\cdot/2}\big\|_{H^{2m}(\mathbb{R})}\\ &= |f(+\infty)| + \big\|L_0(f'e^{w\cdot/2})\big\|_{H^{2m}(\mathbb{R})}\\ &\lesssim |f(+\infty)| + \big\|f'e^{w\cdot/2}\big\|_{H^{2m}(\mathbb{R}_+)}\\ &\lesssim \big\|f\big\|_{\mathsf{D}(A_0^{2m})}. \qquad \square\end{aligned}$$

Let L be the extension operator just introduced and set $v_0 := Lu_0$, $\alpha := L\alpha_0$, and $B := LB_0$, with α_0 and B_0 as in Proposition 5.7, so that

$$v_0 \in L^p(\Omega, \mathscr{F}_0; \mathsf{D}(A^{2m})), \qquad \alpha \in L^p(\Omega; L^1(0,T; \mathsf{D}(A^{2m}))),$$
$$B \in L^p(\Omega; L^2(0,T; \mathscr{L}^2(\ell^2; \mathsf{D}(A^{2m})))),$$

and consider the following stochastic equation in $H(\mathbb{R})$:

$$dv = Av\,dt + \alpha\,dt + B\,dW, \qquad v(0) = v_0, \quad t \geq 0, \tag{21}$$

where A is the generator of the semigroup of translation on $H(\mathbb{R})$ and W is a cylindrical Wiener process on ℓ^2. By the discussion at the beginning of this section, this equation admits a unique mild solution $v \in \mathsf{C}^p(\mathsf{D}(A^{2m}))$, which is thus also a strong solution, i.e. such that

$$v(t) = v_0 + \int_0^t Av(s)\,ds + \int_0^t \alpha(s)\,ds + \int_0^t B(s)\,dW(s),$$

where the equality is in the sense of indistinguishable $H(\mathbb{R})$-valued (hence also $C(\mathbb{R})$-valued) processes. In a more explicit form, one has

$$v(t,x) = v_0(x) + \int_0^t \partial_x v(s,x)\,ds + \int_0^t \alpha(s,x)\,ds + \sum_{j=1}^\infty \int_0^t \sigma_j(s,x)\,dw^j(s)$$

for every $x \in \mathbb{R}$, in particular for every $x \in \mathbb{R}_+$. Since the restrictions of v_0, α and B to $\mathbb{R}_+$ are equal to u_0, α_0 and B_0, respectively, the restriction of v to $\mathbb{R}_+$ must coincide with the unique strong solution in $H(\mathbb{R}_+)$ to the Musiela equation (20). Moreover, the equation in $H(\mathbb{R})$

$$dv_\varepsilon = (A + \varepsilon A^2)v_\varepsilon\,dt + \alpha\,dt + B\,dW, \qquad v_\varepsilon(0) = v_0, \tag{22}$$

also admits a unique mild solution $v_\varepsilon \in \mathsf{C}^p(\mathsf{D}(A^{2m}))$, that converges to v in $\mathsf{C}^p(\mathsf{D}(A^{2m}))$ as $\varepsilon \to 0$. Let $p \in [1, \infty[$. It follows from Proposition 4.1 and Theorem 4.9, setting

$$v_k(t) := t^k S_A(t) A^{2k} u_0 + \int_0^t (t-s)^k S_A(t-s) A^{2k} \alpha(s)\,ds + \sum_{j=1}^\infty \int_0^t (t-s)^k S_A(t-s) A^{2k} \sigma_j(s)\,dw^j(s),$$

that v_ε satisfies an identity of the type

$$v_\varepsilon - v = \sum_{k=1}^{m-1} \frac{v_k}{k!} \varepsilon^k + R_{m,\varepsilon}$$

in $H(\mathbb{R})$, in particular in $C(\mathbb{R})$, where $v_1, \dots, v_{m-1}$, $R_{m,\varepsilon} \in \mathsf{C}^p(H(\mathbb{R}))$ and $R_{m,\varepsilon}/\varepsilon^{m-1}$ tends to zero in $\mathsf{C}^p(H(\mathbb{R}))$ as $\varepsilon \to 0$. Taking the $H(\mathbb{R}_+)$ norm on both sides yields

$$\big\|v_\varepsilon - v\big\|_{H(\mathbb{R}_+)} \le \sum_{k=1}^{m-1} \frac{1}{k!}\varepsilon^k \big\|v_k\big\|_{H(\mathbb{R}_+)} + \big\|R_{m,\varepsilon}\big\|_{H(\mathbb{R}_+)},$$

where all $H(\mathbb{R}_+)$ norms involved are finite because they are dominated by the corresponding ones in $H(\mathbb{R})$, that are finite. We have thus proved the following.

