



陈木法文选

(卷三)

Mu-Fa Chen

Selected Papers

(Volume III)

Beijing Normal University

Contents

Volume III

- [37] Speed of stability for birth–death processes, *Front. Math. China* 2010, 5:3, 379–515 768
- [38] Basic estimates of stability rate for one-dimensional diffusions, Chapter 6 in “Probability Approximations and Beyond”, 75–99, *Lecture Notes in Statistics* 205, Eds: A.D. Barbour, H.P. Chan and D. Siegmund, 2012 900
- [39] General estimate of the first eigenvalue on manifolds, *Front. Math. China* 2011, 6(6): 1025–1043 924
- [40] Lower bounds of the principal eigenvalue in dimension one, *Front. Math. China* 2012, 7(4): 645–668 947
- [41] Mixed principal eigenvalues in dimension one, *Front. Math. China* 2013, 8(2): 317–343 970
- [42] Bilateral Hardy-type inequalities, *Acta Math. Sin. Eng. Ser.* 2013, 29:1, 1–32 1017
- [43] The optimal constant in Hardy-type inequalities, *Acta Math. Sin. Eng. Ser.* 2015, 31(5): 731–754 1063
- [44] Progress on Hardy-type inequalities, Chapter 7 in the book “Festschrift Masatoshi Fukushima”, 131–142. Eds: Z.Q. Chen, N. Jacob, M. Takeda, and T. Uemura, *World Sci.* 2015 1092
- [45] Mixed eigenvalues of discrete p -Laplacian, *Front. Math. China* 2014, 9(6): 1261–1292 1104
- [46] Mixed eigenvalues of p -Laplacian, *Front. Math. China* 2015, 10(2): 249–274 1162

Volume I

- [01] Coupling for jump processes, Acta Math. Sin. New Ser. 1986, 2:2, 123–136. 1
- [02] With S.F. Li. Coupling methods for multidimensional diffusion processes, Ann. of Probab. 1989, 17:1, 151–177..... 14
- [03] Exponential L^2 -convergence and L^2 -spectral gap for Markov processes, Math. Sin. New Ser. 1991, 7:1, 19–37. 43
- [1] With F.Y. Wang. Application of coupling method to the first eigenvalue on manifold, Sci. Sin.(A) 1993, 23:11 (Chinese Edition), 1130–1140; 1994, 37:1 (English Edition), 1–14. 65
- [2] Optimal Markovian couplings and applications, Acta Math. Sin. New Ser. 1994, 10:3, 260–275. 82
- [3] Optimal couplings and application to Riemannian geometry, Prob. Theory and Math. Stat., Vol. 1, Eds. B. Grigelionis et al, 1994 VPS/TEV, 121–142. 102
- [4] On the ergodic region of Schlögl’s model, in Proceedings of International Conference on Dirichlet Forms and Stochastic Processes, Edited by Z.M. Ma, M. Röckner and J.A. Yan, Walter de Gruyter Publishers, 1995, 87–102. 124
- [5] With F.Y. Wang. Estimation of the first eigenvalue of second order elliptic operators, J. Funct. Anal. 1995, 131:2, 345–363. 139
- [6] With F.Y. Wang. Estimates of logarithmic Sobolev constant, J. Funct. Anal. 1997, 144:2, 287–300. 156
- [7] Estimation of spectral gap for Markov chains, Acta Math. Sin. New Series 1996, 12:4, 337–360. 169
- [8] With F.Y. Wang. Estimation of spectral gap for elliptic operators, Trans. Amer. Math. Soc. 1997, 349:3, 1239–1267. 203
- [9] With F.Y. Wang. General formula for lower bound of the first eigenvalue on Riemannian manifolds, Sci. Sin. 1997, 27:1, 34–42 (Chinese Edition); 1997, 40:4, 384–394 (English Edition) 237
- [10] Trilogy of couplings and general formulas for lower bound of spectral gap, in “Probability Towards 2000”, Edited by L. Accardi and C. Heyde, Lecture Notes in Statis. **128**, 123–136, Springer-Verlag, 1998. 249
- [11] Coupling, spectral gap and related topics (I), Chin. Sci. Bulletin, 1997, 42:14, 1472–1477 (Chinese Edition); 1997, 42:16, 1321–1327 (English Edition). 262
- [12] Coupling, spectral gap and related topics (II), Chin. Sci. Bulletin, 1997, 42:15, 1585–1591 (Chinese Edition); 1997, 42:17, 1409–1416 (English Edition). 270

- [13] Coupling, spectral gap and related topics (III), Chin. Sci. Bulletin, 1997, 42:16, 1696–1703 (Chinese Edition); 1997, 42:18, 1497–1505 (English Edition). 279
- [14] Estimate of exponential convergence rate in total variation by spectral gap, Acta Math. Sin. Ser. (A) 1998, 41:1, 1–6; Acta Math. Sin. New Ser. 1998, 14:1, 9–16. 289
- [15] With F.Y. Wang. Cheeger’s inequalities for general symmetric forms and existence criteria for spectral gap. Abstract: Chin. Sci. Bulletin 1998, 43:14 (Chinese Edition), 1475–1477; 43:18 (English Edition), 1516–1519. Ann. Prob. 2000, 28:1, 235–257 305
- [16] Analytic proof of dual variational formula for the first eigenvalue in dimension one, Sci. in China (A) 1999, 29:4 (Chinese Edition), 327–336; 42:8 (English Edition), 805–815. 331
- [17] Nash inequalities for general symmetric forms, Acta Math. Sin. Eng. Ser. 1999, 15:3, 353–370. 348

Volume II

- [18] Equivalence of exponential ergodicity and L^2 -exponential convergence for Markov chains, Stoch. Proc. Appl. 2000, 87:2, 281–297 376
- [19] Logarithmic Sobolev inequality for symmetric forms, Sci. in China (A) 2000, 30:3 (Chinese Edition), 203–209; 43:6 (English Edition), 601–608. 399
- [20] A new story of ergodic theory, in “Applied Probability”, 25–34, eds. R. Chan et al., AMS/IP Studies in Advanced Mathematics, vol. 26, 2002. 411
- [21] Eigenvalues, inequalities and ergodic theory, Chin. Sci. Bulletin, 1999, 44:23 (In Chinese), 2465–2470; 2000, 45:9 (English Edition), 769–774. 424
- [22] Eigenvalues, inequalities and ergodic theory (II), Adv. in Math. (China) 1999, 28:6, 481–505. 432
- [23] The principal eigenvalue for jump processes, Acta Math. Sin. Eng. Ser. 2000, 16:3, 361–368 463
- [24] With Y.Z. Wang. Algebraic Convergence of Markov Chains, Ann. Appl. Prob. 2003, 13:2, 604–627 474
- [25] Explicit bounds of the first eigenvalue, Sci. in China (A) 2000, 39:9 (Chinese Edition), 769–776; 43:10 (English Edition), 1051–1059 497
- [26] Variational formulas and approximation theorems for the first eigenvalue, Sci. in China (A) 2001, 31:1 (Chinese Edition), 28–36; 44:4 (English Edition), 409–418 511

- [27] With E. Scacciatelli and L. Yao. Linear approximation of the first eigenvalue on compact manifolds, *Sci. in China (A)* 2001, 31:9 (Chinese Edition), 807–816; 2002 (English Edition), 45:4, 450–461 535
- [28] With Y.H. Zhang and X.L. Zhao. Dual Variational Formulas for the First Dirichlet Eigenvalue on Half-Line, *Sci. in China (A)* 2003, 33:4 (Chinese Edition), 371–383; (English Edition), 46:6, 847–861 551
- [29] Ergodic Convergence Rates of Markov Processes — Eigenvalues, Inequalities and Ergodic Theory, in *Proceedings of “ICM 2002”*, Vol. III, 41–52, Higher Education Press, Beijing 574
- [30] Variational formulas of Poincaré-type inequalities in Banach spaces of functions on the line, *Acta Math. Sin. Eng. Ser.* 2002, 18:3, 417–436 587
- [31] Variational formulas of Poincaré-type inequalities for birth-death processes, *Acta Math. Sin. Eng. Ser.* 2003, 19:4, 625–644 615
- [32] Variational formulas and explicit bounds of Poincaré-type inequalities for one-dimensional processes, *IMS Lecture Notes – Monograph Series*, Volume 41, 81–96 640
- [33] Ten explicit criteria of one-dimensional processes, *Advanced Studies in Pure Mathematics*, Vol. 39, 89–114, 2004, Mathematical Society of Japan 656
- [34] Capacitary criteria for Poincaré-type inequalities, *Potential Analysis* 2005, 23:4, 303–322 677
- [35] Exponential convergence rate in entropy, *Front. Math. China* 2007, 2:3, 329–358 696
- [36] Spectral gap and logarithmic Sobolev constant for continuous spin systems, *Acta Math. Sin. New Ser.* 2008, 24:5, 705–736 728

SPEED OF STABILITY FOR BIRTH–DEATH PROCESSES

MU-FA CHEN

(Beijing Normal University)

May 25, 2007—November 10, 2009

ABSTRACT. This paper is a continuation of the study on the stability speed for Markov processes. It extends the previous study of the ergodic convergence speed to the non-ergodic one, in which the processes are even allowed to be explosive or to have general killings. At the beginning stage, this paper is concentrated on the birth-death processes. According to the classification of the boundaries, there are four cases plus one having general killings. In each case, some dual variational formulas for the convergence rate are presented, from which the criterion for the positivity of the rate and an approximating procedure of estimating the rate are deduced. As the first step of the approximation, the ratio of the resulting bounds is usually no more than 2. The criteria as well as basic estimates for more general types of stability are also presented. Even though the paper contributes mainly to the non-ergodic case, there are some improvements in the ergodic one. To illustrate the power of the results, a large number of examples are included.

CONTENTS

§ 1. Introduction	2
§ 2. Absorbing (Dirichlet) boundary at infinity: dual variational formulas ..	7
§ 3. Absorbing (Dirichlet) boundary at infinity: criterion, approximating procedure and examples	27
§ 4. Absorbing (Dirichlet) boundary at origin and reflecting (Neumann) boundary at infinity	39
§ 5. Dual approach	50
§ 6. Reflecting (Neumann) boundaries at origin and infinity (ergodic case) ..	59
§ 7. Bilateral absorbing (Dirichlet) boundaries	71
§ 8. Criteria for Poincaré-type inequalities	92
§ 9. General killing	99
§ 10. Notes	121

2000 *Mathematics Subject Classification.* 60J27, 34L15.

Key words and phrases. Birth–death process, speed of stability, first eigenvalue, variational formula, criterion and basic estimates, approximating procedure, duality, killing.

Research supported in part by the Creative Research Group Fund of the National Natural Science Foundation of China (No. 10721091), by the “985” project from the Ministry of Education in China.

Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$ -TEX

References	130

1. INTRODUCTION

Consider a birth-death process on the nonnegative integers $\mathbb{Z}_+$ with birth rates $b_n > 0$ ($n \geq 0$) and death rates $a_n > 0$ ($n \geq 1$). Define

$$\mu_0 = 1, \quad \mu_n = \frac{b_0 \cdots b_{n-1}}{a_1 \cdots a_n}, \quad n \geq 1. \quad (1.1)$$

We say that the birth-death process is nonexplosive if the following Dobrushin's uniqueness criterion holds:

$$\sum_{k=0}^{\infty} \frac{1}{b_k \mu_k} \sum_{i=0}^k \mu_i \left[= \sum_{i=0}^{\infty} \mu_i \sum_{k=i}^{\infty} \frac{1}{b_k \mu_k} \right] = \infty \quad (1.2)$$

(cf. Dobrushin (1952), or Wang and Yang (1992, Corollary 5.2.1), or [10; Corollary 3.18]). This implies a useful condition that

$$\sum_{k=0}^{\infty} \left(\frac{1}{b_k \mu_k} + \mu_k \right) = \infty. \quad (1.3)$$

When $\sum_0^{\infty} \mu_k < \infty$, each of (1.2) and (1.3) is equivalent to the recurrent condition: $\sum_0^{\infty} (b_n \mu_n)^{-1} = \infty$. Otherwise, (1.3) cannot imply (1.2) since one can easily construct a counterexample so that $\sum_0^{\infty} \mu_k = \infty$ but

$$\sum_{i=0}^{\infty} \mu_i \sum_{k=i}^{\infty} \frac{1}{b_k \mu_k} < \infty.$$

Thus, under (1.3), the process may not be unique.

It is well known that for a birth-death process, the transition probabilities $(p_{ij}(t))$ satisfy

$$\lim_{t \rightarrow \infty} p_{ij}(t) =: \pi_j \geq 0 \quad (1.4)$$

for all $i, j \in \mathbb{Z}_+$. We are now interested in the exponential convergence rate

$$\alpha^* = \sup \left\{ \alpha : |p_{ij}(t) - \pi_j| = O(\exp[-\alpha t]) \text{ as } t \rightarrow \infty \text{ for all } i, j \in E \right\}. \quad (1.5)$$

In the ergodic case (i.e., $\lim_{t \rightarrow \infty} p_{ij}(t) > 0$ for all i, j), we have $Z := \sum_{j=0}^{\infty} \mu_j < \infty$ and then $\pi_j := \mu_j/Z > 0$ for all $j \geq 0$. In this case, the problem has been well studied, see, for instance, van Doorn (1981; 2002), Zeifman (1991), Kijima (1997), [2, 12], and the references therein. The problem becomes trivial in the zero-recurrent case for general irreducible Markov chains, since we have on the

one hand $\pi_j = 0$ for all j , and on the other hand, $\int_0^\infty p_{ii}(t)dt = \infty$ for all i . Hence, the exponential decay can only happen in the transient case:

$$\sum_{n=0}^{\infty} \frac{1}{b_n \mu_n} < \infty. \quad (1.6)$$

Since the process is μ -symmetric: $\mu_i p_{ij}(t) = \mu_j p_{ji}(t)$ for all i, j and t , it is natural, as we did in the ergodic case, to use the L^2 -theory. As usual, denote by $\|\cdot\|$ and $(\cdot, \cdot)$ the norm and the inner product on the real Hilbert space $L^2(\mu)$, respectively. Let

$$\mathcal{K} = \{f : f \text{ has finite support}\}. \quad (1.7)$$

Define

$$D(f) = \sum_{i \geq 0} \mu_i b_i (f_{i+1} - f_i)^2 = \sum_{i \geq 1} \mu_i a_i (f_i - f_{i-1})^2$$

with the minimal domain $\mathcal{D}^{\min}(D)$ consisting of the functions in the closure of $\mathcal{K}$ with respect to the norm $\|\cdot\|_D$: $\|f\|_D^2 = \|f\|^2 + D(f)$. Next, define

$$\lambda_0 = \inf\{D(f) : \|f\| = 1, f \in \mathcal{K}\} = \inf\{D(f) : \|f\| = 1, f \in \mathcal{D}^{\min}(D)\}.$$

From now on, we often write f_∞ or $f(\infty)$ as the limit of f at infinity provided it exists. In the definition of λ_0 , it is natural to add the boundary condition $f_\infty = 0$ but this can be ignored since on the one hand, for each $f \in \mathcal{K}$, we have $f_\infty = 0$, and on the other hand $\mathcal{K}$ is a core of the Dirichlet form $(D, \mathcal{D}^{\min}(D))$ (i.e., the form is regular) by [10; Proposition 6.59]. For a large part of the paper, we are dealing with this minimal Dirichlet form or the minimal process.

We now make a connection between α^* and λ_0 . The proofs of the next three propositions are delayed for a moment.

Proposition 1.1. For a general non-ergodic symmetric semigroup $\{P_t\}_{t \geq 0}$ with Dirichlet form $(D, \mathcal{D}(D))$ (not necessarily regular) on $L^2(\mu)$, the parameter λ_0 ,

$$\lambda_0 = \inf\{D(f) : \|f\| = 1, f \in \mathcal{D}(D)\}, \quad (1.8)$$

is the largest ε such that

$$\|P_t f\| \leq \|f\| e^{-\varepsilon t}, \quad t \geq 0, f \in L^2(\mu). \quad (1.9)$$

It was proved in [2; Theorem 5.3] that for birth-death processes, under (1.2), the exponentially ergodic convergence rate α^* coincides with the L^2 -exponential one, denoted by λ_1 :

$$\|P_t f - \pi(f)\| \leq \|f - \pi(f)\| e^{-\lambda_1 t} \quad \text{for all } t \geq 0 \text{ and } f \in L^2(\mu),$$

where $\pi(f) = \int f d\mu / \mu(E)$. For non-ergodic birth-death processes, we have similarly $\alpha^* = \lambda_0$, as mentioned at the end of [2]. Here is a generalization.

Proposition 1.2. For a general non-ergodic μ -symmetric Markov chain with Dirichlet form $(D, \mathcal{D}(D))$, we have $\alpha^* = \lambda_0$ defined by (1.8).

About (1.3), we have the following result.

Proposition 1.3. Let $\mathcal{D}^{\max}(D) = \{f \in L^2(\mu) : D(f) < \infty\}$. Then the Dirichlet form $(D, \mathcal{D}^{\max}(D))$ is regular iff (1.3) holds. In other words, the Dirichlet form corresponding to the rates (a_i) and (b_i) is unique iff (1.3) holds.

Proposition 1.2 reduces the study on α^* to the first (or principal) eigenvalue λ_0 . This is the starting point of this paper. In the two cases we have discussed so far, the state 0 is a reflecting (Neumann) boundary, denoted by code “N”. For λ_1 , since the process starting from any point will certainly come back, the infinity may be regarded as a reflecting (Neumann) boundary. However, for λ_0 , the situation is different. As we will prove in the next section, the corresponding eigenfunction decreases to zero at infinity. Hence, the infinity may be regarded as an absorbing (Dirichlet) boundary, denoted by code “D”. Thus, for the temporary convenience, we rewrite $\lambda_1 = \lambda^{\text{NN}}$ and $\lambda_0 = \lambda^{\text{ND}}$. Replacing the Neumann boundary at 0 by the Dirichlet one (i.e., $b_0 = 0$), we obtain two more cases for which we have the decay rates (eigenvalues) λ^{DN} and λ^{DD} , respectively. The main body of this paper is devoted to study these four cases. Now, the rate α^* coincides with, case by case, one of λ^{NN} , λ^{ND} , λ^{DN} , and λ^{DD} . Here are simple examples to show the difference in the different cases.

Examples 1.4.

- (1) Let $a_i = \delta i$, $b_i = \beta i + \gamma$, $\delta > \beta$. Then $\lambda^{\text{NN}} = \delta - \beta$ if $\gamma > 0$ and so does λ^{DN} if $\gamma = 0$.
- (2) Let $a_i = i$, $b_i = 2(i + \gamma)$. Then $\lambda^{\text{ND}} = \gamma$ if $\gamma > 0$ and $\lambda^{\text{DD}} = 1$ if $\gamma = 0$.

The rate in the first example is the difference of the coefficients of leading terms, independent of γ . This is somehow natural. Surprisingly, the second one is determined by the constant term only except $\gamma = 0$ at which case there is a jump from λ^{ND} to λ^{DD} . Thus, for the convergence rate, the role played by the parameters (a_i, b_i) is mazed and then one may wonder how far we can go (see Theorem 1.5 below for a preliminary answer).

The main body of the paper is devoted to the quantitative study of the convergence rate. For this, our key result (variational formulas) plays a full power. For those readers who are interested only in the qualitative criteria and basic estimates, here is a short statement.

Theorem 1.5 (Criterion and basic estimates). Let (1.3) hold. Then in spite of $b_0 > 0$ or $b_0 = 0$, the exponential convergence rate α^* defined in (1.5) for the unique process is positive

- (1) iff $\delta^{(4.4)} < \infty$ in the case of $\sum_i \mu_i < \infty$; and otherwise,
- (2) iff $\delta^{(3.1)} < \infty$,

where

$$\delta^{(4.4)} = \sup_{n \geq 1} \sum_{i=1}^n \frac{1}{\mu_i a_i} \sum_{j=n}^{\infty} \mu_j, \quad \delta^{(3.1)} = \sup_{n \geq 0} \sum_{i=0}^n \mu_i \sum_{j=n}^{\infty} \frac{1}{\mu_j b_j}.$$

More precisely, we have the basic estimate $\delta^{-1}/4 \leq \alpha^* \leq \delta^{-1}$, where the constant δ is equal to $\kappa^{(6.13)}$ or $\kappa^{(7.5)}$ according to $b_0 > 0$ or $b_0 = 0$, respectively:

$$\begin{aligned} (\kappa^{(6.13)})^{-1} &= \inf_{m \geq n \geq 0} \left[\left(\sum_{i=0}^n \mu_i \right)^{-1} + \left(\sum_{i=m}^{\infty} \mu_i \right)^{-1} \right] \left(\sum_{j=n}^{m-1} \frac{1}{\mu_j b_j} \right)^{-1}, \\ (\kappa^{(7.5)})^{-1} &= \inf_{m \geq n \geq 1} \left[\left(\sum_{i=1}^n \frac{1}{\mu_i a_i} \right)^{-1} + \left(\sum_{i=m}^{\infty} \frac{1}{\mu_i b_i} \right)^{-1} \right] \left(\sum_{j=n}^m \mu_j \right)^{-1}. \end{aligned}$$

Here, the superscript of $\kappa^{(7.5)}$, for instance, means that it is in the case studied in Section 7 and the constant is given in (7.5).

The proof of Theorem 1.5 and its extension are given in Section 7. The more general qualitative results are presented in Section 8 and in Summary 9.12 for the killing case.

To have an impression about the progress made in the paper, let us have a look at the new points made in the well-developed case, Section 6.

- (1) The uniqueness condition (1.2) is replaced by using the maximal process.
- (2) The more complete dual variational formulas are presented in Theorem 6.1.
- (3) Even though the criterion Theorem 6.2 is known before, the upper bound in its improvement (Corollary 6.4) is newly added so that the ratio of the bounds is now no more than 2 as shown by a group of examples. Moreover, a new criterion (Corollary 6.6) which has been expected naturally (in view of Theorem 4.2) for a long time, is now presented.
- (4) A more effective sequence for the upper estimate given in Theorem 6.3 is introduced to replace the original one. The monotonicity of the approximating sequences are proved here for the first time.

We now return to prove the propositions above.

Proof of Proposition 1.1. Replace by $\varepsilon_{\max}$ the largest exponential rate in (1.9). Then we have $\varepsilon_{\max} \geq 0$ because of the contractivity of the semigroup in every L^p -space ($p \geq 1$). We need to show that $\lambda_0 = \varepsilon_{\max}$. The proof of $\lambda_0 \geq \varepsilon_{\max}$ is easier since by an elementary property of the Dirichlet form and (1.9), we have for every f with $\|f\| = 1$,

$$D(f) = \lim_{t \downarrow 0} \uparrow \frac{1}{t} (f - P_t f, f) \geq \lim_{t \downarrow 0} \frac{1}{t} (1 - e^{-\varepsilon_{\max} t}) = \varepsilon_{\max}, \quad (1.10)$$

where $\lim \uparrow$ means an increasing limit. Hence, we have $\lambda_0 \geq \varepsilon_{\max}$. To prove $\varepsilon_{\max} \geq \lambda_0$, assume that $\lambda_0 > 0$. Otherwise, the assertion is trivial. Noticing that $D(f) = (-\Omega f, f)$ for the generator Ω of $\{P_t\}$ on $L^2(\mu)$ and for every $f \in \mathcal{D}(\Omega)$, we have

$$\frac{d}{dt} \|P_t f\|^2 = 2(P_t f, \Omega P_t f) = -2D(P_t f). \quad (1.11)$$

Next, since $P_t f \in \mathcal{D}(D)$ for each $f \in L^2(\mu)$, by the definition of λ_0 , we have

$$-2D(P_t f) \leq -2\lambda_0 \|P_t f\|^2.$$

Thus, $\|P_t f\| \leq \|f\| e^{-\lambda_0 t}$ for all $t \geq 0$ and $f \in \mathcal{D}(\Omega)$, and then for all $f \in L^2(\mu)$ since the density of $\mathcal{D}(\Omega)$ in $L^2(\mu)$ and the contractivity of the semigroup $\{P_t\}_{t \geq 0}$. The assertion now follows since $\varepsilon_{\max}$ is the largest rate. $\square$

Proof of Proposition 1.2. The proof for $\alpha^* \geq \lambda_0$ is rather easy. Simply applying Proposition 1.1 to the indicator function $f = \mathbb{1}_{\{k\}}$, we obtain

$$p_{ik}(t) \leq \sqrt{\mu_k/\mu_i} e^{-\lambda_0 t}.$$

Note that this also provides a non-trivial estimate of the constant in (1.5).

To prove that $\lambda_0 \geq \alpha^*$, we may assume that $\alpha^* > 0$. One may follow the proof of [12; proof of part (4) of Theorem 8.13]. In the last part of the original proof, we have

$$\|P_t f\|^2 = (f, P_{2t} f) \leq \|f\|_\infty^2 e^{-2\alpha^* t} \sum_{i,j \in \text{supp}(f)} \mu_i C_{ij}$$

for every bounded f with compact support. Here, we have used the assumption that $p_{ij}(t) \leq C_{ij} e^{-\alpha^* t}$. $\square$

Proof Proposition 1.3. Since the Q -matrix is conservative, by [10; Lemma 6.52 and Theorem 6.61], $(D, \mathcal{D}^{\max}(D))$ is a Dirichlet form and is indeed the maximal one. Note that in the conservative case, every Q -process (in particular, the semigroup generated by a Dirichlet form) satisfies the backward Kolmogorov's equation by [10; Theorem 1.15 (1)].

(a) Let (1.3) hold. Then the Dirichlet form should be regular. Otherwise, we have two different birth-death semigroups generated by $(D, \mathcal{D}^{\max}(D))$ and the minimal Dirichlet form $(D, \mathcal{D}^{\min}(D))$, respectively. They satisfy first the backward and then also the forward Kolmogorov's equations by [10; Theorem 6.16]. This is impossible since condition (1.3) is the uniqueness criterion for the process satisfying the Kolmogorov's equations simultaneously, due to Karlin and McGregor (1957a, Theorem 15) (cf. Hou et al. (2000, Theorem 6.4.6 (1); 1994, Theorem 12.7.1)). Note that criterion (1.3) is equivalent to the uniqueness for the process satisfying one of the Kolmogorov equations since every symmetric process as well as the minimal one satisfies both of the equations. This is the reason why (1.3) is weaker than (1.2).

(b) Next, let (1.3) fail. Then we have $\sum_i \mu_i < \infty$ and $\sum_i (\mu_i b_i)^{-1} < \infty$. Moreover, (1.2) fails. Note that the birth-death Q -matrix has at most a single exit boundary, and there is precisely one if (1.2) fails. Besides, the non-trivial (maximal) exit solution z_λ is bounded from above by 1. In view of [10; Proposition 6.56], there are infinitely many Dirichlet forms. The minimal one is regular but not the maximal one $(D, \mathcal{D}^{\max}(D))$. $\square$

Actually, Proposition 1.3 is a particular case of a result we will study at the end of Section 9 (Theorem 9.22).

The remainder of the paper is organized as follows. In the next two sections, we study λ^{ND} . Sections 4, 6 and 7 are devoted to λ^{DN} , λ^{NN} and λ^{DD} , respectively. By exchanging N and D, we formally obtain a dual of λ^{ND} and λ^{DN} (resp. λ^{NN} and λ^{DD}) which is studied in Section 5 (resp. 7). In each case, we present a group of dual variational formulas for the first (non-trivial) eigenvalue. By using the

formulas, we then deduce a criterion for the positivity of the eigenvalue and an approximating procedure for estimating the eigenvalue. The criteria and basic estimates in a quite general setup are given in Section 8. A closely related topic, having general killings, is studied in Section 9. In the study of this paper, the author has benefited a great deal from our previous work and from many authors' contribution. A part of the contributions is noted in the context. In the ergodic case under (1.2), a large number of references are given in [10, 12] and the author apologizes for omitting them here. At the end of the paper (Section 10), some remarks on the related results, some open problems or open topics, and so on are discussed. The analog of Theorem 1.5 for one-dimensional diffusions is also included.

Notation 1.6. To be economical, we use the same notation λ_0 , δ , κ , I and II and so on, from time to time in different sections with similar but different meaning. To distinguish them if necessary, we write $\lambda_0^{(\#)}$ for instance to denote the λ_0 defined by formula (#).

2. ABSORBING (DIRICHLET) BOUNDARY AT INFINITY: DUAL VARIATIONAL FORMULAS

This section begins with the study on the property of eigenfunction of λ_0 . It is fundamental in our analysis and has been studied several times before, see, for instance, [3; Lemma 4.2]; [4; proofs of Theorems 3.2 and 3.4]; Chen, Zhang and Zhao (2003, Section 2); Shao and Mao (2007, Proposition 3.1). The main body of this section is devoted to prove a group of variational formulas (Theorem 2.4 and Proposition 2.5). Their applications are given in the next section.

Fix an integer N : $1 \leq N \leq \infty$, and let $E = \{k \in \mathbb{Z}_+ : 0 \leq k < N + 1\}$. Throughout the paper, the infinite case that $N = \infty$ is more essential but the finite case that $N < \infty$ is also included which may be meaningful in matrices theory. To avoid the confusion of these two cases in reading the paper, one may read the infinite case first and then go back to check the modification for the finite case. Besides, note that when $N < \infty$, neither (1.2) nor (1.3) is needed.

Let us start at a general situation. Consider the operator Ω corresponding to the birth-death Q -matrix with birth rates b_i , death rates a_i , and killing rates $c_i \geq 0$ ($i \in E$) as follows.

$$\Omega f(i) = b_i(f_{i+1} - f_i) + a_i(f_{i-1} - f_i) - c_i f_i, \quad i \in E, \quad f_{N+1} = 0 \text{ if } N < \infty. \quad (2.1)$$

In other words, when $N < \infty$, the state $N+1$ is an absorbing (Dirichlet) boundary. When $c_i \neq 0$ for $1 \leq i < N$, unless otherwise stated, we assume that $a_0 = 0$ and $b_N = 0$ if $N < \infty$ (the other a_i and b_i are positive), otherwise, simply replace the original c_0 and c_N by $a_0 + c_0$ and $b_N + c_N$, respectively. Now, since $a_0 = 0$, f_{-1} is free in the last formula. The first eigenvalue λ_0 is now defined by

$$\lambda_0 = \inf\{D(f) : \|f\| = 1, f \in \mathcal{K}\}, \quad (2.2)$$

where

$$D(f) = \sum_{0 \leq i < N} \mu_i b_i (f_{i+1} - f_i)^2 + \sum_{i \in E} \mu_i c_i f_i^2, \quad f_{N+1} = 0 \text{ if } N < \infty. \quad (2.3)$$

We say that g is an “eigenfunction” of $\lambda \in \mathbb{R}$, if g satisfies the “eigenequation”:

$$\Omega g = -\lambda g, \quad g_{N+1} = 0 \text{ if } N < \infty. \quad (2.4)$$

Note that the “eigenvalue” and “eigenfunction” used in this paper are in a generalized sense rather than the standard ones since here we do not require $g \in L^2(\mu)$.

Proposition 2.1.

- (1) Every eigenfunction g of $\lambda \in \mathbb{R}$ satisfies

$$\mu_k b_k(g_k - g_{k+1}) = \sum_{i=0}^k (\lambda - c_i) \mu_i g_i, \quad k \in E, \quad g_{N+1} = 0 \text{ if } N < \infty. \quad (2.5)$$

- (2) If $\lambda_0 > 0$, then $c_i \neq 0$ ($0 \leq i \leq N$) whenever $N < \infty$, and the non-zero eigenfunction g of λ_0 is either positive or negative on E .
 (3) The non-zero eigenfunction g of $\lambda = 0$ is either positive and nondecreasing, or negative and nonincreasing on E . Furthermore, let $g > 0$ for instance. Then $g_{k+1} > g_k$ for all $k : i \leq k < N$ whenever $c_i > 0$.

Proof. (a) Recall the eigenequation

$$\Omega g(i) = b_i(g_{i+1} - g_i) + a_i(g_{i-1} - g_i) - c_i g_i = -\lambda g_i, \quad i \in E, \quad (2.6)$$

or more generally, the Poisson equation

$$b_i(g_i - g_{i+1}) - a_i(g_{i-1} - g_i) = h_i, \quad i \in E, \quad g_{N+1} = 0 \text{ if } N < \infty, \quad (2.7)$$

for a given function h . Multiplying both sides by μ_i , we get

$$\mu_i b_i(g_i - g_{i+1}) - \mu_{i-1} b_{i-1}(g_{i-1} - g_i) = \mu_i h_i, \quad i \in E. \quad (2.8)$$

When $i = 0$, the second term on the left-hand side is set to be zero. Making a summation over i , we obtain

$$\mu_k b_k(g_k - g_{k+1}) = \sum_{i=0}^k \mu_i h_i, \quad k \in E, \quad g_{N+1} = 0 \text{ if } N < \infty. \quad (2.9)$$

With $h_i = (\lambda - c_i)g_i$, this gives us the first assertion of the proposition.

(b) To prove the second assertion, note that $\lambda_0 = 0$ if $c_i \equiv 0$ ($0 \leq i \leq N < \infty$) in which case both 0 and N are reflecting and the process is ergodic. Now, since $\lambda_0 > 0$, one may assume that $g_0 \neq 0$, otherwise $g_i \equiv 0$ by induction. Next, replacing g by g/g_0 if necessary, we can assume that $g_0 = 1$. If g is not positive, then there would exist a $k_0 \in E$, $k_0 \geq 1$ such that $g_i > 0$ for $i < k_0$ and $g_{k_0} \leq 0$. We then modify g from k_0 : set $\tilde{g}_i = g_i$ for $i < k_0$ and $\tilde{g}_i = 0$ for $i \geq k_0 + 1$. By choosing a suitable value $\varepsilon > 0$ at k_0 , the new function $\tilde{g} \in \mathcal{X}$ gives us $D(\tilde{g})/\|\tilde{g}\|^2 < \lambda_0$, which is a contradiction to the definition of λ_0 . Hence, g does not change its sign.

We are now going to specify ε . Note that

$$\begin{aligned} (-\Omega\tilde{g})(k_0 - 1) &= -b_{k_0-1}(\varepsilon - g_{k_0-1}) + a_{k_0-1}(g_{k_0-1} - g_{k_0-2}) + c_{k_0-1}g_{k_0-1} \\ &= (-\Omega g)(k_0 - 1) + b_{k_0-1}(g_{k_0} - \varepsilon) \\ &= \lambda_0 g_{k_0-1} + b_{k_0-1}(g_{k_0} - \varepsilon) \\ &< \lambda_0 g_{k_0-1} \end{aligned}$$

since $\varepsilon > 0 \geq g_{k_0}$. Note also that

$$(-\Omega\tilde{g})(k_0) = -b_{k_0}(0 - \varepsilon) + a_{k_0}(\varepsilon - g_{k_0-1}) + c_{k_0}\varepsilon = \varepsilon(a_{k_0} + b_{k_0} + c_{k_0}) - a_{k_0}g_{k_0-1}.$$

Next, since $D(f) = (f, -\Omega f)$ for every $f \in \mathcal{X}$ and for each i , $\Omega f(i)$ depends on three points i and $i \pm 1$ only, we obtain

$$\begin{aligned} D(\tilde{g}) &= \sum_{0 \leq i \leq k_0-2} \mu_i g_i (-\Omega g)(i) + \mu_{k_0-1} g_{k_0-1} (-\Omega\tilde{g})(k_0 - 1) + \mu_{k_0} \tilde{g}_{k_0} (-\Omega\tilde{g})(k_0) \\ &< \lambda_0 \sum_{i=0}^{k_0-1} \mu_i g_i^2 + \varepsilon \mu_{k_0} [\varepsilon(a_{k_0} + b_{k_0} + c_{k_0}) - a_{k_0}g_{k_0-1}]. \end{aligned}$$

Because

$$\|\tilde{g}\|^2 = \sum_{i=0}^{k_0-1} \mu_i g_i^2 + \mu_{k_0} \varepsilon^2,$$

for $D(\tilde{g})/\|\tilde{g}\|^2 < \lambda_0$, it suffices that

$$\varepsilon[\varepsilon(a_{k_0} + b_{k_0} + c_{k_0}) - a_{k_0}g_{k_0-1}] < \lambda_0 \varepsilon^2.$$

Equivalently,

$$\varepsilon(a_{k_0} + b_{k_0} + c_{k_0} - \lambda_0) < a_{k_0}g_{k_0-1}.$$

This clearly holds for sufficiently small $\varepsilon > 0$.

(c) If $\lambda = 0$, then (2.5) becomes

$$\mu_k b_k (g_{k+1} - g_k) = \sum_{i=0}^k c_i \mu_i g_i, \quad k \in E, \quad g_{N+1} = 0 \text{ if } N < \infty. \quad (2.10)$$

Clearly, if $g_0 = 0$, then $g_i \equiv 0$ by induction. Without loss of generality, assume that $g_0 = 1$. By (2.10) and induction, it follows that $g_{k+1} - g_k \geq 0$ for all $i \in E$. Actually, $g_{k+1} > g_k$ for all k : $i \leq k < N$ provided $c_i > 0$. $\square$

In view of (2.5), the eigenfunction g may not be monotone if $c_i \neq 0$.

For the remainder of this section, we assume that $c_i = 0$ for $i < N$ but $c_N > 0$ if $N < \infty$. However, to simplify our notation, set $c_i \equiv 0$ but let $b_N > 0$ if $N < \infty$. In view of the definition of the state space E , the point $N + 1$ is regarded as a Dirichlet boundary. From now on in the paper, when we talk about $\lambda_0^{(2.2)}$, it is defined by (2.2) but in the present setting.

Proposition 2.2.

- (1) Let g be a non-zero eigenfunction of $\lambda_0 > 0$. Then g is either positive or negative.
- (2) Let g be a positive eigenfunction of $\lambda > 0$. Then g is strictly decreasing. Furthermore,

$$\sum_{k=n}^N \frac{1}{\mu_k b_k} \sum_{i=0}^k \mu_i g_i = \frac{g_n - g_{N+1}}{\lambda}, \quad n \in E. \quad (2.11)$$

In particular,

$$\sum_{n=0}^N \mu_n g_n \nu[n, N] = \sum_{n=0}^N \nu_n \sum_{k=0}^n \mu_k g_k = \frac{g_0 - g_{N+1}}{\lambda} < \infty, \quad (2.12)$$

where $\nu[\ell, m] = \sum_{k=\ell}^m \nu_k$, $\nu_k = (\mu_k b_k)^{-1}$. Moreover, if (1.2) holds, then $g_\infty := \lim_{N \rightarrow \infty} g_N = 0$.

- (3) Let $\lambda_0 = 0$. Then $N = \infty$ and the eigenfunction g must be a constant function.

Proof. (a) The first assertion follows from Proposition 2.1 (2).

(b) Let $\lambda > 0$. Since $g > 0$, by (2.5) with $c_i \equiv 0$, it follows that g_i is strictly decreasing in i . By (2.5) again, we have

$$g_n - g_{N+1} = \sum_{k=n}^N (g_k - g_{k+1}) = \lambda \sum_{k=n}^N \frac{1}{\mu_k b_k} \sum_{i=0}^k \mu_i g_i = \lambda \sum_{i=0}^N \mu_i g_i \nu[i \vee n, N].$$

We obtain formula (2.11) and then (2.12). If $g_\infty > 0$, then by condition (1.2), the left-hand side of (2.11) is bounded below by

$$g_\infty \sum_{k=n}^{\infty} \frac{1}{\mu_k b_k} \sum_{i=0}^k \mu_i = \infty \quad (2.13)$$

which is a contradiction since the right-hand side of (2.11) is bounded from the above by $g_0/\lambda < \infty$. Therefore, we must have $g_\infty = 0$.

With some additional work, condition (1.2) for $g_\infty = 0$ will be removed (see Proposition 2.5 below).

(c) We now prove the last assertion of the proposition. When $N < \infty$, it is well known that $\lambda_0 > 0$. Now, let $\lambda_0 = 0$ and then $N = \infty$. By (2.6) with $c_i \equiv 0$, we have

$$g_{i+1} - g_i = \frac{a_i}{b_i} (g_i - g_{i-1}), \quad i \geq 0.$$

From this and induction, it follows that $g_n = g_0$ for all $n \geq 1$ since $a_0 = 0$. $\square$

We remark that for finite state space with absorbing at $N+1 < \infty$, Proposition 2.2 was actually proved in [4; proof d) of Theorem 3.4] with a change of the order of the state space. Next, when $N = \infty$ and $\lambda_0 > 0$, in contrast with the ergodic case where $g \in L^1(\mu)$ (cf. [12; Proposition 3.5]), here one may have $g \notin L^2(\mu)$ and then $g \notin L^1(\mu)$. However, $g \in L^1(\nu)$ since g_n is strictly decreasing and $\sum_n \nu_n < \infty$, which is a consequence of Theorem 3.1 below.

Corollary 2.3. Let $\lambda_0 > 0$. Then $\lim_{i \rightarrow \infty} P_t f(i) = 0$ for all $t \geq 0$ and $f \in \mathcal{H}$.

Proof. It suffices to show that $\lim_{i \rightarrow \infty} \sum_{k=1}^n p_{ik}(t) = 0$. We now prove a stronger conclusion: $\lim_{i \rightarrow \infty} P_t g(i) = 0$ for all $t \geq 0$, where $g > 0$ with $g_0 = 1$ is the eigenfunction of λ_0 . Since g is bounded, by using the well-known fact that

$$e^{-\lambda_0 t} g_i = P_t g(i), \quad t \geq 0,$$

the conclusion now follows from Propositions 2.2 and 2.5 (2) below. $\square$

For a specialist who does not want to know many details, at the first reading, one may have a glance at the remainder of this section and the next section, especially Proposition 2.7, and then go to Section 4 directly. From here to the end of the next section, we are dealing with a case which is a dual of the one studied in Section 4. However, for the reader who is unfamiliar with this topic, it is better just to follow the context since we present everything in detail in these two sections. A large part of the details in Sections 4 and 6 are omitted since they are supposed to be known.

To state the main results of this section, we need some notation. First, we define two operators as follows.

$$I_i(f) = \frac{1}{\mu_i b_i (f_i - f_{i+1})} \sum_{j \leq i} \mu_j f_j, \quad II_i(f) = \frac{1}{f_i} \sum_{j=i}^N \frac{1}{\mu_j b_j} \sum_{k \leq j} \mu_k f_k. \quad (2.14)$$

They are called an operator of single sum (integral) or double sum, respectively. Here for the first operator, we use a convention: $f_{N+1} = 0$ if $N < \infty$. The second operator can be alternatively expressed as

$$II_i(f) = \frac{1}{f_i} \sum_{k \in E} \mu_k f_k \nu[i \vee k, N], \quad \nu[\ell, m] = \sum_{i=\ell}^m \nu_i, \quad \nu_i = \frac{1}{\mu_i b_i}. \quad (2.15)$$

Next, define a difference operator R as follows.

$$R_i(v) = a_i(1 - v_{i-1}^{-1}) + b_i(1 - v_i), \quad i \in E, \quad v_{-1} > 0 \text{ is free, } v_N := 0 \text{ if } N < \infty. \quad (2.16)$$

The domain of the operators II , I and R are defined, respectively, as follows.

$$\mathcal{F}_{II} = \{f : f > 0 \text{ on } E\},$$

$$\mathcal{F}_I = \{f : f > 0 \text{ on } E \text{ and is strictly decreasing}\},$$

$$\mathcal{V}_1 = \{v : \text{for all } i (0 \leq i < N), v_i \in (0, 1) \text{ if } \sum_j \nu_j < \infty \text{ and } v_i \in (0, 1] \text{ if } \sum_j \nu_j = \infty\}.$$

These sets are used for the lower estimates. For the upper estimates, we need some modifications of them as follows.

$$\widetilde{\mathcal{F}}_{II} = \{f : f > 0 \text{ up to some } m : 1 \leq m < N + 1 \text{ and then vanishes}\},$$

$$\widetilde{\mathcal{F}}_I = \{f : f \text{ is strictly decreasing on some interval } [n, m] (0 \leq n < m < N + 1),$$

$$f_i = f_n \text{ for } i \leq n, f_m > 0, \text{ and } f_i = 0 \text{ for } i > m\},$$

$$\widetilde{\mathcal{V}}_1 = \cup_{m=1}^{N-1} \{v : a_{i+1}(a_{i+1} + b_{i+1})^{-1} < v_i < 1 - a_i(v_{i-1}^{-1} - 1)b_i^{-1} \\ \text{for } i = 0, 1, \dots, m-1 \text{ and } v_i = 0 \text{ for } i \geq m\}.$$

Here and in what follows, to use the above operators on these modified sets, we adopt the usual convention $1/0 = \infty$. Besides, the operator II should be generalized as follows:

$$II_i(f) = \frac{1}{f_i} \sum_{i \leq j \in \text{supp}(f)} \frac{1}{\mu_j b_j} \sum_{k \leq j} \mu_k f_k, \quad i \in \text{supp}(f). \quad (2.17)$$

From now on, we should remember that $II_\bullet(f)$ is defined on $\text{supp}(f)$ only. Fortunately, we need only to consider the following two cases: either $\text{supp}(f) = \{0, 1, \dots, m\}$ for a finite m or $\text{supp}(f) = E$.

To avoid the heavy notation, we now split our main result of this section into a theorem and a proposition below.

Theorem 2.4. The following variational formulas hold for λ_0 defined by (2.2).

(1) Difference form:

$$\inf_{v \in \mathcal{V}_1} \sup_{i \in E} R_i(v) = \lambda_0 = \sup_{v \in \mathcal{V}_1} \inf_{i \in E} R_i(v).$$

(2) Single summation form:

$$\inf_{f \in \tilde{\mathcal{F}}_I} \sup_{i \in E} I_i(f)^{-1} = \lambda_0 = \sup_{f \in \tilde{\mathcal{F}}_I} \inf_{i \in E} I_i(f)^{-1}.$$

(3) Double summation form:

$$\inf_{f \in \tilde{\mathcal{F}}_{II}} \sup_{i \in \text{supp}(f)} II_i(f)^{-1} = \lambda_0 = \sup_{f \in \tilde{\mathcal{F}}_{II}} \inf_{i \in E} II_i(f)^{-1}.$$

Moreover, the supremum on the right-hand side of the above three formulas can be attained.

The next result extends the domain of λ_0 or adds some additional sets of test functions for the operators I and II , respectively. Roughly speaking, a larger set of test functions provides more freedom in practice and a smaller one is helpful for producing a better estimate.

Proposition 2.5.

(1) We have

$$\lambda_0 = \inf\{D(f) : \|f\| = 1, f_{N+1} = 0\}, \quad (2.18)$$

where $f_\infty := \lim_{N \rightarrow \infty} f_N$ in the case of $N = \infty$.

(2) When $\lambda_0 > 0$, the eigenfunction g satisfies $g_{N+1} = 0$.

(3) Moreover, we have

$$\lambda_0 = \inf_{f \in \tilde{\mathcal{F}}'_I} \sup_{i \in E} I_i(f)^{-1} \quad (2.19)$$

$$= \inf_{f \in \tilde{\mathcal{F}}_{II} \cup \tilde{\mathcal{F}}'_{II}} \sup_{i \in \text{supp}(f)} II_i(f)^{-1}, \quad (2.20)$$

$$\inf_{f \in \tilde{\mathcal{F}}_I} \sup_{i \in \text{supp}(f)} II_i(f)^{-1} = \lambda_0 = \sup_{f \in \tilde{\mathcal{F}}_I} \inf_{i \in E} II_i(f)^{-1}, \quad (2.21)$$

where

$$\begin{aligned}\widetilde{\mathcal{F}}'_I &= \{f : f \text{ is strictly decreasing and positive up to some } m : 1 \leq m < N+1 \\ &\quad \text{and then vanishes}\} \subset \widetilde{\mathcal{F}}_I, \\ \widetilde{\mathcal{F}}'_II &= \{f : f > 0 \text{ on } E \text{ and } fII(f) \in L^2(\mu)\}.\end{aligned}$$

Besides, the supremum $\sup_{f \in \mathcal{F}_I}$ in (2.21) can also be attained.

The condition “ $f_{N+1} = 0$ ” in (2.18) explains the meaning of “absorbing (Dirichlet) boundary at infinity” used in the title of this and the next sections.

Among the different groups of variational forms, the difference form is the simplest one in the practical computations. For instance, when $N = \infty$, by choosing $v_i \equiv c < 1$, we obtain the following simple lower estimate:

$$\lambda_0 \geq \inf_{i \in E} [b_i(1 - c) - a_i(c^{-1} - 1)].$$

This is non-trivial and is indeed sharp for a linear model (Example 3.5, $c = 1/2$). The difference form of the variational formulas will be used in Section 5 to deduce a dual representation of λ_0 . In general, the estimates produced by the operator R can be improved by using the operator I and further improved by using II . The price is that more computation is required successively. The single summation form of the variational formulas enables us to deduce a criterion for $\lambda_0 > 0$ (Theorem 3.1). Whereas the double summation form of the variational formulas enables us to deduce an approximating procedure to improve step by step the lower and upper estimates of λ_0 (Theorem 3.2).

Next, we mention that when $N = \infty$, for the upper estimates (the left-hand side of the formulas given in Theorem 2.4 or the formula given in (2.20)), the truncating procedure or the condition “ $fII(f) \in L^2(\mu)$ ” cannot be removed. For instance, the formally dual formula $\inf_{0 < v \leq 1} \sup_{i \in E} R_i(v)$ of the lower estimate $\sup_{0 < v \leq 1} \inf_{i \in E} R_i(v)$ [= $\sup_{v \in \mathcal{V}_1} \inf_{i \in E} R_i(v)$] is not an upper bound of λ_0 , and is indeed trivial. To see this, simply take $\bar{v}_i \equiv 1$ ($i < \infty$). Then $R_i(\bar{v}) \equiv 0$ and so

$$\inf_{0 < v \leq 1} \sup_{i \in E} R_i(v) \leq \sup_{i \in E} R_i(\bar{v}) = 0.$$

More concretely, take $b_i \equiv 2$ and $a_i \equiv 1$. Then for $\bar{v}_i \equiv c < 1$, we have

$$\inf_{v \in \mathcal{V}_1} \sup_{i \in E} R_i(v) \leq \inf_{c < 1} \sup_{i \in E} R_i(\bar{v}) \leq 2 \inf_{c < 1} (1 - c) = 0,$$

but $\lambda_0 = (\sqrt{2} - 1)^2$ as will be seen in the next section (Example 3.4). Therefore, the quantity $\inf_{0 < v \leq 1} \sup_{i \in E} R_i(v)$, as well as $\inf_{v > 0} \sup_{i \in E} R_i(v)$, has no use for an upper estimate of λ_0 .

Proofs of Theorem 2.4 and Proposition 2.5.

Part I. Recall that $\lambda_0^{(\#)}$ denotes the one defined by the formula (#). In particular, the notation λ_0 used from now on in this section is $\lambda_0^{(2.2)}$.

To prove the lower estimates, we adopt the following circle argument:

$$\begin{aligned}\lambda_0 &\geq \lambda_0^{(2.18)} \geq \sup_{f \in \mathcal{F}_I} \inf_{i \in E} II_i(f)^{-1} = \sup_{f \in \mathcal{F}_I} \inf_{i \in E} II_i(f)^{-1} = \sup_{f \in \mathcal{F}_I} \inf_{i \in E} I_i(f)^{-1} \\ &\geq \sup_{v \in \mathcal{V}_1} \inf_{i \in E} R_i(v) \geq \lambda_0.\end{aligned}\quad (2.22)$$

Clearly, $\lambda_0^{(2.18)} = \lambda_0$ if $N < \infty$. However, the identity is not trivial in the case of $N = \infty$. Besides, we will show that each supremum in (2.22) can be attained; and furthermore the eigenfunction g satisfies $g_{N+1} = 0$ whenever $\lambda_0 > 0$.

(a) Prove that $\lambda_0 \geq \lambda_0^{(2.18)} \geq \sup_{f \in \mathcal{F}_I} \inf_{i \in E} II_i(f)^{-1}$.

When $N = \infty$, the first inequality is trivial since

$$\{\|f\| = 1, f \in \mathcal{K}\} \subset \{\|f\| = 1, f_\infty = 0\}.$$

The proof of the second inequality is parallel to the first part of the proof of [4; Theorem 2.1]. Let g satisfy $g_{N+1} = 0$ and $\|g\| = 1$, and let (h_i) be a positive sequence. Then by a good use of the Cauchy-Schwarz inequality, we obtain

$$\begin{aligned}1 &= \sum_i \mu_i g_i^2 \quad (\text{since } \|g\| = 1) \\ &= \sum_i \mu_i \left(\sum_{j=i}^N (g_j - g_{j+1}) \right)^2 \quad (\text{since } g_{N+1} = 0) \\ &\leq \sum_i \mu_i \sum_{j=i}^N \frac{(g_{j+1} - g_j)^2 \mu_j b_j}{h_j} \sum_{k=i}^N \frac{h_k}{\mu_k b_k}.\end{aligned}$$

Exchanging the order of the first two sums on the right-hand side, we get

$$\begin{aligned}1 &\leq \sum_j \mu_j b_j (g_{j+1} - g_j)^2 \frac{1}{h_j} \sum_{i \leq j} \mu_i \sum_{k=i}^N \frac{h_k}{\mu_k b_k} \\ &\leq D(g) \sup_{j \in E} \frac{1}{h_j} \sum_{i \leq j} \mu_i \sum_{k=i}^N \frac{h_k}{\mu_k b_k} \\ &=: D(g) \sup_{j \in E} H_j.\end{aligned}$$

We mention that the right-hand side may be infinite but we do not care at the moment. Now, let $f \in \mathcal{F}_I$ satisfy $c := \sup_{j \in E} II_j(f) < \infty$ and take $h_j = \sum_{i \leq j} \mu_i f_i$. Then $h_j \leq c f_j / v_j < \infty$ for all j . By the proportional property, we have

$$\sup_{j \in E} H_j \leq \sup_{j \in E} \frac{1}{f_j} \sum_{k=j}^N \frac{h_k}{\mu_k b_k} = \sup_{j \in E} \frac{1}{f_j} \sum_{k=j}^N \frac{1}{\mu_k b_k} \sum_{i \leq k} \mu_i f_i = \sup_{j \in E} II_j(f) < \infty.$$

Combining these facts together, we obtain $\lambda_0^{(2.18)} \geq \inf_{j \geq 0} II_j(f)^{-1}$ whenever $\sup_{j \in E} II_j(f) < \infty$. The inequality is trivial if $\sup_{j \in E} II_j(f) = \infty$ and so it holds for all $f \in \mathcal{F}_H$. By making the supremum with respect to $f \in \mathcal{F}_H$, we obtain the required assertion.

(b) Prove that $\sup_{f \in \mathcal{F}_H} \inf_{i \in E} II_i(f)^{-1} = \sup_{f \in \mathcal{F}_I} \inf_{i \in E} II_i(f)^{-1} = \sup_{f \in \mathcal{F}_I} \inf_{i \in E} I_i(f)^{-1}$.

Let $f \in \mathcal{F}_I \subset \mathcal{F}_H$. Without loss of generality, assume that $\sup_{i \in E} I_i(f) < \infty$. By using the proportional property, we obtain

$$\begin{aligned} \sup_{i \in E} II_i(f) &= \sup_{i \in E} \frac{1}{f_i} \sum_{j=i}^N \frac{1}{\mu_j b_j} \sum_{k \leq j} \mu_k f_k \\ &\leq \sup_{i \in E} \sum_{j=i}^N \frac{1}{\mu_j b_j} \sum_{k \leq j} \mu_k f_k \Big/ \sum_{j=i}^N (f_j - f_{j+1}) \quad (\text{since } f_{N+1} \geq 0) \\ &\leq \sup_{i \in E} \frac{1}{f_i - f_{i+1}} \left(\frac{1}{\mu_i b_i} \sum_{k \leq i} \mu_k f_k \right) \quad (\text{note that } f_i > f_{i+1}) \\ &= \sup_{i \in E} I_i(f) < \infty. \end{aligned} \tag{2.23}$$

Making the infimum with respect to $f \in \mathcal{F}_I$, we get

$$\inf_{f \in \mathcal{F}_I} \sup_{i \in E} II_i(f) \leq \inf_{f \in \mathcal{F}_I} \sup_{i \in E} I_i(f).$$

Since $\mathcal{F}_I \subset \mathcal{F}_H$, the left-hand side is bounded below by $\inf_{f \in \mathcal{F}_H} \sup_{i \in E} II_i(f)$. We have thus proved that

$$\sup_{f \in \mathcal{F}_H} \inf_{i \in E} II_i(f)^{-1} \geq \sup_{f \in \mathcal{F}_I} \inf_{i \in E} II_i(f)^{-1} \geq \sup_{f \in \mathcal{F}_I} \inf_{i \in E} I_i(f)^{-1}.$$

There are two ways to prove the inverse inequality. The first one is longer but contains a useful technique. Let $f \in \mathcal{F}_H$ with $c := \sup_{i \in E} II_i(f) < \infty$. Set

$$g_i = \sum_{j=i}^N \frac{1}{\mu_j b_j} \sum_{k \leq j} \mu_k f_k = \sum_{j=i}^N \nu_j \sum_{k \leq j} \mu_k f_k > 0, \quad i \in E, \quad g_{N+1} := 0 \text{ if } N < \infty.$$

Then g_i is strictly decreasing in i , $g_i < g_0 \leq cf_0 < \infty$ for all i . Hence, $g \in \mathcal{F}_I$. Noticing that

$$g_i - g_{i+1} = \sum_{j=i}^N \nu_j \sum_{k \leq j} \mu_k f_k - \sum_{j=i+1}^N \nu_j \sum_{k \leq j} \mu_k f_k = \nu_i \sum_{k \leq i} \mu_k f_k$$

(here and in what follows, $\sum_{k=i}^j$ means $\sum_{i \leq k < j+1}$ and $\sum_{\emptyset} = 0$ by the standard convention), we have

$$\begin{aligned} \Omega g(i) &= b_i(g_{i+1} - g_i) + a_i(g_{i-1} - g_i) \\ &= -b_i \nu_i \sum_{k \leq i} \mu_k f_k + a_i \nu_{i-1} \sum_{k \leq i-1} \mu_k f_k \\ &= -\frac{1}{\mu_i} \sum_{k \leq i} \mu_k f_k + \frac{a_i}{\mu_{i-1} b_{i-1}} \sum_{k \leq i-1} \mu_k f_k \\ &= -f_i, \quad 1 \leq i < N. \end{aligned}$$

Actually, this holds also for $i = 0$ and $i = N$ if $N < \infty$. Applying (2.7) to $h = f$, by (2.9), it follows that

$$\mu_k b_k(g_k - g_{k+1}) = \sum_{j \leq k} \mu_j g_j f_j / g_j \geq \sum_{j \leq k} \mu_j g_j \inf_{i \in E} II_i(f)^{-1}, \quad k \in E.$$

That is,

$$\sup_{i \in E} II_i(f) \geq \frac{1}{\mu_k b_k(g_k - g_{k+1})} \sum_{j \leq k} \mu_j g_j = I_k(g), \quad k \in E.$$

Making the supremum with respect to k , we obtain

$$\inf_{k \in E} I_k(g)^{-1} \geq \inf_{i \in E} II_i(f)^{-1},$$

and hence,

$$\sup_{g \in \mathcal{F}_I} \inf_{k \in E} I_k(g)^{-1} \geq \inf_{i \in E} II_i(f)^{-1}.$$

This lower bound becomes trivial if $\sup_{i \in E} II_i(f) = \infty$, and hence, the inequality holds for all $f \in \mathcal{F}_{II}$. Making the supremum with respect to $f \in \mathcal{F}_{II}$, we obtain

$$\sup_{g \in \mathcal{F}_I} \inf_{k \in E} I_k(g)^{-1} \geq \sup_{f \in \mathcal{F}_{II}} \inf_{i \in E} II_i(f)^{-1}.$$

We have thus proved the required assertion.

The second proof is to show that

$$\sup_{f \in \mathcal{F}_I} \inf_{i \in E} I_i(f)^{-1} \geq \lambda_0$$

and thus completes a smaller circle argument. To do so, without loss of generality, assume that $\lambda_0 > 0$. Let $g > 0$ be the eigenfunction of λ_0 . Applying (2.9) to $h = \lambda_0 g$, we obtain $I_i(g) = \lambda_0^{-1}$ for all $i \in E$, and hence, $\inf_{i \in E} I_i(g)^{-1} = \lambda_0$. Noticing that $g \in \mathcal{F}_I$ by Proposition 2.2, the assertion is now obvious.

(c) Prove that $\sup_{f \in \mathcal{F}_{II}} \inf_{i \in E} II_i(f)^{-1} \geq \sup_{v \in \mathcal{V}_1} \inf_{i \in E} R_i(v)$.

Note that by a change of the sequence $\{v_i\}_{i=0}^{N-1}$:

$$u_i = v_0 v_1 \cdots v_{i-1}, \quad i \in E, \quad v_{-1} > 0 \text{ is free, } v_N := 0 \text{ if } N < \infty,$$

the quantity $R_i(v)$ becomes

$$a_i \left(1 - \frac{u_{i-1}}{u_i} \right) + b_i \left(1 - \frac{u_{i+1}}{u_i} \right), \quad i \in E, \quad u_{-1} > 0 \text{ is free, } u_{N+1} := 0 \text{ if } N < \infty.$$

To save our notation, we use $R_i(u)$ to denote this quantity. Clearly, $\{u_i\}$ is positive and $v_i \leq 1$ for all i mean that $\{u_i\}$ is non-increasing.

Before moving further, we prove that if $\inf_{i \in E} R_i(u) > 0$ for a positive sequence $u = (u_i)$, then u_i must be strictly decreasing in i . To do so, let

$$f_i = (a_i + b_i)u_i - a_i u_{i-1} - b_i u_{i+1}.$$

Then $f_i = u_i R_i(u) > 0$ for all $i \in E$ by assumption, and so $f \in \mathcal{F}_H$. Noticing that

$$\mu_k f_k = \mu_{k+1} a_{k+1} (u_k - u_{k+1}) - \mu_k a_k (u_{k-1} - u_k),$$

we obtain

$$0 < \sum_{k \leq j} \mu_k f_k = \mu_{j+1} a_{j+1} (u_j - u_{j+1}) = \mu_j b_j (u_j - u_{j+1}).$$

Hence, u_i is strictly decreasing in i (equivalently, $v_i := u_{i+1}/u_i < 1$). This proves the required assertion. The reason of using $\mathcal{V}_1$ rather than $\{v : v_i > 0, 0 \leq i < N\}$ should be clear now.

We now return to our main assertion. For this, without loss of generality, assume that $\inf_{i \in E} R_i(u) > 0$ for a given strictly decreasing $u = (u_i)$. Otherwise, the assertion is trivial. From the last formula, we obtain

$$0 < \sum_{j=i}^N \nu_j \sum_{k \leq j} \mu_k f_k = \sum_{j=i}^N (u_j - u_{j+1}) = u_i - u_{N+1} \leq u_i.$$

Therefore,

$$0 < R_i(u) = \frac{f_i}{u_i} \leq f_i \left(\sum_{j=i}^N \nu_j \sum_{k \leq j} \mu_k f_k \right)^{-1} = H_i(f)^{-1}, \quad i \in E.$$

It follows that

$$\inf_{i \in E} R_i(u) \leq \inf_{i \in E} H_i(f)^{-1} \leq \sup_{f \in \mathcal{F}_H} \inf_{i \in E} H_i(f)^{-1}.$$

The assertion now follows by making the supremum with respect to u .

(d) Prove that $\sup_{v \in \mathcal{V}_1} \inf_{i \in E} R_i(v) \geq \lambda_0$.

Assume that $\lambda_0 > 0$ for a moment (in particular, if $N < \infty$). Then by Proposition 2.2, the corresponding eigenfunction g (with $g_0 = 1$) of λ_0 is positive and strictly decreasing. From the eigenequation

$$-\Omega g(i) = \lambda_0 g_i, \quad i \in E, \quad g_{N+1} := 0 \text{ if } N < \infty,$$

it follows that

$$a_i \left(1 - \frac{g_{i-1}}{g_i} \right) + b_i \left(1 - \frac{g_{i+1}}{g_i} \right) = \lambda_0, \quad i \in E.$$

Let $v_i = g_{i+1}/g_i$. Then $v_i \in (0, 1)$ for all $i < N$ and so $v = (v_i) \in \mathcal{V}_1$. Moreover, $R_i(v) = \lambda_0$ for all $i \in E$. Therefore, we certainly have $\sup_{v \in \mathcal{V}_1} \inf_{i \in E} R_i(v) \geq \lambda_0$, as required.

It remains to prove that $\sup_{v \in \mathcal{V}_1} \inf_{i \in E} R_i(v) \geq 0$ when $N = \infty$. First, let $\sum_{i=0}^{\infty} \nu_i < \infty$. Choose a positive f such that

$$\sum_{k=0}^{\infty} \mu_k f_k \varphi_k < \infty, \quad \varphi_k := \sum_{j=k}^{\infty} \nu_j.$$

Define

$$h_i = \sum_{j=i}^{\infty} \nu_j \sum_{k \leq j} \mu_k f_k, \quad i \geq 0.$$

Then

$$h_i = \sum_{k=0}^{\infty} \mu_k f_k \varphi_{i \vee k} \leq \sum_{k=0}^{\infty} \mu_k f_k \varphi_k < \infty.$$

Set $\bar{v}_i = h_{i+1}/h_i$ ($i \geq 0$). Then $\bar{v} \in \mathcal{V}_1$ since h_i is strictly decreasing. A simple computation shows that $R_i(\bar{v}) = \Pi_i(f)^{-1} > 0$ for all $i \geq 0$. Hence, $\sup_{v \in \mathcal{V}_1} \inf_{i \geq 0} R_i(v) \geq 0$. Next, let $\sum_i \nu_i = \infty$ and set $\bar{v}_i \equiv 1$. Then $R_i(\bar{v}) \equiv 0$ and so the same conclusion holds.

The proof of the last paragraph indicates the reason why in $\mathcal{V}_1$ we define “ $v_i \in (0, 1)$ ” and “ $v_i \in (0, 1]$ ” separately according to “ $\sum_i \nu_i < \infty$ ” or “ $\sum_i \nu_i = \infty$ ”. Although we have known from proof (c) that for $\inf_i R_i(v) > 0$, it is necessary that $v < 1$ but this condition may not be sufficient for $\inf_i R_i(v) \geq 0$. The extremal $\bar{v}_i \equiv 1$ is used only in the case of $\sum_i \nu_i = \infty$ in which we indeed have $\lambda_0 = 0$ (cf. Theorem 3.1 below).

We have thus completed the proof of circle (2.22).

(e) We now prove that each supremum in (2.22) can be attained. The case that $\lambda_0 = 0$ is easier since

$$0 = \lambda_0 \geq \inf_{i \in E} \Pi_i(f)^{-1} \geq 0 \quad \text{and} \quad 0 = \lambda_0 \geq \inf_{i \in E} I_i(f)^{-1} \geq 0$$

for every f in the corresponding domain, as an application of (2.22). Similarly, the conclusion holds for the operator R as seen from proof (d): noting that in the degenerated case that $\sum_i \nu_i = \infty$, we have $\lambda_0 = 0$ and then $v_i \equiv 1$ by Proposition 2.2 (3).

Next, we consider the case that $\lambda_0 > 0$ with eigenfunction g : $g_0 = 1$. Then for the operator R , the supremum is attained at $v_i = g_{i+1}/g_i$ as seen from the first paragraph of proof (d). For the operator I , it is attained at $f = g$ as an application of Proposition 2.1 with $c_i \equiv 0$: $I_i(g) \equiv \lambda_0^{-1}$. At the same time, in view of part (2) of Proposition 2.2, we have $\Pi_i(g) \equiv \lambda_0^{-1}$ whenever $g_{N+1} = 0$.

It remains to rule out the possibility that $g_{N+1} > 0$. Otherwise, by part (2) of Proposition 2.2 again, we have $N = \infty$ and

$$M_i := \sum_{j \geq i} \nu_j \sum_{k \leq j} \mu_k \in (0, \infty).$$

Let $\tilde{g} = g - g_\infty$. Then $\tilde{g} \in \mathcal{F}_H$. Noting that

$$\begin{aligned} \sum_{j \geq i} \nu_j \sum_{k \leq j} \mu_k \tilde{g}_k &= \sum_{j \geq i} \nu_j \sum_{k \leq j} \mu_k g_k - g_\infty M_i \\ &= \frac{g_i - g_\infty}{\lambda_0} - g_\infty M_i \quad (\text{by (2.11)}), \end{aligned}$$

we obtain

$$\sup_{i \geq 0} II_i(\tilde{g}) = \sup_{i \geq 0} \left[\frac{1}{\lambda_0} - \frac{g_\infty M_i}{g_i - g_\infty} \right] = \frac{1}{\lambda_0} - g_\infty \inf_{i \geq 0} \frac{M_i}{g_i - g_\infty}.$$

By using the proportional property and (2.5), it follows that

$$\inf_{i \geq 0} \frac{M_i}{g_i - g_\infty} \geq \inf_{i \geq 0} \frac{\nu_i \sum_{k \leq i} \mu_k}{g_i - g_{i+1}} \geq \inf_{i \geq 0} I_i(g) = \frac{1}{\lambda_0}.$$

Thus, we get

$$\sup_{i \geq 0} II_i(\tilde{g}) \leq \frac{1}{\lambda_0} (1 - g_\infty) < \frac{1}{\lambda_0}.$$

Hence, $\inf_{i \geq 0} II_i(\tilde{g})^{-1} > \lambda_0$, which is a contradiction to proof (a): $\lambda_0 \geq \inf_{i \in E} II(\tilde{g})^{-1}$.

We have thus proved that $g_\infty = 0$ whenever $\lambda_0 > 0$. Note that this paragraph uses Proposition 2.2 and proof (a) only.

Part II. Next, to prove the upper estimates, we adopt the following circle argument:

$$\lambda_0 \leq \inf_{f \in \tilde{\mathcal{F}}_H \cup \tilde{\mathcal{F}}'_H} \sup_{i \in \text{supp}(f)} II_i(f)^{-1} \quad (2.24)$$

$$\leq \inf_{f \in \tilde{\mathcal{F}}_H} \sup_{i \in \text{supp}(f)} II_i(f)^{-1} \quad (2.25)$$

$$= \inf_{f \in \tilde{\mathcal{F}}_I} \sup_{i \in \text{supp}(f)} II_i(f)^{-1} = \inf_{f \in \tilde{\mathcal{F}}_I} \sup_{i \in E} I_i(f)^{-1} \quad (2.26)$$

$$\leq \inf_{f \in \tilde{\mathcal{F}}'_I} \sup_{i \in E} I_i(f)^{-1} \quad (2.27)$$

$$\leq \inf_{v \in \tilde{\mathcal{V}}_1} \sup_{i \in E} R_i(v) \quad (2.28)$$

$$\leq \lambda_0. \quad (2.29)$$

Since inequalities (2.25) and (2.27) are obvious, we need only to prove (2.24), (2.26), (2.28) and (2.29).

(f) Prove that $\lambda_0 \leq \inf_{f \in \tilde{\mathcal{F}}_H \cup \tilde{\mathcal{F}}'_H} \sup_{i \in \text{supp}(f)} II_i(f)^{-1}$.

We remark that in the particular case that the eigenfunction f is in $L^2(\mu)$, then the function $g := fII(f)$ is nothing but just $f/\lambda_0 \in L^2(\mu)$. Hence, the infimum in (2.24) is attained at this $f \in \tilde{\mathcal{F}}'_H$ and the equality sign in (2.24) holds.

We now consider the general case. Let $f \in \widetilde{\mathcal{F}}_H$. Then there exists an m such that $f_i > 0$ for $i \leq m$ and $f_i = 0$ for $i > m$. Set $g = \mathbb{1}_{\text{supp}(f)} f H(f)$. That is,

$$g_i = \begin{cases} \sum_{j=i}^m \nu_j \sum_{k \leq j} \mu_k f_k, & i \leq m \\ 0, & i \geq m+1. \end{cases}$$

Clearly, $g \in L^2(\mu)$ and

$$g_i - g_{i+1} = \begin{cases} \nu_i \sum_{k \leq i} \mu_k f_k, & i \leq m \\ 0, & i \geq m+1. \end{cases}$$

We now have

$$D(g) = \sum_{i \leq m} \mu_i b_i (g_{i+1} - g_i)^2 = \sum_{i \leq m} (g_i - g_{i+1}) \sum_{k \leq i} \mu_k f_k = \sum_{k \leq m} \mu_k f_k \sum_{k \leq i \leq m} (g_i - g_{i+1}).$$

Since $g_{m+1} = 0$, we get

$$D(g) = \sum_{k \leq m} \mu_k f_k g_k \leq \sum_{k \leq m} \mu_k g_k^2 \max_{0 \leq i \leq m} (f_i / g_i) = \|g\|^2 \sup_{i \in \text{supp}(f)} H_i(f)^{-1}.$$

Dividing both sides by $\|g\|^2 \in (0, \infty)$, it follows that

$$\lambda_0 \leq D(g) / \|g\|^2 \leq \sup_{i \in \text{supp}(f)} H_i(f)^{-1}, \quad f \in \widetilde{\mathcal{F}}_H. \quad (2.30)$$

For $f \in \widetilde{\mathcal{F}}'_H$, the same conclusion clearly holds if $N < \infty$. When $N = \infty$, since $g \in L^2(\mu)$ by assumption, we have $0 < g < \infty$. As a tail sequence of a convergent series (which sum equals g_0), we certainly have $g_i \downarrow g_\infty = 0$ as $i \uparrow \infty$. Hence, the same proof replacing m with ∞ , plus the fact that $\lambda_0 = \lambda_0^{(2.18)}$ proved in Part I, shows that

$$\lambda_0 = \lambda_0^{(2.18)} \leq \sup_{i \in \text{supp}(f)} H_i(f)^{-1}, \quad f \in \widetilde{\mathcal{F}}'_H.$$

Combining this with (2.30), we prove the required assertion.

The proof indicates the reason why the truncating procedure is used for the upper estimates since in general the eigenfunction g may not belong to $L^2(\mu)$ as shown by Proposition 2.2.

(g) Prove that

$$\inf_{f \in \widetilde{\mathcal{F}}_H} \sup_{i \in \text{supp}(f)} H_i(f)^{-1} = \inf_{f \in \widetilde{\mathcal{F}}_I} \sup_{i \in \text{supp}(f)} H_i(f)^{-1} = \inf_{f \in \widetilde{\mathcal{F}}_I} \sup_{i \in E} I_i(f)^{-1}.$$

Let $f \in \widetilde{\mathcal{F}}_I$. Then there exist $n < m$ such that $f_i = f_{i \vee n} \mathbb{1}_{\{i \leq m\}}$, $f_m > 0$, and f is strictly decreasing on $[n, m]$. Clearly, we have

$$\min_{i \leq m} H_i(f) = \min_{n \leq i \leq m} H_i(f) \quad \text{and} \quad \inf_{i \in E} I_i(f) = \min_{n \leq i \leq m} I_i(f)$$

since, by assumption, $1/0 = \infty$. By the proportional property, first we have

$$\begin{aligned} \min_{n \leq i \leq m} II_i(f) &= \min_{n \leq i \leq m} \sum_{j=i}^m \nu_j \sum_{k \leq j} \mu_k f_k \bigg/ \sum_{j=i}^m (f_j - f_{j+1}) \\ &\geq \min_{n \leq i \leq m} \frac{1}{\mu_i b_i (f_i - f_{i+1})} \sum_{k \leq i} \mu_k f_k \\ &= \min_{n \leq i \leq m} I_i(f), \end{aligned}$$

and then

$$\sup_{f \in \widetilde{\mathcal{F}}_II} \inf_{i \in \text{supp}(f)} II_i(f) \geq \sup_{f \in \widetilde{\mathcal{F}}_I} \inf_{i \in \text{supp}(f)} II_i(f) \geq \sup_{f \in \widetilde{\mathcal{F}}_I} \inf_{i \in E} I_i(f)$$

since $\widetilde{\mathcal{F}}_I \subset \widetilde{\mathcal{F}}_{II}$.

As in proof (b), there are two ways to prove the inverse inequality. First, let $f \in \widetilde{\mathcal{F}}_{II}$. As in proof (f), set $g = \mathbb{1}_{\text{supp}(f)} f II(f)$. Clearly, $g \in \widetilde{\mathcal{F}}'_I \subset \widetilde{\mathcal{F}}_I$ and moreover,

$$\begin{aligned} b_i(g_{i+1} - g_i) + a_i(g_{i-1} - g_i) &= -\frac{1}{\mu_i} \sum_{k \leq i} \mu_k f_k + \frac{a_i}{\mu_{i-1} b_{i-1}} \sum_{k \leq i-1} \mu_k f_k \\ &= -\frac{1}{\mu_i} \sum_{k \leq i} \mu_k f_k + \frac{1}{\mu_i} \sum_{k \leq i-1} \mu_k f_k \\ &= -f_i, \quad i \leq m. \end{aligned}$$

When $i = 0$, the second term on the left-hand side disappears since $a_0 = 0$. It follows that

$$\mu_i b_i (g_{i+1} - g_i) + \mu_i a_i (g_{i-1} - g_i) = -\mu_i f_i, \quad i \leq m,$$

and furthermore,

$$\mu_k b_k (g_k - g_{k+1}) = \sum_{j \leq k} \mu_j g_j f_j / g_j \leq \sum_{j \leq k} \mu_j g_j \max_{0 \leq i \leq m} II_i(f)^{-1}, \quad k \leq m.$$

That is,

$$\min_{0 \leq i \leq m} II_i(f) \leq \frac{1}{\mu_k b_k (g_k - g_{k+1})} \sum_{j \leq k} \mu_j g_j = I_k(g), \quad k \leq m.$$

Making the infimum with respect to k , we obtain

$$\max_{0 \leq k \leq m} I_k(g)^{-1} \leq \max_{0 \leq i \leq m} II_i(f)^{-1}.$$

One may rewrite $\max_{0 \leq k \leq m}$ as $\sup_{k \in E}$ on the left-hand side since $I_k(g) = \infty$ for all $k \geq m + 1$. Since $g \in \widetilde{\mathcal{F}}'_I \subset \widetilde{\mathcal{F}}_I$, we now have

$$\inf_{g \in \widetilde{\mathcal{F}}_I} \sup_{k \in E} I_k(g)^{-1} \leq \inf_{g \in \widetilde{\mathcal{F}}'_I} \sup_{k \in E} I_k(g)^{-1} \leq \sup_{i \in \text{supp}(f)} II_i(f)^{-1}.$$

Next, making the infimum with respect to $f \in \widetilde{\mathcal{F}}_{II}$, we obtain

$$\inf_{g \in \widetilde{\mathcal{F}}_I} \sup_{k \in E} I_k(g)^{-1} \leq \inf_{g \in \widetilde{\mathcal{F}}'_I} \sup_{k \in E} I_k(g)^{-1} \leq \inf_{f \in \widetilde{\mathcal{F}}_{II}} \sup_{i \in \text{supp}(f)} II_i(f)^{-1}.$$

The second proof for the inverse inequality is to show that

$$\inf_{f \in \widetilde{\mathcal{F}}'_I} \sup_{i \in E} I_i(f)^{-1} \leq \lambda_0.$$

For this, recall the definition

$$\lambda_0 = \inf\{D(f) : \|f\| = 1, f_i = 0 \text{ for all } i > \text{some } m : 1 \leq m < N + 1\}.$$

Because of

$$\begin{aligned} & \{\|f\| = 1, f_i = 0 \text{ for all } i > m : 1 \leq m < N + 1\} \\ & \subset \{\|f\| = 1, f_i = 0 \text{ for all } i > m + 1 : 1 \leq m < N + 1\}, \end{aligned}$$

it is clear that

$$\lambda_0^{(m)} := \inf\{D(f) : \|f\| = 1, f_i = 0 \text{ for all } i > m : 1 \leq m < N + 1\} \downarrow \lambda_0$$

as $m \uparrow N$. Note that $\lambda_0^{(m)}$ is just the first eigenvalue of the Dirichlet form $(D, \mathcal{D}(D))$ restricted to $\{0, 1, \dots, m\}$ with Dirichlet (absorbing) boundary at $m + 1$. Now, let $g = g^{(m)}$ be the eigenfunction of $\lambda_0^{(m)} > 0$ with $g_0 = 1$. Extend g to the whole space by setting $g_i = 0$ for all $i > m$. By using Proposition 2.2, it follows that $g \in \widetilde{\mathcal{F}}'_I$ with $\text{supp}(g) = \{0, 1, \dots, m\}$. Furthermore, by (2.9) with $h = \lambda_0 g$, we have $I_i(g)^{-1} = \lambda_0^{(m)} > 0$ for all $i \leq m$, and hence,

$$\sup_{i \in E} I_i(g)^{-1} = \sup_{i \leq m} I_i(g)^{-1} = \lambda_0^{(m)}.$$

Thus,

$$\lambda_0^{(m)} = \sup_{i \in E} I_i(g)^{-1} \geq \inf_{f \in \widetilde{\mathcal{F}}'_I, \text{supp}(f) = \{0, 1, \dots, m\}} \sup_{i \in E} I_i(f)^{-1} \geq \inf_{f \in \widetilde{\mathcal{F}}'_I} \sup_{i \in E} I_i(f)^{-1}.$$

The assertion now follows by letting $m \rightarrow N$.

(h) Prove that $\inf_{f \in \widetilde{\mathcal{F}}_{II}} \sup_{i \in \text{supp}(f)} II_i(f)^{-1} \leq \inf_{v \in \mathcal{V}_1} \sup_{i \in E} R_i(v)$.

Let u with $\text{supp}(u) = \{0, 1, \dots, m\}$ be given such that $v_i := u_{i+1}/u_i \in \widetilde{\mathcal{V}}_1$. Then, the constraint

$$v_i < 1 - a_i(v_{i-1}^{-1} - 1)b_i^{-1}, \quad 0 \leq i \leq m, \quad v_m = 0,$$

is equivalent to $\min_{0 \leq i \leq m} R_i(v) > 0$, and the constraint

$$v_i > a_{i+1}(a_{i+1} + b_{i+1})^{-1}, \quad 0 \leq i \leq m-1,$$

comes from the requirement that $v_i > 0$ for all $i < m$. Since the case of $i = m$ in the first constraint is contained in the second one, we obtain the constraint described in $\widetilde{\mathcal{V}}_1$. In particular, we have

$$a_1(a_1 + b_1)^{-1} < v_0 < 1 - a_0(v_{-1}^{-1} - 1) = 1$$

and so $v_0 \in (0, 1)$. By induction, we have $v_i \in (0, 1)$ for all $i < m$. The existence of such a u is guaranteed since $m < \infty$, as will be shown in proof (i) below. Now, let

$$f_i = \begin{cases} (a_i + b_i)u_i - a_i u_{i-1} - b_i u_{i+1}, & i \leq m, \\ 0, & i > m. \end{cases}$$

Then by assumption, $f_i/u_i = R_i(u) > 0$ for $i \leq m$. Hence, $f \in \widetilde{\mathcal{F}}_H$. Next, we have

$$0 < \sum_{k \leq j} \mu_k f_k = \mu_j b_j (u_j - u_{j+1}), \quad j \leq m.$$

Hence,

$$\sum_{j=i}^m \nu_j \sum_{k \leq j} \mu_k f_k = u_i - u_{m+1} = u_i > 0, \quad i \leq m.$$

Therefore, we obtain

$$R_i(u) = \frac{f_i}{u_i} = f_i \left/ \sum_{j=i}^m \nu_j \sum_{k \leq j} \mu_k f_k \right. = \Pi_i(f)^{-1}, \quad i \leq m$$

and then

$$\sup_{i \in E} R_i(u) = \max_{i \leq m} R_i(u) = \sup_{i \in \text{supp}(f)} \Pi_i(f)^{-1} \geq \inf_{f \in \widetilde{\mathcal{F}}_H} \sup_{i \in \text{supp}(f)} \Pi_i(f)^{-1}.$$

To be consistent with the convention of $R_i(v)$, here we adopt the convention: $R_i(u) = -\infty$ for all $i > m$. The assertion now follows by making the infimum with respect to u .

(i) Prove that $\inf_{v \in \widetilde{\mathcal{V}}_1} \sup_{i \in E} R_i(v) \leq \lambda_0$.

As in the last part of proof (g), denote by g (with $g_0 = 1$) the eigenfunction of $\lambda_0^{(m)} > 0$. Then $\text{supp}(g) = \{0, 1, \dots, m\}$, and g is strictly decreasing on $\{0, 1, \dots, m\}$ by part (2) of Proposition 2.2. The definition of g gives us

$$b_i(g_i - g_{i+1}) - a_i(g_{i-1} - g_i) = \lambda_0^{(m)} g_i, \quad i \leq m, \quad g_{m+1} = 0.$$

That is,

$$a_i \left(1 - \frac{g_{i-1}}{g_i} \right) + b_i \left(1 - \frac{g_{i+1}}{g_i} \right) = \lambda_0^{(m)}, \quad i \leq m.$$

Let $v_i = g_{i+1}/g_i$ for $i \leq m$ and $v_i = 0$ for $i > m$. Then $v_i \in (0, 1)$ for $i \in \{0, 1, \dots, m-1\}$, and $R_i(v) = \lambda_0^{(m)}$ for all $i \leq m$. It is now easy to see that $v \in \tilde{\mathcal{V}}_1$. We have thus constructed a $u (= g)$ required in proof (h). Clearly $R_i(v) = -\infty$ for all $i > m$. Therefore,

$$\begin{aligned} \lambda_0^{(m)} &= \max_{0 \leq i \leq m} R_i(v) \\ &\geq \inf_{v \in \tilde{\mathcal{V}}_1: \text{supp}(v) = \{0, 1, \dots, m-1\}} \max_{0 \leq i \leq m} R_i(v) \\ &\geq \inf_{v \in \tilde{\mathcal{V}}_1: \text{supp}(u) = \{0, 1, \dots, n\} \text{ for some } n \geq 0} \sup_{i \in E} R_i(u) \\ &= \inf_{v \in \tilde{\mathcal{V}}_1} \sup_{i \in E} R_i(v). \end{aligned}$$

Letting $m \rightarrow N$, we obtain the required assertion.

We have thus completed the circle argument of (2.24)–(2.29) and then the proofs of Theorem 2.4 and Proposition 2.5 are finished. $\square$

Before moving further, we mention a technical point in the proof above. Instead of the approximation with finite state space used in Part II of the above proof, it seems more natural to use the truncating procedure for the eigenfunction g . However, the next result shows that this procedure is not practical in general.

Remark 2.6. Let $g \neq 0$ be the eigenfunction of $\lambda_0 > 0$ and define $g^{(m)} = g \mathbb{1}_{\leq m}$. Then

$$\min_{i \in \text{supp}(g^{(m)})} H_i(g^{(m)}) = \frac{1}{\lambda_0} \left[1 - \frac{g_{m+1}}{g_m} \right].$$

In particular, the sequence $\left\{ \min_{i \in \text{supp}(g^{(m)})} H_i(g^{(m)}) \right\}_{m \geq 1}$ may not converge to λ_0^{-1} as $m \uparrow \infty$.

Proof. Note that

$$\begin{aligned} \min_{i \in \text{supp}(g^{(m)})} H_i(g^{(m)}) &= \min_{0 \leq i \leq m} \frac{1}{g_i} \sum_{j=i}^m \nu_j \sum_{k \leq j} \mu_k g_k \\ &= \min_{0 \leq i \leq m} \frac{1}{g_i} \sum_{j=i}^m \nu_j \frac{\mu_j b_j (g_j - g_{j+1})}{\lambda_0} \quad (\text{by (2.5)}) \\ &= \min_{0 \leq i \leq m} \frac{1}{\lambda_0 g_i} (g_i - g_{m+1}) \\ &= \frac{1}{\lambda_0} \left[1 - \frac{g_{m+1}}{g_m} \right]. \end{aligned}$$

This proves the main assertion. For Example 3.4 in the next section, we have

$$\lim_{m \rightarrow \infty} \left(1 - \frac{g_{m+1}}{g_m} \right) = 1 - \sqrt{\frac{a}{b}} < 1,$$

and so

$$\lim_{m \rightarrow \infty} \min_{i \in \text{supp}(g^{(m)})} II_i(g^{(m)}) < \lambda_0^{-1}. \quad \square$$

To conclude this section and also for later use, we introduce a variational formula of λ_0 in a different difference form.

Proposition 2.7. On the set $\mathcal{V} := \{v : v_i > 0, 0 \leq i < N\}$, redefine

$$R_i(v) = a_{i+1} + b_i - a_i/v_{i-1} - b_{i+1}v_i, \quad i \in E, \quad v_{-1} > 0 \text{ is free,}$$

where $a_{N+1} = b_{N+1} = 0$ and v_N is free if $N < \infty$. Then

(1) we have

$$\sup_{v \in \mathcal{V}} \inf_{i \in E} R_i(v) \geq \lambda_0. \quad (2.31)$$

The equality sign holds once $\sum_{i=0}^N \mu_i = \infty$. In this case, we indeed have

$$\lambda_0 = \sup_{v \in \mathcal{V}} \inf_{i \in E} R_i(v) = \sup_{v \in \mathcal{V}_*} \inf_{i \in E} R_i(v),$$

where

$$\mathcal{V}_* = \{v : v_{i-1} > a_i/b_i, 0 \leq i < N+1\}.$$

(2) In general, we have

$$\lambda_0 = \sup_{v \in \mathcal{V}_*} \inf_{i \in E} R_i(v), \quad (2.32)$$

and the supremum in (2.32) can be attained.

Proof. (a) First, we prove that $\sup_{v \in \mathcal{V}_*} \inf_{i \in E} R_i(v) \geq 0$. Given a positive, non-increasing f , $f_{N+1} = 0$ if $N < \infty$, define

$$u_i = (\mu_i b_i)^{-1} \sum_{j \leq i} \mu_j f_j \in (0, \infty), \quad i < N+1.$$

Then

$$b_i u_i - a_i u_{i-1} = f_i > 0, \quad i \in E, \quad u_{-1} > 0 \text{ is free.}$$

This implies that $(v_i := u_{i+1}/u_i : i < N) \in \mathcal{V}_*$. As before, we also use

$$R_i(u) := a_{i+1} + b_i - a_i \frac{u_{i-1}}{u_i} - b_{i+1} \frac{u_{i+1}}{u_i}, \quad i \in E$$

instead of $R_i(v)$. Clearly,

$$R_i(u) = \frac{f_i - f_{i+1}}{u_i} \geq 0, \quad i \in E.$$

Hence $\inf_{i \in E} R_i(u) \geq 0$ and the required assertion is now obvious.

(b) By (a), without loss of generality, assume that $\lambda_0 > 0$. Then by Proposition 2.2, the corresponding eigenfunction g of λ_0 is positive and strictly decreasing. With $u_i := g_i - g_{i+1} > 0$ ($i \in E$), the eigenequation

$$-\Omega g(i) = b_i u_i - a_i u_{i-1} = \lambda_0 g_i, \quad i \in E, \quad g_{N+1} := 0 \text{ if } N < \infty,$$

gives us $(v_i := u_{i+1}/u_i : i < N) \in \mathcal{V}_*$. Next, by making a difference of $-\Omega g(i)$ and $-\Omega g(i+1)$ and noting that $\Omega g(N+1)$ is setting to be zero if $N < \infty$, we obtain

$$(a_{i+1} + b_i)u_i - b_{i+1}u_{i+1} - a_i u_{i-1} = \lambda_0 u_i, \quad i \in E.$$

Thus, we have $R_i(v) = \lambda_0$ for all $i \in E$. Therefore, (2.31) holds.

(c) To prove the equality sign in (2.31) whenever $\sum_{i=0}^N \mu_i = \infty$, in view of Part I of the proofs of Theorem 2.4 and Proposition 2.5 and (b), it suffices to show that

$$\sup_{f \in \mathcal{F}_I} \inf_{i \in E} I_i(f)^{-1} \geq \sup_{v \in \mathcal{V}} \inf_{i \in E} R_i(v).$$

In view of (a), without loss of generality, assume that $\inf_{i \in E} R_i(u) > 0$ for a given $u > 0$. Define $f_i = b_i u_i - a_i u_{i-1}$ for $i \in E$, $f_{N+1} = 0$ if $N < \infty$. Then it is clear that

$$(f_i - f_{i+1})/u_i = R_i(u) > 0, \quad i \in E. \quad (2.33)$$

Hence, f is strictly decreasing.

We now prove that $f \in \mathcal{F}_I$ whenever $\sum_i \mu_i = \infty$. First, we have

$$\sum_{k \leq i} \mu_k f_k = \sum_{k \leq i} \mu_k (b_k u_k - a_k u_{k-1}) = \sum_{k \leq i} (\mu_k b_k u_k - \mu_{k-1} b_{k-1} u_{k-1}) = \mu_i b_i u_i > 0, \quad i \in E. \quad (2.34)$$

In particular, $f_0 > 0$. If $f_{k_0} \leq 0$ for some $k_0 \geq 1$, then $f_{k_0+1} < 0$ and

$$\sum_{k_0+1 \leq i \leq n} \mu_i f_i < f_{k_0+1} \sum_{k_0+1 \leq i \leq n} \mu_i \rightarrow -\infty \quad \text{as } n \rightarrow \infty$$

since $\sum_{i=0}^N \mu_i = \infty$. This implies that

$$\sum_{k_0+1 \leq i \leq n} \mu_i f_i \rightarrow -\infty \quad \text{as } n \rightarrow \infty.$$

Now, by (2.33), we would get

$$0 < \mu_n b_n u_n = \sum_{i \leq n} \mu_i f_i = \sum_{i \leq k_0} \mu_i f_i + \sum_{k_0+1 \leq i \leq n} \mu_i f_i \rightarrow -\infty \quad \text{as } n \rightarrow \infty,$$

which is impossible. Therefore, $f > 0$ and then $f \in \mathcal{F}_I$.

Combining (2.33) with (2.34), we obtain that

$$R_i(u) = \frac{f_i - f_{i+1}}{u_i} = \mu_i b_i (f_i - f_{i+1}) / \sum_{k \leq i} \mu_k f_k = I_i(f)^{-1}, \quad i \in E.$$

Hence, we have first

$$\inf_{i \in E} R_i(u) = \inf_{i \in E} I_i(f)^{-1} \leq \sup_{f \in \mathcal{F}_I} \inf_{i \in E} I_i(f)^{-1},$$

and then

$$\sup_{u > 0} \inf_{i \in E} R_i(u) \leq \sup_{f \in \mathcal{F}_I} \inf_{i \in E} I_i(f)^{-1},$$

as required. We have thus proved the equality in (2.31) under $\sum_i \mu_i = \infty$.

Actually, we have proved in the last paragraph that $(f_i =) b_i u_i - a_i u_{i-1} > 0$ for all $i \in E$ and so $(v_i := u_{i+1}/u_i) \in \mathcal{V}_*$ whenever $\inf_{i \in E} R_i(u) > 0$. This means that the set $\mathcal{V} \setminus \mathcal{V}_*$ is useless since for each $v \in \mathcal{V} \setminus \mathcal{V}_*$, we have $\inf_{i \in E} R_i(u) \leq 0$. Now, because of $\mathcal{V}_* \subset \mathcal{V}$ and (a), using the equality in (2.31), we obtain the last assertion of part (1).

(d) To prove part (2) of the proposition, note that the inequality “ $\leq$ ” is proved in (b). For the inverse inequality, recalling that the main body in proof (c) is to show that the function f_i ($i \in E$) defined there is positive, this is now automatic due to the definition of $\mathcal{V}_*$. The equality sign in (2.32) has already checked in proofs (a) and (b) in the cases $\lambda_0 = 0$ and $\lambda_0 > 0$, respectively. $\square$

Remark 2.8. For the equality in (2.31), the condition $\sum_i \mu_i = \infty$ cannot be removed. For instance, consider the ergodic case for which $\sum_i \mu_i < \infty$ but $\lambda_0 = 0$ by Theorem 3.1 below and so (2.31) is trivial. However, as proved in [3; Theorem 1.1] (cf. Theorem 6.1 below), the left-hand side of (2.31) coincides with another eigenvalue (called λ_1) which can be positive. In this case, the equality in (2.31) fails. This also explains the reason for the use of $\mathcal{V}_*$.

Remark 2.9. The test sequences with the same notation (v_i) used in Theorem 2.4 and Proposition 2.7 are usually different. Corresponding to the eigenfunction (g_i) of λ_0 , the sequence constructed in proof (d) of Theorem 2.4 and Proposition 2.5 is $v_i = g_{i+1}/g_i$, but the one constructed in proof (b) of Proposition 2.7 is

$$v_i = \frac{g_{i+1} - g_{i+2}}{g_i - g_{i+1}} = \frac{1 - g_{i+2}/g_{i+1}}{g_i/g_{i+1} - 1}.$$

Thus, the mapping from the first sequence to the second one is as follows:

$$(v_i)_{0 \leq i < N} \rightarrow \left(\frac{1 - v_{i+1}}{v_i^{-1} - 1} \right)_{0 \leq i < N}, \quad (2.35)$$

where on the right-hand side, v_N is set to be zero if $N < \infty$.

3. ABSORBING (DIRICHLET) BOUNDARY AT INFINITY: CRITERION, APPROXIMATING PROCEDURE AND EXAMPLES

This section is a continuation of the last one. As applications of the variational formulas given in the last section, a criterion for the positivity of λ_0 and an approximating procedure for λ_0 are presented. The section is ended by a class of examples and then the study on the first case of our classification is completed.

