Politechnika Wrocławska

FIELD OF SCIENCE: NATURAL SCIENCES

DISCIPLINE OF SCIENCE: MATHEMATICS

DOCTORAL DISSERTATION

Hardy–Stein identity and Littlewood–Paley theory for non-local operators

Mgr inż. Michał Gutowski

Supervisor:
dr hab. inż. Mateusz Kwaśnicki

Keywords: Hardy–Stein identity, non-local operator, Sobolev–Bregman form, Dirichlet form, Beurling–Deny formula, Littlewood–Paley theory, square function

WROCŁAW, NOVEMBER 2024

Podziękowania

Na wstępie chciałbym podziękować mojemu promotorowi, dr. hab. inż. Mateuszowi Kwaśnickiemu, za jego błyskotliwe uwagi, bardzo duże zaangażowanie oraz pomoc na wielu etapach mojego kształcenia na studiach doktoranckich. Dziękuję mu za jego rozległą wiedzę merytoryczną, którą się ze mną podzielił oraz za współpracę przez te lata. Jestem także wdzięczny za jego wyrozumiałość i cierpliwość, zwłaszcza na końcowym etapie powstawania niniejszej rozprawy. Dziękuję także za świetnie przeprowjozony przez niego kurs analizy matematycznej na pierwszym roku moich studiów, który utwierdził mnie w przekonaniu, że matematyka teoretyczna jest tym, czym chciałbym się w życiu zajmować. Chciałbym także podziękować za jego kurs z teorii potencjałów procesów Markowa na który miałem przyjemność uczęszczać na studiach magisterskich.

Dziękuję także dr. hab. inż. Jackowi Małeckiemu za to, że dostrzegł moje zamiłowanie do matematyki na moich studiach na mechatronice i namówił do rozpoczęcia studiów na drugim kierunku – matematyce teoretycznej. Bez tego niniejsza praca nie mogłaby powstać. Chciałbym podziękować wielu dydaktykom z Wydziału Matematyki, których nie jestem w stanie wszystkich wymienić, a którzy w zaskakująco sympatycznej atmosferze przekazali mi ogrom wiedzy potrzebnej do mojej pracy naukowej.

Dziękuję również prof. dr. hab. inż. Krzysztofowi Bogdanowi za współpracę nad projektem pt. „Regularność i asymptotyka dla wielowymiarowych ewolucji markowskich" w ramach grantu NCN 2017/27/B/ST1/01339. Jestem wdzięczny również za jego pomocne uwagi na dalszych etapach mojego kształcenia. Dziękuję jemu, a także prof. dr. hab. Katarzynie Pietruskiej-Pałubie, za bardzo miłą jak i owocną współpracę nad projektem pt. „Nieliniowe formy całkowe związane ze skokowymi procesami Markowa" w ramach grantu NCN 2018/31/B/ST1/03818. Bardzo miło wspominam tę współpracę i jestem wdzięczny za ich merytoryczne wsparcie.

Chciałbym podziękować osobom, które tworzyły moje życie pozanaukowe. Dziękuję przede wszystkim mojej dziewczynie, Patrycji Burzyńskiej, która wspierała mnie w trakcie moich studiów doktoranckich, podtrzymywała na duchu i motywowała do działania. Zawsze we mnie wierzyła, za co jestem jej ogromnie wdzięczny. Dziękuję mojej mamie, Beacie, za troskę i dopingowanie mnie w trakcie moich studiów. Mama również zawsze we mnie wierzyła i przejmowała się wszelkimi trudnościami, jakie napotykałem, za co również chciałbym szczerze podziękować. Chciałbym podziękować także moim przyjaciółkom, Klaudii Pilichowski i Aleksandrze Szepielak, z którymi spędziłem wiele lat mojego życia studenckiego. Bez wątpienia nasza kilkunastoletnia już znajomość w znacznej mierze ukształtowała moją osobowość.

Dziękuję wielu wartościowym osobom, które poznałem na studiach, zarówno matematycznych, jak i mechatronicznych, a których wszystkich nie jestem w stanie wymienić. W szczególności chciałbym wymienić z nazwiska członków naszej grupy znajomych „Gramy w H3": Macieja Błaszczyka, Michała Cichockiego i Macieja Kowalskiego. Znajomość z nimi uważam za wartościową i znacznie wzbogacającą moje życie towarzyskie.

Na koniec chciałbym podziękować Politechnice Wrocławskiej za przekazane mi wykształcenie. W szczególności chciałbym podziękować dydaktykom ze Studium Języków Obcych, dzięki którym mogłem uczyć się umiłowanego przeze mnie języka hiszpańskiego. ¡Muchas Gracias!

Znaczna część badań przedstawionych w tej rozprawie była realizowana w ramach wyżej wymienionego grantu NCN 2018/31/B/ST1/03818.

Contents

Streszczenie

Tożsamość Hardy'ego–Steina odzwierciedla rozpraszanie się energii danej funkcji w czasie. Przez energię rozumiemy tutaj p-tą normę tej funkcji na przestrzeni L^p. Oryginalnie, tożsamość Hardy'ego–Steina została udowodniona przez Steina w jego książce [99, pp. 86–88] i dotyczyła operatora Laplace'a. Do dowodu użyto reguły łańcuchowej dla laplasjanu. Badania nad przypadkiem nielokalnej tożsamości Hardy'ego–Steina są stosunkowo nowe, a jako przykład można tu podać pracę Bañuelosa, Bogdana i Luksa [5], gdzie wyprowadzono tę tożsamość dla operatorów Lévy'ego na przestrzeni euklidesowej. Badanie tożsamości Hardy'ego–Steina może przebiegać przy użyciu metod zarówno analitycznych jak i probabilistycznych. Autorzy [5] użyli pierwszego z nich. Jako przypadek użycia w tym kontekście metod probabilistycznych, możemy wymienić pracę Bañuelosa i Kima [6], gdzie dowód opiera się o wzór Itô.

Żeby uogólnić uzyskane do tej pory wyniki, autor niniejszej rozprawy oparł się na teorii form Dirichleta. Przez regularną formę Dirichleta rozumiemy domkniętą, nieujemną, symetryczną, markowską oraz posiadającą rdzeń formę dwuliniową $\mathcal{E}(u,v)$ o dziedzinie $\mathcal{D}(\mathcal{E})$, która jest gęsta w przestrzeni L^2. Warto tutaj wspomnieć, że istnieje bijekcja między klasą regularnych form Dirichleta a pewną klasą symetrycznych procesów Markowa nazywaną symetrycznymi procesami Hunta.

Jednym z kluczowych twierdzeń teorii form Dirichleta jest wzór Beurlinga–Deny, który mówi, że każda regularna forma Dirichleta ma jednoznaczny rozkład na trzy składniki: składnik silnie lokalny, składnik skokowy z miarą skoków J i składnik związany z zabijaniem z miarą zabijania k. Składniki te odzwierciedlają kolejno ciągłe, skokowe i związane z zabijaniem zachowanie się procesu Hunta powiązanego z formą Dirichleta $\mathcal{E}$. Wzór ten brzmi następująco:

$$\mathcal{E}(u,v) = \mathcal{E}^c(u,v) + \frac{1}{2}\iint\limits_{E\times E\setminus \mathrm{diag}} (u(y)-u(x))(v(y)-v(x))\, J(\mathrm{d}x,\mathrm{d}y) + \int\limits_E u(x)v(x)\, k(\mathrm{d}x)$$

i zachodzi dla wszystkich ciągłych (w ogólności kwaziciągłych) funkcji u i v z dziedziny $\mathcal{D}(\mathcal{E})$.

W niniejszej rozprawie użyto tak zwanej formy Sobolewa–Bregmana (w skrócie p-formy), która jest uogólnieniem formy Dirichleta na przestrzeń L^p, gdzie $1 < p < \infty$. Pojęcie to pojawiało się wcześniej w literaturze jedynie w kontekście czysto skokowych form Dirichleta; patrz [16, 14]. Autor rozprawy zaproponował ogólną konstrukcję p-formy używając form aproksymujących formę Dirichleta.

Jednym z głównych wyników niniejszej rozprawy jest wzór Beurlinga–Deny dla form Sobolewa–Bregmana. Najważniejszą zaletą tego wyniku jest to, iż wzór ten działa w pełnej ogólności – jest spełniony dla każdej regularnej formy Dirichleta, bez żadnych dodatkowych założeń. Przy użyciu tego wyniku, została udowodniona następująca ogólna tożsamość Hardy'ego–Steina dla dowolnej regularnej formy Dirichleta:

$$\int\limits_E |f(x)|^p\, m(\mathrm{d}x) - \lim_{T\to+\infty} \|P_T f\|_p^p = \frac{4(p-1)}{p} \int\limits_0^{+\infty} \mathcal{E}^c[(P_t f)^{\langle p/2\rangle}]\, \mathrm{d}t$$

$$+ \int\limits_0^{+\infty} \iint\limits_{E\times E\setminus \mathrm{diag}} F_p(P_t f(x), P_t f(y))\, J(\mathrm{d}x,\mathrm{d}y)\mathrm{d}t + p\int\limits_0^{+\infty}\int\limits_E |P_t f(x)|^p\, k(\mathrm{d}x)\mathrm{d}t, \quad f \in L^p(m).$$

Tutaj, funkcja $F_p(a,b) = |b|^p - |a|^p - pa^{\langle p-1\rangle}(b-a)$ to tak zwana *dywergencja Bregmana*, natomiast $a^{\langle\gamma\rangle} = |a|^\gamma \operatorname{sgn}(a)$ oznacza francuską potęgę. Przez operatory P_t rozumiemy półgrupę związaną z formą Dirichleta.

Kolejnym wynikiem rozprawy jest dowód spolaryzowanej tożsamości Hardy'ego–Steina dla czysto skokowych form Dirichleta oraz zbadanie własności spolaryzowanego odpowiednika p-formy.

Ostatni rozdział niniejszej rozprawy zawiera zastosowanie podejścia z prac [5, 63] i przy użyciu uzyskanych wcześniej wyników, otrzymanie oszacowań p-tych norm dla funkcji kwadratowych Littlewooda–Paleya. Aby zaadaptować podejście z wymienionych wyżej prac do bardziej ogólnego przypadku zastosowano miary Revuza i ich relację z funkcjonałami addytywnymi. Dodatkowo, w niniejszej pracy przedstawiono przykłady operatorów, dla których nie da się uzyskać oszacowania p-tych norm oraz wskazano nienaprawialny błąd z artykułu [63].

Abstract

The Hardy–Stein identity captures the disintegration of the energy of a function from the L^p-space. By energy, we mean the p-norm of a function. Originally, the Hardy–Stein identity was derived by Stein in his book [99, pp. 86–88] for the Laplace operator, employing the chain rule for that operator. The study of a non-local instance of the Hardy–Stein identity is quite new. We may refer to Bañuelos, Bogdan, and Luks in [5] for the identity for Lévy-type operators on a Euclidean space. The Hardy–Stein identity can be investigated using both analytical and probabilistic approaches. The authors of [5] utilized the former. For the probabilistic approach, we may refer to Bañuelos and Kim [6], where Itô's formula was employed.

The author employs the theory of regular Dirichlet forms to generalize the previous results. A regular Dirichlet form is a closed, non-negative, symmetric, Markovian bilinear form $\mathcal{E}(u,v)$, with domain $\mathcal{D}(\mathcal{E})$ dense in L^2-space, which possesses a core. The class of regular Dirichlet forms is in a one-to-one correspondence with a specific class of symmetric Markov processes known as symmetric Hunt processes.

One of the main results from the theory of Dirichlet forms is the Beurling–Deny formula, which provides the unique decomposition of a regular Dirichlet form $\mathcal{E}$ into three parts: the strongly local part $\mathcal{E}^c$, the jumping part with the jumping measure J, and the killing part with the killing measure k. These parts respectively describe the diffusive, jumping, and the killing behavior of the Hunt process associated with $\mathcal{E}$. This formula reads

$$\mathcal{E}(u,v) = \mathcal{E}^c(u,v) + \frac{1}{2}\iint\limits_{E\times E\setminus\mathrm{diag}} (u(y)-u(x))(v(y)-v(x))\, J(\mathrm{d}x,\mathrm{d}y) + \int\limits_E u(x)v(x)\, k(\mathrm{d}x)$$

for all continuous (in general quasi-continuous) functions u and v from the domain $\mathcal{D}(\mathcal{E})$.

In the present dissertation, the Sobolev–Bregman form (or shortly p-form) is employed as an extension of the Dirichlet form to L^p for $1 < p < \infty$. This notion has appeared in the literature defined only for pure-jump Dirichlet forms; see [16, 14]. The author proposes a general definition of the p-form using the approximation form of the Dirichlet form.

One of the main results of the dissertation is the Beurling–Deny formula for the Sobolev–Bregman form. The main advantage of this result is its validity for any regular Dirichlet form without requiring additional assumptions. Utilizing this result, the following general Hardy–Stein identity for all regular Dirichlet forms is proved:

$$\int\limits_E |f(x)|^p\, m(\mathrm{d}x) - \lim_{T\to+\infty} \|P_T f\|_p^p = \frac{4(p-1)}{p}\int\limits_0^{+\infty} \mathcal{E}^c[(P_t f)^{\langle p/2\rangle}]\,\mathrm{d}t$$
$$+ \int\limits_0^{+\infty}\iint\limits_{E\times E\setminus\mathrm{diag}} F_p(P_t f(x), P_t f(y))\, J(\mathrm{d}x,\mathrm{d}y)\mathrm{d}t + p\int\limits_0^{+\infty}\int\limits_E |P_t f(x)|^p\, k(\mathrm{d}x)\mathrm{d}t, \quad f\in L^p(m).$$

Here, the function $F_p(a,b) = |b|^p - |a|^p - pa^{\langle p-1\rangle}(b-a)$ is the so-called *Bregman divergence*, where $a^{\langle\gamma\rangle} = |a|^\gamma \operatorname{sgn}(a)$ is the *French power*. The family $(P_t)_{t\geq 0}$ of operators is the semigroup associated with the Dirichlet form.

Another result of the dissertation is the proof of the polarized analog of the Hardy–Stein identity for pure-jump Dirichlet forms. To achieve this, the author study the polarized version of p-form.

In the last chapter of the dissertation, the author follows [5, 63] and utilizes the obtained results to derive L^p-bounds of non-local Littlewood–Paley square functions. To adapt the

approach from the mentioned papers to a more general setting, the Revuz correspondence is utilized. Additionally, the dissertation introduces new examples where Littlewood–Paley estimates do not hold and identifies an irreparable error in the paper [63].

Chapter 1

Introduction

Non-local operators have attracted considerable attention and have been the subject of extensive research in recent years. Throughout this work, the operator will be understood as a certain mapping A acting on an appropriate space of real-valued functions. We consider functions on a topological space E, often the Euclidean space $\mathbb{R}^d$. The name *non-local* refers to the following property of such operators: the value of Au at a point x depends on the values of the function u at points that are located far from the point x. Typical examples of such operators are integral operators. This is in contrast to *local operators*, where the value of $Au(x)$ depends only on the values of u in an arbitrarily small neighborhood of x. Local operators are mainly exemplified by differential operators. The classical example is the Laplace operator Δ. On the other hand, the commonly researched example of a non-local operator is the fractional Laplacian $\Delta^{\alpha/2}$. It may be understood as the operator $-(-\Delta)^{\alpha/2}$ defined by the spectral theory, where Δ is the classical Laplace operator; various equivalent definitions of this operator have been presented by Kwaśnicki in [61]. The fractional Laplacian belongs to the broader class of so-called Lévy-type operators.

If an operator $-A$ is non-negative and self-adjoint on L^2-space, it generates a semigroup of operators $(P_t)_{t\geq 0}$. The connection between this semigroup and such a *generator* can be described by $P_t = e^{tA}$ in the sense of spectral theory.

If the operators P_t are Markovian, then they are transition operators of a Markov process $(X_t)_{t\geq 0}$ with the state space E. This relationship may be understood by the following equality: $P_t f(x) = \mathbb{E}_x f(X_t)$, where $\mathbb{E}_x$ is the conditional expectation under the condition that the process starts at a point x. For instance, the classical Laplace operator Δ is connected with the Brownian motion $(B_t)_{t\geq 0}$ on $\mathbb{R}^d$. As another example, we may mention the connection between the fractional Laplacian $\Delta^{\alpha/2}$ and the rotationally invariant α-stable Lévy process $(X_t)_{t\geq 0}$. Generally, local operators are associated with processes having continuous paths, whereas non-local operators correspond to jump processes. The intensity of jumps of a Markov process can be described by a certain measure J called *jumping measure*.

The author's research focused exclusively on symmetric Markov processes, which are processes in which the probability of transition from a point x to a point y is equal to the probability of transition from y to x. In particular, the jumping measure is symmetric in this case: $J(\mathrm{d}x, \mathrm{d}y) = J(\mathrm{d}y, \mathrm{d}x)$. From the analytical perspective, this symmetry is reflected in the operators P_t being self-adjoint on the L^2-space: $\langle P_t f, g\rangle = \langle f, P_t g\rangle$, $f, g \in L^2(E, m)$. Here, $\langle f, g\rangle = \int_E fg \,\mathrm{d}m$ denotes the inner product of the Hilbert space $L^2(E, m)$. For convenience, we will write later $L^p(m) = L^p(E, m)$.

1.1 Hardy–Stein identity

The main focus of this doctoral dissertation is the Hardy–Stein identity for non-local operators. The typical instance of the identity associated with the purely non-local case reads

$$\int_E |f(x)|^p \, m(\mathrm{d}x) = \int_0^{+\infty} \iint_{E\times E\setminus \mathrm{diag}} F_p(P_t f(x), P_t f(y)) \, J(\mathrm{d}x, \mathrm{d}y)\mathrm{d}t \tag{1.1}$$

for any function f in the Banach space $L^p(m)$, $1 < p < \infty$. Here, the function

$$F_p(a,b) = |b|^p - |a|^p - pa^{\langle p-1\rangle}(b-a)$$

is the so-called *Bregman divergence*, where $a^{\langle\gamma\rangle} = |a|^\gamma \operatorname{sgn}(a)$ is the *French power*. The identity in the form of (1.1) was first proposed by Bañuelos, Bogdan, and Luks in [5] for a certain class of Lévy-type operators (on $E = \mathbb{R}^d$), where the measure J in this context is a translation-invariant Lévy measure ν: $J(\mathrm{d}x, \mathrm{d}y) = \nu(\mathrm{d}y - x)\mathrm{d}x$. The above instance of the Hardy–Stein identity was the starting point of the author's investigation.

In light of the relationship between analysis and probability mentioned above, the Hardy–Stein identity can be investigated using both analytical and probabilistic approaches. The earlier work in [5] utilized the former, while Bañuelos and Kim in [6] later extended these results to non-symmetric Lévy processes, employing Itô's formula.

Originally, the Hardy–Stein identity was derived by Stein in his book [99, pp. 86–88] for the Laplace operator Δ. The approach there relies essentially on the chain rule: $\Delta u^p = p(p-1)u^{p-2}|\nabla u|^2 + pu^{p-1}\Delta u$. This method can be adapted to a broader class of local operators, for which the chain rule is available. This strategy was discussed by Bakry in Section 1.3 of [3]. It is noteworthy that the chain rule is limited to operators associated with Markov processes having continuous trajectories. Therefore, this technique cannot be applied to non-local operators.

The Hardy–Stein identity has applications in Littlewood–Paley theory which is described in Section 1.4 and more broadly in Chapter 8.

The aim of the author's investigation was to generalize the above non-local Hardy–Stein identity to a broader class of operators. The result of this work is the following.

Theorem 1.1. *Let $1 < p < \infty$. Under some standard assumptions about the semigroup $(P_t)_{t\geq 0}$, the following identity holds:*

$$\begin{aligned} \int_E |f(x)|^p \, m(\mathrm{d}x) - \lim_{T\to+\infty} \|P_T f\|_p^p \qquad (1.2) \\ = \frac{4(p-1)}{p} \int_0^{+\infty} \mathcal{E}^c[(P_t f)^{\langle p/2\rangle}] \, \mathrm{d}t \\ + \int_0^{+\infty} \iint_{E\times E\setminus \mathrm{diag}} F_p(P_t f(x), P_t f(y)) \, J(\mathrm{d}x, \mathrm{d}y)\mathrm{d}t \\ + p\int_0^{+\infty} \int_E |P_t f(x)|^p \, k(\mathrm{d}x)\mathrm{d}t, \quad f \in L^p(m). \end{aligned}$$

Here, $\|\cdot\|_p$ is the norm of the space $L^p(m)$. In the above identity k is the killing measure *and $\mathcal{E}^c$ is the* strongly local part *of the Beurling–Deny formula (see* (1.7) *below).*

The strongly local part $\mathcal{E}^c$ and the killing measure k describe, respectively, the diffusive and killing behavior of the Markov process associated with the semigroup $(P_t)_{t\geq 0}$. We stress that in typical cases the term $\lim_{T\to+\infty} \|P_T f\|_p^p$ vanishes. The statement in Theorem 1.1 is rigorously described in Corollary 7.4.

1.2 Sobolev–Bregman form

The strategy employed by the author was to study the inner integral of the right-hand side of (1.1), that is,

$$\mathcal{E}_p[u] = \frac{1}{p} \iint\limits_{E\times E\setminus\text{diag}} F_p(u(x), u(y))\, J(\mathrm{d}x, \mathrm{d}y) = \frac{1}{p} \iint\limits_{E\times E\setminus\text{diag}} H_p(u(x), u(y))\, J(\mathrm{d}x, \mathrm{d}y), \quad u \in L^p(m), \tag{1.3}$$

where the *symmetrized Bregman divergence* is defined by

$$H_p(a,b) = \frac{1}{2}\left(F_p(a,b) + F_p(b,a)\right) = \frac{p}{2}(b-a)\left(b^{\langle p-1\rangle} - a^{\langle p-1\rangle}\right),$$

and is more commonly used in this context in the recent literature. The second inequality in (1.3) follows from the symmetry of the measure J. The form $\mathcal{E}_p$ is the so-called *Sobolev–Bregman form* or shortly *p-form*. Although this notion has appeared earlier, the name *Sobolev–Bregman form* was introduced only recently in Bogdan, Jakubowski, Lenczewska, and Pietruska-Pałuba [16] and Bogdan, Grzywny, Pietruska-Pałuba, and Rutkowski [14]. Let A_p be the generator of the semigroup $(P_t)_{t\geq 0}$ on $L^p(m)$. It turns out that the p-form in the purely non-local case can be written in terms of A_p as follows: $\mathcal{E}_p[u] = \langle -A_p u, u^{\langle p-1\rangle}\rangle$. In particular, $\mathcal{E}_p[u]$ is finite whenever u belongs to the domain $\mathcal{D}(A_p)$ of the generator A_p, which may be written as $\mathcal{D}(A_p) \subseteq \mathcal{D}(\mathcal{E}_p)$. Since the generator A_p is defined by the following limit: $A_p u = \lim_{t\to 0^+} \frac{1}{t}(P_t u - u)$ in $L^p(m)$, one can propose the alternative definition of p-form:

$$\mathcal{E}_p[u] = \lim_{t\to 0^+} \mathcal{E}^{(t)}(u, u^{\langle p-1\rangle}), \quad u \in \mathcal{D}(\mathcal{E}_p), \tag{1.4}$$

with domain

$$\mathcal{D}(\mathcal{E}_p) = \left\{ u \in L^p(m) : \text{finite } \lim_{t\to 0^+} \mathcal{E}^{(t)}(u, u^{\langle p-1\rangle}) \text{ exists} \right\},$$

where $\mathcal{E}^{(t)}(u,v) = \frac{1}{t}\langle u - P_t u, v\rangle$, $t > 0$ is the so-called *approximate form.* It was shown in Lemma 7 in [16], that $\mathcal{D}(\mathcal{E}_p)$ coincides with the class of functions with the integral on the right-hand side of (1.3) finite, at least for the fractional Laplacian. Nevertheless, in contrast to (1.3), the definition (1.4) is more appropriate in the general context, while (1.3) is relevant only for Markov processes with a pure-jump behavior. For instance, in the case of the Laplace operator Δ the p-form has the following form:

$$\mathcal{E}_p[u] = (p-1) \int\limits_{\mathbb{R}^d} |u|^{p-2} |\nabla u|^2 \,\mathrm{d}x, \tag{1.5}$$

at least for u in the Sobolev space $W^{2,p}(\mathbb{R}^d)$; see Metafune and Spina [76] and Lemma E.1 in Bogdan, Gutowski, and Pietruska-Pałuba [15].

With the general definition of the p-form at hand, we can derive a more abstract version of the Hardy–Stein identity.

Theorem 1.2. *Let $1 < p < +\infty$. Under some standard assumptions about the semigroup $(P_t)_{t\geq 0}$, the following identity holds:*

$$\int_E |f(x)|^p \, m(\mathrm{d}x) - \lim_{T\to+\infty} \|P_T f\|_p^p = p \int_0^{+\infty} \mathcal{E}_p[P_t f] \, \mathrm{d}t, \quad f \in L^p(m). \tag{1.6}$$

The precise statement of the above identity is described in Theorem 3.2.

As can be seen, here we include the case where the semigroup $(P_t)_{t\geq 0}$ does not satisfy the *strong stability* condition: $\lim_{T\to+\infty} \|P_T f\|_p = 0$, $f \in L^p(m)$. Identity (1.6) may be written also in the differential form:

$$\frac{\mathrm{d}}{\mathrm{d}t} \|P_t f\|_p^p = -p\mathcal{E}_p[P_t f], \quad f \in L^p(m), \, t > 0.$$

For this, see Proposition 3.4. This last relation has been known since the work of Varopoulos [108]; see equation (1.1). In light of the above observations, the derivation of the explicit form of the Hardy–Stein identity can be reduced to finding an explicit formula for the Sobolev–Bregman form $\mathcal{E}_p$.

The notion of the Sobolev–Bregman form is an extension of the extensively studied concept of a (regular) Dirichlet form. In brief, a (regular) Dirichlet form is a closed, non-negative, symmetric bilinear form $\mathcal{E}(u,v)$, with domain $\mathcal{D}(\mathcal{E}) \subseteq L^2(m)$ satisfying certain additional conditions. In terms of the generator, the Dirichlet form can be characterized by $\mathcal{E}(u,v) = \langle \sqrt{-A_2}u, \sqrt{-A_2}v \rangle$, where the domain is given by $\mathcal{D}(\mathcal{E}) = \mathcal{D}(\sqrt{-A_2})$; here, $\sqrt{-A_2}$ is defined via spectral theory. The *quadratic form* $\mathcal{E}[u] = \mathcal{E}(u,u)$ is the special case of the p-form when $p = 2$, namely $\mathcal{E}_2[u] = \mathcal{E}(u,u)$.

The theory of regular Dirichlet forms is widely explored in the literature. The author primarily follows the book by Fukushima, Oshima, and Takeda [45], however, other references include Ma, Röckner [72], Wang [110], and Fabes et al. [37]. The class of regular Dirichlet forms is in the one-to-one correspondence with a specific class of symmetric Markov processes known as symmetric *Hunt processes*. For a comprehensive description of this correspondence, we refer to Chapter 7 of [45]. For this reason, the results of the present dissertation are relevant to symmetric Hunt processes. Summarizing all the above remarks, generators A, semigroups $(P_t)_{t\geq 0}$, symmetric Hunt processes $(X_t)_{t\geq 0}$, regular Dirichlet forms $\mathcal{E}$ (and, as we will see later, Sobolev–Bregman forms) are all mutually connected.

The central result from this theory, crucial for our consideration, is the Beurling–Deny formula, which provides the following unique decomposition of a regular Dirichlet form $\mathcal{E}$ into three parts:

$$\begin{aligned} \mathcal{E}(u,v) = \mathcal{E}^c(u,v) & \\ & + \frac{1}{2} \iint_{E\times E\setminus \mathrm{diag}} (u(y)-u(x))(v(y)-v(x)) \, J(\mathrm{d}x,\mathrm{d}y) \\ & + \int_E u(x)v(x) \, k(\mathrm{d}x), \end{aligned} \tag{1.7}$$

for all continuous (in general *quasi-continuous*) functions u and v in the domain $\mathcal{D}(\mathcal{E})$. The above three parts are: the *strongly local part*, the *jumping part*, and the *killing part*. These parts respectively describe the diffusive, jumping, and killing behavior of the Hunt process associated with $\mathcal{E}$. The latter two are completely characterized by the so-called *jumping measure* J and the *killing measure* k.

One of the main results of the present work is the Sobolev–Bregman analogue of the Beurling–Deny formula.

Theorem 1.3. *The following characterization of the domain of the Sobolev–Bregman form holds:*

$$\mathcal{D}(\mathcal{E}_p) = \{ u \in L^p(m) : u^{\langle p/2 \rangle} \in \mathcal{D}(\mathcal{E}) \}. \tag{1.8}$$

In fact, the following estimate holds for any $u \in \mathcal{D}(\mathcal{E}_p)$*:*

$$\frac{4(p-1)}{p^2}\mathcal{E}[u^{\langle p/2 \rangle}] \le \mathcal{E}_p[u] \le 2\mathcal{E}[u^{\langle p/2 \rangle}]. \tag{1.9}$$

Moreover, the p*-form* $\mathcal{E}_p$ *is given by the following formula for continuous (in general quasi-continuous)* $u \in \mathcal{D}(\mathcal{E}_p)$*:*

$$\begin{aligned}\mathcal{E}_p[u] = {} & \frac{4(p-1)}{p^2}\mathcal{E}^c[u^{\langle p/2 \rangle}] \\ & + \frac{1}{p}\iint\limits_{E\times E\setminus \mathrm{diag}} F_p(u(x),u(y))\, J(\mathrm{d}x,\mathrm{d}y) + \int\limits_E |u(x)|^p\, k(\mathrm{d}x).\end{aligned} \tag{1.10}$$

Here, the measures J *and* k *are the jumping measure and the killing measure in the classical Beurling–Deny formula for the Dirichlet form* $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$*.*

The above result is presented rigorously in Theorem 7.1. Additionally, less general but chronologically earlier versions of Theorem 7.1 are stated in Theorems 4.6 and 6.2. In fact, the Hardy–Stein identity given in (1.2) is a consequence of (1.6) and (1.10). Observe that in view of (1.10), quantity (1.3) is a special case of the p-form for the so-called *pure-jump* case (i.e., $\mathcal{E}^c = 0$) with no killing (i.e., $k = 0$).

The estimate given in (1.9) for $\mathcal{E}$ and $\mathcal{E}_p$ defined in terms of the generator:

$$\mathcal{E}[u^{\langle p/2 \rangle}] = \langle \sqrt{-A_2}u^{\langle p/2 \rangle}, \sqrt{-A_2}u^{\langle p/2 \rangle}\rangle, \qquad \mathcal{E}_p[u] = \langle -A_p u, u^{\langle p-1 \rangle}\rangle,$$

with $u \in \mathcal{D}(A_p) \subseteq (\mathcal{D}(\sqrt{-A_2}))^{\langle 2/p \rangle}$ was broadly investigated in much earlier works. The lower bound can be found in the book of Stroock [106, Lemma 9.9], and in [108, Lemma on p. 246]. The two-sided estimate was given in Liskevich and Semënov [67, Theorem 1], [68, Theorem 1], Liskevich and Perel'muter [65, Theorem 3.3], and Liskevich, Perel'muter, and Semënov [66, Theorem 3.1]. We also refer to Proposition 8.1 in Kinzebulatov and Semënov [59] for the lower bound. In the context of the fractional Laplacian, the lower bound can also be found in Section 7.6 of Cialdea and Maz'ya [35].

1.3 Polarized Hardy–Stein identity

A separate part of the author's investigation focused on deriving the polarized analog of the purely non-local Hardy–Stein identity (1.1) and the polarized version of the p-form in this context. One of the results of the present dissertation is the following polarized Hardy–Stein identity presented in [15].

Theorem 1.4. *Let $2 \leq p < \infty$ and $f, g \in L^p(m)$. Denote $\Phi := (f, g)$ and $P_t\Phi := (P_t f, P_t g)$. Let $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$ be a pure-jump Dirichlet form (i.e., $\mathcal{E}^c = 0$) with no killing (i.e., $k = 0$). Under some mild assumptions,*

$$\int_E f(x) g^{\langle p-1\rangle}(x)\, m(\mathrm{d}x) = \int_0^{+\infty} \iint_{E\times E\setminus \mathrm{diag}} \mathcal{J}_p(P_t\Phi(x), P_t\Phi(y))\, J(\mathrm{d}x, \mathrm{d}y)\mathrm{d}t. \tag{1.11}$$

Here, $\mathcal{J}_p$ is a polarized version of the Bregman divergence F_p,

$$\begin{aligned}\mathcal{J}_p(w, z) = {} & z_1 z_2^{\langle p-1\rangle} - w_1 w_2^{\langle p-1\rangle} \\ & - w_2^{\langle p-1\rangle}(z_1 - w_1) - (p-1) w_1 |w_2|^{p-2}(z_2 - w_2).\end{aligned}$$

A complete statement of the above result is given in Theorem 5.6.

Additionally, it was proven in [15] that the polarized Sobolev–Bregman form defined by

$$\mathcal{E}_p(u, v) := \frac{1}{p} \iint_{E\times E\setminus \mathrm{diag}} \mathcal{J}_p(\Phi(x), \Phi(y))\, J(\mathrm{d}x, \mathrm{d}y),$$

where $\Phi := (u, v)$, is well-defined for $2 \leq p < \infty$ and u, v in the domain of the generator A_p. In such a situation, the following identity holds.

Theorem 1.5. *Let $2 \leq p < \infty$ and $u, v \in \mathcal{D}(A_p)$. Under the assumptions of Theorem 1.4,*

$$\mathcal{E}_p(u, v) = -\frac{1}{p}\langle A_p u, v^{\langle p-1\rangle}\rangle - \frac{1}{p}\langle A_p v, (p-1) u |v|^{p-2}\rangle. \tag{1.12}$$

The above result is stated rigorously in Theorem 5.10.

1.4 Applications in Littlewood–Paley theory

Revisiting the non-local case, we want to note that Bañuelos, Bogdan, and Luks in [5] applied the Hardy–Stein identity to derive L^p-bounds for the following *square* (or *Littlewood–Paley*) *functions* of $f \in L^p(m)$:

$$G(x) = \left(\frac{1}{2} \int_0^{+\infty} \int_{E\setminus\{x\}} (P_t f(y) - P_t f(x))^2\, J(x, \mathrm{d}y)\mathrm{d}t\right)^{1/2} \tag{1.13}$$

and

$$\widetilde{G}(x) = \left(\int_0^{+\infty} \int_{E\setminus\{x\}} (P_t f(y) - P_t f(x))^2\, \chi(P_t f(x), P_t f(y)) J(x, \mathrm{d}y)\mathrm{d}t\right)^{1/2}, \tag{1.14}$$

where $\chi(s, t) = \mathbb{1}_{\{|s|>|t|\}} + \frac{1}{2}\mathbb{1}_{\{|s|=|t|\}}$ and $J(x, \mathrm{d}y)$ is the kernel of the jumping measure J: $J(\mathrm{d}x, \mathrm{d}y) = J(x, \mathrm{d}y)m(\mathrm{d}x)$. More precisely, they studied the existence of constants $c_p, C_p > 0$ such that

$$c_p \|f\|_p \leq \|G\|_p \leq C_p \|f\|_p \tag{1.15}$$

holds for all $f \in L^p(m)$. The estimates given in (1.15) are called *Littlewood–Paley estimates.*

Square functions were originally introduced by Littlewood and Paley [69], although the concept can be traced back to the works of Kaczmarz [55] and Zygmund [114]; see also Zygmund [113]. These works concerned the application of a certain square function to study the pointwise convergence of Fourier series. Later, Littlewood and Paley presented various results connecting square functions with Fourier series in a series of articles [69, 70, 71, 88, 89]. In later years further works appeared, addressing both earlier and entirely new square functions. These include, in particular, the work of Marcinkiewicz [73] introducing the so-called Marcinkiewicz square function, Marcinkiewicz and Zygmund [75] studying the area integral of Lusin, Zygmund [115, 116] about power series, Marcinkiewicz [74] treating Fourier multipliers, Zygmund [117], where the L^p-boundedness of Marcinkiewicz square function was shown, and Zygmund [118], where the proof of Littlewood's and Paley's results from [70] was simplified. Further works of Zygmund and his students set new directions for the development of square functions. At that point, the theory was generalized to the multidimensional case. For this, we refer to Calderón [26], Stein [95], and Benedek, Calderón, and Panzone [9]; see also Calderón [27]. Moreover, significant progress was achieved in the study of the Marcinkiewicz square function. For this, we may refer to Weiss and Zygmund [112] and to [95]. What is notable, the Marcinkiewicz square function found an application to the study of derivatives (even fractional) in the sense of L^p; see Stein and Zygmund [104, 105]. We also would like to mention the works of Calderón and Zygmund [28] and Zygmund [119, 120]. The later period of research involving square functions brought to life a rich development of the theory of the Hardy spaces. These include the work of Burkholder [23], extending Paley's theorem for Walsh-Paley series to general martingales, and its further extensions: Burkholder and Gundy [24] and Burkholder, Gundy, and Silverstein [25]. We should also mention Stein [96, 97, 98], Fefferman and Stein [40], and Gundy and Stein [49], where, in particular, the characterization of the Hardy spaces was described through Brownian motion. Among other applications of the classical Littlewood–Paley theory, we may also mention Stein [101], Stein and Wainger [103], Nagel, Stein, and Wainger [87], Weiss and Wainger [111, pp. 429–434], Fabes, Jerison, and Kenig [38], and Jones and Kenig [46, pp. 24-90]. Many of the concepts mentioned above were included in the books of Stein [99, 100]. In particular, in [100], square functions played a crucial role in establishing the maximal inequality for semigroups; see Lemma 2.1 below and Chapter 8 for an application of this result in the present dissertation. For a more detailed account of the history of the development of classical Littlewood–Paley theory, outlined in this paragraph, we refer the reader to the brilliant essay of Stein [102].

At this point, it is worth noting that, due to the very broad scope of applications of square functions, the literature features a great variety of square functions, some sharing close similarities, whereas others are more loosely connected. Nevertheless, this dissertation addresses the topic of the family of square functions associated with stochastic processes. This idea was initiated by Meyer in his series of articles [79, 80, 78, 81, 82]; see also further corrections [83, 84]. Briefly speaking, the focus is on those square functions that can be expressed through the carré du champ operator Γ introduced by Meyer in [78]. Since, in the case of Brownian motion, we have $\Gamma[u] = |\nabla u|^2$, the square functions studied by Stein in [99] can be interpreted as square functions corresponding to the Brownian motion and, consequently, to the classical Laplace operator Δ. This area was later further explored by Meyer in [85, 86]. Another probabilistic approach to the Littlewood–Paley functions from [100] was proposed by Varopoulos [109]. Additionally, we highlight the work of

Bañuelos [1], which introduced the Brownian analogue of the Lusin area integral. See also the monograph of Bañuelos and Moore [2]. Nevertheless, closely related square functions for Brownian motion were introduced slightly earlier by Bennett in [10]. Later, Bouleau and Lamberton generalized the aforementioned probabilistic methods from Brownian motion to the α-stable processes in [21]. A continuation of a study of the Littlewood–Paley theory in the α-stable case can be found in more recent works. We refer here to Kim and Kim [58] and to Karlı [56].

The application of the classical Hardy–Stein identity to prove the L^p-boundedness of square functions was provided by Stein in [99, pp. 86–88]. Nevertheless, the main scope of the present dissertation is the L^p-boundedness in the non-local case. For the first implementation of the Hardy–Stein identity in obtaining non-local Littlewood–Paley estimates, we refer again to [5]. The identity (1.2) generalizes the non-local version of the Hardy–Stein identity proposed in [5] not only to a broader class of Markov processes, but also to a class of topological spaces E, significantly broader than Euclidean spaces $\mathbb{R}^d$. For that reason, in this work, we apply the approach from [5] to generalize the Littlewood–Paley estimates for G and $\widetilde{G}$.

Another tool applied by Bañuelos, Bogdan, and Luks in [5] as well as by Li and Wang in [63] to obtain the Littlewood–Paley estimates is the Burkholder–Davies–Gundy inequality. In this work, we propose a nontrivial adaptation of this method for our purpose. To achieve this, we employ the notion of Revuz correspondence.

Moreover, we present some counterexamples to the L^p-boundedness of square functions G and $\widetilde{G}$, and propose alternative definitions of the square function as an attempt to overcome this issue.

For a precise list of Littlewood–Paley estimates and counterexamples, we refer to the introduction to Chapter 8.

1.5 Structure of the Dissertation

The main notions used in this work, such as semigroups of contractions, Dirichlet forms, Sobolev–Bregman forms or Hunt processes, are introduced in Chapter 2. The abstract version of the Hardy–Stein identity (1.6) is proved in Chapter 3. Chapters 4, 6, and 7 cover the results on the p-form analog of the Beurling–Deny formula. The polarized Hardy–Stein formula is given in Chapter 5. The Littlewood–Paley estimates are investigated in Chapter 8.

Most notation is introduced in Chapter 2. The concepts utilized only in Chapters 5 and 8 are described in the introductions to the corresponding chapters.

Chapter 2

Definitions and discussion

In this work we use the following notation. The symbol $x := y$ means that x is defined to be equal to y.

2.1 Topological assumptions

We start with a description of the topological space we will work on. It is a standard structure for the theory of Dirichlet forms, for example, in Fukushima, Oshima, and Takeda; see [45, (1.1.7.)]. The basic notions and properties of the Dirichlet form theory are described in Section 2.4.

Let E be a locally compact separable metric space, and let $\mathcal{B}(E)$ be the σ-algebra of all Borel sets in E. Consider a Radon measure[1] m on E, that is, a measure on the σ-algebra of Borel sets $\mathcal{B}(E)$ which is finite on all compact sets, outer regular on all Borel sets, and inner regular on all open sets. Moreover, we assume that m is of full support, that is, $\operatorname{supp} m = E$. In particular, it follows that m is strictly positive on non-empty open sets.

For $1 \leq p \leq \infty$, we consider the Banach space $L^p(E, \mathcal{B}(E), m)$ of real-valued Borel functions with finite p-norm. For simplicity, we shall write $L^p(m)$. Denote by $\|\cdot\|_p$ the p-norm of $L^p(m)$. Let q be the conjugate exponent of p, that is, $1/p + 1/q = 1$ with $q = 1$ for $p = \infty$ and $q = \infty$ for $p = 1$. For functions $f \in L^p(m)$ and $g \in L^q(m)$ we use the canonical pairing notation

$$\langle f, g \rangle := \int_E f(x)g(x)\, m(\mathrm{d}x).$$

In Chapter 5 we also use the vectorized version of the above notation. Fix a non-negative integer n. Let $f = (f_1, \ldots, f_n)$, where $f_1, \ldots, f_n \in L^p(m)$ and $g = (g_1, \ldots, g_n)$, where $g_1, \ldots, g_n \in L^q(m)$. Then we shall write

$$\langle f, g \rangle := \int_E f(x) \cdot g(x)\, m(\mathrm{d}x) = \sum_{j=1}^{n} \int_E f_j(x)g_j(x)\, m(\mathrm{d}x). \tag{2.1}$$

Here, $\cdot$ is the dot product on $\mathbb{R}^n$.

[1] See Folland [44].

2.2 Strongly continuous semigroup of contractions

In this section, we introduce a typical notion of the family of linear operators on $L^2(m)$ connected with Dirichlet forms. In this procedure, we follow Fukushima, Oshima, and Takeda [45]; see Chapter 1. Regarding L^p-extension, we refer to the monograph by Jacob [53]. The connection with Dirichlet forms is presented in Section 2.4, below.

We consider the family $(P_t)_{t\geq 0}$ of linear operators on $L^2(m)$ with domain $L^2(m)$ satisfying the following conditions:

($P_t.L^2.1$) Each P_t is *symmetric*: $\langle P_t f, g\rangle = \langle f, P_t g\rangle$, $f, g \in L^2(m)$, $t > 0$.

($P_t.L^2.2$) $(P_t)_{t\geq 0}$ is a *semigroup*: $P_s P_t = P_{s+t}$, $s, t > 0$.

($P_t.L^2.3$) *Contraction property*: $\|P_t f\|_2 \leq \|f\|_2$, $f \in L^2(m)$, $t > 0$.

($P_t.L^2.4$) *Strong continuity*: $P_t f \to f$ in $L^2(m)$ as $t \to 0^+$, $f \in L^2(m)$.

($P_t.L^2.5$) *Sub-Markov property*: $0 \leq P_t f \leq 1$ m-a.e. whenever $f \in L^2(m)$, $0 \leq f \leq 1$ m-a.e.

Each operator P_t can be uniquely extended to an operator on $L^\infty(m)$ and also to an operator on $L^1(m)$. These operators are contractions: $\|P_t f\|_p \leq \|f\|_p$, for $f \in L^p(m)$, with sub-Markov property: $0 \leq P_t f \leq 1$ m-a.e., for $f \in L^p(m)$, $0 \leq f \leq 1$ m-a.e., when $p \in \{\infty, 1\}$. For these constructions, we refer to [45, p. 56] and [45, p. 37], respectively. Further, the above family of operators $(P_t)_{t\geq 0}$ on $L^1(m)$ constitutes a strongly continuous semigroup: $P_t f \to f$ in $L^1(m)$ as $t \to 0^+$, for $f \in L^1(m)$; see [45, p. 201]. The above observations imply that the semigroup $(P_t)_{t\geq 0}$ defined as above can be uniquely extended to a family on $L^p(m)$ for $1 \leq p < \infty$, satisfying the following conditions:

($P_t.L^p.1$) Each P_t is *symmetric* on $L^2(m)$:

$$\langle P_t f, g\rangle = \langle f, P_t g\rangle, \quad f, g \in L^p(m) \cap L^2(m),\ t > 0.$$

($P_t.L^p.2$) $(P_t)_{t\geq 0}$ is a *semigroup*: $P_s P_t = P_{s+t}$, $s, t > 0$.

($P_t.L^p.3$) *Contraction property*: $\|P_t f\|_p \leq \|f\|_p$, $f \in L^p(m)$, $t > 0$.

($P_t.L^p.4$) *Strong continuity*: $P_t f \to f$ in $L^p(m)$ as $t \to 0^+$, $f \in L^p(m)$.

($P_t.L^p.5$) *Sub-Markov property*: $0 \leq P_t f \leq 1$ m-a.e. whenever $f \in L^p(m)$, $0 \leq f \leq 1$ m-a.e.

It is noteworthy that the semigroup $(P_t)_{t\geq 0}$ may not be strongly continuous on $L^\infty(m)$ in general. We refer to Section 1 of Farkas, Jacob, and Schilling [39] where a type of Riesz–Thorin interpolation theorem was applied. See also Section 2.6 in [53, pp. 133-138]. These references work (for simplicity) in the context $E = \mathbb{R}^d$, but the same argument holds for a general E.

Strong continuity and sub-Markov property imply together that every P_t is also *positivity preserving* (another name: *conservation of positivity*), i.e., $P_t f \geq 0$ m-a.e. whenever $f \in L^p(m)$, $f \geq 0$ m-a.e. This implication is almost straightforward; nevertheless, for confidence, we may refer, e.g., to Lemma 4.6.2 in the monograph by Jacob [52].

Every operator P_t possess the kernel $P_t(x, \mathrm{d}y)$, namely,

$$P_t f = \int_E f(y)\, P_t(x, \mathrm{d}y), \quad f \in L^p(m).$$

We also denote $P_t(\mathrm{d}x, \mathrm{d}y) := P_t(x, \mathrm{d}y)m(\mathrm{d}x)$. Due to the symmetry property ($P_t.L^p.1$), we have $P_t(\mathrm{d}x, \mathrm{d}y) = P_t(\mathrm{d}y, \mathrm{d}x)$.

In addition, the symmetry also implies that for $f \in L^1(m)$, we may write

$$\begin{aligned}\int_E P_t f(x)\, m(\mathrm{d}x) &= \int_E \left(\int_E f(y)\, P_t(x, \mathrm{d}y) \right) m(\mathrm{d}x) = \iint_{E\times E} f(y)\, P_t(\mathrm{d}x, \mathrm{d}y) \\ &= \int_E f(y) \left(\int_E P_t(y, \mathrm{d}x) \right) m(\mathrm{d}y) = \int_E f(x) P_t 1(x)\, m(\mathrm{d}x).\end{aligned} \tag{2.2}$$

Under assumptions ($P_t.L^p.1$)–($P_t.L^p.4$) of the semigroup $(P_t)_{t\geq 0}$, we obtain the following maximal inequality of Stein. We refer to [100, p. 73]. This inequality will be very useful in Chapter 8.

Lemma 2.1 (Stein's maximal inequality)**.** *Let $1 \leq p < \infty$. For arbitrary $f \in L^p(m)$, if we denote $f^*(x) := \sup_{t\geq 0} |P_t f(x)|$, then*

$$\|f^*\|_p \leq \begin{cases} \frac{p}{p-1} \|f\|_p & \text{if } 1 < p < \infty, \\ \|f\|_\infty & \text{if } p = \infty. \end{cases} \tag{2.3}$$

We define the *infinitesimal generator* of the semigroup $(P_t)_{t\geq 0}$ on $L^p(m)$:

$$\begin{cases} \mathcal{D}(A_p) := \left\{ u \in L^p(m) : \lim\limits_{t\to 0^+} \dfrac{1}{t}(P_t u - u) \text{ exists in } L^p(m) \right\}, \\ \quad A_p u := \lim\limits_{t\to 0^+} \dfrac{1}{t}(P_t u - u) \quad \text{in } L^p(m). \end{cases} \tag{2.4}$$

The contraction property and the symmetry of the semigroup (conditions ($P_t.L^p.1$) and ($P_t.L^p.3$)) imply that $(P_t)_{t\geq 0}$ is analytic on $L^p(m)$ for $1 < p < \infty$, i.e., it extends to the family of operators $(P_{t+ir})_{t+ir\in S_p}$ with $t + ir$ in some sector $S_p \supseteq (0, +\infty)$ on the complex plane for which L^p-valued-function $S_p \ni t + ir \mapsto P_{t+ir} f \in L^p(m)$ is analytic for all $f \in L^p(m)$. This is well-known result provided by Stein in [100, Theorem 1 in Chapter II]. For further results for a wider domain of analyticity than Stein's S_p, we refer to Liskevich and Perel'muter [65].