Theorem 5.9. *Let $p \in [1, \infty[$ and $m \ge 1$ be a positive integer such that*

$$\mathbb{E}\bigg(\sum_{k=1}^{\infty} \int_0^T \big\|\sigma_k(t,\cdot)\big\|^2_{\mathsf{D}(A_0^{2m})}\,dt\bigg)^p < \infty.$$

Then equation (21) has a unique strong solution v in $\mathsf{C}^p(H(\mathbb{R}))$ and its restriction to $H(\mathbb{R}_+)$ coincides with the unique strong solution u in $\mathsf{C}^p(H(\mathbb{R}_+))$ to the Musiela equation (20). Moreover, the restriction to $H(\mathbb{R}_+)$ of the mild solution v_ε to the perturbed extended Musiela equation (22) converges to v in $\mathsf{C}^p(\mathsf{D}(A^{2m}))$ and the estimate

$$\big\|v_\varepsilon - u\big\|_{\mathsf{C}^p(H(\mathbb{R}_+))} \le \sum_{k=1}^{m-1} \frac{1}{k!}\varepsilon^k \big\|v_k\big\|_{\mathsf{C}^p(H(\mathbb{R}_+))} + \big\|R_{m,\varepsilon}\big\|_{\mathsf{C}^p(H(\mathbb{R}_+))}$$

holds, with $\lim_{\varepsilon\to 0} R_{m,\varepsilon}/\varepsilon^{m-1} = 0$ in $\mathsf{C}^p(H(\mathbb{R}_+))$.

We shall now consider bond prices and their approximation in the diffusive correction of Musiela's equation. As already observed, the solutions v and v_ε to the equations (21) and (22) have paths in $H(\mathbb{R})$, hence their restrictions $x \mapsto v(t,x)$ and $x \mapsto v_\varepsilon(t,x)$, $x \in \mathbb{R}_+$, belong to $H(\mathbb{R}_+)$ for every $t \in [0,T]$ and $u(t,x) = v(t,x)$ for every $(t,x) \in [0,T] \times \mathbb{R}_+$. The price of a zero-coupon bond with face value equal to one at time $t \ge 0$ with time to maturity $x \ge 0$ is given by

$$\widehat{P}(t,x) = \exp\bigg(-\int_t^{t+x} v(t,t+y)\,dy\bigg) = \exp\bigg(-\int_0^x v(t,y)\,dy\bigg),$$

and the value at time t of the money market account is given by

$$\beta(t) = \exp\bigg(\int_0^t v(s,0)\,ds\bigg),$$

hence the corresponding discounted price of the zero-coupon bond is

$$P(t,x) := \frac{\widehat{P}(t,x)}{\beta(t)} = \exp\bigg(-\int_0^x v(t,y)\,dy - \int_0^t v(s,0)\,ds\bigg).$$

Let us define the discounted price of the (fictitious) zero coupon bond associated to v_ε as

$$P_\varepsilon(t,x) = \exp\Big(- \int_0^x v_\varepsilon(t,y)\,dy - \int_0^t v_\varepsilon(s,0)\,ds \Big).$$

For fixed $t \in [0,T]$ and $x \geq 0$, let us define the linear map

$$\begin{aligned} F_{t,x} \colon C([0,T]\times\mathbb{R}) &\longrightarrow \mathbb{R} \\ f &\longmapsto \int_0^x f(t,y)\,dy + \int_0^t f(s,0)\,ds \end{aligned}$$

so that $P(t,x) = \exp\big(-F_{t,x}v\big)$ and $P_\varepsilon(t,x) = \exp\big(-F_{t,x}v_\varepsilon\big)$.

Lemma 5.10. *Let $(t,x) \in [0,T]\times\mathbb{R}_+$. The linear map $F_{t,x}$ is continuous*

(i) *from $C([0,T]; H(\mathbb{R}_+))$ to $\mathbb{R}$, hence also from $C([0,T]; H(\mathbb{R}))$ to $\mathbb{R}$, and*
(ii) *from $\mathsf{C}^p(H(\mathbb{R}_+))$ to $L^p(\Omega)$, hence also from $\mathsf{C}^p(H(\mathbb{R}))$ to $L^p(\Omega)$, for every $p > 0$.*

Proof. For any $f \in C([0,T]; H(\mathbb{R}))$ one has

$$\begin{aligned} \Big| \int_0^x f(t,y)\,dy \Big| &\leq \int_0^x |f(t,y) - f(t,+\infty)|\,dy + |f(t,+\infty)|x \\ &\lesssim (1+x)\,\|f(t)\|_{H(\mathbb{R}_+)} \leq (1+x)\,\|f\|_{C([0,T];H(\mathbb{R}_+))} \end{aligned}$$

and

$$\Big| \int_0^t f(s,0)\,ds \Big| \leq \int_0^t \|f(s)\|_{L^\infty(\mathbb{R}_+)}\,ds \lesssim T\,\|f\|_{C([0,T];H(\mathbb{R}_+))},$$

thus proving (i). Raising both sides of both inequalities to the power p and taking expectations proves (ii). □

More generally, it is easy to show that the linear map F defined as

$$\begin{aligned} F \colon C([0,T]\times\mathbb{R}) &\longrightarrow C([0,T]\times\mathbb{R}) \\ f &\longmapsto \Big[(t,x) \mapsto \int_0^x f(t,y)\,dy + \int_0^t f(s,0)\,ds \Big] \end{aligned}$$

is continuous from $C([0,T]; H(\mathbb{R}_+))$ to $C([0,T]\times\mathbb{R})$, endowed with the topology of uniform convergence on compact sets, as well as from $\mathsf{C}^p(H(\mathbb{R}_+))$ to $L^p(\Omega; C([0,T]\times I))$ for every compact set $I \subset \mathbb{R}_+$.