Theorem 3.1 (Criterion and basic estimates). The decay rate $\lambda_0 > 0$ iff $\delta < \infty$, where

$$\delta = \sup_{n \in E} \mu[0, n] \nu[n, N] = \sup_{n \in E} \sum_{j=0}^n \mu_j \sum_{k=n}^N \frac{1}{b_k \mu_k}. \quad (3.1)$$

More precisely, we have $(4\delta)^{-1} \leq \lambda_0 \leq \delta^{-1}$. In particular, when $N = \infty$, we have $\lambda_0 = 0$ if the process is recurrent (i.e., $\nu[1, \infty) = \infty$) and $\lambda_0 > 0$ if the process is explosive (i.e., condition (1.2) does not hold).

Proof. (a) Let $\varphi_n = \sum_{j=n}^N \nu_j =: \nu[n, N]$, $\nu_j = (b_j \mu_j)^{-1}$. To prove the lower estimate, without loss of generality, assume that $\varphi_0 < \infty$. Otherwise, $\delta = \infty$ and so the estimate is trivial. Next, let $M_n = \mu[0, n] := \sum_{k=0}^n \mu_k$. By using the summation by parts formula

$$\sum_{k=0}^n x_k y_k = X_n y_n - \sum_{k=0}^{n-1} X_k (y_{k+1} - y_k), \quad X_n := \sum_{j=0}^n x_j, \quad (3.2)$$

in viewing the definition of δ and using the decreasing property of φ , we get

$$\begin{aligned} \sum_{j=0}^n \mu_j \sqrt{\varphi_j} &= M_n \sqrt{\varphi_n} + \sum_{k=0}^{n-1} M_k (\sqrt{\varphi_k} - \sqrt{\varphi_{k+1}}) \\ &\leq \frac{\delta}{\sqrt{\varphi_n}} + \delta \sum_{k=0}^{n-1} \frac{\sqrt{\varphi_k} - \sqrt{\varphi_{k+1}}}{\varphi_k}. \end{aligned}$$

Noting that

$$(\sqrt{\varphi_k} - \sqrt{\varphi_{k+1}})/\varphi_k \leq 1/\sqrt{\varphi_{k+1}} - 1/\sqrt{\varphi_k},$$

we obtain

$$\sum_{j=0}^n \mu_j \sqrt{\varphi_j} \leq \frac{2\delta}{\sqrt{\varphi_n}}.$$

Therefore,

$$I_n(\sqrt{\varphi}) \leq \frac{1}{\mu_n b_n (\sqrt{\varphi_n} - \sqrt{\varphi_{n+1}})} \cdot \frac{2\delta}{\sqrt{\varphi_n}} = \frac{2\delta}{\sqrt{\varphi_n}} (\sqrt{\varphi_n} + \sqrt{\varphi_{n+1}}) \leq 4\delta.$$

By part (2) of Theorem 2.4, we have $\lambda_0 \geq (4\delta)^{-1}$.

(b) Next, fix arbitrarily $n < m$ and let $f_i = \nu[i \vee n, m] \mathbb{1}_{\{i \leq m\}}$. Then $f \in \widetilde{\mathcal{F}}_I$. To compute $I_i(f)$, note that when $i < n$ or $i > m$, we have $f_i - f_{i+1} = 0$ but $\sum_{j \leq i} \mu_j f_j \geq \mu_0 f_0 > 0$; and when $n \leq i \leq m$, we have $f_i - f_{i+1} = \nu_i = (b_i \mu_i)^{-1}$. Hence, we have

$$I_i(f) = \begin{cases} \mu[0, n] \nu[n, m] + \sum_{n+1 \leq j \leq i} \mu_j \nu[j, m], & n \leq i \leq m, \\ \infty \text{ (by convention, } 1/0 = \infty), & \text{otherwise.} \end{cases}$$

Clearly, $I_i(f)$ achieves its minimum at $i = n$,

$$\inf_{i \in E} I_i(f) = \mu[0, n] \nu[n, m].$$

Since n, m ($n < m$) are arbitrary, by letting $m \rightarrow N$ and making the supremum in n , it follows that

$$\sup_{f \in \tilde{\mathcal{F}}_I} \inf_{i \in E} I_i(f) \geq \sup_{n \in E} \mu[0, n] \nu[n, N] = \delta.$$

By using part (2) of Theorem 2.4 again, we obtain $\lambda_0 \leq \delta^{-1}$. Note that in this proof, we do not preassume that $\delta < \infty$.

(c) The particular assertion for the recurrent case is obvious. The explosive case is also easy since

$$\infty > \sum_{i=0}^{\infty} \mu_i \nu[i, \infty) > \sum_{i=0}^n \mu_i \nu[i, \infty) > \mu[0, n] \nu[n, \infty)$$

for all n , and so $\delta < \infty$. $\square$

The next result is parallel to [7; Theorem 2.2], and is a typical application of parts (2) and (3) of Theorem 2.4. It provides us a way to improve step by step the estimates of λ_0 . In view of Theorem 3.1, the result is meaningful only if $\delta < \infty$.

Theorem 3.2 (Approximating procedure). Write $\nu_j = (\mu_j b_j)^{-1}$ and $\varphi_i = \nu[i, N] := \sum_{j=i}^N \nu_j$, $i \in E$.

- (1) When $\varphi_0 < \infty$, define $f_1 = \sqrt{\varphi}$, $f_n = f_{n-1} II(f_{n-1})$ and $\delta_n = \sup_{i \in E} II_i(f_n)$. Otherwise, define $\delta_n \equiv \infty$. Then δ_n is decreasing in n (denote its limit by δ_∞) and

$$\lambda_0 \geq \delta_\infty^{-1} \geq \dots \geq \delta_1^{-1} \geq (4\delta)^{-1},$$

where δ is defined in Theorem 3.1.

- (2) For fixed $\ell, m \in E$, $\ell < m$ and $m \geq 1$, define

$$\begin{aligned} f_1^{(\ell, m)} &= \nu[\cdot \vee \ell, m] \mathbb{1}_{\leq m}, \\ f_n^{(\ell, m)} &= \mathbb{1}_{\leq m} f_{n-1}^{(\ell, m)} II(f_{n-1}^{(\ell, m)}), \quad n \geq 2, \end{aligned}$$

where $\mathbb{1}_{\leq m}$ is the indicator of the set $\{0, 1, \dots, m\}$, and then define

$$\delta'_n = \sup_{\ell, m: \ell < m} \min_{i \leq m} II_i(f_n^{(\ell, m)}).$$

Then δ'_n is increasing in n (denote its limit by δ'_∞) and

$$\delta^{-1} \geq \delta_1'^{-1} \geq \dots \geq \delta_\infty'^{-1} \geq \lambda_0.$$

Next, define

$$\bar{\delta}_n = \sup_{\ell, m: \ell < m} \frac{\|f_n^{(\ell, m)}\|^2}{D(f_n^{(\ell, m)})}, \quad n \geq 1.$$

Then $\bar{\delta}_n^{-1} \geq \lambda_0$, $\bar{\delta}_{n+1} \geq \delta'_n$ for all $n \geq 1$ and $\bar{\delta}_1 = \delta'_1$.

As the first step of the above approximation, we obtain the following improvement of Theorem 3.1.

Corollary 3.3 (Improved estimates). We have

$$\delta^{-1} \geq \delta_1'^{-1} \geq \lambda_0 \geq \delta_1^{-1} \geq (4\delta)^{-1}, \quad (3.3)$$

where

$$\begin{aligned} \delta_1 &= \sup_{i \in E} \frac{1}{\sqrt{\varphi_i}} \sum_{k \in E} \mu_k \varphi_{i \vee k} \sqrt{\varphi_k} \\ &= \sup_{i \in E} \left[\sqrt{\varphi_i} \sum_{k=0}^i \mu_k \sqrt{\varphi_k} + \frac{1}{\sqrt{\varphi_i}} \sum_{i+1 \leq k < N+1} \mu_k \varphi_k^{3/2} \right]. \end{aligned} \quad (3.4)$$

$$\delta_1' = \sup_{\ell \in E} \frac{1}{\varphi_\ell} \sum_{k \in E} \mu_k \varphi_{k \vee \ell}^2 = \sup_{\ell \in E} \left[\varphi_\ell \mu[0, \ell] + \frac{1}{\varphi_\ell} \sum_{k=\ell+1}^N \mu_k \varphi_k^2 \right] \in [\delta, 2\delta]. \quad (3.5)$$

Proofs of Theorem 3.2 and Corollary 3.3. (a) First, we prove part (1) of Theorem 3.2. Noting that if $\varphi_0 = \infty$, then $\delta = \infty$ and $\delta_n = \infty$ for all $n \geq 1$, the assertion becomes trivial in view of Theorem 3.1. Thus, we can assume that $\varphi_0 < \infty$.

By (2.23), we have

$$\delta_1 = \sup_{i \in E} II_i(f_1) \leq \sup_{i \in E} I_i(f_1).$$

Proof (a) of Theorem 3.1 shows that the last one is bounded from above by 4δ . This gives us the lower bound of δ_1^{-1} as required.

We now prove the monotonicity of $\{\delta_n\}$. By induction, assume that $f_n < \infty$ and $\delta_n < \infty$. Then $f_{n+1} < \infty$. Note that

$$\begin{aligned} \sum_{j \leq k} \mu_j f_{n+1}(j) &= \sum_{j \leq k} \mu_j f_n(j) f_{n+1}(j) / f_n(j) \\ &\leq \sup_{i \in E} II_i(f_n) \sum_{j \leq k} \mu_j f_n(j) \\ &= \delta_n \sum_{j \leq k} \mu_j f_n(j). \end{aligned}$$

Multiplying both sides by ν_k and making a summation of k from i to N , by (2.14), it follows that

$$f_{n+2}(i) \leq \delta_n f_{n+1}(i).$$

Because $\delta_n < \infty$ and $f_{n+1}(i) < \infty$, we obtain $f_{n+2} < \infty$ and $II_i(f_{n+1}) \leq \delta_n < \infty$. Now, making the supremum over i , we obtain $\delta_{n+1} \leq \delta_n < \infty$.

We have thus proved part (1) of Theorem 3.2.

(b) To prove the monotonicity of δ_n' given in part (2) of Theorem 3.2, we use the proportional property twice:

$$\begin{aligned} \min_{i \leq m} [f_{n+1}^{(\ell, m)} / f_n^{(\ell, m)}](i) &= \min_{i \leq m} \sum_{j=i}^m \nu_j \sum_{k \leq j} \mu_k f_n^{(\ell, m)}(k) \bigg/ \sum_{j=i}^m \nu_j \sum_{k \leq j} \mu_k f_{n-1}^{(\ell, m)}(k) \\ &\geq \min_{i \leq m} \sum_{k \leq i} \mu_k f_n^{(\ell, m)}(k) \bigg/ \sum_{k \leq i} \mu_k f_{n-1}^{(\ell, m)}(k) \\ &\geq \min_{i \leq m} f_n^{(\ell, m)}(i) / f_{n-1}^{(\ell, m)}(i). \end{aligned}$$

This implies that $\delta'_{n+1} \geq \delta'_n$.

By part (3) of Theorem 2.4, we also have $\delta'_n \leq \lambda_0^{-1}$ for all $n \geq 1$. The assertion that $\bar{\delta}_n \leq \lambda_0^{-1}$ is obvious. Next, let $f = f_n^{(\ell, m)}$. Then $g := \mathbb{1}_{\text{supp}(f)} f \Pi(f) = f_{n+1}^{(\ell, m)}$. As a consequence of (2.30), we obtain $\bar{\delta}_{n+1} \geq \delta'_n$.

We have thus proved part (2) of Theorem 3.2 except the last assertion that $\bar{\delta}_1 = \delta'_1$.

(c) We now prove (3.4) and $\delta'_1 \geq \delta$. By (2.15), we have

$$\begin{aligned} f_{n+1}(i) &= \sum_{k \in E} \mu_k f_n(k) \nu[i \vee k, N] \\ &= \sum_{k \in E} \mu_k f_n(k) \varphi_{i \vee k} \\ &= \varphi_i \sum_{k=0}^i \mu_k f_n(k) + \sum_{i+1 \leq k < N+1} \mu_k \varphi_k f_n(k). \end{aligned} \quad (3.6)$$

In particular, with $f_1 = \sqrt{\varphi}$, we get

$$f_2(i) = \varphi_i \sum_{k=0}^i \mu_k \sqrt{\varphi_k} + \sum_{i+1 \leq k < N+1} \mu_k \varphi_k^{3/2}. \quad (3.7)$$

From this, we obtain (3.4).

To prove $\delta'_1 \geq \delta$, we need some preparation. As an analog of (3.6), we have

$$f_{n+1}^{(\ell, m)}(i) = \mathbb{1}_{\{i \leq m\}} \sum_{k \leq m} \mu_k f_n^{(\ell, m)}(k) \nu[i \vee k, m]. \quad (3.8)$$

In particular,

$$f_2^{(\ell, m)}(i) = \mathbb{1}_{\{i \leq m\}} \sum_{k \leq m} \mu_k \nu[k \vee \ell, m] \nu[i \vee k, m]. \quad (3.9)$$

Since the right-hand side is decreasing in i for $i \leq \ell < m$, $f_1^{(\ell, m)}(i) = f_1^{(\ell, m)}(\ell)$ for all $i \leq \ell$, and $f_1^{(\ell, m)}(i) = 0$ for $i > m$, it follows that

$$\begin{aligned} \min_{i \leq m} \Pi_i(f_1^{(\ell, m)}) &= \min_{\ell \leq i \leq m} \Pi_i(f_1^{(\ell, m)}) \\ &= \min_{\ell \leq i \leq m} \frac{1}{\nu[i, m]} \sum_{j=i}^m \nu_j \sum_{k \leq j} \mu_k \nu[k \vee \ell, m] \mathbb{1}_{\{k \leq m\}} \\ &= \min_{\ell \leq i \leq m} \sum_{j=i}^m \nu_j \sum_{k \leq j} \mu_k \nu[k \vee \ell, m] \bigg/ \sum_{j=i}^m \nu_j \\ &\geq \min_{\ell \leq i \leq m} \sum_{k \leq i} \mu_k \nu[k \vee \ell, m] \quad \left[= \inf_{i \in E} \Pi_i(f_1^{(\ell, m)}) \right]. \end{aligned}$$

Here in the last step, we have used the proportional property. Since the sum on the right-hand side is increasing in i , it is clear that

$$\min_{\ell \leq i \leq m} \sum_{k \leq i} \mu_k \nu[k \vee \ell, m] = \nu[\ell, m] \sum_{k \leq \ell} \mu_k = \mu[0, \ell] \nu[\ell, m].$$

We have thus proved that

$$\delta'_1 = \sup_{\ell < m} \min_{i \leq m} II_i(f_1^{(\ell, m)}) \geq \sup_{\ell < m} \mu[0, \ell] \nu[\ell, m] \geq \sup_{\ell \in E} \mu[0, \ell] \varphi_\ell = \delta.$$

A different proof of this is given in proof (d) below.

(d) We now compute δ'_1 . Note that by (3.9), we have

$$f_2^{(\ell, m)}(i) = \mathbb{1}_{\{i \leq m\}} \sum_{k=0}^m \mu_k \nu[k \vee \ell, m] \nu[i \vee k, m].$$

Since $f_2^{(\ell, m)}(i)$ is decreasing in i and $f_1^{(\ell, m)}(i)$ is a constant on $\{0, 1, \dots, \ell\}$, it is clear that $\min_{i \leq m} f_2^{(\ell, m)}(i) / f_1^{(\ell, m)}(i) = \min_{\ell \leq i \leq m} f_2^{(\ell, m)}(i) / f_1^{(\ell, m)}(i)$. Besides, when $\ell \leq i \leq m$, we have

$$f_2^{(\ell, m)}(i) = \nu[\ell, m] \nu[i, m] \sum_{k=0}^{\ell} \mu_k + \nu[i, m] \sum_{\ell+1 \leq k \leq i} \mu_k \nu[k, m] + \sum_{i+1 \leq k \leq m} \mu_k \nu[k, m]^2.$$

It follows that

$$\min_{i \leq m} \frac{f_2^{(\ell, m)}(i)}{f_1^{(\ell, m)}(i)} = \min_{\ell \leq i \leq m} \left[\nu[\ell, m] \sum_{k=0}^{\ell} \mu_k + \sum_{k=\ell+1}^i \mu_k \nu[k, m] + \frac{1}{\nu[i, m]} \sum_{k=i+1}^m \mu_k \nu[k, m]^2 \right].$$

We show that the sum on the right-hand side is increasing in i . That is,

$$\begin{aligned} & \sum_{\ell+1 \leq k \leq i} \mu_k \nu[k, m] + \frac{1}{\nu[i, m]} \sum_{k=i+1}^m \mu_k \nu[k, m]^2 \\ & \leq \sum_{k=\ell+1}^{i+1} \mu_k \nu[k, m] + \frac{1}{\nu[i+1, m]} \sum_{i+2 \leq k \leq m} \mu_k \nu[k, m]^2, \quad \ell \leq i \leq m-1. \end{aligned}$$

Collecting the terms, this is equivalent to

$$\frac{1}{\nu[i, m]} \mu_{i+1} \nu[i+1, m]^2 \leq \mu_{i+1} \nu[i+1, m] + \left(\frac{1}{\nu[i+1, m]} - \frac{1}{\nu[i, m]} \right) \sum_{k=i+2}^m \mu_k \nu[k, m]^2.$$

Now, the conclusion becomes obvious because by the decreasing property of $\nu[i, m]$ in i , the first term is controlled by the second, and the last one is nonnegative.

We have thus obtained that

$$\min_{\ell \leq i \leq m} \frac{f_2^{(\ell, m)}(i)}{f_1^{(\ell, m)}(i)} = \nu[\ell, m] \sum_{k=0}^{\ell} \mu_k + \frac{1}{\nu[\ell, m]} \sum_{k=\ell+1}^m \mu_k \nu[k, m]^2. \quad (3.10)$$

As will be seen soon that the right-hand side is increasing in $m (> \ell)$, hence, we obtain

$$\delta'_1 = \sup_{\ell < m} \min_{\ell \leq i \leq m} \frac{f_2^{(\ell, m)}(i)}{f_1^{(\ell, m)}(i)} = \sup_{\ell \in E} \left[\varphi_\ell \sum_{k=0}^{\ell} \mu_k + \frac{1}{\varphi_\ell} \sum_{\ell+1 \leq k < N+1} \mu_k \varphi_k^2 \right]. \quad (3.11)$$

From this, it follows once again that $\delta'_1 \geq \delta$. We now turn to prove the monotone property:

$$\begin{aligned} & \mu[0, \ell] \nu[\ell, m+1] + \frac{1}{\nu[\ell, m+1]} \sum_{i=\ell+1}^{m+1} \mu_i \nu[i, m+1]^2 \\ & \geq \mu[0, \ell] \nu[\ell, m] + \frac{1}{\nu[\ell, m]} \sum_{i=\ell+1}^m \mu_i \nu[i, m]^2. \end{aligned}$$

Equivalently,

$$\mu[0, \ell] \nu_{m+1} + \frac{\mu_{m+1}}{\nu[\ell, m+1]} \nu_{m+1}^2 + \sum_{i=\ell+1}^m \mu_i \left(\frac{\nu[i, m+1]^2}{\nu[\ell, m+1]} - \frac{\nu[i, m]^2}{\nu[\ell, m]} \right) \geq 0.$$

This becomes obvious since the term in the last bracket is positive:

$$\frac{\nu[i, m+1]^2}{\nu[i, m]^2} = \left(1 + \frac{\nu_{m+1}}{\nu[i, m]} \right)^2 > 1 + \frac{\nu_{m+1}}{\nu[\ell, m]} = \frac{\nu[\ell, m+1]}{\nu[\ell, m]}, \quad \ell \leq i \leq m.$$

(e) To show that $\delta'_1 \leq 2\delta$, assume $\delta < \infty$. By using the summation by parts formula (3.2) with $x_k = \mu_k$, $X_k = \sum_{j=0}^k \mu_j$, and $y_k = \varphi_{k \vee i}^2$, we get

$$\begin{aligned} \sum_{k=0}^M \mu_k \varphi_{k \vee i}^2 &= \varphi_M^2 X_M + \sum_{k=0}^{M-1} X_k [\varphi_{k \vee i}^2 - \varphi_{(k+1) \vee i}^2] \\ &= \varphi_M^2 X_M + \sum_{k=i}^{M-1} X_k [\varphi_k^2 - \varphi_{k+1}^2] \\ &= \varphi_M^2 X_M + \sum_{k=i}^{M-1} X_k \nu_k (\varphi_k + \varphi_{k+1}) \\ &< \varphi_M^2 X_M + 2 \sum_{k=i}^{M-1} X_k \nu_k \varphi_k \\ &\leq \delta \varphi_M + 2\delta \sum_{k=i}^{M-1} \nu_k \quad (\text{since } X_k \varphi_k \leq \delta), \quad i < M < N+1. \end{aligned}$$

If $N = \infty$, letting $M \rightarrow N$, it follows that

$$\sum_{k=0}^N \mu_k \varphi_{k \vee i}^2 \leq 2\delta \varphi_i.$$

The same conclusion holds in the case that $N < \infty$ since

$$\delta\varphi_N + 2\delta \sum_{k=i}^{N-1} \nu_k < 2\delta \sum_{k=i}^N \nu_k = 2\delta\varphi_i.$$

Hence,

$$\delta'_1 = \sup_{i \in E} \frac{1}{\varphi_i} \sum_{k=0}^N \mu_k \varphi_{k \vee i}^2 \leq 2\delta.$$

(f) Now, it remains to compute $\bar{\delta}_1$. Since $f_1^{(\ell, m)}(i) = \nu[i \vee \ell, m] \mathbb{1}_{\{i \leq m\}}$, we have

$$\|f_1^{(\ell, m)}\|^2 = \sum_i \mu_i \nu[i \vee \ell, m]^2 \mathbb{1}_{\{i \leq m\}} = \mu[0, \ell] \nu[\ell, m]^2 + \sum_{i=\ell+1}^m \mu_i \nu[i, m]^2,$$

and

$$\begin{aligned} D(f_1^{(\ell, m)}) &= \sum_i \mu_i b_i (f_1^{(\ell, m)}(i+1) - f_1^{(\ell, m)}(i))^2 \\ &= \sum_{i=\ell}^{m-1} \mu_i b_i (\nu[i+1, m] - \nu[i, m])^2 + \mu_m b_m \nu_m^2 \\ &= \sum_{i=\ell}^{m-1} \nu_i + \nu_m \\ &= \nu[\ell, m]. \end{aligned}$$

Thus,

$$\frac{\|f_1^{(\ell, m)}\|^2}{D(f_1^{(\ell, m)})} = \mu[0, \ell] \nu[\ell, m] + \frac{1}{\nu[\ell, m]} \sum_{i=\ell+1}^m \mu_i \nu[i, m]^2.$$

Hence, we have returned to (3.10). Since the right-hand side is increasing in m as we have seen in the proof of (3.11), we obtain

$$\bar{\delta}_1 = \sup_{\ell < m} \frac{\|f_1^{(\ell, m)}\|^2}{D(f_1^{(\ell, m)})} = \delta'_1. \quad \square$$

To conclude this section, we present some examples to illustrate the power of our results. The first one is standard having constant rates.

Example 3.4. Let $b_i \equiv b > 0$ ($i \geq 0$), $a_i \equiv a > 0$ ($i \geq 1$), $b > a$. Then

(1) $\lambda_0 = (\sqrt{a} - \sqrt{b})^2$ with eigenfunction g :

$$g_n = \left(\frac{a}{b}\right)^{n/2} \left(n + 1 - n\sqrt{\frac{a}{b}}\right), \quad n \geq 0, \quad g \notin L^1(\mu) \cup L^2(\mu).$$

(2) $\delta = b(b-a)^{-2}$, $\delta'_1 = (a+b)(b-a)^{-2} = \bar{\delta}_1 > \delta$, and $\delta_1 = \lambda_0^{-1}$ which is exact. Note that $\delta_1/\delta'_1 < 2$ whenever $a \neq b$ and $\lim_{b \rightarrow a} \delta_1/\delta'_1 = 2$. When $a = b$, we have $\lambda_0 = \delta_1^{-1} = \delta'_1^{-1} = 0$.

Proof. (a) First, we have $\mu_n = (b/a)^n$, $n \geq 0$. Hence,

$$\sum_n \mu_n = \infty, \quad \sum_n \mu_n g_n \geq \sum_n \mu_n g_n^2 \geq \sum_n 1 = \infty.$$

Next, since

$$\nu_i = \frac{1}{\mu_i b_i} = \frac{1}{b} \left(\frac{a}{b} \right)^i,$$

we have

$$\varphi_\ell = \sum_{i \geq \ell} \nu_i = \frac{1}{b-a} \left(\frac{a}{b} \right)^\ell$$

and then

$$\sum_{i=0}^{\infty} \mu_i \sum_{k=i}^{\infty} \frac{1}{b_k \mu_k} = \sum_{i=0}^{\infty} \mu_i \varphi_i = \infty.$$

Hence, (1.2) holds. It is easy to check that (2.12) holds:

$$\sum_{n=0}^{\infty} \mu_n g_n \nu[n, \infty) = \frac{1}{b-a} \sum_{n=0}^{\infty} \left(\frac{a}{b} \right)^{n/2} \left(n+1 - n \sqrt{\frac{a}{b}} \right) = \frac{1}{\lambda_0}.$$

(b) To study λ_0 , according to (a), the Dirichlet form is regular and so the condition “ $f \in \mathcal{K}$ ” in the definition of λ_0 can be ignored. Thus,

$$\lambda_0 = \inf_{\|f\|=1} D(f) = b \inf_{\|f\|=1} \sum_{i \geq 0} \mu_i (f_{i+1} - f_i)^2.$$

It suffices to consider the case that $b = 1$. Write $\gamma = b/a > 1$. Then we have

$$g_k = \gamma^{-k/2} (k+1 - k\gamma^{-1/2}), \quad \mu_k = \gamma^k, \quad \nu_k = \gamma^{-k}, \quad \varphi_k = \gamma^{-k+1}/(\gamma-1),$$

and the required quantities are reduced to

$$\lambda_0 = \frac{(\sqrt{\gamma}-1)^2}{\gamma}, \quad \delta = \frac{\gamma^2}{(\gamma-1)^2}, \quad \delta_1 = \frac{\gamma}{(\sqrt{\gamma}-1)^2}, \quad \delta'_1 = \frac{\gamma(\gamma+1)}{(\gamma-1)^2}.$$

Now, to prove part (1) of Example 3.4, write $\xi = (\sqrt{\gamma}-1)^2 \gamma^{-1}$ for distinguishing with λ_0 . Since (g, ξ) satisfies the eigenequation, applying anyone of the variational formulas for the lower estimate given in Theorem 2.4 with $f_i = g_i$ or

$$v_i = \frac{g_{i+1}}{g_i} = \sqrt{\frac{a}{b}} \left(1 + \frac{1 - \sqrt{a/b}}{1 + i(1 - \sqrt{a/b})} \right) = \gamma^{-1/2} \left(1 + \frac{1 - \gamma^{-1/2}}{1 + i(1 - \gamma^{-1/2})} \right),$$

it follows that $\lambda_0 \geq \xi$. We have seen that the equality sign holds once $g \in L^2(\mu)$. Unfortunately, we are now out of this case. Therefore, we need to show that

$\lambda_0 \leq \xi$. To do so, one may use the truncated function of g : $g_i^{(m)} = g_i \mathbb{1}_{\{i \leq m\}}$. Then by the Stolz theorem, we have

$$\lambda_0 \leq \lim_{m \rightarrow \infty} \frac{D(g^{(m)})}{\|g^{(m)}\|^2} = \lim_{m \rightarrow \infty} \left[b_m + \frac{\mu_{m-1} b_{m-1}}{\mu_m} \left(1 - \frac{2g_{m-1}}{g_m} \right) \right]. \quad (3.12)$$

The last limit equals ξ . Alternatively, noting that the leading order of $g \notin L^2(\mu)$ is $\gamma^{-k/2}$, one may adopt the test function $f_i = z^{-i/2}$ for $z > \gamma$. Then $f \in L^2(\mu)$. The required assertion follows by computing $D(f)/\|f\|^2$ and then letting $z \downarrow \gamma$. This proof benefits very much from the explicitly known expression of λ_0 .

(c) The computation of δ is easy:

$$\delta = \sup_{n \geq 0} \varphi_n \sum_{j=0}^n \mu_j = \frac{1}{(\gamma-1)^2} \sup_{n \geq 0} \gamma^{-n+1} (\gamma^{n+1} - 1) = \frac{\gamma^2}{(\gamma-1)^2}.$$

(d) To compute δ_1 , by (3.7), we have

$$\begin{aligned} f_2(i) &= \varphi_i \sum_{k=0}^i \mu_k \sqrt{\varphi_k} + \sum_{k=i+1}^{\infty} \mu_k \varphi_k^{3/2} \\ &= \frac{1}{(\gamma-1)^{3/2}} \left\{ \gamma^{-i+1} \sum_{k=0}^i \gamma^{k/2+1/2} + \sum_{k \geq i+1} \gamma^{-k/2+3/2} \right\} \\ &= \frac{\gamma^{-i/2+3/2}}{(\gamma-1)^{3/2}(\sqrt{\gamma}-1)} (\sqrt{\gamma} - \gamma^{-i/2} + 1). \end{aligned}$$

Therefore, we obtain

$$\delta_1 = \sup_{i \geq 0} \frac{f_2(i)}{f_1(i)} = \frac{\gamma}{(\gamma-1)(\sqrt{\gamma}-1)} (\sqrt{\gamma} + 1) = \frac{\gamma}{(\sqrt{\gamma}-1)^2} = \frac{1}{\lambda_0}.$$

Noting that even if neither f_1 nor f_2 is the eigenfunction, we still obtain the sharp estimate.

(e) To compute δ'_1 , by (3.5), we have

$$\begin{aligned} \delta'_1 &= \sup_{\ell \in E} \left[\varphi_\ell \sum_{k=0}^{\ell} \mu_k + \frac{1}{\varphi_\ell} \sum_{k \geq \ell+1} \mu_k \varphi_k^2 \right] \\ &= \frac{1}{\gamma-1} \sup_{\ell \in E} \left[\gamma^{-\ell+1} \sum_{k=0}^{\ell} \gamma^k + \gamma^{\ell+1} \sum_{k \geq \ell+1} \gamma^{-k} \right] \\ &= \frac{1}{(\gamma-1)^2} \sup_{\ell \in E} [\gamma^2 - \gamma^{-\ell+1} + \gamma] \\ &= \frac{\gamma(\gamma+1)}{(\gamma-1)^2}. \quad \square \end{aligned}$$

The next example is a typical linear model for which, interestingly, we have a very simple and common eigenfunction. Moreover, the eigenvalue λ_0 is determined by the constant term 2γ in the rates, but not the difference of the coefficients of the leading term i , as in the ergodic case (cf. Example 6.8 below).

Example 3.5. Let $b_i = 2(i + \gamma)$ ($i \geq 0$), $\gamma > 0$, $a_i = i$ ($i \geq 1$). Then

- (1) $\lambda_0 = \gamma$, $g_n = 2^{-n}$ for $n \geq 0$, and $g \in L^2(\mu)$.
- (2) When $\gamma = 1$, we have $\delta = \log 2 \approx 0.69$, $\delta'_1 \approx 0.84$, and $\delta_1 \approx 1.09$. Then $\delta_1/\delta'_1 \approx 1.3 < 2$.

Proof. The uniqueness condition (1.2) is trivial since the birth rates are linear:

$$\sum_{k=0}^{\infty} \frac{1}{b_k \mu_k} \sum_{i=0}^k \mu_i = \sum_{k=0}^{\infty} \left[\frac{1}{b_k} + \frac{1}{b_k \mu_k} \sum_{i=0}^{k-1} \mu_i \right] \geq \sum_{k=0}^{\infty} \frac{1}{b_k} = \infty.$$

(a) Because

$$\mu_0 = 1, \quad \mu_n = \frac{2^n \gamma (1 + \gamma) \cdots (n - 1 + \gamma)}{n!}, \quad n \geq 1,$$

it follows that $\mu_n > \gamma 2^n / n$ and so $\sum_n \mu_n = \infty$. Next, since

$$\mu_n b_n = \frac{2^{n+1} \gamma (1 + \gamma) \cdots (n + \gamma)}{n!} > \gamma 2^{n+1},$$

we have $\sum_n (\mu_n b_n)^{-1} < \infty$. Furthermore, we have

$$\sum_{n=0}^{\infty} \mu_n g_n^2 = \sum_{n=0}^{\infty} \frac{2^{-n} \gamma (1 + \gamma) \cdots (n - 1 + \gamma)}{n!}.$$

The ratio test tells us $g \in L^2(\mu)$. Since λ_0 is explicit and $g \in L^2(\mu)$, it is simple to check that (g_n) is the eigenfunction of λ_0 . Hence, the proof of part (1) is done. For this example, the sequence (v_i) takes a simple form: $v_i \equiv 1/2$.

(b) When $\gamma = 1$, we have $\lambda_0 = 1$,

$$\mu_i = 2^i, \quad \mu_i b_i = (i + 1) 2^{i+1}, \quad \varphi_i = \sum_{k \geq i+1} \frac{1}{2^k k}, \quad i \geq 0.$$

In particular, $\varphi_0 = \log 2$, $\varphi_1 = \log 2 - 1/2$. Numerical computations show that the supremum in the definition of δ , δ'_1 and δ_1 are attained at 0, 0 and 1, respectively, and moreover,

$$\begin{aligned} \delta &= \varphi_0 \mu_0 = \varphi_0 = \log 2 \approx 0.69, \\ \delta'_1 &= \varphi_0 \mu_0 + \frac{1}{\varphi_0} \sum_{k \geq 1} \mu_k \varphi_k^2 = \log 2 + \frac{1}{\log 2} \sum_{k \geq 1} 2^k \varphi_k^2 \approx 0.84, \\ \delta_1 &= \sqrt{\varphi_1} (\mu_0 \sqrt{\varphi_0} + \mu_1 \sqrt{\varphi_1}) + \frac{1}{\sqrt{\varphi_1}} \sum_{k \geq 2} \mu_k \varphi_k^{3/2} \\ &= 2 \log 2 - 1 + \frac{1}{2} \sqrt{(2 \log 2)(2 \log 2 - 1)} + \sqrt{\frac{2}{2 \log 2 - 1}} \sum_{k \geq 2} 2^k \varphi_k^{3/2} \\ &\approx 1.09. \end{aligned}$$

We have thus proved part (2) of the conclusion. $\square$

The next example is often used in the study of convergence rates. For which, the first eigenfunction is unknown but λ_0 can still be computed.

Example 3.6. Let $b_i = (i+1)^2$ and $a_i = i^2$. Then $\delta = \pi^2/6 \approx 1.64$, $\delta'_1 \approx 2.19$, and $\delta_1 = 4$ which is sharp ($\lambda_0 = 1/4$). Besides, $\delta_1/\delta'_1 \approx 1.83 < 2$.

Proof. (a) Since $\mu_i \equiv 1$, $\nu_i = (i+1)^{-2}$, we have $\mu[0, i] = i+1$ and $\varphi_i = \sum_{j \geq i+1} j^{-2}$. For δ and δ'_1 , the supremum is attained at 0, therefore,

$$\delta = \varphi_0 = \sum_{k \geq 1} \frac{1}{k^2} = \frac{\pi^2}{6},$$

and

$$\delta'_1 = \frac{1}{\varphi_0} \sum_{k=0}^{\infty} \varphi_k^2 \approx 2.19.$$

(b) For δ_1 , the supremum is attained at ∞ and is equal to 4. By Corollary 3.3, this means that $\lambda_0 \geq 1/4$. This can be also deduced by part (1) of Theorem 2.4 with $v_i = 1 - (2i+4)^{-1}$ for which the minimum of $R_i(v)$ is attained at $i = 0$ and $i = \infty$. It is even more simpler to use $v_i = 1 - (2i+3)^{-1}$. Next, it is known that $\lambda_0 \leq 1/4$ (cf. Example 5.5 below), hence, the estimate is sharp. A direct proof for the upper estimate goes as follows. Since the lower estimate is sharp, it indicates to use the test function

$$f_i = \left(\sum_{j=i}^{\infty} \frac{1}{(j+1)^2} \right)^{1/2} \sim \frac{1}{\sqrt{i+1}}.$$

However, the last function is not in $L^2(\mu)$, and so one needs an approximating procedure. Now, a carefully designed test function is the following:

$$f_i^{(\alpha)} = \frac{1}{\sqrt{(i+1)\alpha^{i+1}}}, \quad \alpha > 1.$$

Then

$$\begin{aligned} \mu(f^{(\alpha)2}) &= \sum_{i=0}^{\infty} \frac{1}{(i+1)\alpha^{i+1}} = \sum_{i=1}^{\infty} \frac{1}{i\alpha^i} = \log[\alpha(\alpha-1)^{-1}] < \infty, \\ D(f^{(\alpha)}) &= \sum_{i=0}^{\infty} (i+1)^2 \left[\frac{1}{\sqrt{(i+2)\alpha^{i+2}}} - \frac{1}{\sqrt{(i+1)\alpha^{i+1}}} \right]^2 \\ &= \sum_{i=1}^{\infty} \frac{i^2}{\alpha^i} \left[\frac{1}{\sqrt{(i+1)\alpha}} - \frac{1}{\sqrt{i}} \right]^2 \\ &= \sum_{i=1}^{\infty} \frac{i}{(i+1)\alpha^{i+1}} \frac{[(i+1)\alpha - i]^2}{[\sqrt{(i+1)\alpha} + \sqrt{i}]^2} \\ &\leq \frac{1}{4} \sum_{i=1}^{\infty} \frac{1}{(i+1)\alpha^{i+1}} [(i+1)\alpha - i]^2 \\ &= \frac{1}{4} (2 + \log[\alpha(\alpha-1)^{-1}]). \end{aligned}$$

The required assertion now follows from

$$\lambda_0 \leq \frac{2 + \log[\alpha(\alpha - 1)^{-1}]}{4 \log[\alpha(\alpha - 1)^{-1}]} \rightarrow \frac{1}{4} \quad \text{as } \alpha \downarrow 1. \quad \square$$

The last example below does not satisfy the non-explosive condition (1.2).

Example 3.7. Let $b_i = (i+1)^4$ and $a_i = i(i-1/2)(i^2+3i+3)$. Then $\sum_i \mu_i < \infty$, $\sum_i \nu_i < \infty$, $\lambda_0 = 1/2$, $\delta \approx 1.83$, $\delta'_1 \approx 1.9$, and $\delta_1 \approx 2$. Moreover, $\delta_1/\delta'_1 \approx 1.05 < 2$.

Proof. A simple computation shows that

$$\mu_i = \frac{i!^3}{\prod_{k=1}^i (k-1/2)(k^2+3k+3)}, \quad \nu_i = \frac{\prod_{k=1}^i (k-1/2)(k^2+3k+3)}{(i+1)(i+1)!^3}.$$

From this, it follows that $\sum_i \mu_i < \infty$ and $\sum_i \nu_i < \infty$, as an application of the typical Kummer's test: for a positive sequence $\{x_n\}$, $\sum_n x_n$ converges or diverges according to $\kappa > 1$ or $\kappa < 1$, respectively, where

$$\kappa = \lim_{n \rightarrow \infty} n \left(\frac{x_n}{x_{n+1}} - 1 \right). \quad (3.13)$$

For each of δ , δ'_1 and δ_1 , the supremum is attained at 0.

To see that $\lambda_0 = 1/2$, first we check that $R_i(v) \equiv 1/2$ for

$$v_i = 1 - \frac{1}{2(i+1)}.$$

This gives us $\lambda_0 \geq 1/2$ by part (1) of Theorem 2.4. Since the corresponding eigenfunction g ,

$$g_i = \prod_{k=0}^{i-1} v_k = \frac{(2i-1)!}{2^{2i-1} i(i-1)!^2}, \quad i \geq 1, \quad g_0 = 1,$$

decreases strictly to 0 and $\sum_i \mu_i < \infty$, we have $g \in L^2(\mu)$. Now, because $-\Omega g = \lambda_0 g$, $g_\infty = 0$, and $D(f) = -(g, \Omega g)$, it follows that $\lambda_0 = 1/2$ by (2.18). $\square$

4. ABSORBING (DIRICHLET) BOUNDARY AT ORIGIN AND REFLECTING (NEUMANN) BOUNDARY AT INFINITY

This section deals with the second case of the boundary conditions. The process has state space $E = \{i : 1 \leq i < N+1\}$ ($N \leq \infty$), birth rates $b_i > 0$ but $b_N = 0$ if $N < \infty$, and death rates $a_i > 0$. The rate $a_1 > 0$ is regarded as a killing from 1. Define

$$\lambda_0 = \inf \{D(f)/\mu(f^2) : f \neq 0, D(f) < \infty\}, \quad (4.1)$$

where $\mu(f) = \sum_{k \in E} \mu_k f_k$, and

$$D(f) = \sum_{k \in E} \mu_k a_k (f_k - f_{k-1})^2, \quad f_0 := 0,$$

$$\mu_1 = 1, \quad \mu_k = \frac{b_1 \cdots b_{k-1}}{a_2 \cdots a_k}, \quad 2 \leq k < N+1.$$

The constant $\lambda_0^{(4,1)}$ describes the optimal constant $C = \lambda_0^{-1}$ in the following *weighted Hardy inequality*:

$$\mu(f^2) \leqslant CD(f), \quad f_0 = 0$$

(cf. [9]). In other words, we are studying the discrete version of the weighted Hardy inequality in this section. To save the notation, in this and the subsequent sections, we use the same notation λ_0 , I , II , R and so on as in Section 2. Each of them plays a similar role but may have different meaning in different sections.

To study λ_0 , as in Section 2, we need some parallel notation originally introduced in [3, 7]:

$$I_i(f) = \frac{1}{\mu_i a_i (f_i - f_{i-1})} \sum_{j=i}^N \mu_j f_j, \quad II_i(f) = \frac{1}{f_i} \sum_{j=1}^i \frac{1}{\mu_j a_j} \sum_{k=j}^N \mu_k f_k.$$

Here, for the first operator, we adopt the convention: $f_0 = 0$. The second one can be re-written as

$$II_i(f) = \frac{1}{f_i} \sum_{k=1}^N \mu_k f_k \nu[1, i \wedge k], \quad \nu[\ell, m] = \sum_{j=\ell}^m \nu_j, \quad \nu_j = \frac{1}{\mu_j a_j}.$$

Next, define

$$R_i(v) = a_i(1 - v_{i-1}^{-1}) + b_i(1 - v_i), \quad i \in E, \quad v_0 := \infty$$

(v_N is free if $N < \infty$ since $b_N = 0$) and

$$\begin{aligned} \mathcal{F}_{II} &= \{f : f_i > 0 \text{ for all } i \in E\}, \\ \mathcal{F}_I &= \{f : f > 0 \text{ and is strictly increasing on } E\}, \\ \mathcal{V}_1 &= \{v : v_i > 1 \text{ for all } i \in E\}. \end{aligned}$$

The modifications of $\mathcal{F}_{II}$ and $\mathcal{F}_I$ are as follows:

$$\begin{aligned} \widetilde{\mathcal{F}}_{II} &= \{f : \text{there exists } m \in E \text{ such that } f_i = f_{i \wedge m} > 0 \text{ for } i \in E\}, \\ \widetilde{\mathcal{F}}_I &= \{f : \text{there exists } m \in E \text{ such that } f_i = f_{i \wedge m} > 0 \text{ for } i \in E \text{ and } f \text{ is} \\ &\quad \text{strictly increasing in } \{1, \dots, m\}\}. \end{aligned}$$

Here, we use again the convention: $1/0 = \infty$. Note that for the localization, f is stopped at m rather than vanishing after m used in Sections 2 and 3. This is due to the fact that the Neumann boundary is imposed at m but not the Dirichlet one. Besides, for the operator II here, the restriction on $\text{supp}(f)$ used in Section 2 is no longer needed. Finally, define a local operator $\widetilde{R}$ (depending on m) acting on

$$\begin{aligned} \widetilde{\mathcal{V}}_1 &= \cup_{m \in E} \{v : 1 < v_i < 1 + a_i(1 - v_{i-1}^{-1})b_i^{-1} \text{ for } i = 1, 2, \dots, m-1 \\ &\quad \text{and } v_i = 1 \text{ for } i \geqslant m\} \end{aligned}$$

by replacing a_m with $\tilde{a}_m := \mu_m a_m / \sum_{k=m}^N \mu_k$ in $R_i(v)$ for the same m as in $\tilde{\mathcal{V}}_1$. Again, the change of a_m is due to the Neumann boundary at m . Note that if $v_i = 1$ for all $i \geq m$, then $\tilde{R}_i(v) = R_i(v) = 0$ for all $i > m$.

Before stating our main results in this section, we mention an exceptional case that $\sum_i \mu_i = \infty$. On the one hand, by choosing $f_0 = 0$ and $f_i = 1$ for $i \geq 1$, it follows that

$$D(f) = \mu_1 a_1 < \infty, \quad \mu(f^2) = \sum_{i \geq 1} \mu_i = \infty$$

and so $\lambda_0 = 0$. On the other hand, if $\sum_{i=1}^N \mu_i < \infty$, then for every f with $\mu(f^2) = \infty$, by setting $f^{(m)} = f_{\cdot \wedge m} \in L^2(\mu)$, we get

$$\begin{aligned} \infty > D(f^{(m)}) &= \sum_{i=1}^m \mu_i a_i (f_i - f_{i-1})^2 \uparrow D(f) \quad \text{as } m \rightarrow \infty, \\ \infty > \mu(f^{(m)2}) &\geq \sum_{i=1}^m \mu_i f_i^2 \rightarrow \infty = \mu(f^2) \quad \text{as } m \rightarrow \infty. \end{aligned}$$

In words, for each non-square-integrable function f , both $\mu(f^2)$ and $D(f)$ can be approximated by a sequence of square-integrable ones. Hence, we can rewrite λ_0 as follows:

$$\lambda_0 = \inf\{D(f) : \mu(f^2) = 1\}. \quad (4.2)$$

In this case, as will be seen soon but not obvious, we also have

$$\lambda_0 = \inf\{D(f) : \mu(f^2) = 1, f_i = f_{i \wedge m} \text{ for some } m \in E \text{ and all } i \in E\}, \quad (4.3)$$

Besides, we mention that the Dirichlet eigenvalue λ_0 is independent of $b_0 \geq 0$ (cf. [4; Theorem 3.4] or [12; Theorem 3.7]).

For a large part of the paper, we do not use the uniqueness condition (1.2) (note that a change of a finite number of the rates a_i and b_i does not interfere in the uniqueness). Under (1.2), the process is ergodic iff $\sum_i \mu_i < \infty$ (see [10; Theorem 4.45 (2)], for instance). If (1.2) fails but $N = \infty$, then the decay rate for the minimal process is delayed to Section 7. In (2.2), the condition “ $f \in \mathcal{K}$ ” means that we deal with the minimal process. This condition is removed in (4.2). It means that we are in this section dealing with the maximal process in the sense that the domain $\mathcal{D}^{\max}(D)$ of D ignored in (4.2) is taken to be the largest one: $\{f \in L^2(\mu) : D(f) < \infty\}$ (that is the maximal process described at the beginning of Section 6 but killed at 1). When $N = \infty$, even though there is now a killing at 1 (i.e., $a_1 > 0$), the regularity for (or the uniqueness of) the Dirichlet form is still equivalent to (1.3):

$$\sum_{k=1}^{\infty} \left(\frac{1}{b_k \mu_k} + \mu_k \right) = \infty \quad (1.3)',$$

since a modification of a finite number of rates does not change the regularity (cf. Theorem 9.22 for further information). In this section and Section 6, starting from any point in E , even though the process can visit every larger state, it will come

back in a finite time. In this sense, the point infinity is regarded as a reflecting boundary.

It is the position to finish the comparison of (4.1) and (4.2). We have seen that $\lambda_0^{(4.1)} = \lambda_0^{(4.2)}$ once $\sum_i \mu_i < \infty$. We now claim that they can be different otherwise. To see this, note that on the one hand, $\lambda_0^{(4.1)} = 0$ if $\sum_i \mu_i = \infty$, as proved above. On the other hand, once (1.3)' holds (in particular, if $\sum_i \mu_i = \infty$, then) by Proposition 1.3, $\lambda_0^{(4.2)}$ coincides with

$$\inf \{D(f) : f \in \mathcal{K}, \mu(f^2) = 1\},$$

which is the one used in (7.1) below and can often be non-zero. Thus, in general, $\lambda_0^{(4.2)} \geq \lambda_0^{(4.1)}$ and they can be different. As will be seen in Theorem 7.1 (2), in the special case that both of the series in (1.3)' are divergent, we have $\lambda_0^{(4.2)} = \lambda_0^{(7.1)} = 0$.

Theorem 4.1. Assume that $\sum_{i=1}^N \mu_i < \infty$. Then the following variational formulas hold for λ_0 defined by one of (4.1)–(4.3).

(1) Difference form:

$$\inf_{v \in \tilde{\mathcal{V}}_1} \sup_{i \in E} \tilde{R}_i(v) = \lambda_0 = \sup_{v \in \mathcal{V}_1} \inf_{i \in E} R_i(v).$$

(2) Single summation form:

$$\inf_{f \in \tilde{\mathcal{F}}_I} \sup_{i \in E} I_i(f)^{-1} = \lambda_0 = \sup_{f \in \mathcal{F}_I} \inf_{i \in E} I_i(f)^{-1},$$

(3) Double summation form:

$$\begin{aligned} \lambda_0 &= \sup_{f \in \tilde{\mathcal{F}}_{II}} \inf_{i \in E} II_i(f)^{-1} = \sup_{f \in \mathcal{F}_I} \inf_{i \in E} II_i(f)^{-1}, \\ \lambda_0 &= \inf_{f \in \tilde{\mathcal{F}}_I} \sup_{i \in E} II_i(f)^{-1} = \inf_{f \in \tilde{\mathcal{F}}_{II}} \sup_{i \in E} II_i(f)^{-1} = \inf_{f \in \tilde{\mathcal{F}}_{II} \cup \tilde{\mathcal{F}}'_I} \sup_{i \in E} II_i(f)^{-1}, \end{aligned}$$

where $\tilde{\mathcal{F}}'_I = \{f : f_i > 0 \text{ for all } i \in E \text{ and } fII(f) \in L^2(\mu)\}$.

The next result was proved in [6] except the exceptional case that $\sum_i \mu_i = \infty$ in which case $\lambda_0 = 0$ (and $\delta = \infty$) and so the assertion is trivial. See also Corollary 5.2 below. Note that (ν_j) below is different from (2.15).

Theorem 4.2 (Criterion and basic estimates). The rate λ_0 defined by (4.1) (or equivalently by (4.2) provided $\sum_{i \in E} \mu_i < \infty$) is positive iff $\delta < \infty$, where

$$\delta = \sup_{n \in E} \nu[1, n] \mu[n, N] = \sup_{n \in E} \sum_{i=1}^n \frac{1}{\mu_i a_i} \sum_{j=n}^N \mu_j. \quad (4.4)$$

More precisely, we have $(4\delta)^{-1} \leq \lambda_0 \leq \delta^{-1}$. In particular, we have $\lambda_0 = 0$ if $\sum_{i \in E} \mu_i = \infty$ and $\lambda_0 > 0$ if either $N < \infty$ or (1.3)' fails.

Theorem 4.3 (Approximating procedure). Assume that $\sum_{i=1}^N \mu_i < \infty$ and $\delta < \infty$. Write $\varphi_0 = 0$, $\varphi_i = \nu[1, i] := \sum_{j=1}^i (\mu_j a_j)^{-1}$, $i \in E$.

- (1) Define $f_1 = \sqrt{\varphi}$, $f_n = f_{n-1} II(f_{n-1})$ and $\delta_n = \sup_{i \in E} II_i(f_n)$. Then δ_n is decreasing in n and

$$\lambda_0 \geq \delta_\infty^{-1} \geq \dots \geq \delta_1^{-1} \geq (4\delta)^{-1}.$$

- (2) For fixed $m \in E$, define

$$\begin{aligned} f_1^{(m)} &= \varphi(\cdot \wedge m), \\ f_n^{(m)} &= [f_{n-1}^{(m)} II(f_{n-1}^{(m)})](\cdot \wedge m), \quad n \geq 2 \end{aligned}$$

and then define $\delta'_n = \sup_{m \in E} \inf_{i \in E} II_i(f_n^{(m)})$. Then δ'_n is increasing in n and

$$\delta^{-1} \geq \delta_1'^{-1} \geq \dots \geq \delta_\infty'^{-1} \geq \lambda_0.$$

Next, define

$$\bar{\delta}_n = \sup_{m \in E} \frac{\mu(f_n^{(m)})^2}{D(f_n^{(m)})}, \quad n \in E.$$

Then $\bar{\delta}_n^{-1} \geq \lambda_0$, $\bar{\delta}_{n+1} \geq \delta'_n$ for all $n \geq 1$ and $\bar{\delta}_1 = \delta'_1$.

As the first step given in Theorem 4.3, we obtain the following improvement of Theorem 4.2.

Corollary 4.4 (Improved estimates). For the rate λ_0 defined by (4.1) (or equivalently by (4.2) provided $\sum_{i \in E} \mu_i < \infty$), we have

$$\delta^{-1} \geq \delta_1'^{-1} \geq \lambda_0 \geq \delta_1^{-1} \geq (4\delta)^{-1},$$

where

$$\begin{aligned} \delta_1 &= \sup_{i \in E} \frac{1}{\sqrt{\varphi_i}} \sum_{k \geq 1} \mu_k \varphi_{i \wedge k} \sqrt{\varphi_k} \\ &= \sup_{i \in E} \left[\frac{1}{\sqrt{\varphi_i}} \sum_{1 \leq k < i} \mu_k \varphi_k^{3/2} + \sqrt{\varphi_i} \sum_{k=i}^N \mu_k \sqrt{\varphi_k} \right], \end{aligned} \quad (4.5)$$

$$\delta_1' = \sup_{m \in E} \frac{1}{\varphi_m} \sum_{k=1}^N \mu_k \varphi_{k \wedge m}^2 = \sup_{m \in E} \left[\frac{1}{\varphi_m} \sum_{k=1}^{m-1} \mu_k \varphi_k^2 + \varphi_m \mu[m, N] \right] \in [\delta, 2\delta]. \quad (4.6)$$

Proof of Theorem 4.1. Note that

$$I_i(f) = \frac{1}{\mu_i a_i (f_i - f_{i-1})} \sum_{j=i}^N \mu_j f_j = \frac{1}{\mu_{i-1} b_{i-1} (f_i - f_{i-1})} \sum_{j=i}^N \mu_j f_j.$$

Hence, $I_i(f)$ coincides with $I_{i-1}(f)$ used in [3, 4, 6, 7, 12], whenever $b_0 > 0$. The same change is made for the operator $II(f)$ in this section.

Throughout this proof, we use $\lambda_0 = \lambda_0^{(4.3)}$ to denote the one given in (4.3). Similar to the proofs of Theorem 2.4 and Proposition 2.5, we adopt the following circle arguments:

$$\lambda_0 \geq \lambda_0^{(4.2)} \quad (4.7)$$

$$\geq \sup_{f \in \mathcal{F}_I} \inf_{i \in E} II_i(f)^{-1} = \sup_{f \in \mathcal{F}_I} \inf_{i \in E} II_i(f)^{-1} = \sup_{f \in \mathcal{F}_I} \inf_{i \in E} I_i(f)^{-1} \quad (4.8)$$

$$\geq \sup_{v \in \mathcal{V}_1} \inf_{i \in E} R_i(v) \quad (4.9)$$

$$\geq \lambda_0 \quad (4.10)$$

and

$$\lambda_0 \leq \inf_{f \in \tilde{\mathcal{F}}_I \cup \tilde{\mathcal{F}}'_I} \sup_{i \in E} II_i(f)^{-1} \quad (4.11)$$

$$\leq \inf_{f \in \tilde{\mathcal{F}}_I} \sup_{i \in E} II_i(f)^{-1} = \inf_{f \in \tilde{\mathcal{F}}_I} \sup_{i \in E} II_i(f)^{-1} = \inf_{f \in \tilde{\mathcal{F}}_I} \sup_{i \in E} I_i(f)^{-1} \quad (4.12)$$

$$\leq \inf_{v \in \tilde{\mathcal{V}}_1} \sup_{i \in E} \tilde{R}_i(v) \quad (4.13)$$

$$\leq \lambda_0 \quad (4.14)$$

Assertion (4.7) is obvious. The following assertions are proved in [4; Theorem 3.3], or [12; §3.8] and [7; §2] (see also the remark given in the next paragraph):

$$\sup_{f \in \mathcal{F}_I} \inf_{i \in E} I_i(f)^{-1} = \sup_{f \in \mathcal{F}_I} \inf_{i \in E} II_i(f)^{-1} = \sup_{f \in \mathcal{F}_I} \inf_{i \in E} II_i(f)^{-1} \leq \lambda_0^{(4.2)}. \quad (4.15)$$

$$\inf_{f \in \tilde{\mathcal{F}}_I} \sup_{i \in E} I_i(f)^{-1} = \inf_{f \in \tilde{\mathcal{F}}_I} \sup_{i \in E} II_i(f)^{-1} = \inf_{f \in \tilde{\mathcal{F}}_I} \sup_{i \in E} II_i(f)^{-1}. \quad (4.16)$$

In particular, we have known (4.8) and (4.12) since the inequality in (4.12) is trivial. It remains to prove (4.9)–(4.11), (4.13) and (4.14).

In [7; §2] and [12; §3.8], only the ergodic case under condition (1.2) is considered. But for (4.15) and (4.16), one does not need (1.2). Actually, one can now follow the proofs of Theorem 2.4 and Proposition 2.5 with a little change. For instance, to prove the last inequality in (4.15), following proof (a) of Theorem 2.4 and Proposition 2.5, let g satisfy $\|g\| = 1$ and $g_0 = 0$. Then

$$\begin{aligned} 1 &= \sum_{i \in E} \mu_i g_i^2 \quad (\text{since } \|g\| = 1) \\ &= \sum_{i \in E} \mu_i \left(\sum_{j=1}^i (g_j - g_{j-1}) \right)^2 \quad (\text{since } g_0 = 0) \\ &\leq \sum_{i \in E} \mu_i \sum_{j=1}^i \frac{(g_j - g_{j-1})^2 \mu_j a_j}{h_j} \sum_{k=1}^i \frac{h_k}{\mu_k a_k}. \end{aligned}$$

Exchanging the order of the first two sums on the right-hand side, we get

$$\begin{aligned} 1 &\leq \sum_{j \in E} \mu_j a_j (g_j - g_{j-1})^2 \frac{1}{h_j} \sum_{i=j}^N \mu_i \sum_{k=1}^i \frac{h_k}{\mu_k a_k} \\ &\leq D(g) \sup_{j \in E} \frac{1}{h_j} \sum_{i=j}^N \mu_i \sum_{k=1}^i \frac{h_k}{\mu_k a_k} \\ &=: D(g) \sup_{j \in E} H_j. \end{aligned}$$

The next step is to choose $h_j = \sum_{i=j}^N \mu_i f_i$ for a given $f \in \mathcal{F}_H$ with $\sup_{j \in E} H_j(f) < \infty$. From these, it should be clear what change is required in order to prove (4.15) and (4.16).

We now begin to work on the additional part of the proof.

(a) Prove that $\sup_{f \in \mathcal{F}_H} \inf_{i \in E} H_i(f)^{-1} \geq \sup_{v \in \mathcal{V}_1} \inf_{i \in E} R_i(v)$.

As in proof (c) of Theorem 2.4 and Proposition 2.5, we use $R_i(u)$,

$$R_i(u) = a_i \left(1 - \frac{u_{i-1}}{u_i} \right) + b_i \left(1 - \frac{u_{i+1}}{u_i} \right), \quad i \in E,$$

(u_{N+1} is free if $N < \infty$ since $b_N = 0$), instead of $R_i(v)$, where $u_i > 0$ for $i \in E$ and $u_0 = 0$. Then $v_i > 1$ ($i \in E$) means that $u_{i+1} > u_i > 0$, and $v_i = 1$ for $i \geq m$ means that $u_i = u_{i \wedge m} > 0$.

Without loss of generality, assume that $\inf_{i \in E} R_i(u) > 0$ for a given strictly increasing u with $u_0 = 0$. Define $f_i = (a_i + b_i)u_i - a_i u_{i-1} - b_i u_{i+1}$ [$= u_i R_i(u)$] for $i \in E$ and $f_0 = 0$. Then by assumption,

$$f_i/u_i = R_i(u) > 0, \quad i \in E.$$

Hence, $f \in \mathcal{F}_H$. Next, since

$$0 < \mu_k f_k = \mu_k a_k (u_k - u_{k-1}) - \mu_{k+1} a_{k+1} (u_{k+1} - u_k)$$

and the strictly increasing property of u_i in i , it follows that

$$0 < \sum_{k=j}^N \mu_k f_k \leq \mu_j a_j (u_j - u_{j-1}),$$

and so

$$u_i = \sum_{j=1}^i (u_j - u_{j-1}) \geq \sum_{j=1}^i \nu_j \sum_{k=j}^N \mu_k f_k > 0.$$

We obtain

$$R_i(u) = f_i/u_i \leq H_i(f)^{-1}, \quad i \in E.$$

Therefore, we have first

$$\inf_{i \in E} R_i(u) \leq \inf_{i \in E} \Pi_i(f)^{-1} \leq \sup_{f \in \mathcal{F}_H} \inf_{i \in E} \Pi_i(f)^{-1},$$

and then

$$\sup_{v \in \mathcal{V}_1} \inf_{i \in E} R_i(v) \leq \sup_{f \in \mathcal{F}_H} \inf_{i \in E} \Pi_i(f)^{-1},$$

as required.

(b) Prove that $\sup_{v \in \mathcal{V}_1} \inf_{i \in E} R_i(v) \geq \lambda_0$.

First, we show that $\sup_{v \in \mathcal{V}_1} \inf_{i \in E} R_i(v) \geq 0$. For a given positive $f \in L^1(\mu)$, let $u = f\Pi(f)$. Then $u_{i+1}/u_i > 1$ and $R_i(u) = f_i/u_i > 0$ for all $i \in E$. With $(v_i = u_{i+1}/u_i) \in \mathcal{V}_1$, this implies $\inf_{i \in E} R_i(v) \geq 0$ and then the required assertion follows.

Alternatively, since $a_1 > 0$, the eigenfunction is still strictly increasing when $\lambda_0 = 0$ by part (3) of Proposition 2.1. Hence the proof in the case of $\lambda_0 = 0$ can be combined into the next paragraph, and then the last paragraph can be omitted.

By assumption, we have $\sum_{i \in E} \mu_i < \infty$. When $\lambda_0 > 0$, it was proved in proof (d) of [12; Theorem 3.7] that the eigenfunction of $\lambda_0^{(4.2)}$ is strictly increasing. Even though λ_0 could formally be bigger than $\lambda_0^{(4.2)}$, the same proof still works for the eigenfunction g of λ_0 since the modified function $\bar{g}$ used there satisfies $\bar{g}_i = \bar{g}_{i \wedge n}$ for some n . Having this at hand, the proof is just a use of the eigenequation:

$$-\Omega g(i) := -b_i(g_{i+1} - g_i) + a_i(g_i - g_{i-1}) = \lambda_0 g_i, \quad i \in E, \quad g_0 := 0$$

(g_{N+1} is free if $N < \infty$ since $b_N = 0$). With $v_i := g_{i+1}/g_i > 1$ for $i < N$, this gives us $v \in \mathcal{V}_1$ and $R_i(v) \equiv \lambda_0$, and so the assertion follows.

We have thus completed the circle argument of (4.7)–(4.10).

(c) Prove that $\lambda_0 \leq \inf_{f \in \mathcal{F}_H \cup \mathcal{F}'_H} \sup_{i \in E} \Pi_i(f)^{-1}$.

In the original proof of [7; Theorem 2.1], when $N = \infty$, from the estimate

$$D(g) \leq \mu(g^2) \sup_{i \in E} \Pi_i(f)^{-1}$$

for $f \in \widetilde{\mathcal{F}}_H$ and $g := [f\Pi(f)](\cdot \wedge m)$ to conclude that $\lambda_0 \leq D(g)/\mu(g^2)$, one requires an additional condition $g \in L^2(\mu)$, provided $m = \infty$ is allowed. This is the reason why the set $\widetilde{\mathcal{F}}'_H$ in part (3) of Theorem 4.1 is added. Anyhow, with the modified conditions, the same proof gives us the required assertion (cf. proof (f) of Theorem 2.4 and Proposition 2.5).

(d) Prove that $\inf_{f \in \mathcal{F}_H} \sup_{i \in E} \Pi_i(f)^{-1} \leq \inf_{v \in \mathcal{V}_1} \sup_{i \in E} \widetilde{R}_i(v)$.

Given u with $u_0 = 0$ and $u_i = u_{i \wedge m}$ for all $i \in E$ so that $(v_i := u_{i+1}/u_i) \in \widetilde{\mathcal{V}}_1$, let

$$f_i = \begin{cases} (a_i + b_i)u_i - a_i u_{i-1} - b_i u_{i+1}, & i \leq m-1 \\ \tilde{a}_m(u_m - u_{m-1}), & i \geq m. \end{cases}$$

It is simple to check that $f_0 = 0$,

$$f_i/u_i = \tilde{R}_i(u) > 0 \text{ for } i \in \{1, \dots, m\} \text{ and } f_i = f_m \text{ for } i > m,$$

and so $f \in \widetilde{\mathcal{F}}_H$. Moreover, since

$$\begin{aligned} \sum_{k=j}^{m-1} \mu_k f_k &= \mu_j a_j (u_j - u_{j-1}) - \mu_m a_m (u_m - u_{m-1}) \\ &= \mu_j a_j (u_j - u_{j-1}) - f_m \sum_{k=m}^N \mu_k, \end{aligned}$$

we get

$$0 < \sum_{k=j}^N \mu_k f_k = \mu_j a_j (u_j - u_{j-1}).$$

It follows that

$$0 < u_i = \sum_{j=1}^i (u_j - u_{j-1}) = \sum_{j=1}^i \nu_j \sum_{k=j}^N \mu_k f_k, \quad i \in \{1, \dots, m\},$$

and then $\tilde{R}_i(u) = f_i/u_i = H_i(f)^{-1}$ for $i \in \{1, 2, \dots, m\}$. Therefore, we have

$$\max_{1 \leq i \leq m} \tilde{R}_i(u) = \max_{1 \leq i \leq m} H_i(f)^{-1} \geq \inf_{f \in \widetilde{\mathcal{F}}_H, f_i = f_{i \wedge m}} \max_{1 \leq i \leq m} H_i(f)^{-1} \geq \inf_{f \in \widetilde{\mathcal{F}}_H} \sup_{i \in E} H_i(f)^{-1},$$

and then

$$\inf_{v \in \mathcal{V}_1} \sup_{i \in E} \tilde{R}_i(v) \geq \inf_{f \in \widetilde{\mathcal{F}}_H} \sup_{i \in E} H_i(f)^{-1}.$$

(e) Prove that $\inf_{v \in \mathcal{V}_1} \sup_{i \in E} \tilde{R}_i(v) \leq \lambda_0$.

Recall the definition of λ_0 :

$$\lambda_0 = \inf \{ D(f) : \mu(f^2) = 1, f_i = f_{i \wedge m} \text{ for some } m \in E \text{ and all } i \in E \}.$$

Clearly, we have

$$\lambda_0^{(m)} := \inf \{ D(f) : \mu(f^2) = 1, f_i = f_{i \wedge m} \text{ for all } i \in E \} \downarrow \lambda_0 \text{ as } m \uparrow N.$$

We now explain the meaning of $\lambda_0^{(m)}$ as follows. Let

$$\begin{aligned} \tilde{\mu}_i &= \mu_i, \quad 1 \leq i < m, \quad \tilde{\mu}_m = \sum_{i=m}^N \mu_i, \\ \tilde{a}_i &= a_i, \quad 1 \leq i < m, \quad \tilde{a}_m = \mu_m a_m / \tilde{\mu}_m, \\ \tilde{D}(f) &= \sum_{i=1}^m \tilde{\mu}_i \tilde{a}_i (f_i - f_{i-1})^2. \end{aligned} \tag{4.17}$$

Then $\tilde{\mu}_i \tilde{a}_i = \mu_i a_i$ for $i = 1, \dots, m$, $\tilde{D}(f) = D(f)$ and $\tilde{\mu}(f^2) = \mu(f^2)$ for every f with $f = f_{\cdot \wedge m}$. Thus, $\lambda_0^{(m)}$ is just the first eigenvalue of the local Dirichlet form $(\tilde{D}, \mathcal{D}(\tilde{D}))$ having the state space $\{1, \dots, m\}$, with Dirichlet (absorbing) boundary at 0 and Neumann (reflecting) boundary at m . Let g ($g_0 = 0$) be the eigenfunction of the local first eigenvalue $\lambda_0^{(m)}$. Extend g to the whole space by setting $g_i = g_{i \wedge m}$. Next, set $u_i = g_i$ for $i < N$. Then

$$\tilde{R}_i(u) = \begin{cases} \lambda_0^{(m)} > 0, & i \in \{1, \dots, m\}, \\ 0, & i > m. \end{cases} \quad (4.18)$$

Furthermore, for $v_i := u_{i+1}/u_i$, we have $v_0 = \infty$, $v_i > 1$ on $\{1, \dots, m-1\}$, and $v_i = 1$ for $i \geq m$. Thus, by (4.18), it is easy to check that $v = (v_i) \in \tilde{\mathcal{V}}_1$. Therefore,

$$\begin{aligned} \lambda_0^{(m)} &= \max_{1 \leq i \leq m} \tilde{R}_i(v) \\ &\geq \inf_{v \in \tilde{\mathcal{V}}_1: v_i=1 \text{ for } i \geq m} \max_{1 \leq i \leq m} \tilde{R}_i(v) \\ &\geq \inf_{v \in \tilde{\mathcal{V}}_1: v_i=1 \text{ for } i \geq \text{some } n > 1} \sup_{i \in E} \tilde{R}_i(v) \\ &= \inf_{v \in \tilde{\mathcal{V}}_1} \sup_{i \in E} \tilde{R}_i(v). \end{aligned}$$

The assertion now follows by letting $m \rightarrow N$.

We have thus completed the circle arguments (4.11)–(4.14), and then completed the proof of Theorem 4.1. $\square$

Proof of Theorem 4.3.

(a) We remark that the sequence $\{f_n^{(m)}\}_{n \in E}$ is clearly contained in $\tilde{\mathcal{F}}_I$. But the modified sequence used in [7; Theorem 2.2],

$$\tilde{f}_1^{(m)} = \varphi(\cdot \wedge m), \quad \tilde{f}_n^{(m)} = \tilde{f}_{n-1}^{(m)}(\cdot \wedge m) \Pi(\tilde{f}_{n-1}^{(m)}(\cdot \wedge m)), \quad n \geq 2,$$

is usually not contained in $\tilde{\mathcal{F}}_{II}$. However,

$$\begin{aligned} \delta'_n &= \sup_{m \in E} \inf_{i \in E} II_i(f_n^{(m)}) \\ &= \sup_{m \in E} \min_{1 \leq i \leq m} II_i(f_n^{(m)}) \\ &= \sup_{m \in E} \min_{1 \leq i \leq m} II_i(\tilde{f}_n^{(m)}(\cdot \wedge m)) \\ &= \sup_{m \in E} \inf_{i \in E} II_i(\tilde{f}_n^{(m)}(\cdot \wedge m)). \end{aligned}$$

Here in the last step, we have used the convention $1/0 = \infty$. Hence, these two sequences produce the same $\{\delta'_n\}$.

(b) The approximating procedure given in Theorem 4.3 is mainly a copy of [7; Theorem 2.2] (cf. the proof of Theorem 3.2). For later use, here we review the proof of part (1). From [6; proof of Theorem 3.5], we have known that

$$I_j(f_1) = \frac{1}{\mu_j a_j (f_1(j) - f_1(j-1))} \sum_{k \geq j} \mu_k f_1(k) \leq 4\delta, \quad j \geq 1.$$

Hence (Alternatively, by the proportional property),

$$f_2(i) = \sum_{j=1}^i \frac{1}{\mu_j a_j} \sum_{k \geq j} \mu_k f_1(k) \leq 4\delta \sum_{j=1}^i (f_1(j) - f_1(j-1)) = 4\delta f_1(i).$$

This gives us the assertion $\delta_1 = \sup_{i \geq 1} II_i(f_1) \leq 4\delta$.

To prove the monotonicity of $\{\delta_n\}$ and $\{f_n\} \subset L^1(\mu)$, we adopt induction. As we have just seen,

$$\delta_1 = \sup_{i \geq 1} \frac{f_2(i)}{f_1(i)} \leq 4\delta.$$

This means that $f_1 \in L^1(\mu)$ (or equivalently, $f_2 < \infty$) and $\delta_1 < \infty$ since $\delta < \infty$ by assumption. Assume that $f_n \in L^1(\mu)$ (or equivalently, $f_{n+1} < \infty$) and $\delta_n < \infty$. Then

$$\sum_{k \geq j} \mu_k f_{n+1}(k) = \sum_{k \geq j} \mu_k f_n(k) [f_{n+1}(k)/f_n(k)] \leq \delta_n \sum_{k \geq j} \mu_k f_n(k).$$

Multiplying both sides by ν_j and making summation from 1 to i , it follows that

$$f_{n+2}(i) \leq \delta_n f_{n+1}(i), \quad i \geq 1.$$

Since $f_{n+1} < \infty$ and $\delta_n < \infty$ by assumption, we have $f_{n+2} < \infty$, and

$$\frac{f_{n+2}(i)}{f_{n+1}(i)} = II_i(f_{n+1}) < \infty, \quad i \geq 1.$$

This proves not only $f_{n+1} \in L^1(\mu)$ but also $\delta_{n+1} \leq \delta_n < \infty$.

The assertion that $\delta_n^{-1} \geq \lambda_0$ is obvious by (4.2). Similar to proof (b) of Theorem 3.2, the assertion $\delta_{n+1} \geq \delta'_n$ is a consequence of the last part of the proof of [7; Theorem 2.1]. $\square$

Proof of Corollary 4.4.

(a) The degenerated case that $\sum_i \mu_i = \infty$ is trivial since $\lambda_0^{(4.1)} = 0$ and $\delta = \delta_1 = \delta'_1 = \infty$. The main assertion of Corollary 4.4 is a consequence of Theorem 4.3. Here, we consider (4.6) only since the proof of (4.5) is easier. Note that

$$II_i(f_1^{(m)}) = \frac{1}{\varphi_{i \wedge m}} \sum_{j=1}^i \frac{1}{\mu_j a_j} \sum_{k=j}^N \mu_k \varphi_{k \wedge m}.$$

The right-hand side is clearly increasing in i for $i \geq m$ and is decreasing (not hard to check) in i when $i \leq m$. Hence, $II_i(f_1^{(m)})$ achieves its minimum at $i = m$. Then, by exchanging the order of the sums, it follows that the minimum is equal to

$$\frac{1}{\varphi_m} \sum_{k=1}^N \mu_k \varphi_{k \wedge m}^2.$$

This observation is due to Sirl, Zhang and Pollett (2007). We have thus proved the first equality in (4.6).

Next, following the proof of [6; Theorem 3.5], we have

$$D(f_1^{(m)}) = \sum_{i=1}^m \mu_i a_i (\varphi_i - \varphi_{i-1})^2 = \varphi_m,$$

and

$$\mu(f_1^{(m)2}) = \sum_{k=1}^N \mu_k \varphi_{k \wedge m}^2.$$

Combining these facts together, it follows that $\bar{\delta}_1 = \delta'_1$.

(b) Finally, we prove the estimates in (4.6). The lower estimate of δ'_1 is rather easy since

$$\frac{1}{\varphi_m} \sum_{k=1}^N \mu_k \varphi_{k \wedge m}^2 \geq \frac{1}{\varphi_m} \sum_{k=m}^N \mu_k \varphi_{k \wedge m}^2 = \varphi_m \sum_{k=m}^N \mu_k.$$

For the upper estimate, use the summation by parts formula:

$$\sum_{k=1}^N \mu_k \varphi_{k \wedge m}^2 = \sum_{k=1}^m [\varphi_k^2 - \varphi_{k-1}^2] \sum_{j=k}^N \mu_j = \sum_{k=1}^m \frac{\varphi_k + \varphi_{k-1}}{\mu_k a_k} \sum_{j=k}^N \mu_j.$$

It follows that

$$\frac{1}{\varphi_m} \sum_{k=1}^N \mu_k \varphi_{k \wedge m}^2 < \frac{2}{\varphi_m} \sum_{k=1}^m \frac{1}{\mu_k a_k} \left[\varphi_k \sum_{j=k}^N \mu_j \right] \leq \frac{2\delta}{\varphi_m} \sum_{k=1}^m \frac{1}{\mu_k a_k} = 2\delta.$$

The estimate now follows by making the supremum with respect to $m \in E$. $\square$

5. DUAL APPROACH

This section is devoted to the duality of the processes studied in the previous sections, as well as a duality to be used in the next two sections. Again, the section is ended by a class of examples.

Suppose that we are given a birth-death process with state space $E = \{i : 0 \leq i < N+1\}$ ($N \leq \infty$), birth rates $b_i > 0$ ($b_0 > 0$, especially) but $b_N \geq 0$ if $N < \infty$, and death rates $a_i > 0$ but $a_0 = 0$. The case that $b_N > 0$ is used in this section while the case of $b_N = 0$ is for use in Section 7. Define a dual chain with state space $\hat{E} = \{i : 1 \leq i < N'+1\}$ and with rates as follows:

$$\hat{b}_0 = 0, \quad \hat{b}_i = a_i, \quad \hat{a}_i = b_{i-1}, \quad i \in \hat{E}, \quad (5.1)$$

where $a_{N+1} = b_{N+1} = 0$ if $N < \infty$ by convention and

$$N' = \begin{cases} N, & N < \infty \text{ and } b_N = 0, \\ N+1, & N < \infty \text{ and } b_N > 0, \\ \infty, & N = \infty. \end{cases}$$

The dual process with rates $(\hat{a}_i, \hat{b}_i)$ has an absorbing at 0. When $N < \infty$, for the dual process, the state $N + 1$ is absorbing if $b_N = 0$ (then $\hat{a}_{N+1} = 0$ but $\hat{b}_N > 0$); otherwise, it is a reflecting boundary since $\hat{a}_{N+1} = b_N > 0$. In a word, the absorbing boundary is dual to the reflecting one and vice versa. This dual technique goes back to Karlin and McGregor (1957b, §6). Next, define

$$\hat{\mu}_1 = 1, \quad \hat{\mu}_n = \frac{\hat{b}_1 \cdots \hat{b}_{n-1}}{\hat{a}_2 \cdots \hat{a}_n}, \quad 2 \leq n < N' + 1. \quad (5.2)$$

When $N < \infty$ and $b_N > 0$, then $\hat{a}_{N+1} > 0$, and so $\hat{\mu}_n$ can be defined up to $n = N + 1$. Otherwise, it can be defined up to $n = N$ only. It is now easy to check (noticing the difference of (ν_j) and $(\hat{\nu}_j)$) that

$$\hat{\mu}_n = \frac{b_0}{\mu_n a_n} = b_0 \nu_{n-1}, \quad \hat{\nu}_n := \frac{1}{\hat{\mu}_n \hat{a}_n} = \frac{1}{b_0} \mu_{n-1}, \quad 1 \leq n < N' + 1. \quad (5.3)$$

Actually, the rates $(\hat{a}_i, \hat{b}_i)$ in (5.1) are determined by the transform given in (5.3): $\hat{\mu}_n = b_0 \nu_{n-1}$ and $\hat{\nu}_n = \mu_{n-1}/b_0$. From this, it follows that

$$\begin{aligned} \mu_n &= b_0 \hat{\nu}_{n+1} = \hat{a}_1 \hat{\nu}_{n+1}, \quad \nu_n = \frac{1}{b_0} \hat{\mu}_{n+1} = \frac{1}{\hat{a}_1} \hat{\mu}_{n+1}, \quad 0 \leq n < N', \\ \mu_N &= \hat{a}_1 (\hat{\mu}_N \hat{b}_N)^{-1} \quad \text{if } N < \infty \text{ and } b_N = 0, \end{aligned} \quad (5.4)$$

and so

$$\begin{aligned} \sum_{n=1}^{N'} \frac{1}{\hat{\mu}_n \hat{a}_n} &= \sum_{n=1}^{N'} \hat{\nu}_n = \frac{1}{b_0} \sum_{n=0}^{N'-1} \mu_n, \\ \sum_{n=1}^{N'} \hat{\mu}_n &= b_0 \sum_{n=1}^{N'} \nu_{n-1} = b_0 \sum_{n=0}^{N'-1} \frac{1}{\mu_n b_n}. \end{aligned} \quad (5.5)$$

Note that by (5.1),

$$a_{i+1} + b_i - \frac{a_i}{v_{i-1}} - b_{i+1} v_i = \hat{b}_{i+1} + \hat{a}_{i+1} - \frac{\hat{b}_i}{v_{i-1}} - \hat{a}_{i+2} v_i.$$

By a change of the variables $(v_i) \in \mathcal{V}$:

$$v_i = \frac{\hat{b}_{i+1}}{\hat{a}_{i+2}} \hat{v}_{i+1}, \quad (5.6)$$

or

$$\hat{v}_i = \frac{\hat{a}_{i+1}}{\hat{b}_i} v_{i-1} = \frac{b_i}{a_i} v_{i-1}, \quad (5.7)$$

we get

$$a_{i+1} + b_i - \frac{a_i}{v_{i-1}} - b_{i+1} v_i = \hat{a}_{i+1} \left(1 - \frac{1}{\hat{v}_i}\right) + \hat{b}_{i+1} (1 - \hat{v}_{i+1}).$$

Since $b_0 > 0$, $v_{-1} > 0$ but $a_0 = 0$, from (5.7), it is clear that we should set $\hat{v}_0 = \infty$. Next, by (5.7) again,

$$v_{i-1} > \frac{a_i}{b_i} \iff \hat{v}_i > 1.$$

It remains to examine the boundary condition on the right-hand side when $N < \infty$.

- (1) First, let $b_N = 0$. Then $v = (v_i > 0 : 0 \leq i < N - 1)$, v_{-1} and v_{N-1} are free. The dual state space is $\hat{E} = \{1, 2, \dots, N\}$. The dual test sequence is $\hat{v} = (\hat{v}_i > 0 : 1 \leq i < N)$, $\hat{v}_N = 0$.
- (2) Next, let $b_N > 0$. Then $v = (v_i > 0 : 0 \leq i < N)$, v_{-1} and v_N are free. The dual state space is $\hat{E} = \{1, 2, \dots, N + 1\}$ with reflecting at $N + 1$. Hence, $\hat{v} = (\hat{v}_i > 0 : 1 \leq i < N + 1)$, $\hat{v}_{N+1} = 0$.

We have thus proved the following result.

Proposition 5.1. For the dual processes defined above, the following identities hold:

$$\begin{aligned} & \sup_v \inf_{0 \leq i < N'} \left[a_{i+1} + b_i - \frac{a_i}{v_{i-1}} - b_{i+1}v_i \right] \\ &= \sup_{\hat{v}} \inf_{1 \leq i < N'+1} \left[\hat{a}_i \left(1 - \frac{1}{\hat{v}_{i-1}} \right) + \hat{b}_i(1 - \hat{v}_i) \right], \end{aligned} \quad (5.8)$$

where $v = (v_i > 0 : 0 \leq i < N' - 1)$ with free v_{-1} , and $\hat{v} = (\hat{v}_i > 0 : 1 \leq i < N')$ with $\hat{v}_0 = \infty$, $v_{N'-1}$ is free and $\hat{v}_{N'} = 0$ if $N < \infty$;

$$\begin{aligned} & \sup_{v \in \mathcal{V}_*} \inf_{0 \leq i < N+1} \left[a_{i+1} + b_i - \frac{a_i}{v_{i-1}} - b_{i+1}v_i \right] \\ &= \sup_{\hat{v} \in \mathcal{V}_1} \inf_{1 \leq i < N+2} \left[\hat{a}_i \left(1 - \frac{1}{\hat{v}_{i-1}} \right) + \hat{b}_i(1 - \hat{v}_i) \right] \end{aligned} \quad (5.9)$$

in the case that $b_N > 0$ if $N < \infty$, where $\mathcal{V}_*$ is given in Proposition 2.7, and $\mathcal{V}_1$ is defined in Theorem 4.1 replacing N by $N + 1$ when $N < \infty$.

In these formulas, $a_{N+1} = b_{N+1} = 0$ if $N < \infty$ by convention.

Corollary 5.2. Given rates (a_i, b_i) as in Section 2 (then $b_N > 0$ if $N < \infty$), let $\lambda_0 = \lambda_0^{(2.2)}$ and define δ by (3.1). Next, define the dual rates $(\hat{a}_i, \hat{b}_i)$ as above. Correspondingly, we have $\hat{\lambda}_0$ and $\hat{\delta}$ defined by (4.1) and (4.4) replacing N by $N + 1$ if $N < \infty$, respectively, in terms of the dual rates. Then we have $\lambda_0 = \hat{\lambda}_0$ and $\delta = \hat{\delta}$.

Proof. Having relationship (5.9) at hand, the assertion that $\lambda_0 = \hat{\lambda}_0$ follows by a combination part (2) of Proposition 2.7 and part (1) of Theorem 4.1, provided $\sum_i \hat{\mu}_i < \infty$.

Next, by (4.4), (5.3), and (3.1), we have

$$\begin{aligned}
 \hat{\delta} &= \sup_{1 \leq i < N+2} \sum_{j=1}^i \frac{1}{\hat{\mu}_j \hat{a}_j} \sum_{k=i}^{N+1} \hat{\mu}_k \\
 &= \sup_{1 \leq i < N+2} \sum_{j=1}^i \hat{\nu}_j \sum_{k=i}^{N+1} \hat{\mu}_k \\
 &= \sup_{1 \leq i < N+2} \sum_{j=1}^i \frac{\mu_{j-1}}{b_0} \sum_{k=i}^{N+1} b_0 \nu_{k-1} \\
 &= \sup_{0 \leq i < N+1} \sum_{j=0}^i \mu_j \sum_{k=i}^N \frac{1}{\mu_k b_k} \\
 &= \delta.
 \end{aligned}$$

This proves that $\delta = \hat{\delta}$. In particular, if $\sum_i \hat{\mu}_i (= \sum_i \nu_i) = \infty$, then by Theorem 3.1 and Corollary 4.4, we get $\lambda_0 = \hat{\lambda}_0 = 0$. We have thus completed the proof of $\lambda_0 = \hat{\lambda}_0$. $\square$

As will be seen in Theorem 7.1 (2), in the degenerated case that $\sum_i \mu_i = \infty$ and $\sum_i (\mu_i b_i)^{-1} = \infty$, the dual of the process studied in Section 2 also goes to the one studied in Section 7.

Before moving further, let us discuss the duality used here. Very recently, Chi Zhang provides us a nice explanation which leads to a deeper understanding of the duality (5.1). Consider a simple example as follows:

$$Q = \begin{pmatrix} -b_0 & b_0 & 0 & 0 \\ a_1 & -a_1 - b_1 & b_1 & 0 \\ 0 & a_2 & -a_2 - b_2 & b_2 \\ 0 & 0 & a_3 & -a_3 - b_3 \end{pmatrix}, \quad a_i, b_i > 0.$$

Introduce an invertible matrix:

$$M = \begin{pmatrix} \mu_0 b_0 & -\mu_0 b_0 & 0 & 0 \\ 0 & \mu_1 b_1 & -\mu_1 b_1 & 0 \\ 0 & 0 & \mu_2 b_2 & -\mu_2 b_2 \\ 0 & 0 & 0 & \mu_3 b_3 \end{pmatrix} \Rightarrow M^{-1} = \begin{pmatrix} \frac{1}{\mu_0 b_0} & \frac{1}{\mu_1 b_1} & \frac{1}{\mu_2 b_2} & \frac{1}{\mu_3 b_3} \\ 0 & \frac{1}{\mu_1 b_1} & \frac{1}{\mu_2 b_2} & \frac{1}{\mu_3 b_3} \\ 0 & 0 & \frac{1}{\mu_2 b_2} & \frac{1}{\mu_3 b_3} \\ 0 & 0 & 0 & \frac{1}{\mu_3 b_3} \end{pmatrix}.$$

Then

$$\begin{aligned}
 MQM^{-1} &= \begin{pmatrix} -a_1 - b_0 & a_1 & 0 & 0 \\ b_1 & -a_2 - b_1 & a_2 & 0 \\ 0 & b_2 & -a_3 - b_2 & a_3 \\ 0 & 0 & b_3 & -b_3 \end{pmatrix} \\
 &= \begin{pmatrix} -\hat{a}_1 - \hat{b}_1 & \hat{b}_1 & 0 & 0 \\ \hat{a}_2 & -\hat{a}_2 - \hat{b}_2 & \hat{b}_2 & 0 \\ 0 & \hat{a}_3 & -\hat{a}_3 - \hat{b}_3 & \hat{b}_3 \\ 0 & 0 & \hat{a}_4 & -\hat{a}_4 \end{pmatrix} \\
 &= \hat{Q}.
 \end{aligned}$$

Hence, the dual matrix $\widehat{Q}$ is just the classical similarity transformation of Q and so they have the same spectrum. In particular, the eigenequation $Qg = -\lambda_0 g$ ($g \neq 0$) is transferred into

$$\widehat{Q}(Mg) = (MQM^{-1})(Mg) = -\lambda_0 Mg = -\widehat{\lambda}_0(Mg).$$

Hence, the eigenfunction g of λ_0 is transformed to $\widehat{g} = Mg$ of $\widehat{\lambda}_0 = \lambda_0$. Correspondingly, the test function f is transformed to $\widehat{f} = Mf$. From this, it should be clear that all the operators R and $\widehat{R}$, I and $\widehat{I}$, II and $\widehat{II}$ are closely related to each other and then so are the variational formulas.

Having these facts at hand, one can simplify a part of the previous proofs. However, we prefer to keep all the details here since they are needed when we go to the more general situation, so called the Poincaré-type inequalities (Section 8), or can be used as a reference for studying the continuous case. For the Poincaré-type inequalities, the current duality seems not available.

By Corollary 5.2, we have two ways to estimate $\lambda_0 = \widehat{\lambda}_0$: using either the rates (a_i, b_i) or $(\widehat{a}_i, \widehat{b}_i)$. The corresponding formulas for δ'_1 , $\widehat{\delta}'_1$, δ_1 and $\widehat{\delta}_1$ are collected in Tables 5.1 and 5.2.

Table 5.1: Expressions of $\delta = \widehat{\delta}$, δ'_1 , $\widehat{\delta}'_1$, δ_1 and $\widehat{\delta}_1$ in terms of the rates (b_i, a_i) :

$$\delta = \widehat{\delta} = \sup_{0 \leq i < N+1} \mu[0, i] \nu[i, N] = \sup_{0 \leq i < N+1} \sum_{j=0}^i \mu_j \sum_{k=i}^N \nu_k, \quad (5.10)$$

$$\begin{aligned} \delta'_1 &= \sup_{0 \leq i < N+1} \frac{1}{\nu[i, N]} \sum_{k=0}^N \mu_k \nu[k \vee i, N]^2 \\ &= \sup_{0 \leq i < N+1} \left[\mu[0, i] \nu[i, N] + \frac{1}{\nu[i, N]} \sum_{k=i+1}^N \mu_k \nu[k, N]^2 \right], \end{aligned} \quad (5.11)$$

$$\begin{aligned} \widehat{\delta}'_1 &= \sup_{0 \leq i < N+1} \frac{1}{\mu[0, i]} \sum_{k=0}^N \nu_k \mu[0, k \wedge i]^2 \\ &= \sup_{0 \leq i < N+1} \left[\mu[0, i] \nu[i, N] + \frac{1}{\mu[0, i]} \sum_{k=0}^{i-1} \nu_k \mu[0, k]^2 \right], \end{aligned} \quad (5.12)$$

$$\begin{aligned} \delta_1 &= \sup_{0 \leq i < N+1} \frac{1}{\sqrt{\nu[i, N]}} \sum_{k=0}^N \mu_k \nu[i \vee k, N] \sqrt{\nu[k, N]} \\ &= \sup_{0 \leq i < N+1} \left[\sqrt{\nu[i, N]} \sum_{k=0}^i \mu_k \sqrt{\nu[k, N]} + \frac{1}{\sqrt{\nu[i, N]}} \sum_{k=i+1}^N \mu_k \nu[k, N]^{3/2} \right], \end{aligned} \quad (5.13)$$

$$\begin{aligned} \widehat{\delta}_1 &= \sup_{0 \leq i < N+1} \frac{1}{\sqrt{\mu[0, i]}} \sum_{k=0}^N \nu_k \mu[0, k \wedge i] \sqrt{\mu[0, k]} \\ &= \sup_{0 \leq i < N+1} \left[\frac{1}{\sqrt{\mu[0, i]}} \sum_{k=0}^{i-1} \nu_k \mu[0, k]^{3/2} + \sqrt{\mu[0, i]} \sum_{k=i}^N \nu_k \sqrt{\mu[0, k]} \right]. \end{aligned} \quad (5.14)$$

Table 5.2: Expressions of $\delta = \hat{\delta}$, δ'_1 , $\hat{\delta}'_1$, δ_1 and $\hat{\delta}_1$ in terms of the rates $(\hat{b}_i, \hat{a}_i)$:

$$\delta = \hat{\delta} = \sup_{1 \leq i < N+1} \hat{\nu}[1, i] \hat{\mu}[i, N] = \sup_{1 \leq i < N+1} \sum_{k=1}^i \hat{\nu}_k \sum_{j=i}^N \hat{\mu}_j, \quad (5.15)$$

$$\begin{aligned} \delta'_1 &= \sup_{1 \leq i < N+1} \frac{1}{\hat{\mu}[i, N]} \sum_{k=1}^N \hat{\nu}_k \hat{\mu}[k \vee i, N]^2 \\ &= \sup_{1 \leq i < N+1} \left[\hat{\mu}[i, N] \hat{\nu}[1, i] + \frac{1}{\hat{\mu}[i, N]} \sum_{k=i+1}^N \hat{\nu}_k \hat{\mu}[k, N]^2 \right], \end{aligned} \quad (5.16)$$

$$\begin{aligned} \hat{\delta}'_1 &= \sup_{1 \leq i < N+1} \frac{1}{\hat{\nu}[1, i]} \sum_{k=1}^N \hat{\mu}_k \hat{\nu}[1, k \wedge i]^2 \\ &= \sup_{1 \leq i < N+1} \left[\hat{\mu}[i, N] \hat{\nu}[1, i] + \frac{1}{\hat{\nu}[1, i]} \sum_{k=1}^{i-1} \hat{\mu}_k \hat{\nu}[1, k]^2 \right]. \end{aligned} \quad (5.17)$$

$$\begin{aligned} \delta_1 &= \sup_{1 \leq i < N+1} \frac{1}{\sqrt{\hat{\mu}[i, N]}} \sum_{k=1}^N \hat{\nu}_k \hat{\mu}[k \vee i, N] \sqrt{\hat{\mu}[k, N]} \\ &= \sup_{1 \leq i < N+1} \left[\sqrt{\hat{\mu}[i, N]} \sum_{k=1}^i \hat{\nu}_k \sqrt{\hat{\mu}[k, N]} + \frac{1}{\sqrt{\hat{\mu}[i, N]}} \sum_{k=i+1}^N \hat{\nu}_k \hat{\mu}[k, N]^{3/2} \right], \end{aligned} \quad (5.18)$$

$$\begin{aligned} \hat{\delta}_1 &= \sup_{1 \leq i < N+1} \frac{1}{\sqrt{\hat{\nu}[1, i]}} \sum_{k=1}^N \hat{\mu}_k \hat{\nu}[1, k \wedge i] \sqrt{\hat{\nu}[1, k]} \\ &= \sup_{1 \leq i < N+1} \left[\frac{1}{\sqrt{\hat{\nu}[1, i]}} \sum_{k=1}^{i-1} \hat{\mu}_k \hat{\nu}[1, k]^{3/2} + \sqrt{\hat{\nu}[1, i]} \sum_{k=i}^N \hat{\mu}_k \sqrt{\hat{\nu}[1, k]} \right], \end{aligned} \quad (5.19)$$

The next four examples are dual of Examples 3.4–3.7, respectively.

Example 5.3. For Example 3.4, we have $\hat{a}_i \equiv b$ ($i \geq 1$), $\hat{b}_i \equiv a$ ($a > 0$), $b \geq a$. Then $\hat{\delta} = \delta = b(a-b)^{-2}$, $\hat{\delta}'_1 = \delta'_1 = (a+b)(a-b)^{-2}$, and $\hat{\delta}_1 = \delta_1 = \lambda_0^{-1} = (\sqrt{a} - \sqrt{b})^{-2}$. In particular, if we take $\hat{a}_i = 4$ and $\hat{b}_i = 1$ ($i \geq 1$), then $\hat{\lambda}_0 = 1$,

$$\begin{aligned} \hat{\delta}'_1 &= 5/9 = 0.\dot{5}, & \hat{\delta}'_2 &= 0.6\dot{4}, & \hat{\delta}'_3 &\approx 0.71, & \hat{\delta}'_4 &\approx 0.755, & \hat{\delta}'_5 &\approx 0.79; \\ \bar{\delta}_1 &= 0.\dot{5}, & \bar{\delta}_2 &\approx 0.71, & \bar{\delta}_3 &\approx 0.79, & \bar{\delta}_4 &\approx 0.835 & \bar{\delta}_5 &\approx 0.8647. \end{aligned}$$

Thus, $\hat{\delta}'_n$ and $\bar{\delta}_n$ are increasing and close to $\hat{\lambda}_0^{-1}$ as $n \uparrow$.