In our study, we will consider only t in real semi-axis $[0, +\infty)$. The above results imply that for $1 < p < \infty$, $t > 0$, and any $f \in L^p(m)$ function $P_t f$ belongs to $\mathcal{D}(A_p)$. This observation will be crucial in the proof of the Hardy–Stein identity.

2.3 Derivatives on L^p

Inspired by Bogdan, Jakubowski, Lenczewska and Pietruska-Pałuba [16], we will employ the calculus of L^p-derivatives. Nevertheless, this approach may be found much earlier. We may refer to the proof of Lemma on page 246 in Varopoulos [108].

The following notions will be very useful to derive the abstract version of the Hardy–Stein identity in Chapter 3 and in the multidimensional adaptation of this approach in Chapter 5.

Let $1 \leq p < \infty$ and fix some interval $I \subseteq \mathbb{R}$. For a mapping $I \ni t \mapsto u(t) \in L^p(m)$ we denote

$$\Delta_h u(t) := u(t + h) - u(t) \quad \text{if } t, t + h \in I.$$

We say that u is *continuous* on I with values in $(L^p(m))^n$ if $\Delta_h u(t) \to 0$ in $(L^p(m))^n$ as $h \to 0$ for every $t \in I$. We say that u is *differentiable* on I with values in $(L^p(m))^n$ if $u'(t) := \lim_{h\to 0} \frac{1}{h}\Delta_h u(t)$ exists in $(L^p(m))^n$ for every $t \in I$. We say that u is *continuously differentiable* (or shortly C^1) on I with values in $(L^p(m))^n$ if u is differentiable and the mapping $I \ni t \mapsto u'(t) \in (L^p(m))^n$ is continuous.

In further considerations, we will pay special attention to the mapping $u(t) := P_t f$, $t \geq 0$, where f is some fixed function in $L^p(m)$. The differentiability of this mapping is known, at least for $1 < p < \infty$. Indeed, since semigroup $(P_t)_{t\geq 0}$ is analytic on $L^p(m)$ for $1 < p < \infty$, the L^p-derivative of $u(t)$ is equal to $A_p u(t)$ for $t > 0$, and also $u'(t) = P_t A_p f$ for $t \geq 0$, when we additionally assume that $f \in \mathcal{D}(A_p)$. In fact, u is C^1 with values in $L^p(m)$ on $(0, +\infty)$ or $[0, +\infty)$ respectively.

Let us define the *French power*

$$a^{\langle\gamma\rangle} := |a|^\gamma \operatorname{sgn} a,$$

for real a and γ, whenever the above expression makes sense. We highlight the following connection between functions: $|a|^\gamma$ and $a^{\langle\gamma\rangle}$

$$(|a|^\gamma)' = \gamma a^{\langle\gamma-1\rangle}, \quad \text{if either } a \in \mathbb{R},\ \gamma > 1 \text{ or } a \in \mathbb{R}\setminus\{0\}, \tag{2.5}$$

$$(a^{\langle\gamma\rangle})' = \gamma\, |a|^{\gamma-1}, \quad \text{if either } a \in \mathbb{R},\ \gamma \geq 1 \text{ or } a \in \mathbb{R}\setminus\{0\}. \tag{2.6}$$

In further chapters, we will utilize the following L^p-counterpart of (2.5) and (2.6) applied to $P_t f$.

Corollary 2.2. *Let $f \in L^p(m)$ and $u(t) := P_t f$. Let $1 < \gamma \leq p$.*

(i) The mapping $|u|^\gamma$ is C^1 on $(0, +\infty)$ with values in $L^{p/\gamma}(m)$ and

$$(|u(t)|^\gamma)' = \gamma u(t)^{\langle\gamma-1\rangle} u'(t) = \gamma u(t)^{\langle\gamma-1\rangle} A_p P_t f, \quad t \geq 0. \tag{2.7}$$

(ii) The mapping $u^{\langle\gamma\rangle}$ is C^1 on $(0, +\infty)$ with values in $L^{p/\gamma}(m)$ and

$$(u(t)^{\langle\gamma\rangle})' = \gamma\, |u(t)|^{\gamma-1} u'(t) = \gamma |u(t)|^{\gamma-1} A_p P_t f, \quad t \geq 0. \tag{2.8}$$

In addition, if $f \in \mathcal{D}(A_p)$, then $|u|^\gamma$ and $u^{\langle\gamma\rangle}$ are C^1 on $[0, +\infty)$ with values in $L^{p/\gamma}(m)$.

The above statement is the corollary of Prop A.2 from Appendix A that derives L^p-derivatives for an arbitrary C^1 mapping in the multidimensional case. The multidimensional version of this statement will be needed only in Chapter 5, therefore here we restrict considerations to the one-dimensional case. This statement appeared previously in Lemmas 15, 16 from [16] and in the joint work written by the author with Bogdan and Pietruska-Pałuba [15, Corollary 2.3].

2.4 Regular Dirichlet forms

In this section, we introduce the notion of Dirichlet form, define a regular Dirichlet form, and present the connection between this notion and the semigroup introduced in Section 2.2. In this framework, we draw upon [45]; see Chapter 1. Some other concepts essential to the results of Chapter 8 are introduced there and are based on Bouleau, Hirsch [20]. For other sources dealing with the theory of Dirichlet forms, we may refer to Ma, Röckner [72], Wang [110], and Fabes et al. [37].

The class of continuous functions defined on a set A will be denoted by $C(A)$. By $C_c(A)$ we will denote the class of functions in $C(A)$ with compact support.

We consider a bilinear form $\mathcal{E}\colon \mathcal{D}(\mathcal{E})\times\mathcal{D}(\mathcal{E})\to\mathbb{R}$ with domain $\mathcal{D}(\mathcal{E})\subseteq L^2(m)$. Denote

$$\mathcal{E}_{(\alpha)}(u,v) := \mathcal{E}(u,v)+\alpha\langle u,v\rangle,\quad \alpha>0,\ u,v\in\mathcal{D}(\mathcal{E}). \tag{2.9}$$

Let $\mathcal{E}[u]:=\mathcal{E}(u,u)$ and $\mathcal{E}_{(\alpha)}[u]:=\mathcal{E}_{(\alpha)}(u,u)$. We say that $v\in L^2(m)$ is a *normal contraction* of $u\in L^2(m)$ if some Borel version $\widetilde{v}$ of v

$$|\widetilde{v}(y)-\widetilde{v}(x)|\le|\widetilde{u}(y)-\widetilde{u}(x)|\,,\quad |\widetilde{v}(x)|\le|\widetilde{u}(x)|\,,\quad x,y\in E,$$

for some Borel version $\widetilde{u}$ of u.

We assume that $\mathcal{E}$ satisfies the following conditions:

($\mathcal{E}$.1) $\mathcal{D}(\mathcal{E})$ is dense in $L^2(m)$

($\mathcal{E}$.2) *Symmetry*: $\mathcal{E}(u,v)=\mathcal{E}(v,u)$, $u,v\in\mathcal{D}(\mathcal{E})$.

($\mathcal{E}$.3) $\mathcal{E}$ is *non-negative*: $\mathcal{E}[u]\ge 0$, $u\in\mathcal{D}(\mathcal{E})$.

($\mathcal{E}$.4) $\mathcal{E}$ is *closed*: if $(u_n)_{n\in\mathbb{N}}$ is a sequence of elements of $\mathcal{D}(\mathcal{E})$ such that $\mathcal{E}_{(1)}[u_n-u_m]\to 0$ as $m,n\to+\infty$, then there exists $u\in\mathcal{D}(\mathcal{E})$ such that $\mathcal{E}_{(1)}[u_n-u]\to 0$ as $n\to+\infty$.

($\mathcal{E}$.5) $\mathcal{E}$ is *Markovian*: For each $\varepsilon>0$ there exists a real function φ_ε such that:

(i) $\varphi_\varepsilon(t)=t$ for $t\in[0,1]$,

(ii) $-\varepsilon(t)\le\varphi_\varepsilon(t)\le 1+\varepsilon$ for $t\in\mathbb{R}$,

(iii) $0\le\varphi_\varepsilon(t)-\varphi_\varepsilon(s)\le t-s$ for $s<t$,

(iv) $\varphi_\varepsilon(u)\in\mathcal{D}(\mathcal{E})$ and $\mathcal{E}[\varphi_\varepsilon(u)]\le\mathcal{E}[u]$ whenever $u\in\mathcal{D}(\mathcal{E})$.

Condition ($\mathcal{E}$.5) is equivalent to the following one:

($\mathcal{E}$.5') *Every normal contraction operates on* $\mathcal{E}$: when $u\in\mathcal{D}(\mathcal{E})$, then for every normal contraction v of u we have $v\in\mathcal{D}(\mathcal{E})$ and $\mathcal{E}[v]\le\mathcal{E}[u]$.

Note that the last one implies that for any Lipschitz function φ with Lipschitz constant L such that $\varphi(0)=0$ the following implication holds:

$$u\in\mathcal{D}(\mathcal{E})\Rightarrow\varphi(u)\in\mathcal{D}(\mathcal{E}),\ L^2\mathcal{E}[\varphi(u)]\le\mathcal{E}[u]. \tag{2.10}$$

A pair $(\mathcal{E},\mathcal{D}(\mathcal{E}))$ satisfying conditions ($\mathcal{E}$.1)–($\mathcal{E}$.5) is called a *Dirichlet form*. In such a case, $\mathcal{D}(\mathcal{E})$ equipped with the inner product $\mathcal{E}_{(\alpha)}$ is a Hilbert space for each $\alpha>0$. We say that a Dirichlet form $(\mathcal{E},\mathcal{D}(\mathcal{E}))$ is *regular*, when it additionally satisfies:

($\mathcal{E}$.6) $(\mathcal{E},\mathcal{D}(\mathcal{E}))$ posses a *core*, i.e., a subset $\mathcal{C}$ of $\mathcal{D}(\mathcal{E})\cap C_c(E)$ such that $\mathcal{C}$ is dense in both spaces: $(\mathcal{D}(\mathcal{E}),\mathcal{E}_{(1)})$ and $(C_c(E),\|\cdot\|_\infty)$.

The principles provided above can be found in Section 1.1 of [45].

One may use the semigroup to construct the approximate form of $\mathcal{E}$. Indeed, let $(P_t)_{t\geq 0}$ be the semigroup associated with a Dirichlet form $(\mathcal{E},\mathcal{D}(\mathcal{E}))$. For any $u,v\in L^2(m)$ (and in general, for $u\in L^p(m)$, $v\in L^q(m)$ in later chapters) we define the following approximate form of $(\mathcal{E},\mathcal{D}(\mathcal{E}))$ by:

$$\mathcal{E}^{(t)}(u,v) := \frac{1}{t}\langle u - P_t u, v\rangle, \qquad t>0. \tag{2.11}$$

We also use the notation $\mathcal{E}^{(t)}[u] := \mathcal{E}^{(t)}(u,u)$. By the approximation we understand the following characterization of $(\mathcal{E},\mathcal{D}(\mathcal{E}))$:

$$\begin{cases} \mathcal{D}(\mathcal{E}) = \left\{ u\in L^2(m) : \text{finite } \lim\limits_{t\to 0^+} \mathcal{E}^{(t)}[u] \text{ exists} \right\}, \\ \mathcal{E}(u,v) = \lim\limits_{t\to 0^+} \mathcal{E}^{(t)}(u,v), \quad u,v\in\mathcal{D}(\mathcal{E}). \end{cases} \tag{2.12}$$

Moreover, for $u\in L^2(m)$ the function $t\mapsto \mathcal{E}^{(t)}[u]$ is non-increasing.

Directly from (2.4) and (2.11) we derive that $\mathcal{D}(A_2)\subseteq\mathcal{D}(\mathcal{E})$ and

$$\mathcal{E}(u,v) = -\langle A_2 u, v\rangle, \qquad u\in\mathcal{D}(A_2), v\in\mathcal{D}(\mathcal{E}). \tag{2.13}$$

The Dirichlet form may be also characterize by the generator, as follows

$$\begin{cases} \mathcal{D}(\mathcal{E}) = \mathcal{D}(\sqrt{-A_2}), \\ \mathcal{E}(u,v) = \langle \sqrt{-A_2}u, \sqrt{-A_2}u\rangle, \quad u,v\in\mathcal{D}(\mathcal{E}), \end{cases} \tag{2.14}$$

which follows from the spectral theorem. In fact, there is one-to-one correspondence between the class of semigroups $(P_t)_{t\geq 0}$ on $L^2(m)$ described in Section 2.2 and the class of Dirichlet forms. This correspondence is given by (2.12). For more details, see Section 1.3. of [45].

It is known that the semigroup $(P_t)_{t\geq 0}$ is strongly continuous on $\mathcal{D}(\mathcal{E})$ with norm $\sqrt{\mathcal{E}[\cdot]}$, i.e.,

$$\mathcal{E}[P_t u - u]\to 0, \quad \text{as } t\to 0^+. \tag{2.15}$$

We refer to Lemma 1.3.3 in [45].

Let $u,v\in L^2(m)$. The approximate form may be rewritten as follows

$$\mathcal{E}^{(t)}[u] = \frac{1}{2t}\iint\limits_{E\times E} (u(y)-u(x))^2\, P_t(\mathrm{d}x,\mathrm{d}y) + \frac{1}{t}\int\limits_E u(x)^2(1-P_t 1(x))\, m(\mathrm{d}x), \tag{2.16}$$

and, by polarization

$$\begin{aligned} \mathcal{E}^{(t)}(u,v) = \frac{1}{2t}&\iint\limits_{E\times E} (u(y)-u(x))(v(y)-v(x))\, P_t(\mathrm{d}x,\mathrm{d}y) \\ &+ \frac{1}{t}\int\limits_E u(x)v(x)(1-P_t 1(x))\, m(\mathrm{d}x). \end{aligned} \tag{2.17}$$

Here, we used the symmetry of the operator P_t on $L^2(m)$. Compare (2.17) with (1.4.8.) in [45].

For arbitrary $u \in \mathcal{D}(\mathcal{E}) \cap L^\infty(m)$ there exists uniquely a positive Radon measure $\mu_{[u]}$ for which

$$\int_E f \, \mathrm{d}\mu_{[u]} = \mathcal{E}(uf, u) - \frac{1}{2}\mathcal{E}(u^2, f), \quad f \in \mathcal{D}(\mathcal{E}) \cap C_c(E).$$

It turns out that the measure $\mu_{[u]}$ can be uniquely extended to any $u \in \mathcal{D}(\mathcal{E})$. We call $\mu_{[u]}$ the *energy measure* (or *carré du champ measure*) of $u \in \mathcal{D}(\mathcal{E})$. In particular,

$$\mathcal{E}[u] = \mu_{[u]}(E), \quad u \in \mathcal{D}(\mathcal{E}).$$

For more details, we refer to [45]; see equation (3.2.14) and the rest of the discussion on that page.

Denote by $\mathcal{O}$ the family of all open subsets of E. For $A \subseteq E$ denote by $\mathcal{O}_A$ the family of all open supersets of A. For $U \in \mathcal{O}$, we define $\mathcal{L}_U := \{ u \in \mathcal{D}(\mathcal{E}) : u \geq 1 \ m\text{-a.e. on } U \}$. The *capacity* of the Dirichlet form $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$ is defined as follows: for $A \in \mathcal{O}$

$$\mathrm{Cap}(A) := \begin{cases} \inf_{u \in \mathcal{L}_U} \mathcal{E}_{(1)}[u] & \text{if } \mathcal{L}_U \neq \emptyset, \\ +\infty & \text{if } \mathcal{L}_U = \emptyset, \end{cases} \tag{2.18}$$

and for any $A \subseteq E$

$$\mathrm{Cap}(A) := \inf_{U \in \mathcal{O}_A} \mathrm{Cap}(U). \tag{2.19}$$

We say that some statement depending on $x \in A$ holds *quasi-everywhere* (or shortly *q.e.*) if there is a set $N \subseteq A$ of zero capacity such that this statement is true for any $x \in A \setminus N$.

We adjoin an extra point Δ to $(E, \mathcal{B}(E))$: we set $E_\Delta := E \cup \{\Delta\}$ and $\mathcal{B}(E_\Delta) = \mathcal{B}(E) \cup \{ B \cup \Delta : B \in \mathcal{B}(E) \}$. We treat $(E_\Delta, \mathcal{B}(E_\Delta))$ as the one-point compactification of $(E, \mathcal{B}(E))$. When $(E, \mathcal{B}(E))$ is already compact, then Δ is an isolated point of $(E_\Delta, \mathcal{B}(E_\Delta))$. We will always treat each function u on E as an extension on E_Δ setting $u(\Delta) = 0$ without mentioning. In particular, every set of zero capacity is of the zero measure m in view of the inequality $m(A) \leq \mathrm{Cap}(A)$.

Let u be a function defined q.e. on E_Δ with values in $[-\infty, +\infty]$. We say that u is *quasi-continuous* if for every $\varepsilon > 0$ there exists an open set $U \subseteq E_\Delta$ such that $\mathrm{Cap}(U) < \varepsilon$ and u is continuous and finite on $E_\Delta \setminus U$. In particular, every continuous function is quasi-continuous. Here, by the *quasi-continuity* we mean the notion called *quasi-continuity in the restricted sense* described in Section 2.1 of [45].

One of the main theorems in the theory of regular Dirichlet forms is the Beurling–Deny formula. Assume that $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$ is a regular Dirichlet form. It is known that every function u in the domain of $\mathcal{D}(\mathcal{E})$ has its *quasi-continuous* modification $\widetilde{u}$ on some set of zero measure m. We refer to Theorem 2.1.7 in [45].

The Beurling–Deny formula provides the following unique decomposition of regular Dirichlet form $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$:

$$\begin{aligned} \mathcal{E}(u, v) = {} & \mathcal{E}^c(u, v) \\ & + \frac{1}{2} \iint_{E \times E \setminus \mathrm{diag}} (\widetilde{u}(y) - \widetilde{u}(x))(\widetilde{v}(y) - \widetilde{v}(x)) \, J(\mathrm{d}x, \mathrm{d}y) \\ & + \int_E \widetilde{u}(x)\widetilde{v}(x) \, k(\mathrm{d}x), \end{aligned} \tag{2.20}$$

where $\tilde{u}, \tilde{v}$ denote quasi-continuous versions of arbitrary $u, v \in \mathcal{D}(\mathcal{E})$ and $\mathcal{E}^c$ is symmetric form on $\mathcal{D}(\mathcal{E}) \times \mathcal{D}(\mathcal{E})$ which satisfies *strongly local condition*:

$$\mathcal{E}^c(u,v) = 0 \quad \text{if } v \text{ is constant on a neighborhood of } \operatorname{supp} u.$$

We will also use the notation $\mathcal{E}^c[u] := \mathcal{E}^c(u,u)$. Measure J is a Radon measure on $E \times E \setminus \mathrm{diag} = \{ (x,y) \in E \times E : x \neq y \}$, called *jumping measure* and k is a Radon measure on E, called *killing measure*. In particular, when $\mathcal{E}^c = 0$, we say that $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$ is a *pure-jump* Dirichlet form.

We emphasize that we cannot identify functions u that are equal m-almost everywhere, because k and J can assign positive measure to sets that are of zero measure under m or $m \otimes m$.

In particular, when $u = v \in \mathcal{D}(\mathcal{E})$,

$$\mathcal{E}[u] = \mathcal{E}^c[u] + \frac{1}{2} \iint\limits_{E \times E \setminus \mathrm{diag}} (\tilde{u}(y) - \tilde{u}(x))^2 \, J(\mathrm{d}x, \mathrm{d}y) + \int\limits_E (\tilde{u}(x))^2 \, k(\mathrm{d}x), \tag{2.21}$$

and for every $u \in \mathcal{D}(\mathcal{E})$ both integrals on the right-hand side of (2.21) are finite. We say that a Dirichlet form $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$ is *maximally defined* when the converse statement is true: for any quasi-continuous function $u \in L^2(m)$, if both integrals on the right-hand side of (2.21) are finite, then $u \in \mathcal{D}(\mathcal{E})$. We stress that many commonly used pure-jump Dirichlet forms are maximally defined (see Schilling, Uemura [92]), and we are unaware of any example of a regular pure-jump Dirichlet form which is not maximally defined in view of the notion of an *(active) reflected Dirichlet form* and *Silverstein extension*. We refer to Chen [30] and Kuwae [60]. However, we also observe that the above definition is essentially applicable only to pure-jump Dirichlet forms: if the strongly local term $\mathcal{E}^c$ is not vanishing, it is expected that $\mathcal{E}$ is not maximally defined in the above sense.

Similarly to previous considerations, for an arbitrary $u \in \mathcal{D}(\mathcal{E}) \cap L^\infty(m)$ there exists uniquely a positive Radon measure $\mu^c_{[u]}$ for which

$$\int\limits_E f \, \mathrm{d}\mu^c_{[u]} = \mathcal{E}^c(uf, u) - \frac{1}{2}\mathcal{E}^c(u^2, f), \quad f \in \mathcal{D}(\mathcal{E}) \cap C_c(E),$$

and the measure $\mu^c_{[u]}$ can be uniquely extended to any $u \in \mathcal{D}(\mathcal{E})$. We call $\mu^c_{[u]}$ the *local part of energy measure* (or *local part of carré du champ measure*) of $u \in \mathcal{D}(\mathcal{E})$. In particular,

$$\mathcal{E}^c[u] = \mu^c_{[u]}(E), \quad u \in \mathcal{D}(\mathcal{E}). \tag{2.22}$$

For more details, we refer to [45]; see equation (3.2.19) and the rest of the discussion on that page.

Furthermore, the local part satisfies so-called *LeJan's formulae*:

$$\mathcal{E}^c(\varphi(u), \psi(u)) = \int\limits_E \varphi'(u(x)) \psi'(u(x)) \, \mu^c_{[u]}(\mathrm{d}x), \quad u \in \mathcal{D}(\mathcal{E}), \tag{2.23}$$

where φ and ψ are any Lipschitz functions equal 0 at point 0. The references of the above statement in that version require more thorough attention. Compare with (2.10). The multidimensional version statement for φ, ψ of class C^1 and bounded u can be found in Theorem 3.2.2 in [45]. Moreover, footnote 8 therein claims that the statement from this work is also true, and refers to Corollaire 4.4 in Bouleau and Hirsch [19]. Therein, the same result is provided under different assumptions, but the proof can be adapted to the present context. We note that everything required in this work is actually covered by the more general Théorème 3.1 in [19], which applies directly to the form $\mathcal{E}^c$.

It is known that for any $u, v \in \mathcal{D}(\mathcal{E})$

$$\lim_{t\to 0^+} \frac{1}{t} \int_E u(x)v(x)(1 - P_t 1(x))\, m(\mathrm{d}x) = \int_E \widetilde{u}(x)\widetilde{v}(x)\, k(\mathrm{d}x). \tag{2.24}$$

See Lemmas 4.5.2 and 4.5.3 in [45] and also equation (1.4) in [31].

Combining (2.12) and (2.17) with (2.24) we know also that for any $u, v \in \mathcal{D}(\mathcal{E})$

$$\begin{aligned} &\lim_{t\to 0^+} \frac{1}{2t} \iint_{E\times E} (u(y) - u(x))(v(y) - v(x))\, P_t(\mathrm{d}x, \mathrm{d}y) = \\ &\mathcal{E}^c(u, v) + \frac{1}{2} \iint_{E\times E\setminus \mathrm{diag}} (\widetilde{u}(y) - \widetilde{u}(x))(\widetilde{v}(y) - \widetilde{v}(x))\, J(\mathrm{d}x, \mathrm{d}y). \end{aligned} \tag{2.25}$$

Moreover,

$$\frac{1}{t} P_t(\mathrm{d}x, \mathrm{d}y) \to J(\mathrm{d}x, \mathrm{d}y) \quad \text{vaguely on } E \times E \setminus \mathrm{diag} \text{ when } t \to 0^+. \tag{2.26}$$

A similar result for the resolvent instead of the semigroup $(P_t)_{t\geq 0}$ was demonstrated in [45]; see equation (3.2.7). The proof is similar, thus we omit it here.

2.5 Sobolev–Bregman forms

It may be non-trivial to point out the first occurrence of the notion of the Sobolev–Bregman form $\mathcal{E}_p$. Our guess is Bakry [4, p. 37], where an unnamed notion $\mathcal{E}_p[u] = \langle -A_2 u, u^{p-1}\rangle$ (for $u > 0$) was introduced; compare this with (2.28) below. The name *Sobolev–Bregman* was introduced only recently in Bogdan, Jakubowski, Lenczewska, and Pietruska-Pałuba [16] and Bogdan, Grzywny, Pietruska-Pałuba, and Rutkowski [14]. For recent applications of p-form, we refer to [13, 14, 16, 57, 17]. In [13] Bogdan, Grzywny, Pietruska-Pałuba, and Rutkowski applied this notion in the theory of non-linear non-local PDEs. The same authors in [14] utilized the Sobolev–Bregman form to derive non-linear non-local Douglas identity. Sobolev–Bregman form was also used to derive the optimal constant of Hardy inequalities for the fractional Laplacian in [16]. This topic was further investigated by Kijaczko and Lenczewska in [57]. For application of Bregman divergence related to a convex function ϕ in the study of a variation of semimartingales, we refer to Bogdan, Kutek, and Pietruska-Pałuba [17].

The Sobolev–Bregman form in the form $\mathcal{E}_p[u] = \langle -A_p u, u^{\langle p-1\rangle}\rangle$ was used, e.g., to study perturbations of semigroups on L^p-spaces; see Theorem 3.2 in Liskevich, Perel'muter and Semënov [66] and also Liskevich and Semënov [68]. We also refer to the notion of L^p-Dirichlet-operator from Farkas, Jacob, and Schilling [39] and from Definition 4.6.7 in Jacob [52].

Whenever we deal with the notion of the Sobolev–Bregman form, we shall assume that $1 < p < \infty$.

Analogously to the characterization (2.12) of Dirichlet form, we define the non-linear form on $L^p(m)$ by

$$\begin{cases} \mathcal{D}(\mathcal{E}_p) := \left\{ u \in L^p(m) : \text{finite } \lim_{t\to 0^+} \mathcal{E}^{(t)}(u, u^{\langle p-1\rangle}) \text{ exists} \right\}, \\ \mathcal{E}_p[u] := \lim_{t\to 0^+} \mathcal{E}^{(t)}(u, u^{\langle p-1\rangle}), \quad u \in \mathcal{D}(\mathcal{E}_p). \end{cases} \tag{2.27}$$

The form $\mathcal{E}_p$ will be called *Sobolev–Bregman form* or just *p-form*.

Note that for $p = 2$ the notion of p-form $\mathcal{E}_2[u]$ reduce to the quadratic Dirichlet form $\mathcal{E}[u]$. We emphasize that, in contrast to Dirichlet forms, we do not know the monotonicity of the function $t \mapsto \mathcal{E}^{(t)}(u, u^{\langle p-1\rangle})$ for $p \neq 2$.

Note that we have the following counterpart of (2.13): the inclusion $\mathcal{D}(A_p) \subseteq \mathcal{D}(\mathcal{E}_p)$ holds and

$$\mathcal{E}_p[u] = -\langle A_p u, u^{\langle p-1\rangle}\rangle, \qquad u \in \mathcal{D}(A_p). \tag{2.28}$$

Recall that $p > 1$. We define *Bregman divergence*: a function $F_p \colon \mathbb{R} \times \mathbb{R} \to \mathbb{R}$, defined by

$$F_p(a, b) := |b|^p - |a|^p - pa^{\langle p-1\rangle}(b - a). \tag{2.29}$$

We use also the *symmetrized Bregman divergence*

$$H_p(a, b) := \frac{1}{2}\left(F_p(a, b) + F_p(b, a)\right) = \frac{p}{2}(b - a)\left(b^{\langle p-1\rangle} - a^{\langle p-1\rangle}\right). \tag{2.30}$$

Note that $F_p(a, b)$ is the second-order Taylor remainder of the convex function $\mathbb{R} \ni a \mapsto |a|^p \in \mathbb{R}$, hence F_p and H_p are non-negative. In particular,

$$F_2(a, b) = H_2(a, b) = (b - a)^2.$$

The following estimate will be useful later.

Lemma 2.3. *Let $1 < p < \infty$. There are constants $c'_p > 0$, $C'_p \geq 1$, and $1 \leq C_p \leq 2$ such that*

$$\frac{2(p-1)}{p}\left(b^{\langle p/2\rangle} - a^{\langle p/2\rangle}\right)^2 \leq H_p(a, b) \leq \frac{pC_p}{2}\left(b^{\langle p/2\rangle} - a^{\langle p/2\rangle}\right)^2 \tag{2.31}$$

and

$$c'_p\left(b^{\langle p/2\rangle} - a^{\langle p/2\rangle}\right)^2 \leq F_p(a, b) \leq C'_p\left(b^{\langle p/2\rangle} - a^{\langle p/2\rangle}\right)^2 \tag{2.32}$$

for any $a, b \in \mathbb{R}$. Here, the constant C_p coincides with the optimal constant $a(p)$ from [67, Lemma 1]. In particular, the following symmetry with exponent $q = p/(p-1)$ conjugated to p holds: $C_q = C_p$.

The second estimate – of the symmetrized Bregman divergence – is well-known. The proof was provided first by Liskevich and Semënov with the optimal upper constant $a(p)$ (here C_p) in [67]; see Lemma 1. We may also refer to [66], Lemma 2.1. Nevertheless, the proof in the case of $a, b \geq 0$, known as *Stroock's inequality*, can be found in previous works. We refer to the proof of Lemma 9.9 in Stroock [106] and also to Lemma on page 246 in [108], to page 269 in Carlen, Kusuoka, and Stroock [29], to inequality (2.2.9) in Davies [36], or to page 39 in [4]. Two-sided bound of non-symmetrized Bregman divergence with unexplicit constants was proven by the author in [50]; see Lemma 2.1. We refer also to the related approach in Lemma 2.3 in [14].

Another estimate for Bregman divergence will be useful in dealing with Littlewood–Paley square functions in Chapter 8.

Lemma 2.4. *Let $1 < p < \infty$. There are constants $c_p, C_p > 0$ such that*

$$c_p \, |b-a|^2 \, (|a| \vee |b|)^{p-2} \leq F_p(a,b) \leq C_p \, |b-a|^2 \, (|a| \vee |b|)^{p-2} \tag{2.33}$$

and therefore

$$c_p \, |b-a|^2 \, (|a| \vee |b|)^{p-2} \leq H_p(a,b) \leq C_p \, |b-a|^2 \, (|a| \vee |b|)^{p-2} \tag{2.34}$$

for any $a, b \in \mathbb{R}$.

The above estimate appeared (in a vectorized version) in Pinchover, Tertikas, and Tintarev [90]; see (2.19). Nevertheless, one-sided bounds may be found earlier in Shafrir [93, Lemma 7.4] or, for $2 \leq p < \infty$, in Barbatis, Filippas, and Tertikas [7, Lemma 3.1]. The one-dimensional case was proven by Bogdan, Dyda, and Luks in [12]; see Lemma 6. We also refer to Lemma 2.3 in [14]. For optimal constants in certain ranges of p, we refer to [16] and [93, Lemma 7.4]. See also Lemma A.3.

Analogously to (2.16), we may rewrite the form approximating p-form and obtain

$$\begin{aligned} \mathcal{E}^{(t)}\left(u, u^{\langle p-1\rangle}\right) = \frac{1}{pt} \iint\limits_{E\times E} & F_p(u(x), u(y)) \, P_t(\mathrm{d}x, \mathrm{d}y) \\ &+ \frac{1}{t} \int\limits_E |u(x)|^p \, (1 - P_t 1(x)) \, m(\mathrm{d}x). \end{aligned} \tag{2.35}$$

Indeed, employing (2.2), we may write

$$\begin{aligned} \mathcal{E}^{(t)}\left(u, u^{\langle p-1\rangle}\right) &= \frac{1}{t} \langle u - P_t u, u^{\langle p-1\rangle} \rangle \\ &= -\frac{1}{t} \iint\limits_{E\times E} u^{\langle p-1\rangle}(x)(u(y) - u(x)) \, P_t(\mathrm{d}x, \mathrm{d}y) \\ &\quad + \frac{1}{t} \int\limits_E |u(x)|^p \, (1 - P_t 1(x)) \, m(\mathrm{d}x) \\ &= \frac{1}{pt} \int\limits_E P_t(|u|^p)(x) \, m(\mathrm{d}x) - \frac{1}{pt} \int\limits_E |u(x)|^p \, P_t 1(x) \, m(\mathrm{d}x) \\ &\quad - \frac{1}{t} \iint\limits_{E\times E} u^{\langle p-1\rangle}(x)(u(y) - u(x)) \, P_t(\mathrm{d}x, \mathrm{d}y) \\ &\quad + \frac{1}{t} \int\limits_E |u(x)|^p \, (1 - P_t 1(x)) \, m(\mathrm{d}x) \\ &= \frac{1}{pt} \iint\limits_{E\times E} F_p(u(x), u(y)) \, P_t(\mathrm{d}x, \mathrm{d}y) \\ &\quad + \frac{1}{t} \int\limits_E |u(x)|^p \, (1 - P_t 1(x)) \, m(\mathrm{d}x). \end{aligned}$$

Observe that since $(P_t)_{t\geq 0}$ is Markovian, $P_t 1 \leq 1$, second integral on the right-hand side is non-negative. Also $F_p \geq 0$, thus first integral is non-negative as well. Therefore, for $u \in L^p(m)$ we have $\mathcal{E}^{(t)}(u, u^{\langle p-1\rangle}) \geq 0$ and so $\mathcal{E}_p[u] \geq 0$ whenever $u \in \mathcal{D}(\mathcal{E}_p)$.

Because of the symmetry of $P_t(\mathrm{d}x, \mathrm{d}y)$ we may also write the approximate form in terms of symmetrized Bregman divergence

$$\begin{aligned}\mathcal{E}^{(t)}\left(u, u^{\langle p-1\rangle}\right) &= \frac{1}{pt}\iint\limits_{E\times E} H_p(u(x), u(y))\, P_t(\mathrm{d}x, \mathrm{d}y) \\ &\quad + \frac{1}{t}\int\limits_E |u(x)|^p\,(1 - P_t1(x))\, m(\mathrm{d}x) \\ &= \frac{1}{2t}\iint\limits_{E\times E} (u(y)-u(x))(u^{\langle p-1\rangle}(y) - u^{\langle p-1\rangle}(x))\, P_t(\mathrm{d}x, \mathrm{d}y) \\ &\quad + \frac{1}{t}\int\limits_E |u(x)|^p\,(1 - P_t1(x))\, m(\mathrm{d}x).\end{aligned} \tag{2.36}$$

This will be more suitable in Chapter 7.

Combining (2.17), (2.36), and the estimate (2.31) from Lemma 2.3, we derive the following estimate between approximating forms:

$$\frac{4(p-1)}{p^2}\mathcal{E}^{(t)}\left[u^{\langle p/2\rangle}\right] \le \mathcal{E}^{(t)}\left(u, u^{\langle p-1\rangle}\right) \le C_p\mathcal{E}^{(t)}\left[u^{\langle p/2\rangle}\right], \tag{2.37}$$

which holds for any $u \in L^p(m)$. Here, we used the fact that the constant C_p from Lemma 2.3 satisfies $C_p \ge 1$.

This observation may be found in the previous works. We refer to [66]; see equation (3.2) in the proof of Theorem 3.1, where the general sub-Markovian semigroups were considered.

2.6 Hunt processes

Below we give a very short introduction to the theory of Hunt processes. For more details on Hunt processes and their connection with Dirichlet forms, we refer to the book of Fukushima, Oshima, and Takeda [45], especially to Appendix A.2. We may also refer to Chapter 5 in the monograph by Jacob [54].

Fix $x \in E$. Let $(\Omega, \mathcal{F}, \mathbb{P}_x)$ be a probability space. We consider a stochastic process $(\Omega, \mathcal{F}, (X_t)_{t\ge 0}, \mathbb{P}_x)$ with a state space $(E, \mathcal{B}(E))$. In particular, $X_t\colon (\Omega, \mathcal{F}) \to (E, \mathcal{B}(E))$ is a measurable mapping for each $t \ge 0$. In this section, when we write $t \ge 0$, we allow $t = \infty$.

Recall that we denote by $E_\Delta = E \cup \{\Delta\}$ and $\mathcal{B}(E_\Delta) = \mathcal{B}(E) \cup \{\, B \cup \Delta : B \in \mathcal{B}(E)\,\}$ the one-point compactification of $(E, \mathcal{B}(E))$. When $(E, \mathcal{B}(E))$ is already compact, then Δ is an isolated point of $(E_\Delta, \mathcal{B}(E_\Delta))$.

For a probabilistic measure μ on $(E_\Delta, \mathcal{B}(E_\Delta))$ we denote $\mathbb{P}_\mu(A) := \int_{E_\Delta} \mathbb{P}_x(A)\,\mu(\mathrm{d}x)$ whenever $x \mapsto \mathbb{P}_x(A)$ is measurable. Let $\zeta(\omega) := \inf\{\, t \ge 0 : X_t(\omega) = \Delta\,\}$ be the *lifetime* of $(X_t)_{t\ge 0}$. We say that the quadruple $(\Omega, \mathcal{F}, (X_t)_{t\ge 0}, (\mathbb{P}_x)_{x\in E_\Delta})$ is a *Hunt process* if there exist an admissible filtration $(\mathcal{F}_t)_{t\ge 0}$ and a family of mappings $(\theta_t)_{t\ge 0}$ such that the following conditions are satisfied:

(X.1) For each $x \in E_\Delta$, the quadruple $(\Omega, \mathcal{F}, (X_t)_{t\ge 0}, \mathbb{P}_x)$ is a stochastic process with a state space $(E, \mathcal{B}(E))$ and $X_\infty(\omega) = \Delta$ for all $\omega \in \Omega$.

(X.2) For each $t \ge 0$ and $B \in \mathcal{B}(E)$, the mapping $E \ni x \mapsto \mathbb{P}_x(X_t \in B)$ is measurable.

(X.3) $(\Omega, \mathcal{F}, (X_t)_{t\geq 0}, (\mathbb{P}_x)_{x\in E_\Delta})$ possesses the *strong Markov property with respect to* $(\mathcal{F}_t)_{t\geq 0}$, i.e., $(\mathcal{F}_t)_{t\geq 0}$ is right continuous and

$$\mathbb{P}_\mu(X_{\sigma+s} \in B|\mathcal{F}_\sigma) = \mathbb{P}_{X_\sigma}(X_s \in B), \quad \mathbb{P}_\mu\text{-a.s.}$$

for any probabilistic measure μ on E, $B \in \mathcal{B}(E_\Delta)$, $s \geq 0$, and $(\mathcal{F}_t)$-stopping time σ.

(X.4) $\mathbb{P}_\Delta(X_t = \Delta) = 1$ for all $t \geq 0$.

(X.5) $\mathbb{P}_x(X_0 = x) = 1$ for all $x \in E$.

(X.6) $X_t(\omega) = \Delta$ for all pairs (t, ω) such that $t \geq \zeta(\omega)$.

(X.7) For each $t \geq 0$, the mapping $\theta_t \colon \Omega \to \Omega$ satisfies $X_s \circ \theta_t = X_{s+t}$ for all $s \geq 0$.

(X.8) For each $\omega \in \Omega$, the sample path $[0, +\infty) \ni t \mapsto X_t(\omega) \in E_\Delta$ is right continuous and has the left limit on $(0, +\infty)$.

(X.9) $(\Omega, \mathcal{F}, (X_t)_{t\geq 0}, (\mathbb{P}_x)_{x\in E_\Delta})$ is *quasi-left-continuous*, i.e., for any sequence of $(\mathcal{F}_t)$-stopping times (σ_n) increasing to $(\mathcal{F}_t)$-stopping time σ and probabilistic measure μ on E_Δ

$$\mathbb{P}_\mu\left(\lim_{n\to+\infty} X_{\sigma_n} = X_\sigma, \sigma < \infty\right) = \mathbb{P}_\mu(\sigma < \infty).$$

The point Δ is called the *cemetery* state. The time ζ is called the *life time*. The mapping θ_t is called the *translation operator*. For simplicity the Hunt process $(\Omega, \mathcal{F}, (X_t)_{t\geq 0}, (\mathbb{P}_x)_{x\in E_\Delta})$ will be denoted shortly as $(X_t)_{t\geq 0}$.

For $(X_t)_{t\geq 0}$ we define

$$\mathcal{F}_t^0 := \begin{cases} \sigma\{\, X_s : s \leq t \,\} & \text{for } t < \infty, \\ \sigma\{\, X_s : s < \infty \,\} & \text{for } t = \infty. \end{cases}$$

We call $(\mathcal{F}_t^0)_{t\geq 0}$ the *minimum admissible filtration*. Let μ be a probabilistic measure on $(E_\Delta, \mathcal{B}(E_\Delta))$. We denote by $\mathcal{F}_t^\mu$ (resp. $\mathcal{F}_\infty^\mu$) the completion of $\mathcal{F}_t^0$ (resp. of $\mathcal{F}_\infty^0$) with respect to $\mathbb{P}_\mu : \mathcal{F}_\infty^0 \to [0, 1]$. The filtration defined by $\mathcal{F}_t := \bigcap_\mu \mathcal{F}_t^\mu$ is called the *minimum completed admissible filtration*. Here, the intersection is taken over all probabilistic measures μ on $(E_\Delta, \mathcal{B}(E_\Delta))$. Usually, it is convenient to consider the Hunt process $(X_t)_{t\geq 0}$ with respect to the minimum completed admissible filtration; see Theorem A.2.1 in [45].

We say that $(\Omega, \mathcal{F}, (X_t)_{t\geq 0}, (\mathbb{P}_x)_{x\in E_\Delta})$ has *Markov property*, if for any $x \in E$, $t, s \geq 0$, and $B \in \mathcal{B}(E)$

$$\mathbb{P}_x(X_{t+s} \in B|\mathcal{F}_t) = \mathbb{P}_{X_t}(X_s \in B), \quad \mathbb{P}_x\text{-a.s.}.$$

Clearly, the strong Markov property is a stronger condition than the Markov property. Markov property imply the following identity:

$$\mathbb{E}_x[\Phi \circ \theta_t|\mathcal{F}_t] = \mathbb{E}_{X_t}\Phi, \quad \mathbb{P}_x\text{-a.s.}, \quad t \geq 0, \tag{2.38}$$

for any $\Phi \colon \Omega \to \mathbb{R}$ measurable with respect to $\mathcal{F}_\infty$.

Let $(X_t)_{t\geq 0}$ be the Hunt process. Then, $(X_t)_{t\geq 0}$ generates the semigroup $P_t f(x) := \mathbb{E}_x f(X_t)$ which satisfies the conditions outlined in Section 2.2, except for the symmetry assumption $(P_t.L^2.1)$. We say that $(X_t)_{t\geq 0}$ is *symmetric* if $(P_t)_{t\geq 0}$ satisfies $(P_t.L^2.1)$. Under $(P_t.L^2.1)$, there exists a regular Dirichlet form $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$ associated with $(P_t)_{t\geq 0}$. We say that $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$ is the Dirichlet form *associated with* the symmetric Hunt process $(X_t)_{t\geq 0}$.

The converse statement is also true in some sense. Let $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$ be a regular Dirichlet form and let $(\tilde{P}_t)_{t\geq 0}$ be the semigroup associated with $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$. Then, there exists a symmetric Hunt process $(X_t)_{t\geq 0}$ with state space E and with a semigroup $(P_t)_{t\geq 0}$ such that for every $f \in L^2(m)$ and $t > 0$ function $P_t f$ is a quasi-continuous version of $\tilde{P}_t f$. The above Hunt process is unique with respect to some equivalence relation. For more details, we refer to Section 4.2, Section 7.2, and Appendix A of [45], especially to Theorem 4.2.3 and Theorem 7.2.1. Later we will not distinguish between P_t and $\tilde{P}_t$, hence we will always write P_t.

Let $(X_t)_{t\geq 0}$ be the (non-symmetric) Hunt process and let $(P_t)_{t\geq 0}$ be the semigroup associated with $(X_t)_{t\geq 0}$. For each $t \geq 0$, an operator P_t has its adjoint operator $\hat{P}_t$ on $L^2(m)$, i.e.,

$$\langle P_t f, g\rangle = \langle f, \hat{P}_t g\rangle,$$

for all $f, g \in L^2(m)$. Clearly, $P_t = \hat{P}_t$ for symmetric $(X_t)_{t\geq 0}$. We denote

$$R_\alpha f := \int_0^{+\infty} e^{-\alpha t} P_t f \,\mathrm{d}t, \qquad \hat{R}_\alpha f := \int_0^{+\infty} e^{-\alpha t} \hat{P}_t f \,\mathrm{d}t,$$

where the above integrals are Bochner integrals. The families of operators $(R_\alpha)_{\alpha>0}$ and $(\hat{R}_\alpha)_{\alpha>0}$ are called the *resolvent* (resp. *coresolvent*) of $(X_t)_{t\geq 0}$. In particular, when $f \in C_c(E)$, $\alpha R_\alpha f \to f$ and $\alpha \hat{R}_\alpha f \to f$ in $L^2(m)$ as $\alpha \to +\infty$.

We say that a non-negative Borel function f is *α-excessive* (resp. *α-coexcessive*) with respect to the semigroup $(P_t)_{t\geq 0}$ (resp. $(\hat{P}_t)_{t\geq 0}$), if for all $t > 0$, $e^{-\alpha t} P_t u \leq u$ m-a.e. (resp. $e^{-\alpha t}\hat{P}_t u \leq u$ m-a.e.). Let f be a non-negative Borel function. The simple example of an α-excessive (resp. α-coexcessive) function is $\alpha R_\alpha f$ (resp. $\alpha \hat{R}_\alpha f$). We say that a Borel measure $\tilde{m}$ is *excessive* if it is σ-finite and $\int_E P_t \mathbb{1}_B \,\mathrm{d}\tilde{m} \leq \tilde{m}(B)$, $B \in \mathcal{B}(E)$.

Assume that the Hunt process $(X_t)_{t\geq 0}$ is symmetric. Then, the following identity holds:

$$\int_E \mathbb{E}_x \varphi(x, X_t)\, m(\mathrm{d}x) = \int_E \mathbb{E}_x \varphi(X_t, x)\, m(\mathrm{d}x), \quad t \geq 0. \tag{2.39}$$

Here, $\varphi\colon E \times E \to \mathbb{R}$ is an arbitrary measurable function such that the above integrals are finite.

The symmetric Hunt process $(X_t)_{t\geq 0}$ is *self-dual* with respect to the reference measure m, which implies that for $T > 0$

$$\int_E \mathbb{E}_x \Phi((X_{(T-t)-} : t \in [0,T]))\, \mathbb{1}_{\{T<\zeta\}}\; m(\mathrm{d}x) = \int_E \mathbb{E}_x \Phi((X_t : t \in [0,T]))\, \mathbb{1}_{\{T<\zeta\}}\; m(\mathrm{d}x). \tag{2.40}$$

In particular, if $(X_t)_{t\geq 0}$ is symmetric, then for non-negative Borel function $\varphi\colon [0,\infty)\times E \to \mathbb{R}$,

$$\begin{aligned}\int_E \mathbb{E}_x \left(\int_0^T \varphi(T-t, X_t)\,\mathrm{d}t\right) \mathbb{1}_{\{T<\zeta\}}\; m(\mathrm{d}x) &= \int_E \mathbb{E}_x \left(\int_0^T \varphi(t, X_{T-t})\,\mathrm{d}t\right) \mathbb{1}_{\{T<\zeta\}}\; m(\mathrm{d}x)\\ &= \int_E \mathbb{E}_x \left(\int_0^T \varphi(t, X_t)\,\mathrm{d}t\right) \mathbb{1}_{\{T<\zeta\}}\; m(\mathrm{d}x). \end{aligned} \tag{2.41}$$

Chapter 3

Hardy–Stein identity – general form

In this short chapter we prove the abstract version of the Hardy–Stein formula introduced in (1.2) which elucidates the role of the Sobolev–Bregman form in the Hardy–Stein-type identities. This result was provided by Bogdan, Pietruska-Pałuba, and the author of this dissertation in [15] in the case of pure-jump Lévy processes; see the proof of Theorem 3.1. A similar approach may be found in Varopoulos [108]; see the proof of Lemma (iii) on page 246. In the present chapter we employ the strategy from therein. This result was proposed later in the full generality by the author. We refer to Theorem 3.1 in [50].

A partial result of the mentioned identity is the equality between the time derivative of a p-th power of a p-norm of $P_t f$ on one side and the Sobolev–Bregman form of $P_t f$ times $-p$ on the other side; see Proposition 3.4 below. This relation illustrates how p-form captures the evolution of the L^p-norm of a function $P_t f$. This connection was utilized to investigate the contractivity of the Feynman–Kac semigroup generated by perturbed fractional Laplacian in Bogdan, Jakubowski, Lenczewska, and Pietruska-Pałuba [16]; see equation (61). A similar result, with p-form defined in terms of the generator of the semigroup (compare with (2.28)), was known much earlier. We refer to the proof of Lemma on page 246 in [108]; see equation (1.1) therein. See also Lemma 2.2.2 in Davies [36] for the context of the logarithmic Sobolev inequalities.

The main identity of this chapter simplifies under the following condition.