The operators $F_{t,x}$ and F, being linear and continuous, are automatically of class C^∞, with $F'(z) = F$ for every z in the domain of F, and higher-order derivatives equal to zero (and completely analogously for $F_{t,x}$). Given an expansion of v_ε around v of the type

$$v_\varepsilon - v = \sum_{k=1}^{m-1} \frac{1}{k!} v_k \, \varepsilon^k + R_{m,\varepsilon},$$

which can be considered as an identity in $\mathsf{C}^p(H(\mathbb{R}))$, as well as in $\mathsf{C}^p(H(\mathbb{R}_+))$ by restriction, it follows immediately that

$$F_{t,x} v_\varepsilon - F_{t,x} v = \sum_{k=1}^{m-1} \frac{1}{k!} F_{t,x} v_k \, \varepsilon^k + F_{t,x} R_{m,\varepsilon}, \tag{23}$$

as an identity in $L^p(\Omega)$. Similar considerations can be made with F in place of $F_{t,x}$. An alternative way to reach the same conclusion is to look at the composition of functions

$$\varepsilon \longmapsto v_\varepsilon \longmapsto F_{t,x} v_\varepsilon,$$

where $\varepsilon \mapsto v_\varepsilon$ is of class C^{m-1} from $\mathbb{R}$ to C^p and $F_{t,x}$ is of class C^∞ from C^p to $L^p(\Omega)$, so that $\varepsilon \mapsto F_{t,x} v_\varepsilon$ is of class C^{m-1} from $\mathbb{R}$ to $L^p(\Omega)$, and the series expansion (23) follows by Taylor's theorem.

To obtain a series expansion for the difference $P_\varepsilon(t,x) - P(t,x)$ we need, however, to work pathwise, i.e. in $L^0(\Omega)$, essentially because it seems difficult to find a (reasonable) Banach space E such that $x \mapsto e^{-x}$ is Fréchet differentiable from $L^p(\Omega)$ to E, so that the chain rule could be applied to obtain a differentiability result for the map $\varepsilon \mapsto P_\varepsilon(t,x)$. We proceed instead as follows: Taylor's theorem yields

$$\begin{aligned} e^{-x} &= 1 + \sum_{j=1}^{m-1} (-1)^j \frac{x^j}{j!} + (-1)^m \int_0^1 \frac{(1-s)^{m-1}}{(m-1)!} e^{-sx} x^m \, ds \\ &=: 1 + J_{m-1}(x) + r_m(x) \end{aligned}$$

for every $x \in \mathbb{R}$, hence

$$\begin{aligned} P_\varepsilon(t,x) &= \exp\bigl(-F_{t,x} v_\varepsilon\bigr) = \exp\bigl(-F_{t,x} v\bigr) \exp\bigl(-F_{t,x}(v_\varepsilon - v)\bigr) \\ &= P(t,x)\Bigl(1 + J_{m-1}\bigl(F_{t,x} v_\varepsilon - F_{t,x} v\bigr) + r_m\bigl(F_{t,x} v_\varepsilon - F_{t,x} v\bigr)\Bigr), \end{aligned}$$

so that the relative pricing error can be written as

$$\eta_\varepsilon(t,x) := \frac{P_\varepsilon(t,x) - P(t,x)}{P(t,x)} = J_{m-1}\bigl(F_{t,x} v_\varepsilon - F_{t,x} v\bigr) + r_m\bigl(F_{t,x} v_\varepsilon - F_{t,x} v\bigr). \tag{24}$$

We are going to show that substituting the expansion of $F_{t,x} v_\varepsilon - F_{t,x} v$ provided by (23), we obtain a representation of $P_\varepsilon(t,x) - P(t,x)$ as a polynomial in ε of degree $m-1$ with coefficients in $L^0(\Omega)$, plus a remainder term of higher degree. To this purpose, we first prove a simple but useful auxiliary result. Given a ring A, the ring of polynomials in the variable x with coefficients in A will be denoted by $A[x]$. Moreover, we shall say that a function $f\colon [0,1] \to E$, with E a topological vector space, is infinitesimal of order higher than α at zero if $\lim_{x\to 0} f(x)/x^\alpha = 0$.

Lemma 5.11. *Let $P \in \mathbb{R}[x]$ and $Q \in L^p(\Omega)[x]$ be polynomials of degree n and m, respectively, and $r\colon [0,1] \to L^p(\Omega)$ a function that is infinitesimal of order higher than m at zero. Then there exists a polynomial $R \in L^{p/n}(\Omega)[x]$ of degree m and a function $s\colon [0,1] \to L^{p/n}(\Omega)$, infinitesimal of order higher than m at zero, such that*

$$P \circ (Q + r) = R + s.$$

Proof. Hölder's inequality implies that the functions $P \circ (Q + r)$ and $P \circ Q$ take values in $L^{p/n}(\Omega)$, thus also $P \circ Q[x] \in L^{p/n}(\Omega)[x]$. Moreover, for any integer $k \geq 1$ the binomial formula yields

$$(Q(x) + r(x))^k = \sum_{j=0}^{k} \binom{k}{j} Q(x)^{k-j} r(x)^j = Q(x)^k + \sum_{j=1}^{k} \binom{k}{j} Q(x)^{k-j} r(x)^j,$$

from which it follows, by Euclidean division, that there exist (unique) polynomials R and R_1 in $L^{p/n}(\Omega)[x]$, such that $P \circ Q(x) = R(x) + x^{m+1} R_1(x)$, with $\deg R \leq m$. Therefore the function s contains the term $x \mapsto x^{m+1} R_1(x)$, which is clearly infinitesimal of order higher than m at zero, and terms of the type $x \mapsto Q(x)^{k-j} r(x)^j$, with $1 \leq j \leq k \leq n$. One has

$$\frac{k-j}{p} + \frac{j}{p} = \frac{1}{p/k},$$

hence, by Hölder's inequality, writing $L^q := L^q(\Omega)$ for every $q > 0$ for notational convenience,

$$\left\| Q^{k-j} r^j \right\|_{L^{p/k}} \leq \left\| Q^{k-j} \right\|_{L^{p/(k-j)}} \left\| r^j \right\|_{L^{p/j}} \leq \left\| Q \right\|_{L^p}^{k-j} \left\| r \right\|_{L^p}^{j}.$$