Proof. To compute $\hat{\delta}'_1$ and $\hat{\delta}_1$, we use Table 5.1. For simplicity, write $\gamma = b/a > 1$. Then

$$\mu_k = \gamma^k, \quad \mu[0, i] = \frac{\gamma^{i+1} - 1}{\gamma - 1}, \quad \nu_k = \frac{1}{b} \gamma^{-k}.$$

(a) Note that

$$\begin{aligned}
& \frac{1}{\mu[0, i]} \sum_{k=0}^{i-1} \nu_k \mu[0, k]^2 + \mu[0, i] \sum_{k=i}^{\infty} \nu_k \\
&= \frac{1}{b} \left[\frac{\gamma-1}{\gamma^{i+1}-1} \sum_{k=0}^{i-1} \gamma^{-k} \left(\frac{\gamma^{k+1}-1}{\gamma-1} \right)^2 + \frac{\gamma^{i+1}-1}{\gamma-1} \sum_{k \geq i} \gamma^{-k} \right] \\
&= \frac{1}{b(\gamma-1)} \left[\frac{1}{\gamma^{i+1}-1} \sum_{k=0}^{i-1} \gamma^{-k} (\gamma^{k+1}-1)^2 + (\gamma^{i+1}-1) \sum_{k \geq i} \gamma^{-k} \right] \\
&= \frac{1}{b(\gamma-1)} \left[\frac{\gamma(1+\gamma)}{\gamma-1} - \frac{2(i+1)\gamma}{\gamma^{i+1}-1} \right].
\end{aligned}$$

Since the second term in the last $[\dots]$ is negative and $\gamma > 1$, the right-hand side attains its supremum at $i = \infty$. By (5.12), we have thus obtained

$$\hat{\delta}'_1 = \frac{\gamma(1+\gamma)}{b(\gamma-1)^2} = \frac{a+b}{(a-b)^2}.$$

(b) Next, note that

$$\begin{aligned}
& \frac{1}{\sqrt{\mu[0, i]}} \sum_{k=0}^{i-1} \nu_k \mu[0, k]^{3/2} + \sqrt{\mu[0, i]} \sum_{k=i}^{\infty} \nu_k \sqrt{\mu[0, k]} \\
&= \frac{1}{b(\gamma-1)} \left[\frac{1}{\sqrt{\gamma^{i+1}-1}} \sum_{k=0}^{i-1} \gamma^{-k} (\gamma^{k+1}-1)^{3/2} + \sqrt{\gamma^{i+1}-1} \sum_{k \geq i} \gamma^{-k} \sqrt{\gamma^{k+1}-1} \right] \\
&\leq \frac{1}{b(\gamma-1)} \left[\frac{1}{\sqrt{\gamma^{i+1}-1}} \sum_{k=0}^{i-1} \gamma^{(k+3)/2} + \sqrt{\gamma^{i+1}-1} \sum_{k \geq i} \gamma^{-k/2+1/2} \right] \\
&= \frac{1}{b(\gamma-1)} \left[\frac{1}{\sqrt{\gamma^{i+1}-1}} \frac{\gamma^{3/2}(\gamma^{i/2}-1)}{\sqrt{\gamma}-1} + \frac{\gamma^{-i/2+1/2} \sqrt{\gamma^{i+1}-1}}{1-1/\sqrt{\gamma}} \right] \\
&\leq \frac{1}{b(\gamma-1)} \left[\frac{\gamma}{\sqrt{\gamma}-1} + \frac{\gamma\sqrt{\gamma}}{\sqrt{\gamma}-1} \right] \\
&= \frac{\gamma}{b(\sqrt{\gamma}-1)^2} \\
&= \frac{1}{(\sqrt{a}-\sqrt{b})^2}.
\end{aligned}$$

By (5.14), this means that $\hat{\delta}_1 \leq \hat{\lambda}_0^{-1}$ and so the equality sign must hold because $\hat{\delta}_1^{-1}$ is a lower estimate: $\hat{\lambda}_0 \geq \hat{\delta}_1^{-1}$.

(c) We now compute the approximating sequences $\{\hat{\delta}'_n\}$ and $\{\bar{\hat{\delta}}_n\}$ for the upper estimate, using the dual rate $(\hat{a}_i, \hat{b}_i)$. In the particular case, we have

$$\hat{\mu}_i = 4^{1-i}, \quad \hat{\nu}_i = 4^{i-2}, \quad \hat{\varphi}_i = \hat{\nu}[1, i] = \frac{4^i - 1}{12}.$$

The approximating sequences can be computed successively by using the following formulas:

$$\begin{aligned} f_1^{(m)}(i) &= \frac{4^i - 1}{12}, \quad i \in \{1, 2, \dots, m\}, \\ f_n^{(m)}(i) &= \frac{1}{3} \left\{ \sum_{k=1}^{i-1} (1 - 4^{-k}) f_{n-1}^{(m)}(k) + (4^i - 1) \sum_{k=i}^{m-1} 4^{-k} f_{n-1}^{(m)}(k) \right. \\ &\quad \left. + \frac{1}{3} (4^i - 1) 4^{1-m} f_{n-1}^{(m)}(m) \right\}, \quad i \in \{1, 2, \dots, m\}, n \geq 2. \end{aligned}$$

Then $\hat{\delta}'_n = \sup_{m \geq 1} \min_{1 \leq i \leq m} f_{n+1}^{(m)}(i) / f_n^{(m)}(i)$. For the first five of $\{\hat{\delta}'_n\}$, the minimum are all attained at m and so the computations become easier.

To compute $\hat{\delta}_n$, simply use the formula

$$\hat{\delta}_n = \sup_{m \geq 1} \frac{\sum_{i=1}^m 4^{1-i} f_n^{(m)}(i)^2 + 3^{-1} 4^{1-m} f_n^{(m)}(m)^2}{\sum_{i=1}^m 4^{2-i} (f_n^{(m)}(i) - f_n^{(m)}(i-1))^2}, \quad f_n^{(m)}(0) := 0. \quad \square$$

Example 5.4. For Example 3.5 with $\gamma = 1$ ($\hat{b}_i = i$, $\hat{a}_i = 2i$), we have $\hat{\delta}'_1 \approx 0.75 < \delta'_1 \approx 0.84$ and $\hat{\delta}_1 \approx 1.12 > \delta_1 \approx 1.09$. Besides, $\hat{\delta}_1 / \hat{\delta}'_1 \approx 1.5$.

Example 5.5. For Example 3.6, we have $\hat{a}_i = \hat{b}_i = i^2$ ($i \geq 1$), $\hat{b}_0 = 0$, $\hat{\delta}'_1 = 2 < \delta'_1 \approx 2.19$ and $\hat{\delta}_1 = \delta_1 = 4$ which is sharp. Besides, $\hat{\delta}_1 / \hat{\delta}'_1 = 2$.

Proof. By Example 3.6 and Corollary 5.2, it follows that $\hat{\lambda}_0 = \lambda_0 = 1/4$. Here, we present an easier proof for the upper estimate. Note that when $\hat{a}_i = \hat{b}_i$ for $i \geq 2$, we have

$$\hat{\mu}_1 = 1, \quad \hat{\mu}_i = \frac{\hat{b}_1 \cdots \hat{b}_{i-1}}{\hat{a}_2 \cdots \hat{a}_i} = \frac{\hat{b}_1}{\hat{a}_i}, \quad i \geq 2; \quad \hat{\mu}_i \hat{b}_i = \hat{b}_1, \quad i \geq 1. \quad (5.20)$$

In the present case, we have $\hat{\mu}_i = i^{-2}$ ($i \geq 1$) and $\hat{\mu}_i \hat{a}_i \equiv 1$. Let $f_i^{(m)} = \sqrt{i \wedge m}$. Then

$$\begin{aligned} \hat{\mu}(f^{(m)2}) &= \sum_{i=1}^m \frac{1}{i} + m \sum_{i \geq m+1} \frac{1}{i^2}, \\ \hat{D}(f^{(m)}) &= \sum_{i=1}^m (\sqrt{i} - \sqrt{i-1})^2 = \sum_{i=1}^m \frac{1}{(\sqrt{i} + \sqrt{i-1})^2} \leq 1 + \frac{1}{4} \sum_{i=1}^{m-1} \frac{1}{i}. \end{aligned}$$

Hence,

$$\hat{\lambda}_0 \leq \lim_{m \rightarrow \infty} \frac{\hat{D}(f^{(m)})}{\hat{\mu}(f^{(m)2})} = \frac{1}{4}. \quad \square$$

Example 5.6. For Example 3.7, we have $\hat{a}_i = i^4$ ($i \geq 1$), $\hat{b}_i = i(i-1/2)(i^2+3i+3)$, $\hat{\lambda}_0 = \lambda_0 = 1/2$, $\hat{\delta}'_1 \approx 1.83 < \delta'_1 \approx 1.9$ and $\hat{\delta}_1 \approx \delta_1 \approx 2$. Besides, $\hat{\delta}_1/\hat{\delta}'_1 = 1.09$.

Proof. First, we have

$$\hat{\mu}_i = \frac{\prod_{k=1}^{i-1} (k-1/2)(k^2+3k+3)}{i!^3}, \quad \hat{\nu}_i = \frac{(i-1)!^3}{\prod_{k=1}^{i-1} (k-1/2)(k^2+3k+3)}, \quad i \geq 1.$$

By (5.5) and Example 3.7, we have $\sum_i \hat{\mu}_i < \infty$ and $\sum_i \hat{\nu}_i < \infty$, and so the minimal dual process is explosive (but here we are dealing with the maximal one). The sharp lower bound can be deduced from part (1) of Theorem 4.1 with the dual test sequence

$$\hat{v}_i = 1 + \frac{1}{i(i^2+3i+3)}, \quad i \geq 1.$$

From this, it follows that the corresponding eigenfunction

$$\hat{g}_i = \prod_{k=1}^{i-1} \hat{v}_k, \quad i \geq 2, \quad \hat{g}_1 = 1,$$

increases strictly to a finite limit since $\sum_{i \geq 1} i^{-1}(i^2+3i+3)^{-1} < \infty$. The sequence $(\hat{v}_i)$ comes from the one computed in Example 3.7 plus a use of (2.35) and (5.7). $\square$

The precise value of λ_0 for the next example is unknown. Its eigenfunction is non-polynomial. It is interesting to compare this example with the ergodic one given in §6 for which $\lambda_1 = 2$, as well as the one with rates $a_i = i+1$ and $b_i = i^2$ ($i \geq 1$) given in §7 for which $\lambda_0 = 2$.

Example 5.7. Let $\hat{b}_0 = 0$, $\hat{b}_i = i+2$ ($i \geq 1$) and $\hat{a}_i = i^2$. It is the dual of the process studied in §2 with rates $a_i = i+2$ ($i \geq 1$) and $b_i = (i+1)^2$ ($i \geq 0$). Then $\hat{\lambda}_0 \in (0.395, 0.399)$, $\hat{\delta}'_1 \approx 2.37 < \delta'_1 \approx 2.48$ and $\hat{\delta}_1 \approx 2.63 > \delta_1 \approx 2.61$. Besides, $\hat{\delta}_1/\hat{\delta}'_1 \approx 1.1$.

It is interesting that for all of Examples 5.3–5.7, we have $\hat{\delta}'_1 \leq \delta'_1$ and $\hat{\delta}_1 \geq \delta_1$ which then means that Corollary 3.3 is more effective than Corollary 4.4. The effectiveness of the bounds δ_1 and δ'_1 given in Corollary 4.4 was also checked by Sirl, Zhang and Pollett (2007) for some models from practice.

Remark 5.8. It is now a suitable position to mention a method for the numerical computation of λ_0 defined in §4. The idea is meaningful in the other cases. From proof (b) of Theorem 4.1, it follows that there is a sequence $(v_i : v_i > 1, 1 \leq i < N)$ such that

$$R_i(v) = a_i(1 - v_{i-1}^{-1}) + b_i(1 - v_i) = \lambda_0, \quad v_0 = \infty, \quad v_N = 0 \text{ if } N < \infty.$$

Hence, we have

$$\begin{cases} v_1 - 1 = (a_1 - \lambda_0)b_1^{-1}, \\ v_i - 1 = [a_i(1 - v_{i-1}^{-1}) - \lambda_0]b_i^{-1}, \quad 2 \leq i < N. \end{cases} \quad (5.21)$$

In other words, replacing $v_i - 1$ by u_i , when $z = \lambda_0$, the equation

$$\begin{cases} u_1 = (a_1 - z)b_1^{-1}, \\ u_i = [a_i u_{i-1} (1 + u_{i-1})^{-1} - z] b_i^{-1}, \quad 2 \leq i < N, \end{cases} \quad (5.22)$$

has a positive solution $(u_i = u_i(z))_{1 \leq i < N}$. Thus, one may use the maximal z so that (5.22) has a positive solution as an approximation of λ_0 (based on part (1) of Theorem 4.1). In this way, we obtain the approximation of $\hat{\lambda}_0$ given in Example 5.7.

6. REFLECTING (NEUMANN) BOUNDARIES AT ORIGIN AND INFINITY (ERGODIC CASE)

We now turn to studying the first non-trivial eigenvalue in the ergodic case. Let $E = \{i : 0 \leq i < N + 1\}$ ($N \leq \infty$), $b_0 > 0$, $b_N = 0$ if $N < \infty$,

$$\lambda_1 = \inf \{D(f) : \mu(f) = 0, \mu(f^2) = 1\}, \quad (6.1)$$

where $\mu(f) = \sum_{i \in E} f_i \mu_i$,

$$D(f) = \sum_{0 \leq i < N} \mu_i b_i (f_{i+1} - f_i)^2 = \sum_{1 \leq i < N+1} \mu_i a_i (f_i - f_{i-1})^2 \quad (6.2)$$

with domain $\mathcal{D}^{\max}(D) = \{f \in L^2(\mu) : D(f) < \infty\}$. In (6.1), we presume that

$$\sum_{i=0}^N \mu_i < \infty. \quad (6.3)$$

Then the Dirichlet form $(D, \mathcal{D}^{\max}(D))$ has a trivial eigenvalue $\lambda_0 = 0$ with constant eigenfunction $\mathbb{1}$, and here we are working on the next “eigenvalue” λ_1 of $(D, \mathcal{D}^{\max}(D))$. If (6.3) does not hold, then $\mathbb{1} \notin L^2(\mu)$ and so λ_1 is not meaningful. Moreover, by (1.3) and Proposition 1.3, the Dirichlet form is unique. In this case, the corresponding process is explosive, or zero-recurrent, or transient. The decay rate is described by λ_0 which has already been treated in Sections 2 and 3. Hence, throughout this section, we assume (6.3).

Note that condition (6.3) plus (1.2) means that the unique process is ergodic. When $N = \infty$ and (1.2) fails, the minimal process was treated in Sections 2 and 3, and in this section, we are dealing with the maximal process (cf. [10; Proposition 6.56]) as in Section 4, it is indeed the unique honest reversible process. Denote by $Q = (q_{ij})$ the birth-death Q -matrix. Then under (6.3), the maximal process $P_{ij}^{\max}(\lambda)$ (Laplace transform) can be expressed as

$$P_{ij}^{\max}(\lambda) = P_{ij}^{\min}(\lambda) + \frac{z_i(\lambda) \mu_j z_j(\lambda)}{\lambda \sum_k \mu_k z_k(\lambda)}, \quad i, j \in E, \quad \lambda > 0,$$

where for each fixed j , $\{P_{ij}^{\min}(\lambda) : i \in E\}$ is the minimal solution to the equations

$$x_i = \sum_{k \neq i} \frac{q_{ik}}{\lambda + q_i} x_k + \frac{\delta_{ij}}{\lambda + q_i}, \quad i \in E,$$

and $(z_i(\lambda) : i \in E)$ is the maximal solution to the equation

$$\begin{cases} (\lambda I - Q)u = 0, \\ 0 \leq u \leq 1, \end{cases} \quad \lambda > 0$$

(cf. [10; Proposition 6.56]). According to a result due to Z.K. Wang (1964) (cf. Wang and Yang (1992, §6.8, Theorem 2)): if $N = \infty$ and (1.2) fails, then every honest process (may be non-symmetric) is ergodic and so is the maximal one. Certainly, within the symmetric context, by using (1.4), it is easy to check directly the ergodicity of the maximal process.

Here, we mention a technical point. If (6.3) fails, then as mentioned before, by (1.3), there is precisely one symmetrizable process (Dirichlet form) which is nothing but the minimal one. Thus, if (1.2) also fails, then the unique process must be explosive and so there is no honest symmetrizable process. This is a different point to the reversible case (i.e., (6.3) holds) for which there exists exactly one honest reversible process as just mentioned above.

We use the same notation $I, II, \mathcal{F}_I, \mathcal{F}_{II}, \widetilde{\mathcal{F}}_I$ and $\widetilde{\mathcal{F}}_{II}$ defined in Section 4 with an addition “ $f_0 = 0$ ” in the last four sets, but redefine R and $\mathcal{V}$ as follows:

$$\begin{aligned} R_i(v) &= a_{i+1} + b_i - a_i/v_{i-1} - b_{i+1}v_i, \quad 0 \leq i < N, \\ v_{-1} &> 0 \text{ is free and so is } v_{N-1} \text{ if } N < \infty, \\ \mathcal{V} &= \{v : v_i > 0 \text{ for all } i : 0 \leq i < N-1\}. \end{aligned}$$

The local operator $\widetilde{R}$ is modified from R , replacing a_m by $\tilde{a}_m := \mu_m a_m / \sum_{k=m}^N \mu_k$ for v with $\text{supp}(v) = \{0, 1, \dots, m-2\}$ in the set

$$\begin{aligned} \widetilde{\mathcal{V}} &= \bigcup_{m=2}^N \left\{ v : \frac{a_{i+1}}{a_{i+2} + b_{i+1}} < v_i < \frac{a_{i+1} + b_i - a_i/v_{i-1}}{b_{i+1}} \text{ for } i = 0, 1, \dots, m-2 \right. \\ &\quad \left. \text{and } v_i = 0 \text{ for } i \geq m-1 \right\}. \end{aligned}$$

Theorem 6.1. Under (6.3), the following variational formulas for λ_1 hold.

(1) Difference form:

$$\inf_{v \in \widetilde{\mathcal{V}}} \sup_{0 \leq i < N} \widetilde{R}_i(v) = \lambda_1 = \sup_{v \in \mathcal{V}} \inf_{0 \leq i < N} R_i(v).$$

(2) Summation form:

$$\inf_{f \in \widetilde{\mathcal{F}}_I \cup \widetilde{\mathcal{F}}'_I} \sup_{1 \leq i \in E} I_i(\bar{f})^{-1} = \lambda_1 = \sup_{f \in \mathcal{F}_I} \inf_{1 \leq i \in E} I_i(\bar{f})^{-1},$$

where

$$\begin{aligned} \widetilde{\mathcal{F}}'_I &= \{f \in L^2(\mu) : f_0 = 0, f \text{ is strictly increasing}\}, \\ \bar{f} &= f - \pi(f), \quad \pi = \mu/Z \quad \text{and} \quad Z = \sum_{i \in E} \mu_i. \end{aligned}$$

The use of $\bar{f}$ in the last line is based on the property $\bar{f} = \overline{f+c}$ for every constant c and so we can fix f_0 to be 0.

Proof of Theorem 6.1. In the ergodic case under (1.2), the assertion

$$\inf_{f \in \tilde{\mathcal{F}}_I \cup \tilde{\mathcal{F}}'_I} \sup_{1 \leq i \in E} I_i(\bar{f})^{-1} \geq \lambda_1$$

was proved in [7; Theorem 2.3] (but in the case that $k = \infty$ in the original proof, one requires the L^2 -integrability condition included in $\tilde{\mathcal{F}}'_I$, as was pointed out in proof (c) of Theorem 4.1). The proof remains the same in the present general situation with an obvious modification when $N < \infty$. Next, in the ergodic case under (1.2), the following result

$$\lambda_1 = \sup_{v \in \mathcal{V}} \inf_{0 \leq i < N} R_i(v) = \sup_{f \in \mathcal{F}_I} \inf_{1 \leq i \in E} I_i(\bar{f})^{-1} \quad (6.4)$$

is just [3; Theorem 1.1]. In the present general situation, the proof for the second equality in (6.4) needs a slight change only (cf. [3; Lemma 2.1]). To prove the first equality in (6.4), we claim that

$$\begin{aligned} \lambda_1 &= \inf \{ D(f) : \mu(|f - \pi(f)|^2) = 1, f_i = f_{i \wedge m} \text{ for some } m \in E, m \geq 1 \} \\ &=: \tilde{\lambda}_1. \end{aligned} \quad (6.5)$$

To see this, first it is clear that $\tilde{\lambda}_1 \geq \lambda_1$. Next, the proof of [4; Theorem 3.2] gives us

$$\lambda_1 \geq \sup_{f \in \mathcal{F}_I} \inf_{1 \leq i \in E} I_i(\bar{f})^{-1},$$

and furthermore, the equality sign with λ_1 replaced by $\tilde{\lambda}_1$ holds. Once again, the key point for the last statement is to show that the eigenfunction of $\tilde{\lambda}_1$ is strictly increasing. For this, the original proof needs only a modification replacing λ_1 by $\tilde{\lambda}_1$ (as indicated in proof (b) of Theorem 4.1). Therefore, (6.4) holds in the present general situation.

Now, we need only to show that

- $\inf_{f \in \tilde{\mathcal{F}}_I} \sup_{1 \leq i \in E} I_i(\bar{f})^{-1} \leq \inf_{v \in \tilde{\mathcal{V}}} \sup_{0 \leq i < N} \tilde{R}_i(v)$, and
- $\inf_{v \in \tilde{\mathcal{V}}} \sup_{0 \leq i < N} \tilde{R}_i(v) \leq \lambda_1$.

(a) Prove that $\inf_{f \in \tilde{\mathcal{F}}_I} \sup_{1 \leq i \in E} I_i(\bar{f})^{-1} \leq \inf_{v \in \tilde{\mathcal{V}}} \sup_{0 \leq i < N} \tilde{R}_i(v)$.

As before, write $\tilde{R}(u)$ instead of $\tilde{R}(v)$. Given u with $\text{supp}(u) = \{0, 1, \dots, m-1\}$ so that $(v_i) \in \tilde{\mathcal{V}}$, where $v_i = u_{i+1}/u_i > 0$ for $i < m-1$ and $v_i = 0$ for $i \geq m-1$, let

$$f_i = \begin{cases} a_i u_{i-1} - b_i u_i, & i < m, \\ \tilde{a}_m u_{m-1}, & i \geq m. \end{cases}$$

Since the constraint in $\tilde{\mathcal{V}}$ is equivalent to $\min_{i \leq m-1} \tilde{R}_i(v) > 0$, it is easy to check that

$$(f_{i+1} - f_i)/u_i = \tilde{R}_i(u) > 0 \text{ for } i < m \text{ and } f_i = f_{i \wedge m},$$

and so $f + b_0 u_0 \in \widetilde{\mathcal{F}}_I$. Moreover, since

$$\begin{aligned} \sum_{k=i}^N \mu_k f_k &= \sum_{k=i}^{m-1} \mu_k f_k + f_m \sum_{j=m}^N \mu_j \\ &= \sum_{k=i}^{m-1} \mu_k (a_k u_{k-1} - b_k u_k) + f_m \sum_{j=m}^N \mu_j \\ &= \mu_i a_i u_{i-1} - \mu_m a_m u_{m-1} + \tilde{a}_m u_{m-1} \sum_{j=m}^N \mu_j \\ &= \mu_i a_i u_{i-1}, \quad i \leq m-1, \end{aligned}$$

we get

$$\mu(f) = \sum_{k=0}^N \mu_k f_k = \mu_0 a_0 u_{-1} = 0,$$

and so

$$\begin{aligned} \sum_{k=i}^N \mu_k \bar{f}_k &= \mu_i a_i u_{i-1}, \quad i \leq m-1, \\ \sum_{k=m}^N \mu_k \bar{f}_k &= f_m \sum_{k=m}^N \mu_k = \tilde{a}_m u_{m-1} \sum_{k=m}^N \mu_k = \mu_m a_m u_{m-1}. \end{aligned}$$

It follows that

$$u_{i-1} = \frac{1}{\mu_i a_i} \sum_{k=i}^N \mu_k \bar{f}_k, \quad i \leq m.$$

Hence,

$$\tilde{R}_{i-1}(u) = \frac{f_i - f_{i-1}}{u_{i-1}} = I_i(\bar{f})^{-1}, \quad 1 \leq i \leq m.$$

Therefore, we have

$$\max_{0 \leq i < m} \tilde{R}_i(u) = \max_{1 \leq i \leq m} I_i(\bar{f})^{-1} \geq \inf_{f \in \widetilde{\mathcal{F}}_I, f_i = f_{i \wedge m}} \max_{1 \leq i \leq m} I_i(\bar{f})^{-1} \geq \inf_{f \in \widetilde{\mathcal{F}}_I} \sup_{1 \leq i \in E} I_i(\bar{f})^{-1},$$

and then

$$\inf_{v \in \widetilde{\mathcal{V}}} \sup_{0 \leq i < N} \tilde{R}_i(v) \geq \inf_{f \in \widetilde{\mathcal{F}}_I} \sup_{1 \leq i \in E} I_i(\bar{f})^{-1}.$$

Here, we have used the fact that $\tilde{R}_i(v) = -\infty$ for $i \geq m-1$ if $\text{supp}(v) = \{0, 1, \dots, m-2\}$ and in the last step, we have returned to the original notation $\tilde{R}(v)$ instead of $\tilde{R}(u)$.

(b) Prove that $\inf_{v \in \widetilde{\mathcal{V}}} \sup_{0 \leq i < N} \tilde{R}_i(v) \leq \lambda_1$.

Because of

$$\{\mu(f) = 0, \mu(f^2) = 1, f_i = f_{i \wedge m}\} \subset \{\mu(f) = 0, \mu(f^2) = 1, f_i = f_{i \wedge (m+1)}\},$$

by (6.5), it is clear that

$$\lambda_1^{(m)} := \inf \{ D(f) : \mu(f) = 0, \mu(f^2) = 1, f_i = f_{i \wedge m} \} \downarrow \lambda_1 \text{ as } m \uparrow N.$$

Actually, this is a special case of an approximation result given in [2; Theorem 4.2 and Corollary 4.3] or [10; Theorem 9.20 and Corollary 9.21]. Note that $\lambda_1^{(m)}$ is just the first non-trivial eigenvalue of the local Dirichlet form $(\tilde{D}, \mathcal{D}(\tilde{D}))$ defined by (4.17) replacing the Dirichlet boundary at 0 by the Neumann one (having the state space $\{0, 1, \dots, m\}$), with Neumann (reflecting) boundary at m . Denote by g the first eigenfunction of $\lambda_1^{(m)}$ and extend it to the whole space by setting $g_i = g_{i \wedge m}$. Now, if we set $u_i = g_{i+1} - g_i$ for $i \in E$, then $u_i > 0$ for $i \leq m-1$, $u_i = 0$ for $i \geq m$, and furthermore,

$$\tilde{R}_i(u) = \lambda_1^{(m)} > 0 \quad \text{for all } i \leq m-1.$$

Moreover, by the definition of g , we have

$$\begin{cases} b_i u_i - a_i u_{i-1} = -\lambda_1^{(m)} g_i, & i \leq m-1, \\ \tilde{a}_m u_{m-1} = \lambda_1^{(m)} g_m. \end{cases}$$

Making a difference of this with the one replacing i by $i+1$, we get $\tilde{R}_i(u) = \lambda_1^{(m)}$ for all $i \leq m-1$ (From this, the reason should be clear why in the definition of $\tilde{\mathcal{V}}$, we use “ $v_i = 0$ for $i \geq m-1$ ” rather than “ $v_i = v_{i \wedge m}$ ”). Thus,

$$\begin{aligned} \lambda_1^{(m)} &= \max_{0 \leq i < m} \tilde{R}_i(u) \\ &\geq \inf_{u: \text{supp}(u) = \{0, 1, \dots, m-1\}; (v_i = u_{i+1}/u_i) \in \tilde{\mathcal{V}}} \max_{0 \leq i < m} \tilde{R}_i(u) \\ &\geq \inf_{u: \text{supp}(u) = \{0, 1, \dots, n\} \text{ for some } n \geq 0, n < N; (v_i = u_{i+1}/u_i) \in \tilde{\mathcal{V}}} \sup_{0 \leq i < N} \tilde{R}_i(u) \\ &= \inf_{v \in \tilde{\mathcal{V}}} \sup_{0 \leq i < N} \tilde{R}_i(v). \end{aligned}$$

Here in the last step, we have returned to the original notation $\tilde{R}(v)$ instead of $\tilde{R}(u)$. Letting $m \rightarrow N$, we obtain the required assertion. $\square$

With the same rates (a_i, b_i) here but endow with the Dirichlet boundary at 0, we return to the situation studied in Section 4. The next result, taken from [7; Theorem 2.2] and [6; Theorem 3.5], is a comparison of λ_1 with the quantities λ_0 , δ , δ_1 and δ'_1 given in Section 4. See also Corollary 6.6 below for an improvement.

Theorem 6.2 (Criterion and basic estimates). Under (6.3), $\lambda_1 > 0$ iff $\delta < \infty$. More precisely, we have

$$\frac{1}{4\delta} \leq \frac{1}{\delta_1} \leq \lambda_0 \leq \lambda_1 \leq \lambda_0 Z \leq \frac{Z}{\delta'_1} \leq \frac{Z}{\delta}. \quad (6.6)$$

The next two results are mainly taken from [7; Theorem 2.4] with an addition on the monotonicity of $\{\eta_m\}$ and $\{\eta'_m\}$.

Theorem 6.3 (Approximating procedure). Let (6.3) hold and $\delta < \infty$. Write $\varphi_0 = 0$, $\varphi_i = \sum_{0 \leq j \leq i-1} (\mu_j b_j)^{-1} =: \nu[0, i-1]$ ($1 \leq i \in E$), $\bar{f} = f - \pi(f)$, $\pi = \mu/Z$, and $Z = \sum_{k \in E} \mu_k =: \mu[0, N]$.

- (1) Define $f_1 = \sqrt{\varphi}$, $f_n = \bar{f}_{n-1} II(\bar{f}_{n-1})$ and $\eta_n = \sup_{1 \leq i \in E} I_i(\bar{f}_n)$. Then η_n is decreasing in n and $\lambda_1 \geq \eta_n^{-1} \geq (4\delta)^{-1}$ for all $n \geq 1$.
- (2) For fixed $m \in E$: $m \geq 1$, define

$$f_1^{(m)} = \varphi_{\cdot \wedge m}, \quad f_n^{(m)} = [\bar{f}_{n-1}^{(m)} II(\bar{f}_{n-1}^{(m)})](\cdot \wedge m), \quad n \geq 2,$$

and then define

$$\eta'_n = \sup_{1 \leq m \in E} \inf_{1 \leq i \in E} I_i(\bar{f}_n^{(m)}), \quad \bar{\eta}_n = \sup_{1 \leq m \in E} \frac{\mu(\bar{f}_n^{(m)})^2}{D(f_n^{(m)})}, \quad n \geq 1.$$

Then η'_n is increasing in n and $\eta'_n{}^{-1} \geq \bar{\eta}_n^{-1} \geq \lambda_1$ for all $n \geq 1$.

The notation “ $\bar{f}_{n-1} II(\bar{f}_{n-1})$ ” used in the theorem may have 0/0 but it should not cost any confusion. Note that here we use the same (ν_j) as in (2.15). In other words, when $b_0 > 0$, we use (2.15). But for its dual, it is more convenient to use $\hat{\nu}_j = (\hat{\mu}_j \hat{a}_j)^{-1}$ as in Section 4 since $b_0 = 0$. This is consistent with the notation used in Section 5.

As a consequence of Theorem 6.3, we have the following improvement of Theorem 6.2.

Corollary 6.4 (Improved estimates). Let (6.3) hold. Then we have

$$(4\delta)^{-1} \leq \eta_1^{-1} \leq \lambda_1 \leq \bar{\eta}_1^{-1}, \quad (6.7)$$

where

$$\eta_1 = \sup_{1 \leq i \in E} (\sqrt{\varphi_i} + \sqrt{\varphi_{i-1}}) \left[\psi_i - \psi_1 \frac{\mu[i, N]}{\mu[0, N]} \right], \quad \psi_i := \sum_{j=i}^N \mu_j \sqrt{\varphi_j}, \quad (6.8)$$

$$\begin{aligned} \bar{\eta}_1 &= \sup_{1 \leq m \in E} \frac{1}{\varphi_m} \left[\sum_{1 \leq k \leq m} \mu_k \varphi_{k \wedge m}^2 - \frac{1}{Z} \left(\sum_{1 \leq k \in E} \mu_k \varphi_{k \wedge m} \right)^2 \right] \\ &= \sup_{1 \leq m \in E} \left\{ \frac{1}{\varphi_m} \left[\sum_{1 \leq k \leq m-1} \mu_k \varphi_k^2 - \frac{1}{\mu[0, N]} \left(\sum_{1 \leq k \leq m-1} \mu_k \varphi_k \right)^2 \right] \right. \\ &\quad \left. + \frac{\mu[m, N]}{\mu[0, N]} \left[\varphi_m \mu[0, m-1] - 2 \sum_{1 \leq k \leq m-1} \mu_k \varphi_k \right] \right\}. \end{aligned} \quad (6.9)$$

Proof of Theorem 6.3.

Part 1. We prove that $\{f_n\} \subset L^1(\mu)$ in three steps. This was missed in the original paper [7]. Certainly, we need only to consider the case that $N = \infty$.

(a) First, we show that the functions $\{h_n\}$,

$$h_0(i) \equiv 1, \quad i \in E, \quad h_n(i) = \sum_{j=1}^i \frac{1}{\mu_j a_j} \sum_{k=j}^{\infty} \mu_k h_{n-1}(k), \quad i \geq 1, n \geq 1,$$

are all in $L^1(\mu)$. Clearly, h_1 (and then h_n for $n \geq 2$) may increase to infinity if the minimal process is recurrent which is the main problem we need to handle. The required assertion says that even though h_n can be unbounded but is still in $L^1(\mu)$. For this, to distinguish with $\{f_n\}$ used in Theorem 6.3, let $\{\tilde{f}_n\}$ be the sequence defined in part (1) of Theorem 4.3:

$$\begin{aligned} \tilde{f}_1(i) &= \left(\sum_{k=1}^i \nu_{k-1} \right)^{1/2} = f_1(i), \quad i \geq 1, \quad \nu_{j-1} := \frac{1}{\mu_j a_j}, \\ \tilde{f}_n(i) &= \sum_{j=1}^i \nu_{j-1} \sum_{k=j}^{\infty} \mu_k \tilde{f}_{n-1}(k), \quad i \geq 1, n \geq 2, \\ \tilde{f}_n(0) &= 0, \quad n \geq 1. \end{aligned}$$

From proof (b) of Theorem 4.3, we have seen that

$$\tilde{f}_2(i) = \sum_{j=1}^i \frac{1}{\mu_j a_j} \sum_{k \geq j} \mu_k \tilde{f}_1(k) \leq 4\delta \tilde{f}_1(i).$$

Because $\tilde{f}_1(i) \geq \tilde{f}_1(1) = a_1^{-1/2}$ for $i \geq 1$, this gives us

$$h_1(i) \leq 4\delta \sqrt{a_1} \tilde{f}_1(i), \quad i \geq 1.$$

By induction, it follows that

$$h_n \leq \sqrt{a_1} (4\delta)^n \tilde{f}_1, \quad n \geq 1.$$

This proves that $h_n \in L^1(\mu)$ for all $n \geq 1$ since $\tilde{f}_1 \in L^1(\mu)$ as mentioned in proof (b) of Theorem 4.3, due to the assumption $\delta < \infty$.

(b) Next, we study the relation between $\{f_n\}$ and $\{\tilde{f}_n\}$. By definition, we have

$$\begin{aligned} f_2(i) &= \sum_{j=1}^i \nu_{j-1} \sum_{k=j}^{\infty} \mu_k \bar{f}_1(k) = \tilde{f}_2(i) - h_1(i) \pi(f_1), \quad i \geq 1, \\ f_3(i) &= \sum_{j=1}^i \nu_{j-1} \sum_{k=j}^{\infty} \mu_k \bar{f}_2(k) = \tilde{f}_3(i) - h_2(i) \pi(f_1) - h_1(i) \pi(f_2), \quad i \geq 1, \\ f_4(i) &= \tilde{f}_4(i) - h_3(i) \pi(f_1) - h_2(i) \pi(f_2) - h_1(i) \pi(f_3), \quad i \geq 1. \end{aligned}$$

Successively, we obtain

$$f_n = \tilde{f}_n - \sum_{k=1}^{n-1} \pi(f_k) h_{n-k}, \quad n \geq 2.$$

(c) Since $f_1 = \tilde{f}_1 \in L^1(\mu)$ as shown in proof (b) of Theorem 4.3. Now, to show that $\{f_n\} \subset L^1(\mu)$, by (a) and (b), it suffices to prove that $\{\tilde{f}_n\} \subset L^1(\mu)$. This is done in proof (b) of Theorem 4.3.

Part 2. We now prove the monotonicity of $\{\eta_n\}$ in two steps. Since $\bar{f}_n$ values both positive and negative or even zero, the proportional property used in the proof of the monotonicity of $\{\delta_n\}$ is currently not available. To overcome this difficulty, a finer technique is needed.

(d) Because

$$\mu_i a_i [f_n(i) - f_n(i-1)] = \sum_{k=i}^N \mu_k \bar{f}_{n-1}(k), \quad n \geq 2,$$

by the definition of $I(f)$, we obtain

$$\eta_n = \sup_{1 \leq i \in E} \sum_{j=i}^N \mu_j \bar{f}_n(j) \Big/ \sum_{k=i}^N \mu_k \bar{f}_{n-1}(k), \quad n \geq 2. \quad (6.10)$$

Since the denominator is positive, the assertion that $\eta_n \leq \eta_{n-1}$ is equivalent to

$$\sum_{j=i}^N \mu_j [\bar{f}_n(j) - \eta_{n-1} \bar{f}_{n-1}(j)] \leq 0, \quad i \in E.$$

That is,

$$\eta_{n-1} \pi(f_{n-1}) - \pi(f_n) \leq \frac{1}{\mu[i, N]} \sum_{j=i}^N \mu_j [\eta_{n-1} f_{n-1}(j) - f_n(j)], \quad i \in E. \quad (6.11)$$

Let us observe the meaning of this inequality: the left-hand side is the infimum (attained at $i = 0$) of the right-hand side.

The monotonicity of $\{\eta_n\}$ now follows once we show that the right-hand side of (6.11) is luckily increasing in i , or equivalently,

$$\mu[i, N] \sum_{j=i+1}^N \mu_j [\eta_{n-1} f_{n-1}(j) - f_n(j)] \geq \mu[i+1, N] \sum_{j=i+1}^N \mu_j [\eta_{n-1} f_{n-1}(j) - f_n(j)].$$

By removing the common term

$$\mu[i+1, N] \sum_{j=i+1}^N \mu_j [\eta_{n-1} f_{n-1}(j) - f_n(j)]$$

in both sides, it is enough to check that

$$\eta_{n-1} \sum_{j=i+1}^N \mu_j [f_{n-1}(j) - f_{n-1}(i)] \geq \sum_{j=i+1}^N \mu_j [f_n(j) - f_n(i)], \quad i \in E, \quad n \geq 2. \quad (6.12)$$

First, let $n \geq 3$. Then by the definition of f_n and (6.10), we have

$$\begin{aligned} f_n(j) - f_n(i) &= \sum_{s=i+1}^j \nu_{s-1} \sum_{k=s}^N \mu_k \bar{f}_{n-1}(k) \\ &\leq \eta_{n-1} \sum_{s=i+1}^j \nu_{s-1} \sum_{k=s}^N \mu_k \bar{f}_{n-2}(k) \\ &= \eta_{n-1} [f_{n-1}(j) - f_{n-1}(i)]. \end{aligned}$$

This certainly implies (6.12) in the case of $n \geq 3$, regarded as an application of the proportional property. Next, let $n = 2$. Then by the definition of f_2 and η_1 , we have

$$\begin{aligned} f_2(j) - f_2(i) &= \sum_{s=i+1}^j \nu_{s-1} \sum_{k=s}^N \mu_k \bar{f}_1(k) \\ &\leq \eta_1 \sum_{s=i+1}^j [f_1(s) - f_1(s-1)] \\ &= \eta_1 [f_1(j) - f_1(i)]. \end{aligned}$$

This also implies (6.12) in the case of $n = 2$. We have thus proved that $\eta_n \leq \eta_{n-1}$ for all $n \geq 2$.

Part 3. To prove the monotonicity of $\{\eta'_n\}$, for each fixed m , as a dual argument (exchanging “sup” and “ $\leq$ ” with “inf” and “ $\geq$ ”, respectively) of the above proofs (d) and (e), we have

$$\inf_{1 \leq i \in E} I_i(\bar{f}_n^{(m)}) \leq \inf_{1 \leq i \in E} I_i(\bar{f}_{n+1}^{(m)}).$$

Then the assertion follows by making supremum with respect to m .

Part 4. The proof of $\bar{\eta}_n \geq \eta'_n$ is given in Lemma 6.5 below. $\square$

In practice, using $\bar{\eta}_n$ rather than η'_n is based on the following result.

Lemma 6.5. For every non-decreasing, and non-constant function f satisfying $f \in L^1(\mu)$ and $D(f) < \infty$, we have

$$\frac{\mu(\bar{f}^2)}{D(f)} \geq \inf_{1 \leq i \in E} I_i(\bar{f}).$$

Similarly, for every nonnegative, non-decreasing, and non-zero function f satisfying $f \in L^1(\mu)$ and $D(f) < \infty$, we have

$$\frac{\mu(f^2)}{D(f)} \geq \inf_{1 \leq i \in E} I_i(f).$$

Proof. (a) Since f is not a constant, we have $\mu(\bar{f}^2) > 0$ and $D(f) > 0$. Moreover, since $f \in L^1(\mu)$ is also non-decreasing, we claim that

$$\infty > \sum_{k=i}^N \mu_k \bar{f}_k > 0 \quad \text{for all } i \in E, \quad i \geq 1.$$

Actually, the non-decreasing sequence $\{\bar{f}_k\}$, starting at $\bar{f}_0 < 0$ (since f is non-trivial) and having mean zero, should be positive for all large enough k . Thus, if $\sum_{k=i_0}^N \mu_k \bar{f}_k \leq 0$ for some $i_0 : 1 \leq i_0 \in E$, then we would have $\bar{f}_{i_0} < 0$ (otherwise $\bar{f}_i \geq 0$ for all $i \geq i_0$ and then $\sum_{k=i_0}^N \mu_k \bar{f}_k > \mu_j \bar{f}_j > 0$ for large enough j). This implies that

$$\sum_{k \leq i_0-1} \mu_k \bar{f}_k \leq \bar{f}_{i_0} \sum_{k \leq i_0-1} \mu_k < 0,$$

and furthermore,

$$0 = \mu(\bar{f}) = \sum_{k=i_0}^N \mu_k \bar{f}_k + \sum_{k \leq i_0-1} \mu_k \bar{f}_k \leq \sum_{k \leq i_0-1} \mu_k \bar{f}_k < 0,$$

which is a contradiction. Because of the assertion we have just proved and using the convention that $1/0 = \infty$, it follows that $\inf_{1 \leq i \in E} I_i(\bar{f}) \in [0, \infty)$.

Let $\gamma = \inf_{1 \leq i \in E} I_i(\bar{f})$. Then we have

$$- \sum_{k \leq i-1} \mu_k \bar{f}_k = \sum_{k=i}^N \mu_k \bar{f}_k \geq \gamma \mu_i a_i (\bar{f}_i - \bar{f}_{i-1})$$

first for those i with $f_i > f_{i-1}$ and then for all $i : 1 \leq i \in E$. Multiplying both sides by $\bar{f}_i - \bar{f}_{i-1} \geq 0$, we obtain

$$-(\bar{f}_i - \bar{f}_{i-1}) \sum_{k \leq i-1} \mu_k \bar{f}_k \geq \gamma \mu_i a_i (\bar{f}_i - \bar{f}_{i-1})^2, \quad i \in E, \quad i \geq 1.$$

Making a summation over i from 1 to m , it follows that

$$- \sum_{i=1}^m (\bar{f}_i - \bar{f}_{i-1}) \sum_{k \leq i-1} \mu_k \bar{f}_k \geq \gamma \sum_{i=1}^m \mu_i a_i (\bar{f}_i - \bar{f}_{i-1})^2.$$

Noticing that the mean of $\bar{f}$ equals zero and exchanging the order of the sums, the left-hand side is equal to

$$\begin{aligned} - \sum_{k=0}^m \mu_k \bar{f}_k \sum_{i=k+1}^m (\bar{f}_i - \bar{f}_{i-1}) &= - \sum_{k=0}^m \mu_k \bar{f}_k (\bar{f}_m - \bar{f}_k) \\ &= -\bar{f}_m \sum_{k=0}^m \mu_k \bar{f}_k + \sum_{k=0}^m \mu_k \bar{f}_k^2 \\ &= \sum_{k=m+1}^N \mu_k \bar{f}_k \bar{f}_m + \sum_{k=0}^m \mu_k \bar{f}_k^2. \end{aligned}$$

As mentioned in the last paragraph, $\bar{f}_m > 0$ first for some m and then for all large enough m since $\bar{f}$ is non-decreasing, the right-hand side is controlled, for large enough m , from above by

$$\sum_{k=m+1}^N \mu_k \bar{f}_k^2 + \sum_{k=0}^m \mu_k \bar{f}_k^2 = \mu(\bar{f}^2).$$

With the assumption $D(f) < \infty$ in mind, the required assertion now follows immediately by passing the limit as $m \rightarrow N$.

(b) For the second assertion, since $f \in L^1(\mu)$ is nonnegative and non-zero, we have

$$\infty > \sum_{j=i}^N \mu_j f_j > 0 \quad \text{for all } i \in E.$$

Now, if $f_0 = 0$, then there is an i_0 such that $f_{i_0-1} = 0$ but $f_{i_0} > 0$ and so $I_{i_0}(f) < \infty$. If $f_0 > 0$ and $\inf_{1 \leq i \in E} I_i(f) = \infty$, then f should be a positive constant, and hence, $D(f) = 0$. In this case, the assertion is trivial since $\mu(f^2) > 0$. Therefore, we may assume that $\gamma := \inf_{1 \leq i \in E} I_i(f) < \infty$. We now have

$$\sum_{k=i}^N \mu_k f_k \geq \gamma \mu_i a_i (f_i - f_{i-1}), \quad i \in E, i \geq 1.$$

Hence,

$$\sum_{i=1}^m (f_i - f_{i-1}) \sum_{k=i}^N \mu_k f_k \geq \gamma \sum_{i=1}^m \mu_i a_i (f_i - f_{i-1})^2.$$

Exchanging the order of the sums, the left-hand side is equal to

$$\sum_{k=1}^N \mu_k f_k \sum_{i=1}^{k \wedge m} (f_i - f_{i-1}) = \sum_{k=1}^N \mu_k f_k (f_{k \wedge m} - f_0) \leq \sum_{k=1}^N \mu_k f_k f_{k \wedge m} \leq \mu(f^2).$$

Combining this with the last inequality, we have obtained the required assertion. $\square$

Having the comparison of $\bar{\eta}_n \geq \eta'_n$ (Lemma 6.5) in mind, one may expect a parallel result for δ'_n and $\bar{\delta}_n$ defined in Theorem 4.3. All the examples we have ever computed support the conjecture that $\bar{\delta}_n \geq \delta'_n$, however, there is still no proof. In general, we have $\bar{\delta}_{n+1} \geq \delta'_n$ only as stated in Theorem 4.3. Note that δ'_n is defined by using $II(f_n)$ rather than $I(f_n)$. If we redefine δ'_n by using $I(f_n)$ as in [7; Theorem 2.2], denoted by $\tilde{\delta}'_n$ for a moment, then by the second assertion of Lemma 6.5, we do have $\bar{\delta}_n \geq \tilde{\delta}'_n$. Besides, by the theorem just quoted, we also have $\delta'_n \geq \tilde{\delta}'_n \geq \delta'_{n-1}$. This remark is also meaningful for those δ'_n and $\bar{\delta}_n$ defined in Section 3.

Note that the factor of the upper and lower bounds of λ_1 given in Theorem 6.2 is $4Z > 4$. The next result has a factor 4 only. A simple comparison of κ below and $\delta^{(4.4)}$ shows that it is not easy to find such a result. Its proof is delayed to the next section.

Corollary 6.6 (Criterion and basic estimates). Let (6.3) hold. Then we have $\kappa^{-1}/4 \leq \lambda_1 \leq \kappa^{-1}$, where

$$\kappa^{-1} = \inf_{0 \leq n < m < N+1} \left[\left(\sum_{i=0}^n \mu_i \right)^{-1} + \left(\sum_{i=m}^N \mu_i \right)^{-1} \right] \left(\sum_{j=n}^{m-1} \frac{1}{\mu_j b_j} \right)^{-1}. \quad (6.13)$$

Furthermore, we have

$$\delta_L \wedge \delta_R \geq \kappa \geq Z^{-1} \delta_L,$$

where $Z = \sum_{i=0}^N \mu_i$,

$$\delta_L = \sup_{1 \leq n < N+1} \sum_{i=1}^n \frac{1}{\mu_i a_i} \sum_{j=n}^N \mu_j = \delta^{(4.4)}, \quad \delta_R = \sup_{0 \leq m < N} \sum_{j=0}^m \mu_j \sum_{k=m}^{N-1} \frac{1}{\mu_k b_k}.$$

In the case that the minimal process is ergodic, since

$$1 < Z < \infty, \quad \sum_j \frac{1}{\mu_j b_j} = \infty,$$

we have $\delta_R = \infty$ and so the second assertion of Corollary 6.6 goes back to Theorem 6.2. However, the first assertion of Corollary 6.6 is clearly finer. An extension of Corollary 6.6 to a more general state space is given in Corollary 7.9 below.

Most of the examples below are taken from [10; Examples 9.27]. The computation of $\bar{\eta}_1$, $\eta_1/\bar{\eta}_1$, and κ is newly added.

Example 6.7. Let $b_i = b$ ($i \geq 0$), and $a_i = a$ ($i \geq 1$), $a > b$. Then

$$\lambda_1 = (\sqrt{a} - \sqrt{b})^2, \quad \delta = \kappa = a(a-b)^{-2}, \quad \bar{\eta}_1 = \delta'_1 = (a+b)/(a-b)^2,$$

and $\eta_1 = \lambda_1^{-1}$ which is sharp. Besides, $\eta_1/\bar{\eta}_1 \leq 2$, the equality sign holds iff $b = a$. Note that λ_1 , η_1^{-1} , and $\bar{\eta}_1^{-1}$ all tend to zero as $b \rightarrow a$. Furthermore, $(\bar{\eta}_1, \eta_1) \subset (\kappa, 4\kappa)$.

Example 6.8. The typical linear model: let $b_i = \beta_1 i + \beta_0$ ($\beta_0 > 0$, $\beta_1 \geq 0$), and $a_i = \gamma_1 i$ ($\gamma_1 > \beta_1$) for $i \geq 0$. Then $\lambda_1 = \gamma_1 - \beta_1$. When $\beta_0 = 0$, we have $\lambda_0^{(4.2)} = \gamma_1 - \beta_1$.

Example 6.9. Let $b_i = b/(i+1)$ ($b > 0$) for $i \geq 0$, $a_i \equiv a > 0$ for $i \geq 1$. Then $\lambda_1 = a - (\sqrt{b^2 + 4ab} - b)/2$.

Example 6.10. Let $b_i \equiv b$ ($b > 0$) for $i \geq 0$, $a_i = (i \wedge k)a$ ($a > 0$) for $i \geq 1$ and some $k \geq 2$ satisfying

$$k^{-1} \leq a/b \leq k(k-1)^{-2}.$$

Then $\lambda_1 = (\sqrt{ak} - \sqrt{b})^2$.

Example 6.11. Let $b_0 = 1$, $b_i = i$, and $a_i = 2i$, $i \geq 1$. Then $\lambda_1 \geq \lambda_0 = 1$ but the precise value is unknown. Moreover,

$$\bar{\eta}_1 \approx 0.55, \quad \eta_1 \approx 0.9986 \quad \text{and} \quad \eta_1/\bar{\eta}_1 \approx 1.82 < 2.$$

Besides, $\kappa \approx 0.4856$ and so $(\bar{\eta}_1, \eta_1) \subset (\kappa, 4\kappa)$.

The next one is a continuation of [6; Example 3.10].

Example 6.12. Let $E = \{0, 1\}$. Then $\lambda_1 = Z\lambda_0 = \bar{\eta}_1^{-1} = \kappa^{-1}$ and $\lambda_0 = \delta^{-1}$. Hence, the last upper bound in (6.6) and the one in Corollary 6.6 are sharp but δ^{-1} is not an upper bound of λ_1 .

The first lower bound in (6.6) and the one in Corollary 6.6 are sharp for the seventh example in Table 6.1 below.

Examples 6.13. Here are some additional examples, given in Table 6.1, for which the quantities $\bar{\eta}_1 \leq \lambda_1^{-1} \leq \eta_1$ and $\kappa \leq \lambda_1^{-1} \leq 4\kappa$ are compared. For all these examples, we have $(\bar{\eta}_1, \eta_1) \subset (\kappa, 4\kappa)$ and so the estimates given in Corollary 6.4 are better than the ones in Corollary 6.6.

Table 6.1 Exact λ_1 and its estimates for eight examples

$b_i (i \geq 0)$	$a_i (i \geq 1)$	λ_1^{-1}	$\bar{\eta}_1$	η_1	$\eta_1/\bar{\eta}_1$	κ
$i+1$	$2i$	1	≈ 0.8	≈ 1.48	≈ 1.85	$2/3$
$i+1$	$2i+3$ $a_0 = 0$	$1/2$	≈ 0.346	≈ 0.638	≈ 1.84	≈ 0.28
$i+1$	$2i+4+\sqrt{2}$ $a_0 = 0$	$1/3$	≈ 0.218	≈ 0.398	≈ 1.83	≈ 0.18
$(i+1)^{-1}$	1	$2(3-\sqrt{5})^{-1}$ ≈ 2.618	≈ 1.92	≈ 3.24	≈ 1.69	≈ 1.6
1	$i \wedge 2$	$(\sqrt{2}-1)^{-2}$ ≈ 5.8284	≈ 3	≈ 5.8284	≈ 1.9	2
$i+2$	i^2	$1/2$	≈ 0.47	≈ 0.85	≈ 1.81	≈ 0.47
i^2 $b_0 = 1$	i^2	4	2	λ_1^{-1}	2	1
$2 + \frac{(-1)^i}{7-\sqrt{33}}$ $b_0 = \frac{7-\sqrt{33}}{2}$	$2[2+(-1)^i]$ $a_0 = 0$	$(6-\sqrt{33})^{-1}$ ≈ 3.9	≈ 2.11	≈ 4.21	≈ 2	≈ 1.56

7. BILATERAL ABSORBING (DIRICHLET) BOUNDARIES

This section deals with the fourth case of boundary conditions. It consists of two parts. The first one is for the ordinary birth-death processes as studied in the previous sections and the second one deals with the bilateral birth-death processes with a more general state space.

First, let us consider the processes with state space $E = \{i : 1 \leq i < N+1\}$ ($N \leq \infty$) with Dirichlet boundaries at 0 ($a_1 > 0$) and $N+1$ if $N < \infty$. Similar to Section 2, define

$$\lambda_0 = \inf \{D(f) : \mu(f^2) = 1, f \in \mathcal{K}\}, \quad (7.1)$$

where the symmetric measure (μ_i) is the same as in Section 4, $\mu(f) = \sum_{k \in E} \mu_k f_k$, and

$$D(f) = \sum_{k \in E} \mu_k a_k (f_k - f_{k-1})^2, \quad f_0 := 0,$$

with domain $\mathcal{D}^{\min}(D)$. Clearly, if one changes only the boundary condition at 0, then the resulting λ_0 is bigger or equal to $\lambda_0^{(2.2)}$. Note that if (1.3) fails, then the eigenvalues $\lambda_0^{(4.2)}$ and $\lambda_0^{(7.1)}$ are different which correspond to the maximal and the minimal Dirichlet forms, respectively. However, as mentioned in Section 4, once (1.3) holds, $\lambda_0^{(4.2)}$ coincides with $\lambda_0^{(7.1)}$. Then there are three cases. The first one is that $\sum_i \mu_i < \infty$ and $\sum_i (\mu_i a_i)^{-1} = \infty$. This case is treated in Section 4. In this section, we are mainly studying the second case that $\sum_i \mu_i = \infty$ but

$$\sum_{k=1}^N \frac{1}{\mu_k a_k} < \infty. \quad (7.2)$$

The third case is that $\sum_i \mu_i = \infty$ and $\sum_i (\mu_i a_i)^{-1} = \infty$ which is treated in the next theorem. In this degenerated case, since there is a killing at 1 (i.e., $a_1 > 0$), the process is transient. Without using duality, by Corollary 7.3 below, we also obtain that $\lambda_0 = 0$. See the comments right after Corollary 7.3.

Theorem 7.1.

- (1) First, let (7.2) hold. Define the dual rates $(\hat{a}_i, \hat{b}_i)$ by (5.1) in the inverse way:

$$\hat{b}_i = a_{i+1}, \quad \hat{a}_i = b_i, \quad 0 \leq i < N+1, \quad (7.3)$$

and denote by $\hat{\lambda}_1$ the eigenvalue defined in Section 6 for the dual process. Then we have $\lambda_0 = \hat{\lambda}_1$.

- (2) Next, let (7.2) fail and $\sum_{i \in E} \mu_i = \infty$. Then λ_0 (as well as $\lambda_0^{(4.2)}$) is equal to its dual $\hat{\lambda}_0 = \lambda_0^{(2.2)} = 0$.

Proof. (a) By (7.3), (7.2) and (5.5), we have

$$\sum_{k=0}^N \hat{\mu}_k < \infty. \quad (7.4)$$

Clearly, the dual process with rates $(\hat{a}_i, \hat{b}_i)$ has the state space $\hat{E} = \{i : 0 \leq i < N+1\}$. By exchanging (a_i, b_i, v_i) and $(\hat{a}_i, \hat{b}_i, \hat{v}_i)$ in part (1) of Theorem 6.1,

$$\hat{\lambda}_1 = \sup_{\hat{v}} \inf_{0 \leq i < N} \left[\hat{a}_{i+1} + \hat{b}_i - \frac{\hat{a}_i}{\hat{v}_{i-1}} - \hat{b}_{i+1} \hat{v}_i \right],$$

and in (5.8) with $N' = N$,

$$\sup_{\hat{v}} \inf_{0 \leq i < N} \left[\hat{a}_{i+1} + \hat{b}_i - \frac{\hat{a}_i}{\hat{v}_{i-1}} - \hat{b}_{i+1} \hat{v}_i \right] = \sup_v \inf_{i \in E} \left[a_i \left(1 - \frac{1}{v_{i-1}} \right) + b_i (1 - v_i) \right],$$

the first assertion of Theorem 7.1 now follows from the variational formula given on the right-hand side of (9.2) in Section 9.

(b) Similarly, replacing the use of Theorem 6.1 by Proposition 2.7 (1), we obtain the second assertion. In this case, as already mentioned at the beginning of Section 4, we have $\lambda_0 = \lambda_0^{(4.2)}$. The fact that $\hat{\lambda}_0 = \lambda_0^{(2.2)} = 0$ comes from (5.5) and Theorem 3.1. $\square$

By Theorem 7.1 (1), all the results obtained in Section 6 can be transformed into the present setup. For instance, by Corollary 6.4, we obtain the following result.

Corollary 7.2. Under (7.2), we have $(4\delta)^{-1} \leq \delta_1^{-1} \leq \lambda_0 \leq \bar{\delta}_1^{-1}$, where

$$\begin{aligned} \delta &= \left[\sup_{1 \leq n < N} \mu[1, n] \left(\nu[n+1, N] + \frac{\mathbb{1}_{\{N < \infty\}}}{\mu_N b_N} \right) \right] \bigvee \frac{\mu[1, N] \mathbb{1}_{\{N < \infty\}}}{\mu_N b_N}, \quad \nu_k = \frac{1}{\mu_k a_k}, \\ \delta_1 &= \sup_{i \in E} (\sqrt{\varphi_i} + \sqrt{\varphi_{i-1}}) \left(\psi_i - \psi_1 \frac{\nu[i+1, N] + (\mu_N b_N)^{-1} \mathbb{1}_{\{N < \infty\}}}{\nu[1, N] + (\mu_N b_N)^{-1} \mathbb{1}_{\{N < \infty\}}} \right), \\ \bar{\delta}_1 &= \sup_{m \in E} \frac{1}{\varphi_m} \left[\sum_{k=1}^{N-1} \nu_{k+1} \varphi_{k \wedge m}^2 + \frac{\varphi_m^2}{\mu_N b_N} \mathbb{1}_{\{N < \infty\}} - \right. \\ &\quad \left. - \frac{1}{\nu[1, N] + (\mu_N b_N)^{-1} \mathbb{1}_{\{N < \infty\}}} \left(\sum_{k=1}^{N-1} \nu_{k+1} \varphi_{k \wedge m} + \frac{\varphi_m}{\mu_N b_N} \mathbb{1}_{\{N < \infty\}} \right)^2 \right], \\ \varphi_i &= \mu[1, i], \quad \psi_i = \sum_{j=i}^{N-1} \nu_{j+1} \sqrt{\varphi_j} + \frac{1}{\mu_N b_N} \sqrt{\varphi_N} \mathbb{1}_{\{N < \infty\}}, \quad i \in E. \end{aligned}$$

Proof. Starting from Corollary 6.4 with its notation, write everything we need in its dual. First by (5.4), we have

$$\mu_n = \hat{a}_1 \hat{\nu}_{n+1}, \quad \nu_n = \frac{1}{\hat{a}_1} \hat{\mu}_{n+1}, \quad 0 \leq n < N, \quad \mu_N = \frac{\hat{a}_1}{\hat{\mu}_N \hat{b}_N} \quad \text{if } N < \infty.$$

Here, recall that $\nu_n = (\mu_n b_n)^{-1}$ but $\hat{\nu}_n = (\hat{\mu}_n \hat{a}_n)^{-1}$. Then the constant δ defined in (4.4) becomes $\delta = \sup_{n \in E} \nu[0, n-1] \mu[n, N]$. Moreover, we have

$$\begin{aligned} \varphi_i &= \sum_{j=0}^{i-1} \nu_j = \frac{1}{\hat{a}_1} \hat{\mu}[1, i], \quad 1 \leq i < N+1, \\ \mu[m, n] &= \sum_{j=m}^n \mu_j = \hat{a}_1 \hat{\nu}[m+1, n+1], \quad 0 \leq m \leq n < N, \\ \mu[m, N] &= \hat{a}_1 \hat{\nu}[m+1, N] + \mu_N \mathbb{1}_{\{N < \infty\}} = \hat{a}_1 \left[\hat{\nu}[m+1, N] + \frac{\mathbb{1}_{\{N < \infty\}}}{\hat{\mu}_N \hat{b}_N} \right], \quad m < N, \\ \psi_i &= \sum_{j=i}^{N-1} \mu_j \sqrt{\varphi_j} + \mu_N \sqrt{\varphi_N} \mathbb{1}_{\{N < \infty\}} \\ &= \sqrt{\hat{a}_1} \left[\sum_{j=i}^{N-1} \hat{\nu}_{j+1} \sqrt{\hat{\mu}[1, j]} + \frac{1}{\hat{\mu}_N \hat{b}_N} \sqrt{\hat{\mu}[1, N]} \mathbb{1}_{\{N < \infty\}} \right]. \end{aligned}$$

Inserting these quantities into (6.8) and (6.9), making a little simplification, and then ignoring the hat everywhere, we obtain Corollary 7.2. $\square$

The next result is a criterion for the positivity of λ_0 , and is a particular case of Corollary 8.4 with $\mathbb{B} = L^1(\mu)$ in the next section. It is not deduced from the last section in terms of duality (Theorem 7.1) but conversely, it provides an improvement of Theorem 6.2 as shown by the proof of Corollary 6.6 below.

Corollary 7.3 (Criterion and basic estimates). Without condition (7.2), we have $\kappa^{-1}/4 \leq \lambda_0 \leq \kappa^{-1}$, where

$$\kappa^{-1} = \inf_{1 \leq n \leq m < N+1} \left[\left(\sum_{i=1}^n \frac{1}{\mu_i a_i} \right)^{-1} + \left(\sum_{i=m}^N \frac{1}{\mu_i b_i} \right)^{-1} \right] \left(\sum_{j=n}^m \mu_j \right)^{-1}. \quad (7.5)$$

Furthermore, we have

$$\delta_L \wedge \delta_R \geq \kappa \geq (\mathbb{1}_{\{S=\infty\}} + (a_1 S)^{-1})(\delta_L \wedge \delta_R),$$

where

$$S = \sum_{i=1}^N \frac{1}{\mu_i a_i} + \frac{1}{\mu_N b_N} \mathbb{1}_{\{N < \infty\}},$$

$$\delta_L = \sup_{1 \leq n < N+1} \sum_{i=1}^n \frac{1}{\mu_i a_i} \sum_{j=n}^N \mu_j, \quad \delta_R = \sup_{1 \leq m < N+1} \sum_{j=1}^m \mu_j \sum_{k=m}^N \frac{1}{\mu_k b_k}.$$

Note that $\delta_L = \delta^{(4.4)}$ and δ_R almost coincides with $\delta^{(3.1)}$, except for δ_R there is a shift of the state space. The second assertion of Corollary 7.3 means that $\lambda_0 > 0$ iff the process goes to either 0 or $N+1$ exponentially fast. This is intuitively clear by (7.1). Obviously, we have $\lambda_0 = \kappa^{-1} = 0$ if $\sum_i \mu_i = \infty$ and $\sum_j (\mu_j a_j)^{-1} = \infty$ since then $\delta_L = \delta_R = \infty$. See also Corollary 8.6 below.

Proof of Corollary 6.6. For given rates (a_i, b_i) in the setup of Section 6, by (5.3), we have

$$\sum_{i=p}^q \hat{\nu}_i = \frac{1}{b_0} \sum_{i=p}^q \mu_{i-1} = \frac{1}{b_0} \sum_{i=p-1}^{q-1} \mu_i, \quad \sum_{j=p}^q \hat{\mu}_j = b_0 \sum_{j=p}^q \nu_{j-1} = b_0 \sum_{j=p-1}^{q-1} \nu_j$$

and

$$\sum_{i=m}^N \frac{1}{\hat{\mu}_i \hat{b}_i} = \sum_{i=m}^{N-1} \hat{\nu}_{i+1} + \frac{1}{\hat{\mu}_N \hat{b}_N} \mathbb{1}_{\{N < \infty\}} = \frac{1}{b_0} \sum_{i=m}^{N-1} \mu_i + \frac{1}{b_0} \mu_N \mathbb{1}_{\{N < \infty\}} = \frac{1}{b_0} \sum_{i=m}^N \mu_i.$$

Regarding the process studied in Corollary 7.3 as a dual of the one given in the last section and then add a hat to each quantity of Corollary 7.3. It follows that

$$\begin{aligned} \hat{\kappa}^{-1} &= \inf_{1 \leq n \leq m < N+1} \left[\left(\sum_{i=1}^n \hat{\nu}_i \right)^{-1} + \left(\sum_{i=m}^N \frac{1}{\hat{\mu}_i \hat{b}_i} \right)^{-1} \right] \left(\sum_{j=n}^m \hat{\mu}_j \right)^{-1} \\ &= \inf_{1 \leq n \leq m < N+1} \left[\left(\sum_{i=0}^{n-1} \mu_i \right)^{-1} + \left(\sum_{i=m}^N \mu_i \right)^{-1} \right] \left(\sum_{j=n-1}^{m-1} \nu_j \right)^{-1} \\ &= \inf_{0 \leq n < m < N+1} \left[\left(\sum_{i=0}^n \mu_i \right)^{-1} + \left(\sum_{i=m}^N \mu_i \right)^{-1} \right] \left(\sum_{j=n}^{m-1} \nu_j \right)^{-1} \\ &= \kappa^{-1}. \end{aligned}$$

Next, we have

$$\hat{a}_1 \hat{S} = b_0 \left(\sum_{i=1}^N \hat{\nu}_i + \frac{1}{\hat{\mu}_N \hat{b}_N} \mathbb{1}_{\{N < \infty\}} \right) = \sum_{i=0}^N \mu_i = Z,$$

$\hat{\delta}_L = \delta_R$, and $\hat{\delta}_R = \delta_L$. Since $\sum_i \mu_i < \infty$, by Theorem 7.1 (1), we have $\lambda_1 = \hat{\lambda}_0$. Thus, Corollary 6.6 now follows from Corollary 7.3 immediately except for a slight change of the lower bound in the second assertion. For which, since $Z < \infty$, the term “ $\wedge \delta_R$ ” is not needed (cf. Proof of Corollary 8.4). $\square$

Proof of Theorem 1.5. (a) Condition (1.3) implies that $N = \infty$ and furthermore the uniqueness of the symmetric process on $L^2(\mu)$ by Proposition 1.3. Now, $\alpha^* = \lambda_1$ or λ_0 by [2; Theorem 5.3] or Proposition 1.2, respectively.

(b) In the case that $\sum_i \mu_i = \infty$ and $\sum_i (\mu_i b_i)^{-1} = \infty$, the process is zero-recurrent and so we have $\alpha^* = 0$. Noting that $\delta^{(3.1)}$, $\delta^{(4.4)}$, $\kappa^{(6.13)}$, and $\kappa^{(7.3)}$ are all equal to infinity, the conclusions of the theorem become obvious. Hence, in what follows, we may assume that only one of $\sum_i \mu_i$ and $\sum_i (\mu_i b_i)^{-1}$ is equal to infinity.

(c) Let $b_0 = 0$. Then the basic estimate follows from Corollary 7.3.

(d) We now prove the first two parts of the theorem under the assumption that $b_0 = 0$. In the case that $\sum_i \mu_i < \infty$ but $\sum_i (\mu_i a_i)^{-1} = \infty$, we have $\kappa^{(7.5)} = \delta^{(4.4)}$ which gives us part (1) of the theorem. Next, if $\sum_i \mu_i = \infty$ but $\sum_i (\mu_i a_i)^{-1} < \infty$, then for δ_L and δ_R given in Corollary 7.3, we have $\delta_L = \delta^{(4.4)} = \infty$ and then $\kappa^{(7.5)} < \infty$ iff $\delta_R < \infty$. Clearly, $\delta_R < \infty$ iff $\delta^{(3.1)} < \infty$ since $\sum_i (\mu_i a_i)^{-1} < \infty$. This gives us part (2) of the theorem.

(e) Finally, let $b_0 > 0$. This is a dual case of that $b_0 = 0$ treated in (c) and (d). By exchanging the measures μ and ν , we obtain the remaining conclusions of the theorem.

Actually, part (1) of the theorem is a combination of Theorems 4.2 and 6.2, and part (2) is a combination of Theorems 3.1 and 7.1. $\square$

We are now ready to prove an extension of Theorem 1.5.

Theorem 7.4 (Criterion and basic estimates). Without condition (1.3), Theorem 1.5 remains true provided

- (1) the process in part (1) is replaced by the maximal one and α^* is replaced by $\alpha^{\max}$: the largest ε such that

$$\sum_j |p_{ij}(t) - \pi_j| \leq C_i e^{-\varepsilon t}, \quad t \geq 0, \quad (7.6)$$

for some $L^1(\pi)$ -locally integrable function C_i depending on i only; and

- (2) the process in part (2) needs no change (i.e., the minimal one).

Proof. Since λ_1 is equivalent to $\lambda_0^{(4.2)}$ (Theorem 6.2) and by duality, $\lambda_1 = \lambda_0^{(7.1)}$ and $\lambda_0^{(4.2)} = \lambda_0^{(2.2)}$, it is clear that $\lambda_0^{(7.1)}$ is equivalent to $\lambda_0^{(2.2)}$. Alternatively, one can use Corollary 7.3 to arrive at the same conclusion. Now, part (2) of

the theorem follows by Proposition 1.2 for which we do not assume (1.3). As mentioned in the last proof, part (1) with the original α^* also follows by [2] provided (1.3) holds.

Even though in the previous study ([12], for instance), we consider only the ergodic processes under (1.2), but λ_1 can be actually identified with some exponentially ergodic convergence rate for more general ergodic processes (reversible Markov chains, in particular). First, the fact that the L^2 -exponential convergence rate is described by λ_1 does not require the regularity of the Dirichlet form (cf. proof of Proposition 1.1, for instance). Next, for a Markov process with state space $(X, \mathcal{X}, \pi)$ and transition probabilities $\{P_t(x, \cdot)\}$, let $\tilde{\varepsilon}_1$ be the largest ε such that

$$\|P_t(x, \cdot) - \pi\|_{\text{Var}} \leq C(x) e^{-\varepsilon t}, \quad x \in X, t \geq 0, \quad (7.7)$$

for some $L^1(X, \pi)$ -locally integrable function $C(x)$ depending on x only. Then for a reversible process having density $p_t(x, y)$, we have $\lambda_1 = \tilde{\varepsilon}_1$ provided

$$p_s(\cdot, \cdot) \in L_{\text{loc}}^{1/2}(X, \pi) \quad \text{for some } s > 0, \quad (7.8)$$

and the set of bounded functions with compact support is dense in $L^2(X, \pi)$. The outline of the proof is as follows.

- (i) Prove that $\tilde{\varepsilon}_1 \geq \lambda_1$.
- (ii) Show that $\|(P_t - \pi)f\|^2 \leq C_f e^{-\tilde{\varepsilon}_1 t}$ for bounded f with compact support.
- (iii) Remove the constant C_f in the last line for fixed f .
- (iv) Extend f to $L^2(X, \pi)$ and then claim that $\lambda_1 \geq \tilde{\varepsilon}_1$.

By assumption, the last step is obvious. The detailed proof for the first three steps is given, respectively, in [12]: (8.6), the last formula in §8.3 replacing ε_1 with $\tilde{\varepsilon}_1$, and the proof of Lemma 8.12. Actually, this is a small correction to [12; Theorem 8.13 (4)] (i.e., replacing ε_1 by $\tilde{\varepsilon}_1$) and its proof. It is known that $\tilde{\varepsilon}_1 > 0$ iff $\varepsilon_1 > 0$ (as well as $\varepsilon_2 > 0$ used in the original proof of the cited theorem). Hence, the exponential ergodicity is kept but the rates $\varepsilon_1 \geq \tilde{\varepsilon}_1 \geq \varepsilon_2$ may be different. By the way, we mention that the change of topology is necessary in many cases. For instance, the pointwise convergence is natural in the discrete case but is not in the continuous case. In the ergodic situation, the total variation norm is good enough in general but it is meaningless in the non-ergodic case.

Having this result at hand, part (1) of the theorem follows since we have $\alpha^{\max} = \tilde{\varepsilon}_1$ in the present context. $\square$

We now introduce an interpretation, similar to Section 5, of the duality used in Theorem 7.1. For the ergodic process with Q -matrix,

$$Q = \begin{pmatrix} -b_0 & b_0 & 0 & 0 \\ a_1 & -a_1 - b_1 & b_1 & 0 \\ 0 & a_2 & -a_2 - b_2 & b_2 \\ 0 & 0 & a_3 & -a_3 \end{pmatrix}, \quad a_i, b_i > 0,$$

we have a simpler transformation matrix

$$M = \begin{pmatrix} \mu_0 & \mu_1 & \mu_2 & \mu_3 \\ 0 & \mu_1 & \mu_2 & \mu_3 \\ 0 & 0 & \mu_2 & \mu_3 \\ 0 & 0 & 0 & \mu_3 \end{pmatrix} \implies M^{-1} = \begin{pmatrix} \mu_0^{-1} & -\mu_0^{-1} & 0 & 0 \\ 0 & \mu_1^{-1} & -\mu_1^{-1} & 0 \\ 0 & 0 & \mu_2^{-1} & -\mu_2^{-1} \\ 0 & 0 & 0 & \mu_3^{-1} \end{pmatrix}.$$

Then

$$MQM^{-1} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ b_0 & -a_1 - b_0 & a_1 & 0 \\ 0 & b_1 & -a_2 - b_1 & a_2 \\ 0 & 0 & b_2 & -a_3 - b_2 \end{pmatrix} \\ = \begin{pmatrix} 0 & 0 & 0 & 0 \\ \hat{a}_1 & -\hat{a}_1 - \hat{b}_1 & \hat{b}_1 & 0 \\ 0 & \hat{a}_2 & -\hat{a}_2 - \hat{b}_2 & \hat{b}_2 \\ 0 & 0 & \hat{a}_3 & -\hat{a}_3 - \hat{b}_3 \end{pmatrix}.$$

We obtain a process having an absorbing state at 0 and being killed at the state 3. The original trivial eigenvalue with non-zero constant eigenfunction is transferred into the trivial one with eigenfunction $\mathbb{1}_{\{0\}}$. Our dual matrix $\hat{Q}$ is now obtained by eliminating the first row and the first column from the matrix on the right-hand side. The elimination is to make the symmetrizability of $\hat{Q}$ and at the same time removes the trivial eigenvalue of the last matrix. Unlike the example given in Section 5 where the size of the state space stays the same: $\{0, 1, 2, 3\} \rightarrow \{1, 2, 3, 4\}$ with a shift for the dual one, here the size of the state space is reduced by one: $\{0, 1, 2, 3\} \rightarrow \{1, 2, 3\}$.

We are now ready to examine some examples.

Examples 7.5. (1) Let $N = 1$. Then the Q -matrix is degenerated to be a single killing $-c$ and so $\lambda_0 = c = \kappa^{-1}$.

(2) Let $N = 2$. Then

$$\lambda_0 = \frac{1}{2} \left(a_1 + a_2 + b_1 + b_2 - \sqrt{(a_1 - a_2 + b_1 - b_2)^2 + 4a_2b_1} \right).$$

The next two examples are taken from Chen, Zhang and Zhao (2003, Examples 2.2 and 2.3)

Examples 7.6. (1) Let $N = 2$, $a_1 = a_2 = 1$, $b_1 = 2$, and $b_2 = 3$. Then $\lambda_0 = 2$, and by Corollary 7.2, we have

$$\bar{\delta}_1 \leq \lambda_0^{-1} = 0.5 \leq \delta_1,$$

where

$$\delta_1 = \frac{4 + \sqrt{3}}{10} \approx 0.573, \quad \bar{\delta}_1 = \frac{7}{15} = 0.4\dot{6}, \quad \frac{\delta_1}{\bar{\delta}_1} \approx 1.23.$$

Next, $\kappa \leq \lambda_0^{-1} \leq 4\kappa$ with $\kappa = 3/7$. Obviously, $(\bar{\delta}_1, \delta_1) \subset (\kappa, 4\kappa)$.

(2) Let $N = 2$, $b_1 = 1$, $b_2 = 2$,

$$a_1 = \frac{2 - \varepsilon^2}{1 + \varepsilon}, \quad \varepsilon \in [0, \sqrt{2}),$$

and $a_2 = 1$. Then $\lambda_0 = 2 - \varepsilon$, and we have

$$\bar{\delta}_1 \leq \lambda_0^{-1} = (2 - \varepsilon)^{-1} \leq \delta_1,$$

where

$$\delta_1 = \frac{4 + \sqrt{2} + (2 + \sqrt{2})\varepsilon - \varepsilon^2}{8 + 2\varepsilon - 3\varepsilon^2} = \frac{1}{\lambda_0} + \frac{(1 + \varepsilon)(\sqrt{2} - \varepsilon)}{8 + 2\varepsilon - 3\varepsilon^2},$$

$$\bar{\delta}_1 = \frac{8 + 6\varepsilon - \varepsilon^2}{16 + 4\varepsilon - 6\varepsilon^2} = \frac{1}{\lambda_0} - \frac{\varepsilon^2}{2(8 + 2\varepsilon - 3\varepsilon^2)}.$$

Hence,

$$\frac{\delta_1}{\bar{\delta}_1} = 2 - \frac{2(4 - \sqrt{2})(1 + \varepsilon)}{8 + 6\varepsilon - \varepsilon^2} < 1.354.$$

Next, $\kappa \leq \lambda_0^{-1} \leq 4\kappa$ with

$$\kappa = \frac{1}{\lambda_0} - \begin{cases} \varepsilon^2(8 - 4\varepsilon^2 + \varepsilon^3)^{-1} & \text{if } \varepsilon \in [0, (\sqrt{13} - 1)/3] \\ (8 + 2\varepsilon - 3\varepsilon^2)^{-1} & \text{if } \varepsilon \in [(\sqrt{13} - 1)/3, \sqrt{2}]. \end{cases}$$

Even though it is not so obvious now but we do have $(\bar{\delta}_1, \delta_1) \subset (\kappa, 4\kappa)$.

Examples 7.7. Because of Theorem 7.1, we can now transfer [10; Examples 9.27] into the present context, see Table 7.1, by using (5.1) and (5.7). Here, for the sixth example, we need a restriction: $1/k < b/a \leq k/(k-1)^2$ ($k \geq 2$).

Table 7.1 Exact λ_0 for nine examples

$\mathbf{a_i} (i \geq 1)$	$\mathbf{b_i} (i \geq 1)$	$\mathbf{\lambda_0}$	$\mathbf{v_i} (i \geq 1)$
a	$b (a < b)$	$(\sqrt{a} - \sqrt{b})^2$	$\sqrt{a/b}$
$\gamma_1(i-1) + \gamma_0$ $\gamma_0 > 0, \gamma_1 \geq 0$	$\beta_1 i (\beta_1 > \gamma_1)$	$\beta_1 - \gamma_1$	$\frac{\gamma_1 i + \gamma_0}{\beta_1 i}$
$i - 1 + \beta_0$ $\beta_0 > 0$	$2(i+1) + \beta_0$	2	$\frac{(i+1)(i+\beta_0)}{i[2(i+1)+\beta_0]}$
i	$2i+4+\sqrt{2}$	3	$\frac{i+1}{2i+4+\sqrt{2}} \left[1 + \frac{2(i+\sqrt{2})}{i(i+2\sqrt{2}-1)} \right]$
$\frac{a}{i}$	b	$b - \frac{\sqrt{a^2+4ab}-a}{2}$	$\frac{\sqrt{a^2+4ab}+a}{2bi}$
a	$(i \wedge k)b$	$(\sqrt{bk} - \sqrt{a})^2$	$\frac{1}{i \wedge k} \sqrt{ak/b}$
$i+1$	i^2	2	i^{-1}
$(i-1)^2 (i \geq 2)$ $a_1 > 0$	i^2	$\frac{1}{4}$	$\frac{2i+1}{2(i+1)}$
$2 + (-1)^{i-1}$ $(i \geq 2)$ $a_1 = \frac{7-\sqrt{33}}{2}$	$2[2+(-1)^i]$	$6 - \sqrt{33}$	$\frac{\sqrt{33} + (-1)^i}{8}$

We now go to the second part of this section. Consider the birth-death processes with a more general state space $E = \{i : -M-1 < i < N+1\}$, $M, N \leq \infty$

and with Dirichlet boundaries at $-M-1$ if $M > -\infty$ and at $N+1$ if $N < \infty$. Its Q -matrix now is $q_{i,i+1} = b_i > 0$, $q_{i,i-1} = a_i > 0$, and $q_{ij} = 0$ if $|i-j| > 1$ for $i, j \in E$. Fix a reference point $\theta \in E$. Define

$$\begin{aligned}\mu_{\theta+n} &= \frac{a_{\theta-1}a_{\theta-2} \cdots a_{\theta+n+1}}{b_{\theta}b_{\theta-1} \cdots b_{\theta+n}}, & -M-1-\theta < n \leq -2, \\ \mu_{\theta-1} &= \frac{1}{b_{\theta}b_{\theta-1}}, & \mu_{\theta} &= \frac{1}{a_{\theta}b_{\theta}}, & \mu_{\theta+1} &= \frac{1}{a_{\theta}a_{\theta+1}}, \\ \mu_{\theta+n} &= \frac{b_{\theta+1}b_{\theta+2} \cdots b_{\theta+n-1}}{a_{\theta}a_{\theta+1} \cdots a_{\theta+n}}, & 2 \leq n < N+1-\theta.\end{aligned}\quad (7.9)$$

Correspondingly,

$$\begin{aligned}D(f) &= \sum_{-M-1 < i \leq \theta} \mu_i a_i (f_i - f_{i-1})^2 + \sum_{\theta \leq i < N+1} \mu_i b_i (f_{i+1} - f_i)^2, \\ f &\in \mathcal{K}, \quad f_{-M-1} = 0 \text{ if } M < \infty \text{ and } f_{N+1} = 0 \text{ if } N < \infty.\end{aligned}\quad (7.10)$$

Let us begin with a particular application of Corollary 8.4 to $\mathbb{B} = L^1(\mu)$.

Corollary 7.8 (Criterion and basic estimates). Let λ_0 be defined by (7.1) with the present state space E . Then we have $\kappa^{-1}/4 \leq \lambda_0 \leq \kappa^{-1}$, where

$$\kappa^{-1} = \inf_{m, n \in E: m \leq n} \left[\left(\sum_{i=-M}^m \frac{1}{\mu_i a_i} \right)^{-1} + \left(\sum_{i=n}^N \frac{1}{\mu_i b_i} \right)^{-1} \right] \left(\sum_{j=m}^n \mu_j \right)^{-1}. \quad (7.11)$$

By the way, we extend Corollary 6.6 to the present general state space.

Corollary 7.9 (Criterion and basic estimates). Let $\sum_{i \in E} \mu_i < \infty$ and define λ_1 as in (6.1). Then we have $\kappa^{-1}/4 \leq \lambda_1 \leq \kappa^{-1}$, where

$$\kappa^{-1} = \inf_{m, n \in E: m < n} \left[\left(\sum_{i=-M}^m \mu_i \right)^{-1} + \left(\sum_{i=n}^N \mu_i \right)^{-1} \right] \left(\sum_{j=m}^{n-1} \frac{1}{\mu_j b_j} \right)^{-1}. \quad (7.12)$$

Proof. When $M < \infty$, the corollary is simply a modification of Corollary 6.6 by shifting the left end-point of the state space from 0 to $-M$. Thus, when $M = \infty$, we can choose a sequence $\{M_p\}_{p=1}^{\infty}$ such that $M_p \uparrow \infty$ as $p \uparrow \infty$ and then the assertion holds if M is replaced by M_p for each p . In which case, the corresponding λ_1 is denoted by $\lambda_1^{(M_p)}$ for a moment. Because $\sum_{i \in E} \mu_i < \infty$, following the proof above (4.2), it follows that

$$\lambda_1 = \inf \{ D(f) : \mu(f^2) = 1, \mu(f) = 0, f_i = f_{(i \vee m) \wedge n} \text{ for some } m, n \in E, m < n \}.$$

Hence, we have $\lambda_1^{(M_p)} \downarrow \lambda_1$ as $p \uparrow \infty$. Similarly, replacing M by M_p , we have the notation $\kappa^{(M_p)}$. The proof will be done once we show that

$$\left(\kappa^{(M_p)} \right)^{-1} \downarrow \kappa^{-1} \quad \text{as } p \uparrow \infty.$$

Obviously, we have

$$(\kappa^{(M_p)})^{-1} \downarrow \quad \text{as } p \uparrow \quad \text{and} \quad (\kappa^{(M_p)})^{-1} \geq \kappa^{-1}.$$

To prove the required assertion, let $\varepsilon > 0$. Then by definition of κ there exist $m_0, n_0 \in E, m_0 < n_0$ such that

$$\left[\left(\sum_{i=-M}^{m_0} \mu_i \right)^{-1} + \left(\sum_{i=n_0}^N \mu_i \right)^{-1} \right] \left(\sum_{j=m_0}^{n_0-1} \frac{1}{\mu_j b_j} \right)^{-1} \leq \kappa^{-1} + \varepsilon.$$

Next, since $\sum_i \mu_i < \infty$, for fixed m_0, n_0 and large enough M_p ($-M_p < m_0$), we have

$$\left[\left(\sum_{i=-M_p}^{m_0} \mu_i \right)^{-1} + \left(\sum_{i=n_0}^N \mu_i \right)^{-1} \right] \left(\sum_{j=m_0}^{n_0-1} \frac{1}{\mu_j b_j} \right)^{-1} \leq \kappa^{-1} + 2\varepsilon.$$

Combining these facts with the definition of $\kappa^{(M_p)}$, we obtain

$$\begin{aligned} \kappa^{-1} &\leq (\kappa^{(M_p)})^{-1} \\ &= \inf_{-M_p \leq m < n < N+1} \left[\left(\sum_{i=-M_p}^m \mu_i \right)^{-1} + \left(\sum_{i=n}^N \mu_i \right)^{-1} \right] \left(\sum_{j=m}^{n-1} \frac{1}{\mu_j b_j} \right)^{-1} \\ &\leq \kappa^{-1} + 2\varepsilon. \end{aligned}$$

Since ε is arbitrary, we have proved that $(\kappa^{(M_p)})^{-1} \rightarrow \kappa^{-1}$ as $p \rightarrow \infty$. $\square$

For the remainder of this section, we study a splitting technique. It provides a different tool to study the problem having bilateral Dirichlet boundaries. This approach is especially meaningful if the duality discussed in Section 5 does not work, such as in studying the processes on the whole $\mathbb{Z}$ or the Poincaré-type inequalities given in the next section. We remark that Corollaries 7.8 and 7.9 use slightly the splitting idea only (cf. Proof (b) of Theorem 8.2 below). The idea is splitting the state space into two parts and then estimating the first (non-trivial) eigenvalue in terms of the local ones. We have used this technique several times before: Chen and Wang (1998) with Dirichlet boundary for the unbounded region, Chen, Zhang and Zhao (2003), as well as Mao and Xia (2009), with Neumann boundary. The first and the third papers work on a very general setup. Here, we follow the second one with some addition.

To state our result, we need to construct two birth-death processes on the left- and the right-hand sides, respectively, for a given birth-death process with rates (a_i, b_i) and state space E . Fix a constant $\gamma > 1$.

- (L) The process on the left-hand side has state space $E^{\theta-} = \{i : -M-1 < i \leq \theta\}$, reflects at θ . Its transition structure is the same as the original one except a_θ is replaced by γa_θ .
- (R) The process on the right-hand side has state space $E^{\theta+} = \{i : \theta \leq i < N+1\}$, reflects at θ . Its transition structure is again the same as the original one except b_θ is replaced by $\gamma(\gamma-1)^{-1}b_\theta$.

For the process on the right-hand side, the state θ is a Neumann boundary. At $N + 1$, it is a Dirichlet boundary if $N < \infty$. For this process, the first eigenvalue, denoted by $\lambda_0^{\theta+, \gamma}$, has already been studied in Sections 2 and 3. With a change of the order of the state space, it follows that the process on the left-hand side has the same boundary condition, denote by $\lambda_0^{\theta-, \gamma}$ its first eigenvalue. Note that ignoring a finite number of the states does not change the positivity of $\lambda_0^{\theta\pm, \gamma}$, in the qualitative case, we simply denote them by $\lambda_0^{(\pm)}$, respectively. In general, according to $\sum_{i=\theta}^N (\mu_i b_i)^{-1}$ and/or $\sum_{i=-M}^{\theta} (\mu_i a_i)^{-1}$, $\sum_{i=\theta}^N \mu_i$ and/or $\sum_{i=-M}^{\theta} \mu_i$ being finite or not, there are eight cases for the processes on $\mathbb{Z}$. For instance, if $\sum_{i=\theta}^N (\mu_i b_i)^{-1} = \infty$, then $\lambda_0^{(+)} = 0$ by Theorem 3.1. Since in this section, we are working on bilateral Dirichlet boundaries, it is natural to assume that $\lambda_0^{(\pm)} > 0$. The other cases may be treated in a parallel way. For instance, when $\lambda_0^{(-)} = 0$, it is more natural to consider the process on $[-M, N + 1)$ with reflecting at some finite $-M$ and then pass to the limit as $M \rightarrow \infty$ (cf. the proof of Corollary 7.8). In this case, the eigenfunction should be strictly decreasing once $\lambda_0 > 0$. Hence, there is no reason to use the splitting technique. Note that the explicit criterion for $\lambda_0^{(\pm)} > 0$ is given by Theorem 3.1. We can now state the main result of the second part of this section as follows.

Theorem 7.10.

- (1) In general, the Dirichlet eigenvalue λ_0 of the birth–death process on $E = \{i : -M - 1 < i < N + 1\}$ satisfies

$$\inf_{\theta \in E} \inf_{\gamma > 1} (\lambda_0^{\theta-, \gamma} \vee \lambda_0^{\theta+, \gamma}) = \lambda_0 \geq \sup_{-M-1 \leq \theta \leq N+1} \sup_{\gamma > 1} (\lambda_0^{\theta-, \gamma} \wedge \lambda_0^{\theta+, \gamma}), \quad (7.13)$$

where on the right-hand, when $\theta = -M - 1$, define $\lambda_0^{\theta-, \gamma} = \infty$, and $\lambda_0^{\theta+, \gamma}$ to be the first eigenvalue of the original process (independent of γ) reflected at $-M$ if $M < \infty$; when $\theta = N + 1$, define $\lambda_0^{\theta+, \gamma} = \infty$, and $\lambda_0^{\theta-, \gamma}$ to be the one reflected at N if $N < \infty$.

- (2) The second equality in (7.13) replacing $\sup_{-M-1 \leq \theta \leq N+1}$ by $\sup_{\theta \in E}$ also holds provided $\lambda_0^{(\pm)} > 0$, and moreover,

$$\sum_{i=-M}^{\theta} \mu_i = \infty \text{ if } M = \infty \quad \text{and} \quad \sum_{i=\theta}^N \mu_i = \infty \text{ if } N = \infty. \quad (7.14)$$

Theorem 7.10 was proved in Chen, Zhang and Zhao (2003) for the half-space (i.e., one of M and N is finite), under the hypotheses that $\sum_i (\mu_i b_i)^{-1} < \infty$ and $\sum_i \mu_i < \infty$ which is essentially the case of having a finite state space.

To prove Theorem 7.10, we need some preparation. First, we couple these two processes on a common state space $\bar{E} = \{i : -M - 1 < i < N + 2\}$. Next, separate the two processes by shifting the state space $E^{\theta+}$ by one to the right: $1 + E^{\theta+}$. Denote by $(\bar{a}_i, \bar{b}_i)$ the rates of the connected process. For this, we need to build a bridge for the processes on the two sides by adding two more rates $\bar{b}_\theta = \gamma - 1$ and $\bar{a}_{\theta+1} = 1$. The construction here will become clear once we have a deeper

understanding about the eigenfunction and it will be explained in Part II of the proof of the theorem. Roughly speaking, there are two possible shapes of the eigenfunction, the construction enables us to transform one of them to the other so that the splitting with Neumann boundaries becomes practical. For which, one needs the parameter γ as shown in Lemma 7.12 below. In detail, we now have

$$\bar{a}_i = \begin{cases} a_i, & -M-1 < i \leq \theta-1, \\ \gamma a_\theta, & i = \theta, \\ 1, & i = \theta+1, \\ a_{i-1}, & \theta+2 \leq i < N+2, \end{cases} \quad \bar{b}_i = \begin{cases} b_i, & -M-1 < i \leq \theta-1, \\ \gamma-1, & i = \theta, \\ \frac{\gamma b_\theta}{\gamma-1}, & i = \theta+1, \\ b_{i-1}, & \theta+2 \leq i < N+2. \end{cases}$$

Applying (7.9) to the present setup and removing the factor $b_\theta(1-\gamma)^{-1}$ (which simplifies the notation but does not change the ratio $\bar{D}(f)/\bar{\mu}(f^2)$), we obtain

$$\begin{aligned} \bar{\mu}_i &= \mu_i, & -M-1 < i \leq \theta-1, \\ \bar{\mu}_\theta &= \frac{1}{\gamma} \mu_\theta, & \bar{\mu}_{\theta+1} &= \frac{\gamma-1}{\gamma} \mu_\theta, \\ \bar{\mu}_i &= \mu_{i-1}, & \theta+2 \leq i < N+2. \end{aligned} \quad (7.15)$$

Then

$$\bar{\mu}_i \bar{a}_i = \mu_i a_i, \quad i \leq \theta, \quad \bar{\mu}_\theta \bar{b}_\theta = \frac{\gamma-1}{\gamma} \mu_\theta, \quad \bar{\mu}_i \bar{b}_i = \mu_{i-1} b_{i-1}, \quad i \geq \theta+1. \quad (7.16)$$

The next two results are basic in using the splitting technique.

Lemma 7.11. Given f on E , define $\bar{f}$ on $\bar{E}$ as follows: $\bar{f}_i = f_i$ for $i \leq \theta$ and $\bar{f}_i = f_{i-1}$ for $i \geq \theta+1$. Then we have $\bar{\mu}(f^2) = \mu(f^2)$ and $\bar{D}(\bar{f}) = D(f)$.