Assumption 3.1 (Strong Stability)**.**

(SS) $\lim_{T\to+\infty} \|P_T f\|_p = 0$ *for every* $f \in L^p(m)$.

The above assumption will be necessary in some of the Littlewood–Paley estimates derived in Chapter 8. It will also be crucial in the study of the polarized Hardy–Stein identity in Chapter 5. For a more explicit form of the Hardy–Stein identity without the strong stability assumption, we refer to the discussion in Appendix A of [50].

Now, we present the main result of this chapter.

Theorem 3.2 (Hardy–Stein identity)**.** *Let* $1 < p < \infty$*. For any* $f \in L^p(m)$ *the following identity holds:*

$$\int_E |f(x)|^p \, m(\mathrm{d}x) - \lim_{T\to+\infty} \|P_T f\|_p^p = p \int_0^{+\infty} \mathcal{E}_p[P_t f] \, \mathrm{d}t. \tag{3.1}$$

Remark 3.3. *Under Assumption 3.1, the identity* (3.2) *reads*

$$\int_E |f(x)|^p \, m(\mathrm{d}x) = p \int_0^{+\infty} \mathcal{E}_p[P_t f] \, \mathrm{d}t. \tag{3.2}$$

Before we prove this theorem, we need the following fact. It is in some sense the differential form of the Hardy–Stein identity.

Proposition 3.4. *Let $1 < p < \infty$ and $f \in L^p(m)$. Then, for $t > 0$*

$$\frac{\mathrm{d}}{\mathrm{d}t}\|P_t f\|_p^p = -p\mathcal{E}_p[P_t f]. \tag{3.3}$$

Moreover, the above quantity is a continuous function with respect to $t > 0$. Additionally, when $f \in \mathcal{D}(A_p)$, then also

$$\frac{\mathrm{d}}{\mathrm{d}t}\|P_t f\|_p^p\Big|_{t=0} = -p\mathcal{E}_p[f] \tag{3.4}$$

and $[0,+\infty) \ni t \mapsto -p\mathcal{E}_p[P_t f] \in (-\infty, 0]$ is continuous.

Proof. Let $u(t) := P_t f$, $t \geq 0$. By Corollary 2.2(i) the mapping $|u(t)|^p$ is C^1 on $(0,+\infty)$ with values in $L^1(m)$. If additionally $f \in \mathcal{D}(A_p)$, then $|u(t)|^p$ is also C^1 on $[0,+\infty)$. In either case

$$(|u(t)|^p)' = pu(t)^{\langle p-1\rangle}A_p u(t).$$

Moreover, since the mapping $L^1(m) \ni g \mapsto \int_E g\,\mathrm{d}m$ is a continuous functional, we may write

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\int_E |u(t)|^p\,\mathrm{d}m &= \int_E (|u(t)|^p)'\,\mathrm{d}m = \int_E pu(t)^{\langle p-1\rangle}A_p u(t)\,\mathrm{d}m \\ &= p\langle A_p u(t), (u(t))^{\langle p-1\rangle}\rangle = -p\mathcal{E}_p[u(t)].\end{aligned}$$

In the last equality we use (2.28). The continuity of $t \mapsto \int_E (|u(t)|^p)'\,\mathrm{d}m$ follows from the continuity of $t \mapsto (|u(t)|^p)' \in L^1(m)$ and $L^1(m) \ni g \mapsto \int_E g\,\mathrm{d}m$. The proof is complete. □

Proof of Theorem 3.2. Let us first assume that $f \in \mathcal{D}(A_p)$. Fix $T > 0$. Integrating both sides of (3.3) from 0 to T we get

$$\|f\|_p^p - \|P_T f\|_p^p = p\int_0^T \mathcal{E}_p[P_t f]\,\mathrm{d}t.$$

Here, we used the fact that $\mathcal{E}_p[P_t f]$ is a continuous function with respect to $t \in [0, T]$, by Proposition 3.4. Since $\mathcal{E}_p[P_t f] \geq 0$, we may pass to the limit as T goes to infinity and obtain

$$\|f\|_p^p - \lim_{T\to+\infty}\|P_T f\|_p^p = p\int_0^{+\infty} \mathcal{E}_p[P_t f]\,\mathrm{d}t. \tag{3.5}$$

In particular, the above limit exists. Indeed, since $(P_t)_{t\geq 0}$ enjoys the contraction property, $\|P_T f\|_p$ is non-increasing as a function of T.

To show the statement for an arbitrary $f \in L^p(m)$, we apply (3.5) for $P_s f$ for some $s > 0$ and get

$$\|P_s f\|_p^p - \lim_{T\to+\infty}\|P_T f\|_p^p = p\int_s^{+\infty} \mathcal{E}_p[P_t f]\,\mathrm{d}t.$$

We may do that because $P_s f \in \mathcal{D}(A_p)$ by the analyticity of $(P_t)_{t\geq 0}$ when $1 < p < \infty$. Since $(P_t)_{t\geq 0}$ is also strongly continuous, $\|P_s f\|_p^p \to \|f\|_p^p$ as $s \to 0^+$. We know that $\mathcal{E}_p[P_t f] \geq 0$, thus applying the monotone convergence theorem, the right-hand side of the above equation converges to the right-hand side of the desired equation (3.1). □

Chapter 4

Sobolev–Bregman form with domination property

In this chapter, we derive an explicit formula for the Sobolev–Bregman form, and so an explicit Hardy–Stein identity, under additional assumptions about jumping and killing measures and their relation to the measure $\frac{1}{t}P_t(\mathrm{d}x, \mathrm{d}y)$. To be precise, we assume the following conditions.

Assumption 4.1 (Domination of the jumping measure)**.**

(PT) *There exists the transition density p_t of the semigroup $(P_t)_{t\geq 0}$, i.e., for every $t > 0$ there exists a measurable function $p_t\colon E \times E \to [0, +\infty)$ such that*

$$P_t(x, \mathrm{d}y) = p_t(x, y)m(\mathrm{d}y), \qquad x \in E.$$

(J) *The jumping measure J is absolute continuous, i.e., there exists a measurable function $J\colon E \times E \setminus \mathrm{diag} \to [0, +\infty)$ such that*

$$J(\mathrm{d}x, \mathrm{d}y) = J(x, y)m(\mathrm{d}x)m(\mathrm{d}y), \qquad (x, y) \in E \times E \setminus \mathrm{diag}.$$

(J1) *There exists a constant $c > 0$ such that*

$$\frac{1}{t}p_t(x, y) \leq cJ(x, y), \qquad t > 0,\ (x, y) \in E \times E \setminus \mathrm{diag}.$$

(J2)

$$\frac{1}{t}p_t(x, y) \to J(x, y), \qquad \textit{when } t \to 0^+ \textit{ for almost all } (x, y) \in E \times E \setminus \mathrm{diag}.$$

Assumption 4.2 (Domination of killing measure)**.**

(K) *The killing measure k is absolute continuous, i.e., there exists a measurable function $k\colon E \to [0, +\infty)$ such that*

$$k(\mathrm{d}x) = k(x)m(\mathrm{d}x), \qquad x \in E.$$

(K1) *There exists a constant $c > 0$ such that*

$$\frac{1}{t}(1 - P_t 1(x)) \leq ck(x), \qquad t > 0,\ x \in E.$$

(K2)

$$\frac{1}{t}(1 - P_t 1(x)) \to k(x), \qquad \text{when } t \to 0^+ \text{ for almost all } x \in E.$$

Recall that J and k are the unique Radon measures from the Beurling–Deny formula. For the details; see Section 2.4. In particular, under the above assumptions about J, the strongly local part $\mathcal{E}^c$ vanishes, i.e., $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$ is a pure-jump Dirichlet form. Moreover, in this configuration the Dirichlet form is maximally defined. The above statement will be shown later in Propositions 4.4 and 4.5, respectively.

Assumption 4.1 is a stronger statement than (2.26). It allows us to apply the dominated convergence theorem and Fatou's lemma to show the convergence of approximate form $\mathcal{E}^{(t)}(u, u^{\langle p-1 \rangle})$. In fact, under these assumptions the following lemma in available to us. This approach was employed in Lemma 6 in [16] and later in Lemma 2.11 in [15]. See also Remark 7 in [12].

Lemma 4.3. *Let f, f_n be non-negative measurable functions on some measure space (X, Σ, μ). If there exists a constant $c > 0$ such that $f_n \le cf$ for almost all $n \in \mathbb{N}$, and $f = \lim_{n\to\infty} f_n$ almost everywhere, then $\lim_{n\to+\infty} \int_X f_n \,\mathrm{d}\mu = \int_X f \,\mathrm{d}\mu$.*

This simple auxiliary fact was observed in [16]. For completeness, we present a proof below.

Proof. When the integral $\int_X f \,\mathrm{d}\mu$ is finite, then the statement follows from the dominance convergence theorem.

Otherwise, we may use Fatou's lemma and write

$$+\infty = \int_X f \,\mathrm{d}\mu \le \liminf_{n\to+\infty} \int_X f_n \,\mathrm{d}\mu.$$

Hence $\lim_{n\to+\infty} \int_X f_n \,\mathrm{d}\mu = +\infty$. □

Proposition 4.4. *Under Assumption 4.1 the Dirichlet form $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$ is pure-jump.*

Proof. It is enough to show $\mathcal{E}^c(u, v) = 0$ for $u = v$. For general u, v, one may then use polarization. According to Lemma 4.3, by (J1) and (J2), we obtain

$$\begin{aligned} \lim_{t\to 0^+} \frac{1}{2t} \iint\limits_{E\times E} (u(y) - u(x))^2 \, p_t(x, y) m(\mathrm{d}x) m(\mathrm{d}y) = \\ \frac{1}{2} \iint\limits_{E\times E \setminus \mathrm{diag}} (u(y) - u(x))^2 \, J(x, y) m(\mathrm{d}x) m(\mathrm{d}y), \end{aligned} \tag{4.1}$$

for all $u \in L^2(m)$. Moreover, when $u \in \mathcal{D}(\mathcal{E})$, the above limit is finite and equal to the right-hand side of (2.25). Comparing (4.1) with (2.25), we conclude that $\mathcal{E}^c = 0$. □

Proposition 4.5. *Under Assumptions 4.1 and 4.2 the Dirichlet form $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$ is maximally defined.*

Proof. By (J1) and (J2), we have (4.1). Similarly, according to Lemma 4.3, we have

$$\lim_{t\to 0^+} \frac{1}{t} \int\limits_E u(x) 2(1 - P_t 1(x)), m(\mathrm{d}x) = \int\limits_E u(x)^2 \, k(x) m(\mathrm{d}x), \tag{4.2}$$

for all $u \in L^2(m)$, whenever (K1) and (K2) hold.

Assume that the right-hand sides of (4.1) and (4.2) are finite for some u in $L^2(m)$. Then $\lim_{t\to0^+} \mathcal{E}^{(t)}[u]$ is finite as a sum of the left-hand sides of (4.1) and (4.2). By characterization (2.12), $u \in \mathcal{D}(\mathcal{E})$.

□

The main result of this chapter is the following.

Theorem 4.6. *Under Assumptions 4.1 and 4.2 there is the following characterization of the domain of the Sobolev–Bregman form:*

$$\mathcal{D}(\mathcal{E}_p) = \{\, u \in L^p(m) : u^{\langle p/2\rangle} \in \mathcal{D}(\mathcal{E}) \,\}. \tag{4.3}$$

In fact, the following estimate holds for any $u \in \mathcal{D}(\mathcal{E}_p)$:

$$\frac{4(p-1)}{p^2}\mathcal{E}[u^{\langle p/2\rangle}] \le \mathcal{E}_p[u] \le C_p\mathcal{E}[u^{\langle p/2\rangle}]. \tag{4.4}$$

Here, $C_p \ge 1$ is the constant from Lemma 2.3. Moreover, p-form $\mathcal{E}_p$ has the following formula for any $u \in \mathcal{D}(\mathcal{E}_p)$:

$$\mathcal{E}_p[u] = \frac{1}{p}\iint\limits_{E\times E\setminus\mathrm{diag}} F_p(u(x),u(y))\,J(x,y)m(\mathrm{d}x)m(\mathrm{d}y) + \int\limits_E |u(x)|^p\,k(x)m(\mathrm{d}x). \tag{4.5}$$

Here, F_p is the Bregman divergence given by (2.29). *The domain $\mathcal{D}(\mathcal{E}_p)$ consists of exactly these functions $u \in L^p(m)$, for which the right-hand side of* (4.5) *is finite.*

Remark 4.7. *The decomposition* (4.5) *is implied by the following convergences of the corresponding parts of approximate form $\mathcal{E}^{(t)}(u, u^{\langle p-1\rangle})$ for an arbitrary $u \in L^p(m)$:*

$$\begin{aligned}\lim_{t\to0^+}\frac{1}{pt}\iint\limits_{E\times E} F_p(u(x),u(y))\,p_t(x,y)(\mathrm{d}x)m(\mathrm{d}y)& \\ = \frac{1}{p}\iint\limits_{E\times E\setminus\mathrm{diag}} F_p(u(x),u(y))\,J(x,y)m(\mathrm{d}x)m(\mathrm{d}y)&\end{aligned} \tag{4.6}$$

and

$$\lim_{t\to0^+}\frac{1}{t}\int\limits_E |u(x)|^p\,(1-P_t1(x))\,m(\mathrm{d}x) = \int\limits_E |u(x)|^p\,k(x)m(\mathrm{d}x). \tag{4.7}$$

Remark 4.8. *Since the jumping density J is symmetric, we may replace F_p in the formula* (4.5) *by the symmetrized Bregman divergence H_p given by* (2.30) *and write*

$$\begin{aligned}\mathcal{E}_p[u] = {}& \frac{1}{2}\iint\limits_{E\times E\setminus\mathrm{diag}} (u(y)-u(x))(u^{\langle p-1\rangle}(y)-u^{\langle p-1\rangle}(x))\,J(x,y)m(\mathrm{d}x)m(\mathrm{d}y) \\ &+ \int\limits_E |u(x)|^p\,k(x)m(\mathrm{d}x).\end{aligned}$$

Proof of Theorem 4.6. Under (J1), (J2), (K1), (K2), we may show formulas (4.6) and (4.7) in the same way as (4.1) and (4.2), by Lemma 4.3. In particular, for any $u \in L^p(m)$ limits $\lim_{t\to0^+}\mathcal{E}^{(t)}(u,u^{\langle p-1\rangle})$ and $\lim_{t\to0^+}\mathcal{E}^{(t)}[u^{\langle p/2\rangle}]$ always exist, possibly infinite. Therefore, (4.3) and (4.4) are clear from (2.37).

Decomposition (4.5) follows immediately from (4.6) and (4.7). □

Combining the above result with Theorem 3.2, the following immediately conclusion holds.

Corollary 4.9 (Hardy–Stein identity). *Let $1 < p < \infty$. Under Assumptions 4.1 and 4.2, for any $f \in L^p(m)$, the following identity holds:*

$$\begin{aligned}\int_E |f(x)|^p \, m(\mathrm{d}x) - \lim_{T\to+\infty} \|P_T f\|_p^p \\ = \int_0^{+\infty} \iint_{E\times E\setminus \mathrm{diag}} F_p(P_t f(x), P_t f(y)) \, J(x,y) m(\mathrm{d}x) m(\mathrm{d}y) \mathrm{d}t \\ + p \int_0^{+\infty} \int_E |u(x)|^p \, k(x) m(\mathrm{d}x) \mathrm{d}t.\end{aligned} \tag{4.8}$$

When we assume in addition the strong stability from Assumption 3.1, then we may rewrite the above formula to the following form:

$$\begin{aligned}\int_E |f(x)|^p \, m(\mathrm{d}x) = \int_0^{+\infty} \iint_{E\times E\setminus \mathrm{diag}} F_p(P_t f(x), P_t f(y)) \, J(x,y) m(\mathrm{d}x) m(\mathrm{d}y) \mathrm{d}t \\ + p \int_0^{+\infty} \int_E |u(x)|^p \, k(x) m(\mathrm{d}x) \mathrm{d}t.\end{aligned} \tag{4.9}$$

Chapter 5

Polarized Hardy–Stein identity and Sobolev–Bregman form

The aim of this chapter is to derive the polarized version of the Hardy–Stein identity for $2 \leq p < \infty$. By this, we mean the identity with the integral $\int_E |f|^p \, \mathrm{d}m$ on the right-hand side replaced by $\int_E f g^{\langle p-1 \rangle} \, \mathrm{d}m$. This main result is presented in Theorem 5.6. The approach makes extensive use of the convexity of certain functions on $\mathbb{R}^2$ that are studied in Appendix B. The proof can be carried out much more easily in the case of $3 \leq p < \infty$. We refer to Appendix D in Bogdan, Gutowski, and Pietruska-Pałuba [15]. Because of difficulties connected with the local and the killing parts of regular Dirichlet form, we focus only on pure-jump forms. We propose also a polarized version of the Sobolev–Bregman form in this configuration in Section 5.4. We believe that this notion can be employed in the study of variational problems of the p-form $\mathcal{E}_p[\cdot]$. As a partial result, in Section 5.2, we introduce a multidimensional version of non-polarized Hardy–Stein identity.

This chapter is based on the joint work [15] with Bogdan and Pietruska-Pałuba. In this work, the results were presented for the symmetric Lévy processes that satisfy the assumptions of this chapter. To maintain compatibility with other chapters, the author presents these results here in a more general setting, with essentially the same proofs.

5.1 Preliminaries

We introduce the vector counterpart of the French power:

$$z^{\langle \gamma \rangle} := |z|^{\gamma-1} z,$$

where $z \in \mathbb{R}^n$ and $\gamma \in \mathbb{R}$ are such that the above expression makes sense. Here, $|\cdot|$ is the Euclidean norm in $\mathbb{R}^n$.

Note that

$$\nabla |z|^\gamma = \gamma z^{\langle \gamma-1 \rangle}, \quad \text{if either } z \in \mathbb{R}^n,\ \gamma > 1 \text{ or } z \in \mathbb{R}^n \setminus \{0\},\ \gamma \in \mathbb{R}. \tag{5.1}$$

Fix $1 < p < \infty$. We define the vector version of the Bregman divergence introduced in Section 2.5. Let $\mathcal{F}_p \colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ be given by

$$\mathcal{F}_p(w, z) := |z|^p - |w|^p - p w^{\langle p-1 \rangle} \cdot (z - w), \tag{5.2}$$

and define its symmetrization by

$$\mathcal{H}_p(w,z) := \frac{1}{2}\left(\mathcal{F}_p(w,z) + \mathcal{F}_p(z,w)\right) = \frac{p}{2}(z-w)\cdot\left(z^{\langle p-1\rangle} - w^{\langle p-1\rangle}\right). \tag{5.3}$$

Recall that by $z \cdot w$ we denote the dot product of z and w on $\mathbb{R}^n$. Note that $\mathcal{F}_p$ is the second-order Taylor remainder of the convex function $\mathbb{R}^n \ni z \mapsto |z|^p \in \mathbb{R}$, hence $\mathcal{F}_p$ and $\mathcal{H}_p$ are non-negative. For instance,

$$\mathcal{F}_2(w,z) = |z-w|^2. \tag{5.4}$$

The function $\mathcal{F}_p$ satisfies the following property. If Q is an $n \times n$ orthogonal matrix, then

$$\mathcal{F}_p(Qw, Qz) = \mathcal{F}_p(w,z). \tag{5.5}$$

We will also need the French power counterpart of the above notions. Let $\gamma > 1$. Denote the Jacobi matrix for the function $\mathbb{R}^n \ni z \mapsto z^{\langle\gamma\rangle} \in \mathbb{R}^n$ at a point z by $J_{\langle\gamma\rangle}(z)$. Then

$$J_{\langle\gamma\rangle}(z) = |z|^{\gamma-1}\left((\gamma-1)\left(\frac{z}{|z|}\otimes\frac{z}{|z|}\right) + I\right), \qquad z \neq 0, \tag{5.6}$$

and $J_{\langle\gamma\rangle}(0) = 0$. Here, I is the identity matrix. In particular, when $n = 1$, then $J_{\langle\gamma\rangle}(z) = \gamma\,|z|^{\gamma-1}$.

Now we may introduce the second-order Taylor remainder of $\mathbb{R}^n \ni z \mapsto z^{\langle\gamma\rangle} \in \mathbb{R}^n$: function $\mathcal{F}_{\langle\gamma\rangle}\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ given by

$$\mathcal{F}_{\langle\gamma\rangle}(w,z) := z^{\langle\gamma\rangle} - w^{\langle\gamma\rangle} - J_{\langle\gamma\rangle}(w)(z-w), \qquad w, z \in \mathbb{R}^n. \tag{5.7}$$

For a fixed positive integer n, we denote by $(L^p(m))^n$ the Banach space of elements of the form $f = (f_1, \ldots, f_n)$, where $f_1, \ldots, f_n \in L^p(m)$, equipped with the norm

$$\|f\|_{L^p} := \left(\int_E |f(x)|^p\, m(\mathrm{d}x)\right)^{1/p}.$$

Recall that $|\cdot|$ is the Euclidean norm in $\mathbb{R}^n$. More details about the calculus on the space $(L^p(m))^n$ are presented in Appendix A.

Let $f = (f_1, \ldots, f_n) \in (L^p(m))^n$. We use the following notation for operators acting on $(L^p(m))^n$:

$$P_t f := (P_t f_1, \ldots, P_t f_n), \quad t \geq 0, \tag{5.8}$$

and, if additionally $f_1, \ldots, f_n \in \mathcal{D}(A_p)$,

$$A_p f := (A_p f_1, \ldots, A_p f_n). \tag{5.9}$$

In particular, $P_t f, A_p f \in (L^p(m))^n$, for $t \geq 0$. We will also use the notation $f \in (\mathcal{D}(A_p))^n$, when $f_1, \ldots, f_n \in \mathcal{D}(A_p)$.

Through this chapter we work under the following assumptions about the semigroup $(P_t)_{t\geq 0}$. We impose Assumption 4.1. Thus, in this setting, the regular Dirichlet form $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$ is pure-jump; see Proposition 4.4. Moreover, we assume the *conservativeness* (another name: *mass conservation*), as follows:

Assumption 5.1 (Conservativeness).

(CS) $P_t 1 = 1$ *for all* $t \geq 0$.

Note that this condition is stronger than Assumption 4.2, because in this case the killing part of the regular Dirichlet form $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$ is identically equal to zero, hence the killing measure k is also zero, according to (2.24). The converse statement is not true in general, i.e., $k = 0$ may not imply the conservativeness if the corresponding Hunt process escapes to infinity in finite time. Moreover, equality (2.2) takes the following form:

$$\int_E P_t f(x)\, m(\mathrm{d}x) = \int_E f(x)\, m(\mathrm{d}x), \quad f \in L^1(m). \tag{5.10}$$

Most of the time, we assume also the strong stability of $(P_t)_{t\geq 0}$ from Assumption 3.1. This condition is helpful in dealing with the limit appearing on the left-hand side of the Hardy–Stein identity.

5.2 Multidimensional Hardy–Stein identity

In this section we present the vectorized version of the identity (4.8). The main result of this section is the following.

Theorem 5.2 (Multidimensional Hardy–Stein identity). *Impose Assumptions 4.1 and 5.1. Let* $1 < p < \infty$ *and* $f = (f_1, \ldots, f_n) \in (L^p(m))^n$. *Then,*

$$\int_E |f(x)|^p\, m(\mathrm{d}x) - \lim_{T\to+\infty} \|P_T f\|_{L^p}^p = \int_0^{+\infty} \iint_{E\times E\setminus \mathrm{diag}} \mathcal{F}_p(P_t f(x), P_t f(y))\, J(x,y) m(\mathrm{d}x) m(\mathrm{d}y) \mathrm{d}t. \tag{5.11}$$

Remark 5.3. *Under additional Assumption 3.1, the identity* (5.11) *reads*

$$\int_E |f(x)|^p\, m(\mathrm{d}x) = \int_0^{+\infty} \iint_{E\times E\setminus \mathrm{diag}} \mathcal{F}_p(P_t f(x), P_t f(y))\, J(x,y) m(\mathrm{d}x) m(\mathrm{d}y) \mathrm{d}t. \tag{5.12}$$

Let us briefly clarify this statement. Under the strong stability from Assumption 3.1, $P_T f_j$ converges to zero as $T \to +\infty$, for every $j = 1, \ldots, n$. Thus, $\|P_T f\|_{L^p} \to 0$, by (A.1) from Appendix A.

Remark 5.4. *Since the jumping density J is symmetric, we may replace $\mathcal{F}_p$ by its symmetrized version $\mathcal{H}_p$ given by* (5.3) *and write* (5.11) *as*

$$\begin{aligned}
&\int_E |f(x)|^p\, m(\mathrm{d}x) - \lim_{T\to+\infty} \|P_T f\|_{L^p}^p \\
&= \frac{p}{2} \int_0^{+\infty} \iint_{E\times E\setminus \mathrm{diag}} (P_t f(y) - P_t f(x)) \cdot ((P_t f(y))^{\langle p-1\rangle} - (P_t f(x))^{\langle p-1\rangle})\, J(x,y) m(\mathrm{d}x) m(\mathrm{d}y) \mathrm{d}t.
\end{aligned}$$

For an arbitrary $f \in (L^p(m))^n$, let $u(t) := P_t f$, $t \geq 0$. Since the semigroup $(P_t)_{t\geq 0}$ is analytic on $L^p(m)$ for $1 < p < \infty$, we get $u'(t) = A_p u(t)$ for $t > 0$, and also $u'(t) = P_t A_p f$ for $t \geq 0$, when $f \in (\mathcal{D}(A_p))^n$. Now, we may present the following multidimensional extension of Corollary 2.2, an immediate consequence of Proposition A.2.

Corollary 5.5. *Let $f \in (L^p(m))^n$ and $u(t) := P_t f$. Let $1 < \gamma \le p$.*

(i) The mapping $|u|^\gamma$ is C^1 on $(0, +\infty)$ with values in $L^{p/\gamma}(m)$ and

$$(|u(t)|^\gamma)' = \gamma u(t)^{\langle \gamma-1\rangle} \cdot u'(t) = \gamma u(t)^{\langle \gamma-1\rangle} \cdot A_p P_t f, \quad t \ge 0. \tag{5.13}$$

(ii) The mapping $u^{\langle\gamma\rangle}$ is C^1 on $(0, +\infty)$ with values in $(L^{p/\gamma}(m))^n$ and

$$(u(t)^{\langle\gamma\rangle})' = \Big(J_{\langle\gamma\rangle} \circ u(t)\Big) A_p u(t) = \Big(J_{\langle\gamma\rangle} \circ u(t)\Big) A_p P_t f, \quad t \ge 0. \tag{5.14}$$

Here, $J_{\langle\gamma\rangle}(z)$ is the Jacobi matrix for the function $\mathbb{R}^n \ni z \mapsto z^{\langle\gamma\rangle} \in \mathbb{R}^n$ at a point z given by (5.6). In addition, if $f \in (\mathcal{D}(A_p))^n$, then $|u|^\gamma$ and $u^{\langle\gamma\rangle}$ are C^1 on $[0, +\infty)$ with values in $L^{p/\gamma}(m)$ and $(L^{p/\gamma}(m))^n$, respectively.

Now, we are ready to present the proof of the main result of this section. The approach is a combination of the methods used in the proof of the general form of the Hardy–Stein identity in Theorems 3.2 and 4.6. The essential tool in this proof is Lemma 4.3 based on the dominated convergence theorem and Fatou's lemma.

Proof of Theorem 5.2. Assume first that $f \in (\mathcal{D}(A_p))^n$ and denote $u(t) := P_t f$. Fix $T > 0$. Then, by Corollary 5.5(i) we know that the mapping $|u(t)|^p$ is C^1 on $[0, T]$ with values in $L^1(m)$ and

$$(|u(t)|^p)' = p u(t)^{\langle p-1\rangle} \cdot A_p u(t), \quad t \in [0, T].$$

Then, since the mapping $L^1(m) \ni g \mapsto \int_E g \,\mathrm{d}m$ is a continuous functional, we have

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t} \int_E |u(t)|^p \,\mathrm{d}m &= \int_E (|u(t)|^p)' \,\mathrm{d}m = \langle A_p u(t), p(u(t))^{\langle p-1\rangle}\rangle \\ &= \lim_{h\to 0^+} \frac{1}{h} \langle P_t u(t) - u(t), p(u(t))^{\langle p-1\rangle}\rangle.\end{aligned}$$

Note that here we use the notation defined in (2.1).

Similarly as in (2.35), we may rewrite the quantity under the limit to obtain

$$\begin{aligned}\frac{1}{h}\langle P_t u(t) - u(t), p(u(t))^{\langle p-1\rangle}\rangle &= \frac{1}{h} \iint_{E\times E} p(u(t)(x))^{\langle p-1\rangle} \cdot (u(t)(y) - u(t)(x))\, P_h(\mathrm{d}x, \mathrm{d}y) \\ &= -\frac{1}{h} \int_E P_h(|u(t)|^p)(x)\, m(\mathrm{d}x) + \frac{1}{h} \int_E |u(t)(x)|^p \, m(\mathrm{d}x) \\ &\quad + \frac{1}{h} \iint_{E\times E} p(u(t)(x))^{\langle p-1\rangle} \cdot (u(t)(y) - u(t)(x))\, P_h(\mathrm{d}x, \mathrm{d}y) \\ &= -\frac{1}{h} \iint_{E\times E} \mathcal{F}_p(u(t)(x), u(t)(y))\, P_h(\mathrm{d}x, \mathrm{d}y).\end{aligned}$$

Here, we used (5.10). Under Assumption 4.1, we may use Lemma 4.3 to obtain

$$\lim_{h\to 0^+} \frac{1}{h} \iint_{E\times E} \mathcal{F}_p(u(t)(x), u(t)(y))\, P_h(\mathrm{d}x, \mathrm{d}y) = \iint_{E\times E\setminus \mathrm{diag}} \mathcal{F}_p(u(t)(x), u(t)(y))\, J(x, y) m(\mathrm{d}x) m(\mathrm{d}y).$$

Here, we utilize the fact that the function $\mathcal{F}_p$ is non-negative. Summarizing,

$$\frac{\mathrm{d}}{\mathrm{d}t}\int\limits_E |u(t)|^p \,\mathrm{d}m = -\iint\limits_{E\times E\setminus\mathrm{diag}} \mathcal{F}_p(u(t)(x), u(t)(y))\, J(x,y)m(\mathrm{d}x)m(\mathrm{d}y). \tag{5.15}$$

Note that this equality is the multidimensional version of (3.3).

As in the proof of Proposition 3.4, the quantity in (5.15) is a continuous function with respect to $t \in [0, T]$ as a composition of continuous mappings $t \mapsto (|u(t)|^p)' \in L^1(m)$ and $L^1(m) \ni g \mapsto \int_E g\,\mathrm{d}m$.

Therefore, we may use the fundamental theorem of calculus, integrate both sides of (5.15) from 0 to T and obtain

$$\int\limits_E |f(x)|^p\, m(\mathrm{d}x) - \|P_T f\|^p_{L^p} = \int\limits_0^T \iint\limits_{E\times E\setminus\mathrm{diag}} \mathcal{F}_p(P_t f(x), P_t f(y))\, J(x,y)m(\mathrm{d}x)m(\mathrm{d}y)\mathrm{d}t.$$

Let $T \to +\infty$. Note that the function $\mathcal{F}_p$ is non-negative. Therefore, by the monotone convergence theorem the integral over $[0, T]$ on the right-hand side converges to the integral over $[0, +\infty)$. In addition, since $(P_t)_{t\geq 0}$ is the semigroup of contractions on $L^p(m)$, norms $\|P_T f_j\|_p$ are non-increasing, the limit of $\|P_T f_j\|_p$ exists as $T \to +\infty$ for every $j = 1, \ldots, n$. Thus, the limit of $\|P_T f\|_{L^p}$ exists as well. We have proved (5.11) for $f \in (\mathcal{D}(A_p))^n$.

For an arbitrary $f \in (L^p(m))^n$, let $s > 0$ and apply (5.11) to $P_s f$ to obtain

$$\int\limits_E |P_s f|^p \,\mathrm{d}m - \lim_{T\to+\infty} \|P_T f\|^p_{L^p} = \int\limits_s^{+\infty} \iint\limits_{E\times E\setminus\mathrm{diag}} \mathcal{F}_p(P_t f(x), P_t f(y))\, J(x,y)m(\mathrm{d}x)m(\mathrm{d}y)\mathrm{d}t.$$

We may do that because $P_s f \in (\mathcal{D}(A_p))^n$, since the semigroup $(P_t)_{t\geq 0}$ is analytic on $L^p(m)$ when $1 < p < \infty$. When $s \to 0^+$, then $\int_E |P_s f|^p \,\mathrm{d}m$ converges to $\int_E |f|^p \,\mathrm{d}m$ because of the strong continuity of the semigroup $(P_t)_{t\geq 0}$ and (A.1) from Appendix A. Of course, since the integrand is non-negative, by the monotone convergence theorem, the right-hand side converges to the desired right-hand side of (5.11). The proof is complete. □

5.3 Polarized Hardy–Stein identity

The polarized Hardy–Stein identity is the result of applying a similar approach as in Chapter 4 and in the previous section of this chapter, but utilized for $\int_E f g^{\langle p-1\rangle}\,\mathrm{d}m$ instead for $\int_E |f|^p \,\mathrm{d}m$, where $2 \leq p < \infty$ and $f, g \in L^p(m)$. In other words, it is an identity for $[f, g]\,\|g\|_p^{p-2}$ instead of $\|f\|_p^p$, where $[f, g] := \|g\|_p^{-(p-2)} \int_E f g^{\langle p-1\rangle}\,\mathrm{d}m$ is the semi-inner product on $L^p(m)$. In this case, the proof is based on a study of the decay of $\int_E P_t f (P_t g)^{\langle p-1\rangle}\,\mathrm{d}m$ with respect to t. Compare this with Proposition 3.4.

To recover the polarized analog of the Bregman divergence F_p, the integrand of the right-hand side of the non-polarized Hardy–Stein identity, we introduce the following function $\mathcal{J}_p\colon \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$:

$$\begin{aligned}\mathcal{J}_p(w, z) := {} & z_1 z_2^{\langle p-1\rangle} - w_1 w_2^{\langle p-1\rangle} \\ & - w_2^{\langle p-1\rangle}(z_1 - w_1) - (p-1) w_1 |w_2|^{p-2}(z_2 - w_2),\end{aligned} \tag{5.16}$$

where $w = (w_1, w_2)$, $z = (z_1, z_2)$ are vectors in $\mathbb{R}^2$. This mapping is the second-order Taylor remainder of the function $\mathbb{R}^2 \ni (z_1, z_2) \mapsto z_1 z_2^{\langle p-1\rangle} \in \mathbb{R}$. In contrast to the non-polarized case, unfortunately, this function is not convex, even on $[0,+\infty)^2$, therefore the remainder $\mathcal{J}_p$ is not non-negative.

We are ready to present the main result of this chapter.

Theorem 5.6 (Polarized Hardy–Stein identity)**.** *Suppose that Assumptions 3.1, 4.1, and 5.1 are satisfied. Let $2 \le p < \infty$. For $f, g \in L^p(m)$, denote $\Phi := (f, g)$. Then,*

$$\int_E f(x) g^{\langle p-1\rangle}(x)\, m(\mathrm{d}x) = \int_0^{+\infty} \iint_{E\times E\setminus \mathrm{diag}} \mathcal{J}_p(P_t\Phi(x), P_t\Phi(y))\, J(x,y) m(\mathrm{d}x) m(\mathrm{d}y) \mathrm{d}t, \tag{5.17}$$

where $\mathcal{J}_p$ is given by (5.16)*. In particular, the signed integral on the right-hand side is absolutely convergent.*

Note that, in particular, if $w_1 = w_2 =: a$ and $z_1 = z_2 =: b$, then $\mathcal{J}_p(w,z) = F_p(a,b)$, so (5.17) with $f = g$ agrees with the non-polarized Hardy–Stein identity (4.9), at least for $p \ge 2$.

Remark 5.7. *For $p = 2$, the identity* (5.17) *reads*

$$\int_E f(x) g(x)\, m(\mathrm{d}x) = \int_0^{+\infty} \iint_{E\times E\setminus \mathrm{diag}} (P_t f(y) - P_t f(x))(P_t g(y) - P_t g(x))\, J(x,y) m(\mathrm{d}x) m(\mathrm{d}y) \mathrm{d}t. \tag{5.18}$$

Note that

$$\mathcal{J}_2(w,z) = z_1 z_2 - w_1 w_2 - w_2(z_1 - w_1) - w_1(z_2 - w_2) = (z_1 - w_1)(z_2 - w_2).$$

In this case (5.17) *is obtained by polarization from the non-polarized Hardy–Stein identity* (4.8)*. Therefore, further we only need to consider the case $p > 2$.*

As we mentioned before, since $\mathbb{R}^2 \ni (z_1, z_2) \mapsto z_1 z_2^{\langle p-1\rangle} \in \mathbb{R}$ is not convex, the integrand $\mathcal{J}_p$ is not non-negative. Therefore, we cannot apply Lemma 4.3 based on the dominated convergence theorem or Fatou's lemma straightforwardly.

Indeed, to see the non-convexity, we calculate the gradient

$$\nabla \left(z_1 z_2^{\langle p-1\rangle} \right) = \begin{bmatrix} z_2^{\langle p-1\rangle} \\ (p-1) z_1 |z_2|^{p-2} \end{bmatrix},$$

and the Hessian matrix

$$\nabla^2 \left(z_1 z_2^{\langle p-1\rangle} \right) = \begin{bmatrix} 0 & (p-1)|z_2|^{p-2} \\ (p-1)|z_2|^{p-2} & (p-1)(p-2) z_1 z_2^{\langle p-3\rangle} \end{bmatrix}. \tag{5.19}$$

Therefore

$$\det \nabla^2 \left(z_1 z_2^{\langle p-1\rangle} \right) = -(p-1)^2 |z_2|^{2p-4} < 0.$$

To overcome this difficulty, we will decompose the function $\mathbb{R}^2 \ni (z_1, z_2) \mapsto z_1 z_2^{\langle p-1\rangle} \in \mathbb{R}$ into a difference of the following convex functions. We recall that for $a \in \mathbb{R}$ we denote $a_+ := a \vee 0$ and $a_- := (-a) \vee 0$. Then, we introduce the functions:

$$Y^{(+)}(z) := z_1 \left((z_2)_+\right)^{p-1} + |z|^p,$$
$$Y^{(-)}(z) := z_1 \left((z_2)_-\right)^{p-1} + |z|^p, \quad z = (z_1, z_2) \in \mathbb{R}^2.$$

According to Lemma B.4, the above functions are convex on $[0,+\infty)\times\mathbb{R}$. We denote also

$$\begin{aligned}\mathcal{J}_p^{(+)}(w,z) &= z_1\left((z_2)_+\right)^{p-1} - w_1\left((w_2)_+\right)^{p-1} - \left((w_2)_+\right)^{p-1}(z_1-w_1)\\ &\quad - (p-1)w_1\left((w_2)_+\right)^{p-2}(z_2-w_2)\\ \mathcal{J}_p^{(-)}(w,z) &= z_1\left((z_2)_-\right)^{p-1} - w_1\left((w_2)_-\right)^{p-1} - \left((w_2)_-\right)^{p-1}(z_1-w_1)\\ &\quad + (p-1)w_1\left((w_2)_-\right)^{p-2}(z_2-w_2),\end{aligned}$$

and note that there are just the second-order Taylor remainders of mappings $\mathbb{R}^2\ni z\mapsto z_1\left((z_2)_+\right)^{p-1}$ and $\mathbb{R}^2\ni z\mapsto z_1\left((z_2)_-\right)^{p-1}$, respectively. In particular, when we denote $\bar z := (z_1,-z_2)$, then

$$\mathcal{J}_p^{(+)}(\bar w,\bar z) = \mathcal{J}_p^{(-)}(w,z). \tag{5.20}$$

Since $z_1((z_2)_+)^{p-1} - z_1((z_2)_-)^{p-1} = z_1 z_2^{\langle p-1\rangle}$, it follows that

$$\mathcal{J}_p = \mathcal{J}_p^{(+)} - \mathcal{J}_p^{(-)} = \left(\mathcal{J}_p^{(+)} + \mathcal{F}_p\right) - \left(\mathcal{J}_p^{(-)} + \mathcal{F}_p\right). \tag{5.21}$$

Here, $\mathcal{F}_p$ is given by (5.2) with $n=2$. Moreover, $\mathcal{J}_p^{(+)}+\mathcal{F}_p\ge 0$ and $\mathcal{J}_p^{(-)}+\mathcal{F}_p\ge 0$ on $([0,+\infty)\times\mathbb{R})^2$ because of the convexity of $Y^{(+)}$ and $Y^{(-)}$ on $[0,+\infty)\times\mathbb{R}$.

Before we prove Theorem 5.6, we show the preliminary result. Since $\mathcal{J}_p^{(+)}+\mathcal{F}_p\ge 0$ and $\mathcal{J}_p^{(-)}+\mathcal{F}_p\ge 0$ on $([0,+\infty)\times\mathbb{R})^2$ we may utilize the approach from the proof of Theorem 5.2.

Proposition 5.8. *Let $2<p<\infty$, $f,g\in L^p(m)$, $f\ge 0$ m-a.e., and $\Phi=(f,g)$. Under the assumptions of Theorem 5.6,*

$$\begin{aligned}&\int_E \left(f(g_+)^{p-1} + |\Phi|^p\right)\mathrm{d}m \\ &\quad= \int_0^{+\infty}\iint_{E\times E\setminus\mathrm{diag}} \left(\mathcal{J}_p^{(+)}+\mathcal{F}_p\right)\left(P_t\Phi(x),P_t\Phi(y)\right)J(x,y)m(\mathrm{d}x)m(\mathrm{d}y)\mathrm{d}t\end{aligned} \tag{5.22}$$

and

$$\begin{aligned}&\int_E \left(f(g_-)^{p-1} + |\Phi|^p\right)\mathrm{d}m \\ &\quad= \int_0^{+\infty}\iint_{E\times E\setminus\mathrm{diag}} \left(\mathcal{J}_p^{(-)}+\mathcal{F}_p\right)\left(P_t\Phi(x),P_t\Phi(y)\right)J(x,y)m(\mathrm{d}x)m(\mathrm{d}y)\mathrm{d}t.\end{aligned} \tag{5.23}$$

Proof. It is enough to show (5.22), because the identity (5.23) follows from (5.22) by substituting $-g$ in place of g and by using (5.5) and (5.20). Assume first that $f,g\in\mathcal{D}(A_p)$. Fix $T>0$. Denote $u(t):=P_tf$, $v(t):=P_tg$, and $w(t):=P_t\Phi$. By Corollary 2.2 with $\gamma=p-1>1$ the mapping $(v(t)_+)^{p-1} = (|v(t)|^{p-1}+v(t)^{\langle p-1\rangle})/2$ is C^1 on $[0,T]$ with values in $L^{p/(p-1)}(m)$ and

$$\begin{aligned}((v(t)_+)^{p-1})' &= \frac{1}{2}\left(|v(t)|^{p-1} + v(t)^{\langle p-1\rangle}\right)' \\ &= \frac{1}{2}\left((p-1)v(t)^{\langle p-2\rangle}A_pv(t) + (p-1)\,|v(t)|^{p-2}A_pv(t)\right) \\ &= (p-1)(v(t)_+)^{p-2}A_pv(t), \quad t\in[0,T].\end{aligned}$$

Also $u(t)$ is C^1 on $[0,T]$ with values in $L^p(m)$ and $u'(t) = A_p u(t)$. Finally, by the product rule from Proposition A.1, $u(t)(v(t)_+)^{p-1}$ is C^1 on $[0,T]$ with values in $L^1(m)$ and

$$(u(t)(v(t)_+)^{p-1})' = A_p u(t)(v(t)_+)^{p-1} + (p-1)u(t)(v(t)_+)^{p-2} A_p v(t), \quad t \in [0,T].$$

As in the proof of Theorem 5.11, the mapping $|w(t)|^p$ is also C^1 on $[0,T]$ with values in $L^1(m)$ and

$$(|w(t)|^p)' = p w(t)^{\langle p-1\rangle} \cdot A_p w(t), \quad t \in [0,T].$$

Since the mapping $L^1(m) \ni r \mapsto \int_E r \, \mathrm{d}m$ is a continuous functional, we have

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t} \int\limits_E \Big(u(t)(v(t)_+)^{p-1} + |w(t)|^p \Big) \, \mathrm{d}m &= \int\limits_E \Big(u(t)(v(t)_+)^{p-1} + |w(t)|^p \Big)' \, \mathrm{d}m = \\
&= \langle A_p u(t), (v(t)_+)^{p-1} \rangle + \langle A_p v(t), (p-1)u(t)(v(t)_+)^{p-2} \rangle \\
&\quad + \langle A_p w(t), p(w(t))^{\langle p-1\rangle} \rangle \\
&= \lim_{h\to 0^+} \Bigg[\frac{1}{h} \langle P_h u(t) - u(t), (v(t)_+)^{p-1} \rangle + \frac{1}{h} \langle P_h v(t) - v(t), (p-1)u(t)(v(t)_+)^{p-2} \rangle \\
&\quad + \frac{1}{h} \langle P_t w(t) - w(t), p(w(t))^{\langle p-1\rangle} \rangle \Bigg]. \qquad (5.24)
\end{aligned}$$

Moreover, the above quantity is a continuous function of $t \in [0,T]$. By Assumption 5.1 and (5.10)

$$\frac{1}{h} \int\limits_E \Big(u(t)(v(t)_+)^{p-1} + |w(t)|^p \Big) \, \mathrm{d}m = \frac{1}{h} \int\limits_E P_h \Big(u(t)(v(t)_+)^{p-1} + |w(t)|^p \Big) \, \mathrm{d}m.$$

Subtracting the right-hand side and adding the left-hand side of the above equation to (5.24), we obtain

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t} \int\limits_E \Big(u(t)(v(t)_+)^{p-1} + |w(t)|^p \Big) \, \mathrm{d}m \\
&\qquad = \lim_{h\to 0^+} \Bigg[-\frac{1}{h} \int\limits_E P_h \Big(u(t)(v(t)_+)^{p-1} + |w(t)|^p \Big)(x) \, m(\mathrm{d}x) \\
&\qquad + \frac{1}{h} \int\limits_E \Big(u(t)(x)(v(t)(x)_+)^{p-1} + |w(t)(x)|^p \Big) \, m(\mathrm{d}x) \\
&\qquad + \frac{1}{h} \iint\limits_{E\times E} (v(t)(x)_+)^{p-1} (u(t)(y) - u(t)(x)) \, P_h(\mathrm{d}x, \mathrm{d}y) \\
&\qquad + \frac{1}{h} \iint\limits_{E\times E} (p-1)u(t)(x)(v(t)(x)_+)^{p-2} (v(t)(y) - v(t)(x)) \, P_h(\mathrm{d}x, \mathrm{d}y) \\
&\qquad + \frac{1}{h} \iint\limits_{E\times E} p(w(t)(x))^{\langle p-1\rangle} \cdot (w(t)(y) - w(t)(y)) \, P_h(\mathrm{d}x, \mathrm{d}y) \Bigg] \\
&\qquad = -\lim_{h\to 0^+} \frac{1}{h} \iint\limits_{E\times E} \Big(\mathcal{J}_p^{(+)} + \mathcal{F}_p \Big) (w(t)(x), w(t)(y)) \, P_h(\mathrm{d}x, \mathrm{d}y).
\end{aligned}$$

As $f \geq 0$ m-a.e., we have also $u(t) \geq 0$ m-a.e. in view of the positivity preservation of P_t. Moreover, $\mathcal{J}_p^{(+)} + \mathcal{F}_p \geq 0$ on $([0,+\infty) \times \mathbb{R})^2$. Thus, under Assumption 4.1, we can apply Lemma 4.3 to derive

$$\lim_{h\to 0^+} \frac{1}{h} \iint\limits_{E\times E} \left(\mathcal{J}_p^{(+)} + \mathcal{F}_p\right)(w(t)(x), w(t)(y))\, P_h(\mathrm{d}x, \mathrm{d}y) = \iint\limits_{E\times E\setminus\mathrm{diag}} \left(\mathcal{J}_p^{(+)} + \mathcal{F}_p\right)(w(t)(x), w(t)(y))\, J(x,y) m(\mathrm{d}x) m(\mathrm{d}y).$$

Taking this into account, we may write

$$\frac{\mathrm{d}}{\mathrm{d}t} \int\limits_E \left(u(t)(v(t)_+)^{p-1} + |w(t)|^p\right) \mathrm{d}m = -\iint\limits_{E\times E\setminus\mathrm{diag}} \left(\mathcal{J}_p^{(+)} + \mathcal{F}_p\right)(w(t)(x), w(t)(y))\, J(x,y) m(\mathrm{d}x) m(\mathrm{d}y). \tag{5.25}$$

Utilizing the fundamental theorem of calculus, we integrate both sides of (5.25) from 0 to T and obtain

$$\int\limits_E \left(f(g_+)^{p-1} + |\Phi|^p\right) \mathrm{d}m - \int\limits_E P_T f((P_T g)_+)^{p-1}\, \mathrm{d}m - \|P_T\Phi\|_{L^p}^p = \int\limits_0^T \iint\limits_{E\times E\setminus\mathrm{diag}} \left(\mathcal{J}_p^{(+)} + \mathcal{F}_p\right)(P_t\Phi(x), P_t\Phi(y))\, J(x,y) m(\mathrm{d}x) m(\mathrm{d}y) \mathrm{d}t.$$

When T converges to infinity, the right-hand side converges to the integral over $[0,+\infty)$, because of the non-negativity of $\mathcal{J}_p^{(+)} + \mathcal{F}_p$ on $([0,+\infty) \times \mathbb{R})^2$. As in the proof of Remark 5.3, $\|P_T\Phi\|_{L^p}^p$ tends to zero by Assumption 3.1. By the same Assumption 3.1, $P_T f$ and $P_T g$ converge to zero in $L^p(m)$, therefore $\int_E P_T f((P_T g)_+)^{p-1}\, \mathrm{d}m \to 0$ by Hölder's inequality. The proof of (5.22) for $f, g \in \mathcal{D}(A_p)$ is complete.