Since $Q(x)$ tends to zero in $L^p(\Omega)$ as $x \to 0$, and r is infinitesimal of order higher than in m in $L^p(\Omega)$, one obtains that

$$\lim_{x\to 0} \frac{Q(x)^{k-j} r(x)^j}{x^m} = 0$$

in $L^{p/k}(\Omega)$, hence in $L^{p/n}(\Omega)$, thus showing that s is infinitesimal of order higher than m in $L^{p/n}(\Omega)$. □

Remark 5.12. One could prove that each term in the expansion of $P \circ (Q + r)$ belongs to $L^{p/n}(\Omega)$ also in a more direct and explicit way. In fact, if

$$Q(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_m x^m,$$

then

$$\begin{aligned}(Q(x)+r(x))^k &= \big(a_0+a_1x+a_2x^2+\cdots+a_mx^m+r(x)\big)^k\\ &= \sum\binom{k}{j_0,\ldots,j_{m+1}} a_0^{j_0}a_1^{j_1}x^{j_1}\cdots a_m^{j_m}x^{mj_m}r(x)^{j_{m+1}},\end{aligned}$$

where the sum is taken over all positive integers $j_0,\ldots j_{m+1}$ such that $j_0+\cdots+j_{m+1}=k$. This implies

$$\frac{j_0}{p}+\cdots+\frac{j_{m+1}}{p}=\frac{1}{p/k},$$

hence, by Hölder's inequality,

$$\begin{aligned}\big\|a_0^{j_0}\cdots a_m^{j_m}r(x)^{j_{m+1}}\big\|_{L^{p/k}} &\le \big\|a_0^{j_0}\big\|_{L^{j_0/p}}\cdots\big\|a_m^{j_m}\big\|_{L^{j_m/p}}\big\|r(x)^{j_{m+1}}\big\|_{L^{j_{m+1}/p}}\\ &= \|a_0\|_{L^p}^{j_0}\cdots\|a_m\|_{L^p}^{j_m}\|r(x)\|_{L^p}^{j_{m+1}}.\end{aligned}$$

This proves that every term in the expansion of $P(Q(x)+r(x))$ belongs to $L^{p/n}(\Omega)$. A completely analogous reasoning shows that $P\circ Q(x)\in L^{p/n}(\Omega)[x]$.

Let us rewrite (23) as

$$F_{t,x}v_\varepsilon - F_{t,x}v = Q_{m-1}(\varepsilon)+F_{t,x}R_{m,\varepsilon}, \tag{25}$$

where

$$Q_{m-1}(\varepsilon):=\sum_{k=1}^{m-1}\frac{1}{k!}F_{t,x}v_k\,\varepsilon^k$$

is a polynomial of degree $m-1$ with coefficients in $L^p(\Omega)$.

Proposition 5.13. *Let $p\in[1,\infty[$. Assume that $u_0\in L^p(\Omega;\mathsf{D}(A^{2m}))$ and*

$$\mathbb{E}\bigg(\sum_{k=1}^{\infty}\int_0^T\|\sigma_k(t,\cdot)\|^2_{\mathsf{D}(A_0^{2m})}\,dt\bigg)^p<\infty.$$

The relative pricing error $\eta_\varepsilon(t,x)$ at time $t\in[0,T]$ and time to maturity $x\in\mathbb{R}_+$ admits an expansion of the type

$$\eta_\varepsilon(t,x)=\sum_{j=1}^{m-1}\Phi_j\varepsilon^j+\rho(\varepsilon), \tag{26}$$

where the coefficients $\Phi_1,\ldots,\Phi_{m-1}$ belong to $L^{p/(m-1)}(\Omega)$, and the remainder $\rho\colon[0,1]\to L^0(\Omega)$ is such that

$$\lim_{\varepsilon\to0}\rho(\varepsilon)/\varepsilon^{m-1}=0$$

in probability.

Proof. Let $(t, x) \in [0, T] \times \mathbb{R}_+$. It follows from (24) and (25) that

$$\eta_\varepsilon(t, x) = J_{m-1}(Q_{m-1}(\varepsilon) + F_{t,x} R_{m,\varepsilon}) + r_m(Q_{m-1}(\varepsilon) + F_{t,x} R_{m,\varepsilon}), \tag{27}$$

where $R_{m,\varepsilon}/\varepsilon^{m-1}$ tends to zero in $\mathsf{C}^p(H(\mathbb{R}))$ as $\varepsilon \to 0$, hence, by Lemma 5.10, $F_{t,x} R_{m,\varepsilon}/\varepsilon^{m-1}$ tends to zero in $L^p(\Omega)$. Lemma 5.11 then yields

$$J_{m-1}(Q_{m-1}(\varepsilon) + F_{t,x} R_{m,\varepsilon}) = \sum_{j=1}^{m-1} \Phi_j \varepsilon^j + \rho_1(\varepsilon),$$