Proof. Clearly, we have $\bar{f}_\theta = \bar{f}_{\theta+1}$. Then

$$\begin{aligned} \bar{\mu}(\bar{f}^2) &= \sum_{-M-1 < i \leq \theta-1} \mu_i f_i^2 + (\bar{\mu}_\theta + \bar{\mu}_{\theta+1}) f_\theta^2 + \sum_{\theta+2 \leq i < N+2} \mu_{i-1} f_{i-1}^2 = \mu(f^2), \\ \bar{D}(\bar{f}) &= \sum_{-M-1 < i \leq \theta} \bar{\mu}_i \bar{a}_i (\bar{f}_i - \bar{f}_{i-1})^2 + \sum_{\theta+1 \leq i < N+2} \bar{\mu}_i \bar{b}_i (\bar{f}_{i+1} - \bar{f}_i)^2 \\ &= \sum_{-M-1 < i \leq \theta} \mu_i a_i (f_i - f_{i-1})^2 + \sum_{\theta+1 \leq i < N+2} \mu_{i-1} b_{i-1} (f_i - f_{i-1})^2 \\ &= D(f) \quad (\text{by (7.10)}). \quad \square \end{aligned} \quad (7.17)$$

Lemma 7.12. For a given birth-death process with state space E and rates (a_i, b_i) , if its eigenfunction g of λ satisfies $g_{\theta-1} < g_\theta > g_{\theta+1}$ (resp. $g_{\theta-1} > g_\theta < g_{\theta+1}$) for some $\theta \in E$ (of course, $g_{-M-1} = 0$ if $M < \infty$, and $g_{N+1} = 0$ if $N = \infty$), let

$$\gamma = 1 + \frac{b_\theta(g_\theta - g_{\theta+1})}{a_\theta(g_\theta - g_{\theta-1})} > 1, \quad (7.18)$$

and let $\bar{g}_i = g_i$ for $i \leq \theta$, $\bar{g}_i = g_{i-1}$ for $i \geq \theta+1$. Then for the $(\bar{a}_i, \bar{b}_i)$ -process, $\bar{g}$ is the eigenfunction of $\bar{\lambda} = \lambda$ having the property $\bar{g}_{\theta+1} = \bar{g}_\theta$. Furthermore, $\bar{g}|_{(-M-1, \theta]}$ is the eigenfunction of $\bar{\lambda}$ of the process on the left-hand side reflecting at θ , and similarly $\bar{g}|_{[\theta+1, N+1)}$ is the eigenfunction of the process on the right-hand side reflecting at $\theta+1$.

Proof. By the construction of $(\bar{a}_i, \bar{b}_i)$ and $\bar{g}$, we have

$$\bar{\Omega} \bar{g}(i) = \begin{cases} \Omega g(i) = -\lambda g_i = -\bar{\lambda} \bar{g}_i, & i \leq \theta - 1, \\ \Omega g(i-1) = -\lambda g_{i-1} = -\bar{\lambda} \bar{g}_i, & i \geq \theta + 2. \end{cases}$$

Next, by (7.15), we have

$$\begin{aligned} \bar{\Omega} \bar{g}(\theta) &= \bar{b}_\theta(\bar{g}_{\theta+1} - \bar{g}_\theta) + \bar{a}_\theta(\bar{g}_{\theta-1} - \bar{g}_\theta) = \bar{a}_\theta(g_{\theta-1} - g_\theta) = \gamma a_\theta(g_{\theta-1} - g_\theta), \\ \bar{\Omega} \bar{g}(\theta+1) &= \bar{b}_{\theta+1}(\bar{g}_{\theta+2} - \bar{g}_{\theta+1}) + \bar{a}_{\theta+1}(\bar{g}_\theta - \bar{g}_{\theta+1}) \\ &= \bar{b}_{\theta+1}(g_{\theta+1} - g_\theta) \\ &= \frac{\gamma}{\gamma-1} b_\theta(g_{\theta+1} - g_\theta). \end{aligned}$$

In the first formula, the term containing $\bar{b}_\theta$ vanishes. This is the reason why we can regard θ as a reflecting boundary for the process on the left-hand side. Similarly, one can regard $\theta+1$ as the one for the process on the right-hand side in view of the second formula. By (7.18), the right-hand sides are the same which is equal to

$$\left[1 + \frac{b_\theta(g_\theta - g_{\theta+1})}{a_\theta(g_\theta - g_{\theta-1})} \right] a_\theta(g_{\theta-1} - g_\theta) = a_\theta(g_{\theta-1} - g_\theta) + b_\theta(g_\theta - g_{\theta+1}) = -\lambda g_\theta = -\bar{\lambda} \bar{g}_\theta.$$

We have thus proved the lemma. $\square$

Proof of Theorem 7.10. Part I. In this part, we prove Theorem 7.10 (1) with the first “=” replaced by “ $\geq$ ”. The proof of this part is relatively easier. Let $f \in \mathcal{K}$, $f \neq 0$. Define $\bar{f}$ as in Lemma 7.11. For fixed $\theta \in E$ and $\gamma > 1$, noting that re-labeling the state space does not change $\lambda_0^{\theta+, \gamma}$, by (7.17), we have

$$\bar{D}(\bar{f}) \geq \lambda_0^{\theta-, \gamma} \sum_{i \leq \theta} \bar{\mu}_i \bar{f}_i^2 + \lambda_0^{\theta+, \gamma} \sum_{i \geq \theta+1} \bar{\mu}_i \bar{f}_i^2 \geq (\lambda_0^{\theta-, \gamma} \wedge \lambda_0^{\theta+, \gamma}) \bar{\mu}(\bar{f}^2).$$

Hence by Lemma 7.11,

$$\frac{D(f)}{\mu(f^2)} = \frac{\bar{D}(\bar{f})}{\bar{\mu}(\bar{f}^2)} \geq \lambda_0^{\theta-, \gamma} \wedge \lambda_0^{\theta+, \gamma}.$$

Making the supremum with respect to γ and θ , it follows that

$$\frac{D(f)}{\mu(f^2)} \geq \sup_{-M-1 < \theta < N+1} \sup_{\gamma > 1} (\lambda_0^{\theta-, \gamma} \wedge \lambda_0^{\theta+, \gamma}).$$

At the boundaries, say $\theta = -M - 1$ for instance, by (7.10) and the convention, we have

$$D(f) \geq \sum_{-M-1 < i < N+1} \mu_i b_i (f_{i+1} - f_i)^2 \geq \lambda_0^{\theta+, \gamma} \mu(f^2) = [\lambda_0^{\theta-, \gamma} \wedge \lambda_0^{\theta+, \gamma}] \mu(f^2).$$

Therefore, we indeed have

$$\frac{D(f)}{\mu(f^2)} \geq \sup_{-M-1 \leq \theta \leq N+1} \sup_{\gamma > 1} (\lambda_0^{\theta-, \gamma} \wedge \lambda_0^{\theta+, \gamma}).$$

Making infimum with respect to f , we obtain

$$\lambda_0 = \inf_{f \in \mathcal{X}, f \neq 0} \frac{D(f)}{\mu(f^2)} \geq \sup_{-M-1 \leq \theta \leq N+1} \sup_{\gamma > 1} (\lambda_0^{\theta-, \gamma} \wedge \lambda_0^{\theta+, \gamma}).$$

This proves the (second) inequality in (7.13).

To prove the upper estimate, fix $\theta \in E$ and $\gamma > 1$ again. As we have seen from the last part of proof (g) of Theorem 2.4 and Proposition 2.5, if we let $\lambda_0^{\theta+, \gamma, n}$ denote the local eigenvalue with Neumann boundary at θ and Dirichlet boundary at $n + 1$, then $\lambda_0^{\theta+, \gamma, n} \downarrow \lambda_0^{\theta+, \gamma}$ as $n \uparrow \infty$. Thus, for each $\varepsilon > 0$, we have $\lambda_0^{\theta+, \gamma, n} < \lambda_0^{\theta+, \gamma} + \varepsilon$ for large enough n . By Proposition 2.2, we can assume that the corresponding eigenfunction $g^{(+, n)}$ of $\lambda_0^{\theta+, \gamma, n}$ satisfies $g_\theta^{(+, n)} = 1$ and $g_i^{(+, n)} = 0$ for all $i > n (> \theta)$. Similarly, we have $\lambda_0^{\theta-, \gamma, m} < \lambda_0^{\theta-, \gamma} + \varepsilon$ for small enough $-m$, and moreover, the eigenfunction $g^{(-, m)}$ of $\lambda_0^{\theta-, \gamma, m}$ satisfies $g_\theta^{(-, m)} = 1$ and $g_i^{(-, m)} = 0$ for all $i < -m (< \theta)$. Let $\bar{f}$ be defined as above, connecting $g^{(-, m)}$ and $g^{(+, n)}$. Then $\bar{f}$ has a finite support, $\bar{f}_\theta = \bar{f}_{\theta+1} = 1$, and moreover by (7.10),

$$\begin{aligned} \overline{D}(\bar{f}) &= \sum_{\bar{E} \ni i \leq \theta} \bar{\mu}_i \bar{a}_i (\bar{f}_i - \bar{f}_{i-1})^2 + \sum_{\bar{E} \ni i \geq \theta+1} \bar{\mu}_i \bar{b}_i (\bar{f}_{i+1} - \bar{f}_i)^2 \\ &= \lambda_0^{\theta-, \gamma, m} \sum_{i \leq \theta} \bar{\mu}_i \bar{f}_i^2 + \lambda_0^{\theta+, \gamma, n} \sum_{i \geq \theta+1} \bar{\mu}_i \bar{f}_i^2 \\ &\leq (\lambda_0^{\theta-, \gamma} \vee \lambda_0^{\theta+, \gamma} + \varepsilon) \bar{\mu}(\bar{f}^2). \end{aligned}$$

By Lemmas 7.11 and 7.12, this gives us

$$\lambda_0 = \bar{\lambda}_0 \leq \lambda_0^{\theta-, \gamma} \vee \lambda_0^{\theta+, \gamma}$$

since ε is arbitrary. Furthermore, we have

$$\lambda_0 \leq \inf_{\theta \in E} \inf_{\gamma > 1} (\lambda_0^{\theta-, \gamma} \vee \lambda_0^{\theta+, \gamma})$$

as required. $\square$

The proof of the equalities in Theorem 7.10 is much harder. For which, we need once again a deeper understanding of the eigenfunction of λ_0 . To have a concrete impression, we mention that the eigenfunction in Examples 7.6 (2) is $g_0 = g_3 = 0$, $g_1 = (1 + \varepsilon)g_2$. Thus, when $\varepsilon = 0$, we have $g_1 = g_2$. Besides, it is rather easy to see the shape of eigenfunction g of the examples given in Table 7.1 since $v_i < 1$ iff $g_{i+1} < g_i$ for all i .

Definition 7.13.

- (1) A function f is said to be *unimodal* if there exists a finite k such that f_i is strictly increasing for $i \leq k$ and strictly decreasing for $i \geq k$.
- (2) A function f is said to be a *simple echelon* if there exists a k such that $f_k = f_{k+1}$, f_i is strictly increasing for $i \leq k$ and strictly decreasing for $i \geq k+1$.

Proposition 7.14. Let g be a positive eigenfunction of $\lambda > 0$ for a birth–death process. Then g is strictly monotone, or unimodal, or a simple echelon.

Proof. (a) Let $g_k \geq g_{k+1}$ for some k . We prove that g is strictly decreasing for $i \geq k+1$. To do so, note that

$$b_{k+1}(g_{k+2} - g_{k+1}) = -\lambda g_{k+1} - a_{k+1}(g_k - g_{k+1}) \leq -\lambda g_{k+1} < 0.$$

Thus, we have $g_{k+2} < g_{k+1}$. Assume that $g_n < g_{n-1}$ for some $n \geq k+2$. Then the eigenequation shows that

$$b_n(g_{n+1} - g_n) = -\lambda g_n - a_n(g_{n-1} - g_n) < -\lambda g_n < 0.$$

By induction, this gives us $g_{n+1} < g_n$ for all $n \geq k+1$.

(b) By symmetry, we can handle with the case that $g_k \leq g_{k+1}$ for some k . One starts at

$$a_k(g_{k-1} - g_k) = -\lambda g_k - b_k(g_{k+1} - g_k) \leq -\lambda g_k < 0.$$

We obtain $g_{k-1} < g_k$ and then $g_{n-1} < g_n$ for all $n \leq k$ by induction.

(c) By (a) and (b), it follows that there is no local convex part of g . Otherwise, there is a k such that either $g_{k-1} > g_k < g_{k+1}$ or $g_{k-1} > g_k = g_{k+1} < g_{k+2}$ which contradict what we proved in (a) and (b).

(d) We claim that for every k , say $k = 0$ for simplicity, the two cases “ $g_{-1} \geq g_0$ ” and “ $g_0 \leq g_1$ ” cannot happen at the same time. Otherwise, there are four situations:

$$g_{-1} = g_0 = g_1, \quad g_{-1} > g_0 < g_1, \quad g_{-1} > g_0 = g_1, \quad \text{and} \quad g_{-1} = g_0 < g_1.$$

The first one cannot happen, otherwise we have $g_i \equiv 0$. By (c), the second case is impossible. The last two cases are also impossible by (b) and (a), respectively.

(e) Having these preparations at hand, we are ready to prove the main assertion of the proposition. Clearly, we need only to consider the case that g is not strictly monotone. Choose a starting point, say 0 for instance. By (d), we have only one possibility: either $g_{-1} \geq g_0$ or $g_0 \leq g_1$. Without loss of generality, assume that $g_0 \leq g_1$. If $g_0 = g_1$, then by (a) and (b), g is a simple echelon. If $g_0 < g_1$, then on the one hand, by (b), g_i is strictly increasing for all $i \leq 1$, and on the other hand, we can find a $k \geq 1$ such that $g_1 < g_2 < \dots < g_k \geq g_{k+1}$ since g is not strictly monotone by assumption. Applying (a) again, it follows that g is strictly decreasing for all $i \geq k+1$. Hence, g is either unimodal or a simple echelon. $\square$

Proposition 7.15. For the birth-death process on $\mathbb{Z}$, the following assertions hold.

(1) The eigenfunction g of λ satisfies the following successive formulas:

$$\begin{aligned} g_{k+1} &= g_k + \frac{1}{\mu_k b_k} \left[\mu_\theta a_\theta (g_\theta - g_{\theta-1}) - \lambda \sum_{i=\theta}^k \mu_i g_i \right], & k \geq \theta, \\ g_{k-1} &= g_k + \frac{1}{\mu_k a_k} \left[\mu_\theta a_\theta (g_{\theta-1} - g_\theta) - \lambda \sum_{i=k}^{\theta-1} \mu_i g_i \right], & k < \theta. \end{aligned} \quad (7.19)$$

(2) If $\lambda = 0$, then the non-trivial eigenfunction g with $g_\theta = 1$ for some $\theta \in \mathbb{Z}$ is given by

$$\begin{aligned} g_n &= 1 + (1 - g_{\theta-1}) \sum_{j=\theta}^{n-1} \prod_{k=\theta}^j \frac{a_k}{b_k}, & n \geq \theta, \\ g_n &= 1 - (1 - g_{\theta-1}) \sum_{j=n}^{\theta-1} \prod_{k=j+1}^{\theta-1} \frac{b_k}{a_k}, & n < \theta. \end{aligned}$$

In this case, the function g is either the constant function $\mathbb{1}$ or strictly monotone on $\mathbb{Z}$.

(3) If $\lambda > 0$ and

$$\sum_{i=-\infty}^{\theta} \mu_i = \sum_{i=\theta}^{\infty} \mu_i = \infty, \quad (7.20)$$

then the non-trivial eigenfunction g of λ cannot be monotone.

Proof. (a) Part (1) of the proposition follows from the eigenequation.

(b) When $\lambda = 0$, with $u_i := g_{i+1} - g_i$ ($i \in \mathbb{Z}$), the eigenequation $b_i u_i = a_i u_{i-1}$ gives us

$$u_j = (1 - g_{\theta-1}) \prod_{k=\theta}^j \frac{a_k}{b_k}, \quad j \geq \theta, \quad u_j = (1 - g_{\theta-1}) \prod_{k=j+1}^{\theta-1} \frac{b_k}{a_k}, \quad j < \theta.$$

It follows that either $g_i \equiv 1$ or g is strictly monotone on $\mathbb{Z}$. Now, part (2) of the proposition follows by making a summation of j from θ to $n-1$ and from n to $\theta-1$, respectively.

(c) Without loss of generality, assume that $g_\theta = 1$ for some $\theta \in \mathbb{Z}$. Suppose that g is non-decreasing, then by the first equation in part (1), we would have

$$\infty > \frac{\mu_\theta a_\theta (g_\theta - g_{\theta-1})}{\lambda} \geq \sum_{k=\theta}^n \mu_k g_k \geq \sum_{k=\theta}^n \mu_k \rightarrow \infty \quad \text{as } n \rightarrow \infty.$$

Otherwise, if g is non-increasing, then by the second equation in part (1), we would have

$$\infty > \frac{\mu_\theta a_\theta (g_{\theta-1} - g_\theta)}{\lambda} \geq \sum_{k=n}^{\theta-1} \mu_k g_k \geq \sum_{k=n}^{\theta-1} \mu_k \rightarrow \infty \quad \text{as } n \rightarrow -\infty.$$

We have thus proved part (3) of the proposition. $\square$

We remark that Proposition 7.15 (2) is different from Proposition 2.2 where the eigenfunction of $\lambda = 0$ must be a constant. Here is a simple example with $\theta = 0$: $a_i = b_i = |i|$ if $i \neq 0$ and $a_0 = b_0 = 1$, then corresponding to $\lambda = 0$, we have a family of linear eigenfunctions $\{g_i^{(\gamma)} = 1 + (1 - \gamma)i : i \in \mathbb{Z}\}_{\gamma \in \mathbb{R}}$ (normalized at 0) with one-parameter γ .

Proposition 7.16. Let (7.14) hold and g be a non-zero eigenfunction of $\lambda_0 > 0$. Then g is either positive or negative on E .

Proof. If one of M or N is finite, then the conclusion follows from Proposition 2.2 (1). From now on in the proof, assume that $M = N = \infty$.

(a) If the conclusion of the proposition does not hold, then there is a k (say) such that $g_k \leq 0$ and either $g_{k-1} > 0$ or $g_{k+1} > 0$. By symmetry, assume that $g_{k+1} > 0$.

(b) We now prove that $g_i > 0$ for all $i \geq k + 1$. Given $m, n \in \mathbb{Z}$ with $m \leq n$, denote by $\lambda_0^{[m,n]}$ the first eigenvalue of the process restricted on the state space $\{i : m \leq i \leq n\}$ with Dirichlet boundaries at $m - 1$ and $n + 1$ in the sense similar to (7.1). If the assertion does not hold, then there is a $k_0 : k_0 > k + 1$ such that $g_{k_0} \leq 0$. Now, let $\tilde{g}$ satisfy $\tilde{g}_k = 0$, $\tilde{g}_i = g_i$ for $i = k + 1, \dots, k_0 - 1$, $\tilde{g}_{k_0} = \varepsilon$ for some $\varepsilon > 0$, $\tilde{g}_i = 0$ for $i \geq k_0 + 1$. Note that

$$\begin{aligned} (-\Omega \tilde{g})(k+1) &= b_{k+1}(\tilde{g}_{k+1} - \tilde{g}_{k+2}) + a_{k+1}(\tilde{g}_{k+1} - \tilde{g}_k) \\ &= b_{k+1}(g_{k+1} - g_{k+2}) + a_{k+1}(g_{k+1} - g_k) + a_{k+1}g_k \\ &= \lambda_0 g_{k+1} + a_{k+1}g_k \\ &\leq \lambda_0 \tilde{g}_{k+1}, \end{aligned}$$

Because of $\lambda_0 > 0$ and following proof (b) of Proposition 2.1, we can choose a suitable $\varepsilon > 0$ such that

$$\sum_{i=k+1}^{k_0} \mu_i \tilde{g}_i (-\Omega \tilde{g})(i) < \lambda_0 \sum_{i=k+1}^{k_0} \mu_i \tilde{g}_i^2.$$

It follows that $\lambda_0^{[k+1, k_0]} < \lambda_0$. However, it is obvious that $\lambda_0 \leq \lambda_0^{[k+1, k_0]}$ and so we get a contradiction. We have thus proved that $g_i > 0$ for all $i \geq k + 1$.

(c) By (7.14) and proof (c) of Proposition 7.15, g cannot be non-decreasing since $\lambda_0 > 0$. Hence, there is a $\theta \geq k + 2$ such that $g_{k+1} < g_{k+2} < \dots < g_\theta \geq g_{\theta+1}$. In the case that $g_\theta > g_{\theta+1}$, by introducing an additional point but keeping the same λ_0 as shown in Lemma 7.12, one can reduce to the case that $g_\theta = g_{\theta+1}$. Hence, one can split the original process into two as in (L) and (R). Now, starting from θ at which $g_\theta > 0$, look at the process on the left-hand side in the inverse way, one finds the point $k < \theta$ at which $g_k \leq 0$. Applying proof (b) above to this process, one may get a contradiction. It follows that $g > 0$ on $(-\infty, \theta] \supset (-\infty, k]$.

Therefore, we should have $g > 0$ on $\mathbb{Z}$. $\square$

Proof of Theorem 7.10. Part II. We now prove the equalities in (7.13). By assumption $\lambda_0^{(\pm)} > 0$ and the second inequality in (7.13), it follows that $\lambda_0 > 0$. If

one of M and N is finite, then the non-trivial eigenfunction g must be positive by Proposition 2.2 (1). In this case, it is helpful to include the boundary into the domain of g for understanding its shape. Then by Proposition 7.14, there are only two possibilities:

- (i) g is unimodal;
- (ii) g is a simple echelon.

Next, if $M = N = \infty$, then by Proposition 7.16, we have $g > 0$. Moreover, by Proposition 7.15, g cannot be monotone. Hence, by Proposition 7.14, g has again one of shapes (i) and (ii) as above.

We now prove the equalities in (7.13) only in the case that $M = N = \infty$. The proof for the other case is simpler.

(a) Case (ii). We use the operator Π defined in Section 2:

$$\overline{\Pi}_i^{\theta+, \gamma}(\bar{f}) = \frac{1}{\bar{f}_i} \sum_{j=i}^{N+1} \frac{1}{\bar{\mu}_j \bar{b}_j} \sum_{k=\theta+1}^j \bar{\mu}_k \bar{f}_k, \quad \theta+1 \leq i < N+2.$$

For each $\bar{f}$ satisfying: $\bar{f}_i = f_i$ for $i \leq \theta$ and $\bar{f}_i = f_{i-1}$ for $i \geq \theta+1$ for some f on E , by (7.15) and (7.16), we have

$$\begin{aligned} \overline{\Pi}_i^{\theta+, \gamma}(\bar{f}) &= \frac{1}{f_{i-1}} \sum_{j=i}^{N+1} \frac{1}{\mu_{j-1} b_{j-1}} \left[\bar{\mu}_{\theta+1} f_{\theta} + \sum_{k=\theta+2}^j \mu_{k-1} f_{k-1} \right] \\ &= \frac{1}{f_{i-1}} \sum_{j=i}^{N+1} \frac{1}{\mu_{j-1} b_{j-1}} \left[\left(1 - \frac{1}{\gamma}\right) \mu_{\theta} f_{\theta} + \sum_{k=\theta+1}^{j-1} \mu_k f_k \right] \\ &= \frac{1}{f_{i-1}} \sum_{j=i-1}^N \frac{1}{\mu_j b_j} \sum_{k=\theta}^j \mu_k f_k - \frac{\mu_{\theta} f_{\theta}}{\gamma f_{i-1}} \sum_{j=i-1}^N \frac{1}{\mu_j b_j} \\ &= \frac{1}{f_{i-1}} \sum_{j=i-1}^N \frac{1}{\mu_j b_j} \sum_{k=\theta+1}^j \mu_k f_k + \left(1 - \frac{1}{\gamma}\right) \frac{\mu_{\theta} f_{\theta}}{f_{i-1}} \sum_{j=i-1}^N \frac{1}{\mu_j b_j} \\ &\quad \theta+1 \leq i < N+2. \end{aligned} \tag{7.21}$$

Similarly, we have

$$\begin{aligned} \overline{\Pi}_i^{\theta-, \gamma}(\bar{f}) &= \frac{1}{\bar{f}_i} \sum_{j=-M}^i \frac{1}{\bar{\mu}_j \bar{a}_j} \sum_{k=j}^{\theta} \bar{\mu}_k \bar{f}_k \\ &= \frac{1}{f_i} \sum_{j=-M}^i \frac{1}{\mu_j a_j} \sum_{k=j}^{\theta} \mu_k f_k - \left(1 - \frac{1}{\gamma}\right) \frac{\mu_{\theta} f_{\theta}}{f_i} \sum_{j=-M}^i \frac{1}{\mu_j a_j}, \\ &\quad -M-1 < i \leq \theta. \end{aligned} \tag{7.22}$$

Because $g_{\theta} = g_{\theta+1}$, we can regard θ as a Neumann boundary of the original process restricted on the left-hand side and at the same time, regard θ as a Neumann boundary of the original process restricted on the right-hand side. Because $\lambda_0^{(\pm)} >$

0, by Proposition 2.5 (2), we have $g_{\pm\infty} = 0$. Hence, by (2.11), (7.21), and (7.22), we obtain

$$\begin{aligned}\overline{H}_i^{\theta+, \gamma}(\bar{g}) &= \frac{1}{\lambda_0} + \left(1 - \frac{1}{\gamma}\right) \frac{\mu_\theta g_\theta}{g_{i-1}} \sum_{j=i-1}^N \frac{1}{\mu_j b_j}, \quad \theta + 1 \leq i < N + 2, \\ \overline{H}_i^{\theta-, \gamma}(\bar{g}) &= \frac{1}{\lambda_0} - \left(1 - \frac{1}{\gamma}\right) \frac{\mu_\theta g_\theta}{g_i} \sum_{j=-M}^i \frac{1}{\mu_j a_j}, \quad -M - 1 < i \leq \theta.\end{aligned}$$

By Proposition 2.2 (2), we have

$$\sup_{\theta \leq i < N+1} \frac{\mu_\theta g_\theta}{g_i} \sum_{j=i}^N \frac{1}{\mu_j b_j} \leq \frac{1}{\lambda_0}, \quad \sup_{-M-1 < i \leq \theta} \frac{\mu_\theta g_\theta}{g_i} \sum_{j=-M}^i \frac{1}{\mu_j a_j} \leq \frac{1}{\lambda_0}.$$

Therefore, by the second inequality in (7.13) and Theorem 2.4 (3), it follows that

$$\begin{aligned}\lambda_0 &\geq \sup_{\theta' \in E} \sup_{\gamma > 1} \left[\lambda_0^{\theta'-, \gamma} \wedge \lambda_0^{\theta'+, \gamma} \right] \\ &\geq \sup_{\gamma > 1} \left[\lambda_0^{\theta-, \gamma} \wedge \lambda_0^{\theta+, \gamma} \right] \\ &\geq \sup_{\gamma > 1} \left[\left(\inf_{-M-1 < i \leq \theta} \overline{H}_i^{\theta-, \gamma}(\bar{g})^{-1} \right) \wedge \left(\inf_{\theta+1 \leq i < N+2} \overline{H}_i^{\theta+, \gamma}(\bar{g})^{-1} \right) \right] \\ &= \sup_{\gamma > 1} \inf_{\theta+1 \leq i < N+2} \overline{H}_i^{\theta+, \gamma}(\bar{g})^{-1} \\ &= \sup_{\gamma > 1} \left\{ \frac{1}{\lambda_0} + \left(1 - \frac{1}{\gamma}\right) \sup_{\theta \leq i < N+1} \frac{\mu_\theta g_\theta}{g_i} \sum_{j=i}^N \frac{1}{\mu_j b_j} \right\}^{-1} \\ &= \lambda_0.\end{aligned}$$

We have thus proved in Case (ii) the second equality in (7.13).

To prove the first equality in (7.13), noting the inequality was proved in Part I, we have dually

$$\begin{aligned}\lambda_0 &\leq \inf_{\theta' \in E} \inf_{\gamma > 1} \left[\lambda_0^{\theta'-, \gamma} \vee \lambda_0^{\theta'+, \gamma} \right] \\ &\leq \inf_{\gamma > 1} \left[\lambda_0^{\theta-, \gamma} \vee \lambda_0^{\theta+, \gamma} \right] \\ &\leq \inf_{\gamma > 1} \left[\left(\sup_{-M-1 < i \leq \theta} \overline{H}_i^{\theta-, \gamma}(\bar{g})^{-1} \right) \vee \left(\sup_{\theta+1 \leq i < N+2} \overline{H}_i^{\theta+, \gamma}(\bar{g})^{-1} \right) \right] \\ &= \inf_{\gamma > 1} \sup_{-M-1 < i \leq \theta} \overline{H}_i^{\theta-, \gamma}(\bar{g})^{-1} \\ &= \sup_{\gamma > 1} \left\{ \frac{1}{\lambda_0} - \left(1 - \frac{1}{\gamma}\right) \sup_{-M-1 < i \leq \theta} \frac{\mu_\theta g_\theta}{g_i} \sum_{j=-M}^i \frac{1}{\mu_j a_j} \right\}^{-1} \\ &= \lambda_0.\end{aligned}$$

However, there is a problem in the second line of the proof. To apply Theorem 2.4 (3), one requires that either $g \in L^2(\mu)$ or g is local. Hence, an additional work is required. Anyhow, the conclusion holds whenever both M and N are finite. We will come back to the proof in proof (c) below.

(b) Case (i). By Lemma 7.12, this case can be reduced to Case (ii). Actually, the proof becomes easier now. With γ given by (7.18), we have

$$\begin{aligned}\overline{H}_i^{\theta+, \gamma}(\bar{g}) &\equiv \frac{1}{\lambda_0}, & \theta + 1 \leq i < N + 2, \\ \overline{H}_i^{\theta-, \gamma}(\bar{g}) &\equiv \frac{1}{\lambda_0}, & -M - 1 < i \leq \theta.\end{aligned}$$

Hence the second equality in (7.13) holds. Moreover, the first equality in (7.13) also holds whenever both M and N are finite.

(c) To complete the proof for the first equality in (7.13), we need to overcome the unbounded problem. For this, choose $M_p, N_p \uparrow \infty$ as $p \rightarrow \infty$. Denote by $\lambda_0^{\theta-, \gamma, p}$, $\lambda_0^{\theta+, \gamma, p}$ and $\lambda_0^{(p)}$, respectively, the quantities $\lambda_0^{\theta-, \gamma}$, $\lambda_0^{\theta+, \gamma}$, and λ_0 when M and N are replaced by M_p and N_p . Note that for a finite state space, we certainly have $\lambda_0^{(p)} > 0$, its eigenfunction is positive (by Proposition 2.2 (1)) and has properties (i) and (ii) mentioned in the above proof (by Proposition 7.14). Clearly, for each fixed θ and γ , we have

$$\lambda_0^{\theta\pm, \gamma, p} \downarrow \lambda_0^{\theta\pm, \gamma}, \quad \lambda_0^{(p)} \downarrow \lambda_0 \quad \text{as } p \rightarrow \infty.$$

Thus, as proved in (a) and (b), whether we are in Case (i) or (ii), we have for each p ,

$$\begin{aligned}\lambda_0^{(p)} &= \inf_{\theta \in [-M_p, N_p]} \inf_{\gamma > 1} [\lambda_0^{\theta-, \gamma, p} \vee \lambda_0^{\theta+, \gamma, p}] \\ &\geq \inf_{\theta \in [-M_p, N_p]} \inf_{\gamma > 1} [\lambda_0^{\theta-, \gamma} \vee \lambda_0^{\theta+, \gamma}] \\ &\geq \inf_{\theta \in E} \inf_{\gamma > 1} [\lambda_0^{\theta-, \gamma} \vee \lambda_0^{\theta+, \gamma}].\end{aligned}$$

Therefore, by the first inequality in (7.13) proved in Part I, it follows that

$$\lambda_0 \leq \inf_{\theta \in E} \inf_{\gamma > 1} [\lambda_0^{\theta-, \gamma} \vee \lambda_0^{\theta+, \gamma}] \leq \lambda_0^{(p)} \downarrow \lambda_0 \quad \text{as } p \rightarrow \infty.$$

We have thus completed the proof of the theorem. $\square$

Here are remarks about the assumption made in part (2) of Theorem 7.10. Similar to the upper estimate, we do have

$$\lambda_0^{(p)} = \sup_{\theta \in [-M_p, N_p]} \sup_{\gamma > 1} [\lambda_0^{\theta-, \gamma, p} \wedge \lambda_0^{\theta+, \gamma, p}].$$

The problem is that $\lambda_0^{\theta\pm, \gamma, p} \downarrow \lambda_0^{\theta\pm, \gamma}$ as $p \rightarrow \infty$ goes to the opposite direction and the approximating sequences $\{M_p\}$ and $\{N_p\}$ depend on θ and γ . Hence, the proof

for the upper estimate does not work for the lower one. Next, to prove the second equality in (7.13), it seems more natural to assume that $\lambda_0^{\theta-, \gamma} \wedge \lambda_0^{\theta+, \gamma} > 0$ for some θ and γ , that is, $\lambda_0^{(-)} \wedge \lambda_0^{(+)} > 0$, rather than $\lambda_0^{(-)} \vee \lambda_0^{(+)} > 0$ as we made. However, if one of them is zero, say $\lambda_0^{(-)} = 0$, then as mentioned before Theorem 7.10, we have a single Dirichlet boundary but not the bilateral Dirichlet ones, and the variational formula takes a different form (i.e., the second inequality in (7.13) at the boundaries). Condition (7.14) is due to the same reason. In particular, when $M = -1$, for instance, if $\sum_i \mu_i < \infty$ and $\sum_i (\mu_i a_i)^{-1} = \infty$, then $\lambda_0^{(+)} = 0$ by Theorem 3.1, and we go back to the case studied in Section 4. In which case, the eigenfunction of λ_0 is strictly increasing.

To conclude this section, we introduce a complement result to [12; Proposition 5.13] about the principal eigenvalue for general Markov chains.

Proposition 7.17. Let $(q_{ij} : i, j \in E)$ be symmetric with respect to (μ_i) on a countable set E , not necessarily conservative (or having killings):

$$d_i := q_i - \sum_{j \neq i} q_{ij} \geq 0, \quad q_i := -q_{ii} \in [0, \infty).$$

Define

$$D(f) = \frac{1}{2} \sum_{i, j \in E} \mu_i q_{ij} (f_j - f_i)^2 + \sum_{i \in E} \mu_i d_i f_i^2$$

and

$$\lambda_0 = \inf \{D(f) : f \text{ has a finite support and } \mu(f^2) = 1\}.$$

Then we have $\inf_{i \in E} q_i \geq \lambda_0$.

Proof. Without loss of generality, assume that $E = \mathbb{Z}_+ = \{0, 1, \dots\}$. Fix $k \in E$ and take $f = \mathbb{1}_{\{k\}}$. Then $\mu(f^2) = \mu_k$ and

$$\begin{aligned} D(f) &= \sum_{i, j: i < j} \mu_i q_{ij} (f_j - f_i)^2 + \sum_{i \in E} \mu_i d_i f_i^2 \\ &= \sum_{j > k} \mu_k q_{kj} (f_k - f_j)^2 + \sum_{i < k} \mu_i q_{ik} (f_i - f_k)^2 + \sum_{i \in E} \mu_i d_i f_i^2 \\ &= \sum_{j > k} \mu_k q_{kj} + \sum_{i < k} \mu_i q_{ik} + \mu_k d_k. \end{aligned}$$

By the symmetry of $\mu_i q_{ij}$, we get

$$D(f) = \sum_{j > k} \mu_k q_{kj} + \sum_{i < k} \mu_k q_{ki} + \mu_k d_k = \mu_k \left(\sum_{j \neq k} q_{kj} + d_k \right) = \mu_k q_k.$$

It follows that

$$\lambda_0 \leq D(f) / \mu(f^2) = q_k.$$

The assertion now follows since $k \in E$ is arbitrary. $\square$

8. CRITERIA FOR POINCARÉ-TYPE INEQUALITIES

As in [9] for the ergodic case having $N < \infty$ or (1.2), the results studied in Sections 2, 3 and 7 can be extended to a more general setup, so called Poincaré-type inequalities. In this way, one obtains various types of stability, not only the L^2 -exponential one studied in the other sections of the paper. Here we consider only the criteria and the basic estimates for the inequalities. In other words, we extend Theorems 3.1, 4.2, and 6.2 to the general setup with some improvement. At the same time, we introduce a criterion for the processes studied in Section 7 in this setup. To do so, we need a class of normed linear spaces $(\mathbb{B}, \|\cdot\|_{\mathbb{B}}, \mu)$ consisting of real Borel measurable functions on a measurable space $(X, \mathcal{X}, \mu)$. We now modify the hypotheses on the normed linear spaces given in [12; Chapter 7] as follows.

- (H1) In the case that $\mu(X) = \infty$, $\mathbb{1}_K \in \mathbb{B}$ for all compact K . Otherwise, $1 \in \mathbb{B}$.
- (H2) If $h \in \mathbb{B}$ and $|f| \leq h$, then $f \in \mathbb{B}$.
- (H3) $\|f\|_{\mathbb{B}} = \sup_{g \in \mathcal{G}} \int_X |f|g d\mu$,

where $\mathcal{G}$, to be specified case by case, is a class of nonnegative $\mathcal{X}$ -measurable functions. A typical example is $\mathcal{G} = \{1\}$ and then $\mathbb{B} = L^1(\mu)$. Throughout this section, we assume (H1)–(H3) for $(\mathbb{B}, \|\cdot\|_{\mathbb{B}}, \mu)$ without mentioning again.

Before moving further, let us mention the following result.

Remark 8.1. Without using (1.2), the results in [9] remain true under the condition $\sum_i \mu_i < \infty$ replacing the original process with the maximal one if necessary.

The key reason is that without condition (1.2), the same conclusion holds in Section 4 on which the cited paper is based on.

In this section, our state space is $E = \{i : -M - 1 < i < N + 1\}$ ($M, N \leq \infty$) as in the second part of Section 7. The next result is the main one in this section; it has several corollaries as we have seen in the last section. Note that the factor 4 in (8.2) below is universal, independent of $\mathbb{B}$.

Theorem 8.2. Consider the minimal birth–death process with Dirichlet boundaries at $-M - 1$ if $M < \infty$ and at $N + 1$ if $N < \infty$. Assume that $\mathcal{G}$ contains a locally positive element. Then the optimal constant $A_{\mathbb{B}}$ in the *Poincaré-type inequality*

$$\|f^2\|_{\mathbb{B}} \leq A_{\mathbb{B}} D(f), \quad f \in \mathcal{D}^{\min}(D), \quad (8.1)$$

satisfies

$$B_{\mathbb{B}} \leq A_{\mathbb{B}} \leq 4B_{\mathbb{B}}, \quad (8.2)$$

where the isoperimetric constant $B_{\mathbb{B}}$ can be expressed as follows:

$$B_{\mathbb{B}}^{-1} = \inf_{m, n \in E: m \leq n} \left[\left(\sum_{i=-M}^m \frac{1}{\mu_i a_i} \right)^{-1} + \left(\sum_{i=n}^N \frac{1}{\mu_i b_i} \right)^{-1} \right] \|\mathbb{1}_{[m, n]}\|_{\mathbb{B}}^{-1}. \quad (8.3)$$

In particular, when $\mathbb{B} = L^{p/2}(\mu)$ ($p \geq 2$) (then (8.1) is called the *Sobolev-type inequality*), we have

$$B_p^{-1} = \inf_{m, n \in E: m \leq n} \left[\left(\sum_{i=-M}^m \frac{1}{\mu_i a_i} \right)^{-1} + \left(\sum_{i=n}^N \frac{1}{\mu_i b_i} \right)^{-1} \right] \left(\sum_{j=m}^n \mu_j \right)^{-2/p}. \quad (8.4)$$

Proof. (a) First consider the transient case, in particular when one of M or N is finite. We use the proof of [11; Corollary 4.1] or [12; Corollary 7.5] with a slight modification. In proof (a) there, it was shown that one can replace “ $f|_K \geq 1$ ” by “ $f|_K = 1$ ” in computing the capacity $\text{Cap}(K)$ for compact K . Without loss of generality, assume that $f \geq 0$. Otherwise, replace f with $|f|$. In the proof just mentioned, the condition “ $\sum_i \mu_i < \infty$ ” was used so that $1 \in \mathcal{D}(D)$. We cannot use this assumption now, but for a given nonnegative $f \in \mathcal{D}^{\min}(D) \cap \mathcal{C}_c(E)$, where $\mathcal{C}_c(E)$ is the set of continuous functions with compact support, we can simply choose a nonnegative smooth $h \in \mathcal{C}_c(E)$ such that $h|_{\text{supp}(f)} = 1$. Then $h \in \mathcal{D}^{\min}(D)$, $f \wedge h \in \mathcal{D}^{\min}(D)$, and so one can use $f \wedge h \in \mathcal{D}^{\min}(D)$ instead of $f \wedge 1$ to arrive at the same conclusion $D(f) \geq D(f \wedge h)$ as in the original proof (a).

The first step in the original proof (b) shows that one can replace a finite number of disjointed finite intervals $\{K_i\}$ by the connected one $[\min \cup_i K_i, \max \cup_i K_i]$. This part of the proof needs no change.

Note that in the original proof, the state space is $\{1, 2, \dots\}$ with Dirichlet boundary at 0. The main body of the original proof (b) is to find a minimizer (actually unique) $f \in \mathcal{C}_c(E)$ for $D(f)$ having the properties $f_0 = 0$ and $f|_K = 1$. Replacing N with q for the consistence with the notation used here and let $K = \{m, m+1, \dots, n\}$ ($1 \leq m \leq n$, here m and n are exchanged from the original proof). Now, within the class of f : $f_0 = 0$, $f|_K = 1$ and $\text{supp}(f) = \{1, \dots, q\}$ ($n \leq q < N+1$), the minimal solution is

$$D(f) = \left(\sum_{i=1}^m \frac{1}{\mu_i a_i} \right)^{-1} + \left(\sum_{i=n}^q \frac{1}{\mu_i b_i} \right)^{-1}. \quad (8.5)$$

To handle with the general state space, one needs to move the original left-end point 1 of the state space to somewhere, say $p > -M-1$. In detail, replace the condition $m \geq 1$ used in defining the compact set K by $m > -M-1$. At the same time, replace $\{1, \dots, q\}$ by $\{p, p+1, \dots, q\}$ with $-M-1 < p \leq m$ for the $\text{supp}(f)$. Then the last formula reads as follows:

$$D(f) = \left(\sum_{i=p}^m \frac{1}{\mu_i a_i} \right)^{-1} + \left(\sum_{i=n}^q \frac{1}{\mu_i b_i} \right)^{-1}.$$

In the original proof, the ergodic condition and (1.2) are mainly used here to remove the second term on the right-hand side. We now keep it. Since the right-hand side is increasing in p and decreasing in q , by making the infimum with respect to f , it follows that

$$\begin{aligned} \text{Cap}(K) &:= \inf \{ D(f) : f \in \mathcal{D}^{\min} \cap \mathcal{C}_c(E) \text{ and } f|_K \geq 1 \} \\ &= \left(\sum_{i=-M}^m \frac{1}{\mu_i a_i} \right)^{-1} + \left(\sum_{i=n}^N \frac{1}{\mu_i b_i} \right)^{-1}, \quad K = \{m, m+1, \dots, n\} =: [m, n]. \end{aligned}$$

The assertion of the theorem now follows by using

$$B_{\mathbb{B}} := \sup_K \frac{\|1_K\|_{\mathbb{B}}}{\text{Cap}(K)} = \sup_{-M-1 < m \leq n < N+1} \frac{\|1_{[m,n]}\|_{\mathbb{B}}}{\text{Cap}([m,n])}$$

and applying [11; Theorem 1.1] or [12; Theorem 7.2]. The last result is an extension of Fukushima and Uemura (2003, Theorem 3.1).

(b) Next, consider the recurrent case: both $\sum_{i<\theta}(\mu_i a_i)^{-1}$ and $\sum_{i>\theta}(\mu_i b_i)^{-1}$ are diverged. Here is actually a direct proof of the lower estimate in (8.2). Without loss of generality, assume that the reference point $\theta = 0$. Fix $m' \geq m \geq 0$ and $n' \geq n \geq 0$. Based on the knowledge about the eigenfunction given in the last section, and similar to proof (b) of Theorem 3.1, define

$$f_i = \begin{cases} \sum_{k=i \vee n}^{n'} \frac{1}{\mu_k b_k} \mathbb{1}_{\{i \leq n'\}}, & i \geq 0, \\ \gamma \sum_{k=-m'}^{i \wedge (-m)} \frac{1}{\mu_k a_k} \mathbb{1}_{\{i \geq -m'\}}, & i \leq 0, \end{cases}$$

where

$$\gamma := \gamma(m', m, n, n') = \sum_{k=n}^{n'} \frac{1}{\mu_k b_k} \bigg/ \sum_{k=-m'}^{-m} \frac{1}{\mu_k a_k}.$$

Here, γ is chosen to make f be a constant on $[-m, n]$. By (7.10), we have

$$\begin{aligned} D(f) &= \sum_{i=-m'}^{-m} \mu_i a_i (f_i - f_{i-1})^2 + \sum_{i=n}^{n'} \mu_i b_i (f_{i+1} - f_i)^2 \\ &= \gamma^2 \sum_{i=-m'}^{-m} \frac{1}{\mu_i a_i} + \sum_{i=n}^{n'} \frac{1}{\mu_i b_i} \\ &= \left(\sum_{i=n}^{n'} \frac{1}{\mu_i b_i} \right) \left[1 + \left(\sum_{k=n}^{n'} \frac{1}{\mu_k b_k} \right) \left(\sum_{k=-m'}^{-m} \frac{1}{\mu_k a_k} \right)^{-1} \right]. \end{aligned}$$

Moreover,

$$\|f^2\|_{\mathbb{B}} \geq \|f\|_{[-m, n]}^2 = \left(\sum_{i=n}^{n'} \frac{1}{\mu_i b_i} \right)^2 \|\mathbb{1}_{[-m, n]}\|_{\mathbb{B}}.$$

Hence,

$$A_{\mathbb{B}} \geq \frac{\|f^2\|_{\mathbb{B}}}{D(f)} \geq \|\mathbb{1}_{[-m, n]}\|_{\mathbb{B}} \left[\left(\sum_{k=-m'}^{-m} \frac{1}{\mu_k a_k} \right)^{-1} + \left(\sum_{j=n}^{n'} \frac{1}{\mu_j b_j} \right)^{-1} \right]^{-1}.$$

From this, we obtain the lower estimate in (8.2). Since $\mathcal{G}$ contains a locally positive element, we have $\|\mathbb{1}_{[-m, n]}\|_{\mathbb{B}} > 0$ for large enough m and n . Letting $m', n' \rightarrow \infty$, by the recurrent assumption, it follows that $A_{\mathbb{B}} = \infty$. Besides, it is obvious that $B_{\mathbb{B}} = \infty$ in this case and so the first and then the second assertion of the theorem becomes trivial in the recurrent case. $\square$

Proof (b) above indicates an easy improvement of the lower bound of $A_{\mathbb{B}}$. Use the same f as above, and define

$$\begin{aligned} h_i^{(m,m',n,n')} &= \left[1 - \sum_{k=i+1}^{-m} \frac{1}{\mu_k a_k} \bigg/ \sum_{k=-m'}^{-m} \frac{1}{\mu_k a_k} \right]^2 \mathbb{1}_{[-m', -m-1]}(i) \\ &\quad + \left[1 - \sum_{k=n}^{i-1} \frac{1}{\mu_k b_k} \bigg/ \sum_{k=n}^{n'} \frac{1}{\mu_k b_k} \right]^2 \mathbb{1}_{[n+1, n']}(i). \end{aligned}$$

Then a simple computation shows that

$$f^2 = \left(\sum_{i=n}^{n'} \frac{1}{\mu_i b_i} \right)^2 (\mathbb{1}_{[-m, n]} + h^{(m,m',n,n')}).$$

Hence

$$\frac{\|f^2\|_{\mathbb{B}}}{D(f)} \geq \|\mathbb{1}_{[-m, n]} + h^{(m,m',n,n')}\|_{\mathbb{B}} \left[\left(\sum_{k=-m'}^{-m} \frac{1}{\mu_k a_k} \right)^{-1} + \left(\sum_{j=n}^{n'} \frac{1}{\mu_j b_j} \right)^{-1} \right]^{-1}.$$

Noting that the right-hand side is increasing in m' and n' , and making a change of the variable $-m \rightarrow m$, we obtain

$$A_{\mathbb{B}} \geq \sup_{m, n \in E: m \leq n} \|\mathbb{1}_{[m, n]} + h^{(-m, M, n, N)}\|_{\mathbb{B}} \left[\left(\sum_{k=-M}^m \frac{1}{\mu_k a_k} \right)^{-1} + \left(\sum_{j=n}^N \frac{1}{\mu_j b_j} \right)^{-1} \right]^{-1}.$$

Denote by $C_{\mathbb{B}}$ the right-hand side. Then the conclusion of Theorem 8.2 can be restated as $B_{\mathbb{B}} \leq C_{\mathbb{B}} \leq A_{\mathbb{B}} \leq 4B_{\mathbb{B}}$. Certainly, this remark is meaningful in other cases but we will not mention again.

The next result is an easier consequence of Theorem 8.2.

Corollary 8.3. Everything in the premise is the same as in Theorem 8.2. Then

(1) we have $B_{\mathbb{B}} \leq B_L \wedge B_R$, where

$$B_L = \sup_{n \in E} \sum_{i=-M}^n \frac{1}{\mu_i a_i} \|\mathbb{1}_{[n, N+1]}\|_{\mathbb{B}}, \quad B_R = \sup_{n \in E} \sum_{i=n}^N \frac{1}{\mu_i b_i} \|\mathbb{1}_{(-M-1, n]}\|_{\mathbb{B}}.$$

The equality sign holds once

$$S := \sum_{i=-M}^N \frac{1}{\mu_i a_i} + \frac{1}{\mu_N b_N} \mathbb{1}_{\{N < \infty\}} = \infty.$$

(2) Next, we have $B_{\mathbb{B}} \geq (B_L \wedge B_R) \mathbb{1}_{\{S=\infty\}} + S^{-1}B$, where

$$B = \sup_{m, n \in E: m \leq n} \left[\left(\sum_{i=-M}^m \frac{1}{\mu_i a_i} \right) \left(\sum_{k=n}^N \frac{1}{\mu_k b_k} \right) \|\mathbb{1}_{[m, n]}\|_{\mathbb{B}} \right].$$

Proof. Clearly, by (8.3), we have

$$B_{\mathbb{B}}^{-1} \geq \inf_{m \leq n} \left(\sum_{i=-M}^m \frac{1}{\mu_i a_i} \|\mathbb{1}_{[m,n]}\|_{\mathbb{B}} \right)^{-1} = \inf_{m \in E} \left(\sum_{i=-M}^m \frac{1}{\mu_i a_i} \|\mathbb{1}_{[m,N+1]}\|_{\mathbb{B}} \right)^{-1},$$

and so $B_{\mathbb{B}} \leq B_L$. The equality sign holds once $\sum_{i=\theta}^N (\mu_i b_i)^{-1} = \infty$. Similarly, we have $B_{\mathbb{B}} \leq B_R$. The equality sign holds once $\sum_{i=-M}^{\theta} (\mu_i a_i)^{-1} = \infty$. Hence, $B_{\mathbb{B}} \leq B_L \wedge B_R$ and the equality sign holds once $S = \infty$.

Next, when $S < \infty$, we have

$$B_{\mathbb{B}}^{-1} \leq S \inf_{m \leq n} \left[\left(\sum_{i=-M}^m \frac{1}{\mu_i a_i} \right) \left(\sum_{k=n}^N \frac{1}{\mu_k b_k} \right) \|\mathbb{1}_{[m,n]}\|_{\mathbb{B}} \right]^{-1} = S B^{-1}.$$

We have thus proved the corollary. $\square$

Of course, one can decompose the constant B in Corollary 8.3 (2). For instance, for fixed m_0 , we have

$$B \geq \left(\sum_{i=-M}^{m_0} \frac{1}{\mu_i a_i} \right) \sup_{m_0 \leq n < N+1} \left[\left(\sum_{k=n}^N \frac{1}{\mu_k b_k} \right) \|\mathbb{1}_{[m_0,n]}\|_{\mathbb{B}} \right].$$

The last factor is close to B_R when m_0 is negative enough. However, when $m_0 \rightarrow -M$, the first term tends to zero since $S < \infty$, unless $M < \infty$. This indicates that bounding $B_{\mathbb{B}}$ by B_L and B_R is rather rough, especially in the case that $E = \mathbb{Z}$ (cf. Example 8.9 below). This is a particularly different point of the processes on the whole $\mathbb{Z}$ or on the half space $\mathbb{Z}_+$ as shown by Corollary 8.4 below.

We now specify Theorem 8.2 and Corollary 8.3 to the half space: either M or N is finite. This corresponds to the processes studied in the first part of Section 7 (see Corollary 7.3).

Corollary 8.4. In Theorem 8.2, let $M = -1$. Then we have

$$B_{\mathbb{B}}^{-1} = \inf_{1 \leq n \leq m < N+1} \left[\left(\sum_{i=1}^n \frac{1}{\mu_i a_i} \right)^{-1} + \left(\sum_{i=m}^N \frac{1}{\mu_i b_i} \right)^{-1} \right] \|\mathbb{1}_{[n,m]}\|_{\mathbb{B}}^{-1}. \quad (8.6)$$

Furthermore, we have

$$B_L \wedge B_R \geq B_{\mathbb{B}} \geq (\mathbb{1}_{\{S=\infty\}} + (a_1 S)^{-1}) (B_L \wedge B_R),$$

where

$$\begin{aligned} B_L &= \sup_{1 \leq n < N+1} \sum_{i=1}^n \frac{1}{\mu_i a_i} \|\mathbb{1}_{[n,N+1]}\|_{\mathbb{B}}, & B_R &= \sup_{1 \leq m < N+1} \sum_{k=m}^N \frac{1}{\mu_k b_k} \|\mathbb{1}_{[1,m]}\|_{\mathbb{B}}, \\ S &= \sum_{i=1}^N \frac{1}{\mu_i a_i} + \frac{1}{\mu_N b_N} \mathbb{1}_{\{N < \infty\}}. \end{aligned} \quad (8.7)$$

Proof. The first assertion follows from Theorem 8.2 with $M = -1$ and an exchange of m and n again. The second one follows from Corollary 8.3 except the last estimate. When $S = \infty$, we have $B_{\mathbb{B}} = B_L$. While when $S < \infty$, we have

$$\begin{aligned} B_{\mathbb{B}}^{-1} &\leq S \inf_{1 \leq n \leq m < N+1} \left[\left(\sum_{i=1}^n \frac{1}{\mu_i a_i} \right) \left(\sum_{k=m}^N \frac{1}{\mu_k b_k} \right) \|\mathbb{1}_{[n,m]}\|_{\mathbb{B}} \right]^{-1} \\ &\leq a_1 S \inf_{1 \leq m < N+1} \left(\sum_{k=m}^N \frac{1}{\mu_k b_k} \|\mathbb{1}_{[1,m]}\|_{\mathbb{B}} \right)^{-1} \\ &= a_1 S B_R^{-1}. \end{aligned}$$

Therefore,

$$B_{\mathbb{B}} \geq B_L \mathbb{1}_{\{S=\infty\}} + (a_1 S)^{-1} B_R \geq (\mathbb{1}_{\{S=\infty\}} + (a_1 S)^{-1}) (B_L \wedge B_R)$$

as required. $\square$

When one of M or N is finite and its Dirichlet boundary is replaced by the Neumann one, the solution becomes simpler. The next result corresponds to the processes studied in Sections 2 and 3.

Theorem 8.5. Let $M = 0$ be the Neumann boundary and assume that $\mathcal{G}$ contains a locally positive element. Then the isoperimetric constant $B_{\mathbb{B}} := \sup_K \|\mathbb{1}_K\|_{\mathbb{B}} / \text{Cap}(K)$ can be expressed as

$$B_{\mathbb{B}} = \sup_{0 \leq n < N+1} \sum_{i=n}^N \frac{1}{\mu_i b_i} \|\mathbb{1}_{[0,n]}\|_{\mathbb{B}}. \quad (8.8)$$

In particular, for the Sobolev-type inequality, we have

$$B_p = \sup_{0 \leq n < N+1} \sum_{i=n}^N \frac{1}{\mu_i b_i} \left(\sum_{j=0}^n \mu_j \right)^{2/p}, \quad p \geq 2. \quad (8.9)$$

Proof. The proof is nearly the same as that of Theorem 8.2 except one point. In proof (b) of [11; Corollary 4.1] or [12; Corollary 7.5], to find a minimizer f for $D(f)$, since the constraint $f_0 = 0$ and $f_n = 1$, f cannot be a constant on $\{0, 1, \dots, n\}$. Now, without the constraint $f_0 = 0$, the minimizer should satisfy $f_j = 1$ for all $j : 0 \leq j \leq n$. Thus, instead of (8.5), the minimal solution becomes

$$D(f) = \left(\sum_{i=n}^q \frac{1}{\mu_i b_i} \right)^{-1}.$$

Then the necessary change of the proof of Theorem 8.2 after (8.5) should be clear. $\square$

Applying Theorem 8.5 to $\mathbb{B} = L^1(\mu)$, we return to Theorem 3.1. Actually, in parallel to [9], one may extend the results in Sections 2 and 3, Theorem 3.1 in particular, to the present setup of normed linear spaces and then deduce Theorem 8.5. The next result is obvious, it says that for a null-recurrent process, the L^p ($p \geq 1$)-Sobolev inequality is still not weak enough.

Corollary 8.6. Consider a birth-death process on $\mathbb{Z}_+$. If $\sum_{i \geq 1} \mu_i = \infty$ and $\sum_{i \geq 1} (\mu_i b_i)^{-1} = \infty$, then $B_p^{(8.4)} = B_p^{(8.9)} = \infty$ for all $p \geq 2$.

Remark 8.7. We now compare (8.6) and (8.8) in the particular case that $\sum_i \mu_i = \infty$. Then the constant $B_{\mathbb{B}}$ given in (8.6) becomes

$$B_{\mathbb{B}}^{(8.6)} = \sup_{1 \leq m < N+1} \sum_{i=m}^N \frac{1}{\mu_i b_i} \|\mathbb{1}_{[1,m]}\|_{\mathbb{B}}.$$

Rewrite the constant $B_{\mathbb{B}}$ given in (8.8) as

$$B_{\mathbb{B}}^{(8.8)} = \left(\sum_{i=0}^N \frac{1}{\mu_i b_i} \|\mathbb{1}_{\{0\}}\|_{\mathbb{B}} \right) \vee \left(\sup_{1 \leq n < N+1} \sum_{i=n}^N \frac{1}{\mu_i b_i} \|\mathbb{1}_{[0,n]}\|_{\mathbb{B}} \right).$$

By (H3), we have

$$\|\mathbb{1}_{[1,n]}\|_{\mathbb{B}} \leq \|\mathbb{1}_{[0,n]}\|_{\mathbb{B}} \leq \|\mathbb{1}_{\{0\}}\|_{\mathbb{B}} + \|\mathbb{1}_{[1,n]}\|_{\mathbb{B}}.$$

Next, by (H1), we have $\|\mathbb{1}_{\{0\}}\|_{\mathbb{B}} < \infty$. It follows that $B_{\mathbb{B}}^{(8.6)} < \infty$ iff $B_{\mathbb{B}}^{(8.8)} < \infty$.

We conclude this section by a simple example to show the role of the Poincaré-type inequalities.

Example 8.8. Consider a birth-death process on $\mathbb{Z}_+$ with $\mu_i = (i+1)^\gamma$ ($\gamma > 1$) and $b_i \equiv 1$. Then $a_i = i^\gamma (i+1)^{-\gamma}$ and

$$B_p^{(8.9)} = \sup_{n \geq 0} \left[\sum_{i=0}^n (i+1)^\gamma \right]^{2/p} \sum_{j=n}^{\infty} \frac{1}{(j+1)^\gamma}, \quad p \geq 2.$$

Hence, $B_p^{(8.9)} < \infty$ iff

$$p \geq 2 \left(1 + \frac{2}{\gamma - 1} \right).$$

However, $\delta^{(3.1)} = B_2^{(8.9)} = \infty$ for all $\gamma > 1$.

Example 8.9. Let $E = \mathbb{Z}$, $b_i \equiv 1$, $\mu_i = e^{i^2}$, and $\mathbb{B} = L^1(\mu)$. Then for the quantities given in Corollary 8.3, we have $B_L = B_R = \infty$ but $B_{\mathbb{B}} < \infty$.

Proof. Obviously, $B_L = B_R = \infty$. To show that $B_{\mathbb{B}} < \infty$, since

$$x \vee y \leq x + y \leq 2(x \vee y),$$

it suffices to prove that

$$\sup_{m \leq n} \left[\left(\sum_{i=-\infty}^m \frac{1}{\mu_i a_i} \right) \wedge \left(\sum_{k=n}^{\infty} \frac{1}{\mu_k b_k} \right) \right] \sum_{j=m}^n \mu_j < \infty.$$

By symmetry, without loss of generality, it is enough to show that

$$\sup_{m \geq n \geq 0} \left(\sum_{i=-\infty}^{-m} \frac{1}{\mu_i a_i} \right) \sum_{j=-m}^n \mu_j < \infty,$$

or

$$\sup_{m \geq 0} \left(\sum_{i=-\infty}^{-m} \frac{1}{\mu_i a_i} \right) \sum_{j=-m}^m \mu_j < \infty.$$

Equivalently,

$$\sup_{m \geq 0} \left(\sum_{i=m}^{\infty} \frac{1}{\mu_i b_i} \right) \sum_{j=0}^m \mu_j < \infty. \quad (8.10)$$

The assertion now follows by using Conte's inequality:

$$x \left(1 + \frac{x}{24} + \frac{x^2}{12} \right) e^{-3x^2/4} < e^{-x^2} \int_0^x e^{y^2} dy \leq \frac{\pi^2}{8x} (1 - e^{-x^2}), \quad x \geq 0$$

and Gautschi's estimate:

$$\frac{1}{2} \left[(x^p + 2)^{1/p} - x \right] < e^{x^p} \int_x^{\infty} e^{-y^p} dy \leq C_p \left[\left(x^p + \frac{1}{C_p} \right)^{1/p} - x \right], \quad x \geq 0,$$

$$C_p = \Gamma(1 + 1/p)^{p/(p-1)}, \quad p > 1; \quad C_2 = \pi/4.$$

Alternatively, one may check directly that the function under supremum on the left-hand side of (8.10) is decreasing in m (≥ 1) and then (8.10) follows easily. $\square$

9. GENERAL KILLING

In Sections 4 and 7, we have studied the special case having a killing at 1 only. We now study the process with general killing, as described by (2.1) with state space shifted by 1: $E = \{i : 1 \leq i < N + 1\}$. We use the same symmetric measure (μ_i) as in Section 4.

The next preliminary result is quite useful. To which it is more convenient to use $a_1 + c_1$ and $b_N + c_N$ for the killing rates at boundaries 1 and N (if $N < \infty$), respectively, rather than c_1 and c_N used in Proposition 2.1. Note that the killing rates in the next proposition are allowed to be zero identically.

Proposition 9.1. Let (a_i) and (b_i) be positive but $a_1 \geq 0$, $b_N \geq 0$ if $N < \infty$, and let (c_i) be nonnegative on E . Define $\lambda_0 = \lambda_0(a_i, b_i, c_i)$ as follows:

$$\lambda_0 = \inf \{ D(f) : \mu(f^2) = 1, f \in \mathcal{K} \},$$

where

$$D(f) = \sum_{i \in E} \mu_i b_i (f_{i+1} - f_i)^2 + \mu_1 a_1 f_1^2 + \sum_{i \in E} \mu_i c_i f_i^2, \quad f_{N+1} = 0 \text{ if } N < \infty.$$

Write $\tilde{\lambda}_0 = \lambda_0(a_i, b_i, 0)$ for simplicity. Then we have

- (1) $\lambda_0(a_i, b_i, c'_i) \geq \lambda_0(a_i, b_i, c_i)$ if $c'_i \geq c_i$ for all $i \in E$.
- (2) $\lambda_0(a_i, b_i, c_i + c) = \lambda_0(a_i, b_i, c_i) + c$ for constant $c \geq 0$.
- (3) $\tilde{\lambda}_0 + \sup_{i \in E} c_i \geq \lambda_0 \geq \tilde{\lambda}_0 + \inf_{i \in E} c_i$ and the equalities hold if c_i is a constant on E .

Proof. Since a change of $\{c_i\}_{i=1}^N$ makes no influence to $\{\mu_i\}_{i=1}^N$, part (1) is simply a comparison of the Dirichlet forms on the same space $L^2(\mu)$ with common core $\mathcal{H}$. Similarly, one can prove the other assertions. $\square$

Note that Proposition 9.1 makes a comparison for the killing rates only. Actually, a more general comparison is available in view of [3; Theorem 3.1]. Next, if (1.3) holds, then by Proposition 1.3 and the remark below (4.3), the Dirichlet is unique, and so the condition $f \in \mathcal{H}$ can be ignored in defining $\tilde{\lambda}_0$.

It is worthy to mention that the principal eigenvalue λ_0 studied here can be extended to a more general class of Schrödinger operators. That is, we may replace the nonnegative potential (c_i) with the one bounded below by a constant: $\inf_i c_i \geq -M > -\infty$. Then we have $c_i + M \geq 0$ for all i and

$$\lambda_0(a_i, b_i, c_i) = \lambda_0(a_i, b_i, c_i + M) - M \geq -M.$$

Having Proposition 9.1 at hand, all the examples for $\tilde{\lambda}_0$ given in Sections 3, 5 and 7, can be translated into the case of λ_0 with constant killing rate. For instance, we have the following example which already shows the complexity of the problem studied in this section.

Example 9.2. Let $a_i \equiv a > 0$ for $i \geq 2$, $b_i \equiv b > 0$ and $c_i \equiv c \geq 0$ for $i \geq 1$.

- (1) If $a_1 = a$, or $a_1 = 0$ but still $a \leq b$, then $\lambda_0 = (\sqrt{a} - \sqrt{b})^2 + c$.
- (2) If $a_1 = 0$ and $a > b$, then $\lambda_0 = c$.

Proof. By Proposition 9.1, we need only to study $\tilde{\lambda}_0$. In the last case, since the process is ergodic, we have $\tilde{\lambda}_0 = 0$. Next, we have $\tilde{\lambda}_0 = (\sqrt{a} - \sqrt{b})^2$ according different cases by

- (i) Example 5.3 if $a_1 = a$ and $a \geq b$,
- (ii) Example 7.7 if $a_1 = a$ and $a \leq b$,
- (iii) Example 3.4 if $a_1 = 0$ and $a \leq b$. $\square$

From now on, we return to a convention made in Section 2, the rates a_1 and b_N are combined into c_1 and c_N if $N < \infty$. Thus, in Theorem 7.1 for instance, we have $a_1 = 0$ and $c_1 > 0$, and moreover, $b_N = 0$ and $c_N > 0$ if $N < \infty$. In general, we assume that $c_i \neq 0$. Otherwise, we will return to what we treated in Sections 2 and 3. Define the operator R :

$$R_i(v) = a_i(1 - v_{i-1}^{-1}) + b_i(1 - v_i) + c_i, \quad i \in E, \quad v_0 = \infty, \quad v_N = 0 \text{ if } N < \infty,$$

for v in the set $\mathcal{V} = \{v_i > 0 : 1 \leq i < N\}$. Next, define $\tilde{\mathcal{V}} = \mathcal{V}$ if $N < \infty$. When $N = \infty$, define

$$\begin{aligned} \tilde{\mathcal{V}} = & \bigcup_{m=1}^{\infty} \left\{ v_i : v_i > 0 \text{ for } i < m, \quad v_i = 0 \text{ for } i \geq m \right\} \\ & \bigcup \left\{ v : v_i > 0 \text{ on } E, \text{ the function } f : f_1 = 1, f_i = \prod_{k=1}^{i-1} v_k \text{ (} i \geq 2 \text{)} \text{ is in } L^2(\mu) \right. \\ & \left. \text{and satisfies } \Omega f / f \leq \eta \text{ on } E \text{ for some constant } \eta \right\}. \end{aligned} \quad (9.1)$$

For $v \in \tilde{\mathcal{V}}$ with finite support, $R_{\bullet}(v)$ is also well defined by setting $1/0 = \infty$.

Theorem 9.3. Assume that $c_i \neq 0$. For λ_0 defined by (2.2) with state space $E = \{i : 1 \leq i < N + 1\}$, the following variational formulas hold:

$$\inf_{v \in \tilde{\mathcal{V}}} \sup_{i \in E} R_i(v) = \lambda_0 = \sup_{v \in \mathcal{V}} \inf_{i \in E} R_i(v). \quad (9.2)$$

Proof. (a) First, we study the lower estimate. In the case that $\sum_{k \in E} \mu_k < \infty$, as a particular consequence of [5; Theorem 1.1], we have

$$\lambda_0 \geq \sup_{g > 0} \inf_{i \in E} \frac{-\Omega g}{g}(i) = \sup_{g > 0} \inf_{i \in E} \left[a_i \left(1 - \frac{g_{i-1}}{g_i} \right) + b_i \left(1 - \frac{g_{i+1}}{g_i} \right) + c_i \right], \quad (9.3)$$

where $g_0 := 0$ and $g_{N+1} = 0$ if $N < \infty$. The proof remains true when $\sum_{k \in E} \mu_k = \infty$, simply using $E_m = \{1, 2, \dots, m\}$ ($m < N + 1$) instead of the original one. Actually, the conclusion holds in a very general setup (cf. Shiozawa and Takeda (2005) and its extension to the unbounded test functions by Zhang (2007)).

Suppose that $\lambda_0 > 0$ for a moment. Then by Proposition 2.1 (with a shift by 1 of the state space), the eigenfunction g of λ_0 is positive. It follows that the first equality sign in (9.3) can be attained and so does the last equality in (9.2) with $v_i = g_{i+1}/g_i > 0$ ($1 \leq i < N$). Next, if $\lambda_0 = 0$, then $N = \infty$ since a_i and b_i are positive for $i : 2 \leq i < N$, and $c_i \neq 0$ (in the case of Theorem 7.1, we have $c_1 > 0$ and also $c_N > 0$ if $N < \infty$). By setting $v_i \equiv 1$ for $i \in E$, we get

$$\inf_{i \in E} \left[a_i \left(1 - \frac{1}{v_{i-1}} \right) + b_i(1 - v_i) + c_i \right] \geq \inf_{i \in E} c_i \geq 0. \quad (9.4)$$

Hence, the last term of (9.2) is nonnegative. Therefore, the last equality in (9.2) is trivial if $\lambda_0 = 0$, in view of (9.3).

(b) Next, we study the upper estimate. We consider only the case that $N = \infty$. Otherwise, the proof is easier. Given $v \in \tilde{\mathcal{V}}$, let $\gamma = \gamma(v) = \sup_{1 \leq i < \infty} R_i(v)$ and as in the definition of $\tilde{\mathcal{V}}$, set

$$f_0 = 0, \quad f_1 = 1, \quad f_i = \prod_{k=1}^{i-1} v_k, \quad i \geq 2. \quad (9.5)$$

First, suppose that $\text{supp}(v) = \{1, 2, \dots, m-1\}$ for a finite m . Then $\text{supp}(f) = \{1, 2, \dots, m\}$ and

$$\frac{-\Omega f}{f}(i) = R_i(v) \leq \gamma, \quad i = 1, 2, \dots, m.$$

Hence,

$$\begin{aligned} \gamma \sum_{k=1}^m \mu_k f_k^2 &\geq \sum_{k=1}^m \mu_k f_k (-\Omega f)(k) \\ &= \sum_{k=2}^{m+1} \mu_k a_k f_{k-1} (f_{k-1} - f_k) - \sum_{k=1}^m \mu_k a_k f_k (f_{k-1} - f_k) + \sum_{k=1}^m \mu_k c_k f_k^2 \\ &= \sum_{k=1}^m \mu_k [a_k (f_{k-1} - f_k)^2 + c_k f_k^2] + \mu_{m+1} a_{m+1} f_m (f_m - f_{m+1}) \\ &= \sum_{k=1}^{m+1} \mu_k [a_k (f_{k-1} - f_k)^2 + c_k f_k^2]. \end{aligned}$$

We have not only $\gamma \geq 0$ (actually $\gamma > 0$ when m is large enough since $c_i \neq 0$) but also

$$\lambda_0 \leq \frac{D(f)}{\|f\|^2} \leq \gamma(v) \quad (9.6)$$

for all $v \in \tilde{\mathcal{V}}$ with finite support.

(c) Next, we are going to prove (9.6) in the case that $v \in \tilde{\mathcal{V}}$ with $v_i > 0$ for all $i \geq 1$. In this case, the positivity condition of v is not enough for the first equality in (9.2), as mentioned in Section 2 (above the proofs of Theorem 2.4 and Proposition 2.5). See also the specific situation given in the proof of Example 9.17 below. This explains why two additional conditions are included in the second union of the definition of $\tilde{\mathcal{V}}$. The condition “ $f \in L^2(\mu)$ ” is essential but not the one “ $\Omega f/f \leq \eta$ ” since the eigenfunction g of λ_0 satisfies “ $\Omega g/g = -\lambda_0$ ”. To prove (9.6), without loss of generality, assume that $\gamma = \gamma(v) < \infty$. Otherwise, (9.6) is trivial. Clearly, $\gamma \geq R_1(v) = b_1(1 - v_1) + c_1 > -\infty$. Note that by assumptions, the function f possesses the following properties:

- (i) $f > 0$ on E .
- (ii) $f \in L^2(\mu)$ and then $P_t f \in L^2(\mu)$, where $P_t = (p_{ij}(t))$ is the minimal semigroup determined by the Dirichlet form.
- (iii) $|\Omega f(i)| = |\sum_j q_{ij} f_j| \leq \max\{|\eta|, |\gamma|\} f_i$ for all $i \in E$.

Here, property (iii) comes from

$$-\eta f \leq -\Omega f \leq \gamma f.$$

Since $(p_{ij}(t))$ satisfies the forward Kolmogorov equation:

$$p_{ij}(t) = \delta_{ij} + \int_0^t \sum_k p_{ik}(s) q_{kj} ds$$

and (i), it follows that

$$P_t f(i) = f_i + \sum_j \int_0^t \sum_k p_{ik}(s) q_{kj} f_j ds.$$

By (ii), $P_t f(i) < \infty$ and is continuous in t . Because of this and (iii), the order of the last two sums and also the integration are exchangeable. This leads to

$$P_t f(i) \geq f_i - \gamma \int_0^t \sum_k p_{ik}(s) f_k ds = f_i - \gamma \int_0^t P_s f(i) ds, \quad i \in E, \quad (9.7)$$

since by assumption $\Omega f \geq -\gamma f$. Therefore, we obtain

$$0 < D(f) = \lim_{t \downarrow 0} \frac{1}{t} (f, f - P_t f) \leq \lim_{t \downarrow 0} \frac{\gamma}{t} \int_0^t (f, P_s f) ds = \gamma(v) \|f\|^2 < \infty.$$

Here, the first limit is due to (ii) and the first equality in (1.10), the last inequality comes from (i) and (9.7). We have thus proved that not only $\gamma > 0$ but also

$f \in \mathcal{D}(D)$ and so we have returned to (9.6). In other words, (9.6) holds for all $v \in \tilde{\mathcal{V}}$. By making infimum with respect to $v \in \tilde{\mathcal{V}}$, we obtain

$$\lambda_0 \leq \inf_{v \in \tilde{\mathcal{V}}} \sup_{i \in E} R_i(v).$$

(d) To prove the equality sign in the last formula holds, in view of proof (b), we have actually proved that for every finite m ,

$$\begin{aligned} \lambda_0^{(m)} &:= \inf\{D(f) : f_0 = 0, f_i = 0 \text{ for all } i \geq m+1, \|f\| = 1\} \\ &\leq \inf_{v \in \tilde{\mathcal{V}}_m} \sup_{1 \leq i \leq m} R_i(v), \end{aligned} \quad (9.8)$$

where

$$\tilde{\mathcal{V}}_m = \{v_i : v_i > 0 \text{ for } i < m, v_m = 0\}.$$

Actually, there is a $\bar{v} \in \tilde{\mathcal{V}}_m$ such that $R_i(\bar{v}) = \lambda_0^{(m)} > 0$ for all i ($1 \leq i \leq m$) since $m < \infty$ and then the equality sign in (9.8) holds. Therefore, the first equality in (9.2) holds since $\lambda_0^{(m)} \downarrow \lambda_0$ as $m \uparrow \infty$. $\square$

We now begin to study the estimate of λ_0 . First, by Proposition 7.17, we have a simple upper bound:

$$\lambda_0 \leq \inf_{i \in E} (a_i + b_i + c_i).$$

Hence, $\lambda_0 = 0$ whenever $\lim_{n \rightarrow \infty} (a_n + b_n + c_n) = 0$. The next result provides us a finer upper bound. It is motivated from Theorem 3.1.

Proposition 9.4. Let $\tilde{c}_i = c_i - \inf_i c_i$. Then

$$\lambda_0 \leq \inf_{i \in E} c_i + \inf_{\ell \in E} \left(\sum_{i=1}^{\ell} \mu_i \right)^{-1} \inf_{E \ni m \geq \ell} \left[\left(\sum_{k=\ell}^m \frac{1}{\mu_k b_k} \right)^{-1} + \sum_{i=1}^m \mu_i \tilde{c}_i \right] \quad (9.9)$$

$$\leq \inf_{i \in E} c_i + \inf_{\ell \in E} \left(\sum_{i=1}^{\ell} \mu_i \right)^{-1} \left[\mu_{\ell} b_{\ell} + \sum_{i=1}^{\ell} \mu_i \tilde{c}_i \right]. \quad (9.10)$$

Proof. By Proposition 9.1, it is enough to consider the case that $\tilde{c}_i \equiv c_i$, i.e., $\inf_i c_i = 0$. Fix $\ell \leq m$ and define

$$\varphi_i = \varphi_i^{(\ell, m)} = \mathbb{1}_{\{i \leq m\}} \sum_{k=i \vee \ell}^m \frac{1}{\mu_k b_k}, \quad i \in E.$$

Then

$$\begin{aligned} \mu(\varphi^2) &= \sum_{i=1}^{\ell} \mu_i \varphi_{\ell}^2 + \sum_{i=\ell+1}^m \mu_i \varphi_i^2 \geq \varphi_{\ell}^2 \sum_{i=1}^{\ell} \mu_i, \\ D(\varphi) &= \sum_{k=\ell}^m \frac{1}{\mu_k b_k} + \varphi_{\ell}^2 \sum_{i=1}^{\ell} \mu_i c_i + \sum_{i=\ell+1}^m \mu_i c_i \varphi_i^2 \leq \varphi_{\ell} + \varphi_{\ell}^2 \sum_{i=1}^m \mu_i c_i. \end{aligned}$$

Hence,

$$\frac{D(\varphi)}{\mu(\varphi^2)} \leq \left(\sum_{i=1}^{\ell} \mu_i \right)^{-1} \left[\varphi_{\ell}^{-1} + \sum_{i=1}^m \mu_i c_i \right].$$

Because $\varphi^{(\ell, m)} \in \mathcal{H}$, it follows that

$$\begin{aligned} \lambda_0 &\leq \inf_{\ell \in E} \inf_{E \ni m \geq \ell} \frac{D(\varphi)}{\mu(\varphi^2)} \\ &\leq \inf_{\ell \in E} \left(\sum_{i=1}^{\ell} \mu_i \right)^{-1} \inf_{E \ni m \geq \ell} \left[\left(\sum_{k=\ell}^m \frac{1}{\mu_k b_k} \right)^{-1} + \sum_{i=1}^m \mu_i c_i \right] \\ &\leq \inf_{\ell \in E} \left(\sum_{i=1}^{\ell} \mu_i \right)^{-1} \inf_{E \ni m \geq \ell} \left[\mu_{\ell} b_{\ell} + \sum_{i=1}^m \mu_i c_i \right] \\ &= \inf_{\ell \in E} \left(\sum_{i=1}^{\ell} \mu_i \right)^{-1} \left[\mu_{\ell} b_{\ell} + \sum_{i=1}^{\ell} \mu_i c_i \right]. \quad \square \end{aligned}$$

As an immediate consequence of (9.10), we obtain the following result.

Corollary 9.5. If $\sum_{i=1}^{\infty} \mu_i = \infty$ and

$$\overline{\lim}_{m \rightarrow \infty} \mu_m b_m \left(\sum_{i=1}^m \mu_i \right)^{-1} = 0,$$

then

$$\lambda_0 \leq \inf_{i \in E} c_i + \overline{\lim}_{m \rightarrow \infty} \sum_{i=1}^m \mu_i \tilde{c}_i / \sum_{i=1}^m \mu_i \leq \inf_{i \in E} c_i + \overline{\lim}_{n \rightarrow \infty} \tilde{c}_n.$$

Proof. Without loss of generality, assume that $\tilde{c}_i \equiv c_i$.

By assumptions, it follows that

$$\begin{aligned} &\overline{\lim}_{\ell \rightarrow \infty} \left(\sum_{i=1}^{\ell} \mu_i \right)^{-1} \left[\mu_{\ell} b_{\ell} + \sum_{i=1}^{\ell} \mu_i c_i \right] \\ &\leq \overline{\lim}_{\ell \rightarrow \infty} \mu_{\ell} b_{\ell} \left(\sum_{i=1}^{\ell} \mu_i \right)^{-1} + \overline{\lim}_{\ell \rightarrow \infty} \left(\sum_{i=1}^{\ell} \mu_i \right)^{-1} \sum_{i=1}^{\ell} \mu_i c_i \\ &= \overline{\lim}_{\ell \rightarrow \infty} \left(\sum_{i=1}^{\ell} \mu_i \right)^{-1} \sum_{i=1}^{\ell} \mu_i c_i. \end{aligned}$$

The first inequality now follows from (9.10).

To prove the second inequality, let $\gamma = \overline{\lim}_{n \rightarrow \infty} c_n \in [0, \infty]$. Then for every $\varepsilon > 0$, we have $\sup_{k \geq n} c_k \leq \gamma + \varepsilon$ for large enough n . Hence,

$$\sum_{i=1}^{\ell} \mu_i c_i = \sum_{i=1}^n \mu_i c_i + \sum_{i=n+1}^{\ell} \mu_i c_i \leq \sum_{i=1}^n \mu_i c_i + (\gamma + \varepsilon) \sum_{i=n+1}^{\ell} \mu_i, \quad \ell > n.$$

We have thus obtained

$$\begin{aligned} \lim_{\ell \rightarrow \infty} \left(\sum_{i=1}^{\ell} \mu_i \right)^{-1} \sum_{i=1}^{\ell} \mu_i c_i &\leq \lim_{\ell \rightarrow \infty} \left(\sum_{i=1}^{\ell} \mu_i \right)^{-1} \left[\sum_{i=1}^n \mu_i c_i + (\gamma + \varepsilon) \sum_{i=n+1}^{\ell} \mu_i \right] \\ &= \gamma + \varepsilon \end{aligned}$$

as required. $\square$

To study the lower estimate of λ_0 , we observe that not every positive sequence (v_i) is useful for the lower estimate given in (9.2) since one may have $\inf_i R_i(v) < 0$. In order for $\inf_i R_i(v) \geq 0$, it is necessary that

$$0 < v_i \leq \frac{1}{b_i} \left(c_i + a_i + b_i - \frac{a_i}{v_{i-1}} \right).$$

From this, we obtain the following necessary condition:

$$\frac{a_{i+1}}{c_{i+1} + a_{i+1} + b_{i+1}} < v_i \leq x_i - \frac{y_i}{x_{i-1} - \frac{y_{i-1}}{x_{i-2} - \frac{y_{i-2}}{\ddots x_3 - \frac{y_3}{x_2 - \frac{y_2}{x_1}}}}},$$

where

$$x_i = \frac{c_i + a_i + b_i}{b_i}, \quad y_i = \frac{a_i}{b_i}.$$

However, the condition is clearly not practical. Because of this reason, we are now going to introduce an alternative variational formula for the lower estimates.

For a given sequence (r_i) , define an operator $II^r = II^{(r_i)}$ of “double sum” on the set of positive functions (f_i) as follows:

$$II_1^r(f) = 0, \quad II_i^r(f) = \sum_{k=1}^{i-1} \frac{1}{\mu_k b_k} \sum_{j=1}^k r_j \mu_j f_j = \sum_{j=1}^{i-1} r_j f_j \mu_j \sum_{k=j}^{i-1} \frac{1}{\mu_k b_k}, \quad E \ni i \geq 2.$$

Write $II(f) = II^{\mathbf{1}}(f)$. For a fixed sequence (c_i) , let $\tilde{c}_i = c_i - \inf_i c_i$ and define

$$\mathcal{F} = \{f > 0 : f_i < f_1 + II_i^{\tilde{c}}(f) \text{ for all } E \ni i \geq 2\}. \quad (9.11)$$

Clearly, if $\tilde{c}_1 > 0$, then every positive constant function belongs to $\mathcal{F}$. Otherwise, every $f > 0$ with $f_i < f_1$ for all $E \ni i \geq 2$ belongs to $\mathcal{F}$.

Theorem 9.6. Let II^r , $(\tilde{c}_i)$ and $\mathcal{F}$ be defined as above. Next, for each fixed $f \in \mathcal{F}$, define

$$\xi = \xi_f = \begin{cases} \inf_{E \ni i \geq 2} \frac{f_1 - f_i + II_i^{\tilde{c}}(f)}{II_i(f)}, & N = \infty, \\ \inf_{E \ni i \geq 2} \frac{f_1 - f_i + II_i^{\tilde{c}}(f)}{II_i(f)} \bigwedge \frac{\sum_{j=1}^N \tilde{c}_j \mu_j f_j}{\sum_{j=1}^N \mu_j f_j}, & N < \infty, \end{cases} \quad (9.12)$$

$$\zeta(\eta, f) = \begin{cases} \inf_{E \ni i \geq 2, \tilde{c}_i < \eta} \left[\tilde{c}_i + \frac{(\eta - \tilde{c}_i) f_i}{f_1 + II_i^{\tilde{c}-\eta}(f)} \right], & \{E \ni i \geq 2 : \tilde{c}_i < \eta\} \neq \emptyset, \\ \eta, & \{E \ni i \geq 2 : \tilde{c}_i < \eta\} = \emptyset, \end{cases} \quad (9.13)$$

$\eta \in [0, \xi].$

Then we have

$$\lambda_0 \geq \inf_{i \in E} c_i + \zeta(\eta, f) \quad \text{and} \quad \eta \geq \zeta(\eta, f), \quad f \in \mathcal{F}, \eta \in [0, \xi]. \quad (9.14)$$

Moreover, for fixed f , $\zeta(\eta, f)$ is increasing in η and furthermore,

$$\lambda_0 = \inf_{i \in E} c_i + \sup_{f \in \mathcal{F}} \zeta(\xi, f). \quad (9.15)$$

Remark 9.7. To indicate the dependence on $(\tilde{c}_i)$, rewrite $\zeta(\eta, f)$ as $\zeta(\tilde{c}_i, \eta, f)$. Similarly, we have $\xi(\tilde{c}_i, f)$. Then for each $f \in \mathcal{F}$ and constant $\gamma \geq 0$, we have a shift property as follows:

$$\xi(\tilde{c}_i + \gamma, f) = \gamma + \xi(\tilde{c}_i, f), \quad \zeta(\tilde{c}_i + \gamma, \eta + \gamma, f) = \gamma + \zeta(\tilde{c}_i, \eta, f). \quad (9.16)$$

Hence, the use of $\inf_{i \in E} c_i$ in Theorem 9.6 is not essential but only for simplifying the computations. The same property holds for (9.10) but not for (9.9).

As will be illustrated later by Examples 9.17 and 9.19, it is not unusual that $\xi_f > \lambda_0$ for some $f \in \mathcal{F}$. In that case, we certainly have $\xi_f > \zeta(\xi_f, f)$. This means that ξ_f may not be a lower bound of λ_0 and so the use of $\zeta(\eta, f)$ in Theorem 9.6 is necessary.

Proof of Theorem 9.6. By Proposition 9.1, for simplicity, we assume that $\tilde{c}_i \equiv c_i$.

(a) First, we prove “ $\lambda_0 \geq$ ” in (9.14). Fix $f \in \mathcal{F}$. Then $\xi = \xi_f \geq 0$. Without loss of generality, assume that $(\xi \geq) \eta > 0$. Otherwise, the assertion is trivial. Let

$$h_i = f_1 + \Pi_i^{c-\eta}(f), \quad i \in E, \eta \in (0, \xi].$$

Since by (9.12),

$$f_1 - f_i + \Pi_i^c(f) \geq \eta \Pi_i(f) > 0$$

for $E \ni i \geq 2$ and $h_1 = f_1 > 0$, we have $h > 0$. Next, define $v_i = h_{i+1}/h_i$ ($v_0 := \infty$ and $v_N = 0$ if $N < \infty$). Then for $i : 2 \leq i < N$, since

$$h_i - h_{i+1} = \Pi_i^{c-\eta}(f) - \Pi_{i+1}^{c-\eta}(f) = \frac{1}{\mu_i b_i} \sum_{j=1}^i (\eta - c_j) \mu_j f_j,$$

we have

$$\begin{aligned} & a_i(1 - v_{i-1}^{-1}) + b_i(1 - v_i) \\ &= \frac{1}{h_i} [a_i(h_i - h_{i-1}) + b_i(h_i - h_{i+1})] \\ &= \frac{1}{h_i} \left[-\frac{a_i}{\mu_{i-1} b_{i-1}} \sum_{j=1}^{i-1} (\eta - c_j) \mu_j f_j + \frac{b_i}{\mu_i b_i} \sum_{j=1}^i (\eta - c_j) \mu_j f_j \right] \\ &= \frac{(\eta - c_i) f_i}{h_i}. \end{aligned}$$

This also holds when $i = 1$ (noting that $a_1 = 0$):

$$b_1(1 - v_i) = \frac{b_1}{h_1}(h_1 - h_2) = \frac{(\eta - c_1)f_1}{h_1} = \eta - c_1.$$

If $N < \infty$, then at $i = N$, by assumption

$$\eta \leq \xi \leq \sum_{j=1}^N c_j \mu_j f_j \bigg/ \sum_{j=1}^N \mu_j f_j,$$

we get

$$a_N(1 - v_{N-1}^{-1}) + b_N(1 - v_N) = -\frac{a_N}{h_N \mu_{N-1} b_{N-1}} \sum_{j=1}^{N-1} (\eta - c_j) \mu_j f_j \geq \frac{(\eta - c_N)f_N}{h_N}.$$

Combining these facts together, we arrive at

$$R_i(v) = c_i + a_i(1 - v_{i-1}^{-1}) + b_i(1 - v_i) \geq c_i + \frac{(\eta - c_i)f_i}{h_i}, \quad i \in E. \quad (9.17)$$

We now show that the right-hand side of (9.17) is nonnegative for all i and so we have ruled out the useless case that $\inf_i R_i(v) < 0$. Since $h > 0$, the assertion is equivalent to

$$c_i h_i \geq (c_i - \eta) f_i, \quad i \in E,$$

or

$$c_i [f_1 - f_i + II_i^c(f)] \geq \eta [c_i II_i(f) - f_i].$$

This is trivial if $c_i II_i(f) \leq f_i$ (in particular if $i = 1$) since $f_1 - f_i + II_i^c(f) \geq 0$ for all $E \ni i \geq 2$ and $f \in \mathcal{F}$. Otherwise, by the definition of ξ and η , we have

$$f_1 - f_i + II_i^c(f) \geq \xi II_i(f) \geq \eta II_i(f) > \eta [II_i(f) - f_i/c_i], \quad E \ni i \geq 2. \quad (9.18)$$

We have thus proved the required assertion.

By Theorem 9.3 and (9.17), we obtain

$$\begin{aligned} \lambda_0 &\geq \sup_{f \in \mathcal{F}} \inf_{i \in E} \left[c_i + \frac{(\eta - c_i)f_i}{f_1 + II_i^{c-\eta}(f)} \right] \\ &= \sup_{f \in \mathcal{F}} \left\{ \eta \wedge \inf_{E \ni i \geq 2} \left[c_i + \frac{(\eta - c_i)f_i}{f_1 + II_i^{c-\eta}(f)} \right] \right\}. \end{aligned} \quad (9.19)$$

Here, the last line is due to the fact that $II_1^r(f) = 0$.

(b) To prove the first assertion of the theorem, we show that for each $i: 2 \leq i \in E$,

$$c_i + \frac{(\eta - c_i)f_i}{f_1 + II_i^{c-\eta}(f)} \geq \eta \quad \text{iff} \quad c_i \geq \eta.$$

Clearly, the inequality is equivalent to

$$(\eta - c_i)f_i \geq (\eta - c_i)[f_1 + \Pi_i^{c-\eta}(f)].$$

The required assertion then follows since by (9.18), we already have

$$f_i \leq f_1 + \Pi_i^{c-\eta}(f).$$

As a consequence of the assertion, we have $\eta \geq \zeta(\eta, f)$. Now, from (9.19), it follows that

$$\lambda_0 \geq \sup_{f \in \mathcal{F}} \eta \wedge \zeta(\eta, f) = \sup_{f \in \mathcal{F}} \zeta(\eta, f).$$

This gives us the first assertion of the theorem.

(c) To prove the monotonicity of $\zeta(\eta, f)$ in η , let $\eta_1 < \eta_2 \leq \xi$. If $\{E \ni i \geq 2 : c_i < \eta_2\} = \emptyset$, then $\{E \ni i \geq 2 : c_i < \eta_1\} = \emptyset$ and so

$$\zeta(\eta_2, f) = \eta_2 > \eta_1 = \zeta(\eta_1, f).$$

If $\{E \ni i \geq 2 : c_i < \eta_1\} \neq \emptyset$, since $\{E \ni i \geq 2 : c_i < \eta_1\} \subset \{E \ni i \geq 2 : c_i < \eta_2\}$, we need only to show that

$$\frac{(\eta_2 - c_i)f_i}{f_1 + \Pi_i^{c-\eta_2}(f)} \geq \frac{(\eta_1 - c_i)f_i}{f_1 + \Pi_i^{c-\eta_1}(f)} \quad \text{on } \{E \ni i \geq 2 : c_i < \eta_2\} \neq \emptyset.$$

Actually, this is enough even if $\{E \ni i \geq 2 : c_i < \eta_1\} = \emptyset$ in view of (b). Now, the required conclusion is trivial on the set $\{E \ni i \geq 2 : \eta_1 \leq c_i < \eta_2\}$. Hence, it suffices to show that

$$\frac{\eta_2 - c_i}{f_1 + \Pi_i^{c-\eta_2}(f)} \geq \frac{\eta_1 - c_i}{f_1 + \Pi_i^{c-\eta_1}(f)} \quad \text{on } \{E \ni i \geq 2 : c_i < \eta_1\}.$$

A simple computation shows that this is equivalent to

$$f_1 + \Pi_i^c(f) \geq c_i \Pi_i(f),$$

which holds on $\{E \ni i \geq 2 : c_i < \eta_1\}$ in view of (9.12) and $\xi > \eta_1$.

(d) To prove (9.15), it suffices to show that the equality in (9.19) holds for $\eta = \xi$. Noting that the right-hand side of (9.19) is nonnegative, without loss of generality, we may assume that $\lambda_0 > 0$. Then, by Proposition 2.1, the eigenfunction $g > 0$ of λ_0 satisfies

$$\mu_k b_k (g_k - g_{k+1}) = \sum_{j=1}^k (\lambda_0 - c_j) \mu_j g_j, \quad k \in E, \quad g_{N+1} = 0 \text{ if } N < \infty.$$

Hence,

$$g_1 - g_i = \Pi_i^{\lambda_0 - c}(g), \quad i \in E, \quad \sum_{j=1}^N (\lambda_0 - c_j) \mu_j g_j = 0 \quad \text{if } N < \infty$$

and furthermore, $g \in \mathcal{F}$. It follows that

$$\begin{aligned} \frac{g_1 - g_i + \Pi_i^c(g)}{\Pi_i(g)} &\equiv \lambda_0, & E \ni i \geq 2, \\ \frac{\sum_{j=1}^N c_j \mu_j g_j}{\sum_{j=1}^N \mu_j g_j} &= \lambda_0 & \text{if } N < \infty, \\ c_i + \frac{(\lambda_0 - c_i)g_i}{g_1 + \Pi_i^{c-\lambda_0}(g)} &\equiv \lambda_0, & i \in E. \end{aligned}$$

Therefore, $\xi_g = \lambda_0$, and furthermore, the equality sign in (9.19) is attained at $(f, \eta) = (g, \lambda_0)$. $\square$

We now make a rough comparison of Theorems 9.6 and 9.3 for the lower estimate. See also the comment below the proof of Corollary 9.9.

Remark 9.8. For a given positive sequence (v_i) such that $\inf_{i \in E} R_i(v) := \gamma_v \geq 0$, corresponding to the sequence (f_i) and ξ_f defined by (9.5) and (9.12), respectively, we have $\xi_f \geq \gamma_v$.

Proof. From the assumption

$$R_i(v) = c_i + a_i(1 - v_{i-1}^{-1}) + b_i(1 - v_i) \geq \gamma_v =: \gamma, \quad i \in E,$$

it follows that

$$f_k - f_{k+1} \geq \frac{1}{\mu_k b_k} \sum_{j=1}^k (\gamma - c_j) \mu_j f_j,$$

and then

$$f_1 - f_i \geq \sum_{k=1}^{i-1} \frac{1}{\mu_k b_k} \sum_{j=1}^k (\gamma - c_j) \mu_j f_j = \Pi_i^{\gamma-c}(f), \quad i \in E.$$

To prove our assertion, without loss of generality, assume that $\gamma > 0$. Then it is clear not only that $f \in \mathcal{F}$ but also $\xi_f \geq \gamma$. $\square$

As a complement to Remark 9.8, it would be nice if we could show that

$$c_i + \frac{(\xi_f - c_i)f_i}{f_1 + \Pi_i^{c-\xi_f}(f)} \geq \gamma_v \quad \text{on the set } \{E \ni i \geq 2 : c_i < \xi_f\}.$$

This holds obviously on the subset $\{\gamma_v \leq c_i < \xi_f\}$, but is not clear on the subset $\{E \ni i \geq 2 : c_i < \gamma_v\}$.