To show the assertion for any $f, g \in L^p(m)$, $f \geq 0$, we apply (5.22) to $P_s f$ and $P_s g$ for some $s > 0$ and obtain

$$\int\limits_E \left(P_s f((P_s g)_+)^{p-1} + |P_s\Phi|^p\right) \mathrm{d}m = \int\limits_s^{+\infty} \iint\limits_{E\times E\setminus\mathrm{diag}} \left(\mathcal{J}_p^{(+)} + \mathcal{F}_p\right)(P_t\Phi(x), P_t\Phi(y))\, J(x,y) m(\mathrm{d}x) m(\mathrm{d}y) \mathrm{d}t.$$

We are allowed to do that because $P_s f, P_s g \in \mathcal{D}(A_p)$, since the semigroup $(P_t)_{t\geq 0}$ is analytic on $L^p(m)$ for $1 < p < \infty$.

We claim that the left-hand side converges to the left-hand side of the desired equation (5.22). It is sufficient to use appropriate lemmas from Appendix A. Indeed, by the strong continuity of the semigroup $(P_t)_{t\geq 0}$, $P_s f \to f$, $P_s g \to g$ in $L^p(m)$ as $s \to 0^+$. Thus, also $((P_s g)_+)^{p-1} \to (g_+)^{p-1}$ in $L^{p/(p-1)}(m)$ by Lemma A.4. Finally, $P_s f((P_s g)_+)^{p-1} \to f(g_+)^{p-1}$ in $L^1(m)$ by Lemma A.5. Similarly, $|P_s\Phi|^p \to |\Phi|^p$ in $L^1(m)$ according to Lemma A.4. Therefore, our claim follows from the continuity of the integral on $L^1(m)$.

Finally, since $\mathcal{J}_p^{(+)} + \mathcal{F}_p$ is non-negative on $([0,+\infty) \times \mathbb{R})^2$, due to the monotone convergence theorem, the integral over $[s,+\infty)$ on the right-hand side tends to the integral over $[0,+\infty)$. The proof is complete. □

Now, when Proposition 5.8 is proved, we may derive the main result of this section.

Proof of Theorem 5.6. According to Remark 5.7, we only need to consider $2 < p < \infty$. Assume first $f \geq 0$. Utilizing Proposition 5.8 and (5.21),

$$\begin{aligned}\int_E f g^{\langle p-1\rangle}\,\mathrm{d}m &= \int_E \left(f(g_+)^{p-1} + |\Phi|^p\right)\mathrm{d}m - \int_E \left(f(g_-)^{p-1} + |\Phi|^p\right)\mathrm{d}m \\ &= \int_0^{+\infty} \iint_{E\times E\setminus\mathrm{diag}} \left(\mathcal{J}_p^{(+)} + \mathcal{F}_p\right)(P_t\Phi(x), P_t\Phi(y))\, J(x,y)m(\mathrm{d}x)m(\mathrm{d}y)\mathrm{d}t \\ &\quad - \int_0^{+\infty} \iint_{E\times E\setminus\mathrm{diag}} \left(\mathcal{J}_p^{(-)} + \mathcal{F}_p\right)(P_t\Phi(x), P_t\Phi(y))\, J(x,y)m(\mathrm{d}x)m(\mathrm{d}y)\mathrm{d}t \\ &= \int_0^{+\infty} \iint_{E\times E\setminus\mathrm{diag}} \mathcal{J}_p(P_t\Phi(x), P_t\Phi(y))\, J(x,y)m(\mathrm{d}x)m(\mathrm{d}y)\mathrm{d}t.\end{aligned}$$

Note that, in particular, all integrals are absolutely convergent.

To relax the assumption $f \geq 0$, we consider an arbitrary $f \in L^p(m)$ and take the decomposition $f = f_+ - f_-$. At this moment we know that the identity holds for $\Phi^{(+)} := (f_+, g)$ and $\Phi^{(-)} := (f_-, g)$. Of course, $\Phi = \Phi^{(+)} - \Phi^{(-)}$. By the linearity of the function $\mathcal{J}_p(w,z)$ in w_1 and z_1, we obtain

$$\begin{aligned}\int_E f g^{\langle p-1\rangle}\,\mathrm{d}m &= \int_E f_+ g^{\langle p-1\rangle}\,\mathrm{d}m - \int_E f_- g^{\langle p-1\rangle}\,\mathrm{d}m \\ &= \int_0^{+\infty} \iint_{E\times E\setminus\mathrm{diag}} \mathcal{J}_p(P_t\Phi^{(+)}(x), P_t\Phi^{(+)}(y))\, J(x,y)m(\mathrm{d}x)m(\mathrm{d}y)\mathrm{d}t \\ &\quad - \int_0^{+\infty} \iint_{E\times E\setminus\mathrm{diag}} \mathcal{J}_p(P_t\Phi^{(-)}(x), P_t\Phi^{(-)}(y))\, J(x,y)m(\mathrm{d}x)m(\mathrm{d}y)\mathrm{d}t \\ &= \int_0^{+\infty} \iint_{E\times E\setminus\mathrm{diag}} \mathcal{J}_p(P_t\Phi(x), P_t\Phi(y))\, J(x,y)m(\mathrm{d}x)m(\mathrm{d}y)\mathrm{d}t.\end{aligned}$$

Here, all the integrals are absolutely convergent as well. □

At the end of this section, we present an explicit upper bound for the absolutely convergence of the integral from the right-hand side of the polarized Hardy–Stein identity.

Proposition 5.9. *Let $2 \leq p < \infty$, $f, g \in L^p(m)$, and $\Phi = (f,g)$. Under the assumptions of Theorem 5.6,*

$$\int_0^{+\infty} \iint_{E\times E\setminus\mathrm{diag}} |\mathcal{J}_p(P_t\Phi(x), P_t\Phi(y))|\; J(x,y)m(\mathrm{d}x)m(\mathrm{d}y)\mathrm{d}t \leq (3 + 2^{p/2})\,\|f\|_p\,\|g\|_p^{p-1}. \tag{5.26}$$

Proof. Following the approach used in the proof of Theorem 5.6, we denote $\Phi^{(+)} = (f_+, g)$ and $\Phi^{(-)} = (f_-, g)$. Then,

$$\mathcal{J}_p(P_t\Phi(x), P_t\Phi(y)) = \mathcal{J}_p(P_t\Phi^{(+)}(x), P_t\Phi^{(+)}(y)) - \mathcal{J}_p(P_t\Phi^{(-)}(x), P_t\Phi^{(-)}(y)),$$

hence

$$|\mathcal{J}_p(P_t\Phi(x), P_t\Phi(y))| \leq \left|\mathcal{J}_p(P_t\Phi^{(+)}(x), P_t\Phi^{(+)}(y))\right| + \left|\mathcal{J}_p(P_t\Phi^{(-)}(x), P_t\Phi^{(-)}(y))\right|. \quad (5.27)$$

According to (5.21), we get

$$|\mathcal{J}_p| \leq \left(\mathcal{J}_p^{(+)} + \mathcal{F}_p\right) + \left(\mathcal{J}_p^{(-)} + \mathcal{F}_p\right),$$

with both terms non-negative on $([0,+\infty)\times\mathbb{R})^2$. Since $\Phi^{(+)}(x), \Phi^{(-)}(x) \in ([0,+\infty)\times\mathbb{R})^2$, we have also $P_t\Phi^{(+)}(x), P_t\Phi^{(-)}(x) \in ([0,+\infty)\times\mathbb{R})^2$ in view of the positivity preservation of P_t. Summarizing,

$$\begin{aligned}\left|\mathcal{J}_p(P_t\Phi^{(+)}(x), P_t\Phi^{(+)}(y))\right| \leq &\left(\mathcal{J}_p^{(+)} + \mathcal{F}_p\right)(P_t\Phi^{(+)}(x), P_t\Phi^{(+)}(y)) \\ &+ \left(\mathcal{J}_p^{(-)} + \mathcal{F}_p\right)(P_t\Phi^{(+)}(x), P_t\Phi^{(+)}(y)),\end{aligned} \quad (5.28)$$

and

$$\begin{aligned}\left|\mathcal{J}_p(P_t\Phi^{(-)}(x), P_t\Phi^{(-)}(y))\right| \leq &\left(\mathcal{J}_p^{(+)} + \mathcal{F}_p\right)(P_t\Phi^{(-)}(x), P_t\Phi^{(-)}(y)) \\ &+ \left(\mathcal{J}_p^{(-)} + \mathcal{F}_p\right)(P_t\Phi^{(-)}(x), P_t\Phi^{(-)}(y)).\end{aligned} \quad (5.29)$$

Utilizing Proposition 5.8, we may write

$$\begin{aligned}\int_0^{+\infty} \iint_{E\times E\setminus\mathrm{diag}} (\mathcal{J}_p^{(+)} + \mathcal{F}_p)(P_t\Phi^{(\pm)}(x), P_t\Phi^{(\pm)}(y))\, J(x,y)m(\mathrm{d}x)m(\mathrm{d}y)\mathrm{d}t \\ = \int_E \left(f_\pm (g_+)^{p-1} + |\Phi^{(\pm)}|^p\right)\mathrm{d}m,\end{aligned}$$

$$\begin{aligned}\int_0^{+\infty} \iint_{E\times E\setminus\mathrm{diag}} (\mathcal{J}_p^{(-)} + \mathcal{F}_p)(P_t\Phi^{(\pm)}(x), P_t\Phi^{(\pm)}(y))\, J(x,y)m(\mathrm{d}x)m(\mathrm{d}y)\mathrm{d}t \\ = \int_E \left(f_\pm (g_-)^{p-1} + |\Phi^{(\pm)}|^p\right)\mathrm{d}m.\end{aligned}$$

Due to (5.28) and (5.29),

$$\begin{aligned}\int_0^{+\infty} \iint_{E\times E\setminus\mathrm{diag}} \left|\mathcal{J}_p(P_t\Phi^{(\pm)}(x), P_t\Phi^{(\pm)}(y))\right| J(x,y)m(\mathrm{d}x)m(\mathrm{d}y)\mathrm{d}t \\ \leq \int_E \left(f_\pm |g|^{p-1} + 2\left|\Phi^{(\pm)}\right|^p\right)\mathrm{d}m.\end{aligned}$$

Summing up, by (5.27), we obtain

$$\begin{aligned}\int_0^{+\infty} \iint_{E\times E\setminus\mathrm{diag}} |\mathcal{J}_p(P_t\Phi(x), P_t\Phi(y))|\, J(x,y)m(\mathrm{d}x)m(\mathrm{d}y)\mathrm{d}t \\ \leq \int_E \left(|f|\,|g|^{p-1} + 2\left|\Phi^{(+)}\right|^p + 2\left|\Phi^{(-)}\right|^p\right)\mathrm{d}m.\end{aligned} \quad (5.30)$$

At this point, we need to estimate the right-hand side of the above inequality. By Hölder's inequality,

$$\int_E |f|\,|g|^{p-1}\ \mathrm{d}m \le \|f\|_p\,\|g\|_p^{p-1}.$$

Moreover, $\left|\Phi^{(+)}\right|^p + \left|\Phi^{(-)}\right|^p = |\Phi|^p + |g|^p$ and

$$|\Phi|^p = (f^2+g^2)^{p/2} \le 2^{p/2-1}(|f|^p + |g|^p).$$

Combining the above estimates with (5.30) we derive

$$\int_0^{+\infty}\iint_{E\times E\setminus\mathrm{diag}} |\mathcal{J}_p(P_t\Phi(x), P_t\Phi(y))|\, J(x,y)m(\mathrm{d}x)m(\mathrm{d}y)\mathrm{d}t$$
$$\le \|f\|_p\,\|g\|_p^{p-1} + 2^{p/2}(\|f\|_p^p + \|g\|_p^p) + 2\,\|g\|_p^p.$$

Assume first that $\|f\|_p = \|g\|_p = 1$. Then

$$\int_0^{+\infty}\iint_{E\times E\setminus\mathrm{diag}} |\mathcal{J}_p(P_t\Phi(x), P_t\Phi(y))|\ J(x,y)m(\mathrm{d}x)m(\mathrm{d}y)\mathrm{d}t \le 3 + 2^{p/2}.$$

If $\|f\|_p = 0$ or $\|g\|_p = 0$, then the desired inequality (5.26) is obvious. Otherwise, notice that $\mathcal{J}_p$ is homogeneous in the first coordinates, and $(p-1)$-homogeneous in the second, that is,

$$\mathcal{J}_p((\alpha w_1, \beta w_2), (\alpha z_1, \beta z_2)) = \alpha\beta^{p-1}\mathcal{J}_p((w_1,w_2),(z_1,z_2)), \quad \alpha,\beta > 0.$$

Therefore, by considering $f/\|f\|_p$ and $g/\|g\|_p$, we get the final result. □

5.4 Polarized Sobolev–Bregman form

In this section we consider the integral expression appearing in the polarized Hardy–Stein identity (5.17) given in Theorem 5.6, namely

$$\mathcal{E}_p(u,v) := \frac{1}{p}\iint_{E\times E\setminus\mathrm{diag}} \mathcal{J}_p(\Phi(x),\Phi(y))\, J(x,y)m(\mathrm{d}x)m(\mathrm{d}y),$$

where $2 \le p < \infty$, $\mathcal{J}_p$ is given by (5.16), $\Phi(x) = (u(x), v(x))$, and $u, v\colon E \to \mathbb{R}$, are some Borel functions. We are interested in the question, whether this integral is well-defined for functions u, v general enough. When $u = v$ then $\mathcal{E}_p(u,v) = \mathcal{E}_p(u,u) = \mathcal{E}_p[u]$ is just the Sobolev–Bregman form defined in Section 2.5. For $p = 2$, we get $\mathcal{E}_2(u,v) = \mathcal{E}(u,v)$, that is, $\mathcal{E}_2(u,v)$ reduces to the usual (bilinear) Dirichlet form defined in Section 2.4. In particular, it is symmetric. When $p > 2$, then in general $\mathcal{E}_p(v,u) \ne \mathcal{E}_p(u,v)$.

The main theorem of this section asserts that $\mathcal{E}_p(u,v)$ is well-defined, when $u, v \in \mathcal{D}(A_p)$ and the formula (2.28) extends to the following polarized version.

Theorem 5.10. *Suppose that Assumptions 4.1 and 5.1 holds. Let $2 \le p < \infty$. If $u, v \in \mathcal{D}(A_p)$, then $\mathcal{E}_p(u,v)$ is well-defined and*

$$\mathcal{E}_p(u,v) = -\frac{1}{p}\langle A_p u, v^{\langle p-1\rangle}\rangle - \frac{1}{p}\langle A_p v, (p-1)u|v|^{p-2}\rangle. \tag{5.31}$$

Before we proceed with the proof of the above theorem, we need a further decomposition of $\mathcal{J}_p^{(+)}$ (and $\mathcal{J}_p$) into a difference of non-negative functions.

Let $\mathbb{1}(a) := (1 + \operatorname{sgn}(a))/2$ be the Heaviside step function. We define

$$\begin{aligned}\mathcal{J}_p^{(++)}(w,z) :=&(z_1)_+ \left((z_2)_+\right)^{p-1} - (w_1)_+ \left((w_2)_+\right)^{p-1} - \mathbb{1}(w_1) \left((w_2)_+\right)^{p-1} (z_1 - w_1) \\ &- (p-1)(w_1)_+ \left((w_2)_+\right)^{p-2} (z_2 - w_2), \end{aligned} \tag{5.32}$$

$$\begin{aligned}\mathcal{J}_p^{(-+)}(w,z) :=&(z_1)_- \left((z_2)_+\right)^{p-1} - (w_1)_- \left((w_2)_+\right)^{p-1} + \mathbb{1}(-w_1) \left((w_2)_+\right)^{p-1} (z_1 - w_1) \\ &- (p-1)(w_1)_- \left((w_2)_+\right)^{p-2} (z_2 - w_2), \end{aligned} \tag{5.33}$$

where $w := (w_1, w_2)$, $z := (z_1, z_2) \in \mathbb{R}^2$. We may treat these functions as the second-order Taylor remainders of the mappings $\mathbb{R}^2 \ni z \mapsto (z_1)_+ \left((z_2)_+\right)^{p-1}$ and $\mathbb{R}^2 \ni z \mapsto (z_1)_- \left((z_2)_+\right)^{p-1}$, respectively, except for non-differentiability of the mappings on the vertical positive semi-axis. For more details; see the proof of Lemma B.5 in Appendix B.

Similarly to (5.21) and (5.20), we get a decomposition of $\mathcal{J}_p^{(+)}$

$$\mathcal{J}_p^{(+)} = \mathcal{J}_p^{(++)} - \mathcal{J}_p^{(-+)} = \left(\mathcal{J}_p^{(++)} + \mathcal{F}_p\right) - \left(\mathcal{J}_p^{(-+)} + \mathcal{F}_p\right) \tag{5.34}$$

and the identity

$$\mathcal{J}_p^{(++)}(-\bar{w}, -\bar{z}) = \mathcal{J}_p^{(-+)}(w,z). \tag{5.35}$$

In Lemma B.5 in Appendix B the following non-negativity was proven:

$$\mathcal{J}_p^{(++)}(w,z) + \mathcal{F}_p(w,z) \geq 0, \quad \mathcal{J}_p^{(-+)}(w,z) + \mathcal{F}_p(w,z) \geq 0$$

for all $z, w \in \mathbb{R}^2$. Therefore, (5.34) is a decomposition of $\mathcal{J}_p^{(+)}$ into a difference of non-negative functions, which is crucial for the proof of Theorem 5.10.

We emphasize that it is crucial to define the Heaviside function so that $\mathbb{1}(0) = 1/2$. Indeed, in that case we have the identity $\mathbb{1}(a) + \mathbb{1}(-a) = 1$ for all $a \in \mathbb{R}$ and the decomposition (5.34) holds.

We shall emphasize that the proof of Theorem 5.10 needs a distinct approach than the one used in Theorem 5.6 (and Proposition 5.8). The explanation for this is that, if $u \in \mathcal{D}(A_p)$, then u_+ and u_- need not be in $\mathcal{D}(A_p)$. Therefore, we cannot show the statement for non-negative u, and then utilize the decomposition $u = u_+ - u_-$. On the other hand, the following approach is not applicable in Theorem 5.6 in view of the fact that the decomposition $P_t f = (P_t f)_+ - (P_t f)_-$ is not of much use there, because in general we have neither differentiability in $L^p(m)$ of $(P_t f)_+$ nor $(P_t f)_-$. In other words, the decomposition $\mathcal{J}_p^{(+)} = \mathcal{J}_p^{(++)} - \mathcal{J}_p^{(-+)}$ is not suitable in Theorem 5.6, due to the fact that $\mathcal{J}_p^{(++)}$ and $\mathcal{J}_p^{(-+)}$ are not second-order Taylor remainders of differentiable functions.

Proof of Theorem 5.10. It is enough to show the statement for $p > 2$, because if $p = 2$, then the desired identity reduces to the already known equality (2.13).

Let $u, v \in \mathcal{D}(A_p)$ and denote $\Phi = (u, v)$. We start with proof of the following identities:

$$\begin{aligned} l_{++} :=& -\langle A_p u, \mathbb{1}(u)(v_+)^{p-1}\rangle - \langle A_p v, (p-1)u_+(v_+)^{p-2}\rangle - \langle A_p \Phi, p\Phi^{\langle p-1\rangle}\rangle \\ =& \iint\limits_{E\times E\setminus\mathrm{diag}} (\mathcal{J}_p^{(++)} + \mathcal{F}_p)(\Phi(x), \Phi(y))\, J(x,y) m(\mathrm{d}x) m(\mathrm{d}y) \end{aligned} \tag{5.36}$$

and

$$\begin{aligned} l_{-+} &:= \langle A_p u, \mathbb{1}(-u)(v_+)^{p-1}\rangle - \langle A_p v, (p-1)u_-(v_+)^{p-2}\rangle - \langle A_p\Phi, p\Phi^{\langle p-1\rangle}\rangle \\ &= \iint\limits_{E\times E\setminus \mathrm{diag}} (\mathcal{J}_p^{(-+)} + \mathcal{F}_p)(\Phi(x),\Phi(y))\, J(x,y)m(\mathrm{d}x)m(\mathrm{d}y). \end{aligned} \tag{5.37}$$

First, we show (5.36). Since $u, v \in \mathcal{D}(A_p)$, the strong limits defining A_pu, A_pv, and $A_p\Phi$ exist. Thus,

$$\begin{aligned} -\langle A_p u, \mathbb{1}(u)(v_+)^{p-1}\rangle &= \lim_{h\to 0^+}\frac{1}{h}\iint\limits_{E\times E} \mathbb{1}(u(x))(v(x)_+)^{p-1}(u(y)-u(x))\, P_h(\mathrm{d}x,\mathrm{d}y), \\ -\langle A_p v, (p-1)u_+(v_+)^{p-2}\rangle &= \lim_{h\to 0^+}\frac{1}{h}\iint\limits_{E\times E} (p-1)u(x)_+(v(x)_+)^{p-2}(v(y)-v(x))\, P_h(\mathrm{d}x,\mathrm{d}y), \\ -\langle A_p\Phi, p\Phi^{\langle p-1\rangle}\rangle &= \lim_{h\to 0^+}\frac{1}{h}\iint\limits_{E\times E} p(\Phi(x))^{\langle p-1\rangle}\cdot(\Phi(y)-\Phi(x))\, P_h(\mathrm{d}x,\mathrm{d}y). \end{aligned}$$

In particular, all limits on the right-hand side exist and are finite.

Summarizing,

$$\begin{aligned} l_{++} = \lim_{h\to 0^+}\frac{1}{h}\iint\limits_{E\times E} \Big[& \mathbb{1}(u(x))(v(x)_+)^{p-1}(u(y)-u(x)) \\ & + (p-1)u(x)_+(v(x)_+)^{p-2}(v(y)-v(x)) \\ & + p(\Phi(x))^{\langle p-1\rangle}\cdot(\Phi(y)-\Phi(x))\Big]\, P_h(\mathrm{d}x,\mathrm{d}y). \end{aligned} \tag{5.38}$$

Similarly as in the proof of Proposition 5.8, under Assumption 5.1 we may employ (5.10) to write

$$\frac{1}{h}\int\limits_E \left(u_+(v_+)^{p-1} + |\Phi|^p\right)\,\mathrm{d}m = \frac{1}{h}\int\limits_E P_h\left(u_+(v_+)^{p-1} + |\Phi|^p\right)\,\mathrm{d}m.$$

Subtracting the right-hand side and adding the left-hand side of the above equation under the limit on the right-hand side of (5.38), we obtain

$$\begin{aligned} l_{++} = \lim_{h\to 0^+}\frac{1}{h}\Bigg[& -\int\limits_E P_h\left(u_+(v_+)^{p-1}+|\Phi|^p\right)(x)\, m(\mathrm{d}x) \\ & + \int\limits_E \left(u(x)_+(v(x)_+)^{p-1} + |\Phi(x)|^p\right)\, m(\mathrm{d}x) + \iint\limits_{E\times E}\mathbb{1}(u(x))(v(x)_+)^{p-1}(u(y)-u(x)) \\ & + (p-1)u(x)_+(v(x)_+)^{p-2}(v(y)-v(x)) + p(\Phi(x))^{\langle p-1\rangle}\cdot(\Phi(y)-\Phi(x))\, P_h(\mathrm{d}x,\mathrm{d}y)\Bigg] \\ = \lim_{h\to 0^+}\frac{1}{h} & \iint\limits_{E\times E}(\mathcal{J}_p^{(++)}+\mathcal{F}_p)(\Phi(x),\Phi(y))\, P_h(\mathrm{d}x,\mathrm{d}y). \end{aligned}$$

By Lemma B.5 in Appendix B, we know that $\mathcal{J}_p^{(++)} + \mathcal{F}_p \geq 0$. Therefore, under Assumption 4.1, we can utilize Lemma 4.3 to derive

$$\begin{aligned} &\lim_{h\to 0^+}\frac{1}{h}\iint\limits_{E\times E}(\mathcal{J}_p^{(++)}+\mathcal{F}_p)(\Phi(x),\Phi(y))\, P_h(\mathrm{d}x,\mathrm{d}y) \\ &\qquad = \iint\limits_{E\times E\setminus\mathrm{diag}}(\mathcal{J}_p^{(++)}+\mathcal{F}_p)(\Phi(x),\Phi(y))\, J(x,y)m(\mathrm{d}x)m(\mathrm{d}y). \end{aligned}$$

This proves (5.36). Equation (5.37) follows from (5.36) by substituting $-u$ in place of u in view of (5.5) and (5.35).

With identities (5.36) and (5.37) at hand we are prepared to show another two equalities

$$
\begin{aligned}
l_+ &:= -\langle A_p u, (v_+)^{p-1}\rangle - \langle A_p v, (p-1)u(v_+)^{p-2}\rangle \\
&= \iint\limits_{E\times E\setminus \mathrm{diag}} \mathcal{J}_p^{(+)}(\Phi(x),\Phi(y))\, J(x,y)m(\mathrm{d}x)m(\mathrm{d}y),
\end{aligned}
\tag{5.39}
$$

and

$$
\begin{aligned}
l_- &:= -\langle A_p u, (v_-)^{p-1}\rangle + \langle A_p v, (p-1)u(v_-)^{p-2}\rangle \\
&= \iint\limits_{E\times E\setminus \mathrm{diag}} \mathcal{J}_p^{(-)}(\Phi(x),\Phi(y))\, J(x,y)m(\mathrm{d}x)m(\mathrm{d}y).
\end{aligned}
\tag{5.40}
$$

Indeed, since the integrals in (5.36), (5.37) are both finite, we use the decomposition (5.34) to derive

$$
\begin{aligned}
l_+ = l_{++} - l_{-+} &= \iint\limits_{E\times E\setminus \mathrm{diag}} (\mathcal{J}_p^{(++)} + \mathcal{F}_p)(\Phi(x),\Phi(y))\, J(x,y)m(\mathrm{d}x)m(\mathrm{d}y) \\
&\quad - \iint\limits_{E\times E\setminus \mathrm{diag}} (\mathcal{J}_p^{(-+)} + \mathcal{F}_p)(\Phi(x),\Phi(y))\, J(x,y)m(\mathrm{d}x)m(\mathrm{d}y) \\
&= \iint\limits_{E\times E\setminus \mathrm{diag}} \mathcal{J}_p^{(+)}(\Phi(x),\Phi(y))\, J(x,y)m(\mathrm{d}x)m(\mathrm{d}y).
\end{aligned}
$$

This yields (5.39). Similarly as before, equality (5.40) follows from (5.39) by substituting $-v$ in place of v, according to (5.5) and (5.35).

Again, both integrals from (5.39), (5.40) are finite. Therefore, we employ (5.39), (5.40) to obtain

$$
\begin{aligned}
&-\langle A_p u, v^{\langle p-1\rangle}\rangle - \langle A_p v, (p-1)u|v|^{p-2}\rangle = l_+ - l_- = \\
&= \iint\limits_{E\times E\setminus \mathrm{diag}} \mathcal{J}_p^{(+)}(\Phi(x),\Phi(y))\, J(x,y)m(\mathrm{d}x)m(\mathrm{d}y) \\
&\quad - \iint\limits_{E\times E\setminus \mathrm{diag}} \mathcal{J}_p^{(-)}(\Phi(x),\Phi(y))\, J(x,y)m(\mathrm{d}x)m(\mathrm{d}y) \\
&= p\mathcal{E}_p(u,v).
\end{aligned}
$$

Here, we used the decomposition (5.21).

In particular, the integral defining $\mathcal{E}_p(u,v)$ is absolutely convergent as a difference of two finite integrals.

This proves the statement for $p > 2$ and completes the proof. □

At the end of this section we present the polarized analogue of Proposition 3.4.

Proposition 5.11. *Let $2 \le p < \infty$ and $f, g \in L^p(m)$. Then, for $t > 0$*

$$
\frac{\mathrm{d}}{\mathrm{d}t}\int\limits_E P_t f (P_t g)^{\langle p-1\rangle}\mathrm{d}m = -p\mathcal{E}_p(P_t f, P_t g).
\tag{5.41}
$$

Moreover, the above quantity is a continuous function with respect to $t > 0$. Additionally, when $f, g \in \mathcal{D}(A_p)$, then also

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_E P_t f (P_t g)^{\langle p-1 \rangle} \mathrm{d}m \bigg|_{t=0} = -p\mathcal{E}_p(f, g) \tag{5.42}$$

and $[0, +\infty) \ni t \mapsto -p\mathcal{E}_p(P_t f, P_t g) \in \mathbb{R}$ is continuous.

Proof. Denote $u(t) := P_t f$ and $v(t) := P_t g$. Since the semigroup $(P_t)_{t\geq 0}$ is analytic on $L^p(m)$ for $1 < p < \infty$, u and v are is C^1 on $(0, +\infty)$ with values in $L^p(m)$. Moreover, according to Corollary 2.2(ii), the mapping $(v(t))^{\langle p-1 \rangle}$ is C^1 on $(0, +\infty)$ with values in $L^{p/(p-1)}(m)$ and

$$((v(t))^{\langle p-1 \rangle})' = (p-1)\, |v(t)|^{p-2}\, v(t).$$

Here, the assumption $p \geq 2$ was crucial. To be precise, Corollary 2.2 requires $p > 2$, but the above formula follows straightforwardly when $p = 2$. Finally, by the product rule from Proposition A.1, it follows that the mapping $u(t)(v(t))^{\langle p-1 \rangle}$ is C^1 on $(0, +\infty)$ with values in $L^1(m)$ and

$$(u(t)(v(t))^{\langle p-1 \rangle})' = (v(t))^{\langle p-1 \rangle} A_p u(t) + (p-1) u(t)\, |v(t)|^{p-2}\, A_p v(t).$$

In addition, if we assume that $f, g \in \mathcal{D}(A_p)$, then the above arguments lead to the conclusion that $u(t)(v(t))^{\langle p-1 \rangle}$ is C^1 on $[0, +\infty)$.

Since mapping $L^1(m) \ni h \mapsto \int_E h\, \mathrm{d}m$ is a continuous functional, we may write

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t} \int_E u(t)(v(t))^{\langle p-1 \rangle} \mathrm{d}m &= \int_E \Big(u(t)(v(t))^{\langle p-1 \rangle} \Big)' \,\mathrm{d}m \\
&= \int_E \Big((v(t))^{\langle p-1 \rangle} A_p u(t) + (p-1) u(t)\, |v(t)|^{p-2}\, A_p v(t) \Big)\, \mathrm{d}m \\
&= \langle A_p u(t), (v(t))^{\langle p-1 \rangle} \rangle + \langle A_p v(t), (p-1) u(t)\, |v(t)|^{p-2} \rangle \\
&= -p\mathcal{E}_p(u(t), v(t)).
\end{aligned}$$

The last equality follows from Theorem 5.10. The continuity of

$$t \mapsto \int_E (u(t)(v(t))^{\langle p-1 \rangle})' \,\mathrm{d}m = \frac{\mathrm{d}}{\mathrm{d}t} \int_E u(t)(v(t))^{\langle p-1 \rangle} \mathrm{d}m$$

follows from the continuity of $t \mapsto (u(t)(v(t))^{\langle p-1 \rangle})' \in L^1(m)$ and $L^1(m) \ni h \mapsto \int_E h\, \mathrm{d}m$. The proof is complete. $\square$

Chapter 6

Sobolev–Bregman form for continuous functions

In the present chapter, we partially relax the assumptions introduced in Chapter 4 and prove the Beurling–Deny formula of the Sobolev–Bregman form without domination conditions. Instead of employing the dominated convergence theorem and Fatou's lemma, the following strategy utilizes the vague convergence of the semigroup measure $\frac{1}{t}P_t(\mathrm{d}x, \mathrm{d}y)$ to the jumping measure $J(\mathrm{d}x, \mathrm{d}y)$, as mentioned in (2.26). In view of that, we obtain an explicit formula for the Sobolev–Bregman form is valid for continuous functions. The main result of this chapter is presented in Theorem 6.2.

This chapter is based on the paper [50] written by the author of this dissertation.

We consider the following assumptions.

Assumption 6.1.

(PJ) The regular Dirichlet form $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$ is pure-jump, that is, the strongly local part $\mathcal{E}^c$ in the Beurling–Deny formula vanishes.

(LF) The function $P_t f$ is continuous on E for every $f \in L^p(m)$ and $t > 0$.

Under assumption (PJ), the formula (2.25) reads

$$\lim_{t\to 0^+} \frac{1}{2t} \iint\limits_{E\times E} (u(y) - u(x))(v(y) - v(x))\, P_t(\mathrm{d}x, \mathrm{d}y) = \frac{1}{2} \iint\limits_{E\times E\setminus \mathrm{diag}} (\widetilde{u}(y) - \widetilde{u}(x))(\widetilde{v}(y) - \widetilde{v}(x))\, J(\mathrm{d}x, \mathrm{d}y), \tag{6.1}$$

where $u, v \in \mathcal{D}(\mathcal{E})$ and $\widetilde{u}, \widetilde{v}$ denote quasi-continuous versions of u and v. For more details, we refer to Section 2.4 in Chapter 2. Condition (PJ) is sufficient to derive an explicit form of the Sobolev–Bregman form for continuous functions. However, to apply it to the Hardy–Stein identity, assumption (LF) is also crucial.

We remark that assumption (LF) is not overly restrictive and the main result of the present chapter can be applied to various examples of pure-jump Dirichlet forms. Ensuring the validity of (LF) requires only that the jumping measure is sufficiently bounded. We may refer to Chen, Kumagai, and Wang [33, 34], for the context of pure-jump Dirichlet forms on a space with volume doubling condition. For instance, Dirichlet forms associated with α-stable-like processes on d-sets are included here; see Chen and Kumagai [32]. We refer to Appendix B of [50] for the verification of the validity of (LF) in the mentioned cases.

We present the main theorem of this chapter.

Theorem 6.2. *Let $u \in \mathcal{D}(\mathcal{E}_p)$. Then $u^{\langle p/2\rangle} \in \mathcal{D}(\mathcal{E})$. Suppose that the Dirichlet form $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$ is pure-jump (see (PJ) from Assumption 6.1) and u is continuous. Then,*

$$\lim_{t\to 0^+} \frac{1}{pt} \iint\limits_{E\times E} F_p(u(x), u(y))\, P_t(\mathrm{d}x, \mathrm{d}y) = \frac{1}{p} \iint\limits_{E\times E\setminus \mathrm{diag}} F_p(u(x), u(y))\, J(\mathrm{d}x, \mathrm{d}y) \tag{6.2}$$

and

$$\mathcal{E}_p[u] = \frac{1}{p} \iint\limits_{E\times E\setminus \mathrm{diag}} F_p(u(x), u(y))\, J(\mathrm{d}x, \mathrm{d}y) + \int\limits_E |u(x)|^p\, k(\mathrm{d}x). \tag{6.3}$$

Nonetheless,

$$\lim_{t\to 0^+} \frac{1}{t} \int\limits_E |u(x)|^p (1 - P_t 1(x))\, m(\mathrm{d}x) = \int\limits_E |\tilde{u}(x)|^p\, k(\mathrm{d}x) \tag{6.4}$$

and

$$\frac{4(p-1)}{p^2} \mathcal{E}[u^{\langle p/2\rangle}] \le \mathcal{E}_p[u] \le C_p \mathcal{E}[u^{\langle p/2\rangle}] \tag{6.5}$$

are valid regardless of (PJ) and the continuity of u. Here, $\tilde{u}$ is a quasi-continuous version of u and $C_p \ge 1$ is the constant from Lemma 2.3.

Compared to the main result of Chapter 4, the above theorem provides only $\subseteq$ inclusion in the characterization (4.3), but for a more general class of Dirichlet forms.

To employ this result to obtain an explicit Hardy–Stein identity, we also need (LF) from Assumption 6.1. Indeed, under (LF) Theorem 6.2 may be combined with the general Hardy–Stein identity from Theorem 3.2 to derive the following result.

Corollary 6.3 (Hardy–Stein identity)**.** *Let $1 < p < \infty$. Impose Assumption 6.1. For any $f \in L^p(m)$, the following identity holds:*

$$\begin{aligned} \int\limits_E |f(x)|^p\, m(\mathrm{d}x) - \lim_{T\to+\infty} \|P_T f\|_p^p = \\ = \int\limits_0^{+\infty} \iint\limits_{E\times E\setminus \mathrm{diag}} F_p(P_t f(x), P_t f(y))\, J(\mathrm{d}x, \mathrm{d}y)\mathrm{d}t \\ + p \int\limits_0^{+\infty} \int\limits_E |P_t f(x)|^p\, k(\mathrm{d}x)\mathrm{d}t. \end{aligned} \tag{6.6}$$

If we in addition assume the strong stability from Assumption 3.1, then the above identity reads

$$\begin{aligned} \int\limits_E |f(x)|^p\, m(\mathrm{d}x) = \int\limits_0^{+\infty} \iint\limits_{E\times E\setminus \mathrm{diag}} F_p(P_t f(x), P_t f(y))\, J(\mathrm{d}x, \mathrm{d}y)\mathrm{d}t \\ + p \int\limits_0^{+\infty} \int\limits_E |P_t f(x)|^p\, k(\mathrm{d}x)\mathrm{d}t. \end{aligned} \tag{6.7}$$

In contrast to Chapter 4, we do not assume pointwise convergence of the transition density $p_t(x, y)$ divided by t to the jumping density $J(x, y)$ as in assumption (J2). We do not even assume the existence of the density of the jumping or the killing measure. Instead of this, we employ vague convergence of measures $\frac{1}{t}P_t(\mathrm{d}x, \mathrm{d}y)$.

Indeed, as we mentioned in (2.26) in Chapter 2,

$$\frac{1}{t}P_t(\mathrm{d}x, \mathrm{d}y) \to J(\mathrm{d}x, \mathrm{d}y) \quad \text{vaguely on } E \times E \setminus \mathrm{diag} \text{ when } t \to 0^+. \tag{6.8}$$

Let us recall some notation. The class of continuous functions defined on a set A is denoted by $C(A)$. By $C_c(A)$ we denote the class of functions in $C(A)$ with compact support.

We consider the class $\mathcal{U}$ of non-negative functions f on $E \times E$ such that

$$\lim_{t \to 0^+} \frac{1}{t} \iint_{E \times E} f(x, y)\, P_t(\mathrm{d}x, \mathrm{d}y) = \iint_{E \times E \setminus \mathrm{diag}} f(x, y)\, J(\mathrm{d}x, \mathrm{d}y) < \infty.$$

By the definition of vague convergence, the class $\mathcal{U}$ contains all continuous functions on $E \times E \setminus \mathrm{diag}$ (extended to $E \times E$ by putting zero on diag) with compact support, i.e., $C_c(E \times E \setminus \mathrm{diag}) \subseteq \mathcal{U}$. Moreover, according to (6.1), $g \in \mathcal{U}$, when $g(x, y) := (u(y) - u(x))^2$ and u is a quasi-continuous function in $\mathcal{D}(\mathcal{E})$. If we want to generalize the above convergence to any $1 < p < \infty$ (compare with (4.6)), we need to show that the function $f(x, y) = F_p(u(x), u(y))$ also belongs to $\mathcal{U}$, at least when u is continuous on E.

To achieve that, we show the following crucial lemma.

Lemma 6.4. *If $0 \le f \le g$, $f = g = 0$ on* diag*, $f, g \in C(E \times E)$, and $g \in \mathcal{U}$, then $f \in \mathcal{U}$.*

Proof. Fix $\varepsilon > 0$. By the definition of $\mathcal{U}$, since $g \in \mathcal{U}$, the integral of g is finite, i.e.,

$$\iint_{E \times E \setminus \mathrm{diag}} g(x, y)\, J(\mathrm{d}x, \mathrm{d}y) < \infty.$$

Therefore also the integral of f is finite, namely

$$\iint_{E \times E \setminus \mathrm{diag}} f(x, y)\, J(\mathrm{d}x, \mathrm{d}y) \le \iint_{E \times E \setminus \mathrm{diag}} g(x, y)\, J(\mathrm{d}x, \mathrm{d}y) < \infty \tag{6.9}$$

which means that there is a compact subset $K \subseteq E \times E \setminus \mathrm{diag}$ such that

$$\iint_{K^c} g(x, y)\, J(\mathrm{d}x, \mathrm{d}y) < \varepsilon. \tag{6.10}$$

Let $\varphi \in C_c(E \times E \setminus \mathrm{diag})$ be such that $0 \le \varphi \le 1$ and $\varphi = 1$ on K. Since f is continuous, also $\varphi \cdot f \in C_c(E \times E \setminus \mathrm{diag})$. Hence from (6.8) we obtain

$$\lim_{t \to 0^+} \frac{1}{t} \iint_{E \times E \setminus \mathrm{diag}} \varphi(x, y) f(x, y)\, P_t(\mathrm{d}x, \mathrm{d}y) = \iint_{E \times E \setminus \mathrm{diag}} \varphi(x, y) f(x, y)\, J(\mathrm{d}x, \mathrm{d}y). \tag{6.11}$$

By (6.10) we have

$$\begin{aligned} \iint_{E \times E \setminus \mathrm{diag}} (1 - \varphi(x, y)) f(x, y)\, J(\mathrm{d}x, \mathrm{d}y) &\le \iint_{E \times E \setminus \mathrm{diag}} (1 - \varphi(x, y)) g(x, y)\, J(\mathrm{d}x, \mathrm{d}y) \\ &\le \iint_{K^c} g(x, y)\, J(\mathrm{d}x, \mathrm{d}y) < \varepsilon. \end{aligned} \tag{6.12}$$

We can also write

$$\begin{aligned}\frac{1}{t}\iint\limits_{E\times E\setminus\mathrm{diag}}(1-\varphi(x,y))f(x,y)\,P_t(\mathrm{d}x,\mathrm{d}y) &\le \frac{1}{t}\iint\limits_{E\times E\setminus\mathrm{diag}}(1-\varphi(x,y))g(x,y)\,P_t(\mathrm{d}x,\mathrm{d}y)\\ &= \frac{1}{t}\iint\limits_{E\times E\setminus\mathrm{diag}} g(x,y)\,P_t(\mathrm{d}x,\mathrm{d}y)\\ &\quad - \frac{1}{t}\iint\limits_{E\times E\setminus\mathrm{diag}}\varphi(x,y)g(x,y)\,P_t(\mathrm{d}x,\mathrm{d}y).\end{aligned}$$

Since g is continuous, we have $\varphi\cdot g\in C_c(E\times E\setminus\mathrm{diag})$, and since $g\in\mathcal{U}$, by (6.8), we find that, when $t\to 0^+$, the right-hand side converges to

$$\begin{aligned}&\iint\limits_{E\times E\setminus\mathrm{diag}} g(x,y)\,J(\mathrm{d}x,\mathrm{d}y) - \iint\limits_{E\times E\setminus\mathrm{diag}}\varphi(x,y)g(x,y)\,J(\mathrm{d}x,\mathrm{d}y)\\ &\quad = \iint\limits_{E\times E\setminus\mathrm{diag}}(1-\varphi(x,y))g(x,y)\,J(\mathrm{d}x,\mathrm{d}y) \le \iint\limits_{K^c} g(x,y)\,J(\mathrm{d}x,\mathrm{d}y) < \varepsilon.\end{aligned}$$

Here, we used (6.10). Summarizing,

$$\limsup_{t\to 0^+}\frac{1}{t}\iint\limits_{E\times E\setminus\mathrm{diag}}(1-\varphi(x,y))f(x,y)\,P_t(\mathrm{d}x,\mathrm{d}y) < \varepsilon. \tag{6.13}$$

Finally, from (6.11), (6.12), and (6.13) we get

$$\begin{aligned}&\limsup_{t\to 0^+}\left|\frac{1}{t}\iint\limits_{E\times E\setminus\mathrm{diag}} f(x,y)\,P_t(\mathrm{d}x,\mathrm{d}y) - \iint\limits_{E\times E\setminus\mathrm{diag}} f(x,y)\,J(\mathrm{d}x,\mathrm{d}y)\right|\\ &\le \limsup_{t\to 0^+}\frac{1}{t}\iint\limits_{E\times E\setminus\mathrm{diag}}(1-\varphi(x,y))f(x,y)\,P_t(\mathrm{d}x,\mathrm{d}y)\\ &\quad + \limsup_{t\to 0^+}\left|\frac{1}{t}\iint\limits_{E\times E\setminus\mathrm{diag}}\varphi(x,y)f(x,y)\,P_t(\mathrm{d}x,\mathrm{d}y) - \iint\limits_{E\times E\setminus\mathrm{diag}}\varphi(x,y)f(x,y)\,J(\mathrm{d}x,\mathrm{d}y)\right|\\ &\quad + \iint\limits_{E\times E\setminus\mathrm{diag}}(1-\varphi(x,y))f(x,y)\,J(\mathrm{d}x,\mathrm{d}y) < \varepsilon + 0 + \varepsilon = 2\varepsilon.\end{aligned}$$

Since $\varepsilon > 0$ was arbitrary, we have shown the desired convergence

$$\lim_{t\to 0^+}\frac{1}{t}\iint\limits_{E\times E\setminus\mathrm{diag}} f(x,y)\,P_t(\mathrm{d}x,\mathrm{d}y) = \iint\limits_{E\times E\setminus\mathrm{diag}} f(x,y)\,J(\mathrm{d}x,\mathrm{d}y).$$

Combining it with (6.9), we conclude that $f\in\mathcal{U}$. □

With the above lemma at hand, we are ready to prove the main result.

Proof of Theorem 6.2. Let $u\in\mathcal{D}(\mathcal{E}_p)$. Without loss of generality, we assume u to be quasi-continuous. By (2.37), in particular

$$\mathcal{E}^{(t)}[u^{\langle p/2\rangle}] \le \frac{p^2}{4(p-1)}\mathcal{E}^{(t)}\left(u, u^{\langle p-1\rangle}\right). \tag{6.14}$$

Now, since $u \in \mathcal{D}(\mathcal{E}_p)$, the right-hand side converges to a finite limit $\frac{p^2}{4(p-1)}\mathcal{E}_p[u]$ as $t \to 0^+$. Since the left-hand side is non-increasing as a function of t, a finite limit $\lim_{t\to 0^+} \mathcal{E}^{(t)}[u^{\langle p/2\rangle}]$ exists, i.e., $u^{\langle p/2\rangle} \in \mathcal{D}(\mathcal{E})$.

Since $u^{\langle p/2\rangle} \in \mathcal{D}(\mathcal{E})$, the equality (6.4) is an immediate consequence of (2.24).

The estimate (6.5) follows again from (2.37).

Now, we assume additionally that $u \in C(E)$ and that the condition (PJ) holds. Then (6.1) is available, from which it follows that $g(x, y) := (u^{\langle p/2\rangle}(y) - u^{\langle p/2\rangle}(x))^2$ belongs to $\mathcal{U}$. Moreover, by the continuity of u, we have $g \in C(E \times E)$, hence g fulfills the assumptions of Lemma 6.4. Similarly, $f(x, y) := F_p(u(x), u(y)$ is continuous and, since it is bounded up to multiplication by a constant by $g(x, y)$ (see Lemma 2.3), f belongs to $\mathcal{U}$ by Lemma 6.4. This proves (6.2).

Identity (6.3) follows immediately from (2.35), (6.4), and (6.2). $\square$

Chapter 7

Sobolev–Bregman form – general case

This chapter is an elaboration of the work [51] written by the author and Mateusz Kwaśnicki. The aim of this chapter is to generalize the Beurling–Deny formula for Sobolev–Bregman forms, and so the Hardy–Stein identity derived in Chapters 4 and 6. In the present analysis, we provide results that do not depend on the earlier assumptions about domination of the measures or continuity of functions. The result is valid for all regular Dirichlet forms. Moreover, we include in our study the strongly local part of a Dirichlet form employing LeJan's formula; see (2.23). Instead of relying on convergence theorems which cannot be used in full generality, our strategy is to utilize the fact that every normal contraction operates on a Dirichlet form. In view of that fact, we approximate the p-form $\mathcal{E}_p[u]$ by the Dirichlet form $\mathcal{E}(u_{1,n}, u_{p-1,n})$, where $u_{1,n}$ and $u_{p,n}$ are normal contractions of $u^{\langle p/2\rangle}$ approximating u and $u^{\langle p-1\rangle}$ respectively. To control the error, we employ a nontrivial technical estimate provided by Lemma 7.5. The main result is contained in Theorem 7.1 below.

Theorem 7.1. *The following characterization of the domain of the Sobolev–Bregman form holds:*

$$\mathcal{D}(\mathcal{E}_p) = \{\, u \in L^p(m) : u^{\langle p/2\rangle} \in \mathcal{D}(\mathcal{E}) \,\}. \tag{7.1}$$

In fact, the following estimate holds for any $u \in \mathcal{D}(\mathcal{E}_p)$:

$$\frac{4(p-1)}{p^2}\mathcal{E}[u^{\langle p/2\rangle}] \le \mathcal{E}_p[u] \le C_p \mathcal{E}[u^{\langle p/2\rangle}]. \tag{7.2}$$

Here, $C_p \ge 1$ is the constant from Lemma 2.3. Moreover, the p-form $\mathcal{E}_p$ is given by the following formula for any $u \in \mathcal{D}(\mathcal{E}_p)$:

$$\begin{aligned}\mathcal{E}_p[u] = {} & \frac{4(p-1)}{p^2}\mathcal{E}^c[u^{\langle p/2\rangle}] \\ & + \frac{1}{p}\iint\limits_{E\times E\setminus\mathrm{diag}} F_p(\tilde{u}(x), \tilde{u}(y))\, J(\mathrm{d}x, \mathrm{d}y) + \int\limits_E |\tilde{u}(x)|^p\, k(\mathrm{d}x),\end{aligned} \tag{7.3}$$

where $\tilde{u}$ is the quasi-continuous modification of u.