with $\Phi_j \in L^{p/(m-1)}(\Omega)$ for every $j = 1, \dots, m-1$ and $\rho_1 \colon [0, 1] \to L^{p/(m-1)}(\Omega)$ is infinitesimal of order higher than $m-1$ at zero. To conclude the proof, it remains to show that the second term on the right-hand side of (27) tends to zero in probability faster than ε^{m-1} as $\varepsilon \to 0$. Note that, for any $s \in [0, 1]$ and $x \in \mathbb{R}$, one has $-sx \le |s(-x)| = s|x| \le |x|$, hence

$$|r_m(x)| \le \frac{1}{(m-1)!}|x|^m \int_0^1 e^{-sx}\, ds \le \frac{1}{(m-1)!}|x|^m e^{|x|},$$

which in turn yields

$$r_m\big(Q_{m-1}(\varepsilon) + F_{t,x} R_{m,\varepsilon}\big) \lesssim_m \big|Q_{m-1}(\varepsilon) + F_{t,x} R_{m,\varepsilon}\big|^m \exp\big(|Q_{m-1}(\varepsilon) + F_{t,x} R_{m,\varepsilon}|\big).$$

Since the polynomial Q_{m-1} does not have term of order zero, a simple variant of Lemma 5.11 shows that

$$\lim_{\varepsilon \to 0} \frac{\big|Q_{m-1}(\varepsilon) + F_{t,x} R_{m,\varepsilon}\big|^m}{\varepsilon^{m-1}} = 0$$

in $L^{p/m}(\Omega)$, in particular in probability. Moreover, since $Q_{m-1}(\varepsilon) + F_{t,x} R_{m,\varepsilon}$ tends to zero in $L^p(\Omega)$, hence in probability, the continuous mapping theorem implies that

$$\lim_{\varepsilon \to 0} \exp\big(|Q_{m-1}(\varepsilon) + F_{t,x} R_{m,\varepsilon}|\big) = 1$$

in probability, thus also that

$$\lim_{\varepsilon \to 0} \frac{r_m\big(Q_{m-1}(\varepsilon) + F_{t,x} R_{m,\varepsilon}\big)}{\varepsilon^{m-1}} = 0$$

in probability. □

Remark 5.14. An expression for the coefficients Φ_j, $j = 1, \dots, m-1$, in (26) could be given in terms of the Faà di Bruno formula. The first three of them are

$$\begin{aligned}
\Phi_1 &= -F_{t,x}v_1, \\
\Phi_2 &= -\frac{1}{2}F_{t,x}v_2 + \frac{1}{2}(F_{t,x}v_1)^2, \\
\Phi_3 &= -\frac{1}{3!}F_{t,x}v_3 + \frac{1}{2}(F_{t,x}v_1)(F_{t,x}v_2) - \frac{1}{3!}(F_{t,x}v_1)^3.
\end{aligned}$$

We shall now discuss conditions under which the order of convergence to zero of the remainder term ρ can be established in topologies stronger than the topology of convergence in probability. In particular, we shall assume that

$$u_0 \in L^{mp}(\Omega; \mathsf{D}(A_0^{2m})), \qquad \mathbb{E}\Bigl(\sum_{k=1}^{\infty}\int_0^T \bigl\|\sigma_k(t,\cdot)\bigr\|^2_{\mathsf{D}(A_0^{2m})}\,dt\Bigr)^{mp} < \infty. \tag{28}$$

Then v and v_ε belong to $\mathsf{C}^{mp}(H(\mathbb{R}))$, which implies that $v_1, \dots, v_{m-1}$ and $R_{m,\varepsilon}$ in (23) belong to $\mathsf{C}^{mp}(H(\mathbb{R}))$. This in turn implies, in complete analogy to the proof of Proposition 5.13, that

$$J_{m-1}(F_{t,x}v_\varepsilon - F_{t,x}v) = \sum_{j=1}^{m-1}\Phi_j\varepsilon^j + \rho_1(\varepsilon),$$

with $\Phi_j \in L^p$ for every $j = 1, \dots, m-1$ and $\rho_1(\varepsilon)/\varepsilon^{m-1}$ converging to zero as $\varepsilon \to 0$. Estimating the second term on the right-hand side of (24) requires further assumptions. For instance, denoting the (Hölder) conjugate exponent to p by p', one has

$$\begin{aligned}
&\bigl\|r_m\bigl(F_{t,x}v_\varepsilon - F_{t,x}v\bigr)\bigr\|_{L^1(\Omega)} \\
&\qquad \le \int_0^1 \frac{(1-s)^{m-1}}{(m-1)!}\bigl\|\exp\bigl(-s(F_{t,x}v_\varepsilon - F_{t,x}v)\bigr)\bigr\|_{L^{p'}(\Omega)}\bigl\|(F_{t,x}v_\varepsilon - F_{t,x}v)^m\bigr\|_{L^p(\Omega)}\,ds,
\end{aligned}$$

where, as already seen,

$$(F_{t,x}v_\varepsilon - F_{t,x}v)^m = \Bigl(\sum_{k=1}^{m-1}\frac{1}{k!}F_{t,x}v_k\,\varepsilon^k + F_{t,x}R_{m,\varepsilon}\Bigr)^m$$

tends to zero in $L^p(\Omega)$ faster than ε^{m-1} as $\varepsilon \to 0$. Another application of Hölder's inequality yields