The next result is a particular application of Theorem 9.6. It is a complement of Corollary 9.5. The combination of Proposition 9.4 and Corollary 9.5 with Corollary 9.9 below indicates that when $\lambda_0(a_i, b_i, 0) = 0$, the condition $\lim_{n \rightarrow \infty} c_n > 0$ is crucial for $\lambda_0(a_i, b_i, c_i) > 0$. This is more or less clear in terms of the Feynman-Kac formula:

$$P_t^c f(x) = \mathbb{E}^x \left[f(X_t) e^{-\int_0^{t \wedge \tau} c_{X_s} ds} \right],$$

where $\{P_t^c\}_{t \geq 0}$ is the minimal semigroup generalized by the operator with rates (a_i, b_i, c_i) , $\{X_t\}_{0 \leq t < \tau}$ is the minimal process with rates (a_i, b_i) , and τ is the life time of $\{X_t\}$. Note that $\lambda_0(a_i, b_i, 0) > 0$, and hence, $\lambda_0(a_i, b_i, c_i) > 0$ if the uniqueness condition (1.2) fails. Otherwise, $\tau = \infty$.

Corollary 9.9. Let $\varepsilon \in (0, 1)$. Define

$$\xi_\varepsilon = \begin{cases} \inf_{E \ni i \geq 2} \frac{1 - \varepsilon + \tilde{c}_1 z_i + \varepsilon x_i}{z_i + \varepsilon y_i}, & N = \infty, \\ \inf_{E \ni i \geq 2} \frac{1 - \varepsilon + \tilde{c}_1 z_i + \varepsilon x_i}{z_i + \varepsilon y_i} \bigwedge \frac{\tilde{c}_1 + \varepsilon \sum_{j=2}^N \tilde{c}_j \mu_j}{1 + \varepsilon \sum_{j=2}^N \mu_j}, & N < \infty, \end{cases} \quad (9.20)$$

$$\zeta_\varepsilon = \begin{cases} \inf_{E \ni i \geq 2: \tilde{c}_i < \xi_\varepsilon} \left[\tilde{c}_i + \frac{\varepsilon(\xi_\varepsilon - \tilde{c}_i)}{1 + \tilde{c}_1 z_i + \varepsilon x_i - \xi_\varepsilon(z_i + \varepsilon y_i)} \right], & \{E \ni i \geq 2: \tilde{c}_i < \xi_\varepsilon\} \neq \emptyset, \\ \xi_\varepsilon, & \{E \ni i \geq 2: \tilde{c}_i < \xi_\varepsilon\} = \emptyset, \end{cases} \quad (9.21)$$

where

$$x_i = \sum_{2 \leq j \leq i-1} \tilde{c}_j \mu_j \nu[j, i-1], \quad y_i = \sum_{2 \leq j \leq i-1} \mu_j \nu[j, i-1], \quad z_i = \nu[1, i-1],$$

and $\nu[i, j] = \sum_{i \leq k \leq j} (\mu_k b_k)^{-1}$. Then we have $\lambda_0 \geq \inf_{i \in E} c_i + \sup_{\varepsilon \in (0, 1)} \zeta_\varepsilon$. The same conclusion holds if ξ_ε in (9.21) is replaced by $\eta \in [0, \xi_\varepsilon]$. In particular, if $\lim_{n \rightarrow \infty} c_n > 0$, then $\lambda_0 > 0$.

Proof. (a) The main assertion of the corollary is an application of Theorem 9.6 to the specific $f \in \mathcal{F}$: $f_1 = 1$, $f_i = \varepsilon \in (0, 1)$ ($E \ni i \geq 2$), for which we have

$$H_1^r(f) = 0, \quad H_i^r(f) = r_1 \sum_{1 \leq k \leq i-1} \frac{1}{\mu_k b_k} + \varepsilon \sum_{2 \leq j \leq i-1} r_j \mu_j \sum_{j \leq k \leq i-1} \frac{1}{\mu_k b_k}, \quad E \ni i \geq 2.$$

Then (9.20) and (9.21) follows from (9.12) and (9.13), respectively.

We now prove the particular assertion for which $N = \infty$.

(b) If (1.2) does not hold, then $\lambda_0(a_i, b_i, 0) > 0$ by Theorem 3.1, and so $\lambda_0 > 0$ by part (3) of Proposition 9.1. Similarly, if $\inf_i c_i > 0$, then we have again $\lambda_0 > 0$. Thus, without loss of generality, assume that

$$\inf_i c_i = 0 \text{ and (1.2) holds.}$$

(c) With the test function f given in (a), by (9.12), we have

$$\xi_\varepsilon = \inf_{i \geq 2} \frac{1 - \varepsilon + H_i^c(f)}{H_i(f)}.$$

By assumption, there exist $\gamma > 0$ and $m \geq 2$ such that $c_i > \gamma$ for all $i \geq m$. Certainly, we have

$$\xi_\varepsilon \geq \inf_{2 \leq i \leq m} \frac{1 - \varepsilon + H_i^c(f)}{H_i(f)} \bigwedge \inf_{i > m} \frac{H_i^c(f)}{H_i(f)}.$$

For $i > m$, we have

$$\begin{aligned} \frac{H_i^c(f)}{H_i(f)} &\geq \sum_{j=m}^{i-1} c_j f_j \mu_j \sum_{k=j}^{i-1} \frac{1}{\mu_k b_k} \bigg/ \sum_{j=1}^{i-1} f_j \mu_j \sum_{k=j}^{i-1} \frac{1}{\mu_k b_k} \\ &> \varepsilon \gamma \sum_{j=m}^{i-1} \mu_j \sum_{k=j}^{i-1} \frac{1}{\mu_k b_k} \bigg/ \sum_{j=1}^{i-1} \mu_j \sum_{k=j}^{i-1} \frac{1}{\mu_k b_k}. \end{aligned}$$

By assumption (1.2), the right-hand side goes to $\varepsilon \gamma > 0$ as $i \rightarrow \infty$. It follows that $\inf_{i>m} H_i^c(f)/H_i(f) > 0$, and furthermore, there exists $\eta \in (0, \gamma)$ such that $\xi_\varepsilon > \eta$.

(d) Noting that the set $\{i \geq 2 : c_i < \eta\} \subset \{i : 2 \leq i < m\}$ is finite, by (9.13), we have

$$\zeta(\eta, f) = \inf_{i \geq 2, c_i < \eta} \left[c_i + \frac{(\eta - c_i) f_i}{f_1 + H_i^{c-\eta}(f)} \right] \geq \min_{i \geq 2, c_i < \eta} \frac{(\eta - c_i) f_i}{f_1 + H_i^{c-\eta}(f)} > 0.$$

Now, by Theorem 9.6 or proof (a) above, we conclude that $\lambda_0 > 0$. $\square$

From proof (c) above, we have seen that when $N = \infty$,

$$\xi_\varepsilon > 0 \quad \text{iff} \quad \inf_{i>m} H_i^{\tilde{c}}(\mathbb{1})/H_i(\mathbb{1}) > 0 \quad \text{for all } m \geq 2. \quad (9.22)$$

Note that

$$\inf_{i>m} \frac{H_i^{\tilde{c}}(\mathbb{1})}{H_i(\mathbb{1})} \geq \inf_{i \geq 1} \sum_{j=1}^i \mu_j \tilde{c}_j \bigg/ \sum_{j=1}^i \mu_j$$

and the right-hand side is positive iff

$$\lim_{m \rightarrow \infty} \sum_{j=1}^m \mu_j \tilde{c}_j \bigg/ \sum_{j=1}^m \mu_j > 0. \quad (9.23)$$

Thus, Corollary 9.9 is qualitatively consistent with Corollary 9.5.

In view of Remark 9.8, it is not obvious that Theorem 9.6 improves Theorem 9.3. An easier way to see the improvement is as follows. Recall that the last assertion of Corollary 9.9 is deduced in terms of the test function f used in its proof (a). For which, the corresponding sequence (v_i) is $v_1 = 1/2$ and $v_i = 1$ for all $i \geq 2$. Inserting this into $R(v)$, we get

$$\inf_{i \geq 1} R_i(v) = (c_1 + b_1/2) \wedge (c_2 - a_2) \wedge \inf_{i \geq 3} c_i.$$

Thus, for $\inf_{i \geq 1} R_i(v) > 0$, it is necessary that $\inf_{i \geq 3} c_i > 0$, which is clearly much stronger than the last condition $\lim_{n \rightarrow \infty} c_n > 0$ used in Corollary 9.9.

As Proposition 9.4, the next result is also motivated from Theorem 3.1.

Corollary 9.10. An explicit lower estimate can be obtained by Theorem 9.6 using the specific test function $f^{(m)}$:

$$f_i^{(m)} = \left(\sum_{j=i \wedge m}^m \frac{1}{\mu_j b_j} \right)^{1/2}, \quad i \in E,$$

where m may be optimized over $\{m \in E : m \geq 2\}$ (or over E if $\tilde{c}_1 > 0$).

We now show that some special killing (or Schrödinger) case can be regarded as a perturbation of the one without killing. To do so, fix constants $\beta, \gamma > 0$, and define

$$\begin{aligned} \hat{a}_i &= b_{i-1}, \quad 2 \leq i < N+1, & \hat{b}_i &= a_{i+1}, \quad 1 \leq i < N, \\ \hat{a}_1 &= \beta, \quad \hat{b}_N = \gamma & \text{if } N < \infty; \\ \check{a}_i &= a_{i+1}, \quad 0 \leq i < N, & \check{b}_i &= b_i, \quad 1 \leq i < N+1. \\ \check{b}_0 &= \beta, \quad \check{a}_N = \gamma & \text{if } N < \infty. \end{aligned} \quad (9.24)$$

Note that $(\hat{a}_i, \hat{b}_i)$ and $(\check{a}_i, \check{b}_i)$ are dual each other in the sense of Section 5 but they are clearly different from (a_i, b_i) . Recall that $a_1 = 0$ and $b_N = 0$ by convention. Next, let (c_i) satisfy

$$c_i \geq \begin{cases} a_{i+1} - a_i - b_i + b_{i-1}, & 2 \leq i < N, \\ a_2 - b_1 + \beta, & i = 1, \\ \gamma - a_N + b_{N-1}, & i = N < \infty. \end{cases} \quad (9.25)$$

Note that the right-hand side of (9.25) can be negative. Conversely, for given rates $(\hat{a}_i, \hat{b}_i)$, the inverse transform is as follows:

$$\begin{aligned} a_i &= \hat{b}_{i-1}, \quad 2 \leq i < N+1, & b_i &= \hat{a}_{i+1}, \quad 1 \leq i < N, \\ c_i &\geq \hat{b}_i - \hat{b}_{i-1} - \hat{a}_{i+1} + \hat{a}_i, & 1 \leq i < N+1 & \text{ (or } i \in E). \end{aligned} \quad (9.26)$$

Proposition 9.11. Suppose that the given rates $(a_i, b_i, c_i : i \in E)$ satisfy (9.25). Define $\lambda_0(a_i, b_i, c_i)$ as in Proposition 9.1 without preassuming that $c_i \geq 0$ for all $i \in E$. Next, define $(\hat{a}_i, \hat{b}_i)$ and $(\check{a}_i, \check{b}_i)$ by (9.24).

- (1) If $\sum_{i=2}^N \mu_i b_{i-1}^{-1} = \infty$, then $\lambda_0(a_i, b_i, c_i) \geq \hat{\lambda}_0$, where $\hat{\lambda}_0$ is defined by (4.1) with rates $(\hat{a}_i, \hat{b}_i)$.
- (2) Otherwise, $\lambda_0(a_i, b_i, c_i) \geq \check{\lambda}_1$, where $\check{\lambda}_1$ is defined by (6.1) with rates $(\check{a}_i, \check{b}_i)$.
- (3) The equality sign of the conclusions in parts (1) and (2) holds provided it does in (9.25).

Proof. (a) As an application of Proposition 9.1, without loss of generality, we may and will assume that the equality sign for c_i in (9.26) holds. Then, we prove that the equality sign of the conclusions in parts (1) and (2) holds.

Clearly, we have

$$\hat{\mu}_1 = 1, \quad \hat{\mu}_1 \hat{a}_1 = \beta, \quad \hat{\mu}_i = \mu_i^{-1}, \quad \hat{\mu}_i \hat{a}_i = \frac{b_{i-1}}{\mu_i}, \quad 2 \leq i < N+1. \quad (9.27)$$

(b) Recall the operators:

$$\Omega f(i) = b_i(f_{i+1} - f_i) + a_i(f_{i-1} - f_i) - c_i f_i,$$

$$\widehat{\Omega} f(i) = \hat{b}_i(f_{i+1} - f_i) + \hat{a}_i(f_{i-1} - f_i), \quad f \in \mathcal{K}, f_0 = 0, f_{N+1} = 0 \text{ if } N < \infty.$$

Clearly, $\lambda_0(a_i, b_i, c_i)$ is the principal eigenvalue of Ω and the idea is describing it in terms of the first eigenvalue $\hat{\lambda}_{\min}$ of $\widehat{\Omega}$. Let U be the diagonal matrix with diagonal elements $(\mu_i : i \in E)$. Then U^{-1} is simply the diagonal matrix with diagonal elements $(\hat{\mu}_i : i \in E)$. For each function h with $h_0 = 0$ and $h_{N+1} = 0$ if $N < \infty$, by (9.27), (9.24) and (9.26), we have

$$\begin{aligned} (\Omega U^{-1} h)(i) &= b_i(\hat{\mu}_{i+1} h_{i+1} - \hat{\mu}_i h_i) + a_i(\hat{\mu}_{i-1} h_{i-1} - \hat{\mu}_i h_i) - c_i \hat{\mu}_i h_i \\ &= \hat{a}_{i+1}(\hat{\mu}_{i+1} h_{i+1} - \hat{\mu}_i h_i) + \hat{b}_{i-1}(\hat{\mu}_{i-1} h_{i-1} - \hat{\mu}_i h_i) \\ &\quad - (\hat{b}_i - \hat{b}_{i-1} - \hat{a}_{i+1} + \hat{a}_i) \hat{\mu}_i h_i \\ &= (\hat{a}_{i+1} \hat{\mu}_{i+1} h_{i+1} - \hat{b}_i \hat{\mu}_i h_i) + (\hat{b}_{i-1} \hat{\mu}_{i-1} h_{i-1} - \hat{a}_i \hat{\mu}_i h_i) \\ &= \hat{\mu}_i \hat{b}_i (h_{i+1} - h_i) + \hat{\mu}_i \hat{a}_i (h_{i-1} - h_i) \\ &= \hat{\mu}_i \widehat{\Omega} h(i) \\ &= (U^{-1} \widehat{\Omega} h)(i), \quad 2 \leq i < N. \end{aligned}$$

It is easy to check that the identity holds also for $i = 1$ and $i = N$, and then for all $i \in E$. Multiplying U from the left on the both sides, we obtain

$$U \Omega U^{-1} = \widehat{\Omega}. \quad (9.28)$$

Furthermore, we get

$$\langle f, \Omega g \rangle_\mu = \langle U^{-1} f, (U \Omega U^{-1}) U g \rangle_\mu = \langle U f, (U \Omega U^{-1}) U g \rangle_{\hat{\mu}} = \langle \hat{f}, \widehat{\Omega} \hat{g} \rangle_{\hat{\mu}}$$

for all $f, g \in \mathcal{K}$, where the mapping $f \rightarrow \hat{f} := U f$ is an isometry from $L^2(\mu)$ to $L^2(\hat{\mu})$. Since $f \in \mathcal{K}$ iff $\hat{f} \in \mathcal{K}$, it follows that the operators Ω and $\widehat{\Omega}$ with the same core $\mathcal{K}$ are isospectral. In particular, $\lambda_0(a_i, b_i, c_i) = \hat{\lambda}_{\min}$.

(c) For assertion (1), since $\sum_i (\hat{\mu}_i \hat{b}_i)^{-1} = \infty$ by assumption, it follows that $N = \infty$ and the Dirichlet form corresponding to $\widehat{\Omega}$ is regular by Proposition 1.3. Hence, the minimal and the maximal domains of the Dirichlet form are coincided. Therefore, $\hat{\lambda}_{\min}$ is equal to $\lambda_0^{(4.1)}$ replacing the original rates (a_i, b_i) by $(\hat{a}_i, \hat{b}_i)$.

For assertion (2), since $\hat{a}_1 > 0$ and $\hat{b}_N > 0$, we come to the setup of Section 7: $\hat{\lambda}_{\min} = \lambda_0^{(7.1)}$ with (a_i, b_i) replaced by $(\hat{a}_i, \hat{b}_i)$. Next, because of $\sum_i (\hat{\mu}_i \hat{b}_i)^{-1} < \infty$, by Theorem 7.1, it turns out $\hat{\lambda}_{\min} = \check{\lambda}_1$ in terms of the dual rates $(\check{a}_i, \check{b}_i)$ of $(\hat{a}_i, \hat{b}_i)$. $\square$

We now summarize our main qualitative result about λ_0 . The three parts given below are obtained by Corollary 9.9, Proposition 9.1 plus Proposition 1.3, and Corollary 9.5, respectively.

Summary 9.12. We have $\lambda_0 > 0$ whenever $N < \infty$. Next, let $N = \infty$. Then

- (1) $\lambda_0 > 0$ if $\varliminf_{n \rightarrow \infty} c_n > 0$ (in particular, if $\inf_{i \in E} c_i > 0$).
- (2) $\lambda_0 = \lambda_0(a_i, b_i, c_i) > 0$ if $\lambda_0(a_i, b_i, c_i \mathbb{1}_{\{1\}}) > 0$ which can be checked case by case by
 - (i) Theorem 3.1 when $c_1 = 0$ and

$$\sum_{i \geq 1} \frac{1}{\mu_i a_i} < \infty; \quad (9.29)$$

- (ii) Theorems 7.1 and 6.2 when $c_1 > 0$ and (9.29) holds;
 - (iii) Theorem 4.2 when $c_1 > 0$ but (9.29) fails.
- (3) $\lambda_0 = 0$ if

$$\varliminf_{m \rightarrow \infty} \sum_{i=1}^m \mu_i c_i / \sum_{k=1}^m \mu_k = 0, \quad \varliminf_{m \rightarrow \infty} \mu_m b_m \left(\sum_{i=1}^m \mu_i \right)^{-1} = 0 \quad \text{and} \quad \sum_i \mu_i = \infty.$$

Open problem 9.13 (Explicit criterion for $\lambda_0 > 0$). As will be seen soon in Example 9.18 below, for $\lambda_0 > 0$, it can happen that $\varliminf_{n \rightarrow \infty} c_n = 0$. Hence, the simple condition “ $\varliminf_{n \rightarrow \infty} c_n > 0$ ” in part (1) is sufficient only but not necessary. Naturally, this condition becomes necessary for the first one in part (3) for which a sufficient condition is $\lim_{n \rightarrow \infty} c_n = 0$. Thus, it is more or less satisfactory whenever (c_n) has a limit. Otherwise, there is a gap. In contrast with the first condition in part (3), condition (9.23) is sufficient for $\xi_\varepsilon > 0$ but there is still a distance to deduce the positivity of λ_0 in view of Corollary 9.9.

Next, since we are dealing with the minimal Dirichlet form, a general criterion for Hardy-type inequalities (cf. [12; Theorems 7.1 and 7.2]) which was successfully used in Section 8, is also available in the present situation, hence, there is already a criterion for $\lambda_0 > 0$ in terms of capacity which is unfortunately not explicit. More seriously, the technique to produce an explicit result used in [12; pages 134–136] does not work at the beginning (replacing a finite number of disjointed finite intervals $\{K_i\}$ by the connected one $[\min \cup_i K_i, \max \cup_i K_i]$) in the present setup. Thus, it is still an unsolved problem to figure out an explicit criterion for $\lambda_0 > 0$ in the present setup.

It is our position to illustrate by examples the application of the results obtained in this section. First, by using Proposition 9.11 and (9.26), it is easy to transfer the examples given in Sections 3 and 6 to the present context. However, most of the resulting killing rates are rather simple. We are now going to construct some new examples, all of them are out of the scope of Proposition 9.11. In the most cases, we use simple (a_i, b_i) and pay more attention on (c_i) . Let us begin with the following simplest case.

Example 9.14. Let

$$Q = \begin{pmatrix} -b_1 - c_1 & b_1 \\ a_2 & -a_2 - c_2 \end{pmatrix}.$$

Then as in Examples 7.5 (2), we have

$$\lambda_0 = \frac{1}{2} \left(a_2 + b_1 + c_1 + c_2 - \sqrt{(a_2 + c_2 - b_1 - c_1)^2 + 4a_2b_1} \right),$$

with eigenvector

$$g = \left(\frac{1}{2a_2} \left[a_2 + c_2 - b_1 - c_1 + \sqrt{(a_2 + c_2 - b_1 - c_1)^2 + 4a_2b_1} \right], 1 \right).$$

Even in such a simple case, the role for λ_0 played by the parameters a_i , b_i , and c_i is ambiguous. For instance, since $\tilde{c}_1 - \tilde{c}_2 = c_1 - c_2$ ($\tilde{c}_k := c_k - c_1 \wedge c_2$), one can separate out the constant $c_1 \wedge c_2$ from the above expression of λ_0 . However, this obvious separation property becomes completely mazed for the next example having three states only.

Example 9.15. Let

$$Q = \begin{pmatrix} -b_1 - c_1 & b_1 & 0 \\ a_2 & -a_2 - b_2 - c_2 & b_2 \\ 0 & a_3 & -a_3 - c_3 \end{pmatrix}.$$

Then

$$\lambda_0 = -\frac{1}{3}\gamma_1 + 2\sqrt{\frac{-U}{3}} \cos \left[\frac{1}{3} \arccos \left(-\frac{V}{2} \left(\frac{-U}{3} \right)^{-3/2} \right) + \frac{2\pi}{3} \right], \quad (9.30)$$

with eigenvector

$$g = \left\{ \frac{b_1(a_3 + c_3 - \lambda_0)}{a_3(b_1 + c_1 - \lambda_0)}, 1 + \frac{c_3 - \lambda_0}{a_3}, 1 \right\},$$

where

$$U = \gamma_2 - \gamma_1^2/3, \quad V = \gamma_3 - \gamma_1\gamma_2/3 + 2(\gamma_1/3)^3,$$

and $\lambda^3 + \gamma_1\lambda^2 + \gamma_2\lambda + \gamma_3$ the eigenpolynomial of $-Q$ with coefficients:

$$\gamma_1 = -a_2 - a_3 - b_1 - b_2 - c_1 - c_2 - c_3,$$

$$\gamma_2 = a_2a_3 + b_1a_3 + c_1a_3 + c_2a_3 + b_1b_2 + a_2c_1 + b_2c_1 + b_1c_2 + c_1c_2 + a_2c_3 \\ + b_1c_3 + b_2c_3 + c_1c_3 + c_2c_3,$$

$$\gamma_3 = -a_2a_3c_1 - a_3b_1c_2 - a_3c_1c_2 - b_1b_2c_3 - a_2c_1c_3 - b_2c_1c_3 - b_1c_2c_3 - c_1c_2c_3.$$

Proof. Since the eigenvalues of $-Q$ are all real, it is easier to write them down. By using the notation given above, the eigenvalues of $-Q$ can be expressed as

$$-\frac{1}{3}\gamma_1 + 2\sqrt{\frac{-U}{3}} \cos \left[\frac{1}{3} \arccos \left(-\frac{V}{2} \left(\frac{-U}{3} \right)^{-3/2} \right) + \frac{2k\pi}{3} \right], \quad k = 0, 1, 2.$$

Among them, the minimal one is λ_0 given in (9.30). Clearly, the solution is indeed rather complicated in view of the coefficients of the eigenpolynomial. $\square$

To see the role played by the killing rate (c_i) , in the following examples, we restrict ourselves to the case that $\lambda_0(a_i, b_i, 0) = 0$ (and then $N = \infty$). The examples are arranged according the increasing order of the polynomial rates (a_i) and (b_i) . Actually, all the examples in the paper are either standard or constructed by using simple rates and simple eigenfunctions. They are used first as a guidance of the study and then to justify the power of the theoretic results.

In contrast to the explosive case (cf. Theorem 3.1), λ_0 can still be zero for the process having positive killing rate, as shown by the following example.

Example 9.16. Let $b_1 = 1$, $a_i = b_i = 1$ for $i \geq 2$, and let (c_i) satisfy $\lim_{n \rightarrow \infty} c_n = 0$. Then we have $\lambda_0 = 0$, even though c_i can be very large locally.

Proof. Apply Corollary 9.5. $\square$

Example 9.17. Let $a_i = b_i = 1$ for $i \geq 2$ and $c_i = \beta^{-1}(\beta - 1)^2$ ($\beta > 0$) for $i \geq 1$. Then for every $a_1 \geq 0$ and $b_1 > 0$, we have $\lambda_0 = \beta^{-1}(\beta - 1)^2$.

Proof. Since $\lambda_0(a_i, b_i, 0) = 0$ and (c_i) is a constant, this is a consequence of part (3) of Proposition 9.1.

Note that the lower estimates given by Proposition 9.1, Theorem 9.3, and (9.4) are all sharp for this example. To see this, simply choose $v_i \equiv 1$ in (9.2) and (9.4). We now consider a more specific situation: $a_1 = 0$, $b_1 = 1 - \beta$ and $\beta \in (0, 1)$. If we set $\bar{v}_i \equiv \beta^{-1}$, then it is easy to check that $R_i(\bar{v}) \equiv 0$ and so $\inf_{v>0} \sup_{i \geq 1} R_i(v) = 0$. This shows that the truncating procedure used in Theorem 9.3 for the upper estimate is necessary in the case that the function f defined by (9.5) does not belong to $L^2(\mu)$, even though $R_i(v)$ is a constant. In the present case, $\bar{v}_i > 1$ for all i and so the corresponding function f is strictly increasing. Since μ_i is a constant for $i \geq 2$, it is clear that $\sum_i \mu_i = \infty$ and then $f \notin L^2(\mu)$. $\square$

Example 9.18. Let $a_1 = 0$, $b_1 = 5/2$, $a_i = 2$ and $b_i = 1$ for $i \geq 2$, $c_i = 0$ for odd i and $c_i = 13/6$ for even i . Then $\lambda_0 = 5/6$. The upper bound provided by (9.9) is approximately 1.03. For the lower estimate, Proposition 9.11 is available but not Corollary 9.9.

Proof. (a) Let $v_i \equiv 1 + (-1)^i/3$. Then it is easy to check that $R_i(v) \equiv 5/6$. Next, define

$$g_1 = 1, \quad g_n = \prod_{k=1}^{n-1} v_k, \quad n \geq 2. \quad (9.31)$$

We claim that $g \in L^2(\mu)$ by using Kummer's test. To do so, note that to study the convergence/divergence of the series $\sum_n \mu_n g_n^2$, the constant κ defined by (3.13) takes a simpler form as follows:

$$\kappa = \lim_{n \rightarrow \infty} n \left(\frac{a_{n+1}}{b_n v_n^2} - 1 \right). \quad (9.32)$$

Now, because $g \in L^2(\mu)$ and

$$-\Omega g/g = R(v) = 5/6,$$

we have $\lambda_0 = 5/6$ by Theorem 9.3. Clearly, this eigenfunction g of λ_0 is not monotone since $g_{i+1}/g_i = v_i = 2/3$ for odd i and $= 4/3$ for even i .

(b) Next, we study the upper estimates of λ_0 . First, we have

$$\mu_1 = 1, \quad \mu_i = \frac{5}{2^i}, \quad i \geq 2; \quad \mu_i b_i = \frac{5}{2^i}, \quad i \geq 1.$$

The upper bound provided by (9.9) is approximately 1.03.

(c) For a lower estimate, we apply Proposition 9.11 (2). The modified rates are as follows: $\check{a}_i \equiv 2$ ($i \geq 1$), $\check{b}_1 = 5/2$, and $\check{b}_i \equiv 1$ ($i \geq 2$). However, (c_i) does not satisfy (9.25) at $i = 2$. We now replace (c_i) by $(\tilde{c}_i := c_i + 1/6)$ and choose $\check{b}_0 = \beta = 2/3$. Then $(\tilde{c}_i)$ satisfy (9.25). With the modified $(\tilde{c}_i)$, we are in the ergodic case, and moreover, $\check{\lambda}_1 = (\sqrt{2} - 1)^2$ with eigenfunction $\check{g}$: $\check{g}_0 = -1$ and

$$\check{g}_i = \frac{1}{20} 2^{i/2} \left[-101 + 60\sqrt{2} + (41 - 25\sqrt{2})i \right], \quad i \geq 1.$$

Therefore, by Proposition 9.11 (2), we obtain $\lambda_0(a_i, b_i, \tilde{c}_i) \geq (\sqrt{2} - 1)^2$. Returning to the original (c_i) by Proposition 9.1 (2), we get a rough lower bound as follows:

$$\lambda_0 = \lambda_0(a_i, b_i, \hat{c}_i) - \frac{1}{6} \geq \frac{17}{6} - 2\sqrt{2} \approx 0.005.$$

Before moving further, let us remark that if only $\check{b}_0$ is changed from $2/3$ to $1/2$, then for the $(\check{a}_i, \check{b}_i)$ -process, we still have $\check{\lambda}_1 = (\sqrt{2} - 1)^2$ with a similar eigenfunction $\check{g}$: $\check{g}_0 = -1$ and

$$\check{g}_i = \frac{1}{10} 2^{i/2} \left[-67 + 42\sqrt{2} + (27 - 17\sqrt{2})i \right], \quad i \geq 1.$$

Now, as an application of Proposition 9.11 (2) with the original (c_i) replacing $c_2 = 4/3$ by $c_2 = 3/2$ only, the resulting λ_0 has a lower bound $(\sqrt{2} - 1)^2 \approx 0.17$.

(d) To apply Corollary 9.9, we write $c_i = 13(1 + (-1)^i)/12$ and use (9.20) and (9.21):

$$\xi_\varepsilon = \inf_{i \geq 2} \frac{1 - \varepsilon + \varepsilon x_i}{z_i + \varepsilon y_i}, \quad \zeta_\varepsilon = \inf_{\text{odd } i \geq 3} \frac{\varepsilon \eta}{1 + \varepsilon x_i - \eta(z_i + \varepsilon y_i)}, \quad \eta \in [0, \xi_\varepsilon], \quad (9.33)$$

$$x_i = \frac{13}{72} (2^{2+i} - 6i - 3 - (-1)^i), \quad y_i = 2^{i-1} - i, \quad z_i = \frac{2^i - 2}{5}.$$

Note that the numerator of ζ_ε given in (9.33) is independent of i but in the denominator, x_i , y_i and z_i all tend to infinity as $i \rightarrow \infty$. To avoid the trivial estimate, one needs to cancel the leading term in i of $\varepsilon x_i - \eta(z_i + \varepsilon y_i)$ in the denominator. This leads to the following solution:

$$\eta = \frac{65\varepsilon}{9(2 + 5\varepsilon)}.$$

Inserting this into $\varepsilon x_i - \eta(z_i + \varepsilon y_i)$, it follows that the denominator of ζ_ε in (9.33) becomes

$$1 - \frac{65(-2i + (-1)^i + 3)\varepsilon^2 + 2(78i + 13(-1)^i - 245)\varepsilon - 144}{72(5\varepsilon + 2)}.$$

Now, in order to remove the leading term in i , the only solution is

$$\varepsilon = 78/65 > 1,$$

which does not belong to the domain of $\varepsilon \in (0, 1)$. Therefore, the test function used in Corollary 9.9 does not provide enough freedom to cover this example.

Note that without the killing rate, the process with rates (a_i) and (b_i) is exponentially ergodic and so $\lambda_0(a_i, b_i, 0) = 0$. $\square$

For the following examples, we assume that $a_i = b_i$ for $i \geq 2$. Then

$$\mu_1 = 1, \quad \mu_i = b_1 a_i^{-1}, \quad i \geq 2; \quad \mu_i b_i = b_1, \quad i \geq 1.$$

The quantities ξ_ε and ζ_ε defined in (9.20) and (9.21), respectively, are now determined by

$$x_i = \sum_{2 \leq j \leq i-1} \frac{i-j}{a_j} \tilde{c}_j, \quad y_i = \sum_{2 \leq j \leq i-1} \frac{i-j}{a_j}, \quad z_i = \frac{i-1}{b_1}.$$

Example 9.19. Let $a_1 = 0$, $b_1 = 2\beta(1-\beta)(1-2\beta)^{-1}$ ($\beta \in (0, 1/2)$), $a_i = b_i = \beta i$ for $i \geq 2$, $c_i = (1-\beta)^2(i-1)$ for $i \geq 1$. Then $\lambda_0 = 2\beta(1-\beta)$. In the special case that $\beta = 1/4$, we have $\lambda_0 = 3/8$. The upper and lower bounds provided by (9.9) and Corollary 9.9 are $3/4$ and approximately 0.274 , respectively.

Proof. Let $v_i = \beta(1+i^{-1})$ for $i \geq 1$. Then $R_i(v) \equiv 2\beta(1-\beta)$. By Kummer's test (cf. (9.32)), the corresponding function g defined by (9.31) belongs to $L^2(\mu)$. Hence, the assertion follows from Theorem 9.3. Note that $v_i < 1$ for all i , and g is strictly decreasing even though $c_1 = 0 < \lambda_0$ and $c_i > \lambda_0$ for all $i > (1+\beta)(1-\beta)^{-1}$ (compare with (2.5)).

As an application of (9.9) with $(\ell, m) = (1, 1)$ or $(3, 4)$, we obtain

$$\lambda_0 \leq (1-\beta) \left\{ \frac{2\beta}{1-2\beta} \bigwedge \frac{23-40\beta+23\beta^2}{2(8-11\beta)} \right\}.$$

To study the lower bound, for simplicity, we let $\beta = 1/4$. Then $\lambda_0 = 3/8$ and the upper bound in the last formula is $3/4$. Choose $\varepsilon = (\sqrt{409} - 5)/24$ so that the infimum $\xi_\varepsilon = (29 - \sqrt{409})/32 \approx 0.274$ is attained simultaneously at $i = 2$ and $i = 3$. Since $\xi_\varepsilon < c_2$, the set $\{i \geq 2 : c_i < \xi_\varepsilon\}$ is empty. Therefore, the lower bound provided by Corollary 9.9 is approximately 0.274 . $\square$

For the following two examples, without the killing rate, the process is exponentially ergodic and so $\lambda_0(a_i, b_i, 0) = 0$.

Example 9.20. Let $a_1 = 0$, $b_1 = 4/5$, $a_i = b_i = i^2$ for $i \geq 2$, and

$$c_i = \frac{8}{9} \left[\frac{8}{3i-8} - \frac{2}{3i-4} + 5 \right], \quad i \geq 1.$$

Then $\lambda_0 = 4$. The upper and lower bounds provided by (9.9) and Corollary 9.9 are $14/3$ and approximately 2.82, respectively.

Proof. The proof is similar as before using

$$v_i = 1 - \frac{1}{3i-4}, \quad i \geq 1.$$

Note that c_i has minimum 0 at $i = 2$. The upper bound provided by (9.9) with $(\ell, m) = (2, 2)$ is $14/3$. The lower bound produced by Corollary 9.9 with $\varepsilon = 1$ is $48/17 \approx 2.82$. Since $c_1 > 0$, the parameter $\varepsilon = 1$ is allowed. Then $\xi_\varepsilon = 48/11$ is attained at $i = 3$, and $\zeta_\varepsilon = 48/17$ is attained at $i = 2$ (noting that the set $\{i \geq 2 : c_i < \xi_\varepsilon\}$ is a singleton $\{2\}$). $\square$

Example 9.21. Let $a_1 = 0$, $b_1 = 3/2$, $c_1 = 15$,

$$\begin{aligned} a_i &= b_i = i(i-4^{-1})(12i^2 - 31i + 27), \quad i \geq 2, \\ c_i &= i^4 - \frac{1}{2}i^3 - \frac{301}{16}i + \frac{227}{8}, \quad i \geq 2. \end{aligned}$$

Then $\lambda_0 = 119/8 = 14.875$. The upper and lower bounds provided by (9.9) and Corollary 9.9 are approximately 15.42 and 13.18, respectively.

Proof. Note that c_i is convex and has its minimum 0 at $i = 2$. For

$$v_i = \frac{3}{4} - \frac{2}{i} + \frac{7}{4i-1}, \quad i \geq 1,$$

we have $R_i(v) \equiv 119/8$. Note that $v_1 > 1$ and $v_i < 1$ for all $i \geq 2$. The function g defined by (9.31) is not monotone but is bounded. Next, since $\mu_i \sim i^{-4}$, we have $g \in L^2(\mu)$. The assertion now follows from Theorem 9.3.

Clearly, $\inf_{i \geq 1} c_i = 11/4$. The upper bound provided by (9.9) with $(\ell, m) = (2, 4)$ is approximately 15.42. To get a lower estimate, we apply Corollary 9.9. Because $\tilde{c}_1 > 0$, we can choose $\varepsilon = 1$. Then $\xi_\varepsilon = 354679/29504$ is attained at $i = 4$. Next, since the set $\{i \geq 2 : c_i < \xi_\varepsilon\}$ is a singleton $\{2\}$, we need only to compute ζ_ε at $i = 2$: $\zeta_\varepsilon \approx 10.43$. Thus, the lower bound produced by Corollary 9.9 is approximately 13.18. $\square$

To conclude this section, we return to the uniqueness problem for birth–death processes with killing of the Dirichlet form as discussed at the end of Section 1. Certainly, the problem is meaningful only if $N = \infty$. Recall that for a given Q -matrix, not necessarily conservative (i.e., may have killing), the exit space $\mathcal{U}_\lambda$ is the set of the solutions (u_i) to the following equation:

$$\begin{cases} (\lambda I - Q)u = 0, \\ 0 \leq u \leq 1, \end{cases} \quad \lambda > 0.$$

Note that the dimension of $\mathcal{U}_\lambda$ is independent of $\lambda > 0$. By (2.5) replacing λ with $-\lambda$, it follows that the non-trivial exit solution, if it exists, is unique and is strictly increasing.

Theorem 9.22 (Uniqueness of the Dirichlet form). Let $N = \infty$.

- (1) The Dirichlet form satisfying the Kolmogorov's equations is unique if $\mathcal{U}_\lambda = \{0\}$. Equivalently,

$$\sum_{n=1}^{\infty} \frac{1}{\mu_n b_n} \sum_{k=1}^n \mu_k (1 + c_k) = \infty. \quad (9.34)$$

- (2) Let $\sum_{i \in E} \mu_i < \infty$. Then the Dirichlet form is unique iff

$$\sum_{i \in E} \mu_i c_i < \infty \quad \text{and} \quad \sum_{i \in E} \frac{1}{\mu_i b_i} = \infty.$$

- (3) Let $\sum_{i \in E} \mu_i = \infty$. Then the Dirichlet form is unique if

$$\text{either } \inf_{i \in \dot{E}} \sum_{j \in E} P_{ij}^{\min}(\lambda) > 0 \quad \text{or} \quad \sum_{i \in E} \mu_i c_i < \infty$$

holds, where $\dot{E} = \{i \in E : c_i > 0\}$.

Here are some comments about the theorem.

- (i) Suppose that only a finite number of c_i are non-zero.

Then condition (9.34) is equivalent to (1.2) [Certainly in this item, we are using the modified (1.2) and (1.3) by removing 0 from the state space]:

$$\sum_{n=1}^{\infty} \frac{1}{\mu_n b_n} \sum_{k=1}^n \mu_k \leq \sum_{n=1}^{\infty} \frac{1}{\mu_n b_n} \sum_{k=1}^n \mu_k (1 + c_k) \leq C \sum_{n=1}^{\infty} \frac{1}{\mu_n b_n} \sum_{k=1}^n \mu_k,$$

where $C = \max_{i: c_i > 0} (1 + c_i) < \infty$. Hence, condition (9.34) is stronger than (1.3). In this case, condition (9.34) is even not needed in part (3) where the last two conditions are automatic.

When $\sum_i \mu_i < \infty$, (1.3) is equivalent to (1.2) which coincides with (9.34). When $\sum_i \mu_i = \infty$, both (1.3) and part (3) hold. Therefore, if only a finite number of c_i are non-zero, then we have

criterion (1.3) $\iff$ one of parts (2) and (3) holds.

- (ii) When $c_i \neq 0$ for infinite number of i , except condition (9.34), an additional condition on the killing rates (c_i) is required. The condition means that if c_i increases very fast, then there exist some Dirichlet forms that do not satisfy the Kolmogorov equations.
- (iii) The second condition in part (3) is the same as the one in part (2). For the first condition in part (3), it is easy to write down some more explicit sufficient conditions. This is due to the following fact. Since for each fixed j , $\{P_{ij}^{\min}(\lambda) : i \in E\}$ is the minimal solution to the equations

$$x_i = \sum_{k \neq i} \frac{q_{ik}}{\lambda + q_i} x_k + \frac{\delta_{ij}}{\lambda + q_i}, \quad i \in E,$$

by the linear combination theorem, $\{\sum_{j \in E} P_{ij}^{\min}(\lambda) : i \in E\}$ is the minimal solution to the equations

$$x_i = \sum_{k \neq i} \frac{q_{ik}}{\lambda + q_i} x_k + \frac{1}{\lambda + q_i}, \quad i \in E.$$

This minimal solution (x_i^*) can be obtained in the following way. Let

$$\begin{aligned} x_i^{(1)} &= \frac{1}{\lambda + q_i}, \quad i \in E, \\ x_i^{(n+1)} &= \sum_{k \neq i} \frac{q_{ik}}{\lambda + q_i} x_k^{(n)} + \frac{1}{\lambda + q_i}, \quad i \in E, \quad n \geq 1. \end{aligned}$$

Then $x_i^{(n)} \uparrow x_i^*$ as $n \rightarrow \infty$ for every $i \in E$ (cf. [10; §2.1 and §2.2]). Hence, for each $n \geq 1$, $x_i^{(n)}$ is a lower bound of $\sum_{j \in E} P_{ij}^{\min}(\lambda)$. From this discussion, it is clear that the first condition in part (3) is also a restriction on the growing of the killing rates (c_i) . This is consistent with the second condition there. It is regretted that we do not know at the moment whether the conditions in part (3) are necessary or not.

Proof of Theorem 9.22. Part (1) follows from [10; Theorem 3.2] and Chen et al. (2005)[1] with a fictitious state 0. The last cited result is an application of the single birth processes. Noting that if $\sum_{i \in E} \mu_i < \infty$ and $\sum_{i \in E} \mu_i c_i < \infty$, then (9.34) holds iff $\sum_{i \in E} (\mu_i b_i)^{-1} = \infty$, hence, part (2) is a special case of [10; Theorem 6.42]. Next, noting that the unique exit solution is strictly increasing, when $\sum_i \mu_i = \infty$, we have $\mathcal{U}_\lambda \cap L^1(\mu) = \{0\}$. Hence, part (3) is a particular application of [10; Theorem 6.41]. $\square$

10. NOTES

10.1 Open problems and basic estimates for diffusions.

Having seen such a long paper, the reader may feel strange if we claim that the story is still incomplete even in the context of birth–death processes. Unfortunately, it is the case.

All of the examples we have done so far show that the following facts hold.

- (1) The ratio of the improved upper and lower bounds belongs to $[1, 2]$.
- (2) The sequence $\{\bar{\delta}_n\}$ is increasing in n and $\bar{\delta}_n \geq \delta'_n$ for all n .
- (3) The sequences $\{\bar{\delta}_n\}$, $\{\delta'_n\}$, and $\{\delta_n\}$ all converge to λ_0^{-1} as $n \rightarrow \infty$.
- (4) The relation $(\bar{\eta}_1, \eta_1) \subset (\kappa, 4\kappa)$ discussed in Section 6 holds.

However, there is still no analytic proof for them. The difficulty for the first question is that the maximum/minimum of $\bar{\delta}_1$ and δ_1 may locate in different places. In the case that (2) would be true, then the story could be simplified since we need the first sequence only. For Questions (2) and (3), the assertions are numerically justified for almost all of the examples in the paper but the results are not included. We have not worked on Question (3) hardly enough since one can go ahead only in a finite number of steps in the symbol computation but

the question is certainly meaningful and in the numerical computation, only in a few steps one achieves the eigenvalue. For the sequences $\{\bar{\eta}_n\}$ and $\{\eta_n\}$, we have similar questions as (1) and (3) about, but the corresponding question (2) is answered by Lemma 6.5.

There is a parallel story for the one-dimensional diffusions. In many cases, one can easily guess what the result should be, even though there may exist a new difficulty in its proof. For instance, as a combination of the proofs of Theorem 8.2 and [12; Corollary 7.6], one may prove the following result.

Theorem 10.1. Consider the minimal diffusion on $(-M, N)$ ($M, N \leq \infty$) with operator

$$L = a(x) \frac{d^2}{dx^2} + b(x) \frac{d}{dx} \quad \left(a(x) > 0, \frac{b(x)}{a(x)} \text{ is locally integrable} \right),$$

and Dirichlet boundaries at $-M$ if $M < \infty$, and at N if $N < \infty$. Let $C(x) = \int_{\theta}^x b/a$ for some fixed reference point $\theta \in (-M, N)$ and assume additionally that e^C/a is also locally integrable. Denote by $A_{\mathbb{B}}$ the optimal constant in Poincaré-type inequality (8.1) with Dirichlet form

$$D(f) = \int_{-M}^N f'^2 e^C, \quad f \in \mathcal{C}_0^\infty(-M, N),$$

Then $A_{\mathbb{B}}$ satisfies $B_{\mathbb{B}} \leq A_{\mathbb{B}} \leq 4B_{\mathbb{B}}$, where

$$B_{\mathbb{B}}^{-1} = \inf_{-M < x < y < N} \left[\left(\int_{-M}^x e^{-C} \right)^{-1} + \left(\int_y^N e^{-C} \right)^{-1} \right] \|\mathbb{1}_{(x,y)}\|_{\mathbb{B}}^{-1}. \quad (10.1)$$

By the way, we prove a dual result of Theorem 10.1 for ergodic diffusions. As discussed in the proof of Theorem 7.5, the exponentially ergodic rate often coincides with the first non-trivial eigenvalue λ_1 defined below. Consider a diffusion process with operator L as in Theorem 10.1, with state space $(-M, N)$ ($M, N \leq \infty$) and reflecting boundaries at $-M$ if $M < \infty$, and at N if $N < \infty$. For convenience, we define two measures as follows:

$$\text{Scale measure: } \nu(dx) = e^{-C(x)} dx, \quad C(x) := \int_{\theta}^x \frac{b}{a},$$

where $\theta \in (-M, N)$ is a fixed reference point.

$$\text{Speed measure: } \mu(dx) = \frac{e^{C(x)}}{a(x)} dx.$$

With these measures, the operator L takes a compact form:

$$L = \frac{d}{d\mu} \frac{d}{d\nu}. \quad (10.2)$$

Next, suppose that $\mu(-M, N) < \infty$, and denote by π the normalized probability measure of μ . Set

$$\mathcal{A} = \{f : f \text{ is absolutely continuous in } (-M, N)\},$$

and define

$$\lambda_1 = \inf\{D(f) : f \in L^2(\mu) \cap \mathcal{A}, \pi(f) = 0, \|f\| = 1\},$$

where

$$D(f) = \int_{-M}^N a f'^2 d\mu, \quad f \in \mathcal{A}.$$

Clearly, in the definition of λ_1 , only those f in the set $\{f \in L^2(\mu) \cap \mathcal{A} : D(f) < \infty\}$ are useful. In other words, we are here using the maximal Dirichlet form, as in Section 6.

Theorem 10.2. Let $a > 0$, a and b be continuous on $[-M, N]$ (or $(-M, N]$ if $M = \infty$, for instance). Assume that $\mu(-M, N) < \infty$. Then for λ_1 , we have the basic estimate: $\kappa^{-1}/4 \leq \lambda_1 \leq \kappa^{-1}$, where

$$\kappa^{-1} = \inf_{-M < x < y < N} \left[\left(\int_{-M}^x d\mu \right)^{-1} + \left(\int_y^N d\mu \right)^{-1} \right] \left(\int_x^y d\nu \right)^{-1}. \quad (10.3)$$

Proof. (a) First we show that for the basic estimate, it suffices to consider the finite M and N with smooth a and b . Since a and b are continuous, if $M = N = \infty$ for instance, we may choose $M_p, N_p \uparrow \infty$ as $p \rightarrow \infty$ such that $\theta \in (-M_p, N_p)$ for all p . Then, by Chen and Wang (1997, Lemma 5.1), we have $\lambda_1^{(M_p, N_p)} \downarrow \lambda_1$ as $p \rightarrow \infty$ (This is parallel to the localizing procedure used in Section 6). At the same time, the isoperimetric constants $\kappa^{(M_p)} \downarrow \kappa$ as $p \rightarrow \infty$ (cf. proof of Corollary 7.9). Hence, in what follows, we may assume that $M, N < \infty$. Next, by using the continuity of a and b again, and using a standard smoothing procedure, we can choose smooth a_p and b_p such that $a_p \rightarrow a$ and $b_p \rightarrow b$ (as $p \rightarrow \infty$) uniformly on finite intervals, and furthermore, we can assume that $a_p > 0$ on each fixed closed finite interval. Clearly, the corresponding κ_p converges to κ as $p \rightarrow \infty$. Therefore, without loss of generality, we assume, unless otherwise stated, that not only $M, N < \infty$ but also a and b are smooth with $a > 0$ on $[-M, N]$.

(b) Recall the following differential form of variational formula for λ_1 :

$$\lambda_1 = \sup_{f \in \mathcal{F}} \inf_{x \in (-M, N)} \left[-b' - \frac{af'' + (a' + b)f'}{f} \right](x), \quad (10.4)$$

where

$$\mathcal{F} = \{f \in \mathcal{C}^1[-M, N] \cap \mathcal{C}^2(-M, N) : f(-M) = f(N) = 0, f|_{(-M, N)} > 0\}.$$

This is an analog of the variational formula for the lower estimate in Theorem 6.1(1). In the original study by Chen and Wang (1997, (2.3)), the state

space is the half-line, not finite, but this is not essential. It works also for finite state spaces. Besides, it was stated as “ $\geq$ ” in (10.4) only. For “ $=$ ”, one simply chooses $f = g'$, where g is the eigenfunction of λ_1 . This gives us the boundary condition: $f(-M) = f(N) = 0$ since $g'(-M) = g'(N) = 0$ by assumption. Here, one requires that $g \in \mathcal{C}^3(-M, N)$ which is satisfied since we are now in a finite interval having smooth a and b . Alternatively, instead of the original coupling proof, one may use the analytic one which leads to (6.4) for birth-death processes.

(c) We are now going to handle with a more general situation: $M, N \leq \infty$, $a, b \in \mathcal{C}^1(-M, N)$ and $a > 0$ on $(-M, N)$. Let us define a dual operator $\hat{L}$ of L . In view of the Karlin and McGregor's construction, the dual of a birth-death process is simply an exchange of the scale and speed measures $\hat{\mu} = \nu$ and $\hat{\nu} = \mu$ up to a constant (cf. (5.3)). Thus, in view of (10.2), the dual operator $\hat{L}$, as was introduced by Cox and Rösler (1983)¹, should be given by

$$\hat{L} = \frac{d}{d\hat{\mu}} \frac{d}{d\hat{\nu}}. \quad (10.5)$$

Again, the speed and scale measures $\hat{\mu}$ and $\hat{\nu}$ of $\hat{L}$ should be expressed as

$$d\hat{\mu} = \frac{e^{\hat{C}}}{\hat{a}} dx, \quad d\hat{\nu} = e^{-\hat{C}} dx$$

in terms of the coefficients $\hat{a}$ and $\hat{b}$ of $\hat{L}$ to be determined now. Because $\hat{\mu} = \nu$ and $\hat{\nu} = \mu$, we have

$$\frac{d\hat{\mu}}{dx} \frac{d\hat{\nu}}{dx} = \frac{d\nu}{dx} \frac{d\mu}{dx}.$$

It follows that $\hat{a} = a$. Then using the equation $\hat{\mu} = \nu$, we get

$$\hat{C} = -C + \log \hat{a} = -C + \log a.$$

Thus, from

$$\frac{\hat{b}}{\hat{a}} = \hat{C}' = -\frac{b}{a} + \frac{a'}{a},$$

we get $\hat{b} = a' - b$. Therefore, the dual operator $\hat{L}$ has the following expression:

$$\hat{L} = a(x) \frac{d^2}{dx^2} + \left(\frac{d}{dx} a(x) - b(x) \right) \frac{d}{dx}. \quad (10.6)$$

For the dual process, the Dirichlet boundary is endowed at $-M$ and N (cf. proof (e) below). Clearly, the dual operator $\hat{L}$ is symmetric on $L^2(e^{-C} dx)$. We remark that the assumption on a and b can be weakened in this paragraph.

¹Added in proof (2011-5-16). This dual appeared earlier in “The equivalence of absorbing and reflecting barrier problems for stochastically monotone Markov processes” by D. Siegmund (Ann. Prob. 4:6 (1976), 914-924).

(d) Define a Schrödinger operator as follows:

$$\begin{aligned} L_S &= a(x) \frac{d^2}{dx^2} + (a'(x) + b(x)) \frac{d}{dx} + b'(x) \\ &= \frac{d}{dx} \left(a(x) \frac{d}{dx} \right) + b(x) \frac{d}{dx} + b'(x), \end{aligned} \quad (10.7)$$

with Dirichlet boundaries at $-M$ and N provided they are finite. Clearly, L_S is symmetric on $L^2(e^C dx)$. Denote by λ_S the principal eigenvalue of L_S :

$$\lambda_S = \left\{ -(f, L_S f)_{L^2(e^C dx)} : f \in \mathcal{C}_0^\infty(-M, N), \int_{-M}^N f^2 e^C = 1 \right\}.$$

In the setup of (b), formula (10.4) becomes

$$\lambda_1 = \sup_{f \in \mathcal{F}} \inf_{x \in (-M, N)} \frac{-L_S f}{f}(x).$$

This leads to the study on λ_S .

(e) An elementary computation shows that

$$e^C L_S e^{-C} = \widehat{L}. \quad (10.8)$$

Note that

$$\int_{-M}^N e^C f L_S g = \int_{-M}^N e^{-C} (e^C f) (e^C L_S e^{-C}) (e^C g) = \int_{-M}^N e^{-C} \widehat{f} \widehat{L} \widehat{g},$$

where the mapping $f \rightarrow \widehat{f} := e^C f$ is an isometry from $L^2(e^C dx)$ to $L^2(e^{-C} dx)$, and that $\widehat{f} \in \mathcal{C}_0^2(-M, N)$ iff $f \in \mathcal{C}_0^2(-M, N)$. Since $\mathcal{C}_0^2(-M, N)$ is also a common core of L_S and $\widehat{L}$ by the assumption on the coefficients a and b , it follows that the operators L_S and $\widehat{L}$ with the same core $\mathcal{C}_0^\infty(-M, N)$ are isospectral. In particular, we have $\lambda_S = \widehat{\lambda}_0$. When $M, N < \infty$, this means that $\widehat{L}$ has Dirichlet boundaries at $-M$ and N since so does L_S . Now, the basic estimates for λ_S can be obtained in terms of the ones for $\widehat{\lambda}_0$, as will be shown in part (f) below.

To go back to λ_1 , noting that by (10.8) again, we also have

$$\frac{-L_S f}{f} = \frac{-(e^C L_S e^{-C})(e^C f)}{e^C f} = \frac{-\widehat{L} \widehat{f}}{\widehat{f}}.$$

By (a), we can assume that $M, N < \infty$ and $a > 0$ on $[-M, N]$. From Shiozawa and Takeda (2005) and X. Zhang (2007), it is known that

$$\widehat{\lambda}_0 \geq \sup_{f \in \mathcal{F}} \inf_{x \in (-M, N)} \frac{-\widehat{L} \widehat{f}}{\widehat{f}}(x)$$

(i.e., Barta's inequality). To see that the equality sign holds, simply choose f to be the eigenfunction $\hat{g}$ of $\hat{\lambda}_0$. The fact that $\hat{g} \in \mathcal{F}$ is guaranteed by the assumptions that $M, N < \infty$, $\hat{a}$ and $\hat{b}$ are continuous, and $\hat{a} > 0$ on $[-M, N]$. This is a standard (regular) Sturm–Liouville eigenvalue problem. The property $\hat{g}|_{(-M, N)} > 0$ is due to the fact that $\hat{\lambda}_0$ is the minimal eigenvalue. We have thus returned to λ_1 from $\hat{\lambda}_0$ through λ_S .

(f) For the dual operator $\hat{L}$ defined in part (c), applying Theorem 10.1 to $\mathbb{B} = L^1(\hat{\mu})$, we obtain $\hat{\kappa}^{-1}/4 \leq \hat{\lambda}_0 \leq \hat{\kappa}^{-1}$, where

$$\hat{\kappa}^{-1} = \inf_{-M < x < y < N} \left[\left(\int_{-M}^x d\hat{\nu} \right)^{-1} + \left(\int_y^N d\hat{\nu} \right)^{-1} \right] \left(\int_x^y d\hat{\mu} \right)^{-1}.$$

Now, the theorem follows by the dual transform $\hat{\mu} = \nu$ and $\hat{\nu} = \mu$.

Finally, the proof of Theorem 10.2 can be summarized as follows:

- λ_1 for general M, N and continuous a, b
- λ_1 for finite M, N and smooth a, b
- (by approximating and smoothing procedure)
- λ_S (by coupling method leading to the Schrödinger operator)
- $\hat{\lambda}_0$ (by isometry in terms of the dual operator)
- basic estimate of $\hat{\lambda}_0$ (by capacity method: Theorem 10.1)
- basic estimate of λ_1 (by duality). □

Actually, we have also proved the following result (cf. parts (c)–(f) in the last proof) which is parallel to Proposition 9.11.

Proposition 10.3. Let $M, N \leq \infty$, $a, b \in \mathcal{C}^1(-M, N)$ and $a > 0$ on $(-M, N)$. Then for the Schrödinger operator L_S on $L^2(e^C dx)$ having the form (10.7) with Dirichlet boundaries at $-M$ if $M < \infty$, and at N if $N < \infty$, we have $\lambda_S = \hat{\lambda}_0$, and furthermore, $\kappa^{-1}/4 \leq \lambda_S \leq \kappa^{-1}$, where κ is defined by (10.3).

The following simplified estimate of $\kappa^{(10.3)}$ is helpful in practice. Recall that by assumption, $\mu(-M, N) < \infty$. Let $m(\mu)$ be the median of μ (i.e., $\mu(-M, m(\mu)) = \mu(m(\mu), N)$). Given $x \in (-M, m(\mu))$, let $y = y(x)$ be the unique solution to the equation: $\mu(y, N) = \mu(-M, x)$. The A-G inequality $\alpha + \beta \geq 2\sqrt{\alpha\beta}$ suggests the use of $y(x)$, which then leads to a simpler bound:

$$\kappa^{(10.3)} \geq 2^{-1} \sup_{x \in (-M, m(\mu))} \mu(-M, x) \nu(x, y(x)).$$

We remark that the equality sign here holds in some cases, but the inequality sign can happen in general. Anyhow, this provides us a guidance in seeking for the infimum in (10.3). Certainly, the similar discussion is meaningful for $\kappa^{(10.1)}$.

Having Theorems 10.1 and 10.2 at hand, the basic estimates in the other cases (λ^{ND} and λ^{DN}) mentioned in Section 1 should be clear.

The study on the one-dimensional case provides a comparison tool for the study on the higher dimensional situation, as we did a lot before. Hence, there is no doubt for the development in the higher dimensional context.

10.2 h -Transform.

In an earlier draft of this paper (roughly speaking, up to Theorem 7.1 plus a part of Theorem 9.3), the author mentioned an open question: how to handle the case that (1.3) fails? Then two answers have appeared. The first one is the use of so-called h -transform by Wang (2008a) where the transient case studied in Section 7 is transferred into the one studied in Section 4. Next, with the help of the duality given in Theorem 7.1, the ergodic case studied in Section 6 can be also transferred into the one studied in Section 4. In this way, with a use of Theorem 4.2, Wang obtains a criterion for λ_1 (Section 6) with a factor 4. To have a taste of this technique, let us quote a particular result here.

Theorem 10.4 (Wang (2008a, Theorem 1.2)). Set $h_i = \sum_{j=i}^N \mu_j$. Then we have $\delta^{-1}/4 \leq \lambda_1 \leq \delta^{-1}$, where

$$\delta = \sup_{1 \leq i < N+1} \left(\frac{1}{h_i} - \frac{1}{h_0} \right) \sum_{j=i}^N \frac{1}{\mu_i a_i} h_i^2. \quad (10.9)$$

Comparing this result with Corollary 6.6, the factor 4 is in common but the isoperimetric constants are quite different. The advantage here is that only one variable is required in the supremum, but in Corollary 6.6 two variables are needed. The price one has to pay to (10.9) is involving a new quantity h . The natural extension of Corollary 6.6 to the whole line (Corollary 7.9) exhibits an interesting symmetry of the left and the right half-lines. Such an extension of Theorem 10.4 with the same factor 4 is unclear to the author. Along the same line and using [9], Wang then extends the results to Poincaré-type inequalities as well as functional inequalities, see Wang (2008b, c). Clearly, Wang's papers show that the h -transform is a powerful tool and may be useful in other cases.

While the author's solution to the above open question is the use of the maximal process as included into this version of the paper. As shown in the paper, Corollary 6.6 comes from the author's previous general result without using the h -transform. An interesting question in mind is to use the variational formulas in Section 6 to derive Corollary 6.6 directly. Besides, a direct generalization of Sections 2, 3, and 7 to the Poincaré-type inequalities is still meaningful in practice since the formulas are quite different (in view of Theorem 10.4) and some of them may be more practicable.

10.3 Remark on some known results.

As mentioned in Section 5, duality (5.1) goes back to Karlin and McGregor (1957b). The author learned this technique mainly from van Doorn (1981; 1985) based on which the proof of the basic result $\lambda_1 = \alpha^*$ was done, cf. [2]. It is now known that such a result holds in a very general setup as indicated in the proof of Theorem 7.4.

We now discuss the situation that (1.3) holds. Then there are three cases.

- (1) $\sum_i \mu_i = \infty$ and $\sum_i (\mu_i b_i)^{-1} < \infty$.
- (2) $\sum_i \mu_i < \infty$ and $\sum_i (\mu_i b_i)^{-1} = \infty$.
- (3) $\sum_i \mu_i = \sum_i (\mu_i b_i)^{-1} = \infty$.

First, let $b_0 > 0$. In cases (1) or (3), by Theorem 2.4 (1) and Proposition 2.7 (1), $\lambda_0^{(2.2)}$ is equal to

$$\sup_{v \in \mathcal{V}} \inf_{i \geq 0} [a_{i+1} + b_i - a_i/v_{i-1} - b_{i+1}v_i]. \quad (10.10)$$

$$\mathcal{V} := \{v : v_{-1} \text{ is free, } v_i > 0 \text{ for all } i \geq 0\}.$$

In case (2), by Theorem 6.1 (1), λ_1 can be expressed by (10.10). Thus, in view of Proposition 1.2 and [2; Theorem 5.3], the convergence rate α^* can be also expressed by (10.10).

Next, let $b_0 = 0$. Then in case (2), by Corollary 5.2, Proposition 2.7 (1), and using (5.8) in an inverse way, it follows that $\lambda_0^{(4.2)}$ is equal to

$$\sup_{v \in \mathcal{V}} \inf_{i \geq 1} \left[a_i \left(1 - \frac{1}{v_{i-1}} \right) + b_i(1 - v_i) \right], \quad (10.11)$$

$$\mathcal{V} := \{v : v_0 = \infty, v_i > 0 \text{ for all } i \geq 1\}.$$

In case (1), $\lambda_0^{(4.2)}$ is equal to $\lambda_0^{(7.1)}$. By Theorem 7.1 (1), in terms of Theorem 6.1 (1) and using (5.8) in an inverse way, we obtain the same expression (10.11) for $\lambda_0^{(7.1)}$. Finally, in the degenerated case (3), we indeed have $\lambda_0^{(4.2)} = \lambda_0^{(7.1)} = 0$ which can be expressed as (10.11) by Theorem 7.1 (2). Hence, by Proposition 1.2, the convergence rate α^* can also be expressed by (10.11). We have thus obtained the following result.

Theorem 10.5 (van Doorn (2002)). Let (1.3) hold. Then the exponential convergence rate α^* is given by (10.10) or (10.11), respectively, according to $b_0 > 0$ or $b_0 = 0$.

With a slightly different expression, this result was given in van Doorn (2002) by the analysis on the extreme zeros of orthogonal polynomials in Karlin and McGregor's representation, and was actually implied in van Doorn's earlier papers (1985; 1987) as mentioned in the paper just cited or in [3]. In the last paper, this result was rediscovered in the study on λ_1 , using the coupling methods. The lower estimate was also obtained by Zeifman (1991) using a different method in the case that the rates of the processes are bounded, with a missing of the equality.

A progress made in the paper is removing Condition (1.3) and even (1.2). In particular, the situation having finite state spaces is included. This is meaningful not only theoretically but also in practice since the infinite situation can be approximated by the finite ones. Besides, when $b_0 > 0$ and $\sum_i \mu_i < \infty$, the duality given by (5.9) is essentially different from (5.8) (cf. Remark 2.8). From the other point of view, the dual of this case goes to $\lambda_0^{(7.1)}$ rather than $\lambda_0^{(4.2)}$. However, we then have to use the maximal process in Section 6, as we did in Theorem 7.1, rather than the minimal one used in Sections 2 and 3, except using (1.3) (which is equivalent to (1.2) if $\sum_i \mu_i < \infty$). From analytical point of view, the use of the maximal process is natural since one looks for the inequality to be held for the largest class of functions, as illustrated by the weighted Hardy inequality in Section 4.1.

In van Doorn (2002), some variational formulas of difference form for the upper bound of α^* are also presented but we do not use them here. As far as we know, the criterion for $\alpha^* > 0$ (Theorem 1.5) has been open for quite a long time; it was answered in the ergodic case only till [6] in terms of the study on the first non-trivial eigenvalue λ_1 . For which, the criterion was obtained independently by Miclo (1999) based on the weighted Hardy's inequality. Criterion 3.1 follows from the variational formulas of single summation form (part (2) of Theorem 2.4), but it is not obvious at all to deduce the criterion from (10.10) (or dually from (10.11)) directly. More clearly, the variational formula of the difference form for the lower bound given in (9.3) which is closely related to (10.11) was known for some years and works in a more general setup, but an explicit criterion for the killing case is still open (Open Problem 9.13). Anyhow, having the duality (Corollary 5.2 and Theorem 7.1) at hand, Theorem 1.5 is essentially known from [6], except the basic estimates in the ergodic case as well as in the setting of Section 7 is presented here for the first time. The technique adopted in this paper depends heavily on the spectral theory, potential theory, and harmonic analysis. In the transient continuous context, Criterion 3.1 was obtained by Maz'ja (1985, §1.3), as a straightforward consequence of Muckenhoupt (1972). The discrete version was proved by Mao (2002, Proposition A.2). In these quoted papers, the problem in a more general (L^q, L^p) -setup was done.

In the continuous context, the Hardy-type or Sobolev-type inequalities (cf. Theorem 10.1) were studied first by P. Gurka and then by Opic and Kufner (1990, Theorem 8.3). Instead of $\mathcal{D}^{\min}(D)$, they considered the following class of functions: the absolutely continuous functions vanishing at $-M$ and N . This seems not essential in view of $\lambda_0^{(2.2)} = \lambda_0^{(2.18)}$. With a different but equivalent isoperimetric constant (i.e., replacing the sum in (10.1) by maximum “ $\vee$ ”), they obtained upper and lower bounds with ratio $2\omega^5 \approx 22$, where $\omega = (\sqrt{5} + 1)/2$ is the gold section number. By the way, we mention that the use of weight functions w and v in (8.6) in the cited book is formally more general than our setup. One can first assume that w and v are positive, otherwise replace them by $w + 1/n$ and $v + 1/n$, respectively, and then pass to the limit as $n \rightarrow \infty$. Next, it is easy to rewrite w and v as e^C/a and e^C for some functions C and $a > 0$. Note that only C and a (without using b) are needed to deduce the basic estimates in our proof. Again, in the continuous context, the splitting technique was also used in Theorem 8.8 of the book just quoted where some basic estimates were derived in terms of an isoperimetric constant, up to a factor 8. Their isoperimetric constant is parallel to the right-hand side of (7.13) replacing $\lambda_0^{\theta\pm}$ by the corresponding $\delta^{(3.1)\pm}$ depending on θ (certainly, without using the parameter γ). Our Example 8.9 is an analog of Examples 6.13 and 8.16 in the quoted book. In contrast with our probabilistic–analytic proof here, their proof is direct, analytic, and works in a more general (L^q, L^p) -setup. We have not seen the discrete analog of their results in the literature. In the (L^p, L^p) -sense ($p \geq 1$), the variational formulas in the continuous context were obtained in Jin (2006) but it remains open for the more general (L^q, L^p) -setup. Even though it is a typical Sturm-Liouville eigenvalue problem having richer literature, we are unable to find an analog of Theorem 10.2.

Finally, in computing the examples in the paper, the author uses the software Mathematica. All the examples were checked by Ling-Di Wang and Chi Zhang using MatLab. Most of the author's papers cited here can be found in [8].

Acknowledgements The first version of the paper dated May 25, 2007. In June, 2008, the author was invited by Professors Tiee-Jian Wu and Yuan-Shih Chow to visit the Center for Theoretical Sciences at Cheng Kung University and Taiwan University. The author was also invited by Professors Mong-Na Lo Haung and Mei-Hui Guo (Sun Yat-sen University), by Professor Wen-Jang Huang (University of Kaohsiung), by Professor Chen-Hai Tsao (Dong Hwa University for the 17th Conference on Statistics in the Southern area and the Annual Meeting of Chinese Institute of Probability and Statistics), by Professors Chii-Ruey Hwang, Tzuu-Shuh Chiang, Yunshyong Chow and Shuenn-Jyi Sheu (Academia Sinica) for visiting their institutes. The scientific communications, the very warm hospitality of the inviters and the financial support from the Center for Theoretical Sciences, and the inviters' universities are deeply acknowledged. The author has also benefited from several discussions with Professor Shuenn-Jyi Sheu in Beijing. A part of the results in the paper were reported during the visit.

Partial results of the paper have also been reported in the following conferences: International Conference on Stochastic Analysis and Related Fields (April 2008, Wuhan), Workshop on Probability (November 2008, Guangzhou), Chinese-French Meeting in Probability and Analysis (September 2008, Marseille), Workshop on IPS in honor of Professor Tom Liggett's 65th Birthday (June 2009, Peking University), IMS-China International Conference on Statistics and Probability (July 2009, Weihai). The author acknowledges the organizers Zhi-Ming Ma and Michael Röckner, Jia-Gang Ren, Dominique Bakry and Yves LeJan, Da-Yue Chen, Jia-An Yan, respectively, for their invitation and financial support.

Because of an honorary professorship from Swansea University, the author was invited to a workshop and visited University of Strathclyde in UK in October, 2009. The recent progress was reported at these two universities. The author appreciates the invitation, the very warm hospitality, and the financial support by Professors N. Jacob, X.R. Mao, F.Y. Wang, J.L. Wu, C.G. Yuan, and their institutes. Special thanks are given to the speakers of the celebration workshop for their kindness: T.J. Lyons, M. Davis, F.Y. Wang, T.S. Zhang, X.R. Mao, and D. Elworthy.

Two drafts of the paper have been reported in a series of seminars organized by Professors Yong-Hua Mao and Yu-Hui Zhang. Especially, the drafts are checked line by line by Zhang's graduate students Ling-Di Wang and Chi Zhang. The corrections and improvement obtained from them improved a lot the quality of the paper. Their serious work is acknowledged. Since May 2007, the author has reported the results from time to time in his seminars and thanks are given to the participants for their discussions and comments.

REFERENCES

1. Chen, A., Pollett, P., Zhang, H. and Cairns, B. (2005), *Uniqueness criteria for continuous-time Markov chains with general transition structure*, Adv. Appl. Prob. 37:4, 1056–1074.

2. Chen, M.F. (1991), *Exponential L^2 -convergence and L^2 -spectral gap for Markov processes*, Acta Math. Sin., New Series 7:1, 19–37.
3. Chen, M.F. (1996), *Estimation of spectral gap for Markov chains*, Acta Math. Sin., New Series 12:4, 337–360.
4. Chen, M.F. (1999), *Analytic proof of dual variational formula for the first eigenvalue in dimension one*, Sci. in China (A) 42:8, 805–815.
5. Chen, M.F. (2000a), *The principal eigenvalue for jump processes*, Acta Math. Sin. Eng. Ser. 16:3, 361–368.
6. Chen, M.F. (2000b), *Explicit bounds of the first eigenvalue*, Sci. Chin. Ser. A 43(10), 1051–1059.
7. Chen, M.F. (2001), *Variational formulas and approximation theorems for the first eigenvalue in dimension one*, Sci. Chin. Ser. A 44(4), 409–418.
8. Chen, M.F. (2001–), *Ergodic Convergence Rates of Markov Processes — Eigenvalues, Inequalities and Ergodic Theory* [Collection of papers, 1993—]
http://math.bnu.edu.cn/~chenmf/main_eng.htm.
9. Chen, M.F. (2003), *Variational formulas of Poincaré-type inequalities for birth-death processes*, Acta Math. Sin. Eng. Ser. 19:4, 625–644.
10. Chen, M.F. (2004), *From Markov Chains to Non-equilibrium Particle Systems*, World Scientific. 2nd ed. (1st ed., 1992).
11. Chen, M.F. (2005a), *Capacitary criteria for Poincaré-type inequalities*, Potential Theory 23:4, 303–322.
12. Chen, M.F. (2005b), *Eigenvalues, Inequalities, and Ergodic Theory*, Springer, London.
13. Chen, M.F. Wang, F.Y. (1997), *Estimation of spectral gap for elliptic operators*, Trans. Amer. Math. Soc. 349:3, 1239–1267.
14. Chen, M.F. and Wang, F.Y. (1998), *Cheeger's inequalities for general symmetric forms and existence criteria for spectral gap*, Abstract. Chin. Sci. Bulletin 43:18, 1516–1519. Ann. Prob. 2000, 28:1, 235–257.
15. Chen, M.F., Zhang, Y.H. and Zhao, X.L. (2003), *Dual variational formulas for the first Dirichlet eigenvalue on half-line*, Sci. China 46:6, 847–861.
16. Cox, J.T. and Rösler, U. (1983), *A duality relation for entrance and exit laws for Markov processes*, Stoch. Proc. Appl. 16, 141–156.
17. Dobrushin, R.L. (1952), *On conditions of regularity of stationary Markov processes with a denumerable number of possible states* (in Russian), Uspehi Matem. Nauk (N.S.) 7:6, 185–191.
18. Fukushima, M. and Uemura, T. (2003), *Capacitary bounds of measures and ultracontractivity of time changed processes*, J. Math. Pure et Appliques 82:5, 553–572.
19. Hou, Z.T., Liu, Z.M., Zhang, H.J., Li, J.P., Zhou, J.Z., and Yuan, C.G. (2000), *Birth-death Processes* (in Chinese), Hunan Sci. Press, Hunan.
20. Hou, Z.T., Zhou, J.Z., Zhang, H.J., Liu, Z.M., Xiao, G.N., Chen, A.Y., and Fei, Z.L. (1994), *The Q-matrix Problem for Markov Chains* (in Chinese), Hunan Sci. Press, Hunan.
21. Jin, H.Y. (2006), *Estimating the constant in L^p -Poincaré inequality* (in Chinese), Master Thesis, Beijing Normal Univ.
22. Karlin, S. and McGregor, J.L. (1957a), *The differential equations of birth-and-death processes, and the Stieltjes moment problem*, Trans. Amer. Math. Soc. **85**, 589–646.
23. Karlin, S. and McGregor, J. (1957b), *The classification of birth and death processes*, Trans. Amer. Math. Soc. 86:2, 366–400.
24. Kijima, M. (1997), *Markov Processes for Stochastic modeling*, Chapman & Hall, London.
25. Mao, Y.H. (2002), *Nash inequalities for Markov processes in dimension one*, Acta. Math. Sin. Eng. Ser. 18:1: 147–156.
26. Mao, Y.H. and Xia, L.Y. (2009), *Spectral gap for jump processes by decomposition method*, Front. Math. China 4:2, 335–347.
27. Maz'ya, V.G. (1985), *Sobolev Spaces*, Springer.
28. Miclo, L. (1999), *An example of application of discrete Hardy's inequalities*, Markov Processes Relat. Fields 5, 319–330.
29. Muckenhoupt, B. (1972), *Hardy's inequality with weights*, Studia Math. XLIV: 31–38.
30. Opic, B. and Kufner, A. (1990), *Hardy-type Inequalities*, Longman, New York.

31. Shao, J.H. and Mao, Y.H. (2007), *Estimation of the Dirichlet eigenvalue of birth-death process on trees*, Acta Math. Sin. Chinese Ser. 50:3, 507–516.
32. Shiozawa, Y. and Takeda, M. (2005), *Variational formula for Dirichlet forms and estimates of principal eigenvalues for symmetric α -stable processes*, Potential Analysis **23**, 135–151.
33. Sirl, D., Zhang, H. and Pollett, P. (2007), *Computable bounds for the decay parameter of a birth-death process*, J. Appl. Prob. 44:2, 476–491.
34. van Doorn, E.A. (1981), *Stochastic Monotonicity and Queuing Applications of Birth-Death Processes*, Lecture Notes in Statistics vol. 4, Springer-Verlag.
35. van Doorn, E.A. (1985), *Conditions for exponential ergodicity and bounds for the decay parameter of a birth-death process*, Adv. Appl. Prob. **17**, 514–530.
36. van Doorn, E.A. (1987), *Representations and bounds for zeros of orthogonal polynomials and eigenvalues of sign-symmetric tri-diagonal matrices*, J. Approx. Th. 51, 254–266.
37. van Doorn, E.A. (2002), *Representations for the rate of convergence of birth-death processes*, Theory Probab. Math. Statist. **65**, 37–43.
38. Wang, J. (2008a), *First Dirichlet eigenvalue of transient birth-death processes*, preprint.
39. Wang, J. (2008b), *Poincaré-type inequalities for transient birth-death processes*, preprint.
40. Wang, J. (2008c), *Functional inequalities for transient birth-death processes and their applications*, preprint.
41. Wang, Z.K. (1964), *The ergodicity and zero-one law for birth and death processes* (in Chinese), Acta Sci. Nankai Univ. 5:5, 93–102.
42. Wang, Z.K. and Yang, X.Q. (1992), *Birth and Death Processes and Markov Chains*, Springer, Berlin and Sci. Press, Beijing.
43. Zeifman, A.I. (1991), *Some estimates of the rate of convergence for birth and death processes*, J. Appl. Prob. **28**, 268–277.
44. Zhang, X. (2007), *The estimation for the first eigenvalue of Schrödinger operators and a class of geometric inequalities* (in Chinese), Ph.D Thesis, Beijing Normal University.

School of Mathematical Sciences, Beijing Normal University, Laboratory of Mathematics and Complex Systems (Beijing Normal University), Ministry of Education, Beijing 100875, The People's Republic of China.

E-mail: mfchen@bnu.edu.cn

Home page: http://math.bnu.edu.cn/~chenmf/main_eng.htm

Received January 6, 2010; accepted March 6, 2010

Chapter 6 in “Probability Approximations and Beyond”, 75–99,
Lecture Notes in Statistics 205,
eds. A.D. Barbour, H.P. Chan and D. Siegmund, 2012

Basic Estimates of Stability Rate for One-dimensional Diffusions

Mu-Fa Chen

(Beijing Normal University, Beijing 100875, China)

Abstract

In the context of one-dimensional diffusions, we present basic estimates (having the same lower and upper bounds with a factor of 4 only) for four Poincaré-type (or Hardy-type) inequalities. The derivation of two estimates have been open problems for quite some time. The bounds provide exponentially ergodic or decay rates. We refine the bounds and illustrate them with typical examples.

1 Introduction

An earlier topic on which Louis Chen has studied is about the Poincaré-type inequalities (see [1, 2], for instance). We now use this good chance to introduction in Section 2 some recent progress on the topic, especially for one-dimensional diffusions (elliptic operators). The basic estimates of exponentially ergodic (or decay) rate and the principal eigenvalue in different cases are presented. Here the term “basic” means that upper and lower bounds are given by an isoperimetric constant up to a factor four. As a consequence, the criteria for the positivity of the rate and the eigenvalue are obtained. The proof of the main result is sketched in Section 3. The materials given in Sections 4, 5, and Appendix are new. In particular, the basic estimates are refined in Section 4 and the results are illustrated through examples in Section 5. The coincidence of the exponentially decay rate and the corresponding principal eigenvalue is proven in Appendix for a large class of symmetric Markov processes.

2000 *Mathematics Subject Classifications.* 60J60, 34L15, 26D10.

Key words and phrases. First nontrivial eigenvalue, Hardy’s inequality, one-dimensional diffusion, principal eigenvalue, Poincaré-type inequality, stability rate.

In the published version as Chapter 6 in the book, Theorem 2.1, Proposition 3.2, Corollary 4.3 and Example 5.1 here are relabeled as Theorem 6.1, Proposition 6.1, Corollary 6.1 and Example 6.1, respectively. Similarly, formulas (1)–(36) here are relabeled as (6.1)–(6.36).

2 The main result and motivation

2.1 Two types of exponential convergence

Let us recall two types of exponential convergence often studied for Markov processes. Let $P_t(x, \cdot)$ be a transition probability on a measurable state space $(E, \mathcal{E})$ with stationary distribution π . Then the process is called *exponentially ergodic* if there exists a constant $\varepsilon > 0$ and a function $c(x)$ such that

$$\|P_t(x, \cdot) - \pi\|_{\text{Var}} \leq c(x)e^{-\varepsilon t}, \quad t \geq 0, x \in E. \quad (1)$$

Denote by $\varepsilon_{\max}$ be the maximal rate ε . For convenience, in what follows, we allow $\varepsilon_{\max} = 0$. Next, let $L^2(\pi)$ be the real $L^2(\pi)$ -space with inner product $(\cdot, \cdot)$ and norm $\|\cdot\|$ respectively, and denote by $\{P_t\}_{t \geq 0}$ the semigroup of the process. Then the process is called to have *L^2 -exponential convergence* if there exists some $\eta (\geq 0)$ such that

$$\|P_t f - \pi(f)\| \leq \|f - \pi(f)\|e^{-\eta t}, \quad t \geq 0, f \in L^2(\pi), \quad (2)$$

where $\pi(f) = \int_E f d\pi$. It is known that $\eta_{\max}$ is described by λ_1 :

$$\lambda_1 = \inf\{(f, -Lf) : f \in \mathcal{D}(L), \pi(f) = 0, \|f\| = 1\}, \quad (3)$$

where L is the generator with domain $\mathcal{D}(L)$ of the semigroup in $L^2(\pi)$. Even though the topologies for these two types of exponential convergence are rather different, but we do have the following result.

Theorem 2.1 ([3, 6]) For a reversible Markov process with symmetric measure π , if with respect to π , the transition probability has a density $p_t(x, y)$ having the property that the diagonal elements $p_s(\cdot, \cdot) \in L_{\text{loc}}^{1/2}(\pi)$ for some $s > 0$, and a set of bounded functions with compact support is dense in $L^2(\pi)$, then we have $\varepsilon_{\max} = \lambda_1$.

As an immediate consequence of the theorem, we obtain some criterion for $\lambda_1 > 0$ in terms of the known criterion for $\varepsilon_{\max} > 0$. In our recent study, we go to the opposite direction: estimating $\varepsilon_{\max}$ in terms of the spectral theory.

We are also going to handle with the non-ergodic case in which (2) becomes

$$\mu((P_t f)^2) \leq \mu(f^2)e^{-2\eta t}, \quad t \geq 0, f \in L^2(\mu), \quad (4)$$

where μ is the invariant measure of the process. Then $\eta_{\max}$ becomes

$$\lambda_0 = \inf\{-\mu(fLf) : f \in \mathcal{C}, \mu(f^2) = 1\}, \quad (5)$$

where $\mathcal{C}$ is a suitable core of the generator, the smooth functions with compact support for instance in the context of diffusions. However, the totally variational norm in (1) may be meaningless unless the process being explosive. Instead of (1), we consider the following exponential convergence:

$$P_t(x, K) \leq c(x, K)e^{-\varepsilon t}, \quad t \geq 0, x \in E, K: \text{compact}, \quad (6)$$

where for each compact K , $c(\cdot, K)$ is locally μ -integrable. Under some mild condition, we still have $\varepsilon_{\max} = \lambda_0$. See Appendix for more details.

2.2 Statement of the result

We now turn to our main object: one-dimensional diffusions. The state space is $E := (-M, N)$ ($M, N \leq \infty$). Consider an elliptic operator

$$L = a(x) \frac{d^2}{dx^2} + b(x) \frac{d}{dx},$$

where $a > 0$ on E . Then define a function $C(x)$ as follows:

$$C(x) = \int_{\theta}^x \frac{b}{a}, \quad x \in E,$$

where $\theta \in E$ is a reference point. Here and in what follows, the Lebesgue measure dx is often omitted. It is convenient for us to define two measures μ and ν as follows.

$$\mu(dx) = \frac{e^{C(x)}}{a(x)} dx, \quad \nu(dx) = e^{-C(x)} dx.$$

The first one has different names: *speed*, or *invariant*, or *symmetrizable measure*. The second one is called *scale measure*. Note that ν is infinite iff the process is recurrent. By using these measures, the operator L takes a very compact form

$$L = \frac{d}{d\mu} \frac{d}{d\nu} \quad \left(\text{i.e., } Lf \equiv a e^{-C} (f' e^C)' \right) \quad (7)$$

which goes back to a series of papers by W. Feller, for instance [12].

Consider first the special case that $M, N < \infty$. Then the ergodic case means that the process has reflection boundaries at $-M$ and N . In analytic language, we have Neumann boundaries at $-M$ and N : the eigenfunction g of λ_1 satisfies $g'(-M) = g'(N) = 0$. Otherwise, in the non-ergodic case, one of the boundaries becomes absorbing. In analytic language, we have Dirichlet boundary at $-M$ (say): the eigenfunction g of λ_0 satisfies $g(-M) = 0$. Let us use codes “D” and “N”, respectively, to denote the Dirichlet and Neumann boundaries. The corresponding minimal eigenvalues of $-L$ are listed as follows.

- λ^{NN} : Neumann boundaries at $-M$ and N ,
- λ^{DD} : Dirichlet boundaries at $-M$ and N ,
- λ^{DN} : Dirichlet at 0 and Neumann at N ,
- λ^{ND} : Neumann at 0 and Dirichlet at N .

We call them *the first non-trivial* or *the principal eigenvalue*. In the last two cases, setting $M = 0$ is for convenience in comparison with other results to be discussed later but it is not necessary. Certainly, this classification is still meaningful if M or N is infinite. For instance, in the ergodic case, the process will certainly come back from any starting point and so one may

imagine the boundaries $\pm\infty$ as reflecting. In other words, the probabilistic interpretation remains the same when $M, N = \infty$. However, the analytic Neumann condition that $\lim_{x \rightarrow \pm\infty} g'(x) = 0$ for the eigenfunction g of λ^{NN} may be lost (cf. the first example given in Section 5). More seriously, the spectrum of the operator may be continuous for unbounded intervals. This is the reason why we need the L^2 -spectral theory. In the Dirichlet case, the analytic condition that $\lim_{x \rightarrow \pm\infty} g(x) = 0$ can be implied by the definition given below, once the process goes $\pm\infty$ exponentially fast. Now, for general $M, N \leq \infty$, let

$$D(f) = \int_{-M}^N f'^2 e^C, \quad M, N \leq \infty, f \in \mathcal{A}(-M, N),$$

$\mathcal{A}(-M, N)$ = the set of absolutely continuous functions on $(-M, N)$,

$\mathcal{A}_0(-M, N) = \{f \in \mathcal{A}(-M, N) : f \text{ has a compact support}\}.$

From now on, the inner product $(\cdot, \cdot)$ and the norm $\|\cdot\|$ are taken with respect to μ (instead of π). Then the principal eigenvalues are defined as follows.

$$\lambda^{\text{DD}} = \inf\{D(f) : f \in \mathcal{A}_0(-M, N), \|f\| = 1\}, \quad (8)$$

$$\lambda^{\text{ND}} = \inf\{D(f) : f \in \mathcal{A}_0(0, N), f(N-) = 0 \text{ if } N < \infty, \|f\| = 1\}, \quad (9)$$

$$\lambda^{\text{NN}} = \inf\{D(f) : f \in \mathcal{A}(-M, N), \mu(f) = 0, \|f\| = 1\}, \quad (10)$$

$$\lambda^{\text{DN}} = \inf\{D(f) : f \in \mathcal{A}(0, N), f(0+) = 0, \|f\| = 1\}. \quad (11)$$

Certainly, the above classification is closely related to the measures μ and ν . For instance, in the DN- and NN-cases, one requires that $\mu(0, N) < \infty$ and $\mu(-M, N) < \infty$, respectively. Otherwise, one gets a trivial result as can be seen from Theorem 2.2 below.

To state the main result of the paper, we need some assumptions. In the NN-case (i.e., the ergodic one), we technically assume that a and b are continuous on $(-M, N)$. For λ^{DN} and λ^{NN} , we allow the process to be explosive since the maximal domain is adopted in definition of λ^{DN} and λ^{NN} . But for λ^{ND} and λ^{DD} , we are working for the minimal process (using the minimal domain) only, assuming that μ and ν are locally finite.

Theorem 2.2 (Basic estimates [9]) Under the assumptions just mentioned, corresponding to each #-case, we have

$$(\kappa^\#)^{-1}/4 \leq \lambda^\# = \varepsilon_{\max} \leq (\kappa^\#)^{-1}, \quad (12)$$

where

$$(\kappa^{\text{NN}})^{-1} = \inf_{x < y} [\mu(-M, x)^{-1} + \mu(y, N)^{-1}] \nu(x, y)^{-1}, \quad (13)$$

$$(\kappa^{\text{DD}})^{-1} = \inf_{x < y} [\nu(-M, x)^{-1} + \nu(y, N)^{-1}] \mu(x, y)^{-1}, \quad (14)$$

$$\kappa^{\text{DN}} = \sup_{x \in (0, N)} \nu(0, x) \mu(x, N) \quad (15)$$

$$\kappa^{\text{ND}} = \sup_{x \in (0, N)} \mu(0, x) \nu(x, N). \quad (16)$$

In particular, $\lambda^\# > 0$ iff $\kappa^\# < \infty$.

In each case, the principal eigenvalue is controlled from above and below by a constant $\kappa^\#$ up to a factor 4 which is universal. Among these cases, the hardest one is the ergodic case. It may be helpful for the reader to show how to write down κ^{NN} step by step.

- We need two parameters, say x and y with $x < y$. The state space is then divided by x and y into three parts: the left-hand part $(-M, x)$, the right-hand part (y, N) , and the middle one (x, y) .
- Measure the left-hand and the right-hand subintervals by μ and the middle one by ν , respectively:

$$\kappa = \kappa^{\text{NN}} : \quad \mu(-M, x) \quad \mu(y, N) \quad \nu(x, y).$$

- Make inverse everywhere:

$$\kappa^{-1} : \quad \mu(-M, x)^{-1} \quad \mu(y, N)^{-1} \quad \nu(x, y)^{-1}.$$

- Finally, summing up the first two terms and making infimum with respect to $x < y$, we get the answer.

Every step is quite natural except the second one: why we use μ but not ν in the first two terms? This is because we are in the ergodic case, μ is a finite measure. If μ is replaced by ν , since $\nu(-\infty, x)$ and $\nu(y, \infty)$ are infinite when $M, N = \infty$, one would get zero for these terms and so the quantity is trivial. A sensitive point here is that we use plus, rather than maximum in the last step. Otherwise, even though the resulting bounds are equivalent to ours but it then would produce a factor 8 rather than 4 as we expected. We have thus completed the first, the most important quantity κ^{NN} . To get κ^{DD} , simply apply the rule: exchanging the codes D and N simultaneously in $\kappa^\#$ leads to the exchange of the measures μ and ν in the formula. Let us now examine (14) more carefully. When $N = \infty$ and $\nu(y, \infty) = \infty$, the second term in the sum of (14) disappeared. In other words, the boundary condition D on the right endpoint is replaced by N. Then the variable y is free and so can be removed. Therefore we obtain formula (15). We remark however that the relation between λ^{DN} and κ^{DN} remains the same even if $\nu(y, \infty) < \infty$. From (15), using again our rule, we obtain (16). We mention that (16) can be formally obtained from (13) by removing the second term in the sum. Actually, (16) is formally a reverse of (15), and so is somehow an easy consequence of (15).

2.3 Short review on the known results

It is the position to say a little about the history of the topic. Clearly, we are in the typical situation of the Sturm–Liouville eigenvalue problem (1836–1837). From which, we learn the general properties of the eigenfunction: the existence and uniqueness, the zeros of the eigenfunction, and so on. Except some very specific cases, the problem is usually not solvable analytically. This leads to the theory of special functions used widely in sciences. The estimation of the principal eigenvalues is usually not included in the Sturm–Liouville theory but is studied in harmonic analysis (especially for λ^{DN}). To see this, rewrite (11) as the *Poincaré inequality*

$$\lambda^{\text{DN}} \|f\|^2 \leq D(f), \quad f(0) = 0.$$

More general, we have *Hardy's inequality*

$$\|f\|_{L^p(\mu)}^p \leq A_p \int_{-M}^N |f'|^p e^C, \quad f(0) = 0, \quad p > 1$$

where A_p denotes the optimal constant in the inequality. Certainly, $A_2 = (\lambda^{\text{DN}})^{-1}$. This was initialed, for the specific operator $L = x^2 d^2/dx^2$, by G.H. Hardy [16] in 1920, motivated from a theorem of Hilbert on double series. To which, several famous mathematicians (H. Weyl, F.W. Wiener, I. Schur, et al.) were involved. After a half-century, the basic estimates in the DN-case were finally obtained by several mathematicians, for instance B. Muckenhoupt (1972). The reason should be now clear why (15) can be so famous in the history. The estimate of λ^{ND} was given in Maz'ya (1985). In the DD-case, the problem was begun by P. Gurka [14] around 1989 and then improved in the book by Opic and Kufner (1990) with a factor ≈ 22 . In terms of a splitting technique, the NN-case can be reduced to the Muckenhoupt's estimate with a factor 8, as shown by Miclo (1999) in the context of birth–death processes. A better estimate can be done in terms of variational formulas given in [4; Theorem 3.3]. It is surprising that in the more complicated DD- and NN-cases, by adding one more parameter only, we can still obtain a compact expression (13) and (14). Note that these two formulas have the following advantage: the left- and the right-hand parts are symmetric; the cases having finite or infinite intervals are unified together without using the splitting technique.

2.4 Motivation and application

Here is a quick overview of our motivation and application of the study on this topic. Consider the φ^4 -model on the d -dimensional lattice $\mathbb{Z}^d$. At each site i , there is a one-dimensional diffusion with operator $L_i = d^2/dx_i^2 - u'(x_i)d/dx_i$, where $u_i(x_i) = x_i^4 - \beta x_i^2$ having a parameter $\beta \geq 0$. Between the nearest neighbors i and j in $\mathbb{Z}^d$, there is an interaction. That is, we have an interaction potential $H(x) = -J \sum_{\langle ij \rangle} x_i x_j$ with parameter $J \geq 0$. For each finite box Λ

(denoted by $\Lambda \in \mathbb{Z}^d$) and $\omega \in \mathbb{R}^{\mathbb{Z}^d}$, let H_Λ^ω denote the conditional Hamiltonian (which acts on those x : $x_k = \omega_k$ for all $k \notin \Lambda$). Then, we have a local operator

$$L_\Lambda^\omega = \sum_{i \in \Lambda} [\partial_{ii} - \partial_i(u + H_\Lambda^\omega)\partial_i].$$

We proved that the first non-trivial eigenvalue $\lambda_1^\beta(\Lambda, \omega)$ (as well as the logarithmic Sobolev constant $\sigma^\beta(\Lambda, \omega)$ which is not touched here) of L_Λ^ω is approximately $\exp[-\beta^2/4] - 4dJ$ uniformly with respect to the boxes Λ and the boundaries ω . The leading rate $\beta^2/4$ is exact which is the only one we have ever known up to now for a continuous model.

Theorem 2.3 ([8]) For the φ^4 -model given above, we have

$$\inf_{\Lambda \in \mathbb{Z}^d} \inf_{\omega \in \mathbb{R}^{\mathbb{Z}^d}} \lambda_1^\beta(\Lambda, \omega) \approx \inf_{\Lambda \in \mathbb{Z}^d} \inf_{\omega \in \mathbb{R}^{\mathbb{Z}^d}} \sigma^\beta(\Lambda, \omega) \approx \exp[-\beta^2/4 - c \log \beta] - 4dJ,$$

where $c \in [1, 2]$. See Figure 1.