In particular, if $u \in \mathcal{D}(\mathcal{E}_p)$, then u has a quasi-continuous modification $\tilde{u}$, and both integrals in (7.3) *are finite. If the Dirichlet form $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$ is maximally defined, then additionally the converse is true: every $u \in L^p(E)$ which has a quasi-continuous modification $\tilde{u}$ such that the two integrals in* (7.3) *are finite, belongs to $\mathcal{D}(\mathcal{E}_p)$.*

Remark 7.2. *The decomposition* (7.3) *is implied by the following convergences of the corresponding parts of approximate form* $\mathcal{E}^{(t)}(u, u^{\langle p-1\rangle})$ *for an arbitrary* $u \in L^p(m)$*:*

$$\lim_{t\to 0^+} \frac{1}{pt} \iint\limits_{E\times E} F_p(u(x), u(y))\, P_t(\mathrm{d}x, \mathrm{d}y) \qquad (7.4)$$

$$= \frac{4(p-1)}{p^2}\mathcal{E}^c[u^{\langle p/2\rangle}] + \frac{1}{p}\iint\limits_{E\times E\setminus \mathrm{diag}} F_p(\tilde{u}(x), \tilde{u}(y))\, J(\mathrm{d}x, \mathrm{d}y)$$

and

$$\lim_{t\to 0^+} \frac{1}{t} \int\limits_E |u(x)|^p\, (1 - P_t 1(x))\, m(\mathrm{d}x) = \int\limits_E |\tilde{u}(x)|^p\, k(\mathrm{d}x). \qquad (7.5)$$

Remark 7.3. *Since the jumping measure* J *is symmetric, we may replace* F_p *in the formula* (7.3) *by the symmetrized Bregman divergence* H_p *given by* (2.30) *and write*

$$\begin{aligned}\mathcal{E}_p[u] = {}& \frac{4(p-1)}{p^2}\mathcal{E}^c[u^{\langle p/2\rangle}] \qquad (7.6)\\ &+ \frac{1}{2}\iint\limits_{E\times E\setminus \mathrm{diag}} (\tilde{u}(y) - \tilde{u}(x))(\tilde{u}^{\langle p-1\rangle}(y) - \tilde{u}^{\langle p-1\rangle}(x))\, J(\mathrm{d}x, \mathrm{d}y)\\ &+ \int\limits_E |\tilde{u}(x)|^p\, k(\mathrm{d}x).\end{aligned}$$

Now, we may present the Hardy–Stein identity in its most general form in this work.

Corollary 7.4 (Hardy–Stein identity)**.** *Let* $1 < p < \infty$*. For any* $f \in L^p(m)$*, the following identity holds:*

$$\begin{aligned}\int\limits_E |f(x)|^p\, m(\mathrm{d}x) - \lim_{T\to+\infty} \|P_T f\|_p^p = {}& \qquad (7.7)\\ &= \frac{4(p-1)}{p}\int\limits_0^{+\infty} \mathcal{E}^c[(P_t f)^{\langle p/2\rangle}]\, \mathrm{d}t\\ &+ \int\limits_0^{+\infty} \iint\limits_{E\times E\setminus \mathrm{diag}} F_p(P_t f(x), P_t f(y))\, J(\mathrm{d}x, \mathrm{d}y)\mathrm{d}t\\ &+ p\int\limits_0^{+\infty}\int\limits_E |P_t f(x)|^p\, k(\mathrm{d}x)\mathrm{d}t,\end{aligned}$$

where $P_t u$ *is assumed to be the quasi-continuous version. If we in addition assume the strong stability from Assumption 3.1, then the above identity reads*

$$\begin{aligned}\int\limits_E |f(x)|^p\, m(\mathrm{d}x) = {}& \frac{4(p-1)}{p}\int\limits_0^{+\infty} \mathcal{E}^c[(P_t f)^{\langle p/2\rangle}]\, \mathrm{d}t \qquad (7.8)\\ &+ \int\limits_0^{+\infty} \iint\limits_{E\times E\setminus \mathrm{diag}} F_p(P_t f(x), P_t f(y))\, J(\mathrm{d}x, \mathrm{d}y)\mathrm{d}t\\ &+ p\int\limits_0^{+\infty}\int\limits_E |P_t f(x)|^p\, k(\mathrm{d}x)\mathrm{d}t.\end{aligned}$$

7.1 Elementary estimates

A number of times in this chapter we employ comparability of symmetrized Bregman divergence H_p introduced in Lemma 2.3 under alternative auxiliary notation. Namely, by substitution $s = a^{\langle p/2 \rangle}$, $t = b^{\langle p/2 \rangle}$, and $\alpha := \frac{2}{p} \in (0, 2)$, we rewrite (2.31) to

$$\alpha(2-\alpha)(t-s)^2 \leq (t^{\langle \alpha \rangle} - s^{\langle \alpha \rangle})(t^{\langle 2-\alpha \rangle} - s^{\langle 2-\alpha \rangle}) \leq D_\alpha (t-s)^2, \tag{7.9}$$

where by D_α we denote the constant $1 \leq D_\alpha := C_p \leq 2$ from Lemma 2.3. In particular, since $C_{p/(p-1)} = C_p$, $D_{2-\alpha} = D_\alpha$.

Lemma 7.5. *For $\alpha \in (0, 2)$ and an integer $n \geq 2$, denote (see Figure 7.1)*

$$\varphi_\alpha(s) = s^{\langle \alpha \rangle},$$

$$\varphi_{\alpha,n}(s) = \begin{cases} s & \text{if } |s| < 1, \\ s^{\langle \alpha \rangle} & \text{if } 1 \leq |s| < n^4, \\ n^{4\alpha} \operatorname{sgn}(s) & \text{if } n^4 \leq |s|, \end{cases}$$

$$\psi_n(s) = \begin{cases} s & \text{if } |s| < n, \\ n \operatorname{sgn}(s) & \text{if } n \leq |s| < n^3, \\ s - (n^3 - n) \operatorname{sgn}(s) & \text{if } n^3 \leq |s|. \end{cases}$$

Then,

$$\begin{aligned} \Big| (\varphi_{\alpha,n}(t) - \varphi_{\alpha,n}(s))(\varphi_{2-\alpha,n}(t) - \varphi_{2-\alpha,n}(s)) \\ - (\varphi_\alpha(t) - \varphi_\alpha(s))(\varphi_{2-\alpha}(t) - \varphi_{2-\alpha}(s)) \Big| \\ \leq 4n^{-\min\{\alpha, 2-\alpha\}}(t-s)^2 + 90 D_\alpha (\psi_n(t) - \psi_n(s))^2 \end{aligned} \tag{7.10}$$

for all $n \geq 2$ and all $s, t \in \mathbb{R}$.

Proof. We divide the present proof into six steps.

Step 1. We begin with elementary simplifications. By the symmetry between s and t, without loss of generality we may assume that $|s| \leq |t|$. Since φ_α, $\varphi_{\alpha,n}$ and ψ_n are odd functions, arguments s, t may be replaced by $-s, -t$. Thus we may additionally assume that $t \geq 0$. Therefore, it is sufficient to consider $s, t \in \mathbb{R}$ such that $-t \leq s \leq t$. This region is shown in Figure 7.2. Furthermore, since the statement of the lemma does not change when α is replaced by $2 - \alpha$, we may consider only $\alpha \in (0, 1]$. These conditions are assumed throughout the rest of the proof.

Because of the way of defining functions $\varphi_{\alpha,n}$ and ψ_n, we split the region $|s| \leq t$ into a number of subregions, as was shown in Figure 7.2.

To further simplify the notation, we introduce the following functions:

$$\begin{aligned} \Phi_\alpha(s,t) &:= (\varphi_\alpha(t) - \varphi_\alpha(s))(\varphi_{2-\alpha}(t) - \varphi_{2-\alpha}(s)), \\ \Phi_{\alpha,n}(s,t) &:= (\varphi_{\alpha,n}(t) - \varphi_{\alpha,n}(s))(\varphi_{2-\alpha,n}(t) - \varphi_{2-\alpha,n}(s)), \\ \Psi_n(s,t) &:= (\psi_n(t) - \psi_n(s))^2. \end{aligned}$$

Thus, the desired inequality (7.10) can be rewritten in the following way:

$$|\Phi_{\alpha,n}(s,t) - \Phi_\alpha(s,t)| \leq 4n^{-\alpha}(t-s)^2 + 90 D_\alpha \Psi_n(s,t). \tag{7.11}$$

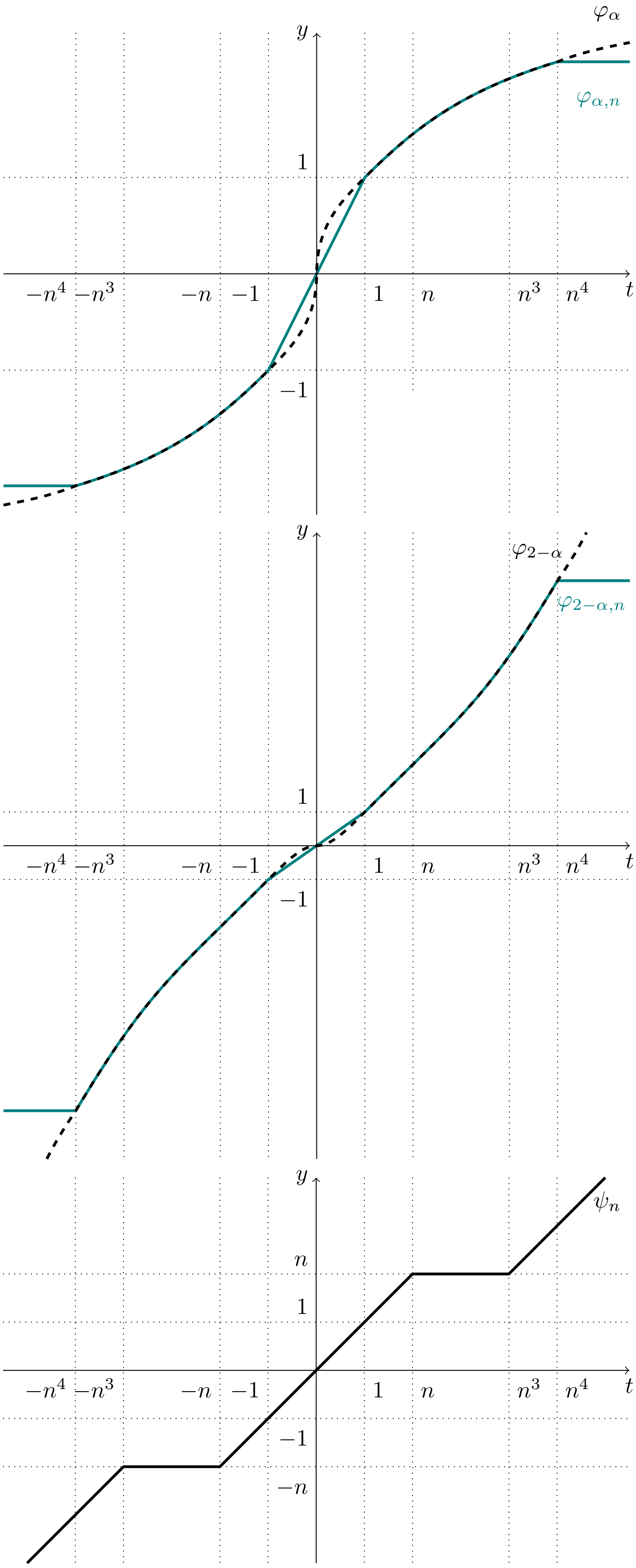

Figure 7.1: Functions defined in Lemma 7.5 (not to scale). Source: [51].

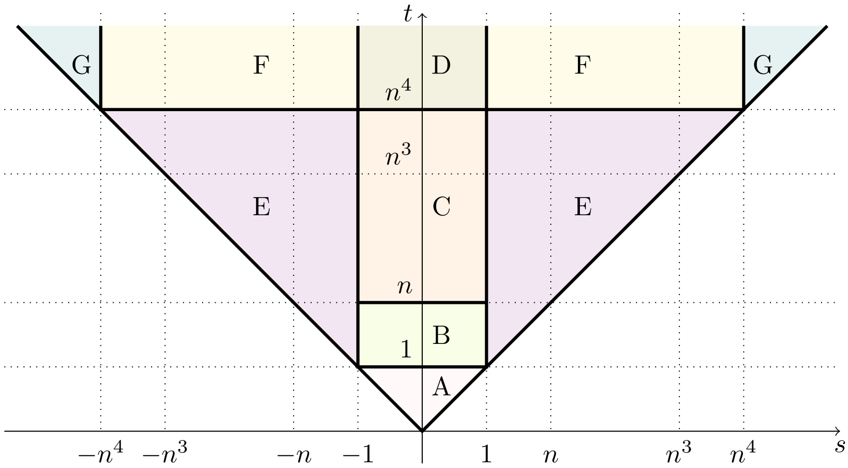

Figure 7.2: Regions considered in the proof of Lemma 7.5 (not to scale). Source: [51].

At the beginning, in Steps 2–4, we gather the necessary estimates of Φ_α and $\Phi_{\alpha,n}$, and then, in Step 5, we estimate Ψ_n, and only then we return to the actual proof of (7.11) in Step 6.

Step 2. We start with the following immediate estimates of Φ_α and $\Phi_{\alpha,n}$. By definitions of $\varphi_{\alpha,n}$ and ψ_n,

$$\Phi_{\alpha,n}(s,t) = \Phi_\alpha(s,t) \qquad \text{if } \underbrace{1 \le |s| \le t \le n^4}_{\text{region E}}, \tag{7.12}$$

and

$$\Phi_{\alpha,n}(s,t) = (s-t)^2 \qquad \text{if } \underbrace{|s| \le t \le 1}_{\text{region A}}. \tag{7.13}$$

By the inequality (7.9),

$$0 \le \Phi_\alpha(s,t) \le D_\alpha (t-s)^2 \qquad \text{if } \underbrace{|s| \le t}_{\text{all regions}}. \tag{7.14}$$

Step 3. We turn to an estimate of $\Phi_{\alpha,n}(s,t)$ by $(t-s)^2$, similar to (7.14) in all regions but A and B. First, if $1 \le |s| \le t$ (regions E, F, G), then

$$0 \le \varphi_{\alpha,n}(t) - \varphi_{\alpha,n}(s) \le \varphi_\alpha(t) - \varphi_\alpha(s),$$
$$0 \le \varphi_{2-\alpha,n}(t) - \varphi_{2-\alpha,n}(s) \le \varphi_{2-\alpha}(t) - \varphi_{2-\alpha}(s),$$

and therefore

$$0 \le \Phi_{\alpha,n}(s,t) \le \Phi_\alpha(s,t).$$

Combining this with (7.14), we get

$$0 \le \Phi_{\alpha,n}(s,t) \le D_\alpha (t-s)^2 \qquad \text{if } \underbrace{1 \le |s| \le t}_{\text{regions E, F, G}}. \tag{7.15}$$

Secondly, when $|s| \le 1$ and $n \le t$ (regions C, D), a similar estimate is valid. In this case we may write

$$0 \le \varphi_{\alpha,n}(t) - \varphi_{\alpha,n}(s) \le \varphi_{\alpha,n}(t) - \varphi_{\alpha,n}(-1),$$
$$0 \le \varphi_{2-\alpha,n}(t) - \varphi_{2-\alpha,n}(s) \le \varphi_{2-\alpha,n}(t) - \varphi_{2-\alpha,n}(-1),$$

and therefore

$$0 \le \Phi_{\alpha,n}(s,t) \le \Phi_{\alpha,n}(-1,t).$$

By (7.15) applied with $s = -1$, we find that

$$\Phi_{\alpha,n}(-1,t) \le D_\alpha(t+1)^2.$$

Moreover, since $t \ge n$ and $n \ge 2$, we may write

$$\begin{aligned} t+1 &\le 3t - 2n + 1 \\ &\le 3(t-1) \\ &\le 3(t-s). \end{aligned}$$

The above estimates together lead to

$$\begin{aligned} \Phi_{\alpha,n}(s,t) &\le \Phi_{\alpha,n}(-1,t) \\ &\le 2(t+1)^2 \\ &\le 9D_\alpha(t-s)^2. \end{aligned}$$

Thus, we have just shown the following analogue of (7.15):

$$0 \le \Phi_{\alpha,n}(s,t) \le 9D_\alpha(t-s)^2 \qquad \text{if } \underbrace{|s| \le 1 < n \le t}_{\text{regions C, D}}. \tag{7.16}$$

Step 4. The most careful estimate of $\Phi_{\alpha,n}(s,t) - \Phi_\alpha(s,t)$ is needed in regions B and C, when $|s| \le 1 \le t \le n^4$. Rearranging this difference, we may write

$$\begin{aligned} \Phi_{\alpha,n}(s,t) - \Phi_\alpha(s,t) &= (\varphi_{\alpha,n}(t) - \varphi_{\alpha,n}(s))(\varphi_{2-\alpha,n}(t) - \varphi_{2-\alpha,n}(s)) \\ &\quad - (\varphi_\alpha(t) - \varphi_\alpha(s))(\varphi_{2-\alpha}(t) - \varphi_{2-\alpha}(s)) \\ &= (t^\alpha - s)(t^{2-\alpha} - s) - (t^\alpha - s^{\langle\alpha\rangle})(t^{2-\alpha} - s^{\langle 2-\alpha\rangle}) \\ &= (t^2 - t^\alpha s - t^{2-\alpha}s + s^2) - (t^2 - t^\alpha s^{\langle 2-\alpha\rangle} - t^{2-\alpha}s^{\langle\alpha\rangle} + s^2) \\ &= t^\alpha s^{\langle 2-\alpha\rangle} + t^{2-\alpha}s^{\langle\alpha\rangle} - t^\alpha s - t^{2-\alpha}s \\ &= s^{\langle\alpha\rangle}t^\alpha(1 - |s|^{1-\alpha})(t^{2-2\alpha} - |s|^{1-\alpha}). \end{aligned}$$

In particular, the quantities in parenthesis are non-negative. Therefore,

$$|\Phi_{\alpha,n}(s,t) - \Phi_\alpha(s,t)| = |s|^\alpha\, t^\alpha(1 - |s|^{1-\alpha})(t^{2-2\alpha} - |s|^{1-\alpha}) \quad \text{if } \underbrace{|s| \le 1 \le t \le n^4}_{\text{regions B, C}}. \tag{7.17}$$

Firstly, consider region C: suppose that $|s| \le 1$ and $n \le t$. Utilizing (7.17), we have

$$|\Phi_{\alpha,n}(s,t) - \Phi_\alpha(s,t)| \le t^{2-\alpha} \le n^{-\alpha}t^2.$$

Since $2 \le n \le t$, we have $t \le 2(t-1) \le 2(t-s)$,

$$|\Phi_{\alpha,n}(s,t) - \Phi_\alpha(s,t)| \le 4n^{-\alpha}(t-s)^2 \qquad \text{if } \underbrace{|s| \le 1 \le n \le t \le n^4}_{\text{region C}}. \tag{7.18}$$

On the other hand, if $|s| \le 1 \le t \le n$ (region B), then again by (7.17),

$$\begin{aligned} |\Phi_{\alpha,n}(s,t) - \Phi_\alpha(s,t)| &\le t^\alpha(1-|s|^{1-\alpha})(t^{2-2\alpha} - |s|^{1-\alpha}) \\ &= t^\alpha(1-|s|^{1-\alpha})((t^{2-2\alpha}-1) + (1-|s|^{1-\alpha})). \end{aligned}$$

We combine $1 - |s|^{1-\alpha} \le 1 - s$ with

$$t^{2-2\alpha} - 1 = (t^{1-\alpha}+1)(t^{1-\alpha}-1) \le 2t^{1-\alpha}(t-1),$$

to continue estimates and write

$$\begin{aligned} |\Phi_{\alpha,n}(s,t) - \Phi_\alpha(s,t)| &\le t^\alpha(1-s)(2t^{1-\alpha}(t-1) + (1-s)) \\ &\le t^\alpha(1-s)(2t^{1-\alpha}(t-1) + 2t^{1-\alpha}(1-s)) \\ &= 2t(1-s)(t-s). \end{aligned}$$

Finally, $t(1-s) = t - st \le t - s$ if $s \ge 0$, and $t(1-s) = t - st \le 2t \le 2(t-s)$ if $s \le 0$. Concluding,

$$|\Phi_{\alpha,n}(s,t) - \Phi_\alpha(s,t)| \le 4(t-s)^2 \qquad \text{if } \underbrace{|s| \le 1 \le t \le n}_{\text{region B}}. \tag{7.19}$$

Step 5. Before we prove (7.11), it remains to estimate Ψ_n. By the definition of ψ_n, we can write immediately that

$$\Psi_n(s,t) = (t-s)^2 \qquad \text{if } \underbrace{|s| \le t \le n}_{\text{includes regions A, B}} \text{ or } \underbrace{n^3 \le |s| \le t}_{\text{includes region G}}. \tag{7.20}$$

Suppose now that $|s| \le n^3$ and $n^4 \le t$. Then

$$\psi_n(t) - \psi_n(s) \ge \psi_n(t) - n = t - n^3.$$

Since $t \ge n^4 \ge 2n^3$,

$$\begin{aligned} 3(\psi_n(t) - \psi_n(s)) &\ge 3t - 3n^3 \\ &\ge t + n^3 \\ &\ge t - s. \end{aligned}$$

By the above estimates, it follows that

$$9\Psi_n(s,t) = 9(\psi_n(t) - \psi_n(s))^2 \ge (t-s)^2.$$

This estimate partially covers regions D and F. The remaining part of these regions is included in the estimate (7.20). Summarizing,

$$9\Psi_n(s,t) \ge (t-s)^2 \qquad \text{if } \underbrace{|s| \le n^4 \le t}_{\text{regions D, F}}. \tag{7.21}$$

Step 6. With the above bounds at hand, we are ready to prove the desired inequality (7.11). We consider one by one the following seven cases, which correspond to regions shown in Figure 7.2.

- Region A: If $|s| \le t \le 1$, then by (7.13), (7.14), and (7.20),

$$\begin{aligned}|\Phi_{\alpha,n}(s,t) - \Phi_\alpha(s,t)| &\le \Phi_{\alpha,n}(s,t) + \Phi_\alpha(s,t)\\ &\le (1+D_\alpha)(t-s)^2\\ &= (1+D_\alpha)\Psi_n(s,t).\end{aligned}$$

- Region B: If $|s| \le 1 \le t \le n$, then by (7.19) and (7.20), we have

$$\begin{aligned}|\Phi_{\alpha,n}(s,t) - \Phi_\alpha(s,t)| &\le 4(t-s)^2\\ &= 4\Psi_n(s,t).\end{aligned}$$

- Region C: If $|s| \le 1 < n \le t \le n^4$, then by (7.18),

$$|\Phi_{\alpha,n}(s,t) - \Phi_\alpha(s,t)| \le 4n^{-\alpha}(t-s)^2.$$

- Region D: If $|s| \le 1 < n^4 \le t$, then by (7.14), (7.16), and (7.21),

$$\begin{aligned}|\Phi_{\alpha,n}(s,t) - \Phi_\alpha(s,t)| &\le \Phi_{\alpha,n}(s,t) + \Phi_\alpha(s,t)\\ &\le 10D_\alpha(t-s)^2\\ &\le 90D_\alpha\Psi_n(s,t).\end{aligned}$$

- Region E: If $1 \le |s| \le t \le n^4$, then by (7.12),

$$|\Phi_{\alpha,n}(s,t) - \Phi_\alpha(s,t)| = 0.$$

- Region F: If $1 \le |s| \le n^4 \le t$, then by (7.14), (7.15), and (7.21),

$$\begin{aligned}|\Phi_{\alpha,n}(s,t) - \Phi_\alpha(s,t)| &\le \Phi_{\alpha,n}(s,t) + \Phi_\alpha(s,t)\\ &\le 2D_\alpha(t-s)^2\\ &\le 18D_\alpha\Psi_n(s,t).\end{aligned}$$

- Region G: If $n^4 \le |s| \le t$, then by (7.14), (7.15), and (7.20),

$$\begin{aligned}|\Phi_{\alpha,n}(s,t) - \Phi_\alpha(s,t)| &\le \Phi_{\alpha,n}(s,t) + \Phi_\alpha(s,t)\\ &\le 2D_\alpha(t-s)^2\\ &= 2D_\alpha\Psi_n(s,t).\end{aligned}$$

The proof is complete. □

7.2 Proof of the main result

With Lemma 7.5 at hand, we are ready to prove Theorem 7.1.

For convenience, we introduce the following notation. Denote by $\mathcal{E}_p^c[u]$, $\mathcal{E}_p^j[u]$, and $\mathcal{E}_p^k[u]$ the strongly local part, the jumping part, and the killing part of the p-form appearing on the right-hand side of (7.3):

$$\mathcal{E}_p^c[u] := \frac{4(p-1)}{p^2}\mathcal{E}^c[u^{\langle p/2\rangle}],$$
$$\mathcal{E}_p^j[u] := \frac{1}{p}\iint\limits_{E\times E\setminus\mathrm{diag}} F_p(u(x),u(y))\,J(\mathrm{d}x,\mathrm{d}y),$$
$$\mathcal{E}_p^k[u] := \int\limits_E |u(x)|^p\,k(\mathrm{d}x).$$

Observe that $\mathcal{E}_p^j[u]$ and $\mathcal{E}_p^k[u]$ are well-defined (possibly infinite) for all Borel functions u due to the non-negativity of the integrands. By now, we cannot verify whether $\mathcal{E}_p^c[u]$ is well-defined unless we assume that $u^{\langle p/2\rangle} \in \mathcal{D}(\mathcal{E})$.

It is important to note that, as we previously mentioned in the case of Dirichlet forms, functions u that are equal almost everywhere cannot be identified here, as k and J may assign positive measure to sets that have zero measure with respect to m or $m \otimes m$.

Proof of Theorem 7.1. First of all, in view of Remark 7.3, it is enough to prove the statement for the symmetrized Bregman divergence H_p instead of F_p, that is:

$$\mathcal{E}_p^j[u] = \frac{1}{p}\iint\limits_{E\times E\setminus\mathrm{diag}} H_p(\widetilde{u}(x),\widetilde{u}(y))\,J(\mathrm{d}x,\mathrm{d}y).$$

Moreover, observe that an arbitrary Borel function u is quasi-continuous if and only if $u^{\langle p/2\rangle}$ is quasi-continuous.

The proof of the inclusion $\subseteq$ in (7.1) follows the same steps as in the proof of Theorem 6.2: by (6.14) and the fact that $\mathcal{E}^{(t)}[u^{\langle p/2\rangle}]$ is non-increasing function with respect to t. Therefore, we omit it here. For the details, we refer to the proof therein.

The proof of the inclusion $\supseteq$ in (7.1) and the formula (7.3) is more complicated. Therefore, we split it into six steps.

Step 1. Let $u^{\langle p/2\rangle} \in \mathcal{D}(\mathcal{E})$. Without loss of generality, we assume $u^{\langle p/2\rangle}$ (and hence u) to be quasi-continuous. Equality (7.5):

$$\lim_{t\to 0^+}\frac{1}{t}\int\limits_E |u(x)|^p\,(1-P_t1(x))\,m(\mathrm{d}x) = \mathcal{E}_p^k[u]$$

is an immediate consequence of (2.24).

In this step, for convenience, we introduce an auxiliary notation and reformulate the problem. Denote $\alpha := 2/p$ and $v := u^{\langle p/2\rangle}$. Thus, $v \in \mathcal{D}(\mathcal{E})$ is quasi-continuous and $\alpha \in (0,2)$.

With the above notation,

$$\mathcal{E}_p^c[u] = \frac{4(p-1)}{p^2}\mathcal{E}^c[u^{\langle p/2\rangle}] = \alpha(2-\alpha)\mathcal{E}^c[v]$$

and

$$\begin{aligned}\mathcal{E}_p^j[u] &= \frac{1}{2}\iint\limits_{E\times E\setminus\mathrm{diag}} (u(y)-u(x))(u^{\langle p-1\rangle}(y)-u^{\langle p-1\rangle}(x))\,J(\mathrm{d}x,\mathrm{d}y)\\ &= \frac{1}{2}\iint\limits_{E\times E\setminus\mathrm{diag}} (v^{\langle\alpha\rangle}(y)-v^{\langle\alpha\rangle}(x))(v^{\langle 2-\alpha\rangle}(y)-v^{\langle 2-\alpha\rangle}(x))\,J(\mathrm{d}x,\mathrm{d}y).\end{aligned}$$

Therefore, our goal (7.4) reads

$$\lim_{t\to0^+}\frac{1}{2t}\iint_{E\times E}(v^{\langle\alpha\rangle}(y)-v^{\langle\alpha\rangle}(x))(v^{\langle2-\alpha\rangle}(y)-v^{\langle2-\alpha\rangle}(x))\,P_t(\mathrm{d}x,\mathrm{d}y) \tag{7.22}$$
$$=\alpha(2-\alpha)\mathcal{E}^c[v]+\frac{1}{2}\iint_{E\times E\setminus\mathrm{diag}}(v^{\langle\alpha\rangle}(y)-v^{\langle\alpha\rangle}(x))(v^{\langle2-\alpha\rangle}(y)-v^{\langle2-\alpha\rangle}(x))\,J(\mathrm{d}x,\mathrm{d}y).$$

To apply Lemma 7.5, we employ the functions φ_α, $\varphi_{\alpha,n}$, and ψ_n $(n\geq 2)$ introduced therein and use the following notation:

$$v_\alpha(x):=n^{-2\alpha}\varphi_\alpha(n^2v(x))=v^{\langle\alpha\rangle}(x),$$
$$v_{\alpha,n}(x):=n^{-2\alpha}\varphi_{\alpha,n}(n^2v(x)),$$
$$w_n(x):=n^{-2}\psi_n(n^2v(x)),$$

and furthermore

$$\Phi_\alpha(x,y):=(v_\alpha(y)-v_\alpha(x))(v_{2-\alpha}(y)-v_{2-\alpha}(x)),$$
$$\Phi_{\alpha,n}(x,y):=(v_{\alpha,n}(y)-v_{\alpha,n}(x))(v_{2-\alpha,n}(y)-v_{2-\alpha,n}(x)).$$

Finally, let $\varepsilon_{\alpha,n}:=4n^{-\min\{\alpha,2-\alpha\}}$. With the above notation, Lemma 7.5 states that

$$|\Phi_\alpha(x,y)-\Phi_{\alpha,n}(x,y)|\leq\varepsilon_{\alpha,n}(v(y)-v(x))^2+90D_\alpha(w_n(y)-w_n(x))^2, \tag{7.23}$$

while our goal (7.22) reads

$$\lim_{t\to0^+}\frac{1}{2t}\iint_{E\times E}\Phi_\alpha(x,y)\,P_t(\mathrm{d}x,\mathrm{d}y)=\alpha(2-\alpha)\mathcal{E}^c[v]+\frac{1}{2}\iint_{E\times E\setminus\mathrm{diag}}\Phi_\alpha(x,y)\,J(\mathrm{d}x,\mathrm{d}y). \tag{7.24}$$

To establish this equality, we approximate Φ_α with $\Phi_{\alpha,n}$ and then estimate the error by (7.23).

Step 2. Observe that, since $v\in\mathcal{D}(\mathcal{E})$ and v is quasi-continuous, by (2.25) we may write

$$\lim_{t\to0^+}\frac{1}{2t}\iint_{E\times E}(v(y)-v(x))^2\,P_t(\mathrm{d}x,\mathrm{d}y)=\mathcal{E}^c[v]+\frac{1}{2}\iint_{E\times E\setminus\mathrm{diag}}(v(y)-v(x))^2\,P_t(\mathrm{d}x,\mathrm{d}y). \tag{7.25}$$

Step 3. At this step and the next step we prepare the necessary results to handle the jumping term.

Recall that $v\in\mathcal{D}(\mathcal{E})$. Since $\varphi_{\alpha,n}$, $\varphi_{2-\alpha,n}$ are Lipschitz functions satisfying $\varphi_{\alpha,n}(0)=\varphi_{2-\alpha,n}(0)=0$, functions $v_{\alpha,n}$, $v_{2-\alpha,n}$ belong to $\mathcal{D}(\mathcal{E})$. We refer to (2.10). Moreover, they are quasi-continuous. Utilizing (2.25), we obtain

$$\lim_{t\to0^+}\frac{1}{2t}\iint_{E\times E}(v_{\alpha,n}(y)-v_{\alpha,n}(x))(v_{2-\alpha,n}(y)-v_{2-\alpha,n}(x))\,P_t(\mathrm{d}x,\mathrm{d}y)$$
$$=\mathcal{E}^c(v_{\alpha,n},v_{2-\alpha,n})+\frac{1}{2}\iint_{E\times E\setminus\mathrm{diag}}(v_{\alpha,n}(y)-v_{\alpha,n}(x))(v_{2-\alpha,n}(y)-v_{2-\alpha,n}(x))\,J(\mathrm{d}x,\mathrm{d}y).$$

In terms of $\Phi_{\alpha,n}$, the above equality reads

$$\lim_{t\to0^+}\frac{1}{2t}\iint_{E\times E}\Phi_{\alpha,n}(x,y)\,P_t(\mathrm{d}x,\mathrm{d}y)$$
$$=\mathcal{E}^c(v_{\alpha,n},v_{2-\alpha,n})+\frac{1}{2}\iint_{E\times E\setminus\mathrm{diag}}\Phi_{\alpha,n}(x,y)\,J(\mathrm{d}x,\mathrm{d}y). \tag{7.26}$$

Step 4. Since v belongs to $\mathcal{D}(\mathcal{E})$ and ψ_n are Lipschitz functions, similarly to the previous step, we find that the functions $w_n(x) = n^{-2}\psi_n(n^2 v(x))$ belong to $\mathcal{D}(\mathcal{E})$ as well in view of (2.10). Moreover, v and w_n are quasi-continuous. Thus, from (2.25) we get

$$\begin{aligned}\lim_{t\to 0^+} \frac{1}{2t} \iint\limits_{E\times E} (w_n(y) - w_n(x))^2\, P_t(\mathrm{d}x, \mathrm{d}y) \\ = \mathcal{E}^c[w_n] + \frac{1}{2} \iint\limits_{E\times E\setminus \mathrm{diag}} (w_n(y) - w_n(x))^2\, J(\mathrm{d}x, \mathrm{d}y).\end{aligned} \tag{7.27}$$

Furthermore, the functions $s \mapsto n^{-2}\psi_n(n^2 s)$ have Lipschitz constant 1 and they converge pointwise to zero as $n \to +\infty$. Thus, $w_n(x) = n^{-2}\psi_n(n^2 v(x)) \to 0$ pointwise as $n \to +\infty$. Moreover, $|w_n(y) - w_n(x)| \le |v(y) - v(x)|$ and $(v(y) - v(x))^2$ is integrable with respect to $J(\mathrm{d}x, \mathrm{d}y)$, therefore we are allowed to utilize the dominated convergence theorem to conclude that

$$\lim_{n\to+\infty} \iint\limits_{E\times E\setminus \mathrm{diag}} (w_n(y) - w_n(x))^2\, J(\mathrm{d}x, \mathrm{d}y) = 0. \tag{7.28}$$

Step 5. At the present step we turn to the strongly local part. To deal with it, we employ LeJan's formula given by (2.23).

Recall that $w_n(x) = n^{-2}\psi_n(n^2 v(x))$ and note that

$$\psi_n'(s) = \begin{cases} 1 & \text{when } |s| < n, \\ 0 & \text{when } n < |s| < n^3, \\ 1 & \text{when } |s| > n^3. \end{cases}$$

Since ψ_n are Lipschitz functions equal to zero at 0, and in view of the fact that $v \in \mathcal{D}(\mathcal{E})$ is quasi-continuous, by LeJan's formula (2.23), we have

$$\begin{aligned}\mathcal{E}^c[w_n] &= \int\limits_E (\psi_n'(n^2 v(x)))^2 \mu^c_{[v]}(\mathrm{d}x) \\ &= \int\limits_E \mathbb{1}_{(0,1/n)\cup(n,+\infty)}(|v(x)|)\mu^c_{[v]}(\mathrm{d}x).\end{aligned}$$

Therefore, by the dominated convergence theorem, we obtain

$$\lim_{n\to+\infty} \mathcal{E}^c[w_n] = 0. \tag{7.29}$$

We follow the same steps with respect to $v_{\alpha,n}$ and $v_{2-\alpha,n}$. Recall that $v_{\alpha,n}(x) = n^{-2}\varphi_{\alpha,n}(n^2 v(x))$ and $v_{2-\alpha,n}(x) = n^{-2}\varphi_{2-\alpha,n}(n^2 v(x))$. Furthermore, note that $\varphi_{\alpha,n}$ and $\varphi_{2-\alpha,n}$ fulfill the assumptions of LeJan's formula. Therefore, again by (2.23)

$$\mathcal{E}^c(v_{\alpha,n}, v_{2-\alpha,n}) = \int\limits_E \varphi_{\alpha,n}'(n^2 v(x))\varphi_{2-\alpha,n}'(n^2 v(x))\mu^c_{[v]}(\mathrm{d}x).$$

However,

$$\varphi_{\alpha,n}'(s) = \begin{cases} 1 & \text{when } |s| < 1, \\ \alpha\,|s|^{\alpha-1} & \text{when } 1 < |s| < n^4, \\ 0 & \text{when } |s| > n^4. \end{cases}$$

Hence,

$$\varphi'_{\alpha,n}(s)\varphi'_{2-\alpha,n}(s) = \begin{cases} 1 & \text{when } |s| < 1, \\ \alpha(2-\alpha) & \text{when } 1 < |s| < n^4, \\ 0 & \text{when } |s| > n^4, \end{cases}$$

and it follows that

$$\mathcal{E}^c(v_{\alpha,n}, v_{2-\alpha,n}) = \int_E \Big(\mathbb{1}_{(0,1/n^2)}(|v(x)|) + \alpha(2-\alpha)\, \mathbb{1}_{(1/n^2,n^2)}(|v(x)|) \Big) \mu^c_{[v]}(\mathrm{d}x).$$

Utilizing the dominated convergence theorem and (2.22), we derive that

$$\lim_{n\to+\infty} \mathcal{E}^c(v_{\alpha,n}, v_{2-\alpha,n}) = \int_E \alpha(2-\alpha)\mu^c_{[v]}(\mathrm{d}x) = \alpha(2-\alpha)\mathcal{E}^c[v]. \tag{7.30}$$

Step 6. After gathering the results from Steps 2–5, we are ready to prove our goal (7.24). We denote

$$W(t) := \frac{1}{2t} \iint_{E\times E} \Phi_\alpha(x,y)\, P_t(\mathrm{d}x,\mathrm{d}y) - \alpha(2-\alpha)\mathcal{E}^c[v] - \frac{1}{2} \iint_{E\times E\setminus\mathrm{diag}} \Phi_\alpha(x,y)\, J(\mathrm{d}x,\mathrm{d}y),$$

for $t > 0$ and $l := \limsup_{t\to 0^+} |W(t)|$.

We claim that $l = 0$. Clearly, for $t > 0$ we have

$$\begin{aligned} |W(t)| \le & \left| \frac{1}{2t} \iint_{E\times E} \Phi_{\alpha,n}(x,y)\, P_t(\mathrm{d}x,\mathrm{d}y) - \mathcal{E}^c(v_{\alpha,n}, v_{2-\alpha,n}) - \frac{1}{2} \iint_{E\times E\setminus\mathrm{diag}} \Phi_{\alpha,n}(x,y)\, J(\mathrm{d}x,\mathrm{d}y) \right| \\ & + \frac{1}{2t} \iint_{E\times E} |\Phi_\alpha(x,y) - \Phi_{\alpha,n}(x,y)|\; P_t(\mathrm{d}x,\mathrm{d}y) \\ & + |\mathcal{E}^c(v_{\alpha,n}, v_{2-\alpha,n}) - \alpha(2-\alpha)\mathcal{E}^c[v]| \\ & + \frac{1}{2} \iint_{E\times E\setminus\mathrm{diag}} |\Phi_{\alpha,n}(x,y) - \Phi_\alpha(x,y)|\; J(\mathrm{d}x,\mathrm{d}y). \end{aligned}$$

By (7.26) from Step 3, the first term on the right-hand side converges to zero as $t \to 0^+$. Thus,

$$\begin{aligned} l \le & \limsup_{t\to 0^+} \frac{1}{2t} \iint_{E\times E} |\Phi_\alpha(x,y) - \Phi_{\alpha,n}(x,y)|\; P_t(\mathrm{d}x,\mathrm{d}y) \\ & + |\mathcal{E}^c(v_{\alpha,n}, v_{2-\alpha,n}) - \alpha(2-\alpha)\mathcal{E}^c[v]| \\ & + \frac{1}{2} \iint_{E\times E\setminus\mathrm{diag}} |\Phi_{\alpha,n}(x,y) - \Phi_\alpha(x,y)|\; J(\mathrm{d}x,\mathrm{d}y). \end{aligned}$$

Applying (7.23) to each of the integrands on the right-hand side yields that

$$\begin{aligned} l \le & \limsup_{t\to 0^+} \left(\frac{\varepsilon_{\alpha,n}}{2t} \iint_{E\times E} (v(y)-v(x))^2\, P_t(\mathrm{d}x,\mathrm{d}y) + \frac{90 D_\alpha}{2t} \iint_{E\times E} (w_n(y)-w_n(x))^2\, P_t(\mathrm{d}x,\mathrm{d}y) \right) \\ & + |\mathcal{E}^c(v_{\alpha,n}, v_{2-\alpha,n}) - \alpha(2-\alpha)\mathcal{E}^c[v]| \\ & + \frac{\varepsilon_{\alpha,n}}{2} \iint_{E\times E\setminus\mathrm{diag}} (v(y)-v(x))^2\, J(\mathrm{d}x,\mathrm{d}y) + 45 D_\alpha \iint_{E\times E\setminus\mathrm{diag}} (w_n(y)-w_n(x))^2\, J(\mathrm{d}x,\mathrm{d}y). \end{aligned}$$

Next, we use (7.25) and (7.27) from Steps 2 and 4 to convert the first two terms on the right-hand side. Then, we combine them with the last two terms and obtain

$$\begin{aligned} l \le \varepsilon_{\alpha,n}\mathcal{E}^c[v] + \varepsilon_{\alpha,n} \iint\limits_{E\times E\setminus \mathrm{diag}} (v(y)-v(x))^2\, J(\mathrm{d}x,\mathrm{d}y) \\ + 90D_\alpha \mathcal{E}^c[w_n] + 90D_\alpha \iint\limits_{E\times E\setminus \mathrm{diag}} (w_n(y)-w_n(x))^2\, J(\mathrm{d}x,\mathrm{d}y) \\ + \left|\mathcal{E}^c(v_{\alpha,n}, v_{2-\alpha,n}) - \alpha(2-\alpha)\mathcal{E}^c[v]\right|. \end{aligned}$$

The above inequality holds for every $n \ge 2$. Moreover, the left-hand side does not depend on n. The first two terms on the right-hand side tend to zero as $n \to +\infty$ by the definition of $\varepsilon_{\alpha,n}$. The third and the fifth terms converge to zero according to (7.29) and (7.30) from Step 5. The fourth one vanishes by (7.28) from Step 4. Summarizing, the left-hand side l is necessarily zero, as claimed. We have just proved our goal (7.24), or equivalently that $u \in \mathcal{D}(\mathcal{E}_p)$ and that $\mathcal{E}_p[u]$ is given by the analogue of the Beurling–Deny formula (7.3).

Rest of the proof. It remains to prove (7.2) and the final claim of the theorem. Let $u \in \mathcal{D}(\mathcal{E}_p)$ be quasi-continuous. Then, in view of (7.1) also $u^{\langle p/2\rangle} \in \mathcal{D}(\mathcal{E})$ and the estimate (7.2) follows from (2.37). In particular, both integrals in (7.3): $\mathcal{E}_p^j[u]$, $\mathcal{E}_p^k[u]$ are finite because $\mathcal{E}_2[u^{\langle p/2\rangle}]$ is finite as well.

Assume now that $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$ is maximally defined, $u \in L^p(m)$ has a quasi-continuous modification $\widetilde{u}$ and the integrals $\mathcal{E}_p^j[\widetilde{u}]$, $\mathcal{E}_p^k[\widetilde{u}]$ in (7.3) are finite. Then,

$$\begin{aligned} \mathcal{E}_2^j[\widetilde{u}^{\langle p/2\rangle}] &\le C_p \mathcal{E}_p^j[\widetilde{u}], \\ \mathcal{E}_2^k[\widetilde{u}^{\langle p/2\rangle}] &= \mathcal{E}_p^k[\widetilde{u}] \end{aligned}$$

are also finite. Here, we used Lemma 2.3. By our maximality assumption on $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$, this implies that $\widetilde{u}^{\langle p/2\rangle} \in \mathcal{D}(\mathcal{E})$ and hence $u^{\langle p/2\rangle} \in \mathcal{D}(\mathcal{E})$. By the characterization (7.1), $u \in \mathcal{D}(\mathcal{E}_p)$, which ends the proof. □

7.3 Example: Euclidean space

Consider a regular Dirichlet form $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$ on a domain E on the Euclidean space $E \subseteq \mathbb{R}^d$ and suppose that $C_c^\infty(E) \subseteq \mathcal{D}(\mathcal{E})$. In this case the strongly local term admits a more explicit form:

$$\mathcal{E}^c(u,v) = \int\limits_E \sum_{i,j=1}^n \frac{\partial u}{\partial x_i}(x)\frac{\partial u}{\partial x_j}(x)\, \nu_{i,j}(\mathrm{d}x), \quad u,v \in C_c^\infty(E)$$

and, consequently, the following version of the Beurling–Deny formula holds:

$$\begin{aligned} \mathcal{E}(u,v) = \int\limits_E \sum_{i,j=1}^d \frac{\partial u}{\partial x_i}(x)\frac{\partial u}{\partial x_j}(x)\, \nu_{i,j}(\mathrm{d}x) \\ + \frac{1}{2} \iint\limits_{E\times E\setminus \mathrm{diag}} (u(y)-u(x))(v(y)-v(x))\, J(\mathrm{d}x,\mathrm{d}y) \\ + \int\limits_E u(x)v(x)\, k(\mathrm{d}x), \quad u,v \in C_c^\infty(E). \end{aligned} \tag{7.31}$$

Here, $\nu_{i,j}$ are some positive Radon measures on E $(1 \le i, j \le d)$ satisfying

$$\sum_{i,j=1}^{d} \xi_i \xi_j \nu_{i,j}(K) \ge 0, \quad \nu_{i,j}(K) = \nu_{j,i}(K), \quad 1 \le i, j \le d$$

for any $\xi \in \mathbb{R}^d$ and any compact set $K \subseteq E$. We refer to Theorem 3.2.3 in Fukushima, Oshima, and Takeda [45] for this fact.

In view of that, according to Theorem 7.1, we obtain the following analogue of the Beurling–Deny formula for the p-form $\mathcal{E}_p$ on the Euclidean space for smooth functions:

$$\begin{aligned} \mathcal{E}_p[u] = (p-1) \int_E |u(x)|^{p-2} \sum_{i,j=1}^{d} \frac{\partial u}{\partial x_i}(x) \frac{\partial u}{\partial x_j}(x)\, \nu_{i,j}(\mathrm{d}x) \\ + \frac{1}{2} \iint_{E \times E \setminus \mathrm{diag}} (u(y) - u(x))(u^{\langle p-1 \rangle}(y) - u^{\langle p-1 \rangle}(x))\, J(\mathrm{d}x, \mathrm{d}y) \\ + \int_E |u(x)|^p \ k(\mathrm{d}x), \qquad u \in C_c^\infty(E). \end{aligned} \tag{7.32}$$

Additionally, whenever (7.31) remains valid for a broader class $\mathcal{D}$ of admissible functions u, v, then (7.32) also holds for every u such that $u^{\langle p/2 \rangle} \in \mathcal{D}$. For example, it is satisfied for an appropriate Sobolev space $W_0^{1,2}(E)$ or $W^{1,2}(E)$ when the strongly local part of $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$ corresponds to a uniformly elliptic second order operator. Compare above observations with (1.5).

According to Corollary 7.4, whenever for an arbitrary $f \in L^p(m)$ we have $P_t f \in \mathcal{D}$ for all $t > 0$, then the following Hardy–Stein identity holds:

$$\begin{aligned} \int_E |f(x)|^p \,\mathrm{d}x - \lim_{T \to +\infty} \|P_T f\|_p^p = \\ = p(p-1) \int_0^{+\infty} \int_E |P_t f(x)|^{p-2} \sum_{i,j=1}^{d} \frac{\partial P_t f}{\partial x_i}(x) \frac{\partial P_t f}{\partial x_j}(x)\, \nu_{i,j}(\mathrm{d}x)\mathrm{d}t \\ + \frac{p}{2} \int_0^{+\infty} \iint_{E \times E \setminus \mathrm{diag}} (P_t f(y) - P_t f(x))((P_t f)^{\langle p-1 \rangle}(y) - (P_t f)^{\langle p-1 \rangle}(x))\, J(\mathrm{d}x, \mathrm{d}y)\mathrm{d}t \\ + p \int_0^{+\infty} \int_E |P_t f(x)|^p \, k(\mathrm{d}x)\mathrm{d}t. \end{aligned} \tag{7.33}$$

Chapter 8

Applications in Littlewood–Paley theory

Bañuelos, Bogdan, and Luks investigate in their paper [5] the following square functions:

$$G(x) := \left(\frac{1}{2}\int_0^{+\infty}\int_{E\setminus\{x\}} (P_tf(y) - P_tf(x))^2\, J(x, \mathrm{d}y)\mathrm{d}t\right)^{1/2}, \tag{8.1}$$

$$\widetilde{G}(x) := \left(\int_0^{+\infty}\int_{E\setminus\{x\}} (P_tf(y) - P_tf(x))^2\, \chi(P_tf(x), P_tf(y))J(x, \mathrm{d}y)\mathrm{d}t\right)^{1/2}, \tag{8.2}$$

where f is a fixed function in $L^p(m)$, $\chi(s,t) := \mathbb{1}_{\{|s|>|t|\}} + \frac{1}{2}\mathbb{1}_{\{|s|=|t|\}}$, and $J(x, \mathrm{d}y)$ is the kernel of the jumping measure J: $J(\mathrm{d}x, \mathrm{d}y) = J(x, \mathrm{d}y)m(\mathrm{d}x)$. Working in the context of a pure-jump Lévy processes, the authors show that the p-norm of G cannot be bounded in terms of the p-norm of f. For this, we refer to Example 2 in [5] or to Example 8.12 below. However, for the square function $\widetilde{G}$ the following Littlewood–Paley estimates hold:

$$c_p \left\|f\right\|_p \leq \left\|\widetilde{G}\right\|_p \leq C_p \left\|f\right\|_p, \quad 1 < p < \infty,$$

for some constants $c_p, C_p > 0$.