$$\begin{aligned}
\bigl\|\exp\bigl(-s(F_{t,x}v_\varepsilon - F_{t,x}v)\bigr)\bigr\|_{L^{p'}(\Omega)} &= \Bigl(\mathbb{E}\exp\bigl(-p's(F_{t,x}v_\varepsilon - F_{t,x}v)\bigr)\Bigr)^{1/p'} \\
&\le \Bigl(\mathbb{E}\exp\bigl(-p'(F_{t,x}v_\varepsilon - F_{t,x}v)\bigr)\Bigr)^{1/p'} \qquad \forall s \in [0,1],
\end{aligned}$$

from which it follows that if $\mathbb{E}\exp\bigl(-p'(F_{t,x}v_\varepsilon - F_{t,x}v)\bigr)$ is bounded for ε in a (right) neighborhood of zero, then

$$\lim_{\varepsilon\to 0} \frac{\bigl\|r_m\bigl(F_{t,x}v_\varepsilon - F_{t,x}v\bigr)\bigr\|_{L^1(\Omega)}}{\varepsilon^{m-1}} = 0.$$

Such uniform bounds of exponential moments, although quite hard to establish in general, can be obtained for the Gaussian HJM model, i.e. assuming that the volatility coefficients (σ_k) are deterministic.

Proposition 5.15. *Assume that u_0 and (σ_k) are non-random with $u_0 \in \mathsf{D}(A_0^{2m})$ and*

$$\sum_{k=1}^{\infty} \int_0^T \bigl\|\sigma_k(t,\cdot)\bigr\|^2_{\mathsf{D}(A_0^{2m})} < \infty.$$

Then, for any $(t,x) \in [0,T] \times \mathbb{R}_+$, the relative pricing error $\eta_\varepsilon(t,x)$ admits an expansion of the type

$$\eta_\varepsilon(t,x) = \sum_{j=1}^{m-1} \Phi_j \varepsilon^j + \rho(\varepsilon),$$

where the coefficients $\Phi_1, \dots, \Phi_{m-1}$ belong to $L^p(\Omega)$, and the remainder $\rho\colon [0,1] \to L^1(\Omega)$ is such that

$$\lim_{\varepsilon\to 0} \frac{\rho(\varepsilon)}{\varepsilon^{m-1}} = 0$$

in $L^1(\Omega)$.

Proof. Note that the assumption (28) is trivially satisfied. We are going to prove that, for any $p \geq 1$, $\mathbb{E}\exp\bigl(-p'(F_{t,x}v_\varepsilon - F_{t,x}v)\bigr)$ is bounded with respect to ε sufficiently small, where, as before, p' is the Hölder conjugate of p. To this purpose, let us first show that $F_{t,x}v$ and $F_{t,x}v_\varepsilon$ are Gaussian random variables. The argument being the same, we consider only $F_{t,x}v$, for which we can write

$$F_{t,x}v = F^1_{t,x}v + F^2_{t,x}v, \qquad F^1_{t,x}v := \int_0^x v(t,y)\,dy, \quad F^2_{t,x}v := \int_0^t v(s,0)\,ds.$$

Since $v(t)$ is Gaussian with values in $H(\mathbb{R})$, and the map $h \mapsto \int_0^x h(y)\,ds$ is linear and continuous from $H(\mathbb{R}_+)$ to $\mathbb{R}$, $F^1_{t,x}v$ is a Gaussian random variable. Moreover, as $v(s,0) = \delta v(s)$, with δ the Dirac measure at zero, which is a linear and continuous map from $H(\mathbb{R})$ to $\mathbb{R}$, we deduce that $s \mapsto v(s,0)$ is a mean-square continuous $\mathbb{R}$-valued Gaussian process. Therefore, elementary properties of Gaussian processes (see, e.g., [8, p. 124]) imply that $F^2_{t,x}v$ is a Gaussian random variable. The Cauchy-Schwarz inequality yields

$$\mathbb{E}\exp\bigl(-p'(F_{t,x}v_\varepsilon - F_{t,x}v)\bigr) \le \Bigl(\mathbb{E}\exp\bigl(-2p'F_{t,x}v_\varepsilon\bigr)\Bigr)^{1/2} \Bigl(\mathbb{E}\exp\bigl(2p'F_{t,x}v\bigr)\Bigr)^{1/2},$$

where both terms on the right-hand side are finite because Gaussian random variables admit finite exponential moments. Moreover, $F_{t,x}v_\varepsilon$ converges to $F_{t,x}v$ in $L^p(\Omega)$ as $\varepsilon \to 0$, hence also in law. Setting

$$m := \mathbb{E}\, F_{t,x}v, \qquad m_\varepsilon := \mathbb{E}\, F_{t,x}v_\varepsilon,$$
$$\varsigma^2 := \mathbb{E}(F_{t,x}v - m)^2, \qquad \varsigma_\varepsilon^2 := \mathbb{E}(F_{t,x}v_\varepsilon - m_\varepsilon)^2,$$

Gaussianity implies that $m_\varepsilon \to m$ and $\varsigma_\varepsilon \to \varsigma$ as $\varepsilon \to 0$. Therefore the well-known expression for the moment generating function of a Gaussian law implies

$$\mathbb{E}\exp\bigl(-2p'F_{t,x}v_\varepsilon\bigr) \le \exp\Bigl(2(p')^2\varsigma_\varepsilon^2 - 2p'm_\varepsilon\Bigr),$$

where the right-hand side is bounded for ε in a right neighborhood of zero. □

Remarks 5.16. (i) Let $L^{\infty-}(\Omega)$ be the Fréchet space defined as the intersection of all $L^p(\Omega)$ spaces with $p \in [1, \infty[$. Since assumption (28) is satisfied for every $p > 0$, a simple variation of the proof shows that $\Phi_1, \dots, \Phi_{m-1}$ belong to $L^{\infty-}(\Omega)$, as well as that $\rho(\varepsilon)/\varepsilon^{m-1}$ converges to zero in $L^{\infty-}(\Omega)$ as $\varepsilon \to 0$.
(ii) One could also try to prove directly that the law of v is Gaussian on $C([0,T]; H(\mathbb{R}))$, from which it would follow directly that $F_{t,x}v$ is also Gaussian thereon, because $F_{t,x}$ is linear and bounded on that space.