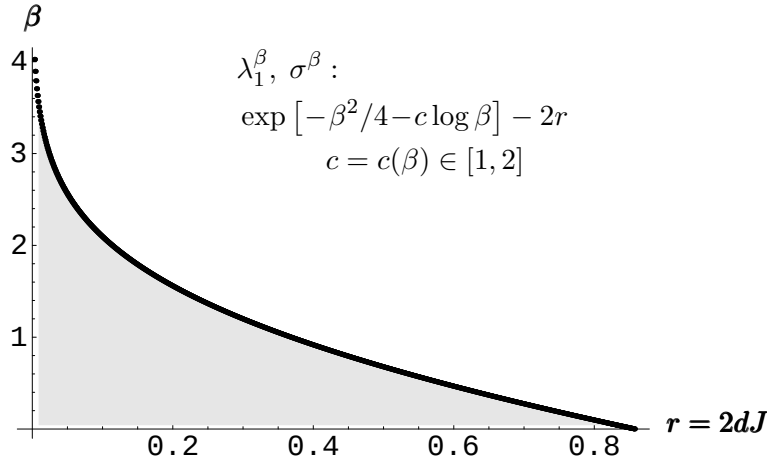


Figure 1 Phase transition of the φ^4 model

The figure says that in the gray region, the system has a positive principal eigenvalue and so is ergodic; but in the region which is a little away above the curve, the eigenvalue vanishes. The picture exhibits a phase transition. The key to prove Theorem 2.2 is a deep understanding about the one-dimensional case. Having one-dimensional result at hand, as far as we know, there are at least three different ways to go to the higher or even infinite dimensions: the conditional technique used in [8]; the coupling method explained in [6; Chapter 2]; and some suitable comparison which is often used in studying the stability rate of interacting particle systems. This explains our original motivation and shows the value of a sharp estimate for the leading eigenvalue in dimension one. The application of the present result to this model should be clear now.

3 Sketch of the proof

The hardest part of Theorem 2.2 is the assertion for λ^{NN} . Here we sketch its proof. Meanwhile, the proof for λ^{DD} is also sketched. The proof for the first assertion consists mainly of three steps by using three methods: the coupling method, the dual method, and the capacity method.

3.1 Coupling method

The next result was proved by using the coupling technique.

Theorem 3.1 (Chen and Wang (1997)) For the operator L on $(0, \infty)$ with reflection at 0, we have

$$\lambda_1 = \lambda^{\text{NN}} \geq \sup_{f \in \mathcal{F}} \inf_{x > 0} \left[-b' - \frac{af'' + (a' + b)f'}{f} \right](x), \quad (17)$$

$$\mathcal{F} = \{f \in \mathcal{C}^2(0, \infty) : f(0) = 0, f|_{(0, \infty)} > 0\}. \quad (18)$$

Actually, the equality sign holds once the eigenfunction of λ_1 belongs to C^3 .

We now rewrite the above formula in terms of an operator, Schrödinger operator L_S .

$$\lambda_1 = \sup_{f \in \mathcal{F}} \inf_{x > 0} \left[-b' - \frac{af'' + (a' + b)f'}{f} \right](x) \quad (19)$$

$$= \sup_{f \in \mathcal{F}} \inf_{x > 0} \left(-\frac{L_S f}{f} \right)(x) =: \lambda_S, \quad (20)$$

$$L_S = a(x) \frac{d^2}{dx^2} + (a'(x) + b(x)) \frac{d}{dx} + b'(x). \quad (21)$$

The original condition $\pi(f) = 0$ in the definition of λ^{NN} means that f has to change its sign. Note that f is regarded as a mimic of the eigenfunction g . The difficulty is that we do not know where the zero-point of g is located. In the new formula (20), the zero-point of $f \in \mathcal{F}$ is fixed at the boundary 0, the function is positive inside of the interval. This is the advantage of formula (20). Now, a new problem appears: there is an additional potential term $b'(x)$. Since $b'(x)$ can be positive, the operator L_S is Schrödinger but may not be an elliptic operator with killing. Up to now, we are still unable to handle with general Schrödinger operator (even with killing one), but at the moment, the potential term is very specific so it gives a hope to go further.

3.2 Dual method

To overcome the difficulty just mentioned, the idea is a use of duality. The dual now we adopted is very simple: just an exchange of the two measures μ

and ν . Recall that the original operator is $L = \frac{d}{d\mu} \frac{d}{d\nu}$ by (7). Hence the dual operator takes the following form

$$L^* = \frac{d}{d\mu^*} \frac{d}{d\nu^*} = \frac{d}{d\nu} \frac{d}{d\mu}, \quad (22)$$

$$L^* = a(x) \frac{d^2}{dx^2} + (a'(x) - b(x)) \frac{d}{dx}, \quad x \in (0, \infty). \quad (23)$$

This dual goes back to Siegmund (1976) and Cox & Rösler (1983) (in which the probabilistic meaning of this duality was explained), as an analog of the duality for birth–death process (cf. [9] for more details and original references). It is now a simple matter to check that the dual operator is a similar transform of the Schrödinger one

$$L^* = e^C L_S e^{-C}. \quad (24)$$

This implies that

$$-\frac{L_S f}{f} = -\frac{L^* f^*}{f^*},$$

where $f^* := e^C f$ is one-to-one from $\mathcal{F}$ into itself. Therefore, we have

$$\lambda_S = \sup_{f \in \mathcal{F}} \inf_{x > 0} \frac{-L_S f}{f}(x) = \sup_{f^* \in \mathcal{F}} \inf_{x > 0} \frac{-L^* f^*}{f^*}(x) = \lambda^{*DD},$$

where the last equality is the so-called *Barta's equality*.

we have thus obtained the following identity.

Proposition 3.2 $\lambda_1 = \lambda_S = \lambda^{*DD}$.

Actually, we have a more general conclusion that L_S and L^* are isospectral from $L^2(e^C dx)$ to $L^2(e^{-C} dx)$. This is because of

$$\int e^C f L_S g = \int e^{-C} (e^C f) (e^C L_S e^{-C}) (e^C g) = \int e^{-C} f^* L^* g^*,$$

and L_S and L^* have a common core. But L on $L^2(\mu)$ and its dual L^* on $L^2(e^{-C} dx)$ are clearly not isospectral.

The rule mentioned in the remark after Theorem 2.2, and used to deduced (14) from (13), comes from this duality. Nevertheless, it remains to compute λ^{DD} for the dual operator.

3.3 Capacitary method

To compute λ^{DD} , we need a general result which comes from a different direction to generalize the Hardy-type inequalities. In contract to what we have talked so far, this time we extend the inequalities to the higher dimensional situation. This leads to a use of the capacity since in the higher dimensions, the boundary may be very complicated. After a great effort by many mathematicians (see for instance Maz'ya 1985; Hasson 1979; Vondraček 1996; Fukushima & Uemura 2003; and [7]), we have finally the following result.

Theorem 3.3 For a regular transient Dirichlet form $(D, \mathcal{D}(D))$ with locally compact state space $(E, \mathcal{E})$, the optimal constant $A_{\mathbb{B}}$ in the Poincaré-type inequality

$$\|f^2\|_{\mathbb{B}} \leq A_{\mathbb{B}} D(f), \quad f \in \mathcal{C}_K^\infty(E)$$

satisfies $B_{\mathbb{B}} \leq A_{\mathbb{B}} \leq 4B_{\mathbb{B}}$, where $\|\cdot\|_{\mathbb{B}}$ is the norm in a normed linear space $\mathbb{B}$ and

$$B_{\mathbb{B}} = \sup_{\text{compact } K} \text{Cap}(K)^{-1} \|\mathbb{1}_K\|_{\mathbb{B}}.$$

The space $\mathbb{B}$ can be very general, for instance $L^p(\mu)$ ($p \geq 1$) or the Orlicz spaces. In the present context, $D(f) = \int_{-M}^N f'^2 e^C$, $\mathcal{D}(D)$ is the closure of $\mathcal{C}_K^\infty(-M, N)$ with respect to the norm $\|\cdot\|_D$: $\|f\|_D^2 = \|f\|^2 + D(f)$, and

$$\text{Cap}(K) = \inf \{D(f) : f \in \mathcal{C}_K^\infty(-M, N), f|_K \geq 1\}.$$

Note that we have the universal factor 4 here and the isoperimetric constant $B_{\mathbb{B}}$ has a very compact form. We now need to compute the capacity only. The problem is that the capacity is usually not computable explicitly. For instance, at the moment, I do not know how to compute it for Schrödinger operators even for the elliptic operators having killings. Very lucky, we are able to compute the capacity for the one-dimensional elliptic operators. The result has a simple expression:

$$B_{\mathbb{B}} = \sup_{-M < x < y < N} [\nu(-M, x)^{-1} + \nu(y, N)^{-1}]^{-1} \|\mathbb{1}_{(x, y)}\|_{\mathbb{B}}.$$

It looks strange to have double inverse here. So, making inverse in both sides, we get

$$B_{\mathbb{B}}^{-1} = \inf_{-M < x < y < N} [\nu(-M, x)^{-1} + \nu(y, N)^{-1}] \|\mathbb{1}_{(x, y)}\|_{\mathbb{B}}^{-1}.$$

Applying this result to $\mathbb{B} = L^1(\mu)$, we obtain the solution to the DD-case: $\lambda^{\text{DD}} = A_{L^1(\mu)}^{-1}$ and

$$(\kappa^{\text{DD}})^{-1} = B_{L^1(\mu)}^{-1} = \inf_{-M < x < y < N} [\nu(-M, x)^{-1} + \nu(y, N)^{-1}] \mu(x, y)^{-1}.$$

3.4 The final step

Applying the last result to the dual process and using Proposition 3.2, we have not only

$$(\kappa^{*\text{DD}})^{-1}/4 \leq \lambda^{\text{NN}} = \lambda_S = \lambda^{*\text{DD}} \leq (\kappa^{*\text{DD}})^{-1},$$

but also

$$\begin{aligned} (\kappa^{*\text{DD}})^{-1} &= \inf_{x < y} [\nu^*(-M, x)^{-1} + \nu^*(y, N)^{-1}] \mu^*(x, y)^{-1} \\ &= \inf_{x < y} [\mu(-M, x)^{-1} + \mu(y, N)^{-1}] \nu(x, y)^{-1} \\ &= (\kappa^{\text{NN}})^{-1}. \end{aligned}$$

This finishes the proof of the main assertion of Theorem 2.2.

3.5 Summery of the proof

Here is the summery of our proof. First, by a change of the topology, we reduce the study on $\varepsilon_{\max}$ to λ^{NN} . Then, by coupling, we reduce λ^{NN} to λ_S . Next, by duality, we reduce λ_S to $\lambda^{*\text{DD}}$. We use capacitary method to compute $\lambda^{*\text{DD}}$. Finally, we use duality again to come back to λ^{NN} . Recall that our original purpose is using $\lambda_1 = \lambda^{\text{NN}}$ to study the phase transition, a basic topic in the study on interacting particle systems (abbrev. IPS). It is very interesting that we now have an opposite interaction. We use the main tools (coupling and duality) developed in the study on IPS to investigate a very classical problem and produce an interesting result.

4 Improvements

The basic estimates given in Theorem 2.2 can be further improved. For half-line at least, we have actually an approximating procedure for each of the principal eigenvalues. Refer to [6, 9] and references therein. Moreover, one may approach the whole line by half-lines. Here we consider an additional method but concentrate on λ^{DD} and λ^{NN} only. As will be seen soon, the resulting bounds are much more complicated, less simple and less symmetry, than those given in Theorem 2.2.

Let us begin with a simpler but effective result.

Proposition 4.1 We have

$$\lambda^{\text{DD}} \leq (\bar{\kappa}^{\text{DD}})^{-1} \leq (\kappa^{\text{DD}})^{-1}$$

and

$$\lambda^{\text{NN}} \leq (\bar{\kappa}^{\text{NN}})^{-1} \leq (\kappa^{\text{NN}})^{-1},$$

where

$$\begin{aligned} (\bar{\kappa}^{\text{DD}})^{-1} &= \inf_{x < y} \left(\nu(-M, x)^{-1} + \nu[y, N)^{-1} \right) \times \\ &\quad \times \left\{ \mu(x, y) + \int_{-M}^x \mu(dz) \left[1 - \frac{\nu(z, x)}{\nu(-M, x)} \right]^2 + \int_y^N \mu(dz) \left[1 - \frac{\nu(y, z)}{\nu(y, N)} \right]^2 \right\}^{-1}, \\ (\bar{\kappa}^{\text{NN}})^{-1} &= \inf_{x < y} \left(\mu(-M, x)^{-1} + \mu[y, N)^{-1} \right) \times \\ &\quad \times \left\{ \nu(x, y) + \int_{-M}^x \nu(dz) \left[1 - \frac{\mu(z, x)}{\mu(-M, x)} \right]^2 + \int_y^N \nu(dz) \left[1 - \frac{\mu(y, z)}{\mu(y, N)} \right]^2 \right\}^{-1}. \end{aligned}$$

Note that if $\nu(-M, N) < \infty$ which is not assumed in Proposition 4.1, then the last two terms in $\{\cdots\}$ in the expression of $(\bar{\kappa}^{\text{DD}})^{-1}$ can be written as

$$\nu(-M, x)^{-2} \int_{-M}^x \mu(dz) \nu(-M, z)^2 + \nu(y, N)^{-2} \int_y^N \mu(dz) \nu(z, N)^2.$$

Otherwise, this expression may be meaningless. Similar comment is meaningful for $(\bar{\kappa}^{\text{NN}})^{-1}$.

Proof. Fix $x < y$. Applying $\lambda^{\text{DD}} \leq D(f)/\mu(f^2)$ to the test function

$$f(z) = \begin{cases} \frac{\nu(y, N)}{\nu(-M, x)} \nu(-M, z \wedge x), & z \leq y \\ \nu(z, N), & z \geq y. \end{cases}$$

we obtain $\lambda^{\text{DD}} \leq (\bar{\kappa}^{\text{DD}})^{-1}$. By duality, we obtain the assertion for $\bar{\kappa}^{\text{NN}}$. Refer to the remark after the proof of [9; Theorem 8.2] for more details. $\square$

To improve the lower estimate Theorem 2.2, we need more work. For a given $f \in \mathcal{C}(-M, N)$ with $f|_{(-M, N)} > 0$, define

$$h^-(z) = h_f^-(z) = \nu\left(\mu(\mathbb{1}_{(\cdot, \theta)} f) \mathbb{1}_{(-M, z)}\right) = \int_{-M}^z e^{-C(x)} dx \int_x^\theta \frac{e^C f}{a}, \quad z \leq \theta, \quad (25)$$

$$h^+(z) = h_f^+(z) = \nu\left(\mu(\mathbb{1}_{(\theta, \cdot)} f) \mathbb{1}_{(z, N)}\right) = \int_z^N e^{-C(x)} dx \int_\theta^x \frac{e^C f}{a}, \quad z > \theta, \quad (26)$$

i.e. (by exchanging the order of the integrals),

$$\begin{aligned} h^-(z) &= \mu(f\nu(-M, \cdot \wedge z)) = \mu\left(f\nu(-M, \cdot) \mathbb{1}_{(-M, z)}\right) + \mu\left(f \mathbb{1}_{(z, \theta)}\right) \nu(-M, z), \quad z \leq \theta, \\ h^+(z) &= \mu(f\nu(\cdot \vee z, N)) = \mu\left(f\nu(\cdot, N) \mathbb{1}_{(z, N)}\right) + \mu\left(f \mathbb{1}_{(\theta, z)}\right) \nu(z, N), \quad z > \theta, \end{aligned}$$

where $x \wedge y = \min\{x, y\}$, $x \vee y = \max\{x, y\}$, and $\theta = \theta(f) \in (-M, N)$ is the unique root of the equation:

$$h^-(\theta) = h^+(\theta)$$

provided $h_f^\pm < \infty$. Next, define $II^\pm(f) = h^\pm/f$.

Theorem 4.2 (Variational formula) Let a and b be continuous and $a > 0$ on $(-M, N)$.

(1) Assume that $\nu(-M, N) < \infty$. Using the notation above, we have

$$\lambda^{\text{DD}} = \sup_{f \in \mathcal{C}_+} \left[\inf_{z \in (-M, \theta)} II^-(f)(z)^{-1} \right] \bigwedge \left[\inf_{z \in (\theta, N)} II^+(f)(z)^{-1} \right], \quad (27)$$

where $\mathcal{C}_+ = \{f \in \mathcal{C}(-M, N) : f > 0 \text{ on } (-M, N)\}$.

(2) Assume that $\mu(-M, N) < \infty$. Then (27) holds replacing λ^{DD} by λ^{NN} provided in definition of $h^\pm$, μ and ν are exchanged.

Proof. By duality, it suffices to prove the first assertion.

(a) Without loss of generality, assume that $h_f^\pm < \infty$. Otherwise, the assertion is trivial. First, we prove “ $\geq$ ”. Let

$$h(z) = \begin{cases} h^-(z), & z \leq \theta, \\ h^+(z), & z > \theta, \end{cases}$$

Clearly, $h|_{(-M,N)} > 0$ and $h \in \mathcal{C}(-M, N)$ in view of definition of θ . Next, note that

$$\begin{aligned} h^{-'}(x) &= e^{-C(x)} \int_x^\theta \frac{e^C}{a} f, & h^{-''}(x) &= e^{-C(x)} \left[-\frac{b}{a} \int_x^\theta \frac{e^C}{a} f - \frac{e^C}{a} f \right], & x < \theta; \\ h^{+'}(x) &= -e^{-C(x)} \int_\theta^x \frac{e^C}{a} f, & h^{+''}(x) &= e^{-C(x)} \left[\frac{b}{a} \int_\theta^x \frac{e^C}{a} f - \frac{e^C}{a} f \right], & x > \theta. \end{aligned}$$

Obviously, $h'(\theta \pm 0) = 0$. Since a , b and f are continuous and $a > 0$ on $(-M, N)$, we also have $h''(\theta+0) = h''(\theta-0)$ and so $h \in \mathcal{C}^2(-M, N)$. Therefore, by Barta's equality, we have

$$\begin{aligned} \lambda^{\text{DD}} &= \sup_{g \in \mathcal{F}} \inf_{z \in (-M, N)} \frac{-Lg}{g}(z) \\ &\geq \inf_{z \in (-M, N)} \frac{-Lh}{h}(z) \\ &= \left[\inf_{z \in (-M, \theta)} \frac{-Lh^-}{h^-}(z) \right] \bigwedge \left[\inf_{z \in (\theta, N)} \frac{-Lh^+}{h^+}(z) \right]. \end{aligned}$$

Now, by (7), required assertion follows by a simple computation.

(b) Next, we show that the equality sign in (27) holds. The assertion becomes trivial if $\lambda^{\text{DD}} = 0$. Otherwise, the eigenfunction g of λ^{DD} should be unimodal (which seems known in the Sturm–Liouville theory and is proved in the discrete context [9; Proposition 7.14]. Actually, the discrete case is even more complex since the eigenfunction can be a simple echelon, not necessarily unimodal). By setting $f = g$ and θ to be the maximum point of g ($g'(\theta) = 0$), it follows that $II^\pm(f)^{-1} \equiv \lambda^{\text{DD}}$ and hence the equality sign holds. $\square$

We now introduce a typical application of Theorem 4.2. Fix $x < y$. Define

$$f^{x,y}(s) = \begin{cases} \sqrt{\frac{\nu(y, N)}{\nu(-M, x)}} \nu(-M, s \wedge x), & s \leq y \\ \sqrt{\nu(s, N)}, & s \geq y \end{cases}$$

and set

$$\underline{\kappa}^{\text{DD}} = \inf_{x < y} \left[\sup_{z \in (-M, \theta)} II^-(f^{x,y})(z) \right] \bigvee \left[\sup_{z \in (\theta, N)} II^+(f^{x,y})(z) \right].$$

By exchanging μ and ν , we obtain $\underline{\kappa}^{\text{NN}}$. Now, by Theorem 4.2, we have the following result.

Corollary 4.3 Under assumptions of Theorem 4.2, we have

$$\lambda^{\text{DD}} \geq (\underline{\kappa}^{\text{DD}})^{-1} \quad \text{and} \quad \lambda^{\text{NN}} \geq (\underline{\kappa}^{\text{NN}})^{-1}.$$

We remark that the assumption in part (1) of Theorem 4.2 is necessary for DD-case (cf. (13)). Recall that (27) is a complete variational formula for the lower estimates of λ^{DD} . Starting at $f_1 = f$ used in Corollary 4.3, replacing f and h used in Theorem 4.2 by f_{n-1} and f_n , respectively, we obtain an approximating procedure from below for λ^{DD} . Dually, we can obtain a variational formula for the upper estimates of λ^{DD} and an approximating procedure from above. Here we omit all of the details. The same remark is meaningful for λ^{NN} , which is especially interesting since here we do not use the property that $\mu(f) = 0$ for the test function f . The new difficulty of (27) is that $\theta(f)$ may not be computable analytically. This costs a question to prove that $\underline{\kappa}^{\text{DD}} \leq 4\kappa^{\text{DD}}$ which should be true in view of our knowledge on the half-line, and is illustrated by examples in the next section. It is noticeable that the method works for the whole line and the use of $\theta(f)$ is essentially different from what used in the splitting technique. Finally, we mention that the method used here is meaningful for birth-death processes, refer to [9; Lemma 7.12].

For convenience in practice, we express $h^\pm$ used in Corollary 4.3 more explicitly. Let $\nu_-(s) = \nu(-M, s)$ and $\nu_+(s) = \nu(s, N)$ for simplicity. Then

$$f(s) = f^{x,y}(s) = \begin{cases} \sqrt{\nu_+(y)\nu_-(s)}/\sqrt{\nu_-(x)}, & s \leq x \\ \sqrt{\nu_+(y)}, & x \leq s \leq y \\ \sqrt{\nu_+(s)}, & s \geq y, \end{cases} \quad (28)$$

and

$$h^-(z) = \mu\left(f\nu_-\mathbb{1}_{(-M,z)}\right) + \nu_-(z)\mu\left(f\mathbb{1}_{(z,\theta)}\right), \quad z \leq \theta, \quad (29)$$

$$h^+(z) = \mu\left(f\nu_+\mathbb{1}_{(z,N)}\right) + \nu_+(z)\mu\left(f\mathbb{1}_{(\theta,z)}\right), \quad z \geq \theta. \quad (30)$$

We now consider the typical case that $\theta \in [x, y]$. Then,

$$\begin{aligned} h^-(\theta) &= \sqrt{\frac{\nu_+(y)}{\nu_-(x)}}\mu\left(\nu_-^{3/2}\mathbb{1}_{(-M,x)}\right) + \sqrt{\nu_+(y)}\mu\left(\nu_-\mathbb{1}_{(x,\theta)}\right), \\ h^+(\theta) &= \mu\left(\nu_+^{3/2}\mathbb{1}_{(y,N)}\right) + \sqrt{\nu_+(y)}\mu\left(\nu_+\mathbb{1}_{(\theta,y)}\right). \end{aligned}$$

Hence the equation $h^-(\theta) = h^+(\theta)$ becomes

$$\begin{aligned} &\frac{1}{\sqrt{\nu_-(x)}}\mu\left(\nu_-^{3/2}\mathbb{1}_{(-M,x)}\right) + \mu\left(\nu_-\mathbb{1}_{(x,\theta)}\right) \\ &= \frac{1}{\sqrt{\nu_+(y)}}\mu\left(\nu_+^{3/2}\mathbb{1}_{(y,N)}\right) + \mu\left(\nu_+\mathbb{1}_{(\theta,y)}\right), \quad \theta \in [x, y]. \end{aligned} \quad (31)$$

Furthermore, by some computations, we obtain the ratio $h^\pm/f^{x,y}$ as follows. We have for z : $z \leq x \leq \theta \leq y$ that

$$\begin{aligned} II^-(f^{x,y})(z) &= \frac{1}{\sqrt{\nu_-(z)}} \mu\left(\nu_-^{3/2} \mathbb{1}_{(-M,z)}\right) + \sqrt{\nu_-(z)} \mu\left(\sqrt{\nu_-} \mathbb{1}_{(z,x)}\right) \\ &\quad + \sqrt{\nu_-(z)\nu_-(x)} \mu(x, \theta), \end{aligned} \quad (32)$$

and for z : $z \geq y \geq \theta$ that

$$\begin{aligned} II^+(f^{x,y})(z) &= \frac{1}{\sqrt{\nu_+(z)}} \mu\left(\nu_+^{3/2} \mathbb{1}_{(z,N)}\right) + \sqrt{\nu_+(z)} \mu\left(\sqrt{\nu_+} \mathbb{1}_{(y,z)}\right) \\ &\quad + \sqrt{\nu_+(z)\nu_+(y)} \mu(\theta, y). \end{aligned} \quad (33)$$

Note that by (25) and (26), h^- is increasing on $[x, \theta]$ and h^+ is decreasing on $[\theta, y]$. Since $f^{x,y}$ is a constant on $[x, y]$, it follows that

$$\max_{z \in [x, \theta]} \frac{h^-(z)}{f^{x,y}(z)} = \frac{h^-(\theta)}{f^{x,y}(x)} \quad \text{and} \quad \max_{z \in [\theta, y]} \frac{h^+(z)}{f^{x,y}(z)} = \frac{h^+(\theta)}{f^{x,y}(x)}.$$

By assumption, $h^-(\theta) = h^+(\theta)$. Hence

$$\begin{aligned} \max_{z \in [x, \theta]} II^-(f^{x,y})(z) &= \max_{z \in [\theta, y]} II^+(f^{x,y})(z) = \frac{h^-(\theta)}{f^{x,y}(x)} \\ &= \frac{1}{\sqrt{\nu_-(x)}} \mu\left(\nu_-^{3/2} \mathbb{1}_{(-M,x)}\right) + \mu\left(\nu_- \mathbb{1}_{(x,\theta)}\right). \end{aligned} \quad (34)$$

Thus, for computing $\underline{\kappa}^{\text{DD}}$, by (32)–(34), we arrive at

$$\begin{aligned} &\left[\sup_{z \in (-M, \theta)} II^-(f^{x,y})(z) \right] \bigvee \left[\sup_{z \in (\theta, N)} II^+(f^{x,y})(z) \right] \\ &= \sup_{z \in (-M, x)} \left[\frac{1}{\sqrt{\nu_-(z)}} \mu\left(\nu_-^{3/2} \mathbb{1}_{(-M,z)}\right) + \sqrt{\nu_-(z)} \mu\left(\sqrt{\nu_-} \mathbb{1}_{(z,x)}\right) \right. \\ &\quad \left. + \sqrt{\nu_-(z)\nu_-(x)} \mu(x, \theta) \right] \\ &\quad \bigvee \left[\frac{1}{\sqrt{\nu_-(x)}} \mu\left(\nu_-^{3/2} \mathbb{1}_{(-M,x)}\right) + \mu\left(\nu_- \mathbb{1}_{(x,\theta)}\right) \right] \\ &\quad \bigvee \sup_{z \in (y, N)} \left[\frac{1}{\sqrt{\nu_+(z)}} \mu\left(\nu_+^{3/2} \mathbb{1}_{(z,N)}\right) + \sqrt{\nu_+(z)} \mu\left(\sqrt{\nu_+} \mathbb{1}_{(y,z)}\right) \right. \\ &\quad \left. + \sqrt{\nu_+(z)\nu_+(y)} \mu(\theta, y) \right]. \end{aligned} \quad (35)$$

Finally, let (x^*, y^*, θ^*) solve equation (31) and two more equations modified from (35) ignoring its left-hand side and replacing the last two “ $\bigvee$ ” with “ $=$ ”. Then we have

$$\underline{\kappa}^{\text{DD}} = \frac{1}{\sqrt{\nu_-(x^*)}} \mu\left(\nu_-^{3/2} \mathbb{1}_{(-M,x^*)}\right) + \mu\left(\nu_- \mathbb{1}_{(x^*, \theta^*)}\right). \quad (36)$$

5 Examples

This section illustrates the application of the basic estimates given in Theorem 2.2 and the improvements given in Proposition 4.1 and Corollary 4.3.

Example 5.1 (OU-processes) The state space is $\mathbb{R}$ and the operator is

$$L = \frac{1}{2} \left(\frac{d^2}{dx^2} - 2x \frac{d}{dx} \right).$$

This is a typical example of the use of special functions. It has discrete eigenvalues $\lambda_n = n$ with eigenfunctions (Hermite polynomials)

$$g_n(x) = (-1)^n e^{x^2} \frac{d^n}{dx^n} (e^{-x^2}), \quad n \geq 0.$$

Then, we have $(\kappa^{\text{DD}})^{-1} = \lambda_0 = 0$, $\lambda^{\text{NN}} = \lambda_1 = 1$ with eigenfunction $g(x) = x$. To compute κ^{NN} , noting that the operator, the eigenfunction are all symmetric with respect to 0 and so does κ^{NN} , one can split the whole line into two parts $(-\infty, 0)$ and $(0, \infty)$ with common Dirichlet boundary at 0. This simplifies the computation and leads to $(\kappa^{\text{NN}})^{-1} = (\kappa^{\text{DN}})^{-1} \approx 2.1$. Note that $g'(x) \equiv 1$ but $\lim_{|x| \rightarrow \infty} (e^C g')(x) = 0$.

For the half-space $(0, \infty)$, as we have just mentioned, $\lambda^{\text{DN}} = \lambda^{\text{DD}} = 1$ with $g(x) = x$, $(\kappa^{\text{DN}})^{-1} = (\kappa^{\text{DD}})^{-1} \approx 2.1$. For λ^{NN} , the symmetry in the whole line is lost. We have $\lambda^{\text{NN}} = 2$ with $g(x) = -1 + 2x^2$, $(\kappa^{\text{NN}})^{-1} \approx 4.367$ which is achieved at $(x, y) \approx (0.316, 1.185)$. Note that $\lim_{x \rightarrow \infty} g'(x) = \infty$ but $\lim_{x \rightarrow \infty} (e^C g')(x) = 0$.

To study $\bar{\kappa}^{\text{NN}}$, recall that we can reduce the NN-case to the DD-one by an exchange of μ and ν . By Proposition 3.1, we have $(\bar{\kappa}^{\text{NN}})^{-1} \approx 2.6$. By Corollary 4.3 and (36), we obtain $(\underline{\kappa}^{\text{NN}})^{-1} \approx 1.83$ with $(x^*, y^*, \theta^*) \approx (0.6405, 0.938, 0.721194)$. For the last conclusion, we use a direct search starting from $(x, y) \approx (0.316, 1.185)$ which leads to κ^{NN} in the last paragraph. The ratio becomes $2.6/1.83 \approx 1.42 < 4$. We mention that similar estimates can also be obtained by using a different approximating procedure in parallel with [9; Theorem 6.3]. Refer to [5; Footnotes 12 and 14].

The following examples are often illustrated in the textbooks on ordinary differential equations, see for instance Hartman (1982), §11.1.

Example 5.2 The equation

$$u'' + \sigma^2 u = 0 \quad (\sigma \neq 0)$$

has the general solution

$$u = c_1 \cos(\sigma x) + c_2 \sin(\sigma x).$$

From this, it should be clear that for the operator $L = d^2/dx^2$ with finite state space (α, β) , we have

$$\begin{aligned}\lambda^{\text{DD}} &= \left(\frac{\pi}{\beta - \alpha}\right)^2, & g(x) &= \sin\left(\frac{\pi(x - \alpha)}{\beta - \alpha}\right); \\ \lambda^{\text{NN}} &= \left(\frac{\pi}{\beta - \alpha}\right)^2, & g(x) &= \cos\left(\frac{\pi(x - \alpha)}{\beta - \alpha}\right); \\ \lambda^{\text{DN}} &= \left(\frac{\pi}{2(\beta - \alpha)}\right)^2, & g(x) &= \sin\left(\frac{\pi(x - \alpha)}{2(\beta - \alpha)}\right); \\ \lambda^{\text{ND}} &= \left(\frac{\pi}{2(\beta - \alpha)}\right)^2, & g(x) &= \cos\left(\frac{\pi(x - \alpha)}{2(\beta - \alpha)}\right).\end{aligned}$$

The corresponding estimates are as follows.

$$(\kappa^{\text{DD}})^{-1} = (\kappa^{\text{NN}})^{-1} = \left(\frac{4}{\beta - \alpha}\right)^2, \quad (\kappa^{\text{DN}})^{-1} = (\kappa^{\text{ND}})^{-1} = \left(\frac{2}{\beta - \alpha}\right)^2.$$

Note that by symmetry, the DD- and NN-cases can be split at $\theta = (\alpha + \beta)/2$ into the DN- and ND-cases. One can then approach λ^{DD} and λ^{NN} by using the known approximating method for λ^{DN} and λ^{ND} (cf. [5; Theorem 1.2]). However, as an illustration of Theorem 4.2 and Corollary 4.3, we now compute $\bar{\kappa}^{\text{DD}}$ and $\underline{\kappa}^{\text{DD}}$.

Consider first the simpler interval $(\alpha, \beta) = (0, 1)$. Since $\mu = \nu = dx$, by symmetry, one may choose $y = 1 - x$. Then $x < 1/2$ and

$$\begin{aligned}(\bar{\kappa}^{\text{DD}})^{-1} &= \inf_{x \in (0, 1/2)} \frac{2}{x} \left[1 - 2x + x^{-2} \int_0^x z^2 dz + x^{-2} \int_{1-x}^1 (1-z)^2 dz \right]^{-1} \\ &= \inf_{x \in (0, 1/2)} \frac{6}{3x - (2x)^2} \\ &= \frac{32}{3} \quad (\text{with } x = 3/8).\end{aligned}$$

To compute $\underline{\kappa}^{\text{DD}}$, set again $y = 1 - x$ with $x \in (0, 1/2)$. Then, the test function $f^{x,y}$ becomes

$$f^x(s) = \begin{cases} \sqrt{s \wedge x} & s \leq 1 - x \\ \sqrt{1 - s} & s \in (1 - x, 1). \end{cases}$$

By symmetry again, we have $\theta = 1/2$. Fix $x \in (0, 1/2)$. For convenience, we express f^x as (f_1, f_2) : $f_1(s) = \sqrt{s}$ for $s \in [0, x]$ and $f_2(s) = \sqrt{x}$ for $s \in [x, 1/2]$. Then by (29) with $\nu_-(s) = s$, we have $h^- = (h_1^-(z), h_2^-(z))$:

$$\begin{aligned}h_1^-(z) &= \int_0^z f_1(s) s ds + z \left[\int_z^x f_1 + \int_x^{1/2} f_2 \right], & z &\in [0, x] \\ h_2^-(z) &= \left[\int_0^x f_1(s) s ds + \int_x^z f_2(s) s ds \right] + z \int_z^{1/2} f_2, & z &\in [x, 1/2].\end{aligned}$$

Hence by (32), we have

$$\Pi^-(f^x)(z) = \frac{h^-(z)}{f^x(z)} = \begin{cases} \left(-\frac{1}{3}x^{3/2} + \frac{1}{2}x^{1/2}\right)\sqrt{z} - \frac{4}{15}z^2, & z \in [0, x], \\ \frac{1}{10}(5z(1-z) - x^2), & z \in [x, 1/2]. \end{cases}$$

Define

$$H(x) = -\frac{1}{3}x^{3/2} + \frac{1}{2}x^{1/2} \quad \text{and} \quad \gamma(z) = H(x)\sqrt{z} - \frac{4}{15}z^2.$$

Then

$$\gamma'(z) = \frac{H(x)}{2\sqrt{z}} - \frac{8}{15}z, \quad \gamma''(z) = -\frac{H(x)}{4z^{3/2}} - \frac{8}{15} < 0.$$

Hence γ achieves its maximum at

$$z^*(x) = \left(\frac{15}{16}H(x)\right)^{2/3}.$$

Furthermore,

$$\gamma(z^*(x)) = H(x)\left(\frac{15}{16}H(x)\right)^{1/3} - \frac{4}{15}\left(\frac{15}{16}H(x)\right)^{4/3} = \frac{3}{8}\left(\frac{15}{2}\right)^{1/3}H(x)^{4/3}.$$

Note that $z^*(x) \leq x$ iff $x \geq 5/14$. Besides, on the subinterval $[x, 1/2]$, $h^-(z)/f^x(z)$ has maximum $1/8 - x^2/10$ by (34). Solving the equation

$$\frac{3}{8}\left(\frac{15}{2}\right)^{1/3}H(x)^{4/3} = \frac{1}{8} - \frac{1}{10}x^2, \quad x \in (5/14, 1/2),$$

we obtain $x^* \approx 0.436273$ and then

$$\inf_{x \in (5/14, 1/2)} \sup_{z \leq 1/2} \frac{h^-(z)}{f^x(z)} = \gamma(z^*(x^*)) \approx 0.105967.$$

From these facts and (36), we conclude that

$$(\kappa^{\text{DD}})^{-1} \approx 1/0.105967 \approx 9.43693.$$

By the way, we mention that a similar but simpler study shows that

$$\inf_{x \in (0, 5/14)} \sup_{z \leq 1/2} \frac{h^-(z)}{f^x(z)} = \frac{1}{8}.$$

This shows that to get a less sharp lower bound $1/8$, the computation becomes much simpler. It needs to study the extremal case that $x = 0$ only; the corresponding test function becomes $f^x \equiv 1$. Return to the original interval (α, β) , by Proposition 4.1 and Corollary 4.3, we obtain

$$\frac{8}{(\beta - \alpha)^2} < \frac{9.4369}{(\beta - \alpha)^2} < \lambda^{\text{DD}} = \left(\frac{\pi}{\beta - \alpha}\right)^2 \leq \frac{32}{3(\beta - \alpha)^2} = \frac{2}{3}\left(\frac{4}{\beta - \alpha}\right)^2.$$

The ratio becomes $\frac{32}{3}/9.4369 \approx 1.13$. The same assertion holds if λ^{DD} is replaced by λ^{NN} because of the symmetry.

It is a good chance to discuss the approximating procedure remarked after Corollary 4.3. Here we consider the lower estimate only. Replacing $f^x = (f_1, f_2)$ by (h_1^-, h_2^-) , one produces a new (h_1^-, h_2^-) and then a new $II^-(f)$ which provides a new lower bound. By using this procedure twice with fixed $\theta = 1/2$ and $x = x^* \approx 0.436273$, we obtain successively the following lower bounds:

$$\frac{9.80392}{(\beta - \alpha)^2}, \quad \frac{9.86193}{(\beta - \alpha)^2}.$$

Clearly, they are quite close to the exact value of λ^{DD} and λ^{NN} :

$$\frac{\pi^2}{(\beta - \alpha)^2} \approx \frac{9.8696}{(\beta - \alpha)^2}.$$

Example 5.3 By a substitute $u = ze^{-bx/2}$, the equation

$$u'' + bu' + \gamma u = 0 \quad (b, \gamma \text{ are real constants})$$

is reduced to

$$z'' + \sigma^2 z = 0 \quad (\sigma^2 = \gamma - b^2/4).$$

From the last example, it follows that the equation has general solutions

$$u = \begin{cases} e^{-bx/2}(c_1 + c_2 x) & \text{if } \gamma = b^2/4 \\ c_1 e^{\xi_1 x} + c_2 e^{\xi_2 x} & \text{if } \gamma < b^2/4 \\ e^{-bx/2} \left(c_1 \cos(x\sqrt{\gamma - b^2/4}) + c_2 \sin(x\sqrt{\gamma - b^2/4}) \right) & \text{if } \gamma > b^2/4, \end{cases}$$

where ξ_1, ξ_2 are solution to the equation

$$\xi^2 + b\xi + \gamma = 0.$$

Thus, for the operator $L = d^2/dx^2 + b d/dx$ (b is a constant) with state space $(0, \infty)$, we have the following principal eigenfunctions

- $g(x) = (2/b+x)e^{-bx/2}$ and $g(x) = xe^{-bx/2}$ in ND- and DD-cases, respectively, when $b > 0$;
- $g(x) = xe^{-bx/2}$ and $g(x) = (1 + bx/2)e^{-bx/2}$ in DN- and NN-cases, respectively, when $b < 0$.

In each of these cases, we have the principal eigenvalue $\lambda^\# = b^2/4$ and $(\kappa^\#)^{-1} = b^2$. Moreover, $(\kappa^{\text{DD}})^{-1}, (\kappa^{\text{NN}})^{-1} = b^2/2$. Clearly, the lower estimate $(\kappa^\#)^{-1}/4$ is sharp in all cases.

Example 5.4 (Cauchy–Euler equation) Consider the operator

$$L = x^2 \frac{d^2}{dx^2} + bx \frac{d}{dx},$$

where b is a constant. By a change of variable $x = e^y$, the equation

$$x^2 u'' + bxu' + \gamma u = 0 \quad (b, \gamma \text{ are constants})$$

is reduced to the last example:

$$\frac{d^2 u}{dy^2} + (b-1) \frac{du}{dy} + \gamma u = 0.$$

Hence the original equation has general solutions

$$u = \begin{cases} x^{(1-b)/2} (c_1 + c_2 \log x) & \text{if } \gamma = (1-b)^2/4 \\ c_1 x^{\xi_1} + c_2 x^{\xi_2} & \text{if } \gamma < (1-b)^2/4 \\ x^{(1-b)/2} \left(c_1 \cos(\sqrt{\gamma - (1-b)^2/4} \log x) + c_2 \sin(\sqrt{\gamma - (1-b)^2/4} \log x) \right) & \text{if } \gamma > (1-b)^2/4, \end{cases}$$

where ξ_1, ξ_2 are solution to the equation $\xi^2 + (b-1)\xi + \gamma = 0$:

$$\xi_1, \xi_2 = (1-b)/2 \pm \sqrt{(1-b)^2/4 - \gamma}.$$

Here we have used Euler's formula:

$$x^{i\sqrt{\xi}} = e^{i\sqrt{\xi} \log x} = \cos(\sqrt{\xi} \log x) + i \sin(\sqrt{\xi} \log x).$$

In particular,

(1) when $b = 0$, we have solutions

$$u = \begin{cases} \sqrt{x} (c_1 + c_2 \log x) & \text{if } \gamma = 1/4 \\ c_1 x^{\xi_1} + c_2 x^{\xi_2} & \text{if } \gamma < 1/4 \\ \sqrt{x} \left(c_1 \cos(\sqrt{\gamma - 1/4} \log x) + c_2 \sin(\sqrt{\gamma - 1/4} \log x) \right) & \text{if } \gamma > 1/4. \end{cases}$$

Now, corresponding to $\gamma = 1/4$, we have

$$\lambda^{\text{DN}} = \frac{1}{4}, \quad g(x) = \begin{cases} \sqrt{x} & \text{if the state space is } (0, \infty) \\ \sqrt{x} \log \sqrt{x} & \text{if the state space is } (1, \infty). \end{cases}$$

The first case is the original Hardy's inequality. Corresponding to $\gamma = 1/4$ again but for state space $(1, \infty)$, we have

$$\lambda^{\text{NN}} = \frac{1}{4}, \quad g(x) = \sqrt{x} (\log \sqrt{x} - 1).$$

Here $\lim_{x \rightarrow \infty} (e^C g')(x) = \lim_{x \rightarrow \infty} g'(x) = 0$. We have $(\kappa^{\text{DN}})^{-1}, (\kappa^{\text{NN}})^{-1} = 1$, $(\bar{\kappa}^{\text{DN}})^{-1}, (\bar{\kappa}^{\text{NN}})^{-1} = 1/2$, respectively. The lower estimate $(\kappa^\#)^{-1}/4$ is sharp in each case. The DN-case is actually a special one of the last example.

(2) When $b = 1$, for finite state space $(1, N)$ with Dirichlet boundaries, we have

$$\lambda_n = \left(\frac{n\pi}{\log N} \right)^2, \quad g(x) = \sin \left(\frac{n\pi}{\log N} \log x \right), \quad n \geq 1.$$

In particular,

$$\lambda^{\text{DD}} = \left(\frac{\pi}{\log N} \right)^2, \quad g(x) = \sin \left(\frac{\pi}{\log N} \log x \right).$$

Next, for Neumann boundaries, we have

$$\lambda^{\text{NN}} = \left(\frac{\pi}{\log N} \right)^2, \quad g(x) = \cos \left(\frac{\pi}{\log N} \log x \right).$$

In both cases, we have $(\kappa^{\text{DD}})^{-1}, (\kappa^{\text{NN}})^{-1} = (4/\log N)^2$. Besides, we have

$$\begin{aligned} \lambda^{\text{DN}} &= \left(\frac{\pi}{2 \log N} \right)^2, & g(x) &= \sin \left(\frac{\pi}{2 \log N} \log x \right); \\ \lambda^{\text{ND}} &= \left(\frac{\pi}{2 \log N} \right)^2, & g(x) &= \cos \left(\frac{\pi}{2 \log N} \log x \right). \end{aligned}$$

In these cases, we have $(\kappa^{\text{DN}})^{-1}, (\kappa^{\text{ND}})^{-1} = (2/\log N)^2$. Note that the present case can be reduced to Example 5.2 under the change of variable $x = e^y$, the results here can be obtained from Example 5.2 replacing $(\alpha - \beta)^2$ by $\log^2 N$. In view of this, we also have

$$(\bar{\kappa}^{\text{DD}})^{-1} = (\bar{\kappa}^{\text{NN}})^{-1} = \frac{32}{3 \log^2 N}, \quad (\underline{\kappa}^{\text{DD}})^{-1} = (\underline{\kappa}^{\text{NN}})^{-1} \approx \frac{9.4369}{\log^2 N}.$$

6 Appendix

The next result is a generalization of [9; Proposition 1.2].

Proposition 6.1 Let $P_t(x, \cdot)$ be symmetric and have density $p_t(x, y)$ with respect to μ . Suppose that the diagonal elements $p_s(\cdot, \cdot) \in L_{\text{loc}}^{1/2}(\mu)$ for some $s > 0$ and a set $\mathcal{H}$ of bounded functions with compact support is dense in $L^2(\mu)$. Then $\lambda_0 = \varepsilon_{\max}$.

Proof. The proof is similar to the ergodic case (cf. [6; Section 8.3] and [9; proof of Theorem 7.4]), and is included here for completeness.

(a) Certainly, the inner product and norm here are taken with respect to μ . First, we have

$$\begin{aligned}
P_t(x, K) &= P_s P_{t-s} \mathbb{1}_K(x) \\
&= \int \mu(dy) \frac{dP_s(x, \cdot)}{d\mu}(y) P_{t-s} \mathbb{1}_K(y) \quad (\text{since } P_s \ll \mu) \\
&= \mu \left(\frac{dP_s(x, \cdot)}{d\mu} P_{t-s} \mathbb{1}_K \right) \\
&= \mu \left(\mathbb{1}_K P_{t-s} \frac{dP_s(x, \cdot)}{d\mu} \right) \quad (\text{by symmetry of } P_t) \\
&\leq \sqrt{\mu(K)} \left\| P_{t-s} \frac{dP_s(x, \cdot)}{d\mu} \right\| \quad (\text{by Cauchy-Schwarz inequality}) \\
&\leq \sqrt{\mu(K)} \left\| \frac{dP_s(x, \cdot)}{d\mu} \right\| e^{-\lambda_0(t-s)} \quad (\text{by } L^2\text{-exponential convergence}) \\
&= \left(\sqrt{\mu(K)} p_{2s}(x, x) e^{\lambda_0 s} \right) e^{-\lambda_0 t} \quad (\text{by [6; (8.3)]}).
\end{aligned}$$

By assumption, the coefficient on the right-hand side is locally μ -integrable. This proves that $\varepsilon_{\max} \geq \lambda_0$.

(b) Next, for each $f \in \mathcal{K}$ with $\|f\| = 1$, we have

$$\begin{aligned}
\|P_t f\|^2 &= (f, P_{2t} f) \quad (\text{by symmetry of } P_t) \\
&\leq \|f\|_\infty \int_{\text{supp}(f)} \mu(dx) P_{2t}|f|(x) \\
&\leq \|f\|_\infty^2 \int_{\text{supp}(f)} \mu(dx) P_{2t}(x, \text{supp}(f)) \\
&\leq \|f\|_\infty^2 \int_{\text{supp}(f)} \mu(dx) c(x, \text{supp}(f)) e^{-2\varepsilon_{\max} t} \\
&=: C_f e^{-2\varepsilon_{\max} t}.
\end{aligned}$$

The technique used here goes back to Hwang et al. (2005).

(c) The constant C_f in the last line can be removed. Following Lemma 2.2 in Wang (2002), by the spectral representation theorem and the fact that $\|f\| = 1$, we have

$$\begin{aligned}
\|P_t f\|^2 &= \int_0^\infty e^{-2\lambda t} d(E_\lambda f, f) \\
&\geq \left[\int_0^\infty e^{-2\lambda s} d(E_\lambda f, f) \right]^{t/s} \quad (\text{by Jensen's inequality}) \\
&= \|P_s f\|^{2t/s}, \quad t \geq s.
\end{aligned}$$

Note that here the semigroup is allowed to be sub-Markovian. Combining this with (b), we have $\|P_s f\|^2 \leq C_f^{s/t} e^{-2\varepsilon_{\max} s}$. Letting $t \rightarrow \infty$, we obtain

$$\|P_s f\|^2 \leq e^{-2\varepsilon_{\max} s},$$

first for all $f \in \mathcal{K}$ and then for all $f \in L^2(\mu)$ with $\|f\| = 1$, because of the denseness of $\mathcal{K}$ in $L^2(\mu)$. Therefore, $\lambda_0 \geq \varepsilon_{\max}$. Combining this with (a), we complete the proof. $\square$

The main result (Theorem 2.2) of this paper is presented in the last section (section 10) of the paper [9], as an analog of birth-death processes. Paper [9], as well as [8] for φ^4 -model, is available on arXiv.org.

Acknowledgments. Research supported in part by the Creative Research Group Fund of the National Natural Science Foundation of China (No. 10721091), by the “985” project from the Ministry of Education in China. The author has been luckily invited by Professor Louis Chen three times with financial support to visit Singapore. Deep appreciation is given to him for his continuous encouragement and friendship in the past 30 years. Sections 2–4 of the paper are based on the talks presented in “Workshop on Stochastic Differential Equations and Applications” (December, 2009, Shanghai), “Chinese-German Meeting on Stochastic Analysis and Related Fields” (May, 2010, Beijing), and “From Markov Processes to Brownian Motion and Beyond — An International Conference in Memory of Kai-Lai Chung” (June, 2010, Beijing). The author acknowledges the organizers of the conferences: Professors Xue-Rong Mao; Zhi-Ming Ma and Michael Röckner; and the Organization Committee headed by Zhi-Ming Ma (Elton P. Hsu and Dayue Chen, in particular), for their kind invitation and financial support.

References

- [1] Chen, L.H.Y. (1985). Poincaré-type inequalities via stochastic integrals, *Z. Wahrsch. Verw. Gebiete.* 69, 251–277.
- [2] Chen, L.H.Y. and Luo, J.H. (1987). Characterization of probability distributions by Poincaré-type inequalities, *Ann. Inst. H. Poincaré Sect. B (N.S.)* 23, 91–110.
- [3] Chen, M.F. (1991). Exponential L^2 -convergence and L^2 -spectral gap for Markov processes, *Acta Math. Sin. New Ser.* 7(1), 19–37.
- [4] Chen, M.F. (2000). Explicit bounds of the first eigenvalue, *Sci. China (A)* 43(10), 1051–1059.
- [5] Chen, M.F. (2001). Variational formulas and approximation theorems for the first eigenvalue, *Sci. China (A)* 44(4), 409–418
The last three and related papers with some complements are collected in book [4] at the author’s homepage.
- [6] Chen, M.F. (2005a). *Eigenvalues, Inequalities, and Ergodic Theory*, Springer, London.
- [7] Chen, M.F. (2005b). Capacitary criteria for Poincaré-type inequalities, *Potential Analysis* 23(4), 303–322
- [8] Chen, M.F. (2008). Spectral gap and logarithmic Sobolev constant for continuous spin systems, *Acta Math. Sin. N.S.* 24(5), 705–736. Available via arXiv.org
- [9] Chen, M.F. (2010). Speed of stability for birth–death processes, *Front. Math. China* 5(3), 379–515.

- [10] Chen, M.F. and Wang, F.Y. (1997). Estimation of spectral gap for elliptic operators, *Trans. Amer. Math. Soc.* 349(3), 1239–1267.
- [11] Cox, J.T. and Rösler, U. (1983). A duality relation for entrance and exit laws for Markov processes, *Stoch. Proc. Appl.* 16, 141–156.
- [12] Feller, W. (1955). On second order differential operators, *Ann. Math.* 2nd Ser., 61(1), 90–105.
- [13] Fukushima, M. and Uemura, T. (2003). Capacitary bounds of measures and ultracontractivity of time changed processes, *J. Math. Pure et Appliquees* 82(5), 553–572.
- [14] Gurka, P. (1989). Generalized Hardy’s inequality for functions vanishing on both ends of the interval, unpublished.
- [15] Hansson, K. (1979). Imbedding theorems of Sobolev type in potential theory, *Math. Scand.* 45, 77–102.
- [16] Hardy, G.H. (1920). Note on a theorem of Hilbert. *Math. Zeitschr.* 6, 314–317.
- [17] Hartman, P. (1982). *Ordinary Differential Equations*, 2nd ed., Birkhäuser.
- [18] Hwang, C.R., Hwang-Ma, S.Y. and Sheu, S.J. (2005). Accelerating diffusions, *Ann. Appl. Prob.* 15(2), 1433–1444.
- [19] Maz’ya, V.G. (1985). *Sobolev Spaces*, Springer-Verlag.
- [20] Miclo, L. An example of application of discrete Hardy’s inequalities, *Markov Processes Relat. Fields* 5, 319–330.
- [21] Muckenhoupt, B. (1972). Hardy’s inequality with weights, *Studia Math.* XLIV, 31–38.
- [22] Opic, B. and Kufner, A. (1990). *Hardy-type Inequalities*, Longman, New York.
- [23] Siegmund, D. (1976). The equivalence of absorbing and reflecting barrier problems for stochastically monotone Markov processes, *Ann. Prob.* 4(6), 914–924.
- [24] Vondraček, Z. (1996). An estimate for the L^2 -norm of a quasi continuous function with respect to a smooth measure, *Arch. Math.* 67, 408–414.
- [25] Wang, F.Y. (2000). Functional inequalities, semigroup properties and spectrum estimates, *Infinite Dim. Anal., Quantum Probab. and related Topics* 3(2), 263–295.

School of Mathematical Sciences, Beijing Normal University, Laboratory of Mathematics and Complex Systems (Beijing Normal University), Ministry of Education, Beijing 100875, The People’s Republic of China.

E-mail: mfchen@bnu.edu.cn

Home page: http://math.bnu.edu.cn/~chenmf/main_eng.htm

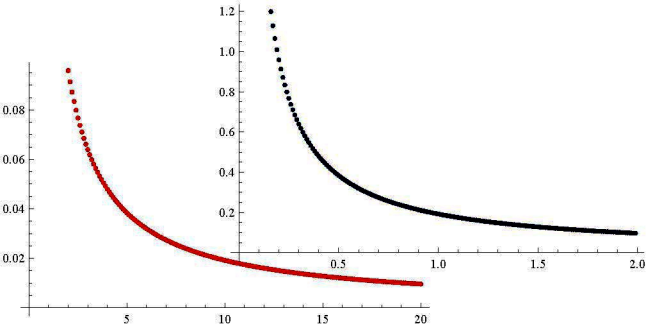


Figure 1-2 Solution of $x^* = x^*(b)$ when b varies on $(0, 2]$ (the curve on right) and on $[2, 20]$ (the curve on left), respectively.

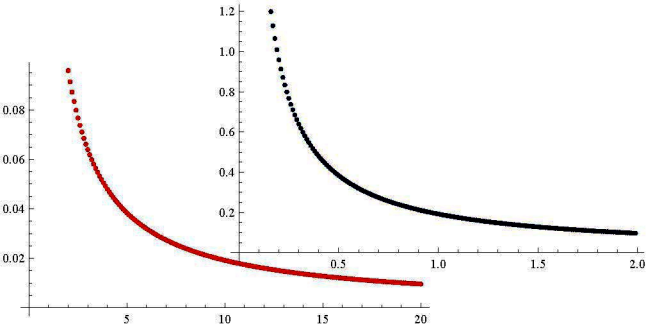


Figure 1-2 Solution of $x^* = x^*(b)$ when b varies on $(0, 2]$ (the curve on right) and on $[2, 20]$ (the curve on left), respectively.

General estimate of the first eigenvalue on manifolds

Mu-Fa Chen

(Beijing Normal University, Beijing 100875, China)

December 30, 2010

Abstract

Ten sharp lower estimates of the first non-trivial eigenvalue of Laplacian on compact Riemannian manifolds are reviewed and compared. An improved variational formula, a general common estimate, and a new sharp one are added. The best lower estimates are now updated. The new estimates provide a global picture of what one can expect by our approach.

1 Introduction

Let M be a compact, connected Riemannian manifold, without or with convex boundary ∂M . When $\partial M \neq \emptyset$, we adopt Neumann boundary condition. Next, let $\text{Ric}_M \geq K$ for some $K \in \mathbb{R}$. Denote by d and D , respectively, the dimension and diameter of M . We are interested in the estimate of the first non-trivial eigenvalue λ_1 of Laplacian. On this topic, there is a great deal of publications (see for instance [5], Schoen and Yau [15], Wang [16], and references therein). Throughout this paper, we use the quantity

$$\alpha = \alpha(K, d, D) = \frac{D}{2} \sqrt{\frac{-K}{d-1}},$$

which involves all of the three geometric quantities: d , D and K . Clearly, $|\alpha| = \alpha(|K|, d, D)$. By the Myers theorem, we have $|\alpha| \leq \pi/2$ whenever $K > 0$ for a complete Riemannian manifold. Therefore, the quantity α is geometrically meaningful. We adopt the convention: $\alpha = 0$ if $d = 1$. With the

Received January 13, 2011; accepted November 2, 2011.

2000 *Mathematics Subject Classifications.* 58C40, 35P15.

Key words and phrases. First non-trivial eigenvalue, sharp estimate, Riemannian manifold.

Research supported in part by the National Natural Science Foundation of China (Nos. 10721091, 11131003), by the “985” project from the Ministry of Education in China.

necessary notation in mind, the main result Theorem 1.4 and its illustrating figures 7–9 should be readable now. One may have a look before going further.

We are going to recall the sharp lower estimates, only the related part of them for saving space. First, for nonnegative curvature, the following sharp lower bounds are perhaps well known.

- Lichnerowicz (1958):

$$\frac{d}{d-1} K = \frac{4d}{D^2} |\alpha|^2, \quad d > 1, K \geq 0. \quad (1)$$

- Bérard, Besson and Gallot (1985):

$$d \left\{ \frac{\int_0^{\pi/2} \cos^{d-1} t dt}{\int_0^{D/2} \cos^{d-1} t dt} \right\}^{2/d} = d \left\{ \frac{\int_0^{\pi/2} \cos^{d-1} t dt}{\int_0^{|\alpha|} \cos^{d-1} t dt} \right\}^{2/d}, \quad K = d - 1 > 0. \quad (2)$$

Here and in what follows, $\cos^k t = (\cos t)^k$.

- Chen and Wang (1997):

$$\frac{dK}{(d-1)(1 - \cos^d |\alpha|)} = \frac{4d|\alpha|^2}{D^2(1 - \cos^d |\alpha|)}, \quad d > 1, K \geq 0. \quad (3)$$

- Zhong and Yang (1984):

$$\pi^2/D^2, \quad K \geq 0. \quad (4)$$

It is clear that (2) improves (1) by the Myers theorem. Even though it is not so obvious but it is true that (3) improves (2). The first three results indicate a long period for the improvements of (1) step by step. All of them are sharp for the unit sphere in two or higher dimensions but fail for the unit circle ($K = 0$). To which, the sharp estimate is given by the last result (4). Note that when $|\alpha| \downarrow 0$, the limit of (3) equals $8D^{-2} \in (0, \pi^2 D^{-2})$.

Secondly, consider the non-positive curvature in which case, the problem becomes harder. Here are the main known sharp lower bounds.

- Yang (1990), Jia (1991), Chen and Wang (1994):

$$\frac{\pi^2}{D^2} e^{-(d-1)\alpha}, \quad K \leq 0. \quad (5)$$

- Chen and Wang (1997):

$$\frac{1}{D^2} \sqrt{\pi^4 + 8(d-1)\alpha^2} \cosh^{1-d} \alpha, \quad d > 1, K \leq 0. \quad (6)$$

- [2; 1994, Theorem 6.6](corrected version):

$$\frac{1}{D^2}((d-1)\alpha \tanh \alpha \operatorname{sech} \theta)^2, \quad d > 1, K \leq 0, \quad (7)$$

where θ is obtained in the following way. Let

$$\theta_1 = 2^{-1}(d-1)\alpha \tanh \alpha, \quad \theta_n = \theta_1 \tanh \theta_{n-1}, \quad n \geq 2,$$

then $\theta_n \downarrow \theta$.

The first two results have the same decay rate but (6) is better than (5) in general. They are sharp for the unit circle but not the last result which is designed for large α . The last two estimates are not comparable. For fixed d , (6) is better than (7) for smaller α but the inverse assertion happens for large α (cf. Fig. 8 below).

Thirdly, consider the optimal linear approximation of the lower estimates with respect to the curvature K . In other words, one looks for a good combination of the optimal estimates (1) and (4). Many authors have contributed to the result (8) below. It was proved by Zhao (1999) under the restriction $-5\pi^2/(3D^2) \leq K \leq 0$ and the remainder gap was covered by Xu and Pang (2001) in the case of $K \leq 0$. The case of $K > 0$ was proved by Xu, Yang and Xu (2002). Independently, the assertion was proved, with computer assisted, for all real K in Chen, Scacciatelli and Yao (2001) (where some refined estimates are included). A more direct analytic proof with some improvement was given by Shi and Zhang (2007). Recently, the result (8) below has been reproved by Ling (2006, 2007) using a different approach.

- The following lower bound is studied/proved in the papers just mentioned:

$$\pi^2/D^2 + K/2, \quad K \in \mathbb{R}. \quad (8)$$

- More precisely, the lower bound given by Shi and Zhang (2007) is as follows:

$$\begin{aligned} & \sup_{s \in (0,1)} s \left[4(1-s)\pi^2/D^2 + K \right] \\ &= \begin{cases} \left(\frac{\pi}{D} + \frac{KD}{4\pi} \right)^2, & -4\pi^2 \leq KD^2 \leq 4\pi^2 \\ K, & KD^2 \in (4\pi^2, (d-1)\pi^2] \\ 0, & KD^2 < -4\pi^2 \end{cases} \end{aligned} \quad (9)$$

- A refined lower bound given by Chen, Scacciatelli and Yao (2001) is the following:

$$\frac{\pi^2}{D^2} + \frac{K}{2} + (10 - \pi^2) \frac{K^2 D^2}{16}, \quad |K|D^2 \leq 4. \quad (10)$$

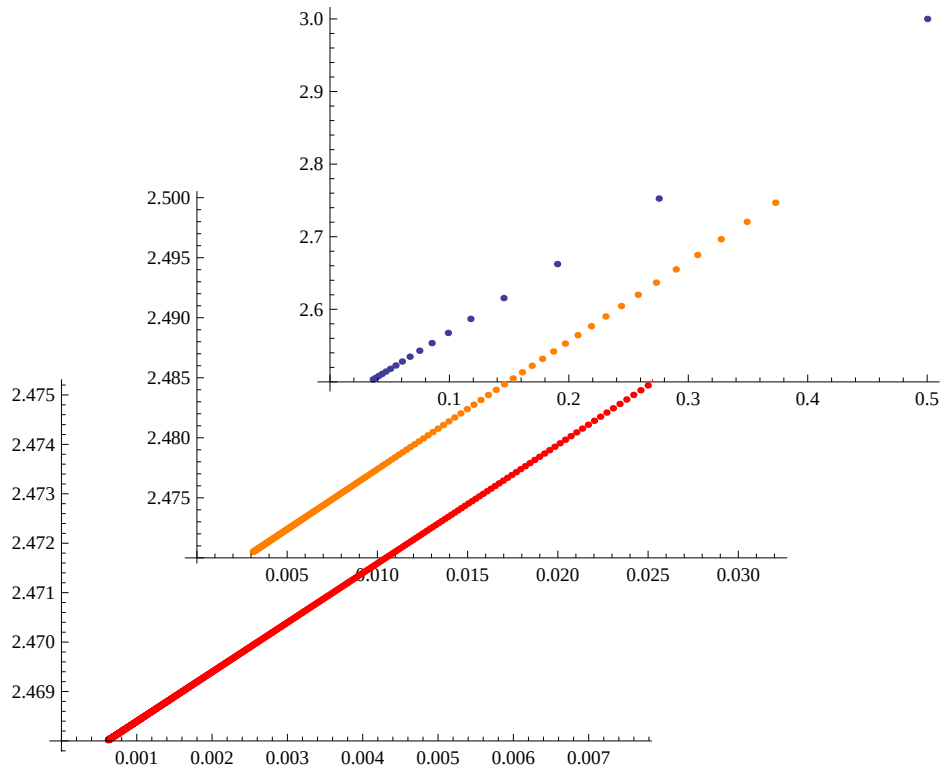
To see that (9) improves (8), simply set $s = 1/2$. On the region

$$\{(K, D) : |K|D^2 \leq 4\},$$

it is obvious that (10) is better than (9). Unlike the results (1)–(7), the bounds given in (8)–(10) are independent of the dimension d .

Before moving further, let us make some remarks on (10). Since one is seeking for the dimension-free estimate and

$$(d-1)\alpha \tanh(\alpha r) \uparrow -KD^2r/4 =: -2\beta r \quad \text{as } d \uparrow \infty$$



Figures 1–3 The first eigenvalue λ_0 corresponding to $\{\beta_n\}_{n=2}^{18}$ (the top curve), $\{\beta_n\}_{n=2}^{200}$ (the middle curve), and $\{\beta_n\}_{n=2}^{1000}$ (the bottom curve), respectively.

(cf. proof of Theorem 1.2 below), the study on λ_1 can be reduced to study the first eigenvalue (say λ_0 , for a moment) of the operator $d^2/dr^2 - 2\beta r d/dr$ on $(0, 1)$ (cf. [7; Lemma 2.3] in which the constant α is replaced by β here). This now becomes a one-parameter problem and the required estimate (10) becomes a product of $4D^{-2}$ and

$$\pi^2/4 + \beta + (10 - \pi^2)\beta^2, \quad |\beta| \leq 1/2 \quad (11)$$

which estimates λ_0 from below. For a sequence of β , say $\{\beta_n\}_{n=2}^\infty$ with $\beta_2 = 1/2$ and $\beta_n \downarrow 0$ as $n \rightarrow \infty$, the problem is numerically solvable and moreover, we

have analytic solutions for the first four of $\{\beta_n\}$. Actually, the corresponding eigenfunctions are all polynomials (cf. [7; Lemma 2.3]. Now, (10) is designed to be exact at $\beta = 0$ and $\beta = 1/2$ with the coefficient of the term K to be $1/2$. The first eigenvalue λ_0 just mentioned corresponding to $\{\beta_n\}$ is shown by three pictures for different region of n (Figures 1–3), and then the difference between the eigenvalue λ_0 and its lower estimate (11) is shown by two pictures (Figures 4 and 5). In Fig.3, even though the whole sequence $\{\beta_n\}_{n=2}^{1000}$ is computed, but the output is restricted to a smaller interval near zero. To see the other part of the interval, one needs two more figures (1 and 2).

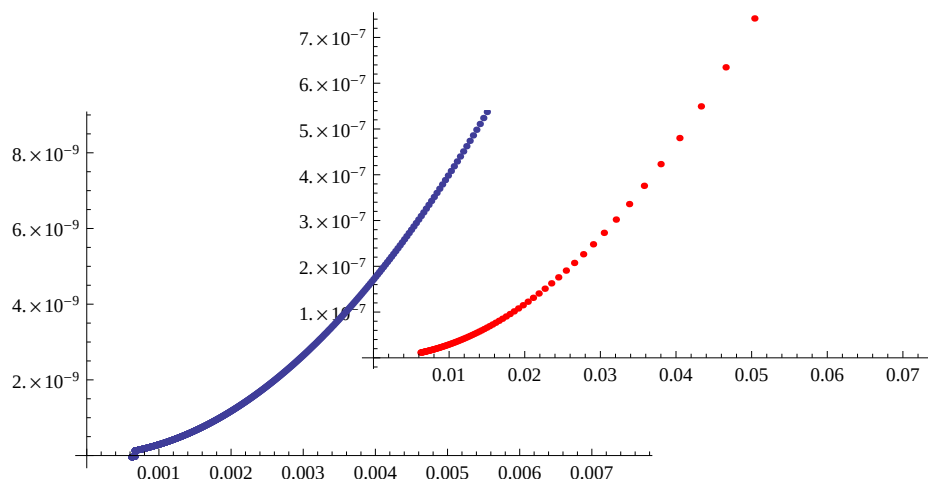


Figure 4–5 The difference of the first eigenvalue λ_0 and its lower estimate (11) corresponding to $\{\beta_n\}_{n=2}^{30}$ (the curve on right) and $\{\beta_n\}_{n=2}^{1000}$ (the curve on left), respectively.

The difficulty is that for each $n \geq 2$, one has to find a root of a polynomial having order $n - 1$. Here, instead of an analytic proof, we have used five figures to show carefully that the lower bound (10) is rather sharp and the coefficient $1/2$ for the linear approximation should be exact. Besides, these figures also indicate that the analytic proofs would be too heavy for the paper and may not be essential, this is the reason why we often use figures, here and in what follows. We mention that by a simple transform, the same conclusions hold for the sequence $\{-\beta_n\}_{n=1}^{\infty}$ (cf. [7; Lemma 2.4]). Therefore, (10) holds for all real K with $|K|D^2 \leq 4$.

Except the results (1)–(3) and (7), all of the above results are sharp only at one point ($K = 0$), some of them can be very poor in some region. For instance, (1)–(3) are meaningless when $K = 0$ and (8)–(10) can be negative or zero for sufficiently large $-K$. The question now is the existence of a universal estimate. Fortunately, the answer is affirmative as shown by Proposition 1.1 below. Its first assertion implies the lower bounds (1)–(10), in terms of $\bar{\lambda}$.

Recall that even though α is an imaginary number when $K > 0$, the quantities $\tanh(\alpha r)$ and $\cosh(\alpha r)$ used below are still meaningful: $\cosh(i\theta) = \cos \theta$, $\tanh(i\theta) = i \tan \theta$ for real θ .

Proposition 1.1 We have $\lambda_1 \geq 4\bar{\lambda}/D^2$ with the following estimates:

$$\frac{1}{4\delta} \leq \frac{1}{\delta_1 \wedge \delta_1^*} \leq \bar{\lambda} \leq \frac{1}{\delta_1' \vee \delta_1^{*'}} \leq \frac{1}{\delta},$$

where

$$\begin{aligned} \delta &= \sup_{r \in (0,1)} [\varphi\psi](r), \\ \delta_1 &= \sup_{r \in (0,1)} \left\{ \frac{1}{\sqrt{\varphi(r)}} \int_0^r C \varphi^{3/2} + \sqrt{\varphi(r)} \int_r^1 C \varphi^{1/2} \right\}, \\ \delta_1' &= \sup_{r \in (0,1)} \left\{ \frac{1}{\varphi(r)} \int_0^r C \varphi^2 + [\varphi\psi](r) \right\}, \\ \delta_1^* &= \sup_{r \in (0,1)} \left\{ \frac{1}{\sqrt{\psi(r)}} \int_r^1 C^{-1} \psi^{3/2} + \sqrt{\psi(r)} \int_0^r C^{-1} \psi^{1/2} \right\}, \\ \delta_1^{*'} &= \sup_{r \in (0,1)} \left\{ \frac{1}{\psi(r)} \int_r^1 C^{-1} \psi^2 + [\varphi\psi](r) \right\} \end{aligned}$$

(here the Lebesgue measure “ du ” is omitted) with

$$C(s) = \cosh^{d-1}(\alpha s), \quad \varphi(r) = \int_0^r C(u)^{-1} du, \quad \psi(r) = \int_r^1 C(u) du.$$

All of the quantities used here depend on d and α .

We remark that originally, the variable r varies on the interval $[0, D]$, the function we are interested is $C(r) = \cosh^{d-1}(\alpha r/D)$, $r \in [0, D]$ (refer to page 2 with $K(V) = 0$ of article [9] in Volume I). Here in Proposition 1.1 and what follows, we adopt a change of variable $r/D \rightarrow r$, then the function $C(r)$ becomes $C(s) = \cosh^{d-1}(\alpha s)$, $s \in [0, 1]$, and then we deduce the main conclusion of Proposition 1.1.

To illustrate the power of Proposition 1.1, consider the simplest case that $K = 0$. Then $\delta = 1/4$, $\delta_1 = \delta_1^* = 5^{1/3}/4 \approx 0.427$, $\delta_1' = \delta_1^{*'} = 3/8$, and so

$$\frac{\delta_1}{\delta_1'} = \frac{\delta_1^*}{\delta_1^{*'}} = \frac{5^{1/3}}{4} \bigg/ \frac{3}{8} \approx 1.14.$$

The sharp estimate for λ_1 is π^2/D^2 and then $\bar{\lambda}^{-1} = 4/\pi^2 \approx 0.405$. Thus, our estimates read as follows.

$$\delta = 0.25 < \delta_1' = \delta_1^{*'} = 0.375 < \bar{\lambda}^{-1} \approx 0.405 < \delta_1 = \delta_1^* \approx 0.427 < 4\delta = 1.$$

For this example, the results that $\delta_1 = \delta_1^*$ and $\delta_1' = \delta_1^{*'}$ are quite natural by symmetry. The not so obvious fact is $\delta_1 \geq \delta_1^*$ in the most cases and the inequality can happen, as shown by numerical computations.

Actually, Proposition 1.1 is deduced from the next result which is an improvement of the main variational formula [9; Theorem 1].

Theorem 1.2 We have $\lambda_1 \geq 4\bar{\lambda}/D^2$ and two variational formulas:

$$\begin{aligned}\bar{\lambda} &= \sup_{f \in \mathcal{F}} \inf_{r \in (0,1)} \frac{f(r)}{\int_0^r C(s)^{-1} ds \int_s^1 C(u) f(u) du} \\ &= \sup_{f \in \mathcal{F}} \inf_{r \in (0,1)} \frac{f(r)}{\int_r^1 C(s) ds \int_0^s C(u)^{-1} f(u) du},\end{aligned}$$

where $\mathcal{F} = \{f \in \mathcal{C}[0,1] : f|_{(0,1)} > 0\}$.

Recall that the estimates given in (8)–(10) are all dimension-free. The next result is an improvement of (9) which depends on dimensions.

Proposition 1.3 For λ_1 , we have lower bound (9) replacing K by KM_α , where

$$M_\alpha = \frac{\pi^2}{4} \int_0^1 (1-y) \cos \frac{\pi y}{2} \operatorname{sech}^2(\alpha y) dy$$

regarding α as the constant α_0 if $K > 0$ and $|\alpha| \in (|\alpha_0|, \pi/2]$, where $|\alpha_0|$ (depending on d) is the first positive root of

$$\left(\frac{\pi}{2\sqrt{d-1}|\alpha|} + \frac{\sqrt{d-1}|\alpha|}{2\pi} \right) \cos |\alpha| = 1. \quad (12)$$

In Proposition 1.3, the improvement of (9) is due to the fact that $M_\alpha < 1$ if $K < 0$ and $M_\alpha > 1$ if $K > 0$. The proposition reduces to (9) once $d \rightarrow \infty$ since then $\alpha \rightarrow 0$, and furthermore, $M_0 = 1$. Here is the picture of M_α (Figure 6).

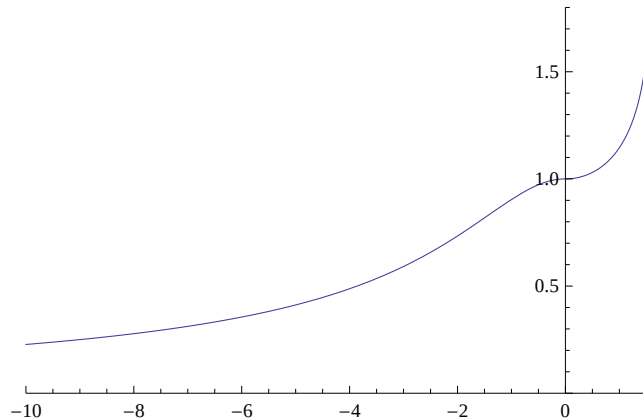


Figure 6 The curve of M_α with $\alpha = \sqrt{-\operatorname{sgn}(x)}|x|$, $x \in (-10, \pi/2)$. Obviously, $K < 0$ iff $x < 0$.

Note that the region on which Proposition 1.3 being available is smaller than that of (9). In view of (10), this is reasonable since a larger lower bound can be held only in a smaller region.

As mentioned in the earlier publication (cf. [5; Chapter 3] for instance) or in Theorem 1.2, our study on λ_1 consists of two steps. The first one is reducing the higher dimensions to dimension one as shown by the first assertion of Proposition 1.1 or Theorem 1.2. The second step is estimating $\bar{\lambda}$. This was started by [9], continued by several papers mentioned before (8), and is also the main aim of the present paper to justify the power of our one-dimensional results [3], [4] and [6]. The lower bound $\delta_1^{-1} \vee \delta_1^{*-1}$ of $\bar{\lambda}$ provided by Proposition 1.1 is universal in the sense that the upper and lower bounds of $\bar{\lambda}$ are the same up to a factor 4. Actually, all of a large number of examples we have ever computed, as well as Figures 7–9 below, show that the ratio δ_1/δ_1' (and $\delta_1^*/\delta_1^{*'}\prime$) is no more than 2. It is somehow unexpected that the lower bound $4D^{-2}\delta_1^{*-1}$ of λ_1 is better than the others except in two cases (cf. Figures 7–9 below). In the case that $K = 0$, we have shown after Proposition 1.1 by an example that $4D^{-2}\delta_1^{*-1}$ is not sharp. When α closes to $\pi/2$, we are near the unit sphere and so $4D^{-2}\delta_1^{*-1}$ can not be better than (3). Thus, we may regard $4D^{-2}\delta_1^{*-1}$ (be careful to distinguish δ_1^* and δ_1) as our general common lower bound, and regard (3), (10) and Proposition 1.3 as an addition. We can now summarize the main result of the paper as follows.

Theorem 1.4 In general, we have the following lower estimate:

$$\lambda_1 \geq \frac{4}{D^2} \left\{ \frac{1}{\delta_1^*} \bigvee_{s \in (0,1)} s \left[(1-s)\pi^2 - (d-1)\alpha^2 M_\alpha \right] \bigvee \left[\mathbb{1}_{\{K>0\}} \frac{d|\alpha|^2}{1 - \cos^d |\alpha|} \right] \right\},$$

where $x \vee y = \max\{x, y\}$, δ_1^* and M_α are given in Propositions 1.1 and 1.3, respectively, with a restriction on $|\alpha|$ for the middle term in the case of $K > 0$: regarding α as the constant α_0 if $|\alpha| \in (|\alpha_0|, \pi/2]$, where $|\alpha_0|$ is the first positive root of (12). Besides, we also have the dimension-free lower bound (10). The middle estimate is better than (10) if $2 \leq d \leq 7$, and conversely if $d \geq 10$.

For the convenience of computation, according to (9), we express the middle term in Theorem 1.4 as follows.

$$\begin{aligned} & \sup_{s \in (0,1)} s \left[(1-s)\pi^2 - (d-1)\alpha^2 M_\alpha \right] \\ &= \begin{cases} \left(\pi/2 - (d-1)\alpha^2 M_\alpha / (2\pi) \right)^2, & -\pi^2 \leq -(d-1)\alpha^2 M_\alpha \leq \pi^2 \\ KM_\alpha, & -(d-1)\alpha^2 M_\alpha \in (\pi^2, (d-1)\pi^2/4] \\ 0, & -(d-1)\alpha^2 M_\alpha < -\pi^2. \end{cases} \end{aligned}$$

However, when $K > 0$, $|\alpha|$ is essentially restricted to the subinterval $(0, |\alpha_0|)$, where $|\alpha_0|$ is the smallest positive root of equation (12).

Fig. 7 illustrates the meaning of Theorem 1.4 ignoring the common factor $4D^{-2}$. Here, we consider only $K \geq 0$ and $d = 2$. Clearly, $\bar{\lambda}$ is located between

the two dotted curves and the ratio of the upper and lower bounds ($\delta_1'^{-1}$ and δ_1^{*-1}) of $\bar{\lambda}$ is obviously less than 2. Note that here we use $*$ twice. The curve, say Curve 1, corresponding to the last term in Theorem 1.4 is sharp at $\pi/2$ but is at the lowest level at origin; and the partially dashed curve, say Curve 2, corresponding to the middle term in Theorem 1.4 is sharp at 0. Note that Curve 2 is located above Curve 1 and is in particular higher than Curve 1 when $|\alpha|$ is close to $\pi/2$, which is impossible since Curve 1 is sharp at $\pi/2$ as we have just mentioned. Hence, a restriction on $|\alpha|$ for Curve 2 is necessary and what we adopted is $|\alpha| \leq 0.97$ (that is ignoring the dashed part of Curve 2). Here for Curve 2, we omit the constant line starting from the endpoint of the non-dashed part of the curve to $\pi/2$ (cf. Fig. 9 below). It is interesting that Curves 1 and 2 together control the most part of the interval $[0, \pi/2]$. Usually, we have $\delta_1 \geq \delta_1^*$. The exceptional is that here $\delta_1/\delta_1^* \in (0.99993, 1)$ when $|\alpha| < 0.87$. Since this part is covered by Curve 2 and so δ_1 is ignored in Theorem 1.4.

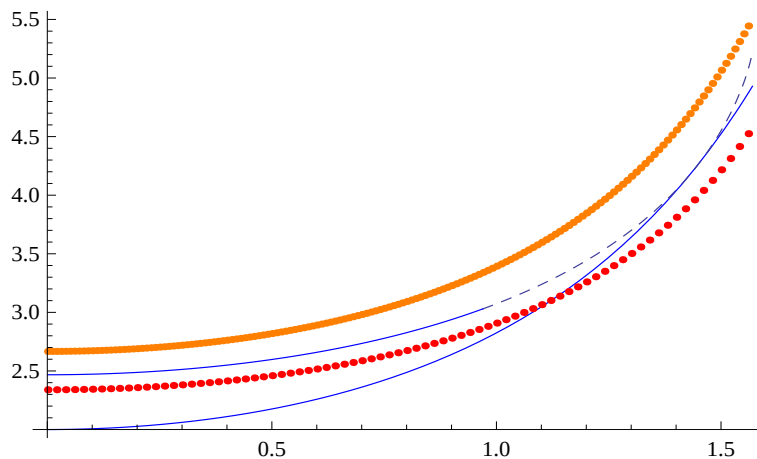


Figure 7 The estimates of $\bar{\lambda}$ in the case of $K \geq 0$ and $d = 2$, $|\alpha| \leq \pi/2$.

Fig. 8 represents the case of $K \leq 0$ and $d = 2$. Again, the upper and lower bounds ($\delta_1'^{-1}$ (but not δ_1^{*-1}) and δ_1^{*-1}) of $\bar{\lambda}$ are given by the two of top dotted curves. The solid curve is determined by the middle term of Theorem 1.4. The figure shows that the bounds $\delta_1'^{-1}$ and δ_1^{*-1} are rather good, even coincide each other for larger α . In a small interval around 0, they are less sharper than the solid curve. The dashed curve corresponds to (6) and dotted part of the triangle corresponds to (7). They are not comparable, and are less powerful than at least one of the others and so are disappeared in Theorem 1.4. For $d > 2$, the picture is similar but each curve decays fast (cf. Fig. 9).

Theorem 1.4 is stated unified in $K \in \mathbb{R}$. Fig. 9 is the result for both negative and positive K , where $\alpha = \sqrt{-\text{sgn}(x)}|x|$ with x varies from -2.5 to $\pi/2$ and $d = 5$. Note that the axes in these figures have different scales. The meaning of each curve should be clear. The top dotted curve is $\delta_1'^{-1}$ for $x < 0$ and is δ_1^{*-1} for $x > 0$. The part of the curves near $\pi/2$ is ignored, otherwise

the left part of the curves would be very mixed. However, the shape of the missed part is very much imaginable, up to 14.75 high.

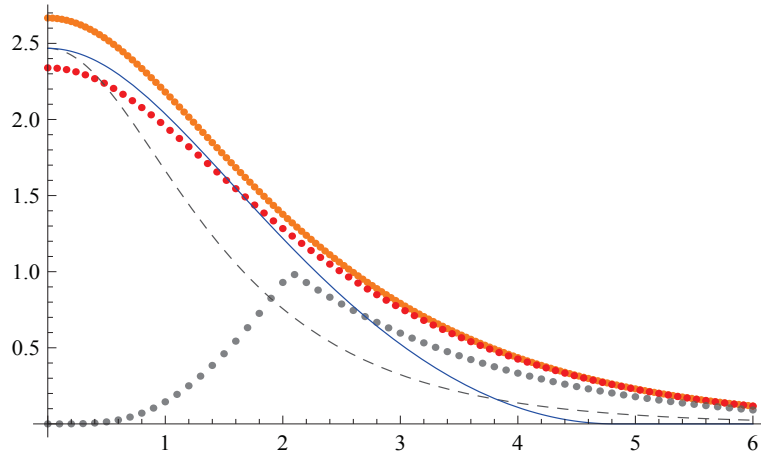


Figure 8 The estimates of $\bar{\lambda}$ in the case of $K \leq 0$ and $d = 2$, $0 \leq \alpha \leq 6$.

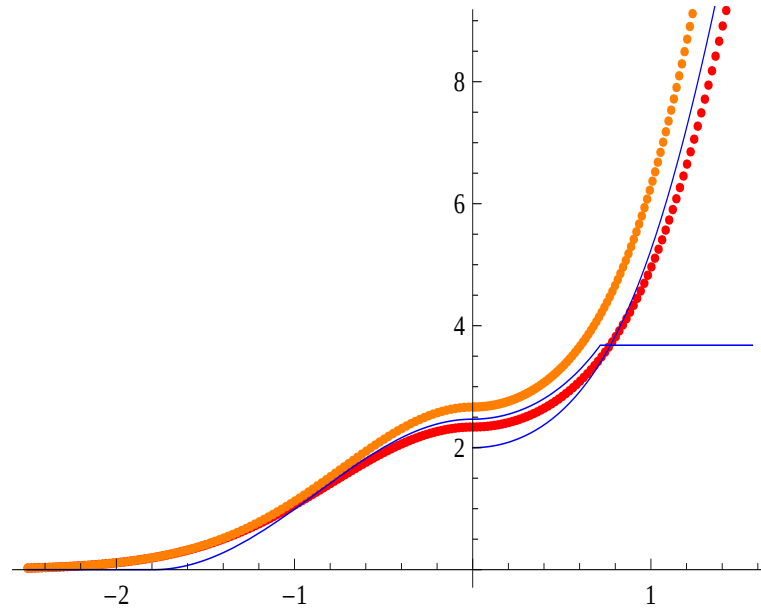


Figure 9 The estimates of $\bar{\lambda}$ when $d = 5$ with $\alpha = \sqrt{-\operatorname{sgn}(x)} |x|$, $x \in (-2.5, \pi/2)$.

Remark 1.5 Figures 7–9 show that the ratio $\delta_1^*/\delta_1^{*'} \in [1, 4]$ (as well as $\delta_1^{*'} - \delta_1^{*-1}$) is controlled by its value at the endpoint $|\alpha| = \pi/2$. The ratio increases quickly from 1.2 to 1.27 and then slowly to 1.334 (< 2) when d varies from 2 to 5 and then to 63.

Remark 1.6 Consider the convex mean $\eta_z = \gamma_z \delta_1^{*-1} + (1 - \gamma_z) \delta_1^{*-1}$ with

$$\gamma_0 = \frac{5^{2/3} - 5 \cdot 16^{-1} \pi^2}{5^{2/3} - 10 \cdot 3^{-1}} \approx 0.39 \quad \text{or} \quad \gamma_{\pi/2} = \frac{4^{-1} d \pi^2 - \delta_1^{*-1}}{\delta_1^{*-1} - \delta_1^{*-1}} \Big|_{|\alpha|=\pi/2}.$$

The latter one depends on d but the former one does not. Here η_0 is designed to be sharp ($= \pi^2/4$) at $\alpha = 0$ and so is $\eta_{\pi/2}$ ($= d \pi^2/4$) at $|\alpha| = \pi/2$ ($K > 0$). In particular, $\gamma_{\pi/2} \approx 0.35$ if $d = 63$. Numerical computations (illustrating figures are given in the author's homepage) exhibit the following unexpected nice conclusion:

$$\bar{\lambda} - 0.056 \leq \eta_{\pi/2} \leq \bar{\lambda} \leq \eta_0 \leq \bar{\lambda} + 1.85, \quad 2 \leq d \leq 63, \quad \forall \alpha.$$

Thus, one may regard η_0 and $\eta_{\pi/2}$ (for each fixed d) as upper and lower bounds of $\bar{\lambda}$, respectively, but they are almost the same since $\bar{\lambda} \approx 155$ when $d = 63$. This illustrates the power of Proposition 1.1, and is independent of (1)–(10).

2 Proofs

Proof of Theorem 1.2. (a) Consider first the case that $\bar{\lambda}$ is defined by its first equality given in the theorem. The first assertion is a comparison of λ_1 with the principal eigenvalue λ_0 of the operator

$$\bar{L} = \frac{d^2}{dr^2} + (d-1)\alpha \tanh(\alpha r) \frac{d}{dr}$$

on $(0, 1)$ with Dirichlet boundary at 0 and Neumann boundary at 1. This was done in [8], as explicitly pointed out by [9; Remark d) after Theorem 1.1]. Actually, the result came out by using the coupling method to a computation of some distance which is certainly valued in the half-line. This reduces the higher dimensions to dimension one. Then it was proved in [9; Theorem 1.1] that $\lambda_0 \geq \bar{\lambda}$. The equality here holds because of [4; Theorem 1.1] noting that one allows the test function f to be positive at origin not necessarily zero (cf. [4; (1.3)]).

(b) Next, let λ_0^* be the principal eigenvalue of the dual operator

$$\bar{L}^* = \frac{d^2}{dr^2} - (d-1)\alpha \tanh(\alpha r) \frac{d}{dr}$$

with Neumann boundary at 0 and Dirichlet boundary at 1. Here and in what follows, the notation “*” is used for dual quantity. Then, we have not only $\lambda_0 = \lambda_0^*$ but also the second equality for $\bar{\lambda}$ given in the theorem, as an analog of [6; §5 and Theorem 2.4 (3)]. $\square$

Proof of Proposition 1.1. The first assertion comes from Theorem 1.2. The result is very helpful since the parameter D is separated out, and then the three parameters d , D , and K are now reduced to two: d and α .

The “ δ part” of the assertion was presented in [5; Corollary 1.4], comes originally from [3; Theorem 2.2] with a change of the variable: $r \rightarrow r/D$ reducing the interval $(0, D)$ to $(0, 1)$. Actually, the original result is more general, including a vector field. We mention that Proposition 1.1 is meaningful in such a general situation.

In view of the author’s knowledge, the “ δ_1 and δ'_1 parts” have not yet published in the geometric context. However, it is indeed the first step of a general approximating procedure for the first eigenvalue, given by [4; Theorem 2.2]. Here we state the procedure in the present context. Define $f_1 = \sqrt{\varphi}$,

$$f_{n+1}(r) = \int_0^r C(s)^{-1} ds \int_s^1 C(u) f_n(u) du, \quad n \geq 1$$

and $\delta_n = \sup_{r \in (0,1)} f_{n+1}(r)/f_n(r)$. Then

$$\bar{\lambda} \geq \dots \geq \delta_n^{-1} \geq \delta_{n-1}^{-1} \geq \dots \geq \delta_1^{-1} \geq (4\delta)^{-1}.$$

Next, fix $r \in (0, 1)$ and define $f_1^{(r)} = \varphi(\cdot \wedge r)$,

$$f_{n+1}^{(r)} = \int_0^{\cdot \wedge r} C(s)^{-1} ds \int_s^1 C(u) f_n^{(r)}(u) du, \quad n \geq 1$$

and $\delta'_n = \sup_{r \in (0,1)} \inf_{s \in (0,1)} f_{n+1}^{(r)}(s)/f_n^{(r)}(s)$. Then

$$\bar{\lambda} \leq \dots \leq \delta_n'^{-1} \leq \delta_{n-1}'^{-1} \leq \dots \leq \delta_1'^{-1} \leq \delta^{-1}.$$

Besides, we also have $\bar{\lambda} \leq \bar{\delta}_n^{-1}$ for all n , where

$$\bar{\delta}_n = \sup_{r \in (0,1)} \int_0^1 f_n^{(r)}(s)^2 C(s) ds \Big/ \int_0^1 (f_n^{(r)}(s))'^2 C(s) ds, \quad n \geq 1.$$

Moreover, $\bar{\delta}_1 = \delta'_1$. All of these approximating procedure comes from some variational formulas [4; Theorem 2.1]. In particular, the sequence $\{\delta_n\}_{n \geq 1}$ comes from the first variational formula given in Theorem 1.2.

The “ δ_1^* and $\delta_1^{*'}$ parts” are parallel to the “ δ_1 and δ'_1 parts” based on the dual formula just mentioned, as an analog of [6; Theorem 3.2 and Corollary 3.3]. The details are omitted here. We mention that the duality is studied more carefully in the author’s forthcoming paper entitled “Basic estimates of stability rate for one-dimensional diffusions”.

Clearly, the estimates given in Proposition 1.1 can be still improved by using the above approximating procedure. $\square$

To prove Proposition 1.3, we need some preparations. The following nice result it due to [14; proof of Lemma 2.2], it is the key leading to (9). As usual, we use $\mathcal{C}^m$ to denote the set of functions having continuous m -th derivative.

Lemma 2.1 Let λ and f satisfy

$$f'' + Ff' = -\lambda f \quad \text{on } [0, \ell], \quad \ell < \infty$$

with boundary conditions $f(0) = 0$ and $f'(\ell) = 0$, where $F \in \mathcal{C}[0, \ell] \cap \mathcal{C}^1(0, \ell)$ having $F(0) = 0$. Then for each $s \in (0, 1)$, the function $g := (f')^{2^{-1}(1-s)^{-1}}$ satisfy

$$4s(1-s) \int_0^\ell g'^2 dr = \int_0^\ell (\lambda + sF')g^2 dr.$$

Moreover, g satisfies the mixed boundary condition: $g'(0) = 0$ and $g(\ell) = 0$.

The proof of Lemma 2.1 goes as follows. Making derivatives of the original equation, we get

$$f''' + Ff'' = -(\lambda + F')f'.$$

Regarding f' as a new function h , one sees that this corresponds to a second order Schrödinger operator (since the appearance of F'). This step is known as an application of the coupling technique (cf. [6; (10.4), (10.7)] and references therein). The next step we have used is to adopt the dual operator $d^2/dx^2 - Fd/dx$ (cf. [6; (10.6)]) which is isospectral to the Schrödinger one. Refer to [6; proof of Theorem 10.2] for more details. However, here the idea is different. Set $t = (1-s)^{-1}$. Multiply the last equation by $(f')^{t-1}$ and then integrate each term in the last equation over $(0, \ell)$. Based on the fact that $F(0) = 0$ and then $f''(0) = 0$, some careful computations lead to the required conclusion.

Having Lemma 2.1 at hand, the assertion (9) is immediate. Simply replace $\lambda + sF'$ by the constant $\lambda + s \sup_{r \in (0, \ell)} F'(r)$ (which then does not depend on the dimension d explicitly) and use the known exact inequality (cf. Proof of Proposition 1.2 below)

$$\int_0^\ell g^2 dr \leq \left(\frac{2\ell}{\pi}\right)^2 \int_0^\ell g'^2 dr.$$

It is at this step, the original proof is improved. This leads to a use of the next result.