Two separate tools were used in [5] to prove this result: the Hardy–Stein identity and the Burkholder–Davies–Gundy inequality. The second one employs the following martingale: $M_t := P_{T-t}f(X_t) - P_t(X_0)$. The authors utilized Itô's formula to connect the square bracket $\langle M\rangle$ with the third square function,

$$H(x) := \left(\frac{1}{2}\int_0^{+\infty}\int_E\int_{E\setminus\{z\}} (P_tf(y) - P_tf(z))^2\, J(z, \mathrm{d}y)P_t(x, \mathrm{d}z)\mathrm{d}t\right)^{1/2}. \tag{8.3}$$

Then, they proved the upper bound $\left\|H\right\|_p \leq C_p \left\|f\right\|_p$ for $2 \leq p < \infty$. Secondly, the authors showed that $\widetilde{G} \leq 2H$ using "time doubling"; see the proof of Lemma 4.2 in [5]. Finally, they obtain the analogous upper bound for $\widetilde{G}$ by combining the aforementioned observations. In Li and Wang [63], a similar concept was attempted, but an error was identified. For details, we refer to the discussion at the end of Section 8.5.1. The lower bound $c_p \left\|f\right\|_p \leq \left\|\widetilde{G}\right\|_p$ in the case of $1 < p < 2$ was provided in [5] utilizing the ultracontractivity of the semigroup.

The study in this chapter is focused on a more general case; see Assumption 8.4 below. In view of this, we cannot use Itô's formula and do not assume ultracontractivity of the semigroup. Instead, we use the Revuz correspondence to establish the connection between the square bracket $\langle M\rangle$ and the square function H.

Moreover, the inequality $\widetilde{G} \leq 2H$ does not hold in general. This is an immediate conclusion from some of our main results in this chapter.

Theorem 8.1. *Let $1 < p < \infty$. Under Assumptions 3.1 and 8.4 there are constants $c_p, C_p > 0$ such that*

$$c_p \left\|f\right\|_p \leq \left\|\widetilde{G}\right\|_p, \quad 2 \leq p < \infty,$$

and

$$\left\|\widetilde{G}\right\|_p \leq C_p \left\|f\right\|_p, \quad 1 < p \leq 2.$$

However, there is no universal constant $\tilde{C}_p > 0$ such that

$$\left\|\widetilde{G}\right\|_p \leq C_p \left\|f\right\|_p, \quad 2 < p < \infty. \tag{8.4}$$

The above estimates are presented rigorously in Section 8.5. In particular, for (8.4) we refer to Example 8.16 below.

Theorem 8.2. *Let $1 < p < \infty$. Under Assumptions 3.1 and 8.4 there are constants $c_p, C_p > 0$ such that*

$$c_p \left\|f\right\|_p \leq \left\|H\right\|_p, \quad 1 < p \leq 2 \text{ or } 3 \leq p < \infty,$$

and

$$\left\|H\right\|_p \leq C_p \left\|f\right\|_p, \quad 2 \leq p < \infty.$$

However, there is no universal constant $\tilde{C}_p > 0$ such that

$$\left\|H\right\|_p \leq \tilde{C}_p \left\|f\right\|_p, \quad 1 < p < 2. \tag{8.5}$$

The above results are presented rigorously in Section 8.6. Particularly, the counterexample disproving (8.5) is analogous to Example 2 from [5]. We prove some parts of these results following the approach from [5] and applying our general Hardy–Stein identity from Corollary 7.4.

In order to overcome the fact that (8.5) is not true, we introduce the fourth square function:

$$\widetilde{H}(x) := \left(\int_0^{+\infty} \int_E \int_{E\setminus\{z\}} (P_t f(y) - P_t f(z))^2 \chi(P_t f(z), P_t f(y)) J(z, \mathrm{d}y) P_t(x, \mathrm{d}z) \mathrm{d}t\right)^{1/2}. \tag{8.6}$$

In Section 8.7 we prove the following Littlewood–Paley estimates of the square function $\widetilde{H}$.

Theorem 8.3. *Let $1 < p < \infty$. Under Assumptions 3.1 and 8.4 there are constants $c_p, C_p > 0$ such that*

$$c_p \|f\|_p \leq \left\|\widetilde{H}\right\|_p, \quad 3 \leq p < \infty,$$

and

$$\left\|\widetilde{H}\right\|_p \leq C_p \|f\|_p, \quad 2 \leq p < \infty.$$

For now, it is an open question whether Littlewood–Paley estimates hold in the other cases.

For more examples of square functions utilized in the literature, we refer to the extensive compilation in Section 1 of [63].

8.1 Preliminaries

For non-negative functions f and g we write $f(x) \lesssim g(x)$ (resp. $f(x) \gtrsim g(x)$) to indicate that there exists a positive constant C_p depending only on $1 \leq p \leq \infty$ such that $f(x) \leq C_p g(x)$ (resp. $f(x) \geq C_p g(x)$) for all the considered arguments x. When $f(x) \lesssim g(x)$ and $f(x) \gtrsim g(x)$, we write $f(x) \asymp g(x)$.

For two sequences $(a_n)_{n\in\mathbb{N}}$ and $(b_n)_{n\in\mathbb{N}}$ of non-negative real numbers we write $a_n \sim b_n$ to indicate that those sequences are asymptotically equal, that is, $a_n/b_n \to 1$ as $n \to +\infty$. Similarly, for non-negative functions f and g, we will write $f(x) \sim g(x)$ as $x \to x_0$ whenever $\lim_{x\to x_0} f(x)/g(x) = 1$ and say that those functions are asymptotically equal as $x \to x_0$.

8.1.1 Conditions on Dirichlet Form

In this chapter we work under the following conditions.

Assumption 8.4.

(JK) There exists the jumping kernel of the jumping measure J, i.e., for every $x \in E$ there exists a Radon measure $J(x, \cdot)$ such that

$$J(\mathrm{d}x, \mathrm{d}y) = J(x, \mathrm{d}y)m(\mathrm{d}x).$$

(PJ) The regular Dirichlet form $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$ is pure-jump, that is, the strongly local part $\mathcal{E}^c$ in the Beurling–Deny formula vanishes.

(K0) The killing measure k vanishes.

The assumption (JK) does not impose significant restrictions. Indeed, in general, the jumping measure J may be written as follows: $J(\mathrm{d}x, \mathrm{d}y) = N(x, \mathrm{d}y)\nu(\mathrm{d}x)$, where $(N(x, \mathrm{d}y), (H_t)_{t\geq 0})$ is the so-called *Lévy system* and ν is the Revuz measure of the AF $(H_t)_{t\geq 0}$; see (5.3.6) in Fukushima, Oshima, and Takeda [45] and the discussion therein. In many cases the Dirichlet form satisfies the stronger condition, for example J is absolute continuous. See (J) from Assumption 4.1.

The assumptions (PJ) and (K0) together ensure that the Beurling-Deny decomposition of the Dirichlet form consists only of the jumping part. Indeed, in view of the above assumptions, the Dirichlet form takes the following form:

$$\mathcal{E}(u,v) = \frac{1}{2} \iint\limits_{E\times E\setminus\mathrm{diag}} (\tilde{u}(y) - \tilde{u}(x))(\tilde{v}(y) - \tilde{v}(x))\, J(x,\mathrm{d}y)m(\mathrm{d}x), \tag{8.7}$$

where $\tilde{u}, \tilde{v}$ denote quasi-continuous versions of arbitrary $u, v \in \mathcal{D}(\mathcal{E})$. Together with (JK) the energy measure $\mu_{[u]}$ defined in Section 2.4 reduces to $\mathrm{d}\mu_{[u]} = \Gamma[u]\mathrm{d}m$, where

$$\Gamma[u](x) := \frac{1}{2} \int\limits_{E\setminus\{x\}} (\tilde{u}(y) - \tilde{u}(x))^2\, J(x,\mathrm{d}y), \quad u \in \mathcal{D}(\mathcal{E}), \tag{8.8}$$

is the *carré du champ operator*. It is known that $\Gamma\colon \mathcal{D}(\mathcal{E}) \to L^1(m)$ is continuous, where $\mathcal{D}(\mathcal{E})$ is the space equipped with the norm $\sqrt{\mathcal{E}[\cdot]}$. We refer to Proposition 4.1.3 in Bouleau and Hirsch [20].

We will also consider

$$\widetilde{\Gamma}[u](x) := \int\limits_{E\setminus\{x\}} (\tilde{u}(y) - \tilde{u}(x))^2\, \chi(\tilde{u}(x), \tilde{u}(y))\, J(x,\mathrm{d}y), \quad u \in \mathcal{D}(\mathcal{E}), \tag{8.9}$$

where

$$\chi(s,t) := \begin{cases} 1 & \text{if } |s| > |t|, \\ \frac{1}{2} & \text{if } |s| = |t|, \\ 0 & \text{if } |s| < |t|, \end{cases} \tag{8.10}$$

is the characteristic function of the set $\{\, (s,t) : |s| > |t| \,\}$, with a minor modification when $|s| = |t|$. Note that $\chi(s,t) + \chi(t,s) = 1$. In particular, by the symmetry of the jumping measure J,

$$\int\limits_E \Gamma[u]\,\mathrm{d}m = \int\limits_E \widetilde{\Gamma}[u]\,\mathrm{d}m = \mathcal{E}[u]. \tag{8.11}$$

Under Assumption 8.4 the following implication is true:

$$u \in \mathcal{D}(A_2) \Rightarrow u^2 \in \mathcal{D}(A_1). \tag{8.12}$$

Moreover, the carré du champ operator on $\mathcal{D}(A_2)$ takes the following form:

$$\Gamma[u] = \frac{1}{2} A_1(u^2) - uA_2u, \quad u \in \mathcal{D}(A_2). \tag{8.13}$$

We refer to Theorem 4.2.2 in [20] for the above statement.

8.1.2 Square functions

Fix $f \in L^p(m)$. In this work, we study the *square functions* (or *Littlewood–Paley functions*) given by (8.1), (8.2), (8.3), and (8.6). Using introduced earlier notation, these equalities reads

$$G(x) := \left(\int_0^{+\infty} \Gamma[P_t f](x)\, \mathrm{d}t \right)^{1/2},$$

$$\widetilde{G}(x) := \left(\int_0^{+\infty} \widetilde{\Gamma}[P_t f](x)\, \mathrm{d}t \right)^{1/2},$$

$$H(x) := \left(\int_0^{+\infty} P_t \Gamma[P_t f](x)\, \mathrm{d}t \right)^{1/2},$$

$$\widetilde{H}(x) := \left(\int_0^{+\infty} P_t \widetilde{\Gamma}[P_t f](x)\, \mathrm{d}t \right)^{1/2}.$$

Here and below, we always assume that the quasi-continuous version of $P_t f$ is taken. Clearly, $\widetilde{G} \le \sqrt{2} G$ and $\widetilde{H} \le \sqrt{2} H$. By (8.11), the L^2-norms of G, $\widetilde{G}$, H, and $\widetilde{H}$ are all equal to

$$\int_0^{+\infty} \int_E \Gamma[P_t f]\, \mathrm{d}m \mathrm{d}t.$$

As an immediate consequence of the Hardy–Stein identity (Corollary 7.4) for $p = 2$, we present the following fact.

Proposition 8.5. *The following equivalence between L^2-norms holds:*

$$\|G\|_2 = \left\|\widetilde{G}\right\|_2 = \|H\|_2 = \left\|\widetilde{H}\right\|_2 = 2 \|f\|_2.$$

8.1.3 Additive functionals and Revuz correspondence

Let $(X_t)_{t\ge 0}$ be a Hunt process (not necessarily symmetric) with respect to the minimum completed admissible filtration $(\mathcal{F}_t)_{t\ge0}$. A set $B \subseteq E$ is *nearly Borel* (relative to $(X_t)_{t\ge0}$) if for each probabilistic measure μ there exist Borel sets $B_1, B_2 \subseteq E$ such that $B_1 \subseteq B \subseteq B_2$ and $\mathbb{P}_\mu(X_t \in B_2 \setminus B_1$ for some $t) = 0$; see Definition (10.21) in Blumenthal and Getoor [11] or [45, p. 392]. Denote by $\sigma_B := \inf\{\, t > 0 : X_t \in B \,\}$ the *hitting time* of a set B for $(X_t)_{t\ge0}$. A nearly Borel set $D \subseteq E$ is *polar* if

$$\int_E \mathbb{P}_x(\sigma_D < +\infty)\, m(\mathrm{d}x) = 0;$$

see Definition (6.3) in Getoor and Sharpe [48]. We say that a nearly Borel set $\widetilde{E} \subseteq E$ is *absorbing* if $\mathbb{P}_x(\sigma_{E\setminus\widetilde{E}} < +\infty) = 0$ for all $x \in \widetilde{E}$; see [48, p. 17]. A nearly Borel set $N \subseteq E$ is *inessential* if N is polar and N^c is absorbing; see Definition (6.8) in [48].

We call a family of functions $(A_t)_{t\ge0}$ on Ω with values in $[-\infty, +\infty]$ an *additive functional* (AF in abbreviation), if there exist a *defining set* $\Lambda \in \mathcal{F}_\infty$ and an inessential set N (called an *exceptional set* for $(A_t)_{t\ge0}$) such that the following conditions hold:

(A.1) A_t is measurable with respect to $\mathcal{F}_t$.

(A.2) $\mathbb{P}_x(\Lambda) = 1$ for all $x \in E \setminus N$.

(A.3) $\theta_t[\Lambda] \subseteq \Lambda$ for all $t > 0$.

(A.4) $A_0(\omega) = 0$ for all $\omega \in \Lambda$.

(A.5) $|A_t(\omega)| < +\infty$ for all $\omega \in \Lambda$, $t < \zeta(\omega)$.

(A.6) $A_t(\omega) = A_{\zeta(\omega)}(\omega)$ for all $\omega \in \Lambda$, $t \geq \zeta(\omega)$.

(A.7) $A_{s+t}(\omega) = A_s(\omega) + A_t(\theta_s\omega)$ for all $\omega \in \Lambda$, $t, s \geq 0$.

(A.8) For each $\omega \in \Lambda$, the function $[0, \zeta(\omega)) \ni t \mapsto A_t(\omega)$ is right continuous and has the left limit.

We say that $(A_t)_{t\geq 0}$ is the *positive additive functional* (PAF in abbreviation), when we replace condition (A.8) in the above definition with the stronger one:

(A.8') For each $\omega \in \Lambda$, the function $[0, +\infty) \ni t \mapsto A_t(\omega)$ is right continuous and non-negative. Moreover, for every $t > 0$, the function $[0, t) \ni s \mapsto A_{t-s}(\theta_s\omega)$ is right continuous.

We say that $(A_t)_{t\geq 0}$ is the *positive continuous additive functional* (PCAF in abbreviation), when we replace condition (A.8) in the above definition with the stronger one:

(A.8") For each $\omega \in \Lambda$, the function $[0, +\infty) \ni t \mapsto A_t(\omega)$ is continuous and non-negative.

We refer to Definitions 3.16(b) in Fitzsimmons [41] and to Chapter 2 of Fukushima [37]. See also Section 5.1 of [45]. We say that the additive functionals $(A_t)_{t\geq 0}$ and $(B_t)_{t\geq 0}$ are *equivalent*, if $\int_E \mathbb{P}_x(A_t = B_t)\, m(\mathrm{d}x) = 1$ for all $t > 0$.

Let $((M_t)_{t\geq 0}, (\mathcal{F}_t)_{t\geq 0}, (\mathbb{P}_x)_{x\in E_\Delta})$ be a martingale. Then there exists a unique integrable and predictable increasing process $(\langle M\rangle_t)_{t\geq 0}$ null at zero such that $((M_t^2 - \langle M\rangle_t)_{t\geq 0}, (\mathcal{F}_t)_{t\geq 0}, (\mathbb{P}_x)_{x\in E_\Delta})$ is a martingale. If, in addition, $(M_t)_{t\geq 0}$ is an AF, then $(\langle M\rangle_t)_{t\geq 0}$ is a unique PCAF (up to the equivalence of additive functionals) such that $((M_t^2 - \langle M\rangle_t)_{t\geq 0}, (\mathcal{F}_t)_{t\geq 0}, (\mathbb{P}_x)_{x\in E_\Delta})$ is a martingale. The process $(\langle M\rangle_t)_{t\geq 0}$ is called the *sharp bracket* (or *predictable quadratic variation*) of $(M_t)_{t\geq 0}$. For more details, we refer to Section A.3 of [45]. See also Section 4 of Trutnau [107] and Théorème 3 in Meyer [77, III].

As we will see later, every PCAF can be associated with a unique Borel measure in the class of so-called smooth measures. But before we define this class of measures, we require some notions following Fitzsimmons and Getoor [43]. Let $\mathcal{B}^*(E)$ be the *universal completion* of $\mathcal{B}(E)$, that is, $\mathcal{B}^*(E) := \bigcap_\mu \mathcal{B}^\mu(E)$, where $\mathcal{B}^\mu(E)$ is the completion of $\mathcal{B}(E)$ with respect to the measure μ and the above intersection is taken over all probabilistic measures on $(E, \mathcal{B}(E))$. Let $\mathcal{B}^e(E)$ be the σ-field generated by $\mathcal{B}^*(E)$-measurable functions which are α-excessive for some $\alpha \geq 0$. A set $B \in \mathcal{B}^e(E)$ is called *semi-polar* provided

$$\int_E \mathbb{P}_x(X_t \in B \text{ for uncountably many } t)\, m(\mathrm{d}x) = 0.$$

Denote by $\tau_B := \inf\{\, t > 0 : X_t \notin B \,\}$ the *exit time* of $(X_t)_{t\geq 0}$ from a set B. In particular, $\tau_B = \sigma_{E_\Delta \setminus B}$. We say that an increasing sequence of sets $(B_n)_{n\in\mathbb{N}}$ in $\mathcal{B}^e(E)$ is a *generalized nest* if

$$\int_E \mathbb{P}_x\left(\lim_{n\to+\infty} \tau_{B_n} < \zeta\right) m(\mathrm{d}x) = 0.$$

A set B is called *finely open* if for each $x \in B$ there exists a nearly Borel set $B_x \supseteq B^c$ such that $\mathbb{P}_x(\sigma_{B_x} > 0) = 1$; see Definition (4.1) in [11]. Finally, let us call a Borel measure ν *smooth* if it charges no semi-polar set and admits an associated generalized nest $(G_n)_{n\in\mathbb{N}}$ of finely open sets with $\nu(G_n) < +\infty$ for each $n \in \mathbb{N}$.

If we additionally assume that the Hunt process $(X_t)_{t\geq 0}$ is symmetric, then it is associated with some regular Dirichlet form $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$, and therefore the class of smooth measures can be characterized more easily using the capacity notion given by (2.18) and (2.19). Indeed, an increasing sequence of sets $(B_n)_{n\in\mathbb{N}}$ is a generalized nest if and only if

$$\lim_{n\to+\infty} \mathrm{Cap}(K \setminus B_n) = 0 \quad \text{for any compact set } K.$$

Then, a Borel measure ν is smooth if and only if it charges no set of zero capacity and admits an associated generalized nest $(F_n)_{n\in\mathbb{N}}$ of closed sets such that $\nu(F_n) < +\infty$ for each $n \in \mathbb{N}$. For these and other equivalent definitions of smooth measure, capacity, and generalized nest, we refer to [43]. See also [45, p. 83] or Appendix B of Li and Ying [64].

Let $(A_t)_{t\geq 0}$ be an additive functional and $\tilde{m}$ be an excessive measure (for $(X_t)_{t\geq 0}$). The *Revuz measure* ν of $(A_t)_{t\geq 0}$ related to a reference measure $\tilde{m}$ is defined by

$$\int_E f \,\mathrm{d}\nu := \lim_{t\to 0^+} \frac{1}{t} \int_E \mathbb{E}_x \left(\int_{(0,t]} f(X_s)\,\mathrm{d}A_s \right) \tilde{m}(\mathrm{d}x), \quad f \in \mathcal{B}^e(E)_+, \tag{8.14}$$

or, equivalently,

$$\int_E f \,\mathrm{d}\nu = \lim_{\beta\to+\infty} \beta \int_E \mathbb{E}_x \left(\int_0^{+\infty} e^{-\beta t} f(X_t)\,\mathrm{d}A_t \right) \tilde{m}(\mathrm{d}x), \quad f \in \mathcal{B}^e(E)_+.$$

Here, $\mathcal{B}^e(E)_+$ is the class of non-negative $\mathcal{B}^e(E)$-measurable functions. We say that such $(A_t)_{t\geq 0}$ and ν are in the *Revuz correspondence* in honor of Revuz; see [91].

The simple example of ν and $(A_t)_{t\geq 0}$ in the Revuz correspondence is provided by a Borel function g. Let

$$A_t(\omega) := \int_0^t g(X_s(\omega))\,\mathrm{d}s, \quad t \geq 0. \tag{8.15}$$

Then $(A_t)_{t\geq 0}$ is AF, PAF, and PCAF when g is bounded, non-negative and bounded non-negative, respectively. Moreover $(A_t)_{t\geq 0}$ is in the Revuz correspondence (related to the reference measure m) with measure $g(x)m(\mathrm{d}x)$; see [37, p. 41].

It is known that the Revuz correspondence is a one-to-one correspondence between the class of smooth measures and the class of PCAF (up to equivalence of additive functionals). For this statement, we refer to Theorem (3.11) in [43]. Extensive research has been dedicated to the Revuz correspondence. Notable works include Getoor [47], [48], and Fitzsimmons and Getoor [42, 43]. We may also refer to Section 75 in the book of Sharpe [94].

Let us assume additionally that the Hunt process $(X_t)_{t\geq 0}$ is symmetric, i.e., it is associated with some regular Dirichlet form and $(A_t)_{t\geq 0}$ is PCAF. In this case, the Revuz correspondence can be characterized by the following formula:

$$\int_E fh\,\mathrm{d}\nu = \lim_{t\to 0^+}\frac{1}{t}\int_E h(x)\mathbb{E}_x\left(\int_0^t f(X_s)\,\mathrm{d}A_s\right)\tilde{m}(\mathrm{d}x) \tag{8.16}$$
$$= \lim_{\beta\to+\infty}\beta\int_E h(x)\mathbb{E}_x\left(\int_0^{+\infty} e^{-\beta t}f(X_t)\,\mathrm{d}A_t\right)\tilde{m}(\mathrm{d}x),$$

where h is any α-excessive function ($\alpha \geq 0$) and f is any non-negative Borel function; see Section 5.1 of [45]. See also Chapter 2 of Fukushima [37] and Appendix B of [64]. For the Revuz correspondence in the case of non-symmetric generalizations of the Dirichlet form, we refer to [41, 64] (for semi-Dirichlet form) and [107, Section 3] (for generalized Dirichlet form).

8.1.4 Space-time process

Let $(X_t)_{t\geq 0}$ be a symmetric Hunt process associated with the regular Dirichlet form $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$. The Hunt process $(X_t)_{t\geq 0}$ is *progressively measurable*, i.e., for each $T \geq 0$,

$$([0,T]\times\Omega, \mathcal{B}([0,T])\times\mathcal{F}_t) \ni (t,\omega) \mapsto X_t(\omega) \in (E_\Delta, \mathcal{B}(E_\Delta))$$

is measurable. Recall that $\mathcal{B}(I)$ is the σ-algebra of Borel sets of $I \subseteq \mathbb{R}$. Therefore, it is permitted to define so-called *space-time process*

$$(\breve{\Omega}, \breve{\mathcal{F}}, (\breve{X}_t)_{t\geq 0}, (\breve{\mathbb{P}}_{(t_0,x)})_{(t_0,x)\in\breve{E}_\Delta}),$$

where

$$\begin{aligned}
\breve{\Omega} &:= (0,\infty)\times\Omega,\\
\breve{\mathcal{F}} &:= \sigma\{\mathcal{B}((0,\infty))\times\mathcal{F}\},\\
\breve{X}_t(\tau,\omega) &:= (\tau+t, X_t(\omega)),\\
\breve{\mathbb{P}}_{(t_0,x)} &:= \delta_{t_0}\otimes\mathbb{P}_x,\\
\breve{E}_\Delta &:= (0,\infty)\times E\cup\{(\infty,\Delta)\}.
\end{aligned}$$

Here, δ_{t_0} is the Dirac delta measure at a point $t_0 \in (0,\infty)$. The point $\{(\infty,\Delta)\}$ is the cemetery state of $(\breve{X}_t)_{t\geq 0}$. The above process is the (non-symmetric) Hunt process with respect to the admissible filtration

$$\breve{\mathcal{F}}_t := \sigma\{\mathcal{B}((0,\infty))\times\mathcal{F}_t\}.$$

Denote $L^2(\mathrm{d}r\otimes m) := L^2((0,\infty)\times E, \mathcal{B}((0,\infty)\times E), \mathrm{d}r\otimes m)$. The semigroup $(\breve{P}_t)_{t\geq 0}$ of the space-time process is given by

$$\breve{P}_t f(r,x) = P_t[f(r+t,\cdot)](x), \quad f\in L^2(\mathrm{d}r\otimes m).$$

Its adjoint operator $\widehat{P}_t$ is given by

$$\widehat{P}_t g(r,x) = \mathbb{1}_{[0,r)}(t)P_t[g(r-t,\cdot)](x), \quad g\in L^2(\mathrm{d}r\otimes m).$$

We also consider the *backward space-time process* $(\widehat{X}_t)_{t\geq 0}$ defined on the same family of probability spaces $(\breve{\Omega}, \breve{\mathcal{F}}, (\breve{\mathbb{P}}_{(t_0,x)})_{(t_0,x)\in \breve{E}_\Delta})$ by

$$\widehat{X}_t(\tau,\omega) := \begin{cases} (\tau - t, X_t(\omega)) & \text{for } t < \tau, \\ (\infty, \Delta) & \text{otherwise.} \end{cases}$$

In other words, $(\widehat{X}_t)_{t\geq 0}$ is a Cartesian product of $(X_t)_{t\geq 0}$ and the uniform motion to the left on $(0,\infty)$ with killing at 0. Note that the process $(\widehat{X}_t)_{t\geq 0}$ is dual to $(\breve{X}_t)_{t\geq 0}$, i.e., $(\widehat{P}_t)_{t\geq 0}$ is the semigroup of $(\widehat{X}_t)_{t\geq 0}$.

For the construction of the above space-time processes, we refer to Section 16 of [94].

Note that for any $\Phi\colon \breve{\Omega} \to \mathbb{R}$ measurable with respect to $\breve{\mathcal{F}}_\infty$,

$$\mathbb{E}_{(t_0,x)}[\Phi|\breve{\mathcal{F}}_t] = \mathbb{E}_x[\Phi(t_0,\cdot)|\mathcal{F}_t], \quad \breve{\mathbb{P}}_{(t_0,x)}\text{-a.s.} \tag{8.17}$$

Indeed, for any $A \in \mathcal{B}((0,\infty))$ and $B \in \mathcal{F}_t$

$$\mathbb{E}_{(t_0,x)}(\Phi\, \mathbb{1}_{A\times B}) = \mathbb{1}_A(t_0)\mathbb{E}_x(\mathbb{E}_x[\Phi(t_0,\cdot)|\mathcal{F}_t]\, \mathbb{1}_B) = \mathbb{E}_{(t_0,x)}(\mathbb{E}_x[\Phi(t_0,\cdot)|\mathcal{F}_t]\, \mathbb{1}_{A\times B}).$$

At the end of this section we show the characterization of the Revuz correspondence for the particular case of a space-time process. This result will be beneficial in Section 8.2. Compare this with Appendix B of [64].

Lemma 8.6. *Let $(A_t)_{t\geq 0}$ be a PCAF related to the space-time process $(\breve{X}_t)_{t\geq 0}$ described above. The Revuz measure ν of $(A_t)_{t\geq 0}$ (related to the reference measure* $\mathrm{d}r \otimes m$*) satisfies*

$$\int\limits_{(0,+\infty)\times E} f\hat{h}\,\mathrm{d}\nu = \lim_{t\to 0^+} \frac{1}{t} \int\limits_{(0,+\infty)\times E} \hat{h}(r,x)\mathbb{E}_{(r,x)}\left(\int\limits_0^t f(\breve{X}_s)\,\mathrm{d}A_s\right) \mathrm{d}r m(\mathrm{d}x) \tag{8.18}$$

for any $f \in \mathcal{B}^e((0,\infty)\times E)_+$ and α-coexcessive function $\hat{h}$ ($\alpha \geq 0$).

Proof. Denote by $(\breve{X}^\alpha_t)_{t\geq 0}$ the α-subprocess of $(\breve{X}_t)_{t\geq 0}$, that is, the process obtained by killing $(\breve{X}_t)_{t\geq 0}$ with intensity α, i.e., at a time T_α, where T_α has the exponential distribution with parameter α. Similarly, let $(\widehat{X}^\alpha_t)_{t\geq 0}$ be the α-subprocess of $(\widehat{X}_t)_{t\geq 0}$. The semigroup of $(\breve{X}^\alpha_t)_{t\geq 0}$ is $(e^{-\alpha t}\breve{P}_t)_{t\geq 0}$. Therefore, since $\hat{h}$ is α-coexcessive for $(\breve{X}_t)_{t\geq 0}$, the measure $\hat{h}m$ is excessive for $(\breve{X}^\alpha_t)_{t\geq 0}$. Moreover, the mapping $[0,+\infty) \ni t \mapsto \hat{h}(\widehat{X}^\alpha_t)$ is $\mathbb{P}_m$-almost surely right continuous; see Proposition (4.2) and Theorem (4.8) in [11]. Here, $\mathbb{P}_m(\Lambda) := \int_E \mathbb{P}_x(\Lambda)\, m(\mathrm{d}x)$. By self-duality of $(X_t)_{t\geq 0}$ (see (2.40)) the mapping $(0,\zeta^\alpha) \ni t \mapsto \hat{h}(\breve{X}^\alpha_{t-})$ is $\mathbb{P}_m$-almost surely left continuous. Here, ζ^α is the lifetime of $(\breve{X}^\alpha_t)_{t\geq 0}$.

According to the above observations, we are allowed to use Corollary (8.11) from [48] and write

$$\begin{aligned} &\lim_{t\to 0^+} \frac{1}{t} \int\limits_{(0,+\infty)\times E} \mathbb{E}_{(r,x)}\left(\int\limits_0^t f(\breve{X}^\alpha_{s-})\hat{h}(\breve{X}^\alpha_{s-})\,\mathrm{d}A^\alpha_s\right) \mathrm{d}r m(\mathrm{d}x) \\ &\qquad = \lim_{t\to 0^+} \frac{1}{t} \int\limits_{(0,+\infty)\times E} \hat{h}(r,x)\mathbb{E}_{(r,x)}\left(\int\limits_0^t f(\breve{X}^\alpha_{s-})\,\mathrm{d}A^\alpha_s\right) \mathrm{d}r m(\mathrm{d}x). \end{aligned}$$

Here, $(A_t^\alpha)_{t\geq 0}$ is the PCAF $(A_t)_{t\geq 0}$ killed at time T_α.

Since the discontinuity points of the sample paths $t \mapsto \breve{X}_t^\alpha$ are at most countable and $(A_t^\alpha)_{t\geq 0}$ is continuous, we may rewrite the above formula as follows:

$$\lim_{t\to 0^+} \frac{1}{t} \int\limits_{(0,+\infty)\times E} \mathbb{E}_{(r,x)} \left(\int\limits_0^t f(\breve{X}_s^\alpha)\hat{h}(\breve{X}_s^\alpha)\, \mathrm{d}A_s^\alpha \right) \mathrm{d}r m(\mathrm{d}x)$$

$$= \lim_{t\to 0^+} \frac{1}{t} \int\limits_{(0,+\infty)\times E} \hat{h}(r,x) \mathbb{E}_{(r,x)} \left(\int\limits_0^t f(\breve{X}_s^\alpha)\, \mathrm{d}A_s^\alpha \right) \mathrm{d}r m(\mathrm{d}x).$$

Moreover, it is easy to see, that we may replace $\breve{X}_s^\alpha$ and A_s^α by $\breve{X}_s$ and A_s:

$$\lim_{t\to 0^+} \frac{1}{t} \int\limits_{(0,+\infty)\times E} \mathbb{E}_{(r,x)} \left(\int\limits_0^t f(\breve{X}_s)\hat{h}(\breve{X}_s)\, \mathrm{d}A_s \right) \mathrm{d}r m(\mathrm{d}x)$$

$$= \lim_{t\to 0^+} \frac{1}{t} \int\limits_{(0,+\infty)\times E} \hat{h}(r,x) \mathbb{E}_{(r,x)} \left(\int\limits_0^t f(\breve{X}_s)\, \mathrm{d}A_s \right) \mathrm{d}r m(\mathrm{d}x).$$

Finally, the left-hand side is equal to $\int_{(0,+\infty)\times E} f\hat{h}\, \mathrm{d}\nu$ by (8.14). □

8.1.5 Martingale

Let $(M_t)_{t\geq 0}$ be an arbitrary martingale. Denote by $([M]_t)_{t\geq 0}$ the *square bracket* of $(M_t)_{t\geq 0}$ given by

$$[M]_t := \langle M^c \rangle_t + \sum_{s\leq t} \Delta M_s^2,$$

where $\Delta X_t := X_t - X_{t^-}$ and $(M_t^c)_{t\geq 0}$ is the continuous part of $(M_t)_{t\geq 0}$; see (A.3.6) in [45]. Note that $[M]_t = \langle M \rangle_t$ for continuous $(M_t)_{t\geq 0}$. Let $1 \leq p < \infty$. According to the Burkholder–Davies–Gundy inequality:

$$\mathbb{E}_x |M_T - M_0|^p \asymp \mathbb{E}_x([M]_T)^{p/2}, \quad T > 0. \tag{8.19}$$

We emphasize that, in general, we cannot replace the square bracket $[M]_T$ by the sharp bracket $\langle M \rangle_T$. Nevertheless, it is known that for $2 \leq p \leq \infty$,

$$c_p \mathbb{E}_x(\langle M \rangle_t)^{p/2} \leq \mathbb{E}_x([M]_t)^{p/2}, \quad t > 0, \tag{8.20}$$

and, for $1 \leq p \leq 2$,

$$\mathbb{E}_x([M]_t)^{p/2} \leq C_p \mathbb{E}_x(\langle M \rangle_t)^{p/2} \quad t > 0, \tag{8.21}$$

for some constants $c_p, C_p > 0$. We refer to Remarque 4.2 in Lenglart, Lépingle, and Pratelli [62]. See also item (4.b'), Table 4.1, on p. 162 in Barlow, Jacka, and Yor [8].

In this chapter, our aim is to consider an auxiliary martingale. For a fixed $T > 0$ and $f \in L^p(m)$ we define

$$M_t := P_{(T-t)\vee 0} f(X_{t\wedge T}) - P_T f(X_0). \tag{8.22}$$

We will see later in this subsection that $(M_t)_{t\geq 0}$ is a martingale stopped at T with $M_t = f(X_T) - P_T f(X_0)$ for $t \geq T$. Note that $M_0 = 0$. We will also consider the space-time version of $(M_t)_{t\geq 0}$:

$$\breve{M}_t(\tau,\omega) := \begin{cases} P_{T-\tau-t}f(X_t(\omega)) - P_{T-\tau}f(X_0(\omega)) & \text{if } t+\tau \leq T, \\ f(X_{T-\tau}(\omega)) - P_{T-\tau}f(X_0(\omega)) & \text{if } \tau \leq T < t+\tau, \\ 0 & \text{if } T < \tau, \end{cases} \tag{8.23}$$

or, concisely,

$$\breve{M}_t(\tau,\omega) = P_{(T-\tau-t)\vee 0}f(X_{[t\wedge(T-\tau)(\omega)]\vee 0}) - P_{(T-\tau)\vee 0}f(X_0(\omega)).$$

Note that $\breve{M}_t(0,\cdot) = M_t$ and $\breve{M}_0 = 0$.

Proposition 8.7. *The process* $((\breve{M}_t)_{t\geq 0}, (\breve{\mathcal{F}}_t)_{t\geq 0}, (\breve{\mathbb{P}}_{(t_0,x)})_{(t_0,x)\in(0,+\infty]\times E_\Delta})$ *given by* (8.23) *is a martingale.*

Proof. Since $(\breve{M}_t)_{t\geq 0}$ is just a process $P_{T-\tau-t}f(X_t) - P_{T-\tau}f(X_0)$ stopped at $t = T-\tau$, it is enough to consider $t+\tau \leq T$. By the Markov property (2.38), we may write

$$P_{T-t_0-t}f(X_t) = \mathbb{E}_{X_t}f(X_{T-t_0-t}) = \mathbb{E}_x[f(X_{T-t_0})|\mathcal{F}_t], \quad \mathbb{P}_x\text{-a.s.} \tag{8.24}$$

for each $x \in E$. Let $0 \leq s \leq t \leq T - t_0$. By (8.17) and (8.24),

$$\begin{aligned} \mathbb{E}_{(t_0,x)}[\breve{M}_t|\breve{\mathcal{F}}_s] &= \mathbb{E}_x[P_{T-t_0-t}f(X_t)|\mathcal{F}_s] - P_{T-t_0}f(x) \\ &= \mathbb{E}_x[\mathbb{E}_x[f(X_{T-t_0})|\mathcal{F}_t]|\mathcal{F}_s] - P_{T-t_0}f(x) \\ &= \mathbb{E}_x[f(X_T)|\mathcal{F}_s] - P_{T-t_0}f(x), \quad \mathbb{P}_{(t_0,x)}\text{-a.s.} \end{aligned}$$

Again, by (8.24),

$$\mathbb{E}_x[f(X_T)|\mathcal{F}_s] - P_{T-t_0}f(x) = P_{T-t_0-t}f(X_s) - P_{T-t_0}f(x) = \breve{M}_s, \quad \mathbb{P}_{(t_0,x)}\text{-a.s.}$$

The proof is complete. □

8.2 Sharp bracket of martingale

In this section we consider martingales given by (8.22) and (8.23). The goal of this part is to obtain the following result.

Theorem 8.8. *Impose Assumption 8.4. Let* $f \in \mathcal{D}(\mathcal{E})$. *Let* $(\breve{M}_t)_{t\geq 0}$ *be the martingale given by* (8.23). *Then the sharp bracket* $\langle \breve{M}\rangle_t(\tau,\omega)$ *of this martingale is equivalent to*

$$\int_0^t 2\Gamma[P_{T-\tau-s}f](X_s(\omega))\, \mathbb{1}_{(0,T]}(\tau+s)\,\mathrm{d}s, \tag{8.25}$$

where Γ *is the carré du champ operator given by* (8.8).

Due to the fact that $\langle \breve{M}\rangle_t(0,\omega) = \langle M\rangle_t(\omega)$, the above result imply the following representation of $\langle M\rangle_t$.

Corollary 8.9. *Impose Assumption 8.4. Let $f \in \mathcal{D}(\mathcal{E})$. Let $(M_t)_{t\geq 0}$ be the martingale given by* (8.22). *Then the sharp bracket $\langle M\rangle_t(\omega)$ of this martingale is equivalent to*

$$\int_0^{t\wedge T} 2\Gamma[P_{T-s}f](X_s(\omega))\,\mathrm{d}s, \tag{8.26}$$

where Γ is the carré du champ operator given by (8.8).

We stress that even if we only need the representation of $\langle M\rangle_t$, the space-time version $\langle \breve{M}\rangle_t$ is necessary. The reason for this is the following. Our approach employs the notion of the Revuz correspondence. In view of the fact that $(M_t)_{t\geq 0}$ (and thus $(\langle M\rangle_t)_{t\geq 0}$) is not an additive functional, it is useful to move to the space-time, where the sharp bracket $(\langle \breve{M}\rangle_t)_{t\geq 0}$ is a PCAF.

Before we start the proof of Theorem 8.8, we require auxiliary facts about the calculus on $L^p(m)$.

Lemma 8.10. *Let $1 < p < \infty$. Let u be differentiable on $[0,+\infty)$ with values in $L^p(m)$ (in the sense of Section 2.3). Then $P_t u(t)$ is differentiable on $(0,+\infty)$ and*

$$(P_t u(t))' = A_p P_t u(t) + P_t u'(t), \quad t > 0.$$

In addition, when we assume that for each $t \geq 0$, $P_t u(t) \in \mathcal{D}(A_1)$, then $P_t u(t)$ is differentiable on $[0,+\infty)$ with values in $L^1(m)$ and

$$(P_t u(t))' = A_1 P_t u(t) + P_t u'(t), \quad t \geq 0. \tag{8.27}$$

Proof. Let $1 \leq p < \infty$. Recall that $\Delta_h u(t) := u(t+h) - u(t)$. Note that

$$\frac{1}{h}\Delta_h(P_t u(t)) = P_{t+h}u'(t) + P_{t+h}\left(\frac{1}{h}\Delta_h u(t) - u'(t)\right) + \frac{1}{h}(P_{t+h} - P_t)u(t).$$

The first term converges to $P_t u'(t)$ as $h \to 0$, by the strong continuity of $(P_t)_{t\geq 0}$. The second part tends to zero, since

$$\left\| P_{t+h}\left(\frac{1}{h}\Delta_h u(t) - u'(t)\right)\right\|_p \leq \left\|\frac{1}{h}\Delta_h u(t) - u'(t)\right\|_p \to 0, \quad \text{as } h \to 0.$$

Here, we used the contraction property of $(P_t)_{t\geq 0}$. The last part converges to $A_p P_t u(t)$ by the definition of the generator A_p. In the case of $p > 1$, we utilize the analyticity of $(P_t)_{t\geq 0}$. When $p = 1$, the convergence follows from the assumption $P_t u(t) \in \mathcal{D}(A_1)$. Summarizing,

$$\frac{1}{h}\Delta_h(P_t u(t)) \to P_t u'(t) + 0 + A_p P_t u(t)$$

in $L^p(m)$ as $h \to 0$. This completes the proof. □

Lemma 8.11. *Fix $T > 0$. Let $f \in L^2(m)$. Under Assumption 8.4, the mapping $[0,T) \ni t \mapsto P_t[(P_{T-t}f)^2]$ is C^1 with values in $L^1(m)$ and*

$$(P_t[(P_{T-t}f)^2])' = 2P_t(\Gamma[P_{T-t}f]). \tag{8.28}$$

Proof. We apply (8.27) from Lemma 8.10 with $u(t) := (P_{T-t}f)^2$ and obtain

$$(P_t[(P_{T-t}f)^2])' = A_1 P_t[(P_{T-t}f)^2] + P_t u'(t).$$

Let us explain that we are allowed to do that. For every $t \in [0,T)$, $P_{T-t}f \in \mathcal{D}(A_2)$ due to the fact that the semigroup $(P_t)_{t\geq 0}$ is analytic on $L^2(m)$. Under Assumption 8.4 implication (8.12) is true, hence $(P_{T-t}f)^2 \in \mathcal{D}(A_1)$ and also $P_t[(P_{T-t}f)^2] \in \mathcal{D}(A_1)$. By Corollary 2.2(i), $u'(t) = -2P_{T-t}f A_2 P_{T-t}f$.

Summarizing,

$$\begin{aligned}(P_t[(P_{T-t}f)^2])' &= A_1 P_t[(P_{T-t}f)^2] - 2P_t[P_{T-t}f A_2 P_{T-t}f] \\ &= P_t[A_1[(P_{T-t}f)^2] - 2P_{T-t}f A_2 P_{T-t}f] \\ &= 2P_t(\Gamma[P_{T-t}f]).\end{aligned}$$

In the last line we used (8.13).

Recall that $\Gamma\colon \mathcal{D}(\mathcal{E}) \to L^1(m)$ is continuous. Moreover, the semigroup $(P_t)_{t\geq 0}$ is continuous on $L^1(m)$ and $\mathcal{D}(\mathcal{E})$; see (2.15). Thus, the mapping $[0,T) \ni t \mapsto P_t(\Gamma[P_{T-t}f]) \in L^1(m)$ is continuous. □

Proof of Theorem 8.8. Note that the additive functional given by (8.25) is of the form (8.15) with $\breve{X}_s$ and $g(t,x) := 2\Gamma[P_{T-t}f](x)\,\mathbb{1}_{(0,T]}(t)$. Thus, its Revuz measure is

$$2\Gamma[P_{T-t}f](x)\,\mathbb{1}_{(0,T]}(t)m(\mathrm{d}x)\mathrm{d}t.$$

Let ν be the Revuz measure of $(\langle\breve{M}\rangle_t)_{t\geq 0}$. Let $\hat{f}$ be an arbitrary bounded α-coexcessive function with respect to the space-time process $(\breve{X}_t)_{t\geq 0}$. Applying (8.18), we get

$$\begin{aligned}\int\limits_{(0,+\infty)\times E} \hat{f}\,\mathrm{d}\nu &= \lim_{h\to 0^+}\frac{1}{h}\int\limits_0^{+\infty}\int\limits_E \hat{f}(t,x)\mathbb{E}_{(t,x)}\langle\breve{M}\rangle_h\, m(\mathrm{d}x)\mathrm{d}t \qquad (8.29)\\ &= \lim_{h\to 0^+}\frac{1}{h}\int\limits_0^{+\infty}\int\limits_E \hat{f}(t,x)\mathbb{E}_{(t,x)}\breve{M}_h^2\, m(\mathrm{d}x)\mathrm{d}t.\end{aligned}$$

When $h+t \leq T$, then $\mathbb{E}_x P_{T-t-h}f(X_h) = P_h P_{T-t-h}f(x) = P_{T-t}f(x)$, and therefore, by (8.17),

$$\begin{aligned}\mathbb{E}_{(t,x)}\breve{M}_h^2 &= \mathbb{E}_x\breve{M}_h^2(t,\cdot) = \mathbb{E}_x(P_{T-t-h}f(X_h))^2 - 2P_{T-t}f(x)\mathbb{E}_x P_{T-t-h}f(X_h) + (P_{T-t}f(x))^2 \\ &= P_h[(P_{T-t-h}f)^2](x) - (P_{T-t}f(x))^2. \qquad (8.30)\end{aligned}$$

Further, because of Lemma 8.11 we obtain

$$\begin{aligned}P_h[(P_{T-t-h}f)^2] - (P_{T-t}f)^2 &= \int\limits_0^h 2P_s(\Gamma[P_{T-t-s}f])\,\mathrm{d}s \qquad (8.31)\\ &= h\int\limits_0^1 2P_{hs}(\Gamma[P_{T-t-hs}f])\,\mathrm{d}s.\end{aligned}$$

Here, the integral on the right-hand side is the Bochner integral.

Similarly, when $t \leq T < h+t$,

$$\mathbb{E}_{(t,x)}\breve{M}_h^2 = P_{T-t}[f^2](x) - (P_{T-t}f(x))^2 \qquad (8.32)$$

and

$$P_{T-t}[f^2] - (P_{T-t}f)^2 = \int_0^{T-t} 2P_s(\Gamma[P_{T-t-s}f])\,\mathrm{d}s \tag{8.33}$$
$$= h\int_0^1 2P_{hs}(\Gamma[P_{T-t-hs}f])\,\mathbb{1}_{[0,T-t]}(hs)\,\mathrm{d}s.$$

Clearly, $\mathbb{E}_{(t,x)}\breve{M}_h^2 = 0$ for $T < t$.