An alternative estimate on the pricing error can be obtained under a positivity assumption.

Proposition 5.17. *Assume that u_0 and (σ_k) satisfy (28). If $u \ge 0$ and $u_\varepsilon \ge 0$, then, for any $(t,x) \in [0,T] \times \mathbb{R}_+$, there exist $\Phi_j \in L^p(\Omega)$, $j = 1, \dots, m-1$, and $\rho_1,\ \rho_2 \colon [0,1] \to L^p(\Omega)$ such that*

$$P_\varepsilon(t,x) - P(t,x) = P(t,x)\sum_{j=1}^{m-1}\Phi_j\varepsilon^j + P(t,x)\,\rho_1(\varepsilon) + \rho_2(\varepsilon)$$

with $\lim_{\varepsilon\to 0}\rho_1(\varepsilon)/\varepsilon^{m-1} = \lim_{\varepsilon\to 0}\rho_2(\varepsilon)/\varepsilon^{m-1} = 0$ in $L^p(\Omega)$.

Proof. Let $(t,x) \in [0,T] \times \mathbb{R}_+$. It follows from (24) that

$$P_\varepsilon(t,x) - P(t,x) = P(t,x)J_{m-1}\bigl(F_{t,x}v_\varepsilon - F_{t,x}v\bigr) + P(t,x)\,r_m\bigl(F_{t,x}v_\varepsilon - F_{t,x}v\bigr),$$

where, as in the proof of Proposition 5.13,

$$J_{m-1}\bigl(F_{t,x}v_\varepsilon - F_{t,x}v\bigr) = \sum_{j=1}^{m-1}\Phi_j\varepsilon^j + \rho_1(\varepsilon),$$

with $\lim_{\varepsilon\to 0}\rho_1(\varepsilon)/\varepsilon^{m-1} = 0$ in $L^p(\Omega)$. Moreover, writing

$$
\begin{aligned}
&P(t,x)\, r_m\big(F_{t,x}v_\varepsilon - F_{t,x}v\big)\\
&\qquad = (-1)^m \int_0^1 \frac{(1-s)^{m-1}}{(m-1)!} \exp\big(-sF_{t,x}v_\varepsilon - (1-s)F_{t,x}v\big)(F_{t,x}v_\varepsilon - F_{t,x}v)^m\, ds,
\end{aligned}
$$

taking into account that $v \geq 0$ and $v_\varepsilon \geq 0$, and that $F_{t,x}$ is positivity preserving, one has

$$
\exp\big(-sF_{t,x}v_\varepsilon - (1-s)F_{t,x}v\big) \leq 1 \qquad \forall s \in [0,1].
$$

This in turn implies

$$
\big\| P(t,x)\, r_m\big(F_{t,x}v_\varepsilon - F_{t,x}v\big)\big\|_{L^p(\Omega)} \lesssim \big\|(F_{t,x}v_\varepsilon - F_{t,x}v)^m\big\|_{L^p(\Omega)},
$$

where the right hand side converges to zero faster than ε^{m-1}. □

The positivity of forward rates is a natural assumption from the financial perspective, which is, however, not guaranteed by the general HJM model (an example is the Gaussian HJM model). The positivity of forward rates, seen as mild solutions to the Musiela SPDE, is discussed, e.g., in [11,15], and in [17,21] in the more general context of positivity of mild solutions to stochastic evolution equations.

Without knowing a priori that u and u_ε are positive, expansions of the pricing errors that hold on the set

$$
A_\varepsilon := \big\{\omega \in \Omega :\ u(\omega,\cdot),\ u_\varepsilon(\omega,\cdot) \geq 0 \quad \text{on } [0,t]\times[0,x]\big\}
$$

can be obtained. In fact, as in the proof of the previous proposition, one has

$$
P_\varepsilon(t,x) - P(t,x) = P(t,x)\sum_{j=1}^{m-1} \Phi_j \varepsilon^j + P(t,x)\rho_1(\varepsilon) + P(t,x)\, r_m\big(F_{t,x}v_\varepsilon - F_{t,x}v\big),
$$

hence, multiplying both sides by 1_{A_ε}, setting

$$
\widetilde{\Phi}_j := 1_{A_\varepsilon}\Phi_j, \qquad \widetilde{\rho}_1(\varepsilon) := 1_{A_\varepsilon}\rho_1(\varepsilon), \quad \widetilde{\rho}_2(\varepsilon) := 1_{A_\varepsilon}\rho_2(\varepsilon),
$$

and noting that

$$
\widetilde{\rho}_2(\varepsilon) := \big\|1_{A_\varepsilon} P(t,x)\, r_m\big(F_{t,x}v_\varepsilon - F_{t,x}v\big)\big\|_{L^p(\Omega)} \lesssim \big\|(F_{t,x}v_\varepsilon - F_{t,x}v)^m\big\|_{L^p(\Omega)},
$$

where the right-hand side converges to zero faster than ε^{m-1}, one finally concludes that

$$
1_{A_\varepsilon}\big(P_\varepsilon(t,x) - P(t,x)\big) = P(t,x)\sum_{j=1}^{m-1} \widetilde{\Phi}_j \varepsilon^j + P(t,x)\widetilde{\rho}_1(\varepsilon) + \widetilde{\rho}_2(\varepsilon)
$$

where $\widetilde{\Phi}_j$ are random variables in $L^p(\Omega)$ and $\widetilde{\rho}_1$, $\widetilde{\rho}_2$ converge to zero in $L^p(\Omega)$ faster than ε^{m-1}.