Lemma 2.2 Let $a > 0$ on $(0, 1)$ and $a \in \mathcal{C}^2(0, 1)$. Next, let $\hat{\lambda}$ be the principal eigenvalue of the operator $\hat{L} = a(x)^{-1}d^2/dx^2$ on $(0, 1)$ with Neumann boundary at 0 and Dirichlet boundary at 1. Then $\hat{\lambda}$ obeys the following estimates

$$\inf_{x \in (0, 1)} h(x)^{-1} \leq \hat{\lambda} \leq \sup_{x \in (0, 1)} h(x)^{-1},$$

where

$$h(x) = \left(\frac{2}{\pi}\right)^2 \left\{ a(x) - \sec \frac{\pi x}{2} \left[(1-x)a'(0) + (1-x) \int_0^x a''(y) \cos \frac{\pi y}{2} dy \right. \right. \\ \left. \left. + \int_x^1 [-2a'(y) + (1-y)a''(y)] \cos \frac{\pi y}{2} dy \right] \right\}.$$

Proof. We are now in the case which is the continuous analog of [6; §2]. Noticing that when $a(x)$ is a constant, the assertion of the lemma is exact. In this case, the eigenfunction is cosine and so our approximation should start at the function cosine, because we are looking for such estimates which are sharp when $a(x)$ is a constant. Define $f_1(x) = \cos \frac{\pi x}{2}$ and

$$f_n(x) = \int_x^1 dy \int_0^y a(u) f_{n-1}(u) du, \quad x \in (0, 1), \quad n \geq 1.$$

Here, the proof of the lemma is mainly for simplifying f_2/f_1 and so is rather elementary. Exchanging the integrals, we get

$$\begin{aligned} f_2(x) &= \int_x^1 dy \int_0^y a(u) \cos \frac{\pi u}{2} du \\ &= \int_0^1 (1-x \vee u) a(u) \cos \frac{\pi u}{2} du \\ &= (1-x) \int_0^x a(y) \cos \frac{\pi y}{2} dy + \int_x^1 (1-y) a(y) \cos \frac{\pi y}{2} dy. \end{aligned}$$

As an application of the integration by parts formula, we have

$$\begin{aligned} f_2(x) &= \frac{2}{\pi} \left[(1-x) \int_0^x a(y) d \sin \frac{\pi y}{2} + \int_x^1 (1-y) a(y) d \sin \frac{\pi y}{2} \right] \\ &= -\frac{2}{\pi} \left[(1-x) \int_0^x a'(y) \sin \frac{\pi y}{2} dy + \int_x^1 [-a(y) + (1-y)a'(y)] \sin \frac{\pi y}{2} dy \right]. \end{aligned}$$

Using the integration by parts formula again, we get

$$\begin{aligned} f_2(x) &= \left(\frac{2}{\pi} \right)^2 \left\{ a(x) \cos \frac{\pi x}{2} - (1-x)a'(0) - (1-x) \int_0^x a''(y) \cos \frac{\pi y}{2} dy \right. \\ &\quad \left. - \int_x^1 [-2a'(y) + (1-y)a''(y)] \cos \frac{\pi y}{2} dy \right\}. \end{aligned}$$

Unlike the original one, no double integral appears in the last formula. Obviously, $f_2(x)/f_1(x) = h(x)$. The required assertion now follows by [4; Theorem 1.1]. $\square$

Proof of Proposition 1.3. Applying Lemma 2.1 to $\bar{L}$ and $\bar{\lambda}$ defined in the proof of Theorem 1.2 with $F(r) = (d-1)\alpha \tanh(\alpha r)$, we obtain

$$\begin{aligned} 4s(1-s) \int_0^1 g'^2 dr &= \int_0^1 [\bar{\lambda} + s(d-1)\alpha^2 \operatorname{sech}^2(\alpha r)] g^2 dr \\ &= \bar{\lambda} \int_0^1 g^2 dr + s(d-1)\alpha^2 \int_0^1 \operatorname{sech}^2(\alpha r) g^2 dr. \end{aligned} \quad (13)$$

(a) Let $K \leq 0$ and set $a(r) = \operatorname{sech}^2(\alpha r)$. Then $a > 0$ on $[0, 1]$ and

$$\begin{aligned} a'(r) &= -2\alpha \operatorname{sech}^2(\alpha r) \tanh(\alpha r), \\ a''(r) &= -2\alpha^2 \operatorname{sech}^4(\alpha r) [2 - \cosh(2\alpha r)]. \end{aligned}$$

Applying Lemma 2.2 to this $a(r)$, we obtain, replacing h by h_α , that

$$\begin{aligned} \frac{\pi^2}{4} h_\alpha(x) &= \operatorname{sech}^2(\alpha x) + 2\alpha \sec \frac{\pi x}{2} \left[\alpha(1-x) \int_0^x q_\alpha(y) \cos \frac{\pi y}{2} dy \right. \\ &\quad \left. + \int_x^1 [-2p_\alpha(y) + \alpha(1-y)q_\alpha(y)] \cos \frac{\pi y}{2} dy \right], \end{aligned}$$

where

$$p_\alpha(y) = \operatorname{sech}^2(\alpha y) \tanh(\alpha y) \quad \text{and} \quad q_\alpha(y) = \operatorname{sech}^4(\alpha y) [2 - \cosh(2\alpha y)].$$

Clearly, we have

$$\frac{\pi^2}{4} h_\alpha(0) = 1 + 2\alpha \int_0^1 [-2p_\alpha(y) + \alpha(1-y)q_\alpha(y)] \cos \frac{\pi y}{2} dy. \quad (14)$$

Note that $\alpha^2 > 0$ iff $K < 0$ in which case, $\sup_{x \in (0,1)} h_\alpha(x) = h_\alpha(0)$ (see part (c) of the proof below). Now, as an application of Lemma 2.2, we have

$$\int_0^1 \operatorname{sech}^2(\alpha r) g^2 dr \leq h_\alpha(0) \int_0^1 g'^2 dr, \quad K \leq 0.$$

In particular, letting $\alpha \downarrow 0$, it follows that

$$\int_0^1 g^2 dr \leq \frac{4}{\pi^2} \int_0^1 g'^2 dr.$$

By Lemma 2.2 again, we have for all $\alpha \geq 0$,

$$\alpha^2 \int_0^1 \operatorname{sech}^2(\alpha r) g^2 dr \leq \alpha^2 h_\alpha(0) \int_0^1 g'^2 dr,$$

and then by (13),

$$\pi^2 s(1-s) \leq \bar{\lambda} + s(d-1)\alpha^2 \frac{\pi^2}{4} h_\alpha(0) = \bar{\lambda} - s \frac{KD^2}{4} \frac{\pi^2}{4} h_\alpha(0). \quad (15)$$

Noticing that among q_α , p_α , $\operatorname{sech}^2(\alpha r)$, and $\tanh(\alpha r)$, the former comes from the derivative of the latter, starting from (14), by using the integral by parts formula three times, one finally sees that

$$\frac{\pi^2}{4} h_\alpha(0) = \frac{\pi^2}{4} \int_0^1 \left(\cos \frac{\pi y}{2} + \frac{\pi}{2} (1-y) \sin \frac{\pi y}{2} \right) \frac{\tanh(\alpha y)}{\alpha} dy.$$

The right-hand side gives us M_α in terms of the integration by parts formula ($\sin = -d \cos$). Applying this to (15), Proposition 1.3 now follows from the first assertion of Proposition 1.1 in the case of $K \leq 0$.

(b) Next, let $K > 0$ and set

$$a(r) = \bar{\lambda} + s(d-1)\alpha^2 \operatorname{sech}^2(\alpha r) = \bar{\lambda} - s(d-1)|\alpha|^2 \sec^2(|\alpha|r).$$

In other words, we do not separate this a into two parts as in (13). By (9), in order that $a > 0$ on $(0, 1)$, it suffices that

$$\left(\frac{\pi}{2\sqrt{d-1}|\alpha|} + \frac{\sqrt{d-1}|\alpha|}{2\pi} \right) \cos |\alpha| > 1.$$

Once the assertion is proved under this assumption, one may replace “ $>$ ” here with “ $\geq$ ” by a limiting procedure. For the other part of α , simply regard α as the constant α_0 (or $\alpha_0 - \varepsilon$ if necessary) since then $a(r)$ is upper bounded uniformly in r by a positive constant on that subinterval of α . We remark that even though the restriction here can be relaxed a little we do not do so since the estimate is mainly essential for small K (or for small $|\alpha|$). Applying Lemma 2.2 to this a , we obtain

$$\hat{\lambda}^{-1} \leq \sup_{x \in (0,1)} h(x) : \quad h(x) = 4\bar{\lambda}/\pi^2 + s(d-1)\alpha^2 h_\alpha(x),$$

where h_α is the same as used in proof (a). Note that in the present case, we have $\alpha^2 < 0$ and then

$$\sup_{x \in (0,1)} h(x) = 4\bar{\lambda}/\pi^2 + s(d-1)\alpha^2 \inf_{x \in (0,1)} h_\alpha(x).$$

This is the main different point to the previous case of $K \leq 0$. Luckily, we then have $\inf_{x \in (0,1)} h_\alpha(x) = h_\alpha(0)$ (see part (c) of the proof below). In this case, since $\alpha = i|\alpha|$, it may be more convenient to rewrite h_α as

$$\begin{aligned} \frac{\pi^2}{4} h_\alpha(x) &= \sec^2(|\alpha|x) - 2|\alpha| \sec \frac{\pi x}{2} \left[|\alpha|(1-x) \int_0^x q_\alpha^+(y) \cos \frac{\pi y}{2} dy \right. \\ &\quad \left. + \int_x^1 [-2p_\alpha^+(y) + |\alpha|(1-y)q_\alpha^+(y)] \cos \frac{\pi y}{2} dy \right], \end{aligned}$$

where

$$p_\alpha^+(y) = \sec^2(|\alpha|y) \tan(|\alpha|y) \quad \text{and} \quad q_\alpha^+(y) = \sec^4(|\alpha|y) [2 - \cos(2|\alpha|y)].$$

From this, one sees that h_α is always real for any $K \in \mathbb{R}$. The remainder of the proof is similar to proof (a) above.

(c) To see that $\sup_{x \in (0,1)} h_\alpha(x) = h_\alpha(0)$ when $K \leq 0$ and $\inf_{x \in (0,1)} h_\alpha(x) = h_\alpha(0)$ when $K \geq 0$, it is helpful to look at first three figures (Figures 10–12)

for the latter case. Fig. 10 shows that the surface is rather regular. But it may not be very clear that the minimum is attached at $x = 0$ for each fixed $|\alpha|$, so two more figures (11 and 12) are included. The pictures in case of $K \leq 0$ are parallel. Based on the observation, it should not be hard to present an analytic proof but we prefer to omit the details here for saving the space. $\square$

Once again, as indicated in the proof of Lemma 2.2, Proposition 1.3 uses only the first step of our general approximating procedure for the first eigenvalue $\hat{\lambda}$.

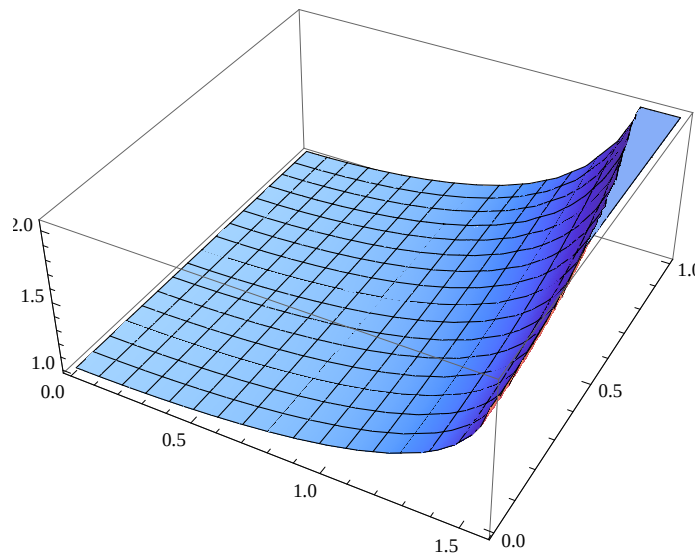


Figure 10 When $K \geq 0$, the surface of $\pi^2 h_\alpha(x)/4$ in the interval $|\alpha| \in (0, 1.57075)$ and $x \in (0, 1)$.

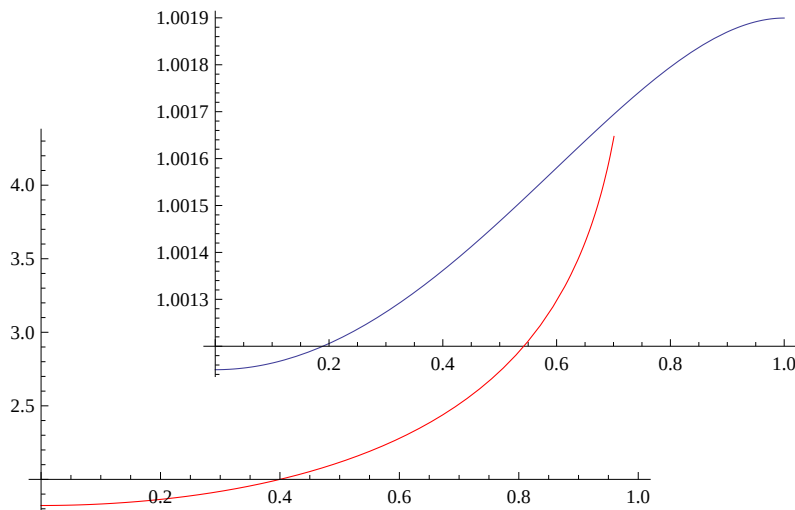


Figure 11–12 The curve of $\pi^2 h_\alpha(x)/4$, $x \in (0, 1)$ when $K \geq 0$ for $|\alpha| = 0.1$ or $|\alpha| = 1.57$, respectively.

To conclude the paper, we make some remarks about the methods used in the paper. Recall that all of the results (3), (5)–(8), and (10) are an application of a coupling to a carefully designed (case by case) distance (equivalent to the Riemannian one). As mentioned in [2; Theorem 6.2], the method works for more general “cost” functions, not necessarily a distance. With coupling method in mind, the boundaries of the reduced process with operator $\bar{L}$ given in the proof of Theorem 1.2 is natural. Then the “ δ_1 ” (resp. “ δ'_1 ”) part of Proposition 1.1 says that we do have a universal distance which is equivalent to the Riemannian one and provides us a universal lower bound $4D^{-2}\delta_1^{-1}$. Thus, all of these results can be regarded as an application of the general variational formula given in [9]. However, for part “ δ_1^* ”, as a dual of $\bar{L}$, it has a different probabilistic meaning. The use of the dual technique is an essential new point of the present paper. Finally, in Lemma 2.1 (or Proposition 1.3), the original boundary conditions are also dualled, at the same time, the operator is changed. It is mainly a specific comparison result, one can not say that the two operators used in Lemma 2.1 have the same principal eigenvalue.

Acknowledgments. The main results of the paper have been presented at “International Conference on Stochastic Partial Differential Equations and Related Topics (April 25 - 30, 2011; Chern Institute of Math.)” and at “Stochastic Analysis and Application to Financial Mathematics, in honor of Professor Jia-An Yan’s 70th Birthday (July 4–6, 2011; AMSS, CAS)”. The author acknowledges the organizing groups for their kind invitation and financial support.

References

- [1] Bérard, P.H., Besson, G. and Gallot, S. (1985). Sur une inégalité isopérimétrique qui généralise celle de Paul Lévy-Gromov. *Invent. Math.*, 80:295–308.
- [2] Chen, M.F. (1994). Optimal Markovian couplings and applications. *Acta Math. Sin. New Ser.* 10:3, 1994, 260–275.
- [3] Chen, M.F. (2000). Explicit bounds of the first eigenvalue. *Sci. China (A)* 43:10, 1051–1059.
- [4] Chen, M.F. (2001). Variational formulas and approximation theorems for the first eigenvalue. *Sci. China (A)* 44:4, 409–418
The last three and related papers with some complements are collected in book [4] at the author’s homepage.
- [5] Chen, M.F. (2005). *Eigenvalues, Inequalities, and Ergodic Theory*. Springer, London.
- [6] Chen, M.F. (2010). Speed of stability for birth–death processes. *Front. Math. China* 5:3, 379–515.
- [7] Chen, M.F., Scacciatelli, E. and Yao, L. (2002). Linear approximation of the first eigenvalue on compact manifolds. *Sci. Sin.(A)*, 45:4, 450–461.
- [8] Chen, M.F. and Wang, F.Y. (1993). Application of coupling method to the first eigenvalue on manifold. *Sci. Sin. (A)*, 23:11, 1130–1140, 1993 (Chinese Edition); 37:1, 1–14, 1994 (English Edition).

- [9] Chen, M.F. and Wang, F.Y. (1997). General formula for lower bound of the first eigenvalue on Riemannian manifolds. *Sci. Sin.* 40:4, 384–394.
- [10] Jia, F. (1991). Estimate of the first eigenvalue of a compact Riemannian manifold with Ricci curvature bounded below by a negative constant (in Chinese). *Chin. Ann. Math.* 12A:4, 496–502.
- [11] Lichnerowicz, A. (1958). *Géométrie des Groupes des Transformations*. Dunod, Paris.
- [12] Ling, J. (2006). Lower bounds of the eigenvalues of compact manifolds with positive Ricci curvature. *Proc. Amer. Math. Soc.* 134:10, 3071–3079.
- [13] Ling, J. (2008). An exact solution to an equation and the first eigenvalue of a compact manifold. *Illinois J. Math.* 51:3, 853–860.
- [14] Shi, Y.M. and Zhang, H.C. (2007). Lower bounds for the first eigenvalue on compact manifolds (in Chinese). *Chin. Ann. Math.* 28A:6, 863–866.
- [15] Schoen, S. and Yau, S.T. (1988). *Differential Geometry* (in Chinese). Science Press, Beijing, China. English Translation: *Lectures on Differential Geometry*, International Press (1994).
- [16] Wang, F.Y. (2004). *Functional Inequalities, Markov Processes, and Spectral Theory*. Science Press, Beijing.
- [17] Xu S.L. and Pang H.D. (2001). Estimate of the first eigenvalue on compact manifolds. *Math. Appl.* 14:1, 116–119.
- [18] Xu S.L., Yang F.Y. and Xu, X. (2002). Estimate of the first eigenvalue of compact manifold with positive Ricci curvature (in Chinese). *Math. Appl.* 15:2, 85–88.
- [19] Yang, H.C. (1990). Estimate of the first eigenvalue for a compact Riemannian manifold. *Sci. Sin. (A)* 33:1, 39–51.
- [20] Zhao, D. (1999). Eigenvalue estimate on a compact Riemannian manifold. *Sci. in China, Ser. A* 42:9, 897–904.
- [21] Zhong, J.Q. and Yang, H.C. (1984). Estimates of the first eigenvalue of a compact Riemannian manifold. *Sci. Sin. Ser. A* 27:12, 1265–1273.

School of Mathematical Sciences, Beijing Normal University,
 Laboratory of Mathematics and Complex Systems (Beijing Normal University),
 Ministry of Education,
 Beijing 100875, The People's Republic of China.
 E-mail: mfchen@bnu.edu.cn
 Home page: http://math.bnu.edu.cn/~chenmf/main_eng.htm

Remark in addition^[2012/9/16] As mentioned in the proof of Proposition 1.1, the results in this paper, except Remark 1.6, can be generalized to the general setup studied in [9]. In particular, the results remain the same for the first nontrivial Neumann eigenvalue of Laplacian on a compact manifold with convex boundary. Interestingly, the same results (with obvious change of notation) also hold for Finsler-Laplacian. Refer to

Wang, G. and Xia, C. A sharp lower bound for the first nontrivial eigenvalue on Finsler manifolds [arXiv, 2011]

and [5; the second paragraph on page 58] for details.

Appendix (Additional figures)

We begin with the comparison of (3) and (2), and of (6) and (5). It is easy to check that estimate (3) $\geq$ estimate (2) iff

$$1 - (1 - \cos^d |\alpha|) \left[\int_0^{\pi/2} \cos^{d-1} t \, dt \right] / \left[\int_0^{|\alpha|} \cos^{d-1} t \, dt \right] \geq 0.$$

Fig. 13 shows the curves of the function on the left-hand side for different d .

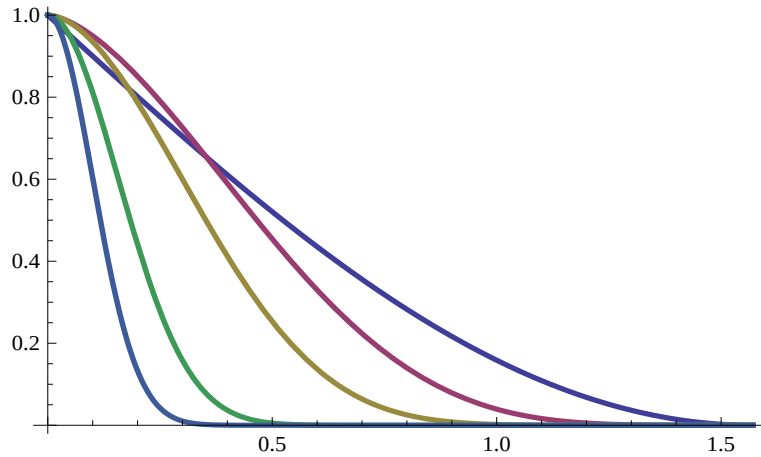


Figure 13 The curves of $|\alpha|$ on $(0, \pi/2)$, from right to left viewing at bottom, correspond to $d = 2, 5, 10, 40$ and 100 , successively.

To see that (6) improves (5), it suffices that

$$\sqrt{1 + \frac{8(d-1)\alpha^2}{\pi^4}} - e^{-(d-1)\alpha} \cosh^{d-1} \alpha \geq 0.$$

Fig. 14 shows the curves of the function on the left-hand side for different d .

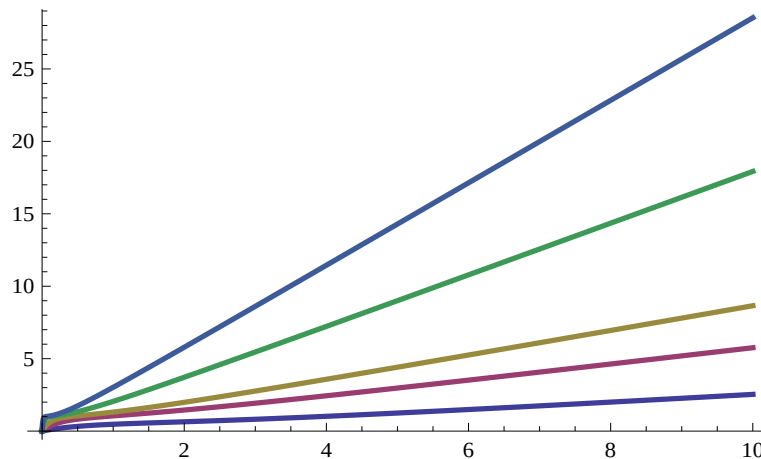


Figure 14 The curves of α on $(0, 10)$, from top to bottom, correspond to $d = 2, 5, 10, 40$ and 100 , successively.

The reason we do not publish the figures below is that some of them (Fig. 19 for instance) can not be printed out clearly without using colors.

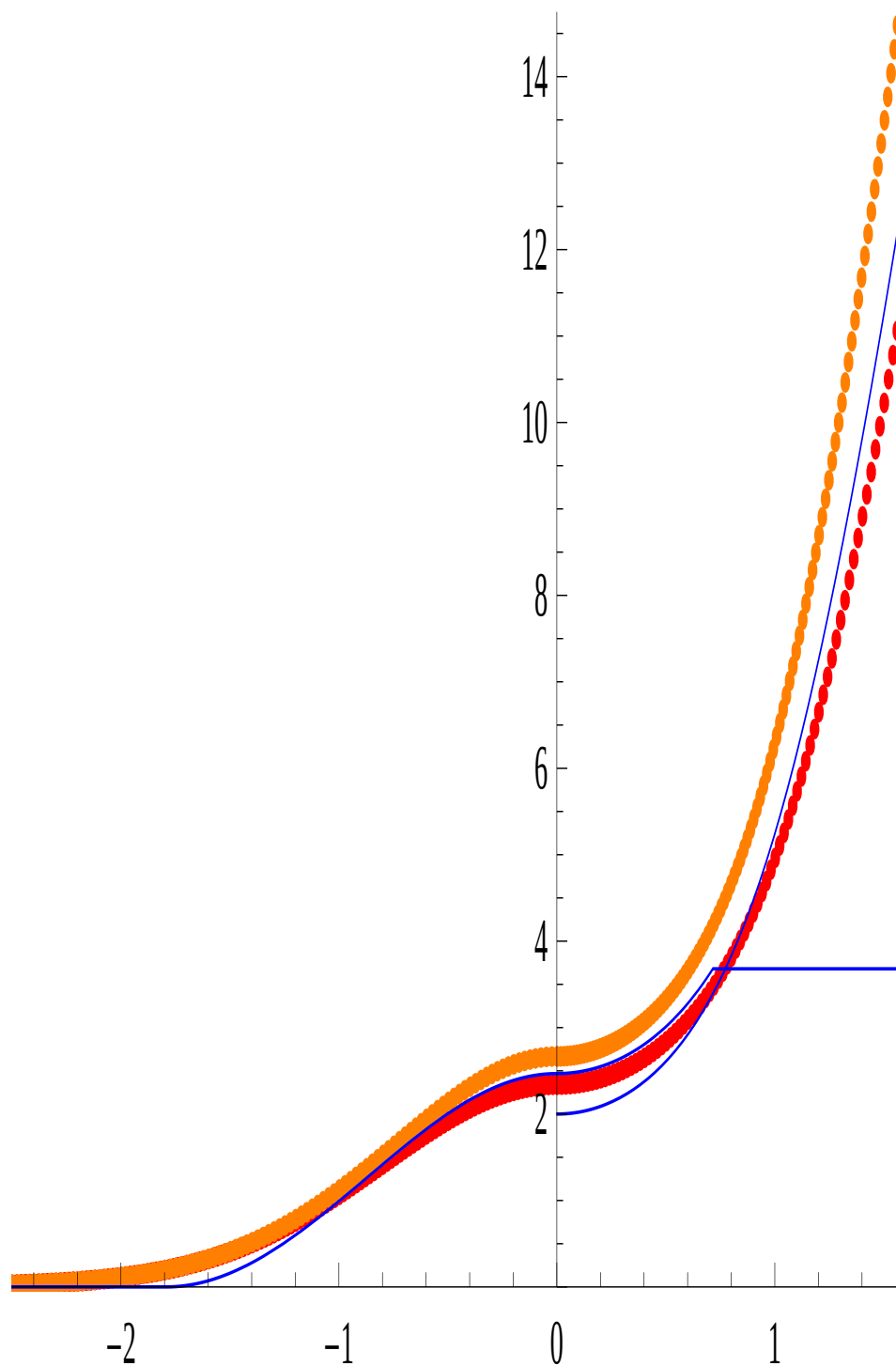


Figure 15 (Extension of Fig.9) $d = 5$, $\alpha = \sqrt{-\text{sgn}(x)|x|}$, $x \in (-2.5, \pi/2)$.

The next figure is used for Remark 1.5 and the last three figures are for Remark 1.6. Certainly, “ $|\alpha| = \pi/2$ ” here means at the right-endpoint ($K > 0$).

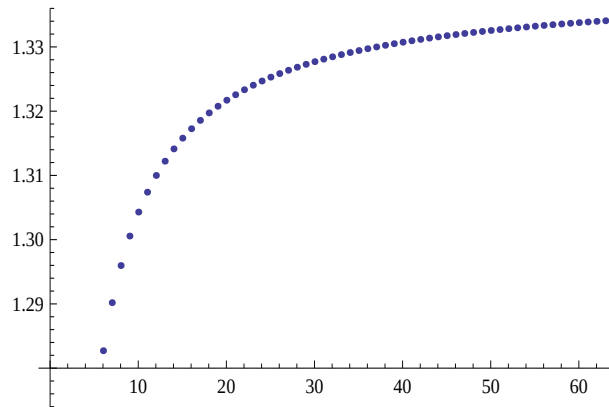


Figure 16 The ratio $\delta_1^*/\delta_1^{*'}$ at $|\alpha| = \pi/2$ when d varies over $\{2, 3, \dots, 63\}$.

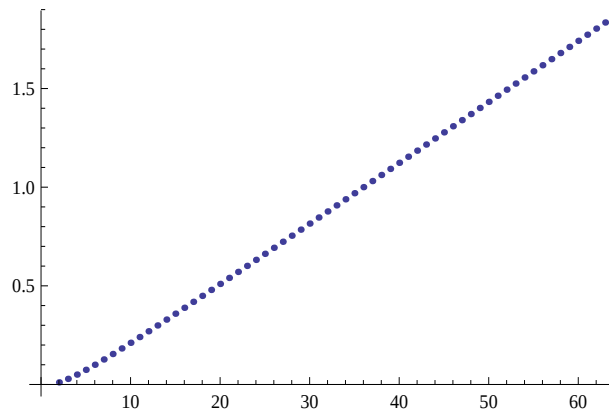


Figure 17 The difference $\eta_0 - \bar{\lambda} \in [0, 1.85]$ at $|\alpha| = \pi/2$ when d varies over $2, 3, \dots, 63$. $\bar{\lambda} \approx 155$ when $d = 63$.

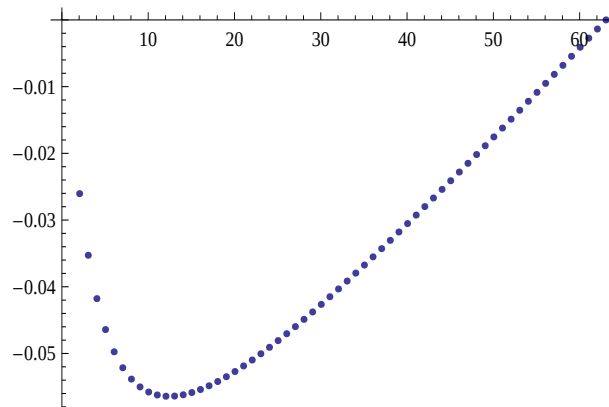


Figure 18 Let $\eta_{\pi/2}^{63}$ be the $\eta_{\pi/2}$ corresponding to that $\gamma_{\pi/2}$ defined at $d = 63$. Then the figure is the difference $\eta_{\pi/2}^{63} - \bar{\lambda} \in [-0.056, 0]$ at $|\alpha| = \pi/2$ when d varies over $2, 3, \dots, 63$. $\bar{\lambda} \approx 155$ when $d = 63$.

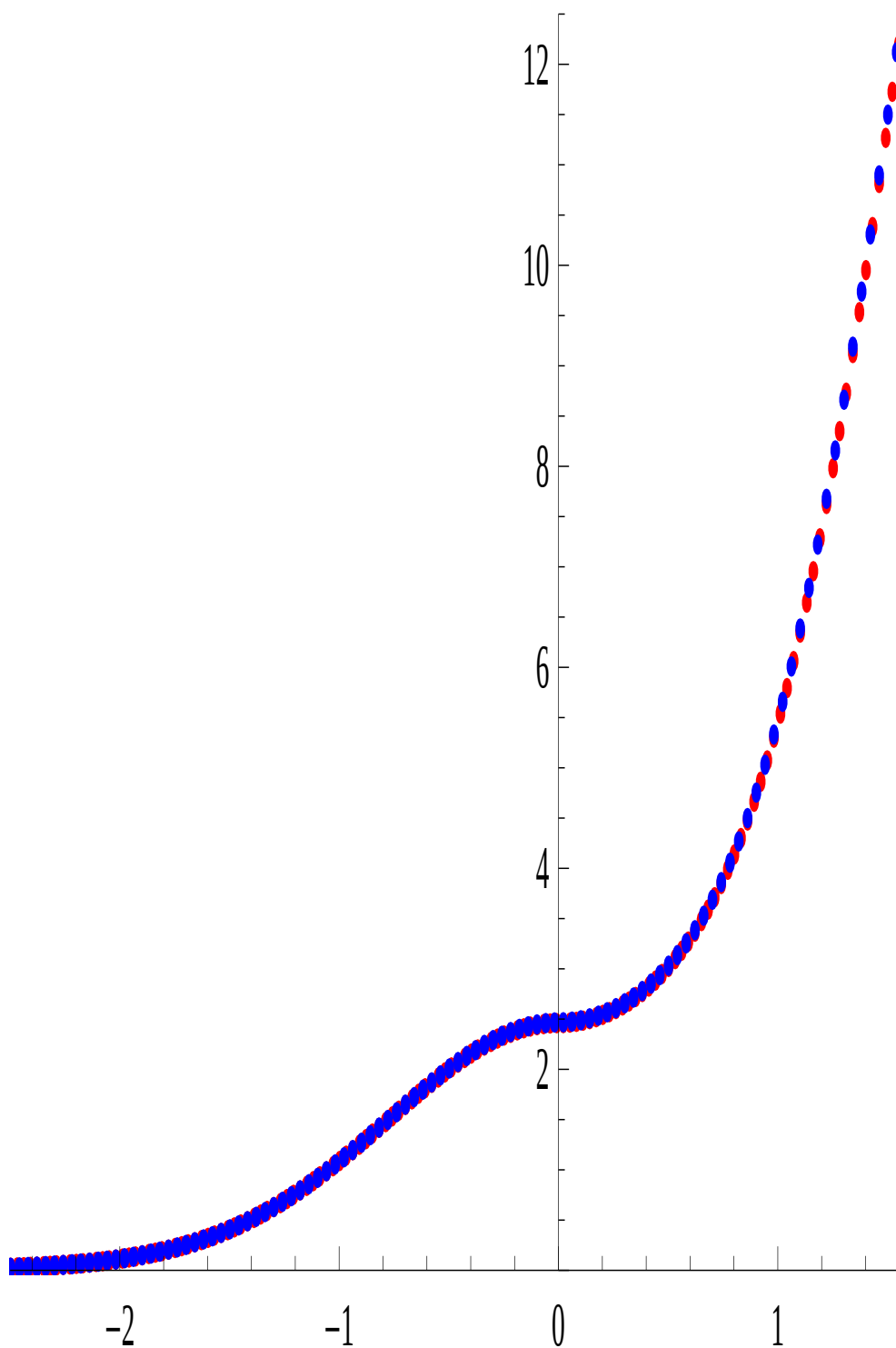


Figure 19 In the case of Fig.15, the upper bound η_0 (blue) and lower bound $\eta_{\pi/2}$ (red) of $\bar{\lambda}$. Here $\gamma_{\pi/2} \approx 0.367$ since $d = 5$.

Lower bounds of principal eigenvalue in dimension one

Mu-Fa Chen

(Beijing Normal University, Beijing 100875, China)

Abstract

For the principal eigenvalue with bilateral Dirichlet boundary condition, the so-called basic estimates were originally obtained by capacity method. The Neumann case (i.e., the ergodic case) is even harder, and was deduced from the Dirichlet one plus a use of duality and the coupling method. In this paper, an alternative and more direct proof for the basic estimates is presented. The estimates in the Dirichlet case are then improved by a typical application of a recent variational formula. As a dual of the Dirichlet case, the refine problem for bilateral Neumann boundary condition is also treated. The paper starts with the continuous case (one-dimensional diffusions) and ends at the discrete one (birth–death processes). Possible generalization of the results studied here is discussed at the end of the paper.

1 Introduction (continuous case)

Consider an elliptic operator

$$L = a(x) \frac{d^2}{dx^2} + b(x) \frac{d}{dx} \quad (1)$$

(with $a > 0$) on $E := (-M, N)$ ($M, N \leq \infty$). Define a function $C(x)$:

$$C(x) = \int_o^x \frac{b}{a}, \quad x \in E,$$

Received November 30, 2011; accepted May 31, 2012

2000 *Mathematics Subject Classifications.* 60J60, 34L15, 26D10.

Key words and phrases. Principal eigenvalue, lower estimate, variational formula, one-dimensional diffusion, birth–death process.

Research supported in part by the National Natural Science Foundation of China (No. 11131003), and by the “985” project from the Ministry of Education in China.

where $o \in E$ is a reference point. Here and in what follows, the Lebesgue measure dx is often omitted. It is convenient for us to define two measures μ and ν :

$$\mu(dx) = \frac{e^{C(x)}}{a(x)} dx, \quad \nu(dx) = e^{-C(x)} dx. \quad (2)$$

As usual, the norm on $L^2(\mu)$ is denoted by $\|\cdot\|$. Define

$\mathcal{A}(-M, N)$ = the set of absolutely continuous functions on $(-M, N)$,

$\mathcal{A}_0(-M, N) = \{f \in \mathcal{A}(-M, N) : f \text{ has a compact support}\},$

$$D(f) = \int_{-M}^N f'^2 e^C, \quad f \in \mathcal{A}(-M, N), \quad M, N \leq \infty.$$

Here $D(f)$ is allowed to be ∞ . We are interested in the following eigenvalues:

$$\lambda^{\text{DD}} = \inf\{D(f) : f \in \mathcal{A}_0(-M, N), \|f\| = 1\}, \quad (3)$$

$$\lambda^{\text{NN}} = \inf\{D(f) : f \in \mathcal{A}(-M, N), \mu(f) = 0, \|f\| = 1\}, \quad (4)$$

where $\mu(f) = \int_E f d\mu$.

The basic estimates, of λ^{DD} for instance, given in [3] are as follows:

$$(4\kappa^{\text{DD}})^{-1} \leq \lambda^{\text{DD}} \leq (\kappa^{\text{DD}})^{-1}, \quad (5)$$

where

$$(\kappa^{\text{DD}})^{-1} = \inf_{x < y} [\nu(-M, x)^{-1} + \nu(y, N)^{-1}] \mu(x, y)^{-1}, \quad \mu(x, y) := \int_x^y d\mu. \quad (6)$$

The proof for the upper estimate is already straightforward, simply using the classical variational formula for λ^{DD} (cf. [3; Proof (b) of Theorem 8.2]). However, the proof for the lower estimate is much harder and deeper, using capacity theory (cf. [3; Sections 8, 10]). Even through the capacity tool is suitable in a general setup (cf. [6], [2; Theorems 7.1 and 7.2], [7; Chapter 2]), it is still expected to have a direct proof (avoiding capacity) in such a concrete situation. This is done at the beginning of the next section. Surprisingly, the simple proof also works in the ergodic case for which the original proof is based on (5) plus a use of the duality and the coupling technique. The main body of the paper is devoted to an improvement of the basic lower estimate given in (5), as stated in Corollary 1.1 below. The result can be regarded as a typical application of a recent variational formula ([4; Theorem 4.2] or Theorem 2.1 below). This note is an addition to the recent papers [3, 4] from which one can find the motivation of the study on the topic and further references. It is remarkable that the new result makes the whole analytic proof for the basic estimates more elementary.

Here is our first main result which is a refinement of [4; Corollary 4.3].

Corollary 1.1 (1) We have

$$\lambda^{\text{DD}} \geq (\underline{\kappa}^{\text{DD}})^{-1} \geq (4\kappa^{\text{DD}})^{-1},$$

where κ^{DD} is given in (6) and $\underline{\kappa}^{\text{DD}}$ is defined by (12) below.

(2) Let $\mu(-M, N) < \infty$. Then assertion (1) holds if the codes DD are replaced by NN (for instance, $\lambda^{\text{NN}} \geq (\underline{\kappa}^{\text{NN}})^{-1}$) and the measures μ and ν are exchanged.

The remainder of the paper is organized as follows. In the next section, we present shortly an alternative proof of the estimates in (5). The proof shows one of the main new ideas of the paper. Then we prove Corollary 1.1. Two illustrating examples are also included in this section. The discrete analog of Corollary 1.1 is presented in the third section.

2 Proofs and Examples

Proof of (5).

Let $\theta \in (-M, N)$ be a reference point. Define

$$\delta_{\theta}^{-} = \sup_{z \in (-M, \theta)} \nu(-M, z) \mu(z, \theta), \quad \delta_{\theta}^{+} = \sup_{z \in (\theta, N)} \mu(\theta, z) \nu(z, N).$$

As will be remarked in the next section, we may assume that $\delta_{\theta}^{\pm} < \infty$. Otherwise, the problem becomes either trivial or degenerated. Next, denote by $\lambda_{\theta}^{\pm}$ the principal eigenvalue on $(-M, \theta)$ and (θ, N) , respectively, with common reflecting (Neumann) boundary at θ and absorbing (Dirichlet) boundary at $-M$ (and N) provided $M < \infty$ ($N < \infty$). Actually, by an approximating procedure, one may assume that $M, N < \infty$ (cf. [3; Proof of Corollary 7.9]). Next, by a splitting technique, one may choose $\theta = \bar{\theta}$ to be the unique solution to the equation $\lambda_{\bar{\theta}}^{-} = \lambda_{\bar{\theta}}^{+}$. Then they coincide with λ^{DD} since by [5; Theorem 1.1], we have

$$\lambda_{\theta}^{-} \wedge \lambda_{\theta}^{+} \leq \lambda^{\text{DD}} \leq \lambda_{\theta}^{-} \vee \lambda_{\theta}^{+}$$

for every $\theta \in (-M, N)$, where $x \wedge y = \min\{x, y\}$ and dually $x \vee y = \max\{x, y\}$. Alternatively, $\bar{\theta}$ is the root of the derivative of the eigenfunction of λ^{DD} by [5; Proposition 1.3] and the monotonicity of the eigenfunctions of $\lambda_{\theta}^{\pm}$. From now on in this proof, we fix this $\bar{\theta}$. For given $\varepsilon > 0$, let $\bar{x} < \bar{\theta}$ and $\bar{y} > \bar{\theta}$ satisfy

$$\nu(-M, \bar{x}) \mu(\bar{x}, \bar{\theta}) \geq \delta_{\bar{\theta}}^{-} - \varepsilon, \quad \mu(\bar{\theta}, \bar{y}) \nu(\bar{y}, N) \geq \delta_{\bar{\theta}}^{+} - \varepsilon,$$

respectively. As a continuous analog of [1; Theorem 1.1], we have

$$[(\lambda^{\text{DD}})^{-1} =] \quad (\lambda_{\bar{\theta}}^{+})^{-1} \leq 4\delta_{\bar{\theta}}^{+} \quad [\leq 4\mu(\bar{\theta}, \bar{y}) \nu(\bar{y}, N) + 4\varepsilon].$$

Hence,

$$[(\lambda^{\text{DD}})^{-1} - 4\varepsilon] \nu(\bar{y}, N)^{-1} \leq 4\mu(\bar{\theta}, \bar{y}).$$

In parallel, we have

$$[(\lambda^{\text{DD}})^{-1} - 4\varepsilon] \nu(-M, \bar{x})^{-1} \leq 4\mu(\bar{x}, \bar{\theta}).$$

Summing up the last two inequalities, it follows that

$$[(\lambda^{\text{DD}})^{-1} - 4\varepsilon] [\nu(-M, \bar{x})^{-1} + \nu(\bar{y}, N)^{-1}] \leq 4\mu(\bar{x}, \bar{y}).$$

That is,

$$(\lambda^{\text{DD}})^{-1} - 4\varepsilon \leq 4[\nu(-M, \bar{x})^{-1} + \nu(\bar{y}, N)^{-1}]^{-1} \mu(\bar{x}, \bar{y}).$$

In view of (6), the right-hand side is bounded from above by $4\kappa^{\text{DD}}$. Since ε is arbitrary, we have proved the lower estimate in (5). A direct proof for the upper one in (5) is presented in [3; Proof (b) of Theorem 8.2]. $\square$

Proof of the dual of (5):

$$(4\kappa^{\text{NN}})^{-1} \leq \lambda^{\text{NN}} \leq (\kappa^{\text{NN}})^{-1},$$

where

$$(\kappa^{\text{NN}})^{-1} = \inf_{x < y} [\mu(-M, x)^{-1} + \mu(y, N)^{-1}] \nu(x, y)^{-1}.$$

By exchanging “Neumann” and “Dirichlet”, the splitting point $\theta = \bar{\theta}$ is now a common Dirichlet boundary and $-M$ becomes Neumann boundary if $M < \infty$ (and so is N). In other words, $\bar{\theta}$ is the unique root of the eigenfunction of λ^{NN} . Now, in the proof above, we need only to use [1; Theorem 3.3] instead of [5; Theorem 1.1] and making the exchange of μ and ν . We have thus returned to the role mentioned in [4]: exchanging the boundary condition “Neumann” and “Dirichlet” simultaneously leads to the exchange of the measures μ and ν .

Here is a direct proof for the upper estimate. Given $x, y \in (-M, N)$ with $x < y$, let $\bar{\theta} = \bar{\theta}(x, y)$ be the unique solution to the equation

$$\begin{aligned} & \mu(-M, x)\nu(x, \theta) + \int_x^\theta \mu(dz)\nu(z, \theta) \\ &= \mu(y, N)\nu(\theta, y) + \int_\theta^y \mu(dz)\nu(\theta, z), \quad \theta \in (x, y). \end{aligned}$$

Next, define

$$f(z) = -\mathbb{1}_{\{z \leq \bar{\theta}\}} \nu(x \vee z, \bar{\theta}) + \mathbb{1}_{\{z > \bar{\theta}\}} \nu(\bar{\theta}, y \wedge z).$$

Then $\mu(f) = 0$ by the definition of $\bar{\theta}$. We have

$$\int_{-M}^N |f'|^2 e^C = \nu(x, \bar{\theta}) + \nu(\bar{\theta}, y) = \nu(x, y).$$

Moreover,

$$\begin{aligned} \int_{-M}^N (f - \pi(f))^2 d\mu &= \int_{-M}^N f^2 d\mu \\ &> \int_{-M}^x f^2 d\mu + \int_y^N f^2 d\mu \\ &= \mu(-M, x) \nu(x, \bar{\theta})^2 + \mu(y, N) \nu(\bar{\theta}, y)^2. \end{aligned}$$

Note that the function

$$\gamma(x) = \alpha x^2 + \beta(1-x)^2, \quad x \in (0, 1), \quad \alpha, \beta > 0$$

achieves its minimum $(\alpha^{-1} + \beta^{-1})^{-1}$ at $x^* = (1 + \beta/\alpha)^{-1}$. As an application of this result with

$$\alpha = \mu(-M, x), \quad \beta = \mu(y, N), \quad x = \nu(x, \bar{\theta})/\nu(x, y),$$

we get

$$\int_{-M}^N (f - \pi(f))^2 d\mu \geq \frac{\nu(x, y)^2}{\mu(-M, x)^{-1} + \mu(y, N)^{-1}}.$$

Hence

$$\frac{\int_{-M}^N (f - \pi(f))^2 d\mu}{\int_{-M}^N |f'|^2 e^C} \geq \frac{\nu(x, y)}{\mu(-M, x)^{-1} + \mu(y, N)^{-1}}.$$

Making supremum with respect to $x < y$, we obtain the required κ^{NN} . $\square$

It is remarkable that although the last proof is in parallel to the previous one, it does not depend on (5). This is rather lucky since in other cases, part (2) of Corollary 1.1 for instance, we do not have such a direct proof.

From now on, unless otherwise stated, we restrict ourselves to the Dirichlet case. For fixed θ , much knowledge on $\lambda_\theta^\pm$ is known (variational formulas, approximating procedure and so on, refer to [2, 3] for instance). Of which, only a little is used in the proof above. For instance, by [5; Corollary 1.5], we have

$$\left(\sup_\theta [\delta_\theta^- \wedge \delta_\theta^+] \right)^{-1} \geq \lambda^{\text{DD}} \geq \left(4 \inf_\theta [\delta_\theta^- \vee \delta_\theta^+] \right)^{-1}.$$

Thus, if we choose $\bar{\theta}$ to be the solution of equation $\delta_{\bar{\theta}}^- = \delta_{\bar{\theta}}^+$, then we obtain

$$(\delta_{\bar{\theta}}^-)^{-1} \geq \lambda^{\text{DD}} \geq (4\delta_{\bar{\theta}}^-)^{-1}$$

which is even more compact than (5) in view of the comparison of κ^{DD} and $\delta_\theta^\pm$. The problem is that $\bar{\theta}$, especially the one used in the first proof of this section, is usually not explicitly known and so a large part of the known results for $\lambda_\theta^\pm$ are not practical. To overcome this difficulty, the first proof above uses two parameters x and y to get κ^{DD} and then to obtain the explicit lower estimates

(5). For our main result Corollary 1.1, the fixed point $\bar{\theta}$ used in the proof of (5) is replaced by its mimic given in (9) below for suitable test function f . The difference is that equation (9) is explicit but not the one for $\bar{\theta}$ used in the first proof above.

Proof of Corollary 1.1 (1).

By [3] or [4], we have known that part (2) of Corollary 1.1 is a dual of part (1). Hence in what follows, we need study part (1) only.

The first inequality in part (1) comes from [4; Corollary 4.3]. Thus, it suffices to prove the last inequality in part (1).

Even though it is not completely necessary, we assume that $M, N < \infty$ until the last paragraph of the proof.

For a given $f \in \mathcal{C}_+$:

$$\mathcal{C}_+ = \{f \in \mathcal{C}(-M, N) : f > 0 \text{ on } (-M, N), f(-M+0) = 0 \text{ and } f(N-0) = 0\},$$

define

$$h^-(z) = h_f^-(z) = \int_{-M}^z e^{-C(u)} du \int_u^\theta \frac{e^C f}{a}, \quad z \leq \theta, \quad (7)$$

$$h^+(z) = h_f^+(z) = \int_z^N e^{-C(u)} du \int_\theta^u \frac{e^C f}{a}, \quad z > \theta, \quad (8)$$

where $\theta = \theta(f) \in (-M, N)$ is the unique root of the equation:

$$h^-(\theta) = h^+(\theta) \quad (9)$$

provided $h_f^\pm < \infty$. The uniqueness of θ should be clear since on $(-M, N)$, as a function of θ , $h^-(\theta)$ is continuously increasing from zero to $h^-(N-0) > 0$ and $h^+(\theta)$ is continuously decreasing from $h^+(-M+0) > 0$ to zero. Next, define

$$H^\pm(f) = h^\pm / f.$$

Then we have the following variational formula.

Theorem 2.1 [4; Theorem 4.2 (1)] Assume that $\nu(-M, N) < \infty$. Then

$$\lambda^{\text{DD}} = \sup_{f \in \mathcal{C}_+} \left\{ \left[\inf_{z \in (-m, \theta)} H^-(f)(z)^{-1} \right] \bigwedge \left[\inf_{z \in (\theta, N)} H^+(f)(z)^{-1} \right] \right\}. \quad (10)$$

We remark that in the original statement of [4; Theorem 4.2 (1)], the boundary condition “ $f(-M+0) = 0$ and $f(N-0) = 0$ ” is ignored. The condition is added here for the use of the operators $I^\pm$ (different from $H^\pm$) to be defined later. However, the conclusion (10) remains true since the eigenfunction of λ^{DD} does satisfy this condition.

We now fix $x < y$ and let $f = f^{x,y}$:

$$f^{x,y}(s) = \begin{cases} \sqrt{\varphi^+(y)\varphi^-(s \wedge x)/\varphi^-(x)}, & s \leq y \\ \sqrt{\varphi^+(s)}, & s \geq y, \end{cases} \quad (11)$$

where

$$\varphi^-(s) = \nu(-M, s) \quad \text{and} \quad \varphi^+(s) = \nu(s, N).$$

Certainly, here we assume that $\varphi^\pm < \infty$ (which is automatic whenever $M, N < \infty$). Clearly, $f^{x,y} \in \mathcal{C}_+$. Here we are mainly interested in those pair $\{x, y\}$ having the property $x < \theta < y$. As proved in [4], the quantity $\underline{\kappa}^{\text{DD}}$:

$$\underline{\kappa}^{\text{DD}} = \inf_{x < y} \left[\sup_{z \in (-M, \theta)} H^-(f^{x,y})(z) \right] \bigvee \left[\sup_{z \in (\theta, N)} H^+(f^{x,y})(z) \right] \quad (12)$$

used in Corollary 1.1 (1) has an explicit expression:

$$\begin{aligned} \inf_{x < y} \left\{ \sup_{z \in (-M, x)} \left[\frac{1}{\sqrt{\varphi^-(z)}} \mu \left((\varphi^-)^{3/2} \mathbb{1}_{(-M, z)} \right) + \sqrt{\varphi^-(z)} \mu \left(\sqrt{\varphi^-} \mathbb{1}_{(z, x)} \right) \right. \right. \\ \left. \left. + \sqrt{\varphi^-(z) \varphi^-(x)} \mu(x, \theta) \right] \right. \\ \bigvee \left[\frac{1}{\sqrt{\varphi^-(x)}} \mu \left((\varphi^-)^{3/2} \mathbb{1}_{(-M, x)} \right) + \mu \left(\varphi^- \mathbb{1}_{(x, \theta)} \right) \right] \\ \left. \bigvee \sup_{z \in (y, N)} \left[\frac{1}{\sqrt{\varphi^+(z)}} \mu \left((\varphi^+)^{3/2} \mathbb{1}_{(z, N)} \right) + \sqrt{\varphi^+(z)} \mu \left(\sqrt{\varphi^+} \mathbb{1}_{(y, z)} \right) \right. \right. \\ \left. \left. + \sqrt{\varphi^+(z) \varphi^+(y)} \mu(\theta, y) \right] \right\}. \end{aligned}$$

We have thus sketched the original attempt (cf. [4; Corollary 4.3]) to prove Corollary 1.1 (1). The study was stopped here since we were unable to compare this long expression with $4 \kappa^{\text{DD}}$.

Before moving further, let us make a remark on (9). As proved in [4; (31)], for fixed x and y , equation (9) is equivalent to the following one.

$$\frac{\mu \left((\varphi^-)^{3/2} \mathbb{1}_{(-M, x)} \right)}{\sqrt{\varphi^-(x)}} + \mu \left(\varphi^- \mathbb{1}_{(x, \theta)} \right) = \frac{\mu \left((\varphi^+)^{3/2} \mathbb{1}_{(y, N)} \right)}{\sqrt{\varphi^+(y)}} + \mu \left(\varphi^+ \mathbb{1}_{(\theta, y)} \right). \quad (13)$$

The quantity in (13) is actually the ratio

$$\frac{h^-(\theta)}{f^{x,y}(x)} = \frac{h^+(\theta)}{f^{x,y}(y)}$$

(cf. [4; (34)]) noting that $f^{x,y}$ is a constant on $[x, y]$:

$$f^{x,y}(x) = f^{x,y}(y) = \sqrt{\varphi^+(y)}.$$

Next, note that the left-hand and the right-hand sides of (13) are monotone, with respect to x and y respectively, since each of their derivatives does not change its sign:

$$-\frac{e^{-C(x)}}{2(\varphi^-)^{3/2}(x)} \mu \left((\varphi^-)^{3/2} \mathbb{1}_{(-M, x)} \right) < 0 \quad \text{and} \quad \frac{e^{-C(y)}}{2(\varphi^+)^{3/2}(y)} \mu \left((\varphi^+)^{3/2} \mathbb{1}_{(y, N)} \right) > 0.$$

The unique solution θ to (9), or equivalently (13), should satisfy

$$\begin{aligned} \lim_{x \rightarrow -M} \frac{\mu((\varphi^-)^{3/2} \mathbb{1}_{(-M, x)})}{\sqrt{\varphi^-(x)}} + \mu(\varphi^- \mathbb{1}_{(-M, \theta)}) &\geq \frac{\mu((\varphi^+)^{3/2} \mathbb{1}_{(\theta, N)})}{\sqrt{\varphi^+(\theta)}} \quad \text{and} \\ \lim_{y \rightarrow N} \frac{\mu((\varphi^+)^{3/2} \mathbb{1}_{(y, N)})}{\sqrt{\varphi^+(y)}} + \mu(\varphi^+ \mathbb{1}_{(\theta, N)}) &\geq \frac{\mu((\varphi^-)^{3/2} \mathbb{1}_{(-M, \theta)})}{\sqrt{\varphi^-(\theta)}}. \end{aligned} \quad (14)$$

As just mentioned above (cf. [4; (34)]), we also have

$$\begin{aligned} \max_{z \in [x, \theta]} II^-(f^{x,y})(z) &= \max_{z \in [\theta, y]} II^+(f^{x,y})(z) = \frac{h^-(\theta)}{f^{x,y}(x)} = \frac{h^+(\theta)}{f^{x,y}(y)} \\ &= \frac{1}{\sqrt{\varphi^+(y)}} \mu((\varphi^+)^{3/2} \mathbb{1}_{(y, N)}) + \mu(\varphi^+ \mathbb{1}_{(\theta, y)}). \end{aligned} \quad (15)$$

Hence we have arrived at

$$\begin{aligned} &\left[\sup_{z \in (-M, \theta)} II^-(f^{x,y})(z) \right] \vee \left[\sup_{z \in (\theta, N)} II^+(f^{x,y})(z) \right] \\ &= \left[\sup_{z \in (-M, x)} II^-(f^{x,y})(z) \right] \vee \frac{h^+(\theta)}{\sqrt{\varphi^+(y)}} \vee \left[\sup_{z \in (y, N)} II^+(f^{x,y})(z) \right] \end{aligned} \quad (16)$$

which is also known from [4]. Define

$$I^-(f)(x) = \frac{e^{-C(x)}}{f'(x)} \int_x^\theta \frac{e^C}{a} f, \quad I^+(f)(x) = -\frac{e^{-C(x)}}{f'(x)} \int_\theta^x \frac{e^C}{a} f$$

and

$$\delta_{x, \theta}^- = \sup_{z \in (-M, x)} \varphi_z^- \mu(z, \theta), \quad \delta_{y, \theta}^+ = \sup_{z \in (y, N)} \varphi_z^+ \mu(\theta, z).$$

Then we have first by the mean value theorem (both h^- and $f^{x,y}$ are vanished at $-M$) that

$$\sup_{z \in (-M, x)} II^-(f^{x,y})(z) \leq \sup_{z \in (-M, x)} I^-(f^{x,y})(z)$$

and then by [1; Lemma 1.2] or [2; page 97] that

$$\sup_{z \in (-M, x)} I^-(f^{x,y})(z) \leq 4 \delta_{x, \theta}^-.$$

Here we remark that the supremum in the definition of $\delta_{x, \theta}^-$ is taken over $(-M, x)$ rather than $(-M, \theta) \supset (-M, x)$. Hence the original proof for the last estimate needs a slight modification using the fact that the function $f^{x,y}$ is a constant on $[x, \theta]$. In parallel, since h^+ and $f^{x,y}$ vanish at N , we have

$$\sup_{z \in (y, N)} II^+(f^{x,y})(z) \leq \sup_{z \in (y, N)} I^+(f^{x,y})(z) \leq 4 \delta_{y, \theta}^+.$$

Therefore, we have arrived at

$$\begin{aligned}
 \underline{\kappa}^{\text{DD}} &\leq \inf_{x < y} \left\{ \left[\sup_{(-M, x)} II^-(f^{x,y}) \right] \vee \frac{h^+(\theta)}{\sqrt{\varphi^+(y)}} \vee \left[\sup_{(y, N)} II^+(f^{x,y}) \right] \right\} \\
 &\leq \inf_{x < \theta < y} \left\{ \left[\sup_{(-M, x)} II^-(f^{x,y}) \right] \vee \frac{h^+(\theta)}{\sqrt{\varphi^+(y)}} \vee \left[\sup_{(y, N)} II^+(f^{x,y}) \right] \right\} \\
 &\leq \inf_{x < \theta < y} \left\{ [4\delta_{x,\theta}^-] \vee \frac{h^+(\theta)}{\sqrt{\varphi^+(y)}} \vee [4\delta_{y,\theta}^+] \right\} \\
 &=: \inf_{x < \theta < y} R(x, y, \theta) \\
 &=: \alpha.
 \end{aligned} \tag{17}$$

The restriction $\theta \in (x, y)$ is due to the fact that the eigenfunction of λ^{DD} is unimodal and θ is a mimic of its maximum point. The use of $II^\pm$, $I^\pm$ and $\delta^\pm$ is now standard (cf. [2]–[4], for instance).

We now go to the essential new part of the proof. First, we claim that for each small ε , there exist $\bar{x} \in (-M, \theta)$ and $\bar{y} \in (\theta, N)$ (may depend on ε) such that

$$\varphi_{\bar{x}}^- \mu(\bar{x}, \theta) \geq \frac{R(x_0, y_0, \theta_0)}{4} - \varepsilon, \quad \varphi_{\bar{y}}^+ \mu(\theta, \bar{y}) \geq \frac{R(x_0, y_0, \theta_0)}{4} - \varepsilon \tag{18}$$

for some point (x_0, y_0, θ_0) . In the present continuous case, the conclusion is clear since the infimum $\alpha = R(x^*, y^*, \theta^*)$ is achieved at a point (x^*, y^*, θ^*) with $x^* \leq \theta^* \leq y^*$, at which we have not only $h^-(\theta^*) = h^+(\theta^*)$ but also

$$4\delta_{x^*, \theta^*}^- = 4\delta_{y^*, \theta^*}^+ = \frac{h^+(\theta^*)}{\sqrt{\varphi^+(y^*)}}. \tag{19}$$

To see this, suppose that at the point (x, y, θ) with $x < \theta < y$, we have

$$\frac{h^+(\theta)}{\sqrt{\varphi^+(y)}} > 4[\delta_{x,\theta}^- \vee \delta_{y,\theta}^+]. \tag{20}$$

Without loss of generality, assume that $\delta_{x,\theta}^- \geq \delta_{y,\theta}^+$. We now fix y and let $\tilde{\theta} \in (\theta, y]$. Then $\delta_{y,\theta}^+ \geq \delta_{y,\tilde{\theta}}^+$ by definition. In view of (15), we have

$$\frac{h^+(\theta)}{\sqrt{\varphi^+(y)}} > \frac{h^+(\tilde{\theta})}{\sqrt{\varphi^+(y)}}.$$

Next, to keep $h^-(\tilde{\theta}) = h^+(\tilde{\theta})$, one has a new $\tilde{x} > x$ by using (13) (the left-hand side of (13) is decreasing in x). Correspondingly, we have $\delta_{\tilde{x}, \tilde{\theta}}^- \geq \delta_{x,\theta}^-$.

In particular, for $\tilde{\theta}$ closed enough to θ such that

$$\frac{h^+(\tilde{\theta})}{\sqrt{\varphi^+(y)}} \geq 4\delta_{\tilde{x}, \tilde{\theta}}^-,$$

we obtain

$$\begin{aligned} [4\delta_{x,\theta}^-] \bigvee \frac{h^+(\theta)}{\sqrt{\varphi^+(y)}} \bigvee [4\delta_{y,\theta}^+] &= \frac{h^+(\theta)}{\sqrt{\varphi^+(y)}} \\ &> \frac{h^+(\tilde{\theta})}{\sqrt{\varphi^+(y)}} \\ &= [4\delta_{\tilde{x},\tilde{\theta}}^-] \bigvee \frac{h^+(\tilde{\theta})}{\sqrt{\varphi^+(y)}} \bigvee [4\delta_{y,\tilde{\theta}}^+]. \end{aligned}$$

Thus, once (20) holds, we can find a new point $(\tilde{x}, y, \tilde{\theta})$ such that $R(x, y, \theta) > R(\tilde{x}, y, \tilde{\theta})$. In other words, if the infimum α is attained at (x^*, y^*, θ^*) , we should have

$$\frac{h^+(\theta^*)}{\sqrt{\varphi^+(y^*)}} \leq 4[\delta_{x^*,\theta^*}^- \vee \delta_{y^*,\theta^*}^+]. \quad (21)$$

One may handle with the other two cases and finally arrive at (19). Note that instead of (19), the following weaker condition is still enough for our purpose. If at some point (x, y, θ) ,

$$4\delta_{x,\theta}^- = 4\delta_{y,\theta}^+ =: \alpha' \geq \frac{h^+(\theta)}{\sqrt{\varphi^+(y)}}, \quad (22)$$

then we have not only $\alpha' \geq \alpha$ but also (18) for suitable $\bar{x} \leq x$ and $\bar{y} \geq y$. To check (22), we first mention that the equation $\delta_{x,\theta}^- = \delta_{y,\theta}^+$ is solvable, at least in the case that $M, N < \infty$. Because $\delta_{x,\theta}^-$ starts from zero at $x = -M$ and then increases as $x \uparrow$; $\delta_{y,\theta}^+$ also starts from zero at $y = N$ and then increases as $y \downarrow$. Therefore, there are a lot of (x, y) satisfying the required equation. Next, by (9), we can regard θ as a function of x and y . Then, determine y in terms of x by the equation $\delta_{y,\theta(x,y)}^+ = \delta_{x,\theta(x,y)}^-$. Now there is only one free variable x . We claim that (22) holds for some x (and then for some (x, y, θ)). Otherwise, the inverse inequality of (22) would hold for all x which contradicts with (21).

What we actually need is not the pair $\{x, y\}$ satisfying (22) but the pair $\{\bar{x}, \bar{y}\}$ satisfying (18). From which, the remainder of the proof is very much the same as the one given at the beginning of this section. First, we have

$$(\alpha/4 - \varepsilon) \varphi^-(\bar{x})^{-1} \leq \mu(\bar{x}, \theta), \quad (\alpha/4 - \varepsilon) \varphi^+(\bar{y})^{-1} \leq \mu(\theta, \bar{y}).$$

Summing up these inequalities, we get

$$(\alpha/4 - \varepsilon) [\varphi^-(\bar{x})^{-1} + \varphi^+(\bar{y})^{-1}] \leq \mu(\bar{x}, \bar{y}).$$

Therefore

$$\begin{aligned} \alpha/4 - \varepsilon &\leq [\varphi^-(\bar{x})^{-1} + \varphi^+(\bar{y})^{-1}]^{-1} \mu(\bar{x}, \bar{y}) \\ &\leq \sup_{x < y} [\varphi^-(x)^{-1} + \varphi^+(y)^{-1}]^{-1} \mu(x, y) \\ &= \kappa^{\text{DD}} \quad (\text{by (6)}). \end{aligned}$$

Combining this fact with (17), we obtain

$$\underline{\kappa}^{\text{DD}} \leq \alpha \leq 4\kappa^{\text{DD}} + 4\varepsilon.$$

Letting $\varepsilon \downarrow 0$, we have thus proved that $\underline{\kappa}^{\text{DD}} \leq 4\kappa^{\text{DD}}$ as required. The main part of the proof is done since the first Dirichlet eigenvalue is based on compact sets.

Finally, consider the general case that $M, N \leq \infty$. First, we can rule out the degenerated situation that $\delta_{x,\theta}^- = \delta_{y,\theta}^+ = \infty$. To see this, rewrite κ^{DD} as follows

$$(\kappa^{\text{DD}})^{-1} = \inf_{x < y} [(\varphi^-(x) \mu(x, y))^{-1} + (\mu(x, y) \varphi^+(y))^{-1}].$$

It is clear that $(\kappa^{\text{DD}})^{-1} = 0$ and then $\lambda^{\text{DD}} = 0$ by (5). The corollary becomes trivial. Next, if one of $\delta_{x,\theta}^-$ or $\delta_{y,\theta}^+$ is ∞ , say $\delta_{x,\theta}^- = \infty$ for instance, then

$$(\kappa^{\text{DD}})^{-1} = \left(\sup_y \mu(-M, y) \varphi^+(y) \right)^{-1},$$

i.e.,

$$\kappa^{\text{DD}} = \sup_y \mu(-M, y) \varphi^+(y).$$

This becomes the essentially known one-side Dirichlet problem. In the case that both of $\delta_{x,\theta}^-$ and $\delta_{y,\theta}^+$ are finite, one may adopt an approximating procedure with finite M and N . This was done in the discrete context, refer to [3; Proof of Corollary 7.9 and Proof (c) of Theorem 7.10]. $\square$

To illustrate what was going on in the proof above and the computation/estimation of $\underline{\kappa}^{\text{DD}}$, we consider two examples to conclude this section.

Example 2.2 [4; Example 5.2] Consider the simplest example, i.e. the Laplacian operator on $(0, 1)$. It was proved in [4] that $\lambda^{\text{DD}} = \pi^2$, $(\kappa^{\text{DD}})^{-1} = 16$ and $(\underline{\kappa}^{\text{DD}})^{-1} \approx 9.43693$. The eigenfunction of λ^{DD} is $g(x) = \sin(\pi x)$ for which $g'(1/2) = 0$ and so $\bar{\theta} = 1/2$ is the root of equation $\lambda_{\bar{\theta}}^- = \lambda_{\bar{\theta}}^+$. Because of the symmetry, we have $\theta^* = 1/2$ and $y^* = 1 - x^*$. Since $\mu = \nu = dx$, we have $\varphi_z^- \mu(z, 1/2) = (1/2 - z)z$. Thus,

$$\delta_x^- = \sup_{z \in (0, x)} \varphi_z^- \mu(z, 1/2) = \begin{cases} (1/2 - x)x & \text{if } x \leq 1/4 \\ 1/16 & \text{if } x \in (1/4, 1/2). \end{cases}$$

By (15) and (13), we have

$$\frac{h^-(1/2)}{f^{x, 1-x}(x)} = \frac{1}{8} - \frac{x^2}{10}.$$

Therefore, each

$$x^* \in \left[\frac{20 - \sqrt{205}}{78}, \frac{20 + \sqrt{205}}{78} \right]$$

is a solution to the inequality $4\delta_x^- \geq h^-(1/2)/f^{x,1-x}(x)$. Correspondingly, we have

$$4\delta_{x^*}^- = \begin{cases} 2(1 - 2x^*)x^* & \text{if } x^* \in [(20 - \sqrt{205})/78, 1/4] \\ 1/4 & \text{if } x^* \in [1/4, (20 + \sqrt{205})/78]. \end{cases}$$

Using this, our conclusion that

$$\underline{\kappa}^{\text{DD}} \leq 4\delta_{x^*}^- \leq 4\kappa^{\text{DD}}$$

can be refined as follows:

$$(4\kappa^{\text{DD}})^{-1} = 4 \leq (4\delta_{x^*}^-)^{-1} \leq \frac{35 - \sqrt{41/5}}{4} \approx 8.034 < 9.43693 \approx (\underline{\kappa}^{\text{DD}})^{-1}.$$

It follows that there are many solutions x^* , and so we have a lot of freedom in choosing (θ^*, x^*, y^*) for (22). However, the maximum of $(4\delta_{x^*}^-)^{-1}$ is attained only at the point x^* which is the smaller root of equation: $4\delta_x^- = h^-(1/2)/f^{x,1-x}(x)$.

The next example is unusual since for which the lower bound $(4\kappa^{\text{DD}})^{-1}$ is sharp. Hence, there is no room for the improvement $(\underline{\kappa}^{\text{DD}})^{-1}$. The proof above seems rather dangerous for this example since at each step

$$\begin{aligned} (\lambda^{\text{DD}})^{-1} &\leq \underline{\kappa}^{\text{DD}} \leq \text{Est}(I^\pm(f)) \leq \text{Est}(\delta_{\cdot, \theta}^\pm) \\ &\leq 4(\Phi(\bar{x}, \bar{y}) + \varepsilon) \leq 4 \sup_{x < y} (\Phi(x, y) + \varepsilon) = 4(\kappa^{\text{DD}} + \varepsilon) \end{aligned}$$

for some Φ , where $\text{Est}(H)$ means the estimate using H , one may lose something. Here we have also explained the reason why $\underline{\kappa}^{\text{DD}}$ is often much better than $4\kappa^{\text{DD}}$ as shown in the last example.

Example 2.3 [4; Example 5.3] Consider the operator $L = d^2/dx^2 + bd/dx$ with $b > 0$ on $(0, \infty)$. It was checked in [4] that $\lambda^{\text{DD}} = b^2/4$, $(\kappa^{\text{DD}})^{-1} = b^2$ and so the lower estimate $(4\kappa^{\text{DD}})^{-1}$ is sharp. The eigenfunction of λ^{DD} is $g(x) = xe^{-bx/2}$ for which $g'(2/b) = 0$ and so $\bar{\theta} = 2/b$ solves the equation $\lambda_{\bar{\theta}}^- = \lambda_{\bar{\theta}}^+$. We have $C(x) = bx$, $\mu(dx) = e^{bx}dx$,

$$\varphi^-(s) = \int_0^s e^{-bz} dz = \frac{1}{b}(1 - e^{-bs}) \quad \text{and} \quad \varphi^+(s) = \int_s^\infty e^{-bz} dz = \frac{1}{b}e^{-bs}.$$

We begin our study on the equation $\delta_{x, \theta}^- = \delta_{y, \theta}^+$ rather than Eq.(13) since the former one is simpler. Note that the function

$$\varphi_z^- \mu(z, \theta) = \frac{1}{b^2}(1 - e^{-bz})(e^{b\theta} - e^{bz}), \quad z \in (0, \theta)$$

achieves its maximum $b^{-2}(e^{b\theta/2} - 1)^2$ at $z = \theta/2$ and the function

$$\mu(\theta, z)\varphi_z^+ = \frac{1}{b^2}(1 - e^{b(\theta-z)}), \quad z \geq \theta$$

achieves its maximum $1/b^2$ at ∞ . Hence

$$\delta_{x,\theta}^- = \frac{1}{b^2}(e^{b\theta/2} - 1)^2 \quad \forall x \in [\theta/2, \theta] \quad \text{and} \quad \delta_{y,\theta}^+ = \frac{1}{b^2} \quad \forall y \geq \theta.$$

Solving the equation

$$\frac{1}{b^2}(e^{b\theta/2} - 1)^2 = \frac{1}{b^2},$$

we get $\theta^* = 2b^{-1} \log 2$. To study (14), note that

$$\begin{aligned} & \frac{1}{\sqrt{\varphi^+(y)}} \mu((\varphi^+)^{3/2} \mathbb{1}_{(y,\infty)}) + \mu(\varphi^+ \mathbb{1}_{(\theta,y)}) = \frac{2}{b^2} + \frac{1}{b}(y - \theta) \\ & \frac{1}{\sqrt{\varphi^-(x)}} \mu((\varphi^-)^{3/2} \mathbb{1}_{(0,x)}) + \mu(\varphi^- \mathbb{1}_{(x,\theta)}) \\ & = \frac{1}{b^2} \left\{ 2 - b\theta + e^{b\theta} + bx - \frac{3(bx + \log(1 + \sqrt{1 - e^{-bx}}))}{2\sqrt{1 - e^{-bx}}} \right\}. \end{aligned}$$

Then the second inequality in (14) is trivial and the first one there becomes

$$\frac{1}{b^2}(2 - b\theta + e^{b\theta}) \geq \frac{2}{b^2}.$$

It is now easy to check that $\theta^* = 2b^{-1} \log 2$ does not satisfy this inequality. In other words, there is no required solution (x^*, y^*, θ^*) under the restriction $x^* \in [\theta^*/2, \theta^*]$. Thus, unlike the last example, there is not much freedom in choosing (x^*, y^*, θ^*) for (22). However, this does not finish the story since the solution x^* may belong to $[0, \theta^*/2)$.

We are now looking for a solution x^* in the interval $[0, \theta^*/2)$. When $x \leq \theta/2$, the maximum of the function $\sup_{z \leq x} \varphi_z^- \mu(z, \theta)$ on $[0, x]$ is achieved at x . Hence

$$\delta_{x,\theta}^- = \frac{1}{b^2}(1 - e^{-bx})(e^{b\theta} - e^{bx}) \quad \forall x \in (0, \theta/2] \quad \text{and} \quad \delta_{y,\theta}^+ = \frac{1}{b^2} \quad \forall y \geq \theta.$$

Solving the equation

$$\frac{1}{b^2}(1 - e^{-bx})(e^{b\theta} - e^{bx}) = \frac{1}{b^2},$$

we obtain $\theta^* = x - b^{-1} \log(1 - e^{-bx})$. Besides, solving the equation $4\delta_{y^*, \theta^*}^+ = h^+(\theta^*)/\sqrt{\varphi^+(y^*)}$, we get $y^* = 2/b + \theta^*$. Inserting these into Eq.(13), we obtain

$$\frac{e^{2bx}}{e^{bx} - 1} + \log(1 - e^{-bx}) = 2 + \frac{3}{2\sqrt{1 - e^{-bx}}} \left(bx + 2 \log(\sqrt{1 - e^{-bx}} + 1) \right).$$

From this, we obtain the required solution x^* as shown by Figures 1 and 2 below, noting that the constraint that $x^* \leq \theta^*/2$ is equivalent to $x^* \leq b^{-1} \log 2$. Having x^* at hand, it is clear that the solution θ^* here is very different from $2/b$.

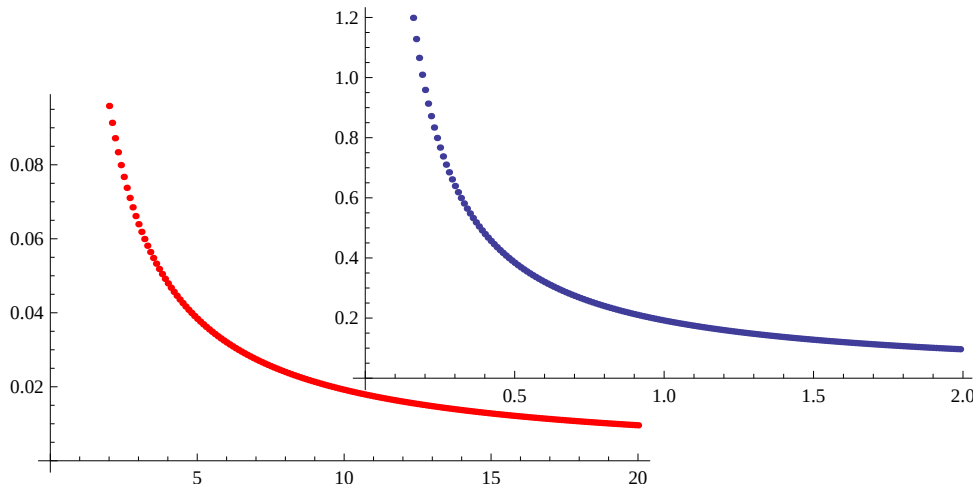


Figure 1–2 Solution of $x^* = x^*(b)$ when b varies on $(0, 2]$ (the curve on right) and on $[2, 20]$ (the curve on left), respectively.

To see that the solutions $(\bar{x}, \bar{y})$ to (18) may not be unique, keeping θ^* to be the same as in the last paragraph but replace y^* with a smaller one $\bar{y} = b^{-1} + \theta^*$, then one can find a point $\bar{x}$ satisfying Eq. (13).

3 Birth–death processes (discrete case)

This section deals with the discrete case which is parallel in principal to the continuous one studied above, but it is quite involved and so is worth to write down some details here.

The state space is

$$E = \{i \in \mathbb{Z} : -M - 1 < i < N + 1\}, \quad M, N \leq \infty.$$

The transition rates $Q = (q_{ij})$ are as follows: $b_i := q_{i,i+1} > 0$, $a_i := q_{i,i-1} > 0$, $q_{ii} = -(a_i + b_i)$, $i \in E$. $q_{ij} = 0$ for other $i \neq j$. Thus, we have $a_{-M} > 0$ if $M < \infty$ and similarly for b_N . The operator of the process becomes

$$\Omega f(i) = b_i(f_{i+1} - f_i) + a_i(f_{i-1} - f_i), \quad i \in E$$

with a convention $f_{-M-1} = 0$ if $M < \infty$ and $f_{N+1} = 0$ if $N < \infty$. Next, define the speed (or invariant, or symmetric) measure μ as follows. Fix a reference point $o \in E$ and set

$$\mu_{o+n} = \frac{a_{o-1}a_{o-2} \cdots a_{o+n+1}}{b_o b_{o-1} \cdots b_{o+n}}, \quad -M - 1 - o < n \leq -2,$$

$$\begin{aligned}\mu_{o-1} &= \frac{1}{b_o b_{o-1}}, & \mu_o &= \frac{1}{a_o b_o}, & \mu_{o+1} &= \frac{1}{a_o a_{o+1}}, \\ \mu_{o+n} &= \frac{b_{o+1} b_{o+2} \cdots b_{o+n-1}}{a_o a_{o+1} \cdots a_{o+n}}, & 2 \leq n &< N+1-o.\end{aligned}$$

A change of the reference point o leads to a constant factor only to the sequence (μ_i) and so does not make any influence to the results below. Corresponding to Ω , the Dirichlet form is

$$\begin{aligned}D(f) &= \sum_{-M-1 < i \leq o} \mu_i a_i (f_i - f_{i-1})^2 + \sum_{o \leq i < N+1} \mu_i b_i (f_{i+1} - f_i)^2, \\ f &\in \mathcal{K}, \quad f_{-M-1} = 0 \text{ if } M < \infty \text{ and } f_{N+1} = 0 \text{ if } N < \infty,\end{aligned}$$

where $\mathcal{K}$ is the set of functions on E with compact supports. Having these preparations at hand, one can define the eigenvalues λ^{DD} and λ^{NN} on $L^2(\mu)$ as in the first section.

To state our main result in this context, we need more notation. Define

$$\mathcal{C}_+ = \{f|_E > 0 : f_{-M-1} = 0 \text{ if } M < \infty \text{ and } f_{N+1} = 0 \text{ if } N < \infty\}.$$

Given $f \in \mathcal{C}_+$, define $h^\pm = h_f^\pm$ as follows.

$$\begin{aligned}h_i^- &= \sum_{k=-M}^i \frac{1}{\mu_k a_k} \sum_{\ell=k}^\theta \mu_\ell f_\ell = \sum_{\ell=-M}^\theta \mu_\ell f_\ell \varphi_{\ell \wedge i}^-, \quad i \leq \theta, \\ h_i^+ &= \sum_{k=i}^N \frac{1}{\mu_k b_k} \sum_{\ell=\theta}^k \mu_\ell f_\ell = \sum_{\ell=\theta}^N \mu_\ell f_\ell \varphi_{\ell \vee i}^+, \quad i \geq \theta,\end{aligned}$$

where

$$\varphi_i^- = \sum_{k=-M}^i \frac{1}{\mu_k a_k}, \quad \varphi_k^+ = \sum_{\ell=k}^N \frac{1}{\mu_\ell b_\ell}$$

and $\theta \in (-M-1, N+1)$ will be specified soon. Applying $h_f^\pm$ to the test function $f = f^{m,n}$ ($m, n \in E$, $m \leq n$):

$$f_i^{m,n} = \begin{cases} \sqrt{\varphi_n^+ \varphi_{i \wedge m}^- / \varphi_m^-} & i \leq n \\ \sqrt{\varphi_i^+} & i \geq n, \end{cases}$$

we obtain a condition for θ which is an analog of (14):

$$\begin{aligned}\lim_{m \rightarrow -M} \frac{1}{\sqrt{\varphi_m^-}} \sum_{k=-M}^{m-1} (\varphi_k^-)^{3/2} \mu_k + \sum_{k=-M}^\theta \varphi_k^- \mu_k &\geq \frac{1}{\sqrt{\varphi_\theta^+}} \sum_{k=\theta}^N (\varphi_k^+)^{3/2} \mu_k \quad \text{and} \\ \lim_{n \rightarrow N} \frac{1}{\sqrt{\varphi_n^+}} \sum_{k=n+1}^N (\varphi_k^+)^{3/2} \mu_k + \sum_{k=\theta}^N \varphi_k^+ \mu_k &\geq \frac{1}{\sqrt{\varphi_\theta^-}} \sum_{k=-M}^\theta (\varphi_k^-)^{3/2} \mu_k.\end{aligned} \tag{23}$$

However, in the discrete situation, one can not expect (9). This leads to a serious change. To explain the main idea, let us return to Theorem 2.1. Because the derivative of the eigenfunction of λ^{DD} has uniquely one zero point, say θ . We can split the interval $(-M, N)$ into two parts having a common boundary θ . Thus, the original process is divided into processes having a common reflecting boundary θ . Theorem 2.1 says that the original λ^{DD} can be represented by using the principal eigenvalues of these sub-processes. This idea is the starting point of [5], as already used in the first proof in Section 2. Since the maximum point θ is unknown in advance, in the original formulation, θ is free and then there is an additional term $\sup_{\theta}$ in the expression of Theorem 2.1. This term was removed in [4], choosing θ as a mimic of the maximum point of the eigenfunction. Unfortunately, such a mimic still does not work in the discrete case, we may lost (9) and more seriously, the eigenfunction may be a simple echelon but not a unimodal (cf. [3; Definition 7.13]). Therefore, more work is required. Again, the idea goes back to [5] except here the choice of θ is based on (23). The first key step of the method is constructing two birth-death processes on the left- and the right-hand sides, separately. As before, the two processes have Dirichlet boundaries at $-M-1$ and $N+1$ but they now have a common Neumann boundary at $\theta \in E$. Let us start from the birth-death process with rates (a_i, b_i) and state space E . Fix a constant $\gamma > 1$.

- (L) The process on the left-hand side has state space $E^{\theta-} = \{i : -M-1 < i \leq \theta\}$, reflects at θ (and so $b_{\theta} = 0$). Its transition structure is the same as the original one except a_{θ} is replaced by $a_{\theta}^{-,\gamma} := \gamma a_{\theta}$. Then for this process, the sequence $(\mu_i : i \in E^{\theta-})$ is the same as the original one except the original μ_{θ} is replaced by μ_{θ}/γ . Hence, the sequence $(\mu_i a_i : i \in E^{\theta-})$ keeps the same as original.
- (R) The process on the right-hand side has state space $E^{\theta+} = \{i : \theta \leq i < N+1\}$, reflects at θ (and then $a_{\theta} = 0$). Its transition structure is again the same as the original one except b_{θ} is replaced by $b_{\theta}^{+,\gamma} := \gamma(\gamma-1)^{-1}b_{\theta}$. Then for this process, the sequence $(\mu_i : i \in E^{\theta+})$ is the same as the original one except the original μ_{θ} is replaced by $(1-\gamma^{-1})\mu_{\theta}$. Hence, the sequence $(\mu_i b_i : i \in E^{\theta+})$ remains the same as original.

Noting that $a_{\theta}^{-,\gamma} \downarrow a_{\theta}$ and $b_{\theta}^{+,\gamma} \uparrow \infty$ as $\gamma \downarrow 1$, $a_{\theta}^{-,\gamma} \uparrow \infty$ and $b_{\theta}^{+,\gamma} \downarrow b_{\theta}$ as $\gamma \uparrow \infty$, the constant γ plays a balance role for the principal eigenvalues of these processes. From here, following the first proof given in Section 2 and using [5] and [3; Theorem 7.10], one can prove the basic estimate $\lambda^{\text{DD}} \geq (4\kappa^{\text{DD}})^{-1}$ in the present context. Certainly, the parallel proof works also in the ergodic case.

We now continue our study on the discrete analog of Corollary 1.1 (1). The

quantity $\varphi^\pm$ needs no change. But $h^\pm$ has to be modified as follows.

$$\begin{aligned} h_i^{-,\gamma} &= \sum_{k=-M}^i \frac{1}{\mu_k a_k} \left[\sum_{k \leq \ell \leq \theta-1} \mu_\ell f_\ell + \frac{1}{\gamma} \mu_\theta f_\theta \right], \quad i \leq \theta, \\ h_i^{+,\gamma} &= \sum_{k=i}^N \frac{1}{\mu_k b_k} \left[\frac{\gamma-1}{\gamma} \mu_\theta f_\theta + \sum_{\theta+1 \leq \ell \leq k} \mu_\ell f_\ell \right], \quad i \geq \theta. \end{aligned}$$

Finally, define $\Pi^{\pm,\gamma}(f) = h^{\pm,\gamma}/f$. It is now more convenient to write the test functions on $E_\theta^\pm$ separately:

$$f_i^{-,m} = \sqrt{\varphi_{i \wedge m}^-}, \quad i \leq \theta, \quad f_i^{+,n} = \sqrt{\varphi_{i \vee n}^+}, \quad i \geq \theta.$$

Comparing with the original $f^{m,n}$, here a factor acting on $f^{-,m}$ is ignored (the reason why one needs the factor in the original case is for $f_\theta^{-,m} = f_\theta^{+,n}$).

Corollary 3.1 We have

$$\lambda^{\text{DD}} \geq (\underline{\kappa}^{\text{DD}})^{-1} \geq (4\kappa^{\text{DD}})^{-1}, \quad (24)$$

where

$$\underline{\kappa}^{\text{DD}} = \inf_{\theta: (23) \text{ holds}} \inf_{m \leq \theta \leq n} \inf_{\gamma > 1} \left\{ \left[\sup_{E \ni i \leq \theta} \Pi_i^{-,\gamma}(f^{-,m}) \right] \vee \left[\sup_{\theta \leq i \in E} \Pi_i^{+,\gamma}(f^{+,n}) \right] \right\}, \quad (25)$$

$$(\kappa^{\text{DD}})^{-1} = \inf_{m,n \in E: m \leq n} \left[\left(\sum_{i=-M}^m \frac{1}{\mu_i a_i} \right)^{-1} + \left(\sum_{i=n}^N \frac{1}{\mu_i b_i} \right)^{-1} \right] \left(\sum_{j=m}^n \mu_j \right)^{-1}. \quad (26)$$

Proof. By using an approximating procedure, one may assume that $M, N < \infty$ (cf. [3; Proof of Corollary 7.9 and Proof (c) of Theorem 7.10]). Fix $\theta \in [m, n]$ and define

$$\begin{aligned} I_i^{-,\gamma}(f) &= \frac{1}{\mu_i a_i (f_i - f_{i-1})} \left[\frac{1}{\gamma} \mu_\theta + \sum_{i \leq \ell \leq \theta-1} \mu_\ell \right], \quad i \leq \theta \\ I_i^{+,\gamma}(f) &= \frac{1}{\mu_i b_i (f_i - f_{i+1})} \left[\frac{\gamma-1}{\gamma} \mu_\theta + \sum_{\theta+1 \leq \ell \leq i} \mu_\ell \right], \quad i \geq \theta \\ \delta_{m,\theta}^{-,\gamma} &= \sup_{i \leq m} \varphi_i^- \left[\frac{1}{\gamma} \mu_\theta + \sum_{i \leq \ell \leq \theta-1} \mu_\ell \right], \quad \delta_{n,\theta}^{+,\gamma} = \sup_{i \geq n} \varphi_i^+ \left[\frac{\gamma-1}{\gamma} \mu_\theta + \sum_{\theta+1 \leq \ell \leq i} \mu_\ell \right]. \end{aligned}$$

We have

$$\frac{h_\theta^{-,\gamma}}{f_m^{-,m}} = \frac{1}{\sqrt{\varphi_m^-}} \sum_{k=-M}^{m-1} (\varphi_k^-)^{3/2} \mu_k + \sum_{k=m}^{\theta-1} \varphi_k^- \mu_k + \frac{1}{\gamma} \varphi_\theta^- \mu_\theta, \quad (27)$$

$$\frac{h_{\theta}^{+, \gamma}}{f_n^{+, n}} = \frac{1}{\sqrt{\varphi_n^+}} \sum_{k=n+1}^N (\varphi_k^+)^{3/2} \mu_k + \sum_{k=\theta+1}^n \varphi_k^+ \mu_k + \frac{\gamma-1}{\gamma} \varphi_{\theta}^+ \mu_{\theta}. \quad (28)$$

For simplicity, let

$$H(m, n, \theta, \gamma) = \max \left\{ \frac{h_{\theta}^{-, \gamma}}{f_m^{-, m}} \mathbb{1}_{\{m < \theta\}}, \frac{h_{\theta}^{+, \gamma}}{f_n^{+, n}} \mathbb{1}_{\{\theta < n\}} \right\}.$$

By [3; Theorem 7.10 (1), Sections 4, 2 and 3], we obtain

$$\begin{aligned} \underline{\kappa}^{\text{DD}} &\leq \inf_{\substack{\gamma > 1 \\ [m, n] \ni \theta \\ \theta: (23) \text{ holds}}} \left\{ \left[\bigvee_{i \leq m} \Pi_i^{-, \gamma}(f^{-, m}) \right] \bigvee H(m, n, \theta, \gamma) \bigvee \left[\bigvee_{i \geq n} \Pi_i^{+, \gamma}(f^{+, n}) \right] \right\} \\ &\leq \inf_{\substack{\gamma > 1 \\ [m, n] \ni \theta \\ \theta: (23) \text{ holds}}} \left\{ \left[\bigvee_{i \leq m} I_i^{-, \gamma}(f^{-, m}) \right] \bigvee H(m, n, \theta, \gamma) \bigvee \left[\bigvee_{i \geq n} I_i^{+, \gamma}(f^{+, n}) \right] \right\} \\ &\leq \inf_{\substack{\gamma > 1 \\ [m, n] \ni \theta \\ \theta: (23) \text{ holds}}} \left\{ \left[4\delta_{m, \theta}^{-, \gamma} \right] \bigvee H(m, n, \theta, \gamma) \bigvee \left[4\delta_{n, \theta}^{+, \gamma} \right] \right\} \\ &=: \inf_{\theta: (23) \text{ holds}} \inf_{[m, n] \ni \theta} \inf_{\gamma > 1} R(m, n, \theta, \gamma) \\ &=: \alpha. \end{aligned} \quad (29)$$

The point we need two terms in the expression of H , rather than one only in the continuous case, is the loss of an analog of (9): here we may not have $h_{\theta}^{-, \gamma} = h_{\theta}^{+, \gamma}$. We now choose a candidate of θ^* (independent of m, n) from (23) and then choose $\{m^*, n^*\}$ with $m^* \leq \theta^* \leq n^*$ (may not be unique) so that (m^*, n^*, θ^*) satisfies the following inequalities

$$\begin{aligned} \left(\delta_{m, \theta}^{-, 1} = \right) \sup_{i \leq m} \varphi_i^- \left[\mu_{\theta} + \sum_{i \leq \ell \leq \theta-1} \mu_{\ell} \right] &\geq \sup_{i \geq n} \varphi_i^+ \sum_{\theta+1 \leq \ell \leq i} \mu_{\ell} \left(= \delta_{n, \theta}^{+, 1} \right) \quad \text{and} \\ \left(\delta_{m, \theta}^{-, \infty} = \right) \sup_{i \leq m} \varphi_i^- \sum_{i \leq \ell \leq \theta-1} \mu_{\ell} &\leq \sup_{i \geq n} \varphi_i^+ \left[\mu_{\theta} + \sum_{\theta+1 \leq \ell \leq i} \mu_{\ell} \right] \left(= \delta_{n, \theta}^{+, \infty} \right). \end{aligned} \quad (30)$$

Roughly speaking, the condition (9) in the continuous case is replaced by a much weaker one (23) and the condition $\delta_{m, \theta}^{-} = \delta_{n, \theta}^{+}$ is replaced by (30). Instead, let γ^* be the unique solution to the equation

$$\delta_{m^*, \theta^*}^{-, \gamma^*} = \delta_{n^*, \theta^*}^{+, \gamma^*}, \quad \gamma \in [1, \infty]$$

for each fixed pair $\{m^*, n^*\}$: $m^* \leq \theta^* \leq n^*$. Here is the balance role played by γ as mentioned before. As an analog of the continuous case, we are interested in those $\{m^*, n^*\}$: $[m^*, n^*] \ni \theta^*$ having the property

$$\delta_{m^*, \theta^*}^{-, \gamma^*} = \delta_{n^*, \theta^*}^{+, \gamma^*} \geq \frac{1}{4} \max \left\{ \frac{h_{\theta^*}^{-, \gamma^*}}{f_{m^*}^{-, m^*}} \mathbb{1}_{\{m^* < \theta^*\}}, \frac{h_{\theta^*}^{+, \gamma^*}}{f_{n^*}^{+, n^*}} \mathbb{1}_{\{\theta^* < n^*\}} \right\}. \quad (31)$$

Unlike the continuous case, here we may have to repeat the procedure in choosing (m^*, n^*, γ^*) since θ^* suggested by (23) may not be unique. Note that the right-hand side of (31) is trivial in the particular case that $m^* = n^* = \theta^*$. Thus, for sufficiently small $\varepsilon > 0$, we may choose $(\bar{m}, \bar{n})$ with $[\bar{m}, \bar{n}] \ni \theta^*$ and $\bar{\gamma} \in (1, \infty)$ such that

$$\begin{aligned} \varphi_{\bar{m}}^- \left[\frac{1}{\bar{\gamma}} \mu_{\theta^*} + \sum_{\ell=\bar{m}}^{\theta^*-1} \mu_\ell \right] &\geq \frac{R(m^*, n^*, \theta^*, \gamma^*)}{4} - \varepsilon \quad \text{and} \\ \varphi_{\bar{n}}^+ \left[\frac{\bar{\gamma}-1}{\bar{\gamma}} \mu_{\theta^*} + \sum_{\ell=\theta^*+1}^{\bar{n}} \mu_\ell \right] &\geq \frac{R(m^*, n^*, \theta^*, \gamma^*)}{4} - \varepsilon. \end{aligned}$$

Therefore, we have

$$\left(\frac{\alpha}{4} - \varepsilon \right) (\varphi_{\bar{m}}^-)^{-1} \leq \frac{1}{\bar{\gamma}} \mu_{\theta^*} + \sum_{\ell=\bar{m}}^{\theta^*-1} \mu_\ell, \quad \left(\frac{\alpha}{4} - \varepsilon \right) (\varphi_{\bar{n}}^+)^{-1} \leq \frac{\bar{\gamma}-1}{\bar{\gamma}} \mu_{\theta^*} + \sum_{\ell=\theta^*+1}^{\bar{n}} \mu_\ell.$$

Summing up these inequalities, it follows that

$$\left(\frac{\alpha}{4} - \varepsilon \right) \left\{ (\varphi_{\bar{m}}^-)^{-1} + (\varphi_{\bar{n}}^+)^{-1} \right\} \leq \sum_{\ell=\bar{m}}^{\theta^*-1} \mu_\ell + \frac{1}{\bar{\gamma}} \mu_{\theta^*} + \frac{\bar{\gamma}-1}{\bar{\gamma}} \mu_{\theta^*} + \sum_{\ell=\theta^*+1}^{\bar{n}} \mu_\ell = \sum_{\ell=\bar{m}}^{\bar{n}} \mu_\ell.$$

The remainder of the proof is the same as in the continuous situation. $\square$

The following example is almost the simplest one but is indeed very helpful to understand Corollary 3.1 and its proof.

Example 3.2 [5; Example 2.3] and [3; Example 7.6 (2)] Let $M = -1$, $N = 2$, $b_1 = 1$, $b_2 = 2$,

$$a_1 = \frac{2 - \varepsilon^2}{1 + \varepsilon}, \quad \varepsilon \in [0, \sqrt{2}) \quad \text{and} \quad a_2 = 1.$$

Then $\lambda^{\text{DD}} = 2 - \varepsilon$. It is known that

$$\kappa^{\text{DD}} = \frac{1}{\lambda_0} - \begin{cases} \varepsilon^2(8 - 4\varepsilon^2 + \varepsilon^3)^{-1} & \text{if } \varepsilon \in [0, (\sqrt{13} - 1)/3] \\ (8 + 2\varepsilon - 3\varepsilon^2)^{-1} & \text{if } \varepsilon \in [(\sqrt{13} - 1)/3, \sqrt{2}). \end{cases}$$

We are now going to compare $\underline{\kappa}^{\text{DD}}$ with $4\kappa^{\text{DD}}$. First, we have $\mu_1 = \mu_2 = 1$, $\mu_1 a_1 = a_1$, $\mu_1 b_1 = b_1$ and $\mu_2 b_2 = b_2$. Next, (23) holds for a small part of ε when $\theta = 1$ but never holds if $\theta = 2$. Hence, we choose $\theta = 1$. Then $m = 1$ and

$$\varphi_1^- = \frac{1}{a_1} = \frac{1 + \varepsilon}{2 - \varepsilon^2}, \quad \varphi_1^+ = \frac{1}{b_1} + \frac{1}{b_2} = \frac{3}{2}, \quad \varphi_2^+ = \frac{1}{b_2} = \frac{1}{2}.$$

Furthermore

$$\delta_{m,\theta}^{-,\gamma} = \frac{1}{\gamma a_1} = \frac{1 + \varepsilon}{\gamma(2 - \varepsilon^2)}.$$

By (30), we have $n = 1$ or 2 .