Combining (8.29), (8.30), (8.31), (8.32), and (8.33), we obtain

$$\begin{aligned}\int_{(0,+\infty)\times E} \hat{f}\,\mathrm{d}\nu &= \lim_{h\to 0^+}\int_0^T\int_E \hat{f}(t,\cdot)\int_0^1 2P_{hs}(\Gamma[P_{T-t-hs}f])\times\\ &\quad\times\left[\mathbb{1}_{(0,T-h]}(t) + \mathbb{1}_{[0,T-h]}(hs)\,\mathbb{1}_{(T-h,T]}(t)\right]\,\mathrm{d}s\mathrm{d}m\mathrm{d}t\\ &= \lim_{h\to 0^+}\int_0^T\int_0^1\int_E 2\hat{f}(t,\cdot)P_{hs}(\Gamma[P_{T-t-hs}f])\times\\ &\quad\times\left[\mathbb{1}_{(0,T-h]}(t) + \mathbb{1}_{[0,T-h]}(hs)\,\mathbb{1}_{(T-h,T]}(t)\right]\,\mathrm{d}m\mathrm{d}s\mathrm{d}t.\end{aligned}$$

Here, we utilized Tonelli's theorem. We claim that

$$\begin{aligned}&\lim_{h\to 0^+}\int_0^T\int_0^1\int_E 2\hat{f}(t,\cdot)P_{hs}(\Gamma[P_{T-t-hs}f])\times\\ &\qquad\times\left[\mathbb{1}_{(0,T-h]}(t) + \mathbb{1}_{[0,T-h]}(hs)\,\mathbb{1}_{(T-h,T]}(t)\right]\,\mathrm{d}m\mathrm{d}s\mathrm{d}t\\ &\quad= \int_0^T\int_0^1 \lim_{h\to 0^+}\int_E 2\hat{f}(t,\cdot)P_{hs}(\Gamma[P_{T-t-hs}f])\,\mathrm{d}m\times\\ &\qquad\times\left[\mathbb{1}_{(0,T-h]}(t) + \mathbb{1}_{[0,T-h]}(hs)\,\mathbb{1}_{(T-h,T]}(t)\right]\,\mathrm{d}s\mathrm{d}t.\end{aligned}$$

For all enough small $h \geq 0$, by the boundedness of $\hat{f}$,

$$\begin{aligned}\int_E 2\hat{f}(t,\cdot)P_{hs}(\Gamma[P_{T-t-hs}f])\,\mathrm{d}m &\leq 2\int_E \Gamma[P_{T-t-hs}f]\,\mathrm{d}m \sup_{(r,x)\in(0,T]\times E}\left|\hat{f}(r,x)\right|\\ &= 2\mathcal{E}[P_{T-t-hs}f]\sup_{(r,x)\in(0,T]\times E}\left|\hat{f}(r,x)\right|\\ &\leq 2\mathcal{E}[f]\sup_{(r,x)\in(0,T]\times E}\left|\hat{f}(r,x)\right| < +\infty.\end{aligned}$$

Here, we employed the fact that $f \in \mathcal{D}(\mathcal{E})$. Our claim follows from the dominated convergence theorem.

Due to Lemma 8.11, the mapping $[0, T-t) \ni r \mapsto P_r(\Gamma[P_{T-t-r}f]) \in L^1(m)$ is continuous. Hence, $P_{hs}(\Gamma[P_{T-t-hs}f]) \to \Gamma[P_{T-t}f]$ in $L^1(m)$ as $h \to 0^+$. Thus,

$$\lim_{h\to 0^+}\int_E 2\hat{f}(t,\cdot)P_{hs}(\Gamma[P_{T-t-hs}f])\,\mathrm{d}m = \int_E 2\hat{f}(t,\cdot)\Gamma[P_{T-t}f]\,\mathrm{d}m.$$

Clearly,

$$\mathbb{1}_{(0,T-h]}(t) + \mathbb{1}_{[0,T-h]}(hs)\,\mathbb{1}_{(T-h,T]}(t) \to \mathbb{1}_{(0,T]}(t),$$

for almost every $t \in (0,T]$ as $h \to 0^+$.

On account of the above observations, we may write

$$\begin{aligned}\int\limits_{(0,+\infty)\times E} \hat{f}\,\mathrm{d}\nu &= \int\limits_0^T \int\limits_0^1 \int\limits_E 2\hat{f}(t,\cdot)\Gamma[P_{T-t}f]\,\mathrm{d}m\mathrm{d}s\mathrm{d}t \\ &= \int\limits_0^{+\infty} \int\limits_E \hat{f}(t,x)\cdot 2\Gamma[P_{T-t}f](x)\,\mathbb{1}_{(0,T]}(t)\,m(\mathrm{d}x)\mathrm{d}t.\end{aligned}$$

The above identity may be generalized to any non-negative $\hat{f} \in C_c((0,+\infty)\times E)$. Indeed, $\hat{f}$ may be approximated by the bounded n-coexcessive functions $\hat{f}_n := n\hat{R}_n\hat{f}$, where $\hat{R}_n$ is the coresolvent of the space-time process $(\check{X}_t)_{t\geq 0}$. In such a case, there exists a subsequence $(\hat{f}_{n_k})_{k\in\mathbb{N}}$ such that $\hat{f}_{n_k} \to \hat{f}$ almost everywhere as $k \to +\infty$.

Finally, we have the equality of the Revuz measures

$$\nu(\mathrm{d}t,\mathrm{d}x) = 2\Gamma[P_{T-t}f](x)\,\mathbb{1}_{(0,T]}(t)m(\mathrm{d}x)\mathrm{d}t.$$

The statement of the theorem follows from the uniqueness of the PCAF. □

8.3 Applications of Hardy–Stein identity

The Hardy–Stein identity finds application in the estimation of the square functions. We apply this identity, given in Corollary 7.4, to the setting under Assumption 8.4 and the strong stability of the semigroup $(P_t)_{t\geq 0}$ from Assumption 3.1. In this configuration, the Hardy–Stein identity reads

$$\int\limits_E |f(x)|^p\,m(\mathrm{d}x) = \int\limits_0^{+\infty} \iint\limits_{E\times E\setminus\mathrm{diag}} F_p(P_tf(x),P_tf(y))\,J(\mathrm{d}x,\mathrm{d}y)\mathrm{d}t, \tag{8.34}$$

for all $1<p<\infty$ and $f \in L^p(m)$. Yet,

$$\int\limits_E |f(x)|^p\,m(\mathrm{d}x) \geq \int\limits_0^{+\infty} \iint\limits_{E\times E\setminus\mathrm{diag}} F_p(P_tf(x),P_tf(y))\,J(\mathrm{d}x,\mathrm{d}y)\mathrm{d}t, \tag{8.35}$$

remains true even in the absence of Assumption 3.1. Due to Lemma 2.4, for all $1<p<\infty$,

$$F_p(a,b) \asymp |b-a|^2\,(|a|\vee|b|)^{p-2} \asymp |b-a|^2\,(|a|+|b|)^{p-2}, \quad a,b\in\mathbb{R}. \tag{8.36}$$

If $p \geq 2$, then moreover

$$F_p(a,b) \asymp |b-a|^2\,(|a|\vee|b|)^{p-2} \asymp |b-a|^2\,(|a|^{p-2}+|b|^{p-2}), \quad a,b\in\mathbb{R}. \tag{8.37}$$

Combining the Hardy–Stein identity from equality (8.34) (or (8.35)) with estimates (8.36) and (8.37) we get some auxiliary estimates of the p-norm of $f \in L^p(m)$. They will be useful later to obtain the estimation of the norms of the square functions.

Indeed, by (8.34), (8.36), and the symmetry of the jumping measure J:

$$\begin{aligned}\int_E |f(x)|^p\, m(\mathrm{d}x) &\asymp \int_0^{+\infty} \iint_{E\times E\setminus\mathrm{diag}} (P_tf(y)-P_tf(x))^2(|P_tf(x)|\vee|P_tf(y)|)^{p-2}\, J(\mathrm{d}x,\mathrm{d}y)\mathrm{d}t\\ &= \int_0^{+\infty} \iint_{E\times E\setminus\mathrm{diag}} (P_tf(y)-P_tf(x))^2|P_tf(x)|^{p-2}\chi(P_tf(x),P_tf(y))\, J(\mathrm{d}x,\mathrm{d}y)\mathrm{d}t\\ &= \int_0^{+\infty}\int_E \widetilde{\Gamma}[P_tf](x)|P_tf(x)|^{p-2}\, m(\mathrm{d}x)\mathrm{d}t.\end{aligned}$$

Recall that $\widetilde{\Gamma}[\cdot]$ and χ are defined by (8.9) and (8.10). If $2\le p<\infty$, then by (8.34), (8.37), and the symmetry of J:

$$\begin{aligned}\int_E |f(x)|^p\, m(\mathrm{d}x) &\asymp \int_0^{+\infty} \iint_{E\times E\setminus\mathrm{diag}} (P_tf(y)-P_tf(x))^2(|P_tf(x)|^{p-2}+|P_tf(y)|^{p-2})\, J(\mathrm{d}x,\mathrm{d}y)\mathrm{d}t\\ &= 2\int_0^{+\infty} \iint_{E\times E\setminus\mathrm{diag}} (P_tf(y)-P_tf(x))^2|P_tf(x)|^{p-2}\, J(\mathrm{d}x,\mathrm{d}y)\mathrm{d}t\\ &= \int_0^{+\infty}\int_E \Gamma[P_tf](x)|P_tf(x)|^{p-2}\, m(\mathrm{d}x)\mathrm{d}t.\end{aligned}$$

For the definition of $\Gamma[\cdot]$, we refer to (8.8). When $1<p\le 2$, we only have a one-sided ($\lesssim$) bound in the above estimate.

Summing up, if $2\le p<\infty$, then

$$\int_E |f(x)|^p\, m(\mathrm{d}x) \asymp \int_0^{+\infty}\int_E \widetilde{\Gamma}[P_tf](x)|P_tf(x)|^{p-2}\, m(\mathrm{d}x)\mathrm{d}t \tag{8.38}$$

$$\asymp \int_0^{+\infty}\int_E \Gamma[P_tf](x)|P_tf(x)|^{p-2}\, m(\mathrm{d}x)\mathrm{d}t, \tag{8.39}$$

while if $1<p\le 2$, then:

$$\begin{aligned}\int_E |f(x)|^p\, m(\mathrm{d}x) &\asymp \int_0^{+\infty}\int_E \widetilde{\Gamma}[P_tf](x)|P_tf(x)|^{p-2}\, m(\mathrm{d}x)\mathrm{d}t\\ &\le \int_0^{+\infty}\int_E \Gamma[P_tf](x)|P_tf(x)|^{p-2}\, m(\mathrm{d}x)\mathrm{d}t.\end{aligned}$$

In the absence of Assumption 3.1, the same estimates holds when we substitute $\int_E |f|^p\,\mathrm{d}m - \lim_{T\to+\infty}\|P_Tf\|_p^p$ in place of $\int_E |f|^p\,\mathrm{d}m$. In such a case, if $2\le p<\infty$, then

$$\begin{aligned}\int_E |f(x)|^p\, m(\mathrm{d}x) &\gtrsim \int_0^{+\infty}\int_E \widetilde{\Gamma}[P_tf](x)|P_tf(x)|^{p-2}\, m(\mathrm{d}x)\mathrm{d}t\\ &\asymp \int_0^{+\infty}\int_E \Gamma[P_tf](x)|P_tf(x)|^{p-2}\, m(\mathrm{d}x)\mathrm{d}t,\end{aligned}$$

while if $1 < p \le 2$, then

$$\int_E |f(x)|^p \, m(\mathrm{d}x) \gtrsim \int_0^{+\infty} \int_E \widetilde{\Gamma}[P_t f](x) |P_t f(x)|^{p-2} \, m(\mathrm{d}x) \mathrm{d}t. \tag{8.40}$$

8.4 Square function G

The first considered square function,

$$G(x) = \left(\int_0^{+\infty} \Gamma[P_t f](x) \, \mathrm{d}t \right)^{1/2},$$

does not posses the desired boundedness of p-norm for all $1 < p < \infty$. It was shown in [5], that in the following example G is too large for certain $f \in L^p(m)$, $1 < p < 2$.

Example 8.12. The following example was provided in Example 2 in [5]. Consider Dirichlet form on the Euclidean space $E = \mathbb{R}^d$ ($d \ge 2$) associated with the Cauchy process (α-stable process with $\alpha = 1$), that is, $J(x, \mathrm{d}y) = J(x, y)\mathrm{d}y$, where

$$J(x, y) = \frac{\mathcal{A}_{d,1}}{|y - x|^{d+1}}$$

and $\mathcal{A}_{d,1}$ is some constant. The semigroup is given by the following transition density:

$$p_t(x, y) = c_d \frac{t}{(t^2 + |y - x|^2)^{(d+1)/2}}$$

and $c_d > 0$ is some constant.

Let $h(x) := |x|^{-(d+1)/2}$ and $f(x) := h(x) \, \mathbb{1}_{B(0,1)}(x)$. Here, $B(x, r)$ is the ball of radius r centered at a point x. Observe that f belongs to $L^p(m)$ for $1 < p < 2d/(d+1) < 2$. We claim that for the function f the square function G is identically equals to infinity. The essence of the problem lies in the fact that $\Gamma[P_t f](x)$ grows too fast when t approaches zero. To make the example clear and to prepare some notation for Example 8.20 below, we present the full computation.

Introduce the following function:

$$v(x, t) := \begin{cases} P_t h(x), & \text{for } t > 0, \\ h(x), & \text{for } t = 0. \end{cases}$$

Observe that, since h is locally integrable on $\mathbb{R}^d$ and vanishes at infinity, the function v is well-defined and continuous except at the point $(x, t) = (0, 0)$. Denote also

$$v_s(x, t) := \int_{B(0,1/s)} h(y) p_t(x, y) \, \mathrm{d}y.$$

Calculate $P_t f(x)$. Observe firstly that $p_t(x, y)$ satisfy the following scaling property

$$p_t(x, y) = \frac{1}{t^d} p_1 \left(\tfrac{x}{t}, \tfrac{y}{t} \right).$$

Also $h(x) = t^{-(d+1)/2} h(x/t)$. Therefore,

$$\begin{aligned} P_t f(x) &= c_d \int\limits_{B(0,1)} \frac{1}{t^{(d+1)/2}} h\left(\tfrac{y}{t}\right) \cdot \frac{1}{t^d} p_1\left(\tfrac{x}{t}, \tfrac{y}{t}\right) \mathrm{d}y \\ &= c_d \int\limits_{B(0,1/t)} \frac{1}{t^{(d+1)/2}} h\left(y\right) \cdot \frac{1}{t^d} p_1\left(\tfrac{x}{t}, y\right) t^d \,\mathrm{d}y \\ &= t^{-(d+1)/2} v_t\left(\tfrac{x}{t}, 1\right). \end{aligned}$$

Similarly,

$$v(x,t) = t^{-(d+1)/2} v\left(\tfrac{x}{t}, 1\right), \quad t > 0,\ x \in \mathbb{R}^d.$$

We have

$$\begin{aligned} (G(x))^2 &= \mathcal{A}_{d,1} \int\limits_0^\infty \int\limits_{\mathbb{R}^d} \frac{(P_t f(y) - P_t f(x))^2}{|x-y|^{d+1}} \mathrm{d}y\mathrm{d}t \\ &= \mathcal{A}_{d,1} \int\limits_0^\infty \int\limits_{\mathbb{R}^d} \frac{1}{t^{d+1}} \cdot \frac{(v_t(\frac{y}{t},1) - v_t(\frac{x}{t},1))^2}{|x-y|^{d+1}} \mathrm{d}y\mathrm{d}t \\ &= \mathcal{A}_{d,1} \int\limits_{\mathbb{R}^d} \int\limits_0^\infty \frac{1}{t} \cdot \frac{(v_t(y,1) - v_t(\frac{x}{t},1))^2}{|x-ty|^{d+1}} \mathrm{d}t\mathrm{d}y. \end{aligned} \tag{8.41}$$

In the last line we used Tonelli's theorem. Note that $v_t(y,1) \to v(y,1) > 0$ as $t \to 0^+$ for all $y \in \mathbb{R}^d$. Moreover, when $x \neq 0$,

$$v_t\left(\tfrac{x}{t}, 1\right) \leq v\left(\tfrac{x}{t}, 1\right) = t^{(d+1)/2} v(x,t) \to 0 \cdot h(x) = 0, \quad \text{as } t \to 0^+.$$

Summarizing, when $x \neq 0$, then the integrand of (8.41)

$$\frac{1}{t} \cdot \frac{(v_t(y,1) - v_t(\frac{x}{t},1))^2}{|x-ty|^{d+1}} \sim \frac{(v(y,1))^2}{|x|^{d+1}} \cdot \frac{1}{t}, \quad \text{as } t \to 0^+,$$

and when $x = 0$,

$$\frac{1}{t} \cdot \frac{(v_t(y,1) - v_t(\frac{x}{t},1))^2}{|x-ty|^{d+1}} \sim \frac{(v(y,1) - v(0,1))^2}{|y|^{d+1}} \cdot \frac{1}{t^{d+2}}, \quad \text{as } t \to 0^+.$$

We conclude that the inner integral of (8.41) is equal to infinity, hence $G \equiv +\infty$.

Nevertheless, the authors shown in [5], that $\|G\|_p \asymp \|f\|_p$ when $2 < p < \infty$ for the semigroups $(P_t)_{t \geq 0}$ associated with the symmetric Lévy processes in the Euclidean space. The results were generalized to the non-symmetric case in [6] by Bañuelos and Kim.

For our more general class of processes, when $2 < p < \infty$, the lower bound of $\|G\|_p$ is actually valid. As we will see in the next section, the Hardy–Stein identity implies the following estimate.

Proposition 8.13. *Impose Assumptions 3.1 and 8.4. Let $2 < p < \infty$ and $f \in L^p(m)$. Then,*

$$\|f\|_p \lesssim \|G\|_p.$$

The above results follows immediately from $\widetilde{G} \leq \sqrt{2} G$ and the estimate of $\widetilde{G}$, which is presented in Theorem 8.15 below.

Unfortunately, the upper bound of $\|G\|_p$ does not stay valid in a general case. We will see this case in Example 8.16 in the next section.

8.5 Square function $\widetilde{G}$

In this section we consider the following square function:

$$\widetilde{G}(x) := \left(\int_0^{+\infty} \widetilde{\Gamma}[P_t f](x) \, \mathrm{d}t \right)^{1/2}.$$

The square function $\widetilde{G}$ resolves the issue of unboundedness presented in Example 8.12 that addresses the case $1 < p < 2$. Indeed, utilizing the Hardy–Stein identity, we derive the following two estimates.

Theorem 8.14. *Impose Assumption 8.4. Let* $1 < p \leq 2$ *and* $f \in L^p(m)$*. Then,*

$$\left\| \widetilde{G} \right\|_p \lesssim \|f\|_p .$$

Proof. Let $f^*(x) := \sup_{t \geq 0} |P_t f(x)|$. Observe that for $p \leq 2$

$$\begin{aligned}
(\widetilde{G}(x))^2 &= \int_0^{+\infty} \widetilde{\Gamma}[P_t f](x) \mathrm{d}t \\
&= (f^*(x))^{2-p} \int_0^{+\infty} \widetilde{\Gamma}[P_t f](x) |f^*(x)|^{p-2} \, \mathrm{d}t \\
&\leq (f^*(x))^{2-p} \int_0^{+\infty} \widetilde{\Gamma}[P_t f](x) |P_t f(x)|^{p-2} \, \mathrm{d}t.
\end{aligned}$$

Due to Hölder's inequality, inequality (8.40), and Stein's maximal theorem (2.3),

$$\begin{aligned}
\int_E (\widetilde{G}(x))^p \, m(\mathrm{d}x) &\leq \int_E (f^*(x))^{(2-p)p/2} \left(\int_0^{+\infty} \widetilde{\Gamma}[P_t f](x) |P_t f(x)|^{p-2} \, \mathrm{d}t \right)^{p/2} m(\mathrm{d}x) \\
&\leq \left(\int_E (f^*(x))^p \, m(\mathrm{d}x) \right)^{1-p/2} \left(\int_0^{+\infty} \int_E \widetilde{\Gamma}[P_t f](x) |P_t f(x)|^{p-2} \, m(\mathrm{d}x) \mathrm{d}t \right)^{p/2} \\
&\asymp \left(\int_E (f^*(x))^p \, m(\mathrm{d}x) \right)^{1-p/2} \left(\int_E |f(x)|^p \, m(\mathrm{d}x) \right)^{p/2} \\
&\asymp \int_E |f(x)|^p \, m(\mathrm{d}x).
\end{aligned}$$

Therefore,

$$\int_E (\widetilde{G}(x))^p \, m(\mathrm{d}x) \lesssim \int_E |f(x)|^p \, m(\mathrm{d}x).$$

□

Theorem 8.15. *Impose Assumptions 3.1 and 8.4. Let* $2 \leq p < \infty$ *and* $f \in L^p(m)$*. Then,*

$$\|f\|_p \lesssim \left\| \widetilde{G} \right\|_p .$$

Proof. Let $f^*(x) := \sup_{t\geq 0} |P_t f(x)|$. By (8.38), Hölder's inequality, and Stein's maximal theorem (2.3),

$$\begin{aligned}
\int_E |f(x)|^p \, m(\mathrm{d}x) &\asymp \int_0^{+\infty} \int_E \widetilde{\Gamma}[P_t f](x) |P_t f(x)|^{p-2} \, m(\mathrm{d}x)\mathrm{d}t \\
&\leq \int_0^{+\infty} \int_E \widetilde{\Gamma}[P_t f](x) |f^*(x)|^{p-2} \, m(\mathrm{d}x)\mathrm{d}t \\
&= \int_E (\widetilde{G}(x))^2 |f^*(x)|^{p-2} \, m(\mathrm{d}x) \\
&\leq \left(\int_E (\widetilde{G}(x))^p \, m(\mathrm{d}x) \right)^{2/p} \left(\int_E |f^*(x)|^p \, m(\mathrm{d}x) \right)^{1-2/p} \\
&\asymp \left(\int_E (\widetilde{G}(x))^p \, m(\mathrm{d}x) \right)^{2/p} \left(\int_E |f(x)|^p \, m(\mathrm{d}x) \right)^{1-2/p}.
\end{aligned}$$

This yields

$$\int_E |f(x)|^p \, m(\mathrm{d}x) \lesssim \int_E (\widetilde{G}(x))^p \, m(\mathrm{d}x).$$

□

8.5.1 Brownian motion on interval with removed segment

Example 8.16. The present example was proposed and constructed by the supervisor of the author, Mateusz Kwaśnicki.

Idea. We construct an approximation of the reflected Brownian motion on the interval $[0, \frac{1}{4}] \cup [\frac{3}{4}, 1]$ with additional jumps occurring between $\frac{1}{4}$ and $\frac{3}{4}$. The intensity of these jumps are chosen so that the function $\cos x$ is an eigenfunction of the generator.

This process is analogous to the reflected Brownian motion $(B_t)_{t\geq 0}$ in $[0, 1]$, but with segments of paths lying on the part $(\frac{1}{4}, \frac{3}{4})$ removed.

Below we consider a discrete version of the above process: a Markov chain with continuous time which can be viewed as a reflected symmetric nearest-neighbor random walk $(X_t)_{t\geq 0}$ on $\{1, 2, \ldots, 4n\}$, with parts of paths corresponding to $\{t : X_t \in \{n+1, n+2, \ldots, 3n\}\}$ deleted.

To use the standard notation for the generator matrix of a Markov chain, we reindex the state space: we shift the states $3n+1, 3n+2, \ldots, 4n$ to the left and call them $n+1, n+2, \ldots, 2n$. Therefore $E = \{1, 2, \ldots, 2n\}$. We equip E with the counting measure m.

Generator. The generator of the process $(X_t)_{t\geq 0}$ is given by the following $2n \times 2n$ matrix $A = [a_{i,j}]_{i,j=1}^{2n}$:

$$A = \begin{pmatrix}
-1 & 1 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\
1 & -2 & 1 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\
0 & 1 & -2 & 1 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & -2 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & 0 & \cdots & -2 & 1 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \cdots & 1 & \alpha_n^* & \alpha_n & 0 & \cdots & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \cdots & 0 & \alpha_n & \alpha_n^* & 1 & \cdots & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \cdots & 0 & 0 & 1 & -2 & \cdots & 0 & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & -2 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 1 & -2 & 1 & 0 \\
0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 1 & -2 & 1 \\
0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 1 & -1
\end{pmatrix},$$

where

$$\alpha_n := \frac{1}{1 + \cot \frac{\pi}{8n}}$$

and $\alpha_n^* := -1 - \alpha_n$. In other words, $J(x, \mathrm{d}y) = J(x,y)m(\mathrm{d}y)$, $J(i,j) := a_{i,j}$, and

$$a_{i,j} = \begin{cases} 1 & \text{when } |i-j| = 1 \text{ and } \{i,j\} \neq \{n, n+1\}, \\ \alpha_n & \text{when } i = n,\ j = n+1 \text{ or } i = n+1,\ j = n, \\ 0 & \text{when } |i-j| > 1. \end{cases}$$

Eigenfunction. Consider the following function on E:

$$f(k) := \begin{cases} \cos \frac{(2k-1)\pi}{8n} & \text{if } k \leq n, \\ \cos \frac{(2k-1+4n)\pi}{8n} & \text{if } k > n. \end{cases}$$

In this part we will show that f is an eigenfunction of A, namely

$$Af(k) = -\lambda_n f(k),$$

where

$$\lambda_n := 2(1 - \cos \tfrac{\pi}{4n}) = 4(\sin \tfrac{\pi}{8n})^2.$$

Note that $f(2n+1-k) = -f(k)$ and $Af(2n+1-k) = -Af(k)$, hence it suffices to consider $k \leq n$. By the sum-to-product formula, for $k = 2, 3, \ldots, n-1$, we obtain

$$\begin{aligned} Af(k) &= f(k+1) + f(k-1) - 2f(k) \\ &= \cos \tfrac{(2k+1)\pi}{8n} + \cos \tfrac{(2k-3)\pi}{8n} - 2\cos \tfrac{(2k-1)\pi}{8n} \\ &= 2\cos \tfrac{(2k-1)\pi}{8n} \cos \tfrac{\pi}{4n} - 2\cos \tfrac{(2k-1)\pi}{8n} = -\lambda_n f(k). \end{aligned}$$

Similarly, when $k = 1$,

$$\begin{aligned}Af(1) &= f(2) - f(1)\\ &= \cos\tfrac{3\pi}{8n} - \cos\tfrac{\pi}{8n}\\ &= \cos\tfrac{3\pi}{8n} + \cos\tfrac{-\pi}{8n} - 2\cos\tfrac{\pi}{8n}\\ &= 2\cos\tfrac{\pi}{8n}\cos\tfrac{\pi}{4n} - 2\cos\tfrac{\pi}{8n}.\end{aligned}$$

Thus, again $Af(1) = -\lambda_n f(1)$. Finally, for $k = n$ we have

$$\begin{aligned}Af(n) &= \alpha_n(f(n+1) - f(n)) + (f(n-1) - f(n))\\ &= \alpha_n(\cos\tfrac{(6n+1)\pi}{8n} - \cos\tfrac{(2n-1)\pi}{8n}) + \cos\tfrac{(2n-3)\pi}{8n} - \cos\tfrac{(2n-1)\pi}{8n}\\ &= -2\alpha_n\cos\tfrac{(2n-1)\pi}{8n} + 2\sin\tfrac{(n-1)\pi}{4n}\sin\tfrac{\pi}{8n}\\ &= 2\left(-\alpha_n + \frac{\sin\frac{(n-1)\pi}{4n}\sin\frac{\pi}{8n}}{\cos\frac{(2n-1)\pi}{8n}}\right)\cos\tfrac{(2n-1)\pi}{8n}.\end{aligned}$$

A parameter α_n is chosen so that the right-hand side is equal to $-\lambda_n f(n)$. Indeed,

$$\begin{aligned}&\left(-\alpha_n + \frac{\sin\frac{(n-1)\pi}{4n}\sin\frac{\pi}{8n}}{\cos\frac{(2n-1)\pi}{8n}}\right)\\ &= \frac{-\sin\frac{\pi}{8n}}{\sin\frac{\pi}{8n} + \cos\frac{\pi}{8n}} + \frac{\frac{\sqrt{2}}{2}(\cos\frac{\pi}{4n} - \sin\frac{\pi}{4n})\sin\frac{\pi}{8n}}{\frac{\sqrt{2}}{2}(\cos\frac{\pi}{8n} + \sin\frac{\pi}{8n})}\\ &= -\frac{(1 - \cos\frac{\pi}{4n} + \sin\frac{\pi}{4n})\sin\frac{\pi}{8n}}{\cos\frac{\pi}{8n} + \sin\frac{\pi}{8n}}\\ &= -\frac{(2(\sin\frac{\pi}{8n})^2 + 2\sin\frac{\pi}{8n}\cos\frac{\pi}{8n})\sin\frac{\pi}{8n}}{\cos\frac{\pi}{8n} + \sin\frac{\pi}{8n}}\\ &= -2(\sin\tfrac{\pi}{8n})^2 = -\frac{\lambda_n}{2}.\end{aligned}$$

Square functions. Due to the previous part the semigroup operator P_t acting on f is given by $P_t f(k) = e^{-\lambda_n t} f(k)$. Recall that the square function $\widetilde{G}$ acting on f is given by

$$(\widetilde{G}(x))^2 = \int_0^\infty \int_{E\setminus\{x\}} (P_t f(y) - P_t f(x))^2 \chi(P_t f(x), P_t f(y))\, J(x,y) m(\mathrm{d}y)\mathrm{d}t.$$

Since $\int_0^{+\infty} e^{-\lambda_n t}\,\mathrm{d}t = 1/\lambda_n$ and m is the counting measure,

$$(\widetilde{G}(k))^2 = \frac{1}{\lambda_n}\sum_{l\neq k}(f(l) - f(k))^2\chi(f(k), f(l))J(k,l).$$

Additionally, $\widetilde{G}(2n+1-k) = \widetilde{G}(k)$. When $k = 1, 2, \ldots, n-1$, we have $f(k) > f(k+1) > 0$ and

$$(\widetilde{G}(k))^2 = \frac{1}{\lambda_n}(f(k+1) - f(k))^2 = \frac{4(\sin\frac{k\pi}{4n}\sin\frac{\pi}{8n})^2}{4(\sin\frac{\pi}{8n})^2} = (\sin\tfrac{k\pi}{4n})^2.$$

On the other hand, if $k = n$, then $|f(n)| = |f(n+1)| = -f(n)$, hence

$$\begin{aligned}(\widetilde{G}(n))^2 &= \frac{\alpha_n}{2\lambda_n}(f(n+1) - f(n))^2 \\ &= \frac{(2\cos\frac{(2n-1)\pi}{8n})^2}{8(\sin\frac{\pi}{8n})^2(1+\cot\frac{\pi}{8n})} \\ &= \frac{(\frac{\sqrt{2}}{2}(\cos\frac{\pi}{8n} + \sin\frac{\pi}{8n}))^2}{2\sin\frac{\pi}{8n}(\sin\frac{\pi}{8n} + \cos\frac{\pi}{8n})} \\ &= \frac{\sin\frac{\pi}{8n} + \cos\frac{\pi}{8n}}{4\sin\frac{\pi}{8n}} = \frac{1+\cot\frac{\pi}{8n}}{4}.\end{aligned}$$

Now, the p-norm of $\widetilde{G}$ is given by

$$\left\|\widetilde{G}\right\|_p^p = 2\sum_{k=1}^{n-1}(\sin\tfrac{k\pi}{4n})^p + 2(\widetilde{G}(n))^p. \tag{8.42}$$

The first term may be written as

$$2\sum_{k=1}^{n-1}(\sin\tfrac{k\pi}{4n})^p = 2n\sum_{k=1}^{n}\frac{1}{n}(\sin\tfrac{k\pi}{4n})^p - 2\left(\frac{\sqrt{2}}{2}\right)^p,$$

where the sum on the right-hand side is the Riemann sum of the function $2(\sin\frac{\pi x}{4})^p$ over $[0,1]$. Therefore, the first term on the right-hand side of (8.42) is asymptotically equal to $c_1 n$, where $c_1 := 2\int_0^1(\sin\frac{\pi x}{4})^p\,\mathrm{d}x > 0$.

The second term in (8.42) is asymptotically equal to $2(\frac{1}{4}\cdot\frac{8n}{\pi})^{p/2} = 2(\frac{2n}{\pi})^{p/2}$. Thus, for $p > 2$,

$$\left\|\widetilde{G}\right\|_p^p \sim \frac{2^{1+p/2}}{\pi^{p/2}}\, n^{p/2}.$$

In particular, we see that $\widetilde{G}(n)$ is the dominating term in the p-norm of $\widetilde{G}$. On the other hand,

$$\|f\|_p^p = 2\sum_{k=1}^{n}(\cos\tfrac{k\pi}{4n})^p \sim c_2 n,$$

where $c_2 := 2\int_0^1(\cos\frac{\pi x}{4})^p\,\mathrm{d}x > 0$. Finally,

$$\frac{\left\|\widetilde{G}\right\|_p}{\|f\|_p} \sim \frac{2^{1/p+1/2}}{\pi^{1/2}c_2^{1/p}}\, n^{1/2-1/p}.$$

This implies that there is no universal constant $C_p > 0$ such that

$$\left\|\widetilde{G}\right\|_p \le C_p\,\|f\|_p\,.$$

Additionally, since $\widetilde{G} \le \sqrt{2}G$, the above statement holds also for the square function G.

We shall discuss the above result. The present example demonstrates that the statement of Theorem 3.4 in [63] is not valid. The problem lies in the proof of Lemma 3.3, where it was unjustifiably assumed that $P_t[P_t f(k(\cdot,y))](x)$ is equal to $P_{2t}f(k(x,y))$ for some function k. The above calculations confirm, that this error is irreparable.

8.6 Square function H

The next square function we will consider is the following:

$$H(x) := \left(\int_0^{+\infty} P_t \Gamma[P_t f](x)\, \mathrm{d}t \right)^{1/2}.$$

The square function H overcomes problems demonstrated for $2 < p < +\infty$ in Example 8.16.

Indeed, employing the martingale $(M_t)_{t \geq 0}$ introduced in Subsection 8.1.5 and the Burkholder–Davies–Gundy inequality, we obtain the following estimate.

Theorem 8.17. *Impose Assumption 8.4. Let $2 \leq p \leq \infty$ and $f \in L^p(m)$. Then,*

$$\|H\|_p \lesssim \|f\|_p .$$

Proof. Assume first $f \in \mathcal{D}(\mathcal{E})$. Since $p \geq 2$, by Jensen's inequality, we get

$$\begin{aligned}
\int_E (H(x))^p \, m(\mathrm{d}x) &= \int_E \left(\int_0^{+\infty} P_t \Gamma[P_t f](x)\, \mathrm{d}t \right)^{p/2} m(\mathrm{d}x) \\
&= \int_E \left(\mathbb{E}_x \int_0^{+\infty} \Gamma[P_t f](X_t)\, \mathrm{d}t \right)^{p/2} m(\mathrm{d}x) \\
&\leq \int_E \mathbb{E}_x \left(\int_0^{+\infty} \Gamma[P_t f](X_t)\, \mathrm{d}t \right)^{p/2} m(\mathrm{d}x).
\end{aligned}$$

Utilizing the monotone convergence theorem, we may write

$$\int_E (H(x))^p \, m(\mathrm{d}x) \leq \lim_{T \to +\infty} \int_E \mathbb{E}_x \left(\int_0^{T} \Gamma[P_t f](X_t) \mathrm{d}t \right)^{p/2} m(\mathrm{d}x).$$

By the self-duality of the process $(X_t)_{t \geq 0}$ (formula (2.41)),

$$\int_E (H(x))^p \, m(\mathrm{d}x) \leq \lim_{T \to +\infty} \int_E \mathbb{E}_x \left(\int_0^{T} \Gamma[P_{T-t} f](X_t) \mathrm{d}t \right)^{p/2} m(\mathrm{d}x).$$

Since $f \in \mathcal{D}(\mathcal{E})$, in view of Corollary 8.9,

$$\int_E \mathbb{E}_x \left(\int_0^{T} \Gamma[P_{T-t} f](X_t) \mathrm{d}t \right)^{p/2} m(\mathrm{d}x) = 2^{-p/2} \int_E \mathbb{E}_x (\langle M \rangle_T)^{p/2} \, m(\mathrm{d}x),$$

where $\langle M \rangle_T$ is the angle bracket of the martingale $(M_t)_{t \geq 0}$ given by (8.22). Thus, by the Burkholder–Davies–Gundy inequality (8.19) and inequality (8.20) we get

$$\begin{aligned}
\int_E (H(x))^p \, m(\mathrm{d}x) &\lesssim \lim_{T \to +\infty} \int_E \mathbb{E}_x (\langle M \rangle_T)^{p/2} \, m(\mathrm{d}x) \\
&\lesssim \lim_{T \to +\infty} \int_E \mathbb{E}_x ([M]_T)^{p/2} \, m(\mathrm{d}x) \\
&\lesssim \lim_{T \to +\infty} \int_E \mathbb{E}_x |M_T - M_0|^p \, m(\mathrm{d}x) \\
&= \lim_{T \to +\infty} \int_E \mathbb{E}_x |f(X_T) - P_T f(x)|^p \, m(\mathrm{d}x).
\end{aligned}$$

Employing $|a+b|^p \le 2^{p-1}(|a|^p + |b|^p)$, we obtain

$$\begin{aligned}\int_E (H(x))^p \, m(\mathrm{d}x) &\lesssim \lim_{T\to+\infty} \int_E (\mathbb{E}_x|f(X_T)|^p + |P_T f(x)|^p) \, m(\mathrm{d}x) \\ &= \lim_{T\to+\infty} \int_E (P_T[|f|^p](x) + |P_T f(x)|^p) \, m(\mathrm{d}x) \\ &= \lim_{T\to+\infty} \int_E (|f(x)|^p + |P_T f(x)|^p) \, m(\mathrm{d}x).\end{aligned}$$

By the contraction property of $(P_t)_{t\ge 0}$, we have $\|P_T f\|_p^p \le \|f\|_p^p$. Hence,

$$\int_E (H(x))^p \, m(\mathrm{d}x) \lesssim \int_E |f(x)|^p \, m(\mathrm{d}x).$$

Now, we relax the assumption $f \in \mathcal{D}(\mathcal{E})$. Let f be an arbitrary function belonging to $L^p(m)$. Let $s > 0$. Denote by $H[P_s f]$ the square function H acting on $P_s f$, namely,

$$H[P_s f](x) = \left(\int_0^{+\infty} P_t \Gamma[P_t P_s f](x) \, \mathrm{d}t \right)^{1/2} = \left(\int_0^{+\infty} P_t \Gamma[P_{s+t} f](x) \, \mathrm{d}t \right)^{1/2}. \tag{8.43}$$

Note that, by the monotone convergence theorem,

$$\begin{aligned}(H(x))^2 &= \lim_{s\to 0^+} \int_s^{+\infty} P_t \Gamma[P_t f](x) \, \mathrm{d}t = \lim_{s\to 0^+} \int_0^{+\infty} P_s P_t \Gamma[P_t P_s f](x) \, \mathrm{d}t \\ &= \lim_{s\to 0^+} P_s(H[P_s f]^2)(x),\end{aligned} \tag{8.44}$$

where the above limit converges monotonically. Since the statement holds for $P_s f$,

$$\int_E (H[P_s f](x))^p \, m(\mathrm{d}x) \lesssim \int_E |P_s f(x)|^p \, m(\mathrm{d}x).$$

Therefore, again by the monotone convergence theorem,

$$\begin{aligned}\int_E (H(x))^p \, m(\mathrm{d}x) &= \lim_{s\to 0^+} \int_E \left(P_s(H[P_s f]^2)(x) \right)^{p/2} m(\mathrm{d}x) \le \lim_{s\to 0^+} \int_E (H[P_s f](x))^p \, m(\mathrm{d}x) \\ &\lesssim \lim_{s\to 0^+} \int_E |P_s f(x)|^p \, m(\mathrm{d}x) = \int_E |f(x)|^p \, m(\mathrm{d}x).\end{aligned}$$

Here, we used Jensen's inequality. The last line follows from the strong continuity of the semigroup $(P_t)_{t\ge 0}$. □

The next estimate addresses the case of $1 < p \le 2$.

Theorem 8.18. *Impose Assumptions 3.1 and 8.4. Let* $1 < p \le 2$ *and* $f \in L^p(m)$*. Then,*

$$\|f\|_p \lesssim \|H\|_p .$$

Proof. First, we assume $f \in \mathcal{D}(\mathcal{E})$. Since $p < 2$, by Jensen's inequality, we get

$$\begin{aligned}\int_E (H(x))^p\, m(\mathrm{d}x) &= \int_E \Big(\int_0^{+\infty} P_t\Gamma[P_t f](x)\,\mathrm{d}t \Big)^{p/2} m(\mathrm{d}x)\\ &= \int_E \Big(\mathbb{E}_x \int_0^{+\infty} \Gamma[P_t f](X_t)\,\mathrm{d}t \Big)^{p/2} m(\mathrm{d}x)\\ &\geq \int_E \mathbb{E}_x \Big(\int_0^{+\infty} \Gamma[P_t f](X_t)\,\mathrm{d}t \Big)^{p/2} m(\mathrm{d}x).\end{aligned}$$

Employing the monotone convergence theorem, we get

$$\int_E (H(x))^p\, m(\mathrm{d}x) \geq \lim_{T\to+\infty} \int_E \mathbb{E}_x \Big(\int_0^{T} \Gamma[P_t f](X_t)\,\mathrm{d}t \Big)^{p/2} m(\mathrm{d}x).$$

By the self-duality of the process $(X_t)_{t\geq 0}$ (formula (2.41)),

$$\int_E (H(x))^p\, m(\mathrm{d}x) \geq \lim_{T\to+\infty} \int_E \mathbb{E}_x \Big(\int_0^{T} \Gamma[P_{T-t} f](X_t)\,\mathrm{d}t \Big)^{p/2} m(\mathrm{d}x).$$

Due to the fact that $f \in \mathcal{D}(\mathcal{E})$, Corollary 8.9 yields

$$\int_E \mathbb{E}_x \Big(\int_0^{T} \Gamma[P_{T-t} f](X_t)\mathrm{d}t \Big)^{p/2} m(\mathrm{d}x) = 2^{-p/2} \int_E \mathbb{E}_x (\langle M\rangle_T)^{p/2}\, m(\mathrm{d}x),$$

where $\langle M\rangle_T$ is the angle bracket of the martingale $(M_t)_{t\geq 0}$ given by (8.22). Therefore, by the Burkholder–Davies–Gundy inequality (8.19) and inequality (8.21), we obtain

$$\begin{aligned}\int_E (H(x))^p\, m(\mathrm{d}x) &\gtrsim \lim_{T\to+\infty} \int_E \mathbb{E}_x (\langle M\rangle_T)^{p/2}\, m(\mathrm{d}x)\\ &\gtrsim \lim_{T\to+\infty} \int_E \mathbb{E}_x ([M]_T)^{p/2}\, m(\mathrm{d}x)\\ &\gtrsim \lim_{T\to+\infty} \int_E \mathbb{E}_x |M_T - M_0|^p\, m(\mathrm{d}x)\\ &= \lim_{T\to+\infty} \int_E \mathbb{E}_x |f(X_T) - P_T f(x)|^p\, m(\mathrm{d}x).\end{aligned}$$

Utilizing $|a+b|^p \geq 2^{1-p}|a|^p - |b|^p$, we get

$$\begin{aligned}\int_E (H(x))^p\, m(\mathrm{d}x) &\gtrsim \lim_{T\to+\infty} \int_E (\mathbb{E}_x |f(X_T)|^p - 2^{p-1}|P_T f(x)|^p)\, m(\mathrm{d}x)\\ &= \lim_{T\to+\infty} \int_E (P_T[|f|^p](x) - 2^{p-1}|P_T f(x)|^p)\, m(\mathrm{d}x)\\ &= \lim_{T\to+\infty} \int_E (|f(x)|^p - 2^{p-1}|P_T f(x)|^p)\, m(\mathrm{d}x).\end{aligned}$$

In the last line we used (5.10). Finally, under Assumption 3.1,

$$\int_E (H(x))^p \, m(\mathrm{d}x) \gtrsim \int_E |f(x)|^p \, m(\mathrm{d}x).$$

Now, we relax the assumption $f \in \mathcal{D}(\mathcal{E})$. Let f be an arbitrary function belonging to $L^p(m)$. The rest of the proof proceeds analogously to the proof of Theorem 8.15. Let $s > 0$. Denote by $H[P_s f]$ the square function H acting on $P_s f$, namely, define $H[P_s f]$ by (8.43). Note that (8.44) is valid. Since the statement holds for $P_s f$,

$$\int_E (H[P_s f](x))^p \, m(\mathrm{d}x) \gtrsim \int_E |P_s f(x)|^p \, m(\mathrm{d}x),$$

the monotone convergence theorem and Jensen's inequality implies

$$\begin{aligned}\int_E (H(x))^p \, m(\mathrm{d}x) &= \lim_{s\to 0^+} \int_E \left(P_s(H[P_s f]^2)(x)\right)^{p/2} \, m(\mathrm{d}x) \geq \lim_{s\to 0^+} \int_E (H[P_s f](x))^p \, m(\mathrm{d}x) \\ &\gtrsim \lim_{s\to 0^+} \int_E |P_s f(x)|^p \, m(\mathrm{d}x) = \int_E |f(x)|^p \, m(\mathrm{d}x).\end{aligned}$$

The last line follows from the strong continuity of the semigroup $(P_t)_{t\geq 0}$. □

The last result of this section stay valid for $3 \leq p < \infty$. The validity of this statement is still an open question for the range $2 < p < 3$.

Proposition 8.19. *Impose Assumption 8.4. Let* $3 \leq p \leq \infty$ *and* $f \in L^p(m)$. *Then,*

$$\|f\|_p \lesssim \|H\|_p \, .$$

This inequality is straightforward consequence of Proposition 8.21 below, due to $\widetilde{H} \leq \sqrt{2} H$.

Nevertheless, the square function H shares the same issue as the function G when $1 < p < 2$. To see this, we employ the same example proposed by Bañuelos, Bogdan, and Luks in Example 2 in [5].

Example 8.20. We continue Example 8.12. We consider the same function f as therein.

Similarly as (8.41), we can write

$$\begin{aligned}(H(x))^2 &= \mathcal{A}_{d,1} \int_0^\infty \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \frac{1}{t} \cdot \frac{(v_t(y,1) - v_t(\frac{z}{t},1))^2}{|z-ty|^{d+1}} p_t(x,z) \, \mathrm{d}y\mathrm{d}z\mathrm{d}t \\ &= \mathcal{A}_{d,1} \int_0^\infty \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} t^{d-1} \cdot \frac{(v_t(y,1) - v_t(z,1))^2}{|tz-ty|^{d+1}} p_t(x,tz) \, \mathrm{d}y\mathrm{d}z\mathrm{d}t \\ &= \mathcal{A}_{d,1} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \int_0^\infty \frac{1}{t^2} \cdot \frac{c_d (v_t(y,1) - v_t(z,1))^2}{|z-y|^{d+1} (t^2 + |tz-x|^2)^{(d+1)/2}} \, \mathrm{d}t\mathrm{d}y\mathrm{d}z. \end{aligned} \tag{8.45}$$

Recall that $v_t(1,y) \to v(1,y) > 0$ as $t \to 0^+$. We have the following asymptotic equivalence of the integrand. When $x \neq 0$,

$$\frac{1}{t^2} \cdot \frac{c_d (v_t(y,1) - v_t(z,1))^2}{|z-y|^{d+1} (t^2 + |tz-x|^2)^{(d+1)/2}} \sim \frac{c_d (v(y,1) - v(z,1))^2}{|z-y|^{d+1} \, |x|^{d+1}} \cdot \frac{1}{t^2}, \quad \text{as } t \to 0^+,$$

and when $x = 0$,

$$\frac{1}{t^2} \cdot \frac{c_d(v_t(y,1) - v_t(z,1))^2}{|z-y|^{d+1}(t^2 + |tz - x|^2)^{(d+1)/2}} \sim \frac{c_d(v(y,1) - v(z,1))^2}{|z-y|^{d+1}(1 + |z|^2)^{(d+1)/2}} \cdot \frac{1}{t^{d+3}}, \quad \text{as } t \to 0^+.$$

In both cases the inner integral of (8.41) is equal to infinity. Thus, $H \equiv +\infty$.

8.7 Square function $\widetilde{H}$

Recall that

$$\widetilde{H}(x) = \left(\int_0^{+\infty} P_t \widetilde{\Gamma}[P_t f](x)\,\mathrm{d}t \right)^{1/2}.$$

Although the lower bound has been proven for $3 \le p \le \infty$, whether this statement holds true remains unresolved for $2 < p < 3$.

Proposition 8.21. *Impose Assumptions 3.1 and 8.4. Let $3 \le p \le \infty$ and $f \in L^p(m)$. Then,*

$$\|f\|_p \lesssim \left\|\widetilde{H}\right\|_p.$$

Proof. Since $p \ge 3$, the function $t \mapsto t^{p-2}$ is convex. Thus, by Jensen's inequality $|P_t f|^{p-2} \le P_t(|f|^{p-2})$ a.e. Therefore, by (8.38), the symmetry of the operator P_t, and Hölder's inequality

$$\begin{aligned}
\int_E |f(x)|^p \, m(\mathrm{d}x) &\asymp \int_0^{+\infty} \int_E \widetilde{\Gamma}[P_t f](x) |P_t f(x)|^{p-2} \, m(\mathrm{d}x)\mathrm{d}t \\
&\le \int_0^{+\infty} \int_E \widetilde{\Gamma}[P_t f](x) P_t(|f|^{p-2})(x) \, m(\mathrm{d}x)\mathrm{d}t \\
&= \int_0^{+\infty} \int_E P_t \widetilde{\Gamma}[P_t f](x) |f(x)|^{p-2} \, m(\mathrm{d}x)\mathrm{d}t \\
&= \int_E (\widetilde{H}(x))^2 |f(x)|^{p-2} \, m(\mathrm{d}x) \\
&\le \left(\int_E (\widetilde{H}(x))^p \, m(\mathrm{d}x) \right)^{2/p} \left(\int_E |f(x)|^p \, m(\mathrm{d}x) \right)^{1-2/p}.
\end{aligned}$$

This implies

$$\int_E |f(x)|^p \, m(\mathrm{d}x) \lesssim \int_E (\widetilde{H}(x))^p \, m(\mathrm{d}x).$$

□

The next result is the straightforward implication of Theorem 8.17 and the fact that $\widetilde{H} \le \sqrt{2}H$.

Proposition 8.22. *Impose Assumption 8.4. Let* $2 \leq p < \infty$ *and* $f \in L^p(m)$. *Then,*

$$\left\|\widetilde{H}\right\|_p \lesssim \|f\|_p .$$

The behavior of the square function $\widetilde{H}$ for $1 < p < 2$ is still an open problem. The methods used earlier appear to be unsuitable for obtaining the desired estimates, at least when applied directly.