Data availability

No data was used for the research described in the article.

References

[1] M. Barski, J. Zabczyk, Heath-Jarrow-Morton-Musiela equation with Lévy perturbation, J. Differ. Equ. 253 (2012) 2657–2697.
[2] V.I. Bogachev, O.G. Smolyanov, Topological Vector Spaces and Their Applications, Springer, Cham, 2017.
[3] H. Brézis, Functional Analysis, Sobolev Spaces and Partial Differential Equations, Springer, New York, 2011.
[4] Z. Brzeźniak, T. Kok, Stochastic evolution equations in Banach spaces and applications to the Heath-Jarrow-Morton-Musiela equations, Finance Stoch. 22 (2018) 959–1006.
[5] P.L. Butzer, H. Berens, Semi-Groups of Operators and Approximation, Springer, New York, 1967.
[6] R.A. Carmona, M.R. Tehranchi, Interest Rate Models: an Infinite Dimensional Stochastic Analysis Perspective, Springer, Berlin, 2006.
[7] R. Cont, Modeling term structure dynamics: an infinite dimensional approach, Int. J. Theor. Appl. Finance 8 (2005) 357–380.
[8] G. Da Prato, J. Zabczyk, Stochastic Equations in Infinite Dimensions, second ed., Cambridge University Press, Cambridge, 2014.
[9] K.J. Engel, R. Nagel, One-Parameter Semigroups for Linear Evolution Equations, Springer, New York, 2000.
[10] D. Filipović, Consistency Problems for Heath-Jarrow-Morton Interest Rate Models, Lecture Notes in Mathematics, vol. 1760, Springer, Berlin, 2001.
[11] D. Filipović, S. Tappe, J. Teichmann, Term structure models driven by Wiener processes and Poisson measures: existence and positivity, SIAM J. Financ. Math. 1 (2010) 523–554.
[12] D.C. Heath, R.A. Jarrow, A. Morton, Bond pricing and the term structure of interest rates: a new methodology for contingent claims valuation, Econometrica 60 (1992) 77–105.
[13] T. Kato, Perturbation Theory for Linear Operators, Springer, Berlin, 1980.
[14] M. Kunze, J. van Neerven, Continuous dependence on the coefficients and global existence for stochastic reaction diffusion equations, J. Differ. Equ. 253 (2012) 1036–1068.
[15] S. Kusuoka, Term structure and SPDE, in: Advances in Mathematical Economics, vol. 2, Springer, Tokyo, 2000, pp. 67–85.
[16] S. Lang, Real and Functional Analysis, third ed., Springer, New York, 1993.
[17] C. Marinelli, Positivity of mild solution to stochastic evolution equations with an application to forward rates, https://arxiv.org/abs/1912.12472.
[18] C. Marinelli, Local well-posedness of Musiela's SPDE with Lévy noise, Math. Finance 20 (2010) 341–363.
[19] C. Marinelli, On well-posedness of semilinear stochastic evolution equations on L_p spaces, SIAM J. Math. Anal. 50 (2018) 2111–2143.
[20] C. Marinelli, M. Röckner, On the maximal inequalities of Burkholder, Davis and Gundy, Expo. Math. 34 (2016) 1–26.
[21] C. Marinelli, L. Scarpa, On the positivity of local mild solutions to stochastic evolution equations, in: S. Ugolini, et al. (Eds.), Geometry and Invariance in Stochastic Dynamics, in: Springer Proceedings in Mathematics & Statistics, vol. 378, Springer, Cham, 2022, pp. 231–245, https://arxiv.org/abs/1912.13259.
[22] C. Marinelli, G. Ziglio, L. Di Persio, Approximation and convergence of solutions to semilinear stochastic evolution equations with jumps, J. Funct. Anal. 264 (2013) 2784–2816.
[23] M. Musiela, Stochastic PDEs and term structure models, J. Int. Financ. (1993).
[24] M. Musiela, M. Rutkowski, Martingale Methods in Financial Modelling, second ed., Springer, Berlin, 2005.
[25] S. Peszat, J. Zabczyk, Stochastic Partial Differential Equations with Lévy Noise, Cambridge University Press, Cambridge, 2007.
[26] E.M. Stein, Singular Integrals and Differentiability Properties of Functions, Princeton University Press, Princeton, N.J., 1970.
[27] M. Tehranchi, A note on invariant measures for HJM models, Finance Stoch. 9 (2005) 389–398.
[28] T. Vargiolu, Invariant measures for the Musiela equation with deterministic diffusion term, Finance Stoch. 3 (1999) 483–492.
[29] I.I. Vrabie, C_0-Semigroups and Applications, North-Holland, Amsterdam, 2003.