Appendix A

Calculus of multidimensional L^p space

In this appendix we extend the study of derivatives in L^p space to a multidimensional case. This part was written based on Appendix B from the joint work [15] with Bogdan and Pietruska-Pałuba. Similar results may also be found in Bogdan, Jakubowski, Lenczewska, and Pietruska-Pałuba [16]; see Lemmas 13 and 15.

Let $1 \leq p < \infty$. For fixed positive integer n, we denote by $(L^p(m))^n$ the Banach space of elements of the form $f = (f_1, \ldots, f_n)$, where $f_1, \ldots, f_n \in L^p(m)$, equipped with the norm

$$\|f\|_{L^p} := \left(\int_E |f(x)|^p \, m(\mathrm{d}x) \right)^{1/p}.$$

Recall that $|\cdot|$ is the Euclidean norm in $\mathbb{R}^n$. We shall also consider another equivalent norm

$$\|f\|_{\ell_2^n(L^p)} := \left(\sum_{j=1}^{n} \|f_j\|_p^2 \right)^{1/2}.$$

Here, $\|\cdot\|_p$ is the norm of $L^p(m)$. In particular, for $f^{(k)} = (f_1^{(k)}, \ldots, f_n^{(k)})$ in $(L^p(m))^n$

$$f^{(k)} \to f \text{ in } (L^p(m))^n \text{ if and only if } f_j^{(k)} \to f \text{ in } L^p(m) \text{ for every } j = 1, \ldots, n \tag{A.1}$$

when $k \to +\infty$.

Fix interval $I \subseteq \mathbb{R}$. For a mapping $I \ni t \mapsto u(t) \in (L^p(m))^n$ we denote

$$\Delta_h u(t) := u(t+h) - u(t) \quad \text{if } t, t+h \in I.$$

We say that u is *continuous* on I with values in $(L^p(m))^n$ if $\Delta_h u(t) \to 0$ in $(L^p(m))^n$ as $h \to 0$ for every $t \in I$. We say that u is *differentiable* on I with values in $(L^p(m))^n$ if $u'(t) := \lim_{h\to 0} \frac{1}{h} \Delta_h u(t)$ exists in $(L^p(m))^n$ for every $t \in I$. We say that u is *continuously differentiable* (or shortly C^1) on I with values in $(L^p(m))^n$ if u is differentiable and the mapping $I \ni t \mapsto u'(t) \in (L^p(m))^n$ is continuous.

The main goal of this section is to derive the following rules of differentiation, which hold for $1 < p < \infty$.

Proposition A.1 (Product rule)**.** *Let $1 < p < \infty$ and $r \in [q, \infty)$, where $q = p/(p-1)$. If the mappings $I \ni t \mapsto u(t) \in (L^p(m))^n$ and $I \ni t \mapsto v(t) \in (L^r(m))^n$ are C^1 with values in $(L^p(m))^n$ and $(L^r(m))^n$, respectively, then $u \cdot v$ is C^1 with values in $L^{pr/(p+r)}(m)$ and $(u \cdot v)' = u' \cdot v + u \cdot v'$.*

Proposition A.2 (Derivatives of power functions)**.** *Let $1<\gamma\le p<\infty$. Assume that $I\ni t\mapsto u(t)$ is differentiable with values in $(L^p(m))^n$. Then:*

(i) The mapping $|u|^\gamma$ is differentiable with values in $L^{p/\gamma}(m)$ and

$$(|u|^\gamma)' = \gamma u^{\langle\gamma-1\rangle}\cdot u'. \tag{A.2}$$

(ii) The mapping $u^{\langle\gamma\rangle}$ is differentiable with values in $(L^{p/\gamma}(m))^n$ and

$$\left(u^{\langle\gamma\rangle}\right)' = \left(J_{\langle\gamma\rangle}\circ u\right)u', \tag{A.3}$$

where $J_{\langle\gamma\rangle}(z)$ is the Jacobi matrix for the function $\mathbb{R}^n\ni z\mapsto z^{\langle\gamma\rangle}\in\mathbb{R}^n$ at a point z given by (5.6)*.*

In addition, if u is C^1, then $|u|^\gamma$ and $u^{\langle\gamma\rangle}$ are C^1 on I with values in $L^{p/\gamma}(m)$ and $(L^{p/\gamma}(m))^n$, respectively.

Before we prove the above results, we need several necessary facts. First, we need some estimates provided in [15]; see Appendix A.

Lemma A.3. *There are constants $C_\gamma, C'_\gamma, c_\gamma, c'_\gamma>0$ such that for all $w,z\in\mathbb{R}^n$,*

$$0\le \mathcal{F}_\gamma(w,z)\le C_\gamma|z-w|^\lambda(|w|\vee|z|)^{\gamma-\lambda},\qquad \lambda\in[0,2],\gamma>1, \tag{A.4}$$

$$|\mathcal{F}_{\langle\gamma\rangle}(w,z)|\le C'_\gamma|z-w|^\lambda(|w|\vee|z|)^{\gamma-\lambda},\qquad \lambda\in[0,2],\gamma>1, \tag{A.5}$$

$$||z|^\gamma-|w|^\gamma|\le c_\gamma|z-w|^\lambda(|w|\vee|z|)^{\gamma-\lambda},\qquad \lambda\in[0,1],\gamma>0, \tag{A.6}$$

$$|z^{\langle\gamma\rangle}-w^{\langle\gamma\rangle}|\le c'_\gamma|z-w|^\lambda(|w|\vee|z|)^{\gamma-\lambda},\qquad \lambda\in[0,1],\gamma>0. \tag{A.7}$$

Here, $\mathcal{F}_\gamma$ and $\mathcal{F}_{\langle\gamma\rangle}$ are defined in (5.2) *and* (5.7)*, respectively.*

Lemma A.4. *Let $0<\gamma\le p$. Then the following mappings are continuous:*

$$(L^p(m))^n\ni u\mapsto |u|^\gamma\in L^{p/\gamma}(m), \tag{A.8}$$

$$(L^p(m))^n\ni u\mapsto u^{\langle\gamma\rangle}\in (L^{p/\gamma}(m))^n. \tag{A.9}$$

Proof. In the first place, note that for $u\in(L^p(m))^n$, $|u|^\gamma$ and $u^{\langle\gamma\rangle}$ are in $L^{p/\gamma}(m)$ and $(L^{p/\gamma}(m))^n$, respectively.

To prove (A.8) and (A.9), we employ inequalities (A.6) and (A.7), respectively.

We start with the proof of (A.8). Choose $\lambda\in(0,1)$ such that $\gamma-\lambda>0$ and suppose that $u_k\to u$ in $(L^p(m))^n$ as $k\to\infty$. From (A.7) we get,

$$||u_k|^\gamma-|u|^\gamma|\le c_\gamma\,|u_k-u|^\lambda\,(|u_k|\vee|u|)^{\gamma-\lambda}$$

pointwise. Applying Hölder's inequality with exponents γ/λ and $\gamma/(\gamma-\lambda)$, we obtain

$$\begin{aligned}\||u_k|^\gamma-|u|^\gamma\|_{L^{\frac{p}{\gamma}}}^{\frac{p}{\gamma}} &= \int_E ||u_k|^\gamma-|u|^\gamma|^{\frac{p}{\gamma}}\,\mathrm{d}m\\ &\le c_\gamma^{\frac{p}{\gamma}}\int_E |u_k-u|^{\frac{\lambda p}{\gamma}}\,(|u_k|\vee|u|)^{\frac{(\gamma-\lambda)p}{\gamma}}\,\mathrm{d}m\\ &\le c_\gamma^{\frac{p}{\gamma}}\,\|u_k-u\|_{L^p}^{\frac{\lambda p}{\gamma}}\left(\|u_k\|_{L^p}^{\frac{(\gamma-\lambda)p}{\gamma}}+\|u\|_{L^p}^{\frac{(\gamma-\lambda)p}{\gamma}}\right).\end{aligned}$$

The right-hand side converges to zero, when we pass to the limit as $k\to+\infty$. This proves (A.8).

The proof of (A.9) goes similar, but utilize inequality (A.7) instead of (A.6). □

The following lemma is a multidimensional generalization of Lemma 13 from [16]. We present the proof for completeness.

Lemma A.5. *Let* $1 < p < \infty$ *and* $r \in [q, \infty)$*, where* $q = p/(p-1)$*. Let* $f \in (L^p(m))^n$ *and* $g \in (L^r(m))^n$*. Then*

$$\|f \cdot g\|_{\frac{pr}{p+r}} \leq \|f\|_{L^p} \|g\|_{L^r} .$$

Moreover, if $f_k \to f$ *in* $(L^p(m))^n$ *and* $g_k \to g$ *in* $(L^r(m))^n$*, then* $f_k \cdot g_k \to f \cdot g$ *in* $L^{pr/(p+r)}(m)$*, as* $k \to +\infty$*.*

Proof. Let $f \in (L^p(m))^n$ and $g \in (L^r(m))^n$. Then, by Hölder's inequality with exponents $(p+r)/r$ and $(p+r)/p$

$$\int_E |f \cdot g|^{\frac{pr}{p+r}} \, \mathrm{d}m \leq \int_E |f|^{\frac{pr}{p+r}} |g|^{\frac{pr}{p+r}} \, \mathrm{d}m \leq \left(\int_E |f|^p \, \mathrm{d}m \right)^{\frac{r}{p+r}} \left(\int_E |g|^r \, \mathrm{d}m \right)^{\frac{p}{p+r}} .$$

This proves desired inequality. We utilize it to prove the second statement.

Let $f_k \to f$ in $(L^p(m))^n$ and $g_k \to g$ in $(L^r(m))^n$. Then,

$$\begin{aligned} \|f_k \cdot g_k - f \cdot g\|_{\frac{pr}{p+r}} &= \|f_k \cdot (g_k - g) + (f_k - f) \cdot g\|_{\frac{pr}{p+r}} \\ &\leq \|f_k\|_{L^p} \|g_k - g\|_{L^r} + \|f_k - f\|_{L^p} \|g\|_{L^r} \to 0 \end{aligned}$$

as $k \to +\infty$. □

With Lemma A.5 at hand, we are able to prove Proposition A.1.

Proof of Proposition A.1. Fix $t \in I$. To prove the statement, we need to prove that

$$W_h(t) := \frac{1}{h}\Delta_h(u \cdot v)(t) - \frac{1}{h}\Delta_h u(t) \cdot v(t) - \frac{1}{h} u(t) \cdot \Delta_h v(t) \to 0 \quad \text{in } L^{pr/(p+r)}(m),$$

as $h \to 0$. Note that, $\Delta_h(u \cdot v)(t) = u(t+h) \cdot \Delta_h v(t) + \Delta_h u(t) \cdot v(t)$. Hence, after rearranging, we get

$$W_h(t) = \frac{1}{h}\Delta_h u(t) \cdot \Delta_h v(t) \to 0 \quad \text{in } L^{pr/(p+r)}(m),$$

by Lemma A.5. Here, we utilized the continuity and differentiability of u and v. We have proved that $(u \cdot v)' = u' \cdot v + u \cdot v'$. Again, by Lemma A.5 we derive that the mapping $u' \cdot v + u \cdot v'$ is continuous due to the fact that u and v are C^1. □

At the end we present the proof of Proposition A.2.

Proof of Proposition A.2. Both statements should be proved similarly, thus we only prove (A.3), as it is the slightly more complicated of the two.

Fix $t \in I$. To prove identity (A.3), it is enough to show that

$$\frac{1}{h}\mathcal{F}_{\langle\gamma\rangle}\left(u(t), u(t+h)\right) = \frac{1}{h}\Delta_h\left(u^{\langle\gamma\rangle}\right)(t) - (J_{\langle\gamma\rangle} \circ u(t))\frac{1}{h}\Delta_h u(t) \to 0 \tag{A.10}$$

in $(L^{p/\gamma}(m))^n$ as $h \to 0$.

We utilize (A.5) with $\lambda \in (1,2]$ such that $\gamma - \lambda > 0$ and obtain

$$\begin{aligned}\left|\frac{1}{h}\mathcal{F}_{\langle\gamma\rangle}\left(u(t),u(t+h)\right)\right| &\le \frac{1}{|h|}C'_\gamma \left|u(t+h)-u(t)\right|^{\lambda} \left(|u(t+h)| \vee |u(t)|\right)^{\gamma-\lambda} \\ &= C'_\gamma \left|h\right|^{\lambda-1} \left|\frac{1}{h}\Delta_h u(t)\right|^{\lambda} \left(|un(t+h)| \vee |u(t)|\right)^{\gamma-\lambda}\end{aligned}$$

pointwise. Thus, by Hölder's inequality with exponents γ/λ and $\gamma/(\gamma-\lambda)$,

$$\begin{aligned}\left\|\frac{1}{h}\mathcal{F}_{\langle\gamma\rangle}\left(u(t),u(t+h)\right)\right\|_{L^{\frac{p}{\gamma}}} &\le C'_\gamma \left|h\right|^{\lambda-1} \left\|\left|\frac{1}{h}\Delta_h u(t)\right|^{\lambda} \left(|un(t+h)| \vee |u(t)|\right)^{\gamma-\lambda}\right\|_{L^{\frac{p}{\gamma}}} \\ &\le C'_\gamma \left|h\right|^{\lambda-1} \left\|\frac{1}{h}\Delta_h u(t)\right\|_{L^p}^{\lambda} \left\||u(t+h)|+|u(t)|\right\|_{L^p}^{\gamma-\lambda}.\end{aligned}$$

Due to the differentiability of u with values in $(L^p(m))^n$, $\left\|\frac{1}{h}\Delta_h u(t)\right\|_{L^p}$ is bounded, therefore the right-hand side of the above inequality converges to zero as $h \to 0$. This proves (A.10) and so (A.3).

Now, we assume that u is C^1. The continuity of the mapping

$$t \mapsto \left(u^{\langle\gamma\rangle}\right)' = \left(J_{\langle\gamma\rangle} \circ u\right) u' \in (L^p(m))^n$$

follows from Lemmas A.4 and A.5.

□

Appendix B

Convexity of $z_1 z_2^{\langle p-1 \rangle} + |z|^p$

In this section we derive proofs of convexity properties required in Sections 5.3 and 5.4.

First, we recall some basic notions and facts from the theory of convex functions. Let $T\colon A \to \mathbb{R}$, where the set $A \subset \mathbb{R}^n$ is convex. We say that the vector $d(w) \in \mathbb{R}^n$ is a *subgradient* of T at a point $w \in A$ if

$$T(z) \geq T(w) + d(w) \cdot (z - w) \quad \text{for all } z \in A. \tag{B.1}$$

The function T is convex on A if and only if, for every $w \in A$, there exists a subgradient $d(w)$, not necessarily unique. In particular, if T is convex and the first-order partial derivatives of T exist at some $w \in A$, then T has exactly one subgradient at a point w, which is equal to its gradient $\nabla T(w)$.

Denote the directional derivative of T by $\frac{\partial T}{\partial v}(w)$ at a point $w \in A$ along the given vector $v \in \mathbb{R}^n$. Since T is a convex function, finite $\frac{\partial T}{\partial v}(w)$ exists for all directions $v \in \mathbb{R}^n$ and all w in the interior of A. In such a case, the vector $d(w)$ is a subgradient of function T at a point w if and only if

$$\frac{\partial T}{\partial v}(w) \geq d(w) \cdot v, \quad \text{for all } v \in \mathbb{R}^n.$$

In particular, when $\nabla T(w)$ exists, then the above inequality reduces to the well-known equality, i.e.,

$$\frac{\partial T}{\partial v}(w) = \nabla T(w) \cdot v, \quad v \in \mathbb{R}^n.$$

For more details see, for example, Borwein and Lewis [18, Chapter 3].

We start by showing the convexity of the investigated function on the first quadrant $[0, +\infty)^2$.

Lemma B.1. *Let $p \geq 2$. The function*

$$Y(z) := z_1 \, (z_2)^{p-1} + |z|^p$$

is convex on $[0, +\infty)^2$.

Proof. Since Y is continuous on $[0, +\infty)^2$, it is enough to prove the convexity on $(0, +\infty)^2$. The proof relies on investigating the Hessian $\nabla^2 Y(z)$ of Y.

Recall (5.1) and note that

$$\nabla^2 |z|^p = p(p-2)|z|^{p-4} \begin{bmatrix} z_1^2 & z_1 z_2 \\ z_1 z_2 & z_2^2 \end{bmatrix} + p|z|^{p-2}\mathrm{Id}, \quad z \in \mathbb{R}^2 \setminus \{0\}.$$

The Hessian $\nabla^2 \left(z_1 z_2^{p-1}\right)$ is calculated in (5.19). Thus, the Hessian $\nabla^2 Y(z)$ is equal to

$$\begin{bmatrix} p|z|^{p-2} + p(p-2)z_1^2|z|^{p-4} & (p-1)z_2^{p-2} + p(p-2)z_1 z_2 |z|^{p-4} \\ (p-1)z_2^{p-2} + p(p-2)z_1 z_2|z|^{p-4} & (p-1)(p-2)z_1 z_2^{p-3} + p|z|^{p-2} + p(p-2)z_2^2|z|^{p-4} \end{bmatrix}.$$

We claim that the matrix is positive semi-definite for $z \in (0, +\infty)^2$. Clearly,

$$\left[p|z|^{p-2} + p(p-2)z_1^2|z|^{p-4}\right] > 0.$$

Furthermore, after extensive yet straightforward calculations, we obtain

$$\begin{aligned} \det \nabla^2 Y(z) &= \left[p|z|^{p-2} + p(p-2)z_1^2|z|^{p-4}\right](p-1)(p-2)z_1 z_2^{p-3} \\ &\quad + p^2|z|^{2p-4} + p^2(p-2)|z|^{2p-4} \\ &\quad - (p-1)^2 z_2^{2p-4} - p(p-1)(p-2)z_1 z_2^{p-1}|z|^{p-4} \\ &= p^2(p-1)|z|^{2p-4} - (p-1)^2 z_2^{2p-4} \\ &\quad + p(p-1)(p-2)|z|^{p-4}\left((p-1)z_1^3 z_2^{p-3} - z_1 z_2^{p-1}\right). \end{aligned}$$

Clearly $z_2 \le |z|$, thus applying Young's inequality with exponents p and $q = p/(p-1)$ to the product $z_1 z_2^{p-1}$ we obtain

$$z_1 z_2^{p-1} \le \frac{z_1^p}{p} + \frac{(p-1)z_2^p}{p} = \frac{1}{p}(z_1^2)^{\frac{p}{2}} + \frac{p-1}{p}(z_2^2)^{\frac{p}{2}} \le |z|^p.$$

Summarizing,

$$\begin{aligned} \det \nabla^2 Y(z) &\ge p(p-1)^2(p-2)|z|^{p-4} z_1^3 z_2^{p-3} \\ &\quad + |z|^{2p-4}(p-1)\left(p^2 - p(p-2) - (p-1)\right) \\ &= p(p-1)^2(p-2)|z|^{p-4}z_1^3 z_2^{p-3} + |z|^{2p-4}(p-1)(p+1) > 0. \end{aligned}$$

□

Now, we require additional notions and facts from the convex analysis. If $w_1 \le z_1, w_2 \le z_2, \dots, w_n \le z_n$ implies $T(w_1, \dots, w_n) \le T(z_1, \dots, z_n)$ in the domain of a real-valued function T, then we say that T is *coordinate-wise non-decreasing.* The following fact is straightforward, see also Boyd and Vandenberghe [22, Section 3.2.4].

Lemma B.2. *Let $S\colon A \to \mathbb{R}^n$, $S(A) \subset B$, and $T\colon B \to \mathbb{R}$, where $A \subset \mathbb{R}^m$ and $B \subset \mathbb{R}^n$ are convex. If each coordinate of S is convex and T is coordinate-wise non-decreasing and convex, then the composition $T \circ S\colon A \to \mathbb{R}$ is convex.*

We need also the following property. We may refer for example to [22, Section 3.2.3].

Lemma B.3. *Let $T, S\colon A \to \mathbb{R}$ be convex. Then* $\max\{T, S\}$ *is also convex.*

With Lemma B.1 at hand we are prepare to provide the following two lemmas, which are crucial to prove the results in Chapter 5. The first one derives the convexity of auxiliary functions on the positive half-plane $[0,+\infty)\times\mathbb{R}$. This provides the non-negativity of theirs second-order Taylor remainders. This fact is essential to prove the polarized Hardy–Stein identity from Section 5.3.

Lemma B.4. *Let $p \geq 2$. Functions*

$$Y^{(+)}(z) := z_1\left((z_2)_+\right)^{p-1} + |z|^p, \qquad Y^{(-)}(z) := z_1\left((z_2)_-\right)^{p-1} + |z|^p$$

are convex on $[0,+\infty)\times\mathbb{R}$.

Proof. Define $T\colon [0,+\infty)\times\mathbb{R} \to [0,+\infty)^2$ as

$$T(z) := (z_1, (z_2)_+),$$

and define $Y\colon [0,+\infty)^2 \to \mathbb{R}$ as in Lemma B.1. Due to the fact that each coordinate of T is convex and the function Y is convex and coordinate-wise non-decreasing, the composition

$$(Y\circ T)(z) = z_1((z_2)_+)^{p-1} + \Big((z_1)^2 + ((z_2)_+)^2\Big)^{p/2}$$

is convex on $[0,+\infty)\times\mathbb{R}$ in view of Lemma B.2. Therefore,

$$Y^{(+)}(z) = \max\{(Y\circ T)(z), |z|^p\}$$

is also convex on $[0,+\infty)\times\mathbb{R}$ as the maximum of convex functions, by Lemma B.3.

The convexity of $Y^{(-)}$ follows from the convexity of $Y^{(+)}$ by the fact that $Y^{(-)}(z_1,z_2) = Y^{(+)}(z_1,-z_2)$. □

Lemma B.5. *If $p > 2$, then for all $z, w \in \mathbb{R}^2$,*

$$\mathcal{J}_p^{(++)}(w,z) + \mathcal{F}_p(w,z) \geq 0, \quad \mathcal{J}_p^{(-+)}(w,z) + \mathcal{F}_p(w,z) \geq 0,$$

where $\mathcal{F}_p$, $\mathcal{J}_p^{(++)}$, and $\mathcal{J}_p^{(-+)}$ are given in (5.2), (5.32), *and* (5.33).

Proof. In view of (5.5) and (5.35), we only need to show the first inequality. We rewrite it as follows

$$Y^{(++)}(z) \geq Y^{(++)}(w) + d(w)\cdot(z-w), \quad z, w \in \mathbb{R}^2, \tag{B.2}$$

where $Y^{(++)}(z) := (z_1)_+\left((z_2)_+\right)^{p-1} + |z|^p$ and

$$d(w) := \Big(\mathbb{1}(w_1)\left((w_2)_+\right)^{p-1}, (p-1)(w_1)_+\left((w_2)_+\right)^{p-2}\Big) + pw^{\langle p-1\rangle}.$$

Therefore, the proof of (B.2) is equivalent to verifying that $d(w)$ is a subgradient of the function $Y^{(++)}$ at a point $w\in\mathbb{R}^2$.

Hence, we first establish the convexity of $Y^{(++)}$. Define $T\colon \mathbb{R}^2 \to [0,+\infty)^2$ as

$$T(z) := ((z_1)_+, (z_2)_+).$$

Let $Y\colon [0,+\infty)^2 \to \mathbb{R}$ as in Lemma B.1. By the fact that each coordinate of T is convex and the function Y is convex and coordinate-wise non-decreasing, the composition

$$(Y\circ T)(z) = (z_1)_+((z_2)_+)^{p-1} + \Big(((z_1)_+)^2 + ((z_2)_+)^2\Big)^{p/2}$$

is convex by Lemma B.2. Since

$$Y^{(++)}(z) = \max\{(Y \circ T)(z), |z|^p\},$$

it is convex on $\mathbb{R}^2$ as a maximum of two convex functions, according to Lemma B.3.

Now, we show that $d(w)$ is a subgradient of $Y^{(++)}$.

If $w = 0$, then $Y^{(++)}(w) = 0$ and $d(w) = 0$. Hence, (B.2) is clear.

In case of $w \neq 0$, to show that $d(w)$ is a subgradient of $Y^{(++)}$ at w, we will prove that

$$\frac{\partial Y^{(++)}}{\partial v}(w) \geq d(w) \cdot v, \quad w \in \mathbb{R}^2 \setminus \{0\}, \quad \text{for every } v = (v_1, v_2) \in \mathbb{R}^2.$$

Denote $B := \{ (w_1, w_2) \in \mathbb{R}^2 : w_1 = 0, w_2 > 0 \}$ – the vertical positive semi-axis. The function $Y^{(++)}$ is differentiable everywhere, but on B. Therefore, when $w \notin B$, the gradient of $Y^{(++)}$ exists, is equal to $d(w)$, thus

$$\frac{\partial Y^{(++)}}{\partial v}(w) = \nabla Y^{(++)}(w) \cdot v = d(w) \cdot v.$$

In the remaining case, where $w \in B$, we have two possibilities. First, when $v_1 \geq 0$, then

$$\begin{aligned} \frac{\partial Y^{(++)}}{\partial v}(w) &= ((w_2)_+)^{p-1} v_1 + p w^{\langle p-1\rangle} \cdot v \\ &\geq \frac{1}{2} ((w_2)_+)^{p-1} v_1 + p w^{\langle p-1\rangle} \cdot v = d(w) \cdot v. \end{aligned}$$

Otherwise, if $v_1 < 0$, then

$$\frac{\partial Y^{(++)}}{\partial v}(w) = p w^{\langle p-1\rangle} \cdot v \geq \frac{1}{2} ((w_2)_+)^{p-1} v_1 + p w^{\langle p-1\rangle} \cdot v = d(w) \cdot v.$$

The proof is complete. □

Bibliography

[1] Rodrigo Bañuelos. Brownian motion and area functions. *Indiana Univ. Math. J.*, 35(3):643–668, 1986.

[2] Rodrigo Bañuelos and Charles N. Moore. *Probabilistic behavior of harmonic functions*, volume 175 of *Progress in Mathematics*. Birkhäuser Verlag, Basel, 1999.

[3] Dominique Bakry. Functional inequalities for Markov semigroups. In *Probability measures on groups: recent directions and trends*, pages 91–147. Tata Inst. Fund. Res., Mumbai, 2006.

[4] Dominique Bakry, Richard D. Gill, and Stanislav A. Molchanov. *Lectures on probability theory*, volume 1581 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1994. Lectures from the Twenty-second Saint-Flour Summer School held July 9–25, 1992.

[5] Rodrigo Bañuelos, Krzysztof Bogdan, and Tomasz Luks. Hardy–Stein identities and square functions for semigroups. *J. Lond. Math. Soc. (2)*, 94(2):462–478, 2016.

[6] Rodrigo Bañuelos and Daesung Kim. Hardy–Stein identity for non-symmetric Lévy processes and Fourier multipliers. *J. Math. Anal. Appl.*, 480(1):20, 2019. Id/No 123383.

[7] Gerassimos Barbatis, Stathis Filippas, and Achilles Tertikas. A unified approach to improved L^p Hardy inequalities with best constants. *Trans. Amer. Math. Soc.*, 356(6):2169–2196, 2004.

[8] Martin T. Barlow, Saul D. Jacka, and Marc Yor. Inequalities for a pair of processes stopped at a random time. *Proc. London Math. Soc. (3)*, 52(1):142–172, 1986.

[9] Agnes I. Benedek, Alberto P. Calderón, and Rafael Panzone. Convolution operators on Banach space valued functions. *Proc. Nat. Acad. Sci. U.S.A.*, 48:356–365, 1962.

[10] Andrew G. Bennett. Probabilistic square functions and a priori estimates. *Trans. Amer. Math. Soc.*, 291(1):159–166, 1985.

[11] Robert M.C. Blumenthal and Ronald K. Getoor. *Markov processes and potential theory*, volume Vol. 29 of *Pure and Applied Mathematics*. Academic Press, New York-London, 1968.

[12] Krzysztof Bogdan, Bartłomiej Dyda, and Tomasz Luks. On Hardy spaces of local and nonlocal operators. *Hiroshima Math. J.*, 44(2):193–215, 2014.

[13] Krzysztof Bogdan, Tomasz Grzywny, Katarzyna Pietruska-Pałuba, and Artur Rutkowski. Extension and trace for nonlocal operators. *J. Math. Pures Appl. (9)*, 137:33–69, 2020.

[14] Krzysztof Bogdan, Tomasz Grzywny, Katarzyna Pietruska-Pałuba, and Artur Rutkowski. Nonlinear nonlocal Douglas identity. *Calc. Var. Partial Differ. Equ.*, 62(5):31, 2023. Id/No 151.

[15] Krzysztof Bogdan, Michał Gutowski, and Katarzyna Pietruska-Pałuba. Polarized Hardy–Stein identity, 2023. arXiv:2309.09856.

[16] Krzysztof Bogdan, Tomasz Jakubowski, Julia Lenczewska, and Katarzyna Pietruska-Pałuba. Optimal Hardy inequality for the fractional Laplacian on L^p. *J. Funct. Anal.*, 282(8):31, 2022. Id/No 109395.

[17] Krzysztof Bogdan, Dominik Kutek, and Katarzyna Pietruska-Pałuba. Bregman variation of semimartingales, 2023. In preparation.

[18] Jonathan M. Borwein and Adrian S. Lewis. *Convex analysis and nonlinear optimization*, volume 3 of *CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC*. Springer, New York, second edition, 2006. Theory and examples.

[19] Nicolas Bouleau and Francis Hirsch. Formes de Dirichlet générales et densité des variables aléatoires réelles sur l'espace de Wiener. (General Dirichlet forms and density of real random variables on Wiener space). *J. Funct. Anal.*, 69:229–259, 1986.

[20] Nicolas Bouleau and Francis Hirsch. *Dirichlet forms and analysis on Wiener space*, volume 14 of *De Gruyter Studies in Mathematics*. Walter de Gruyter & Co., Berlin, 1991.

[21] Nicolas Bouleau and Damien Lamberton. Théorie de Littlewood-Paley-Stein et processus stables. In *Séminaire de Probabilités, XX, 1984/85*, volume 1204 of *Lecture Notes in Math.*, pages 162–185. Springer, Berlin, 1986.

[22] Stephen Boyd and Lieven Vandenberghe. *Convex optimization*. Cambridge University Press, Cambridge, 2004.

[23] Donald L. Burkholder. Martingale transforms. *Ann. Math. Statist.*, 37:1494–1504, 1966.

[24] Donald L. Burkholder and Richard F. Gundy. Extrapolation and interpolation of quasi-linear operators on martingales. *Acta Math.*, 124:249–304, 1970.

[25] Donald L. Burkholder, Richard F. Gundy, and Martin L. Silverstein. A maximal function characterization of the class H^p. *Trans. Amer. Math. Soc.*, 157:137–153, 1971.

[26] Alberto P. Calderón. On a theorem of Marcinkiewicz and Zygmund. *Trans. Amer. Math. Soc.*, 68:55–61, 1950.

[27] Alberto P. Calderón. On the behaviour of harmonic functions at the boundary. *Trans. Amer. Math. Soc.*, 68:47–54, 1950.

[28] Alberto P. Calderón and Antoni Zygmund. On the existence of certain singular integrals. *Acta Math.*, 88:85–139, 1952.

[29] Eric A. Carlen, Shigeo Kusuoka, and Daniel W. Stroock. Upper bounds for symmetric Markov transition functions. *Ann. Inst. Henri Poincaré, Probab. Stat.*, 23:245–287, 1987.

[30] Zhen-Qing Chen. On reflected Dirichlet spaces. *Probab. Theory Related Fields*, 94(2):135–162, 1992.

[31] Zhen-Qing Chen and Masatoshi Fukushima. A localization formula in Dirichlet form theory. *Proc. Amer. Math. Soc.*, 140(5):1815–1822, 2012.

[32] Zhen-Qing Chen and Takashi Kumagai. Heat kernel estimates for stable-like processes on d-sets. *Stochastic Process. Appl.*, 108(1):27–62, 2003.

[33] Zhen-Qing Chen, Takashi Kumagai, and Jian Wang. Stability of heat kernel estimates for symmetric non-local Dirichlet forms. *Mem. Amer. Math. Soc.*, 271(1330):v+89, 2021.

[34] Zhen-Qing Chen, Takashi Kumagai, and Jian Wang. Heat kernel estimates for general symmetric pure jump Dirichlet forms. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)*, 23(3):1091–1140, 2022.

[35] Alberto Cialdea and Vladimir Maz'ya. *Semi-bounded differential operators, contractive semigroups and beyond*, volume 243 of *Operator Theory: Advances and Applications*. Birkhäuser/Springer, Cham, 2014.

[36] Edward B. Davies. *Heat Kernels and Spectral Theory*. Cambridge Tracts in Mathematics. Cambridge University Press, 1989.

[37] Eugene Fabes, Masatoshi Fukushima, Leonard Gross, Carlos Kenig, Michael Röckner, and Daniel W. Stroock. *Dirichlet forms*, volume 1563 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1993. Lectures given at the First C.I.M.E. Session held in Varenna, June 8–19, 1992.

[38] Eugene B. Fabes, David S. Jerison, and Carlos E. Kenig. Multilinear square functions and partial differential equations. *Amer. J. Math.*, 107(6):1325–1368, 1985.

[39] Walter Farkas, Niels Jacob, and René L. Schilling. Feller semigroup, L^p-sub-Markovian semigroups, and applications to pseudo-differential operators with negative definite symbols. *Forum Math.*, 13(1):51–90, 2001.

[40] Charles L Fefferman and Elias M. Stein. H^p spaces of several variables. *Acta Math.*, 129(3-4):137–193, 1972.

[41] Patrick J. Fitzsimmons. On the quasi-regularity of semi-Dirichlet forms. *Potential Anal.*, 15(3):151–185, 2001.

[42] Patrick J. Fitzsimmons and Ronald K. Getoor. Revuz measures and time changes. *Math. Z.*, 199(2):233–256, 1988.

[43] Patrick J. Fitzsimmons and Ronald K. Getoor. Smooth measures and continuous additive functionals of right Markov processes. In *Itô's stochastic calculus and probability theory*, pages 31–49. Springer, Tokyo, 1996.

[44] Gerald B. Folland. *Real analysis: Modern techniques and their applications*. Pure and Applied Mathematics (New York). John Wiley & Sons, Inc., New York, second edition, 1999. A Wiley-Interscience Publication.

[45] Masatoshi Fukushima, Yoichi Oshima, and Masayoshi Takeda. *Dirichlet forms and symmetric Markov processes*, volume 19 of *De Gruyter Studies in Mathematics*. Walter de Gruyter & Co., Berlin, extended edition, 2011.

[46] J. García-Cuerva, editor. *Harmonic analysis and partial differential equations*, volume 1384 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1989.

[47] Ronald K. Getoor. Multiplicative functionals of dual processes. *Ann. Inst. Fourier (Grenoble)*, 21(2):43–83, 1971.

[48] Ronald K. Getoor and Michael J. Sharpe. Naturality, standardness, and weak duality for Markov processes. *Z. Wahrsch. Verw. Gebiete*, 67(1):1–62, 1984.

[49] Richard F. Gundy and Elias M. Stein. H^p theory for the poly-disc. *Proc. Nat. Acad. Sci. U.S.A.*, 76(3):1026–1029, 1979.

[50] Michał Gutowski. Hardy–Stein identity for pure-jump Dirichlet forms. *Bull. Pol. Acad. Sci. Math.*, 71(1):65–84, 2023.

[51] Michał Gutowski and Mateusz Kwaśnicki. Beurling–Deny formula for Sobolev–Bregman forms, 2023. arXiv:2312.10824.

[52] Niels Jacob. *Pseudo differential operators and Markov processes. In 3 vol. Vol. 1: Fourier analysis and semigroups*. Imperial College Press, London, 2001.

[53] Niels Jacob. *Pseudo differential operators and Markov processes. In 3 vol. Vol. 2: Generators and their potential theory*. Imperial College Press, London, 2002.

[54] Niels Jacob. *Pseudo differential operators and Markov processes. In 3 vol. Vol. 3: Markov processes and applications*. Imperial College Press, London, 2005.

[55] Stefan Kaczmarz. Über die Konvergenz der Reihen von Orthogonalfunktionen. *Math. Z.*, 23(1):263–270, 1925.

[56] Deniz Karlı. Harnack inequality and regularity for a product of symmetric stable process and Brownian motion. *Potential Anal.*, 38(1):95–117, 2013.

[57] Michał Kijaczko and Julia Lenczewska. Sharp Hardy inequalities for Sobolev–Bregman forms. *Mathematische Nachrichten*, 2023. In press.

[58] Ildoo Kim and Kyeong-Hun Kim. A generalization of the Littlewood-Paley inequality for the fractional Laplacian $(-\Delta)^{\alpha/2}$. *J. Math. Anal. Appl.*, 388(1):175–190, 2012.

[59] Damir Kinzebulatov and Yuliy A. Semënov. Fractional Kolmogorov operator and desingularizing weights. *Publ. Res. Inst. Math. Sci.*, 59(2):339–391, 2023.

[60] Kazuhiro Kuwae. Reflected Dirichlet forms and the uniqueness of Silverstein's extension. *Potential Anal.*, 16(3):221–247, 2002.

[61] Mateusz Kwaśnicki. Ten equivalent definitions of the fractional Laplace operator. *Fract. Calc. Appl. Anal.*, 20(1):7–51, 2017.

[62] Erik Lenglart, Dominique Lépingle, and Maurizio Pratelli. Présentation unifiée de certaines inégalités de la théorie des martingales. In *Seminar on Probability, XIV (Paris, 1978/1979) (French)*, volume 784 of *Lecture Notes in Math.*, pages 26–52. Springer, Berlin, 1980. With an appendix by Lenglart.

[63] Huaiqian Li and Jian Wang. Littlewood–Paley–Stein estimates for non-local Dirichlet forms. *J. Anal. Math.*, 143(2):401–434, 2021.

[64] Liping Li and Jiangang Ying. Bivariate Revuz measures and the Feynman–Kac formula on semi-Dirichlet forms. *Potential Anal.*, 42(4):775–808, 2015.

[65] Vitali A. Liskevich and Motti A. Perel'muter. Analyticity of sub-Markovian semigroups. *Proc. Amer. Math. Soc.*, 123(4):1097–1104, 1995.

[66] Vitali A. Liskevich, Motti A. Perel'muter, and Yuliy A. Semënov. Form-bounded perturbations of generators of sub-Markovian semigroups. *Acta Appl. Math.*, 44(3):353–377, 1996.

[67] Vitali A. Liskevich and Yuliy A. Semënov. Some inequalities for sub-Markovian generators and their applications to the perturbation theory. *Proc. Amer. Math. Soc.*, 119(4):1171–1177, 1993.

[68] Vitali A. Liskevich and Yuliy A. Semënov. Some problems on Markov semigroups. In *Schrödinger operators, Markov semigroups, wavelet analysis, operator algebras*, volume 11 of *Math. Top.*, pages 163–217. Akademie Verlag, Berlin, 1996.

[69] John E. Littlewood and Raymond E. A. C. Paley. Theorems on Fourier Series and Power Series. *J. London Math. Soc.*, 6(3):230–233, 1931.

[70] John E. Littlewood and Raymond E. A. C. Paley. Theorems on Fourier Series and Power Series (II). *Proc. London Math. Soc. (2)*, 42(1):52–89, 1936.

[71] John E. Littlewood and Raymond E. A. C. Paley. Theorems on Fourier Series and Power Series (III). *Proc. London Math. Soc. (2)*, 43(2):105–126, 1937.

[72] Zhi Ming Ma and Michael Röckner. *Introduction to the theory of (nonsymmetric) Dirichlet forms*. Universitext. Springer-Verlag, Berlin, 1992.

[73] Józef Marcinkiewicz. Sur quelques intégrales du type de Dini. *Annales de la Société Polonaise de Mathématique*, 17:42–50, 1938.

[74] Józef Marcinkiewicz. Sur les multiplicateurs des séries de Fourier. *Studia Math.*, 8(1):78–91, 1939.

[75] Józef Marcinkiewicz and Antoni Zygmund. A theorem of Lusin. *Duke Math. J.*, 4(3):473–485, 1938.

[76] Giorgio Metafune and Chiara Spina. An integration by parts formula in Sobolev spaces. *Mediterr. J. Math.*, 5(3):357–369, 2008.

[77] Paul-A. Meyer. Intégrales stochastiques. I, II, III, IV. In *Séminaire de Probabilités (Univ. Strasbourg, Strasbourg, 1966/67), Vol. I*, pages 72–94, 95–117, 118–141, 142–162. Springer, Berlin-New York, 1967.

[78] Paul-A. Meyer. Correction aux "Inégalités de Littlewood-Paley. II. L'opérateur carré du champ". In *Séminaire de Probabilités, X*, volume Vol. 511 of *Lecture Notes in Math.*, pages 162–163. Springer, Berlin-New York, 1976. (Première partie, Univ Strasbourg, Strasbourg, année universitaire 1974/1975).

[79] Paul-A. Meyer. Démonstration probabiliste de certaines inégalités de Littlewood-Paley. I. Les inégalités classiques. In *Séminaire de Probabilités, X*, volume Vol. 511 of *Lecture Notes in Math.*, pages 125–141. Springer, Berlin-New York, 1976. (Première partie, Univ Strasbourg, Strasbourg, année universitaire 1974/1975).

[80] Paul-A. Meyer. Démonstration probabiliste de certaines inégalités de Littlewood-Paley. II. L'opérateur carré du champ. In *Séminaire de Probabilités, X*, volume Vol. 511 of *Lecture Notes in Math.*, pages 142–161. Springer, Berlin-New York, 1976. (Première partie, Univ Strasbourg, Strasbourg, année universitaire 1974/1975).

[81] Paul-A. Meyer. Démonstration probabiliste de certaines inégalités de Littlewood-Paley. III. Les inégalités générales. In *Séminaire de Probabilités, X*, volume Vol. 511 of *Lecture Notes in Math.*, pages 164–174. Springer, Berlin-New York, 1976. (Première partie, Univ Strasbourg, Strasbourg, année universitaire 1974/1975).

[82] Paul-A. Meyer. Démonstration probabiliste de certaines inégalités de Littlewood-Paley. IV. Semi-groupes de convolution symétriques. In *Séminaire de Probabilités, X*, volume Vol. 511 of *Lecture Notes in Math.*, pages 175–183. Springer, Berlin-New York, 1976. (Première partie, Univ Strasbourg, Strasbourg, année universitaire 1974/1975).

[83] Paul-A. Meyer. Correction à: "Démonstration probabiliste de certaines inégalités de Littlewood-Paley, IV" [Séminaire de Probabilités, X (Première partie, Univ. Strasbourg, Strasbourg, année universitaire 1974/1975), pp. 175–183, Lecture Notes in Math., 511, Springer, Berlin, 1976; MR **58** #18751e]. In *Séminaire de Probabilités, XII (Univ. Strasbourg, Strasbourg, 1976/1977)*, volume 649 of *Lecture Notes in Math.*, page p. 741. Springer, Berlin, 1978.

[84] Paul-A. Meyer. Retour sur la théorie de Littlewood-Paley. In *Seminar on Probability, XV (Univ. Strasbourg, Strasbourg, 1979/1980) (French)*, volume 850 of *Lecture Notes in Math.*, pages 151–166. Springer, Berlin, 1981.

[85] Paul-A. Meyer. Transformations de Riesz pour les lois gaussiennes. In *Seminar on probability, XVIII*, volume 1059 of *Lecture Notes in Math.*, pages 179–193. Springer, Berlin, 1984.

[86] Paul-A. Meyer. Sur la théorie de Littlewood-Paley-Stein (d'après Coifman-Rochberg-Weiss et Cowling). In *Séminaire de probabilités, XIX, 1983/84*, volume 1123 of *Lecture Notes in Math.*, pages 113–129. Springer, Berlin, 1985.

[87] Alexander Nagel, Elias M. Stein, and Stephen Wainger. Differentiation in lacunary directions. *Proc. Nat. Acad. Sci. U.S.A.*, 75(3):1060–1062, 1978.

[88] Raymond E. A. C. Paley. A Remarkable Series of Orthogonal Functions (I). *Proc. London Math. Soc. (2)*, 34(4):241–264, 1932.

[89] Raymond E. A. C. Paley. A Remarkable Series of Orthogonal Functions (II). *Proc. London Math. Soc. (2)*, 34(4):265–279, 1932.

[90] Yehuda Pinchover, Achilles Tertikas, and Kyril Tintarev. A Liouville-type theorem for the p-Laplacian with potential term. *Ann. Inst. H. Poincaré C Anal. Non Linéaire*, 25(2):357–368, 2008.

[91] Daniel Revuz. Mesures associées aux fonctionnelles additives de Markov. I. *Trans. Amer. Math. Soc.*, 148:501–531, 1970.

[92] René L. Schilling and Toshihiro Uemura. On the structure of the domain of a symmetric jump-type Dirichlet form. *Publ. Res. Inst. Math. Sci.*, 48(1):1–20, 2012.

[93] Itai Shafrir. Asymptotic behaviour of minimizing sequences for Hardy's inequality. *Commun. Contemp. Math.*, 2(2):151–189, 2000.

[94] Michael J. Sharpe. *General theory of Markov processes*, volume 133 of *Pure and Applied Mathematics.* Academic Press, Inc., Boston, MA, 1988.

[95] Elias M. Stein. On the functions of Littlewood-Paley, Lusin, and Marcinkiewicz. *Trans. Amer. Math. Soc.*, 88:430–466, 1958.

[96] Elias M. Stein. Classes H^p, multiplicateurs et fonctions de Littlewood-Paley. *C. R. Acad. Sci. Paris Sér. A-B*, 263:A716–A719, 1966.

[97] Elias M. Stein. Classes H^p, multiplicateurs et fonctions de Littlewood-Paley. Applications de résultats antérieurs. *C. R. Acad. Sci. Paris Sér. A-B*, 263:A780–A781, 1966.

[98] Elias M. Stein. Classes H^p et multiplicateurs: Cas n-dimensionnel. *C. R. Acad. Sci. Paris Sér. A-B*, 264:A107–A108, 1967.

[99] Elias M. Stein. *Singular integrals and differentiability properties of functions*, volume 30 of *Princeton Math. Ser.* Princeton University Press, Princeton, NJ, 1970.

[100] Elias M. Stein. *Topics in harmonic analysis related to the Littlewood–Paley theory*, volume 63 of *Ann. Math. Stud.* Princeton University Press, Princeton, NJ, 1970.

[101] Elias M. Stein. Maximal functions. I. Spherical means. *Proc. Nat. Acad. Sci. U.S.A.*, 73(7):2174–2175, 1976.

[102] Elias M. Stein. The development of square functions in the work of A. Zygmund. In *Conference on harmonic analysis in honor of Antoni Zygmund, Vol. I, II (Chicago, Ill., 1981)*, Wadsworth Math. Ser., pages 2–30. Wadsworth, Belmont, CA, 1983.

[103] Elias M. Stein and Stephen Wainger. Problems in harmonic analysis related to curvature. *Bull. Amer. Math. Soc.*, 84(6):1239–1295, 1978.

[104] Elias M. Stein and Antoni Zygmund. On the differentiability of functions. *Studia Math.*, 23:247–283, 1964.

[105] Elias M. Stein and Antoni Zygmund. On the fractional differentiability of functions. *Proc. London Math. Soc. (3)*, 14a:249–264, 1965.

[106] Daniel W. Stroock. *An introduction to the theory of large deviations*. Universitext. Springer-Verlag, New York, 1984.

[107] Gerald Trutnau. Stochastic calculus of generalized Dirichlet forms and applications to stochastic differential equations in infinite dimensions. *Osaka J. Math.*, 37(2):315–343, 2000.

[108] Nicholas Th. Varopoulos. Hardy–Littlewood theory for semigroups. *J. Funct. Anal.*, 63:240–260, 1985.

[109] Nicolas Th. Varopoulos. Aspects of probabilistic Littlewood-Paley theory. *J. Functional Analysis*, 38(1):25–60, 1980.

[110] Feng-Yu Wang. *Functional Inequalities, Markov Semigroups and Spectral Theory*. Science Press, Beijing, 2004.

[111] Guido Weiss and Stephen Wainger, editors. *Harmonic analysis in Euclidean spaces. Part 2*, volume XXXV, Part 2 of *Proceedings of Symposia in Pure Mathematics*. American Mathematical Society, Providence, RI, 1979. Dedicated to Nestor M. Rivière.

[112] Mary Weiss and Antoni Zygmund. A note on smooth functions. *Indag. Math.*, 21:52–58, 1959. Nederl. Akad. Wetensch. Proc. Ser. A **62**.

[113] Antoni Zygmund. Une remarque sur un théorème de M. Kaczmarz. *Math. Z.*, 25(1):297–298, 1926.

[114] Antoni Zygmund. Sur l'application de la première moyenne arithmétique dans la théorie des séries de fonctions orthogonales. *Fund. Math.*, 10(1):356–362, 1927.

[115] Antoni Zygmund. On the Convergence and Summability of Power Series on the Circle of Convergence (I). *Fund. Math.*, 30:170–196, 1938.

[116] Antoni Zygmund. On the Convergence and Summability of Power Series on the Circle of Convergence (II). *Proc. London Math. Soc. (2)*, 47:326–350, 1942.

[117] Antoni Zygmund. On certain integrals. *Trans. Amer. Math. Soc.*, 55:170–204, 1944.

[118] Antoni Zygmund. Proof of a theorem of Littlewood and Paley. *Bull. Amer. Math. Soc.*, 51:439–446, 1945.

[119] Antoni Zygmund. On the Littlewood-Paley function $g^*(\theta)$. *Proc. Nat. Acad. Sci. U.S.A.*, 42:208–212, 1956.

[120] Antoni Zygmund. *Trigonometric series. 2nd ed. Vols. I, II.* Cambridge University Press, New York, 1959.