(1) When $n = 1$, we have

$$\delta_{n,\theta}^{+,\gamma} = \left[\varphi_1^+ \left(1 - \frac{1}{\gamma} \right) \right] \vee \left[\varphi_2^+ \left(2 - \frac{1}{\gamma} \right) \right] = \begin{cases} \frac{3}{2} \left(1 - \frac{1}{\gamma} \right) & \text{if } \gamma \geq 2 \\ 1 - \frac{1}{2\gamma} & \text{if } \gamma \in (1, 2). \end{cases}$$

Clearly, the equation $\delta_{m,\theta}^{-,\gamma} = \delta_{n,\theta}^{+,\gamma}$ has a unique solution

$$\gamma = \begin{cases} \frac{8 + 2\varepsilon - 3\varepsilon^2}{3(2 - \varepsilon^2)} \geq 2 & \text{if } \varepsilon \in \left[\frac{\sqrt{13} - 1}{3}, \sqrt{2} \right) \\ \frac{4 + 2\varepsilon - \varepsilon^2}{2(2 - \varepsilon^2)} \in (1, 2) & \text{if } \varepsilon \in \left(0, \frac{\sqrt{13} - 1}{3} \right). \end{cases}$$

Correspondingly, with $m = n = \theta = 1$, we have

$$\delta_{m,\theta}^{-,\gamma} = \delta_{n,\theta}^{+,\gamma} = \begin{cases} \frac{3(1 + \varepsilon)}{8 + 2\varepsilon - 3\varepsilon^2} & \text{if } \varepsilon \in \left[\frac{\sqrt{13} - 1}{3}, \sqrt{2} \right) \\ \frac{2(1 + \varepsilon)}{4 + 2\varepsilon - \varepsilon^2} & \text{if } \varepsilon \in \left(0, \frac{\sqrt{13} - 1}{3} \right). \end{cases}$$

It is interesting that the last quantity coincides with $4\kappa^{\text{DD}}$. We have thus arrived at (31) since we are in the particular case: $m^* = n^* = \theta^*$.

(2) When $n = 2$, we have

$$\frac{h_{\theta}^{+,\gamma}}{f_n^{+,n}} = \varphi_2^+ + \frac{\gamma - 1}{\gamma} \varphi_1^+ = 2 - \frac{3}{2\gamma}, \quad \delta_{n,\theta}^{+,\gamma} = 1 - \frac{1}{2\gamma}.$$

Clearly,

$$\frac{h_{\theta}^{+,\gamma}}{f_n^{+,n}} \leq 4\delta_{n,\theta}^{+,\gamma} \quad \text{iff } \gamma \geq 1/4.$$

As we have seen above, the solution to the equation $\delta_{m,\theta}^{-,\gamma} = \delta_{n,\theta}^{+,\gamma}$ is

$$\gamma = \frac{4 + 2\varepsilon - \varepsilon^2}{2(2 - \varepsilon^2)} > 1 \quad \text{on } (0, \sqrt{2}).$$

Then

$$\delta_{m,\theta}^{-,\gamma} = \delta_{n,\theta}^{+,\gamma} = \frac{2(1 + \varepsilon)}{4 + 2\varepsilon - \varepsilon^2} > \frac{h_{\theta}^{+,\gamma}}{4f_n^{+,n}} \quad (m = \theta = 1, n = 2).$$

Hence (31) holds. Combining this case with the last one (i.e., $n = 1$), it follows that $\underline{\kappa}^{\text{DD}} < 4\kappa^{\text{DD}}$ for $\varepsilon \in ((\sqrt{13} - 1)/3, \sqrt{2})$.

Example 3.3 [3; Examples 7.7 (5)] Let $E = \{1, 2, \dots\}$, $a_i = 1/i$ and $b_i = 1$ for all $i \geq 1$. Then $\lambda^{\text{DD}} = (3 - \sqrt{5})/2 \approx 0.38$ and $(\lambda^{\text{DD}})^{-1} \approx 2.618$. We have $\mu_i = i!$, $\mu_i a_i = (i-1)!$ and $\mu_i b_i = i!$ for all $i \geq 1$. Furthermore, we have

$$(\kappa^{\text{DD}})^{-1} = \left(\left[\sum_{k=1}^1 \frac{1}{(k-1)!} \right]^{-1} + \left[\sum_{k=4}^{\infty} \frac{1}{k!} \right]^{-1} \right) \left[\sum_{\ell=1}^4 \ell! \right]^{-1} = \frac{1}{33} \left[1 + \frac{3}{3e-8} \right] \approx 0.6174.$$

And so $\kappa^{\text{DD}} \approx 1.62$. With

$$\varphi_i^- = \sum_{k=1}^i \frac{1}{\mu_k a_k} = \sum_{k=1}^i \frac{1}{(k-1)!} \quad \text{and} \quad \varphi_k^+ = \sum_{\ell=k}^{\infty} \frac{1}{\mu_{\ell} b_{\ell}} = \sum_{\ell=k}^{\infty} \frac{1}{\ell!},$$

we have

$$\begin{aligned} \delta_{m,\theta}^{-,\gamma} &= \sup_{i \leq m} \left[\frac{1}{\gamma} \theta! + \sum_{i \leq \ell \leq \theta-1} \ell! \right] \varphi_i^-, & \delta_{n,\theta}^{+,\gamma} &= \sup_{i \geq n} \left[\frac{\gamma-1}{\gamma} \theta! + \sum_{\theta+1 \leq \ell \leq i} \ell! \right] \varphi_i^+. \\ \frac{h_{\theta}^{-,\gamma}}{f_m^{-,m}} &= \frac{1}{\sqrt{\varphi_m^{-,\gamma}}} \sum_{k=1}^{m-1} (\varphi_k^-)^{3/2} k! + \sum_{k=m}^{\theta-1} \varphi_k^- k! + \frac{1}{\gamma} \varphi_{\theta}^- \theta!, \\ \frac{h_{\theta}^{+,\gamma}}{f_n^{+,n}} &= \frac{1}{\sqrt{\varphi_n^{+,\gamma}}} \sum_{k=n+1}^{\infty} (\varphi_k^+)^{3/2} k! + \sum_{k=\theta+1}^n \varphi_k^+ k! + \frac{\gamma-1}{\gamma} \varphi_{\theta}^+ \theta!. \end{aligned}$$

For convenience, let (L) , (R) , (M_-) and (M_+) denote the last four quantities.

The candidates given by (23) are $\theta = 2, 3$. The case of $\theta = 3$ is ruled out by (30) and so we fix $\theta = 2$. Then with $m = 1, 2$, (m, n) satisfies (30) for every n : $2 \leq n \leq 17$. For the simplest choice $m = n = \theta$, $\delta_{m,\theta}^{-,\gamma}$ is attained at $i = 1$ once $\gamma \geq 2$, $\delta_{n,\theta}^{+,\gamma}$ is attained at $i = 4$ whenever $\gamma \geq 5/4$, and then the solution to the equation $\delta_{m,\theta}^{-,\gamma} = \delta_{n,\theta}^{+,\gamma}$ is $\gamma \approx 3.2273$. Therefore,

$$\delta_{m,\theta}^{-,\gamma} = \delta_{n,\theta}^{+,\gamma} \approx 1.62.$$

A better choice is $(m, n) = (1, 5)$. Then $\gamma \approx 3.944$ and

$$4 \times (R) = 4 \times (L) \approx 6.042, \quad (M_-) \approx 2.014, \quad (M_+) \approx 5.54.$$

This certainly implies (31).

Example 3.4 [3; Examples 7.7 (8)] Let $E = \{1, 2, \dots\}$, $a_1 = 1$, $a_i = (i-1)^2$ for $i \geq 2$ and $b_i = i^2$ for $i \geq 1$. Then $\lambda^{\text{DD}} = 1/4 = (4\kappa^{\text{DD}})^{-1}$. Once again, this example is dangerous. Clearly, $\mu_i = 1$, $\mu_i a_i = a_i$ and $\mu_i b_i = b_i$ for all $i \geq 1$. We have

$$\varphi_i^- = 1 + \sum_{k=1}^{i-1} \frac{1}{k^2}, \quad \varphi_i^+ = \sum_{\ell=i}^{\infty} \frac{1}{\ell^2}$$

and

$$\begin{aligned}\delta_{m,\theta}^{-,\gamma} &= \sup_{i \leq m} \left[\frac{1}{\gamma} + \theta - i \right] \varphi_i^-, & \delta_{n,\theta}^{+,\gamma} &= \sup_{i \geq n} \left[\frac{\gamma-1}{\gamma} + i - \theta \right] \varphi_i^+, \\ \frac{h_{\theta}^{-,\gamma}}{f_m^{-,m}} &= \frac{1}{\sqrt{\varphi_m^-}} \sum_{k=1}^{m-1} (\varphi_k^-)^{3/2} + \sum_{k=m}^{\theta-1} \varphi_k^- + \frac{1}{\gamma} \varphi_{\theta}^-, \\ \frac{h_{\theta}^{+,\gamma}}{f_n^{+,n}} &= \frac{1}{\sqrt{\varphi_n^+}} \sum_{k=n+1}^{\infty} (\varphi_k^+)^{3/2} + \sum_{k=\theta+1}^n \varphi_k^+ + \frac{\gamma-1}{\gamma} \varphi_{\theta}^+.\end{aligned}$$

As in the last example, we use (L) , (R) , (M_-) and (M_+) to denote the last four quantities. The only candidate by (23) is $\theta = 2$ which is fixed now. Then (30) holds for all $m = 1, 2$ and $n \geq 2$. The key for this example is that $4 \times (R) = 4$, independent of θ and γ . With $m = 2$, the maximum of (L) is achieved at $i = 1$, it tends to 1 as $\gamma \rightarrow \infty$. Since $m = \theta$, the term (M_-) is ignored. Besides, we have $(M_+) < 4$ for all n : $2 \leq n \leq 58$. Therefore, (31) holds.

To conclude the paper, we make a remark on the generalization of the results given here.

Remark 3.5 By using a known technique (cf. [2; Section 6.7]), the variational formula and its corollaries for the lower estimate of λ^{DD} can be extended to a more general setup (Poincaré-type inequalities). The upper estimate is easier and was given in [3; the remark above Corollary 8.3].

References

- [1] Chen, M.F. (2000). Explicit bounds of the first eigenvalue, *Sci. China (A)* 43(10), 1051–1059.
- [2] Chen, M.F. (2005). *Eigenvalues, Inequalities, and Ergodic Theory*. Springer, London.
- [3] Chen, M.F. (2010). Speed of stability for birth–death processes. *Front. Math. China* 5(3), 379–515.
- [4] Chen, M.F. (2012). Basic estimates of stability rate for one-dimensional diffusions. Chapter 6 in “Probability Approximations and Beyond”. Eds: A. Barbour, H.P. Chan and D. Siegmund. *Lect. Notes in Stat.* 205. Springer
- [5] Chen, M.F., Zhang, Y.H. and Zhao, X.L. (2003). Dual variational formulas for the first Dirichlet eigenvalue on half-line, *Sci. China* 46(6), 847–861.
- [6] Fukushima, M. and Uemura, T. (2003). Capacitary bounds of measures and ultracontractivity of time changed processes, *J. Math. Pure et Appliquees* 82(5), 553–572.
- [7] Maz’ya, V. (2011). *Sobolev Spaces with applications to elliptic partial differential equations* (2nd Ed.), Springer.

School of Mathematical Sciences, Beijing Normal University,
Laboratory of Mathematics and Complex Systems (Beijing Normal University),
Ministry of Education,
Beijing 100875, The People's Republic of China.
E-mail: mfchen@bnu.edu.cn
Home page: http://math.bnu.edu.cn/~chenmf/main_eng.htm

Bilateral Hardy-type Inequalities

Mu Fa CHEN

School of Mathematical Sciences, Beijing Normal University, Beijing 100875, P. R. China

and

Laboratory of Mathematics and Complex Systems, Beijing Normal University,

Ministry of Education, Beijing 100875, P. R. China

E-mail: mfchen@bnu.edu.cn

Abstract This paper studies the Hardy-type inequalities on the intervals (may be infinite) with two weights, either vanishing at two endpoints of the interval or having mean zero. For the first type of inequalities, in terms of new isoperimetric constants, the factor of upper and lower bounds becomes smaller than the known ones. The second type of the inequalities is motivated from probability theory and is new in the analytic context. The proofs are now rather elementary. Similar improvements are made for Nash inequality, Sobolev-type inequality, and the logarithmic Sobolev inequality on the intervals.

Keywords Hardy-type inequality, vanishing at two endpoints, mean zero, splitting technique, normed linear space, Nash inequality, logarithmic Sobolev inequality

MR(2000) Subject Classification 26D10, 60J60, 34L15

A large number of results on Hardy-type inequalities have been already collected and explored in the books [1–5]. This paper makes two additions. The first one is for the functions vanishing at two endpoints of the interval. This type of inequalities was included in [5]. The contribution here is some improvement, not only on the isoperimetric constant but also on the factor of the upper and lower bounds. The second addition is for the case where the functions have mean zero, which is motivated from a probabilistic consideration and is not included in the books cited above. These two cases are studied in the next two sections separately. The main result in each case is stated as a theorem (Theorems 1.6 and 2.6). Their extensions to more general setup are presented as Theorems 1.11 and 2.9. As applications of the results or ideas developed in the first two sections, in the third section, we study the Nash inequality, the Sobolev-type inequality, and the logarithmic Sobolev inequality. The paper can be regarded as an extension of the L^2 -case studied in [6, 7].

1 The Case of Vanishing Boundaries

Consider an interval $[-M, N]$ with $M, N \leq \infty$. Certainly, here $[-M, N]$ means $[-M, N)$ if $N = \infty$. This costs no confusion. In this section, we study the case that the functions vanish

at two endpoints of the interval. That is the Hardy-type inequality:

$$\left(\int_{-M}^N |f|^q d\mu \right)^{1/q} \leq A \left(\int_{-M}^N |f'|^p d\nu \right)^{1/p}, \quad f(-M) = 0 \text{ and } f(N) = 0, \quad (1.1)$$

where $f(N) = 0$ for instance means that $\lim_{x \rightarrow \infty} f(x) = 0$ if $N = \infty$. Throughout the paper, all of the functions involved in the Hardy-type inequalities are assumed to be absolutely continuous without mentioned time by time in what follows. In the special case that $p = q = 2$, the results in this and the next sections are proved in [6, 7] using much advanced methods. The present study is motivated from seeking for more direct proofs for the results. At the moment, it is unclear how the capacity technique used in [6, 7] can be applied in the present general setup. It may be helpful to the reader by studying the problem step by step to show how to find out the main result. The study consists of five steps. At each step, we have either a proposition or a lemma. If one is in hurry, who may jump from here to the main results, Theorems 1.11 and 2.9.

Our first step is using the splitting technique (which we have used several times before, cf. [8], [9, Theorems 3.3 and 3.4], and [10]). To do so, fix $\theta \in (-M, N)$ and denote by A_θ^+ and A_θ^- , respectively, the optimal constant in the following inequalities.

$$\begin{aligned} \left(\int_\theta^N |f|^q d\mu \right)^{1/q} &\leq A_\theta^+ \left(\int_\theta^N |f'|^p d\nu \right)^{1/p}, \quad f(N) = 0, \\ \left(\int_{-M}^\theta |f|^q d\mu \right)^{1/q} &\leq A_\theta^- \left(\int_{-M}^\theta |f'|^p d\nu \right)^{1/p}, \quad f(-M) = 0. \end{aligned}$$

Clearly, these inequalities are different from (1.1) since only one-side boundary condition is endowed. Here and in what follows the superscript “ $-$ ” means on the left-hand side of θ and “ $+$ ” means on the right-hand side of θ .

The next result shows that we can describe the optimal constant A in (1.1) in terms of $A_\theta^\pm$ which are the optimal constants on half-spaces with different boundary conditions.

Proposition 1.1 *For $1 \leq p \leq q < \infty$, we have*

$$2^{1/q-1/p} \sup_{\theta \in [-M, N]} (A_\theta^- \wedge A_\theta^+) \leq A \leq \inf_{\theta \in [-M, N]} (A_\theta^- \vee A_\theta^+),$$

where $A_N^+ = 0$ and $A_{-M}^- = 0$ by convention, $\alpha \wedge \beta = \min\{\alpha, \beta\}$, and $\alpha \vee \beta = \max\{\alpha, \beta\}$.

Proof (a) For each fixed $\theta \in [-M, N]$ and f with $f(-M) = 0$ and $f(N) = 0$, by the inequalities on the half-spaces, we have

$$\begin{aligned} \int_{-M}^N |f'|^p d\nu &= \int_{-M}^\theta |f'|^p d\nu + \int_\theta^N |f'|^p d\nu \\ &\geq (A_\theta^-)^{-p} \left(\int_{-M}^\theta |f|^q d\mu \right)^{p/q} + (A_\theta^+)^{-p} \left(\int_\theta^N |f|^q d\mu \right)^{p/q} \\ &\geq [(A_\theta^-)^{-p} \wedge (A_\theta^+)^{-p}] \left[\left(\int_{-M}^\theta |f|^q d\mu \right)^{p/q} + \left(\int_\theta^N |f|^q d\mu \right)^{p/q} \right] \\ &\geq (2^{(p/q-1)\vee 0})^{-1} [(A_\theta^-)^{-p} \wedge (A_\theta^+)^{-p}] \left(\int_{-M}^N |f|^q d\mu \right)^{p/q} \quad (\text{by } c_r\text{-inequality}). \end{aligned}$$

Since f is arbitrary, we have

$$A^p \leq 2^{(p/q-1) \vee 0} [(A_\theta^-)^p \vee (A_\theta^+)^p].$$

Now, since θ is arbitrary, we obtain

$$A \leq 2^{(1/q-1/p) \vee 0} \inf_{\theta \in [-M, N]} (A_\theta^- \vee A_\theta^+).$$

This conclusion holds for general $p, q \in [1, \infty)$.

(b) Again, fix θ . Suppose for a moment that we can construct two absolutely continuous functions f_- and f_+ having the following properties: $f_-(-M) = 0$, $f'_-(\theta) = 0$, $f_-(\theta) > 0$,

$$\int_{-M}^{\theta} |f_-|^q d\mu = 1 \quad \text{and} \quad \left(\int_{-M}^{\theta} |f'_-|^p d\nu \right)^{1/p} < (A_\theta^-)^{-1} + \varepsilon;$$

$$f_+(N) = 0, \quad f'_+(\theta) = 0, \quad f_+(\theta) > 0,$$

$$\int_{\theta}^N |f_+|^q d\mu = 1 \quad \text{and} \quad \left(\int_{\theta}^N |f'_+|^p d\nu \right)^{1/p} < (A_\theta^+)^{-1} + \varepsilon.$$

Set $f = cf_- \mathbb{1}_{[-M, \theta]} + f_+ \mathbb{1}_{(\theta, N]}$, where $c = f_+(\theta)/f_-(\theta)$. Then

$$\begin{aligned} 1 + |c|^q &= \int_{-M}^{\theta} |cf_-|^q d\mu + \int_{\theta}^N |f_+|^q d\mu = \int_{-M}^N |f|^q d\mu, \\ \int_{-M}^N |f'|^p d\nu &= |c|^p \int_{-M}^{\theta} |f'_-|^p d\nu + \int_{\theta}^N |f'_+|^p d\nu \\ &\leq |c|^p ((A_\theta^-)^{-1} + \varepsilon)^p + ((A_\theta^+)^{-1} + \varepsilon)^p \\ &\leq ((A_\theta^-)^{-1} \vee (A_\theta^+)^{-1} + \varepsilon)^p (1 + |c|^p). \end{aligned}$$

Hence

$$\begin{aligned} \left(\int_{-M}^N |f'|^p d\nu \right)^{1/p} &\leq ((A_\theta^-)^{-1} \vee (A_\theta^+)^{-1} + \varepsilon) (1 + |c|^p)^{1/p} \\ &\leq 2^{1/p-1/q} ((A_\theta^-)^{-1} \vee (A_\theta^+)^{-1} + \varepsilon) (1 + |c|^q)^{1/q} \\ &\quad \text{(by Jensen's inequality requiring } q \geq p) \\ &= 2^{1/p-1/q} ((A_\theta^-)^{-1} \vee (A_\theta^+)^{-1} + \varepsilon) \left(\int_{-M}^N |f|^q d\mu \right)^{1/q}. \end{aligned}$$

Thus, whenever $q \geq p$, we have

$$A \geq 2^{1/q-1/p} (A_\theta^- \wedge A_\theta^+).$$

Therefore, we obtain

$$A \geq 2^{1/q-1/p} \sup_{\theta \in [-M, N]} (A_\theta^- \wedge A_\theta^+), \quad 1 \leq p \leq q < \infty.$$

Combining this with (a), we arrive at the conclusion of the proposition.

(c) To complete the proof, it remains to construct the functions f_- and f_+ used in (b). For this, we need consider f_- only by symmetry. The problem is the condition at θ : $f'_-(\theta) = 0$ and $f_-(\theta) > 0$. The proof given below is modified from [11, Proof (ii) of Theorem 1.1]. If necessary, by modifying f_- properly on a sufficiently small neighborhood of θ , we can assume

that $f'_-(\theta) = 0$. The main point here is to modify f_- so that we also have $f_-(\theta) \neq 0$. Otherwise, suppose that $f_-(\theta) = 0$. Since f_- is absolutely continuous, $f_-(-M) = 0$ and $f_-(\theta) = 0$, there exists $x_1 \in (-M, \theta)$ such that $|f_-(x_1)| = \sup_{x \in (-M, \theta)} |f_-(x)|$. Then $f_-(x_1) \neq 0$ (otherwise, $f \equiv 0$ which contradicts with the norm 1 assumption). Let $\tilde{f}_- = f_- \mathbb{1}_{[-M, x_1]} + f_-(x_1) \mathbb{1}_{[x_1, \theta]}$. Then $\tilde{f}_-$ is absolutely continuous,

$$c^q := \int_{-M}^{\theta} |\tilde{f}_-|^q d\mu \geq \int_{-M}^{\theta} |f_-|^q d\mu = 1,$$

$$\left(\int_{-M}^{\theta} |\tilde{f}_-|^p d\nu \right)^{1/p} \leq \left(\int_{-M}^{\theta} |f'_-|^p d\nu \right)^{1/p} < (A_{\theta}^-)^{-1} + \varepsilon.$$

Set $\bar{f}_- = \frac{1}{c} \tilde{f}_-$. Now it follows that

$$\bar{f}_-(-M) = 0, \quad \bar{f}'_-(\theta) = 0, \quad \bar{f}_-(\theta) \neq 0, \quad \int_{-M}^{\theta} |\bar{f}_-|^q d\mu = 1,$$

and

$$\left(\int_{-M}^{\theta} |\bar{f}_-|^p d\nu \right)^{1/p} = \frac{1}{c} \left(\int_{-M}^{\theta} |\tilde{f}_-|^p d\nu \right)^{1/p} < (A_{\theta}^-)^{-1} + \varepsilon.$$

Hence, we can replace f_- by $\bar{f}_-$ when $f_-(\theta) = 0$. $\square$

Having Proposition 1.1 at hand, it is ready to write down some estimates of the optimal constant A in (1.1), as we did in [9, 10], in terms of $B_{\theta}^{\pm}$ given below (cf. [5, Theorem 6.2] and [4, §1.3, Theorem 3] in which the factor $k_{q,p}$ may be different):

$$B_{\theta}^{\pm} \leq A_{\theta}^{\pm} \leq k_{q,p} B_{\theta}^{\pm}, \quad 1 < p \leq q < \infty, \quad (1.2)$$

$$B_{\theta}^+ = \sup_{r \in (\theta, N)} \mu[\theta, r]^{1/q} \left[\int_r^N \left(\frac{d\nu^*}{dx} \right)^{-1/(p-1)} dx \right]^{(p-1)/p}, \quad (1.3)$$

$$B_{\theta}^- = \sup_{r \in (-M, \theta)} \mu[r, \theta]^{1/q} \left[\int_{-M}^r \left(\frac{d\nu^*}{dx} \right)^{-1/(p-1)} dx \right]^{(p-1)/p}, \quad (1.4)$$

where ν^* is the absolutely continuous part of ν and $k_{q,p}$ is a universal constant will be used often in this paper:

$$k_{q,p} = \left(1 + \frac{q}{p'} \right)^{1/q} \left(1 + \frac{p'}{q} \right)^{1/p'}, \quad (1.5)$$

where p' is the conjugate number of p and similarly for q' . In particular, $k_{p,p} = p^{1/p} p'^{1/p'}$. On the half-line when $q > p$, the constant is improved as follows:

$$k_{q,p} = \left[\frac{\Gamma(\frac{pq}{q-p})}{\Gamma(\frac{q}{q-p}) \Gamma(\frac{p(q-1)}{q-p})} \right]^{1/p-1/q} = \left[\frac{q-p}{pq B(\frac{q}{q-p}, \frac{p(q-1)}{q-p})} \right]^{1/p-1/q}, \quad q > p,$$

where $\Gamma(x)$ and $B(x, y) = \Gamma(x)\Gamma(y)/\Gamma(x+y)$ are Gamma and Beta functions, respectively (cf. [12, Theorem 8], [13, Theorem 2], and also [2, pp. 45–47] for historical remarks). According to Lebesgue's decomposition theorem, each measure ν can be decomposed into three parts:

$$\nu = \nu^* + \nu_{\text{sing}} + \nu_{\text{pp}},$$

where ν_{sing} is the singular continuous part and ν_{pp} is the pure point part (a discrete measure).

We are now going to present some more explicit estimates. To do so, we need the following simple result.

Lemma 1.2 *For a given Borel measure μ and positive functions φ and ψ on $[-M, N]$, we have*

$$\sup_{(x,y): x \leq y} \frac{\mu[x, y]}{\varphi(x) + \psi(y)} \geq \sup_{\theta} \left\{ \left[\sup_{x \leq \theta} \frac{\mu[x, \theta]}{\varphi(x)} \right] \wedge \left[\sup_{y \geq \theta} \frac{\mu(\theta, y]}{\psi(y)} \right] \right\}.$$

Proof For fixed $x \leq y$ and $(x, y) \ni \theta$, we have by proportional property that

$$\frac{\mu[x, y]}{\varphi(x) + \psi(y)} = \frac{\mu[x, \theta] + \mu(\theta, y]}{\varphi(x) + \psi(y)} \geq \frac{\mu[x, \theta]}{\varphi(x)} \wedge \frac{\mu(\theta, y]}{\psi(y)}$$

and furthermore

$$\frac{\mu[x, y]}{\varphi(x) + \psi(y)} \geq \sup_{\theta \in [x, y]} \left\{ \frac{\mu[x, \theta]}{\varphi(x)} \wedge \frac{\mu(\theta, y]}{\psi(y)} \right\}.$$

Thus,

$$\begin{aligned} \sup_{x \leq y} \frac{\mu[x, y]}{\varphi(x) + \psi(y)} &\geq \sup_{x \leq y} \sup_{\theta \in [x, y]} \left\{ \frac{\mu[x, \theta]}{\varphi(x)} \wedge \frac{\mu(\theta, y]}{\psi(y)} \right\} \\ &= \sup_{\theta} \sup_{[x, y] \ni \theta} \{ \dots \} \\ &= \sup_{\theta} \left\{ \left[\sup_{x \leq \theta} \frac{\mu[x, \theta]}{\varphi(x)} \right] \wedge \left[\sup_{y \geq \theta} \frac{\mu(\theta, y]}{\psi(y)} \right] \right\} \end{aligned}$$

as required. $\square$

It is remarkable that we do not have an expected dual result of the above lemma. At beginning, we do have the dual

$$\frac{\mu[x, y]}{\varphi(x) + \psi(y)} \leq \frac{\mu[x, \theta] + \mu(\theta, y]}{\varphi(x) + \psi(y)} \leq \frac{\mu[x, \theta]}{\varphi(x)} \vee \frac{\mu(\theta, y]}{\psi(y)}.$$

Hence,

$$\sup_{x \leq \theta \leq y} \frac{\mu[x, y]}{\varphi(x) + \psi(y)} \leq \left[\sup_{x \leq \theta} \frac{\mu[x, \theta]}{\varphi(x)} \right] \vee \left[\sup_{y \geq \theta} \frac{\mu(\theta, y]}{\psi(y)} \right]$$

and furthermore

$$\inf_{\theta} \sup_{x \leq \theta \leq y} \frac{\mu[x, y]}{\varphi(x) + \psi(y)} \leq \inf_{\theta} \left[\sup_{x \leq \theta} \frac{\mu[x, \theta]}{\varphi(x)} \right] \vee \left[\sup_{y \geq \theta} \frac{\mu(\theta, y]}{\psi(y)} \right].$$

Clearly, this is somehow a dual of Lemma 1.2 but it is still a distance to what we expect:

$$\sup_{x \leq y} \frac{\mu[x, y]}{\varphi(x) + \psi(y)} \leq \inf_{\theta} \left[\sup_{x \leq \theta} \frac{\mu[x, \theta]}{\varphi(x)} \right] \vee \left[\sup_{y \geq \theta} \frac{\mu(\theta, y]}{\psi(y)} \right].$$

Alternatively, let $\bar{\theta}$ satisfy

$$\sup_{x \leq \bar{\theta}} \frac{\mu[x, \bar{\theta}]}{\varphi(x)} = \sup_{y \geq \bar{\theta}} \frac{\mu(\bar{\theta}, y]}{\psi(y)}.$$

Then we have

$$\sup_{x \leq \bar{\theta} \leq y} \frac{\mu[x, y]}{\varphi(x) + \psi(y)} \leq \sup_{x \leq \bar{\theta}} \frac{\mu[x, \bar{\theta}]}{\varphi(x)} = \left[\sup_{x \leq \bar{\theta}} \frac{\mu[x, \bar{\theta}]}{\varphi(x)} \right] \vee \left[\sup_{y \geq \bar{\theta}} \frac{\mu(\bar{\theta}, y]}{\psi(y)} \right]. \quad (1.6)$$

Very often, the right-hand side coincides with

$$\inf_{\theta} \left\{ \left[\sup_{x \leq \theta} \frac{\mu[x, \theta]}{\varphi(x)} \right] \vee \left[\sup_{y \geq \theta} \frac{\mu(\theta, y]}{\psi(y)} \right] \right\},$$

but one can not remove $\bar{\theta}$ from the left-hand side and keep the inequality.

Throughout this paper, we mainly restrict ourselves to the case that $1 < p \leq q < \infty$. The limit case that either $p = 1$ or $q = \infty$ are easier and so are omitted here. For simplicity, throughout this paper, we set

$$h(x) = \left(\frac{d\nu^*}{dx} \right)^{-1/(p-1)}, \quad \hat{\nu}(dx) = h(x)dx.$$

Clearly, h and $\hat{\nu}$ depend on $p > 1$. The measure $\hat{\nu}$ comes, but different from ν . In what follows, almost every estimate is expressed by using the pair $(\mu, \hat{\nu})$ but not (μ, ν) . Besides, we may assume that

$$\hat{\nu}(-M, N) := \int_{-M}^N h = \int_{-M}^N \left(\frac{d\nu^*}{dx} \right)^{-1/(p-1)} dx < \infty. \quad (1.7)$$

This technical assumption can often be avoided by replacing $d\nu^*/dx$ with $d\nu^*/dx + \varepsilon \exp[(p-1)x^2]$ and then passing to the limit as $\varepsilon \downarrow 0$. Alternatively, one may start at $M, N < \infty$, replace $d\nu^*/dx$ with $d\nu^*/dx + \varepsilon$. Then pass to the limit as $\varepsilon \downarrow 0$, and then as $M, N \rightarrow \infty$ if necessary. In parallel, without loss of generality, we can also assume that μ is positive on each subinterval.

Next, define a constant B^* by

$$(B^{*q})^{-1} = \inf_{-M \leq x \leq y \leq N} [\hat{\nu}[-M, x]^{-\frac{q(p-1)}{p}} + \hat{\nu}[y, N]^{-\frac{q(p-1)}{p}}] \mu[x, y]^{-1}. \quad (1.8)$$

Let us now discuss the boundary condition in the definition of B^* above (or B_* below), when $M = \infty$, here $x = -M$ means that $x \rightarrow -\infty$:

$$\begin{aligned} & \overline{\lim}_{x \rightarrow -\infty} \mu[x, y]^{1/q} [\hat{\nu}[-M, x]^{-\frac{q(p-1)}{p}} + \hat{\nu}[y, N]^{-\frac{q(p-1)}{p}}]^{-1/q} \\ &= \overline{\lim}_{x \rightarrow -\infty} \mu[x, y]^{1/q} \hat{\nu}[-M, x]^{(p-1)/p}, \end{aligned}$$

which is the type $\infty \cdot 0$ of limit provided $\mu[-\infty, y] = \infty$. Otherwise, the limit is zero and so the boundary $-M$ can be ignored in computing B^* . When $M = \infty = N$, we need to compute the iterated limit only. To which, the main reason is that the optimal constant A is increasing as either $N \uparrow$ or $-M \downarrow$. Hence, the general case can be regarded as the limit of finite M and N . In other words, we do not need to consider the other types of double limits as $M, N \rightarrow \infty$.

Here is our upper estimate.

Lemma 1.3 *Let $\mu_{pp} = 0$. Then for $1 < p \leq q < \infty$, we have $A \leq k_{q,p} B^*$, where B^* is defined by (1.8).*

Proof As mentioned above, without loss of generality, we can assume (1.7). Rewrite B^{*q} as

$$B^{*q} = \sup_{x \leq y} \frac{\mu[x, \theta] + \mu(\theta, y)}{\hat{\nu}[-M, x]^{-q(p-1)/p} + \hat{\nu}[y, N]^{-q(p-1)/p}}.$$

As an application of Lemma 1.2, we have

$$\begin{aligned} B^{*q} &\geq \sup_{\theta} \left\{ \left[\sup_{x \leq \theta} \frac{\mu[x, \theta]}{\hat{\nu}[-M, x]^{-q(p-1)/p}} \right] \wedge \left[\sup_{y \geq \theta} \frac{\mu(\theta, y)}{\hat{\nu}[y, N]^{-q(p-1)/p}} \right] \right\} \\ &= \sup_{\theta} [B_{\theta}^{-} \wedge B_{\theta}^{+}]^q. \end{aligned}$$

Here in the last step, we have used the condition $\mu_{pp} = 0$. Since we can represent

$$\mu_{\text{sing}}[x, \theta] = \mu_{\text{sing}}[-M, \theta] - \mu_{\text{sing}}[-M, x]$$

(the last term is continuous in x) and $\mu_{pp} = 0$, it follows that the function $\mu[x, \theta]$ is continuous in x and θ . By choosing $\bar{\theta}$ such that $B_{\bar{\theta}}^- = B_{\bar{\theta}}^+$, it follows that $B_{\bar{\theta}}^- \leq B^*$ (just proved) and furthermore,

$$\begin{aligned} A &\leq \inf_{\theta \in [-M, N]} (A_{\theta}^- \vee A_{\theta}^+) \quad (\text{by Proposition 1.1}) \\ &\leq k_{q,p} \inf_{\theta \in [-M, N]} (B_{\theta}^- \vee B_{\theta}^+) \quad (\text{by (1.2)}) \\ &\leq k_{q,p} B_{\bar{\theta}}^- \quad (\text{by definition of } \bar{\theta}) \\ &\leq k_{q,p} B^*. \end{aligned} \quad \square$$

Note that the parameter θ is used temporary in the proof above. Thus, the splitting procedure is a bridge to go to the upper estimate but our final result does not depend on the splitting points. The simple technique used in proving the upper estimate above is in common, and will be used several times later (Lemma 2.3 and Theorem 3.2).

Before moving further, let us discuss the technical assumption that $\mu_{pp} = 0$. In the study of $B_{\theta}^{\pm}$ for half-spaces, one first handles with the case that $\mu \ll dx$ and $\nu \ll dx$ and then removes this restriction by the following technique. Without loss of generality, assume that $M, N < \infty$. Besides, we may also assume that $f' \geq 0$ in the study of the upper estimate. The idea is to use an approximating procedure (cf. [14] or [4, p. 45]). Note that

$$\left(\int_{-M}^N f^q d\mu \right)^{1/q} = \left(\int_{-M}^N \mu[x, N] df(x)^q \right)^{1/q}.$$

Now, we can approximate $\mu[x, N]$ by a sequence of absolutely continuous, decreasing functions $\{g_n\}$ having the property: $g_n \leq \mu[\cdot, N]$ for every n ; as $n \rightarrow \infty$, $g_n(x)$ converges to $\mu[x, N]$ for almost all x . Thus, we can first replace $\mu[x, N]$ by absolutely continuous g_n and then pass to the limit as $n \rightarrow \infty$. Actually, now $\mu[\cdot, N]$ consists of three parts: the absolutely continuous part, the singular continuous one plus a step function. Each of them is decreasing. There is nothing to do about the absolutely continuous part. The singular decreasing continuous function can be approximated from below by decreasing step functions. Furthermore, each of the step functions can be approximated from below almost everywhere by absolutely continuous decreasing functions. The new difficulty arises: even though we have the control $g_n(x) \leq \mu[x, N]$ for all x , but we still do not know how to construct a sequence $\{g_n\}$ as above having the control $(0 \leq) g_n(x) - g_n(y) \leq \mu[x, y]$ for every pair $\{x, y\}$ with $x < y$ and each n .

For the lower bound of A , the dual proof of Lemma 1.2 does not work well as remarked below Lemma 1.2. More precisely, what we can obtain by Proposition 1.1 and (1.6) is as follows.

$$A \geq 2^{1/q-1/p} \sup_{x \leq \bar{\theta} \leq y} \left\{ \mu[x, y]^{1/q} [\hat{\nu}[-M, x]^{-\frac{q(p-1)}{p}} + \hat{\nu}[y, N]^{-\frac{q(p-1)}{p}}]^{-1/q} \right\}, \quad (1.9)$$

where $\bar{\theta}$ is the solution of the equation $B_{\bar{\theta}}^- = B_{\bar{\theta}}^+$. The result is less satisfactory since $\bar{\theta}$ (unknown explicitly) is included. Fortunately, there is a direct technique (cf. [6, Proof (b) of Theorem 8.2]) to handle with the lower estimate.

Lemma 1.4 For $1 < p, q < \infty$, we have

$$A \geq \sup_{-M \leq x \leq y \leq N} \left\{ \mu[x, y]^{1/q} (\hat{\nu}[-M, x]^{1-p} + \hat{\nu}[y, N]^{1-p})^{-1/p} \right\} =: B_*.$$

Proof Given $m', m, \theta, n, n' \in [-M, N]$ with $m' < m < \theta < n < n'$, define

$$f(x) = \gamma \mathbb{1}_{\{m' \leq x \leq \theta\}} \hat{\nu}[m', x \wedge m] + \mathbb{1}_{\{\theta < x \leq n'\}} \hat{\nu}[x \vee n, n'],$$

where $\gamma = \hat{\nu}[n, n'] / \hat{\nu}[m', m]$. Clearly, f is absolutely continuous. We have

$$\left(\int_{-M}^N |f|^q d\mu \right)^{1/q} = \left(\int_{m'}^{n'} |f|^q d\mu \right)^{1/q} \geq \left(\int_m^n |f|^q d\mu \right)^{1/q} = \mu[m, n]^{1/q} \hat{\nu}[n, n']$$

and

$$\begin{aligned} \left(\int_{-M}^N |f'|^p d\nu \right)^{1/p} &= \left(\gamma^p \int_{m'}^m h^p d\nu + \int_n^{n'} h^p d\nu \right)^{1/p} \\ &= (\gamma^p \hat{\nu}[m', m] + \hat{\nu}[n, n'])^{1/p}. \end{aligned}$$

Here in the last step, we have ignored the singular part of ν since the original inequality is equivalent to the one having $\nu = \nu^*$. To see this, simply set $f' = 0$ on the singular part of ν . Thus, the optimal constant A satisfies

$$A \geq (\mu[m, n]^{1/q} \hat{\nu}[n, n']) (\gamma^p \hat{\nu}[m', m] + \hat{\nu}[n, n'])^{-1/p}.$$

But

$$\begin{aligned} &(\gamma^p \hat{\nu}[m', m] + \hat{\nu}[n, n']) \hat{\nu}[n, n']^{-p} \\ &= \{ \hat{\nu}[n, n']^p \hat{\nu}[m', m]^{-p} \hat{\nu}[m', m] + \hat{\nu}[n, n'] \} \hat{\nu}[n, n']^{-p} \\ &= \hat{\nu}[m', m]^{1-p} + \hat{\nu}[n, n']^{1-p}, \end{aligned}$$

it follows that

$$A \geq \mu[m, n]^{1/q} (\hat{\nu}[m', m]^{1-p} + \hat{\nu}[n, n']^{1-p})^{-1/p}.$$

Let $m' \downarrow -M$, $n' \uparrow N$ and then make supremum with respect to $m = x \leq y = n$. We get the required assertion. $\square$

On the comparison of B_* and B^* , it is obvious that $B^* = B_*$ if $p = q$. In general, we have the following result.

Lemma 1.5 Let $q \geq p$. Then we have $B_* \leq B^* \leq 2^{1/p-1/q} B_*$.

Proof Simply apply the c_r -inequality:

$$(\alpha + \beta)^r \leq 2^{(r-1) \vee 0} (\alpha^r + \beta^r).$$

(a) Set

$$\alpha = \hat{\nu}[-M, x]^{\frac{q(1-p)}{p}}, \quad \beta = \hat{\nu}[y, N]^{\frac{q(1-p)}{p}}, \quad r = \frac{p}{q} \in (0, 1].$$

It follows that

$$(\alpha + \beta)^{1/q} \leq (\alpha^{p/q} + \beta^{p/q})^{1/p},$$

and then $B^* \geq B_*$.

(b) Set

$$\alpha = \hat{\nu}[-M, x]^{1-p}, \quad \beta = \hat{\nu}[y, N]^{1-p}, \quad r = \frac{q}{p} \geq 1.$$

We have

$$(\alpha + \beta)^{1/p} \leq 2^{1/p-1/q} (\alpha^{q/p} + \beta^{q/p})^{1/p},$$

and then $B_* \geq 2^{1/q-1/p} B^*$. Certainly, in this case the assertion can also be deduced by Jensen's inequality. $\square$

We mention that even though their supremums are equivalent but in the proofs of Lemmas 1.3 and 1.4, the expressions of B^* and B_* are not exchangeable, because B_* does not own the homogeneous of that of B^* .

We are now ready to state our first main result.

Theorem 1.6 *The optimal constant A in the Hardy-type inequality (1.1) satisfies*

- (1) $A \leq k_{q,p} B^*$ for $1 < p \leq q < \infty$ once $\mu_{pp} = 0$, where $k_{q,p}$ is defined by (1.5), and
- (2) $A \geq B_*$ for $1 < p, q < \infty$, where

$$B^* = \sup_{-M \leq x \leq y \leq N} \left\{ \mu[x, y]^{1/q} (\hat{\nu}[-M, x]^{q(1-p)/p} + \hat{\nu}[y, N]^{q(1-p)/p})^{-1/q} \right\},$$

$$B_* = \sup_{-M \leq x \leq y \leq N} \left\{ \mu[x, y]^{1/q} (\hat{\nu}[-M, x]^{1-p} + \hat{\nu}[y, N]^{1-p})^{-1/p} \right\}.$$

Moreover, we have $B_* \leq B^* \leq 2^{1/p-1/q} B_*$ when $q \geq p$.

Proof The conclusions are combination of Lemmas 1.3–1.5. $\square$

It is interesting to have a look at the factor $k_{q,p}$ in Theorem 1.6. When $p = q$, the factor becomes

$$\left(\frac{q}{q-1} \right)^{(q-1)/q} q^{1/q}.$$

On $(1, \infty)$, it is unimodal having maximum 2 at $q = 2$ and decreases to 1 as $q \rightarrow 1$ or ∞ . More generally, the rough ratio of the upper and lower bounds is no more than

$$\left(1 + \frac{q}{p'} \right)^{1/q} \left(1 + \frac{p'}{q} \right)^{1/p'} 2^{1/p-1/q}$$

which is again ≤ 2 (for every $q \geq p$), having equality sign iff $p = q = 2$.

The study on the inequality (1.1) was began by Gurka in an unpublished paper using a common constant

$$\begin{aligned} \tilde{B} &= \sup_{-M \leq x \leq y \leq N} \left\{ \mu[x, y]^{\frac{1}{q}} (\hat{\nu}[-M, x]^{1-p} \vee \hat{\nu}[y, N]^{1-p})^{-\frac{1}{p}} \right\} \\ &= \sup_{-M \leq x \leq y \leq N} \left\{ \mu[x, y]^{\frac{1}{q}} (\hat{\nu}[-M, x]^{\frac{p-1}{p}} \wedge \hat{\nu}[y, N]^{\frac{p-1}{p}}) \right\} \end{aligned}$$

and having a universal factor 8. This $\tilde{B}$ is closely related to B_* : replacing “+” with “ $\vee$ ”, we obtain $\tilde{B}$ from B_* . Gurka's result was then improved in [5, Theorem 8.2] with a smaller factor (unexplicit one ≈ 4.71 and explicit one $= 2\sqrt{6}$ in the case of $p = q = 2$). Note that using the inequalities

$$\alpha \vee \beta \leq \alpha + \beta \leq 2(\alpha \vee \beta),$$

from Theorem 1.6, it follows that we have lower and upper bounds replacing B^* and B_* by the same $\tilde{B}$ with an additional factor $2^{-1/p}$ for the lower estimate. Then the factor becomes $2\sqrt{2}$ in the case of $p = q = 2$. Replacing $\alpha \vee \beta$ with $\alpha + \beta$ is an essential difference of the present

paper from the previous ones in the bilateral situation. Besides, the inequality (1.1) was also proved in [5, Theorem 8.8] with a common constant

$$\tilde{B} = \inf_{-M \leq \theta \leq N} \left\{ \left[\sup_{-M \leq x \leq \theta} \mu[x, \theta]^{\frac{1}{q}} \hat{\nu}[-M, x]^{\frac{p-1}{p}} \right] \vee \left[\sup_{\theta \leq y \leq N} \mu[\theta, y]^{\frac{1}{q}} \hat{\nu}[y, N]^{\frac{p-1}{p}} \right] \right\}$$

having a factor

$$2^{\frac{1}{p}} \left(1 + \frac{q}{p'} \right)^{\frac{1}{q}} \left(1 + \frac{p'}{q} \right)^{\frac{1}{p'}}$$

which has an additional factor is $2^{1/p}$ than (1.5). The last result is related to our splitting technique. All of these results use the assumption that $\mu \ll dx$ and $\nu \ll dx$.

Before moving further, we want to describe B_* and B^* more carefully. It also leads some quantities which are easier in practical computations. For this, we need some preparation. Assume that (1.7) holds. For each $x \in (-M, N)$, let $y(x)$ be the unique solution of the equation

$$\hat{\nu}[-M, x] = \hat{\nu}[y, N].$$

Next, let $m(\hat{\nu})$ be a solution to the equation

$$y(x) = x, \quad x \in (-M, N).$$

Thus, $m(\hat{\nu})$ is actually the median of the measure $\hat{\nu}$ (but not ν):

$$\hat{\nu}[-M, m] = \hat{\nu}[m, N].$$

Set

$$H_{\mu, \nu}(x, y) = \mu[x, y]^{1/q} [\hat{\nu}[-M, x]^{1-p} + \hat{\nu}[y, N]^{1-p}]^{-1/p}, \quad -M \leq x \leq y \leq N.$$

Define

$$H^o = 2^{-1/p} \sup_{x \in (-M, m(\hat{\nu}))} \mu[x, y(x)]^{1/q} \hat{\nu}[-M, x]^{(p-1)/p}. \quad (1.10)$$

Denote by Γ be the limiting points of $H_{\mu, \nu}(x, y)$ as $\mu[y, N] = \infty$ or $\mu[-M, x] = \infty$, as well as the iterated limits if $\mu[-M, N] = \infty$ when $M = \infty = N$. Set

$$H^\partial = \begin{cases} \sup\{\gamma : \gamma \in \Gamma\}, & \text{if } \Gamma \neq \emptyset, \\ 0, & \text{if } \Gamma = \emptyset. \end{cases} \quad (1.11)$$

Clearly, $H^\partial = 0$ if $M, N < \infty$.

We are now ready to describe B^* and B_* in terms of H^o and H^∂ .

Lemma 1.7 *Let (1.7) hold. Then we have*

$$H^o \vee H^\partial \leq B_* \leq (2^{1/p} H^o) \vee H^\partial.$$

Proof Rewrite H as

$$H_{\mu, \nu}(x, y) = \left[\frac{\hat{\nu}[-M, x]^{1-p} + \hat{\nu}[y, N]^{1-p}}{\mu[x, y]^{p/q}} \right]^{-1/p}.$$

under (1.7), because for finite x and y with $x \leq y$, we have

$$\frac{\hat{\nu}[-M, x]^{1-p} + \hat{\nu}[y, N]^{1-p}}{\mu[x, y]^{p/q}} \geq \frac{2}{\mu[x, y]^{p/q}} [\hat{\nu}[-M, x] \hat{\nu}[y, N]]^{(1-p)/2},$$

and the equality sign holds iff $\hat{\nu}[-M, x] = \hat{\nu}[y, N]$ which gives us the solution $y(x)$. Thus, we obtain

$$\inf_{x \leq y} \frac{\hat{\nu}[-M, x]^{1-p} + \hat{\nu}[y, N]^{1-p}}{\mu[x, y]^{p/q}} \leq 2 \inf_{x \leq m(\hat{\nu})} \frac{\hat{\nu}[-M, x]^{1-p}}{\mu[x, y(x)]^{p/q}}, \quad (1.12)$$

since $\{(x, y(x)) : x \leq m(\hat{\nu})\} \subset \{(x, y) : x, y \in (-M, N)\}$. This gives us a lower bound of the supremum of $H_{\mu, \nu}(x, y)$ over the set $\{(x, y) : x < y, \mu[x, y] < \infty\}$. Next, we have

$$\begin{aligned} \inf_{x \leq y} \frac{\hat{\nu}[-M, x]^{1-p} + \hat{\nu}[y, N]^{1-p}}{\mu[x, y]^{p/q}} &= \inf_{x \leq y} \left\{ \frac{\hat{\nu}[-M, x]^{1-p}}{\mu[x, y]^{p/q}} + \frac{\hat{\nu}[y, N]^{1-p}}{\mu[x, y]^{p/q}} \right\} \\ &\geq \inf_{x \leq y} \left\{ \frac{\hat{\nu}[-M, x]^{1-p}}{\mu[x, y]^{p/q}} \vee \frac{\hat{\nu}[y, N]^{1-p}}{\mu[x, y]^{p/q}} \right\} \\ &=: \xi. \end{aligned}$$

Without loss of generality, assume that $M, N < \infty$. Because of the continuity of the involved functions, the minimum ξ can be achieved at some pair (x_0, y_0) . We now prove that (x_0, y_0) should be located at the surface where the two terms in the last $\{\dots\}$ are equal. Otherwise, without loss of generality, assume that

$$\varepsilon := \mu[x_0, y_0]^{-p/q} \hat{\nu}[-M, x_0]^{1-p} - \mu[x_0, y_0]^{-p/q} \hat{\nu}[y_0, N]^{1-p} > 0.$$

Let $\bar{y} > y_0$ be sufficiently close to y_0 . Then we have

$$\mu[x_0, \bar{y}]^{-p/q} \hat{\nu}[-M, x_0]^{1-p} < \mu[x_0, y_0]^{-p/q} \hat{\nu}[-M, x_0]^{1-p}$$

(here we have used the preassumption that μ is positive on each subinterval) and

$$\mu[x_0, \bar{y}]^{-p/q} \hat{\nu}[\bar{y}, N]^{1-p} < \mu[x_0, y_0]^{-p/q} \hat{\nu}[y_0, N]^{1-p} + \varepsilon/2,$$

due to the continuity of the involved functions. We have thus obtained a pair $(x_0, \bar{y})$ with $x_0 < \bar{y}$ such that

$$\left\{ \left[\mu[x_0, \bar{y}]^{-p/q} \hat{\nu}[-M, x_0]^{1-p} \right] \vee \left[\mu[x_0, \bar{y}]^{-p/q} \hat{\nu}[\bar{y}, N]^{1-p} \right] \right\} < \xi.$$

This is a contradiction to the minimum property of ξ . Therefore, we obtain

$$\inf_{x \leq y} \mu[x, y]^{-p/q} [\hat{\nu}[-M, x]^{1-p} + \hat{\nu}[y, N]^{1-p}] \geq \inf_{x \leq m(\hat{\nu})} \frac{\hat{\nu}[-M, x]^{1-p}}{\mu[x, y(x)]^{p/q}}. \quad (1.13)$$

From this, we obtain an upper bound of the supremum of $H_{\mu, \nu}(x, y)$ over the set $\{(x, y) : x < y, \mu[x, y] < \infty\}$ in terms of H^o up to a factor $2^{-1/p}$. In other words, we have worked out the case that the supremum is achieved inside of the interval. In general, it may be achieved at the ∞ -boundaries (at which $\mu[y, N] = \infty$ or $\mu[-M, x] = \infty$). This leads to the boundary condition H^∂ , when one of M and N is infinite. Combining these two parts together, we get the estimates of B_* under (1.7). $\square$

An easier way to understand what was going on in the last proof is look at the following simple example. Consider functions $f(x) = 2x$ and $g(x) = 3 - x$ on $[0, 2]$. They intersect uniquely at the point $x^* = 1$. Then we have

$$2\sqrt{f(x^*)g(x^*)} = 4 > \inf_{x \in [0, 2]} (f(x) + g(x)) = 3 > \inf_{x \in [0, 2]} [f(x) \vee g(x)] = f(x^*) = 2.$$

If we rewrite the first term as $2 \inf_{x \in [0,2]} (f(x) \vee g(x))$, then it becomes obvious that the middle term can be bounded by the first and the last ones. Certainly, the bounds are usually not sharp.

Lemma 1.8 *Let (1.7) hold. Then we have*

$$(2^{1/p-1/q} H^o) \vee H^\partial \leq B^* \leq (2^{1/p} H^o) \vee H^\partial.$$

Proof Note that the difference of B^* and B_* is only the summation terms. If one of the terms in the sum is ignored, then the remaining terms coincide with each other. Thus, the boundary condition H^∂ is the same for B^* and B_* . When $M, N < \infty$, the proof of the comparison of B^* with H^o is similar to the last one. $\square$

We remark that in the degenerated case that (1.7) does not hold, say $\hat{\nu}[y, N] = \infty$, then we have obviously that $B_* = B^*$.

As a combination of the last two lemmas, we obtain the following simple criterion.

Corollary 1.9 *The Hardy-type inequality (1.1) holds iff $H^o \vee H^\partial < \infty$.*

Proof When (1.7) holds, the assertion follows from the last two lemmas. Note that if $\hat{\nu}[y, N] = \infty$ for instance, we have $B^* = B_*$ and so the assertion is described by H^∂ only. $\square$

We now extend Theorem 1.6 to a more general setup which is mainly used in interpolation of L^p -spaces. For this, we need a class of normed linear spaces $(\mathbb{B}, \|\cdot\|_{\mathbb{B}}, \mu)$ consisting of real Borel measurable functions on a measurable space $(X, \mathcal{X}, \mu)$. We now modify the hypotheses on the normed linear spaces given in [10, Chapter 7] as follows.

Hypotheses 1.10 (H1) In the case that $\mu(X) = \infty$, $\mathbf{1}_K \in \mathbb{B}$ for all compact K . Otherwise, $\mathbf{1} \in \mathbb{B}$.

(H2) If $h \in \mathbb{B}$ and $|f| \leq h$, then $f \in \mathbb{B}$.

(H3) $\|f\|_{\mathbb{B}} = \sup_{g \in \mathcal{G}} \int_X |f|g d\mu$,

where $\mathcal{G}$, to be specified case by case, is a class of nonnegative $\mathcal{X}$ -measurable functions. A typical example is $\mathcal{G} = \{\mathbf{1}\}$ and then $\mathbb{B} = L^1(\mu)$. In what follows, the measure space $(X, \mathcal{X}, \mu)$ is fixed to be $([-M, N], \mathcal{B}([-M, N]), \mu)$. We often use the dual representation of the norm

$$\left(\int_{-M}^N |f|^r d\mu \right)^{1/r} = \sup_{g \in \text{The unit ball in } L^{r'}(\mu)} \int_{-M}^N |f|g d\mu,$$

where r' is the conjugate number of $r (\geq 1)$. Throughout this paper, we assume (H1)–(H3) for $(\mathbb{B}, \|\cdot\|_{\mathbb{B}}, \mu)$ without mentioned again.

For simplicity, we write the L^p -norm with respect to μ as $\|\cdot\|_{\mu,p}$. If necessary, we also write $\|\cdot\|_{\alpha,\beta;\mu,p}$ to indicate the interval $[\alpha, \beta]$.

Theorem 1.11 *Let $\mathcal{G}$ satisfy Hypotheses 1.10 and consider the Hardy-type inequality*

$$\|f^q\|_{\mathbb{B}}^{1/q} \leq A_{\mathbb{B}} \|f'\|_{\nu,p}, \quad f(-M) = 0 \text{ and } f(N) = 0.$$

(1) *Then the optimal constant $A_{\mathbb{B}}$ satisfies*

$$A_{\mathbb{B}} \leq k_{q,p} B_{\mathbb{B}}^* \quad \text{for } 1 < p \leq q < \infty \text{ once } \mu_{pp} = 0, \text{ and}$$

(2) $A_{\mathbb{B}} \geq B_{\mathbb{B}*}$ for $1 < p, q < \infty$, where

$$B_{\mathbb{B}}^* = \sup_{-M \leq x \leq y \leq N} \left\{ \|\mathbb{1}_{[x,y]}\|_{\mathbb{B}}^{\frac{1}{q}} \left(\hat{\nu}[-M, x]^{\frac{q(1-p)}{p}} + \hat{\nu}[y, N]^{\frac{q(1-p)}{p}} \right)^{-\frac{1}{q}} \right\},$$

$$B_{\mathbb{B}*} = \sup_{-M \leq x \leq y \leq N} \left\{ \|\mathbb{1}_{[x,y]}\|_{\mathbb{B}}^{\frac{1}{q}} \left(\hat{\nu}[-M, x]^{1-p} + \hat{\nu}[y, N]^{1-p} \right)^{-\frac{1}{p}} \right\}.$$

Moreover, we have $B_{\mathbb{B}*} \leq B_{\mathbb{B}}^* \leq 2^{1/p-1/q} B_{\mathbb{B}*}$ whenever $q \geq p$.

Proof Let $g \in \mathcal{G}$. Without loss of generality, assume that $g > 0$. For the pair $\mu_g := g\mu$ and ν , by Theorem 1.6, we know that the corresponding optimal constant A_g in (1.1) satisfies

$$B_{g*} \leq A_g \leq k_{q,p} B_{g*}^*,$$

where

$$B_g^* = \sup_{-M \leq x \leq y \leq N} \left\{ \mu_g[x, y]^{\frac{1}{q}} \left(\hat{\nu}[-M, x]^{\frac{q(1-p)}{p}} + \hat{\nu}[y, N]^{\frac{q(1-p)}{p}} \right)^{-\frac{1}{q}} \right\},$$

$$B_{g*} = \sup_{-M \leq x \leq y \leq N} \left\{ \mu_g[x, y]^{\frac{1}{q}} \left(\hat{\nu}[-M, x]^{1-p} + \hat{\nu}[y, N]^{1-p} \right)^{-\frac{1}{p}} \right\}.$$

Hence

$$\begin{aligned} \sup_{g \in \mathcal{G}} B_{g*} &= \sup_{g \in \mathcal{G}} \sup_{x \leq y} \left\{ \mu_g[x, y]^{1/q} \left(\hat{\nu}[-M, x]^{1-p} + \hat{\nu}[y, N]^{1-p} \right)^{-1/p} \right\} \\ &= \sup_{x \leq y} \left\{ \left(\sup_g \mu_g[x, y] \right)^{1/q} \left(\hat{\nu}[-M, x]^{1-p} + \hat{\nu}[y, N]^{1-p} \right)^{-1/p} \right\} \\ &= \sup_{x \leq y} \left\{ \|\mathbb{1}_{[x,y]}\|_{\mathbb{B}}^{1/q} \left(\hat{\nu}[-M, x]^{1-p} + \hat{\nu}[y, N]^{1-p} \right)^{-1/p} \right\} \\ &= B_{\mathbb{B}*}. \end{aligned}$$

Similarly, we have $\sup_{g \in \mathcal{G}} B_g^* = B_{\mathbb{B}}^*$. From these facts, we obtain the estimates of $A_{\mathbb{B}} = \sup_{g \in \mathcal{G}} A_g$ immediately. The last assertion then follows from Lemma 1.5 (or its proof). $\square$

As an application of Theorem 1.11, it follows that the optimal constant $A_{\mathbb{B}}$ in the inequality

$$\|f^p\|_{\mathbb{B}}^{1/p} \leq A_{\mathbb{B}} \|f'\|_{\nu,p}, \quad f(-M) = 0 \text{ and } f(N) = 0$$

satisfies

$$B_{\mathbb{B}*} \leq A_{\mathbb{B}} \leq k_{p,p} B_{\mathbb{B}*}, \quad (1.14)$$

where

$$B_{\mathbb{B}*} = \sup_{-M \leq x \leq y \leq N} \left\{ \|\mathbb{1}_{[x,y]}\|_{\mathbb{B}}^{\frac{1}{p}} \left(\hat{\nu}[-M, x]^{1-p} + \hat{\nu}[y, N]^{1-p} \right)^{-\frac{1}{p}} \right\}.$$

Clearly, this result is simpler than Theorem 1.11. Now, applying this result to $\mathbb{B} = L^{q/(q-p)}$ ($q > p$, where $q/(q-p)$ is the conjugate number of q/p), we return to Theorem 1.6 with a factor $k_{p,p}$ different from $k_{q,p}$. In other words, we have arrived at a conclusion that the optimal constant A in the Hardy-type inequality (1.1) (with $q \geq p$) satisfies

$$B_* \leq A \leq k_{p,p} B_*, \quad (1.15)$$

where B_* is the same as in Theorem 1.6. However, as we will see from Example 1.13 below that this result (1.15) may be less sharp than Theorem 1.6. Thus, lifting the left-hand side of the Hardy-type inequality from $L^p(\mu)$ to $\mathbb{B}$ does keep the constant $k_{p,p}$ but can not improve it. We have thus explained the reason why on the left-hand side of the first inequality in Theorem 1.11,

we use q but not p . In the discrete context and $p = 2$, the conclusion (1.14) was presented by [6, Theorem 8.2].

In view of the direct proof for the lower estimate of A , the restriction $\mu_{pp} = 0$ coming from our splitting technique seems unnecessary. As just mentioned, the conclusion (1.14) was proved in the discrete situation when $p = 2$ by [6, Theorem 8.2]. Its proof should be meaningful in the present continuous case (cf. [10, Proof of Corollary 7.6], simply replacing the function $e^{C(x)}$ there by the one $h(x)$ here). Hence for (1.14) in the case of $p = 2$, the condition $\mu_{pp} = 0$ is again not needed. Thus, one may remove this condition by the capacitary method first for $q = p$, if possible, and then extend to general normed linear space $\mathbb{B}$ as shown above. The present proof depends heavily on the splitting property as shown by the first step of the proof of Lemma 1.3 and the comments below Lemma 1.3. This is one of the reasons why the case of $q < p$ is missed here. Actually, the last case is a rather different story, refer to [5, Theorem 8.17] or [4, pp. 50–51].

To illustrate the application of Theorem 1.6, we study two examples.

Example 1.12 Let $(-M, N) = (0, 1)$ and $d\mu = d\nu = dx$. Then $H^\partial = 0$ and

$$B_* = H^o = \frac{1}{2} \left(\frac{p}{p-q+pq} \right)^{\frac{1}{q}} \left(\frac{(p-1)q}{p-q+pq} \right)^{\frac{p-1}{p}},$$

$$B^* = 2^{1/p-1/q} H^o, \quad 1 < p \leq q < \infty.$$

When $p = q = 2$, it is known that $A = \pi^{-1}$ and $B^* = B_* = 1/4$ (cf. [7, Example 5.2]). More generally, when $q = p$, it is known that $A = (2\pi)^{-1} p(p-1)^{-1/p} \sin \frac{\pi}{p}$, and we have $B^* = B_* = 2^{-1} p^{-1/p} \left(\frac{p-1}{p} \right)^{\frac{p-1}{p}}$ and $k_{p,p} B^* = 2^{-1}$.

Proof We have $h \equiv 1$, $y(x) = 1 - x$, and $m(\hat{\nu}) = 1/2$. The function

$$H_{\mu,\nu}(x, y(x)) = 2^{-\frac{1}{p}} \mu[x, y(x)]^{\frac{1}{q}} \hat{\nu}[0, x]^{\frac{p-1}{p}} = 2^{-\frac{1}{p}} (1-2x)^{\frac{1}{q}} x^{\frac{p-1}{p}}$$

achieves its maximum at

$$x = \frac{1}{2} \left(1 + \frac{p}{(p-1)q} \right)^{-1} < \frac{1}{2}.$$

From this, we obtain H^o . Next, note that

$$B_* = \sup_{x \leq y} \frac{(1 - (1-y) - x)^{1/q}}{(x^{1-p} + (1-y)^{1-p})^{1/p}},$$

$$B^* = \sup_{x \leq y} \frac{(1 - (1-y) - x)^{1/q}}{(x^{q(1-p)/p} + (1-y)^{q(1-p)/p})^{1/q}}.$$

Both of them are symmetric in x and $1 - y$. Hence $B_* = H^o$ and furthermore, $B^* = 2^{1/p-1/q} H^o$. $\square$

The next example illustrates the role played by H^∂ .

Example 1.13 Let $\mu(dx) = dx$ and $\nu(dx) = x^2 dx$ on $(1, \infty)$. Assume that $p \in (1, 3)$. Then the inequality does not hold if $q \in [p, p/(3-p))$. Otherwise, the inequality holds with

$$B_* = B^* = H^\partial = \left(\frac{p-1}{3-p} \right)^{\frac{p-1}{p}} \quad \text{if } q = \frac{p}{3-p} \text{ and } p \in [2, 3);$$

B_* and B^* are bounded in terms of H^o if $q > \frac{p}{3-p}$ and $q \geq p$,

where

$$H^o = 2^{-\frac{1}{p}} \left(\frac{p-1}{3-p} \right)^{\frac{p-1}{p}} \sup_{x \in (1, 2^{(3-p)/(p-1)}]} \left[\left(1 - x^{\frac{p-3}{p-1}} \right)^{\frac{p-1}{p-3}} - x \right]^{\frac{1}{q}} \left(1 - x^{\frac{p-3}{p-1}} \right)^{\frac{p-1}{p}}.$$

When $p = q = 2$, we have $B^* = B_* = 1$. In this case, $A = 2$ and so the upper estimate $2B^*$ in Theorem 1.6 is exact (cf. [7, Example 5.4]). For fixed $p = 2$, when q varies from 2.01 to 4.8, the five quantities we have worked so far are shown in Figure 1. The ratio of the upper and lower bounds is decreasing in q and is no more than 2. Note that we have a common lower bound B_* in Theorem 1.6 and (1.15), but the upper bound in Theorem 1.6 is better than that in (1.15).

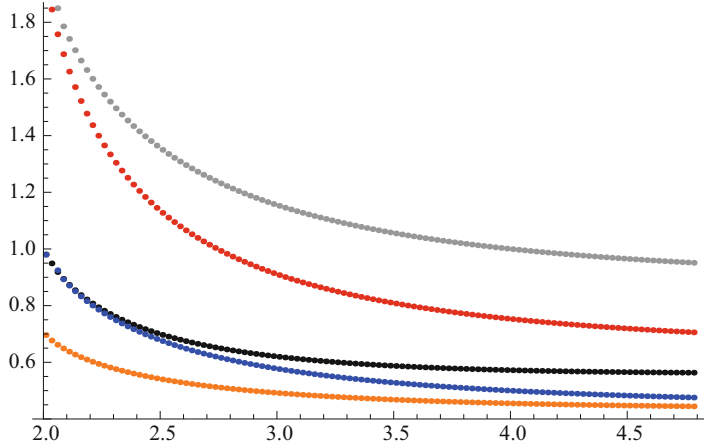


Figure 1 The curves from bottom to top are H^o , B_* , B^* , $k_{q,p}B^*$, and $k_{p,p}B_*$ respectively

Proof We have $h(x) = x^{-2/(p-1)}$. Then

$$\int_1^x h = \frac{p-1}{3-p} (1 - x^{\frac{p-3}{p-1}}), \quad \int_y^\infty h = \frac{p-1}{3-p} y^{\frac{p-3}{p-1}},$$

where and in what follows, the Lebesgue measure dz is omitted. Hence

$$y(x) = \left(1 - x^{\frac{p-3}{p-1}} \right)^{\frac{p-1}{p-3}}, \quad m(\hat{\nu}) = 2^{\frac{3-p}{p-1}},$$

$$H_{\mu,\nu}(x, y) = \left(\frac{p-1}{3-p} \right)^{\frac{p-1}{p}} (y-x)^{\frac{1}{q}} \left[\left(1 - x^{\frac{p-3}{p-1}} \right)^{1-p} + y^{3-p} \right]^{-\frac{1}{p}}.$$

In particular,

$$\overline{\lim}_{y \rightarrow \infty} H_{\mu,\nu}(x, y) = \lim_{y \rightarrow \infty} \left(\frac{p-1}{3-p} \right)^{\frac{p-1}{p}} y^{1+\frac{1}{q}-\frac{3}{p}} = \begin{cases} \left(\frac{p-1}{3-p} \right)^{\frac{p-1}{p}}, & \text{if } 1 + \frac{1}{q} = \frac{3}{p}, \\ \infty, & \text{if } 1 + \frac{1}{q} > \frac{3}{p}, \\ 0, & \text{if } 1 + \frac{1}{q} < \frac{3}{p}. \end{cases}$$

The right-hand side is our H^∂ . Thus, if $1 + \frac{1}{q} > \frac{3}{p}$, then $B_* = \infty$. Next, we have

$$H_{\mu,\nu}(x, y(x)) = 2^{-\frac{1}{p}} \left(\frac{p-1}{3-p} \right)^{\frac{p-1}{p}} \left[\left(1 - x^{\frac{p-3}{p-1}} \right)^{\frac{p-1}{p-3}} - x \right]^{\frac{1}{q}} \left(1 - x^{\frac{p-3}{p-1}} \right)^{\frac{p-1}{p}}.$$

Then

$$H^o = 2^{-\frac{1}{p}} \left(\frac{p-1}{3-p} \right)^{\frac{p-1}{p}} \sup_{x \in (1, 2^{(3-p)/(p-1)})} \left[\left(1 - x^{\frac{p-3}{p-1}} \right)^{\frac{p-1}{p-3}} - x \right]^{\frac{1}{q}} \left(1 - x^{\frac{p-3}{p-1}} \right)^{\frac{p-1}{p}}.$$

The point here is that $H^o \leq H^\partial$ and $2^{1/p} H^o \leq H^\partial$ in the case of $q = p/(3-p)$ and $p \geq 2$. Besides, we have

$$B_* = \left(\frac{p-1}{3-p} \right)^{\frac{p-1}{p}} \sup_{x < y} \left\{ (y-x)^{\frac{1}{q}} \left[\left(1 - x^{\frac{p-3}{p-1}} \right)^{1-p} + y^{3-p} \right]^{-\frac{1}{p}} \right\},$$

$$B^* = \left(\frac{p-1}{3-p} \right)^{\frac{p-1}{p}} \sup_{x < y} \left\{ (y-x)^{\frac{1}{q}} \left[\left(1 - x^{\frac{p-3}{p-1}} \right)^{\frac{q(1-p)}{p}} + y^{\frac{q(3-p)}{p}} \right]^{-\frac{1}{q}} \right\}.$$

Finally, numerical computation gives us the quantities B^* et al., as shown in Figure 1. $\square$

2 The Case of Mean Zero

Throughout this section, we assume that $\mu[-M, N] < \infty$ and define a probability measure $\pi = (\mu[-M, N])^{-1} \mu$. In probabilistic language, we are going to study the ergodic case of the corresponding processes. Corresponding to the three inequalities given at the beginning of the last section, we now study the inequality

$$\left(\int_{-M}^N |f - \pi(f)|^q d\mu \right)^{1/q} \leq A \left(\int_{-M}^N |f'|^p d\nu \right)^{1/p}, \quad (2.1)$$

where $\pi(f) = \int f d\pi$, in terms of

$$\left(\int_{\theta}^N |f|^q d\mu \right)^{1/q} \leq A_{\theta}^{+} \left(\int_{\theta}^N |f'|^p d\nu \right)^{1/p}, \quad f(\theta) = 0,$$

$$\left(\int_{-M}^{\theta} |f|^q d\mu \right)^{1/q} \leq A_{\theta}^{-} \left(\int_{-M}^{\theta} |f'|^p d\nu \right)^{1/p}, \quad f(\theta) = 0.$$

To save our notation, without any confusion, we use the same notation A , $A_{\theta}^{\pm}$ and so on as in the last section.

Before moving further, let us mention the spectral meaning of (1.1) and (2.1). Suppose that $\mu \ll dx$ and $\nu \ll dx$, denote by $u = d\mu/dx$ and $v = d\nu/dx$. Then the inverse of the optimal constant A in (1.1) and (2.1), when $q = p$, corresponds to the infimum $\lambda^{1/p}$ of the nontrivial spectrum of

$$(v|f'|^{p-1} \text{sgn}(f'))' = -\lambda u|f|^{q-1} \text{sgn}(f)$$

with boundary condition $f(-M) = 0 = f(N)$ and $f'(-M) = 0 = f'(N)$ (when $M, N < \infty$), respectively. The word “bilateral” in the title means that a same boundary condition is endowed at two endpoints of the interval. The spectral point of view has played a crucial role in our previous study. For instance, it appears in each of the papers [6–11] and [15].

To study (2.1), we start again at the splitting technique. We begin with the easier case: the lower estimate. It is indeed easier than the one studied in the last section.

Lemma 2.1 *Let $1 \leq p \leq q < \infty$. Then we have*

$$A \geq 2^{1/q-1/p} \sup_{\theta \in [-M, N]} (A_{\theta}^{-} \wedge A_{\theta}^{+}).$$

Proof Fix $\theta \in [-M, N]$. Let f_- satisfy $f_-|_{[\theta, N]} = 0$,

$$\int_{-M}^{\theta} |f_-|^q d\mu = 1, \quad \text{and} \quad \left(\int_{-M}^{\theta} |f'_-|^p d\nu \right)^{1/p} < (A_{\theta}^-)^{-1} + \varepsilon.$$

Let f_+ satisfy $f_+|_{[-M, \theta]} = 0$,

$$\int_{\theta}^N |f_+|^q d\mu = 1, \quad \text{and} \quad \left(\int_{\theta}^N |f'_+|^p d\nu \right)^{1/p} < (A_{\theta}^+)^{-1} + \varepsilon.$$

Set $f = cf_- + f_+$, where $c = -\pi(f_+)/\pi(f_-)$. Then $\pi(f) = 0$,

$$1 + |c|^q = \int_{-M}^{\theta} |cf_-|^q d\mu + \int_{\theta}^N |f_+|^q d\mu = \int_{-M}^N |f|^q d\mu,$$

and

$$\begin{aligned} \int_{-M}^N |f'|^p d\nu &= |c|^p \int_{-M}^{\theta} |f'_-|^p d\nu + \int_{\theta}^N |f'_+|^p d\nu \\ &\leq |c|^p ((A_{\theta}^-)^{-1} + \varepsilon)^p + ((A_{\theta}^+)^{-1} + \varepsilon)^p \\ &\leq ((A_{\theta}^-)^{-1} \vee (A_{\theta}^+)^{-1} + \varepsilon)^p (1 + |c|^p). \end{aligned}$$

Hence

$$\begin{aligned} \left(\int_{-M}^N |f'|^p d\nu \right)^{1/p} &\leq ((A_{\theta}^-)^{-1} \vee (A_{\theta}^+)^{-1} + \varepsilon) (1 + |c|^p)^{1/p} \\ &\leq 2^{1/p-1/q} ((A_{\theta}^-)^{-1} \vee (A_{\theta}^+)^{-1} + \varepsilon) (1 + |c|^q)^{1/q} \\ &\quad \text{(by Jensen's inequality requiring } q \geq p) \\ &= 2^{1/p-1/q} ((A_{\theta}^-)^{-1} \vee (A_{\theta}^+)^{-1} + \varepsilon) \left(\int_{-M}^N |f|^q d\mu \right)^{1/q}. \end{aligned}$$

Thus

$$A \geq 2^{1/q-1/p} (A_{\theta}^- \wedge A_{\theta}^+).$$

Since θ is arbitrary, we obtain the lower bound of A . □

The upper bound of A is harder than the lower one just studied. But the first step is still easy. Given f and $\theta \in (-M, N)$, let $\tilde{f} = f - f(\theta)$. Then

$$\begin{aligned} \int_{-M}^N |f'|^p d\nu &= \int_{-M}^{\theta} |\tilde{f}'|^p d\nu + \int_{\theta}^N |\tilde{f}'|^p d\nu \\ &\geq (A_{\theta}^-)^{-p} \left(\int_{-M}^{\theta} |\tilde{f}|^q d\mu \right)^{p/q} + (A_{\theta}^+)^{-p} \left(\int_{\theta}^N |\tilde{f}|^q d\mu \right)^{p/q} \\ &\geq [(A_{\theta}^-)^{-p} \wedge (A_{\theta}^+)^{-p}] \left[\left(\int_{-M}^{\theta} |\tilde{f}|^q d\mu \right)^{p/q} + \left(\int_{\theta}^N |\tilde{f}|^q d\mu \right)^{p/q} \right] \\ &\geq (2^{(p/q-1) \vee 0})^{-1} [(A_{\theta}^-)^{-p} \wedge (A_{\theta}^+)^{-p}] \left(\int_{-M}^N |\tilde{f}|^q d\mu \right)^{p/q} \quad (\text{by } c_r\text{-inequality}). \end{aligned}$$

Our aim is to replace $|\tilde{f}|^q$ on the right-hand side with $|f - \pi(f)|^q$. This is true in the case of $q = 2$ since

$$\inf_{c \in \mathbb{R}} \int_{-M}^N (f - c)^2 d\mu = \int_{-M}^N (f - \pi(f))^2 d\mu.$$

Unfortunately, it does not work for general q . Anyhow, when $q = 2$, we have

$$\int_{-M}^N |f'|^p d\nu \geq (2^{(p/2-1)\vee 0})^{-1} [(A_\theta^-)^{-p} \wedge (A_\theta^+)^{-p}] \left(\int_{-M}^N |f - \pi(f)|^2 d\mu \right)^{p/2}.$$

Since θ is arbitrary, we obtain

$$\begin{aligned} (2^{(1/2-1/p)\vee 0}) \left(\int_{-M}^N |f'|^p d\nu \right)^{1/p} &\geq \sup_{\theta \in (-M, N)} [(A_\theta^-)^{-1} \wedge (A_\theta^+)^{-1}] \\ &\quad \times \left(\int_{-M}^N |f - \pi(f)|^2 d\mu \right)^{1/2}. \end{aligned}$$

Next, since f is arbitrary, it follows that

$$A \leq 2^{(1/2-1/p)\vee 0} \inf_{\theta \in (-M, N)} (A_\theta^- \vee A_\theta^+).$$

Up to now, the proof is similar to [9, Theorems 3.3 and 3.4] in the specific case that $q = 2$. For general $q \geq 2$, we have luckily a different approach (cf. [10, Chapter 6] and references therein). Note that we have already proved that if the measure μ is replaced by $\mu_g := g\mu$ for a nonnegative function g on $[-M, N]$, then the optimal constant A_g in the inequality

$$\left(\int_{-M}^N |f - \pi(f)|^2 d\mu_g \right)^{1/2} \leq A_g \left(\int_{-M}^N |f'|^p d\nu \right)^{1/p}$$

obeys

$$A_g \leq 2^{(1/2-1/p)\vee 0} \inf_{\theta \in (-M, N)} (A_g^{\theta, -} \vee A_g^{\theta, +}),$$

where $A_g^{\theta, \pm}$ is obtained from $A_\theta^\pm$ replacing μ with μ_g . From now on in this proof, the constants A_g , $A_g^{\theta, \pm}$, and $B_g^{\theta, \pm}$ are used for μ_g in the specified case that $q = 2$ only. Note that for $A_g^{\theta, +}$ for instance, the function g can be replaced by $g\mathbb{1}_{[\theta, N]}$.

Even though we are now mainly working on the L^q -case to which $\mathcal{G}$ is the set of functions in the unit ball of $L^{\frac{q}{q-2}}(\mu)$ (where $\frac{q}{q-2}$ is the conjugate number of $q/2$):

$$\left(\int_{-M}^N |f - \pi(f)|^q d\mu \right)^{2/q} = \sup_{g \in \mathcal{G}} \int_{-M}^N |f - \pi(f)|^2 g d\mu = \sup_{g \in \mathcal{G}} \int_{-M}^N |f - \pi(f)|^2 d\mu_g,$$

at the moment, we allow $\mathcal{G}$ to be general in the setup of Hypotheses 1.10:

$$\|(f - \pi(f))^2\|_{\mathbb{B}}^{1/2} \leq A_{\mathbb{B}} \|f'\|_{\nu, p}. \quad (2.2)$$

We have

$$\begin{aligned} A_{\mathbb{B}} &= \sup_{g \in \mathcal{G}} A_g \leq \sup_{g \in \mathcal{G}} \inf_{\theta \in (-M, N)} (A_g^{\theta, -} \vee A_g^{\theta, +}) \\ &\leq \inf_{\theta \in (-M, N)} \left[\left(\sup_{g \in \mathcal{G}_\theta^-} A_g^{\theta, -} \right) \vee \left(\sup_{g \in \mathcal{G}_\theta^+} A_g^{\theta, +} \right) \right], \end{aligned}$$

where

$$\mathcal{G}_\theta^- = \{g|_{[-M, \theta]} : g \in \mathcal{G}\}, \quad \mathcal{G}_\theta^+ = \{g|_{[\theta, N]} : g \in \mathcal{G}\}.$$

Here is a technical point. Because on the left-hand side of (2.2), we start at $q = 2$. This leads to the restriction that $p \in (1, 2]$ since we need $q \geq p$ in order to use the basic estimates in terms of $B_\theta^\pm$ given in the next proof. Anyhow, we have proved the first assertion of the next result.

Lemma 2.2 *Let $\mathcal{G}$ satisfy Hypotheses 1.10. Then for $p \in (1, 2]$, we have*

$$A_{\mathbb{B}} \leq \inf_{\theta \in (-M, N)} \left[\left(\sup_{g \in \mathcal{G}_{\theta}^{-}} A_g^{\theta, -} \right) \vee \left(\sup_{g \in \mathcal{G}_{\theta}^{+}} A_g^{\theta, +} \right) \right].$$

Moreover,

$$\sup_{g \in \mathcal{G}_{\theta}^{\pm}} A_g^{\theta, \pm} \leq k_{2,p} B_{\mathbb{B}}^{\theta, \pm},$$

where

$$\begin{aligned} B_{\mathbb{B}}^{\theta, +} &= \sup_{r \in (\theta, N)} \left\| \mathbb{1}_{[r, N]} \right\|_{\mathbb{B}}^{1/2} \hat{\nu}[\theta, r]^{(p-1)/p}, \\ B_{\mathbb{B}}^{\theta, -} &= \sup_{r \in (-M, \theta)} \left\| \mathbb{1}_{[-M, r]} \right\|_{\mathbb{B}}^{1/2} \hat{\nu}[r, \theta]^{(p-1)/p}. \end{aligned}$$

Proof By [5, Theorem 1.14] and [4, §1.3, Theorem 1] (see also [12, Theorem 8] and [13, Theorem 2] in which the factor $k_{q,p}$ may be different), we have for general $1 < p \leq q < \infty$ that

$$\begin{aligned} B_{\theta}^{\pm} &\leq A_{\theta}^{\pm} \leq k_{q,p} B_{\theta}^{\pm}, \quad 1 < p \leq q < \infty, \\ B_{\theta}^{+} &= \sup_{r \in (\theta, N)} \mu[r, N]^{1/q} \hat{\nu}[\theta, r]^{(p-1)/p}, \\ B_{\theta}^{-} &= \sup_{r \in (-M, \theta)} \mu[-M, r]^{1/q} \hat{\nu}[r, \theta]^{(p-1)/p}, \end{aligned}$$

where $\hat{\nu}$ is the same as in the last section. It remains to estimate $\sup_{g \in \mathcal{G}_{\theta}^{+}} A_g^{\theta, +}$ for instance. First, we have for $q = 2$ that

$$\sup_{g \in \mathcal{G}_{\theta}^{+}} A_g^{\theta, +} \leq k_{2,p} \sup_{g \in \mathcal{G}_{\theta}^{+}} B_g^{\theta, +}.$$

Next, we have

$$\begin{aligned} \sup_{g \in \mathcal{G}_{\theta}^{+}} B_g^{\theta, +} &= \sup_{g \in \mathcal{G}_{\theta}^{+}} \sup_{r \in (\theta, N)} \mu_g[r, N]^{1/2} \hat{\nu}[\theta, r]^{(p-1)/p} \\ &= \sup_{r \in (\theta, N)} \left(\sup_{g \in \mathcal{G}_{\theta}^{+}} \mu_g[r, N] \right)^{1/2} \hat{\nu}[\theta, r]^{(p-1)/p} \\ &= B_{\mathbb{B}}^{\theta, +}. \end{aligned}$$

Similar computation holds for $\sup_{g \in \mathcal{G}_{\theta}^{-}} B_g^{\theta, -}$. Combining these facts with the first assertion gives us the second one of the lemma. $\square$

As mentioned in the last section, without loss of generality, we can assume that μ is positive on each subinterval.

Here is our upper estimate.

Lemma 2.3 *Let $\mu_{\text{pp}} = 0$ and $\mathcal{G}$ satisfy Hypotheses 1.10. Then for $p \in (1, 2]$, we have*

$$A_{\mathbb{B}} \leq k_{2,p} B_{\mathbb{B}}^{*},$$

where the constant $B_{\mathbb{B}}^{*}$ is defined by

$$B_{\mathbb{B}}^{*- \frac{p}{p-1}} = \inf_{x < y} \left[\left\| \mathbb{1}_{[-M, x]} \right\|_{\mathbb{B}}^{- \frac{p}{2(p-1)}} + \left\| \mathbb{1}_{[y, N]} \right\|_{\mathbb{B}}^{- \frac{p}{2(p-1)}} \right] \hat{\nu}[x, y]^{-1}.$$

Proof Write

$$B_{\mathbb{B}}^{*\frac{p}{p-1}} = \sup_{x < y} \frac{\hat{\nu}[x, \theta] + \hat{\nu}(\theta, y)}{\|\mathbb{1}_{[-M, x]}\|_{\mathbb{B}}^{-\frac{p}{2(p-1)}} + \|\mathbb{1}_{[y, N]}\|_{\mathbb{B}}^{-\frac{p}{2(p-1)}}}.$$

Similar to the proof of Lemma 1.3, the assertion follows by using Lemmas 1.2 and 2.2. Here we may need the approximating procedure by finite M and N . $\square$

The following result is on the lower estimate of A . Its proof is new even in the special case that $p = q = 2$. Note that $\mu[x, y] = \mu(x, y)$ whenever $\mu_{pp} = 0$.

Lemma 2.4 *Let $\mu_{pp} = 0$. Then for $1 < p, q < \infty$, the optimal constant A in (2.1) satisfies*

$$\begin{aligned} A &\geq \sup_{-M \leq x < y \leq N} \left\{ [\mu[-M, x]^{\frac{1}{1-q}} + \mu[y, N]^{\frac{1}{1-q}}]^{\frac{1-q}{q}} \hat{\nu}[x, y]^{\frac{p-1}{p}} \right\} \\ &=: B_*. \end{aligned} \quad (2.3)$$

Proof Given $m, n \in (-M, N)$ with $m < n$, let $\bar{\theta} = \bar{\theta}(m, n)$ be the unique solution to the equation

$$\mu[-M, m] \hat{\nu}[m, \theta] + \int_m^\theta \mu(dx) \hat{\nu}[x, \theta] = \mu[n, N] \hat{\nu}[\theta, n] + \int_\theta^n \mu(dx) \hat{\nu}[\theta, x], \quad \theta \in (m, n).$$

The existence of the solution is clear since μ is continuous, when θ varies from m to n , the left-hand side goes from 0 to a positive number and the right-hand side goes from a positive number to zero. Next, define

$$f(x) = -\mathbb{1}_{\{x \leq \bar{\theta}\}} \hat{\nu}[m \vee x, \bar{\theta}] + \mathbb{1}_{\{x > \bar{\theta}\}} \hat{\nu}[\bar{\theta}, n \wedge x], \quad x \in [-M, N].$$

Then $\mu(f) = 0$ by definition of $\bar{\theta}$. Clearly, f is absolutely continuous. On the one hand, we have

$$\begin{aligned} \left(\int_{-M}^N |f'|^p d\nu \right)^{1/p} &= \left(\int_m^{\bar{\theta}} h^p d\nu + \int_{\bar{\theta}}^n h^p d\nu \right)^{1/p} \\ &= (\hat{\nu}[m, \bar{\theta}] + \hat{\nu}[\bar{\theta}, n])^{1/p} \\ &= \hat{\nu}[m, n]^{1/p}. \end{aligned} \quad (2.4)$$

Here in the second step, we have once again ignored the singular part of ν . On the other hand, we have

$$\begin{aligned} \int_{-M}^N |f - \pi(f)|^q d\mu &= \int_{-M}^N |f|^q d\mu \\ &> \int_{-M}^m |f|^q d\mu + \int_n^N |f|^q d\mu \\ &= \mu[-M, m] \hat{\nu}[m, \bar{\theta}]^q + \mu[n, N] \hat{\nu}[\bar{\theta}, n]^q. \end{aligned} \quad (2.5)$$

Now we have naturally, as in proof (a) of Proposition 1.1, that

$$\begin{aligned} \text{RHS of (2.5)} &\geq (\mu[-M, m] \wedge \mu[n, N]) (\hat{\nu}[m, \bar{\theta}]^q + \hat{\nu}[\bar{\theta}, n]^q) \\ &\geq 2^{1-q} (\mu[-M, m] \wedge \mu[n, N]) \hat{\nu}[m, n]^q. \end{aligned}$$

However, such a lower bound is quite rough for our purpose so we need a different approach. Note that the function

$$\gamma(x) = \alpha x^q + \beta(1-x)^q, \quad x \in (0, 1), \quad \alpha > 0, \beta > 0, \quad q \in (1, \infty)$$

achieves its minimum

$$\left(\alpha^{\frac{1}{1-q}} + \beta^{\frac{1}{1-q}}\right)^{1-q} \quad (\text{resp., } \alpha \wedge \beta \text{ in the case of } q = 1)$$

at

$$x^* = \left[1 + \left(\frac{\alpha}{\beta}\right)^{\frac{1}{q-1}}\right]^{-1} = \beta^{\frac{1}{q-1}} [\alpha^{\frac{1}{q-1}} + \beta^{\frac{1}{q-1}}]^{-1} \in (0, 1).$$

Applying this result with

$$\alpha = \mu[-M, m], \quad \beta = \mu[n, N], \quad x = \hat{\nu}[m, \bar{\theta}] / \hat{\nu}[m, n]$$

to (2.5), we get

$$\left(\int_{-M}^N |f - \pi(f)|^q d\mu\right)^{1/q} \geq \left\{\mu[-M, m]^{\frac{1}{1-q}} + \mu[n, N]^{\frac{1}{1-q}}\right\}^{\frac{1-q}{q}} \hat{\nu}[m, n].$$

Because

$$A \geq \left(\int_{-M}^N |f - \pi(f)|^q d\mu\right)^{1/q} \left(\int_{-M}^N |f'|^p d\nu\right)^{-1/p},$$

the estimate given in the lemma now follows immediately. $\square$

Now, one may ask the possibility using the idea in the last part of the proof above to improve the estimate produced by proof (a) of Proposition 1.1. The answer is yes if $p > q$ and no if $p \leq q$. Note that here we have power $q > 1$ and in proof (a) of Proposition 1.1, the power is p/q . Thus, if $p > q$, we can follow the proof here to have an improvement. However, in this paper, we are mainly interested in the case that $p \leq q$. Then the function $\alpha x^\gamma + \beta(1-x)^\gamma$ ($\gamma \leq 1$) is concave, its minimum is achieved at the boundaries: either at $x = 0$ or at $x = 1$. That is, $\min_{x \in (0,1)} \{\alpha x^\gamma + \beta(1-x)^\gamma\} = \alpha \wedge \beta$. In this case, we have thus returned to the original result given in proof (a) of Proposition 1.1. We mention that this remark is also meaningful for the first step of the proof of Lemma 2.2 given right below the proof of Lemma 2.1.

As an analog of Lemma 1.5, we have the following result.

Lemma 2.5 *Let $q \geq p$. Then we have $B_* \leq B^* \leq 2^{1/p-1/q} B_*$, where the constant B^* is defined by*

$$B^*{}^{\frac{p}{1-p}} = \inf_{x < y} \left\{ \mu[-M, x]^{\frac{p}{(1-p)q}} + \mu[y, N]^{\frac{p}{(1-p)q}} \right\} \hat{\nu}[x, y]^{-1}. \quad (2.6)$$

Proof Applying $\mathbb{B}$ to $L^{\frac{q}{q-2}}(\mu)$, the constant $B_{\mathbb{B}}^*$ given in Lemma 2.3 is reduced to B^* defined by (2.6).

(a) Part $B^* \geq B_*$ follows from the c_r -inequality by setting

$$\alpha = \mu[-M, x]^{\frac{p}{(1-p)q}}, \quad \beta = \mu[y, N]^{\frac{p}{(1-p)q}}, \quad r = \frac{(p-1)q}{p(q-1)} \in (0, 1].$$

(b) Part $B_* \geq 2^{1/q-1/p} B^*$ follows from the inequality by setting

$$\alpha = \mu[-M, x]^{\frac{1}{1-q}}, \quad \beta = \mu[y, N]^{\frac{1}{1-q}}, \quad r = \frac{p(q-1)}{(p-1)q} \geq 1. \quad \square$$

As a combination of Lemmas 2.3–2.5, we obtain the following result.

Theorem 2.6 *Let $\mu[-M, N] < \infty$ and $\mu_{pp} = 0$. Then*

(1) *for $1 < p \leq 2 \leq q < \infty$, the optimal constant A in (2.1) satisfies*

$$A \leq k_{2,p} B^*, \text{ where } k_{q,p} \text{ is defined by (1.5), and}$$

(2) *for $1 < p, q < \infty$, we have $A \geq B_*$,*

where B_* and B^* are defined in Lemmas 2.4 and (2.6), respectively. Moreover, we have $B_* \leq B^* \leq 2^{1/p-1/q} B_*$ once $q \geq p$.

Proof As an application of Lemma 2.3, we get the upper estimate of A . The lower estimate of A is due to Lemma 2.4. The comparison of B^* and B_* comes from Lemma 2.5. $\square$

When $p = q = 2$, from Theorem 2.6, it follows that

$$B^* \leq A \leq 2B^*.$$

We have thus returned to [6, Theorem 10.2]. It is interesting that in the special case of $p = q = 2$, the duality given in [3, p. 13 and (1.17)] coincides with that used in [6, 7]. The former duality exchanges the (single-side but not bilateral) boundary conditions $f(-M) = 0$ and $f(N) = 0$. This is clearly different from a dual of (1.1) and (2.1). To prove the last duality, in [6, 7], several techniques were adopted: coupling, duality, and capacity. Thus, the proofs given here are essentially different from that presented in [6, 7], much direct and elementary. Besides, it is unclear how these advanced techniques can be applied to the present setup.

In parallel to the last section, define $y(x)$ to be the solution to the equation $\mu[-M, x] = \mu[y, N]$ and denote by $m(\mu)$ to be the median of μ . Set

$$H_{\mu,\nu}(x, y) = [\mu[-M, x]^{\frac{1}{1-q}} + \mu[y, N]^{\frac{1}{1-q}}]^{\frac{1-q}{q}} \hat{\nu}[x, y]^{\frac{p-1}{p}}.$$

Define

$$H^o = 2^{1/q-1} \sup_{x \in (-M, m(\mu)]} \mu[-M, x]^{1/q} \hat{\nu}[x, y(x)]^{(p-1)/p}. \quad (2.7)$$

Denote by Γ be the limiting points of $H_{\mu,\nu}(x, y)$ as

$$\hat{\nu}[-M, x] = \infty \quad \text{or} \quad \hat{\nu}[y, N] = \infty,$$

as well as the iterated limits if $\hat{\nu}[-M, N] = \infty$ when $M = \infty = N$. Set

$$H^\partial = \begin{cases} \sup\{\gamma : \gamma \in \Gamma\}, & \text{if } \Gamma \neq \emptyset, \\ 0, & \text{if } \Gamma = \emptyset. \end{cases} \quad (2.8)$$

Similar to Lemmas 1.7 and 1.8, we have the following result.

Lemma 2.7 *Let $\mu[-M, N] < \infty$ and $\mu_{pp} = 0$. Define H^o and H^∂ as above. Then we have*

$$H^o \vee H^\partial \leq B_* \leq (2^{1-1/q} H^o) \vee H^\partial$$

and

$$(2^{1/p-1/q} H^o) \vee H^\partial \leq B^* \leq (2^{1-1/q} H^o) \vee H^\partial.$$

Proof Consider B_* for instance. Recalling that $\mu[-M, x] = \mu[y(x), N]$, we have

$$\begin{aligned} & \sup_{-M \leq x < y \leq N} \left\{ \mu[-M, x]^{\frac{1}{1-q}} + \mu[y, N]^{\frac{1}{1-q}} \right\}^{\frac{1-q}{q}} \hat{\nu}[x, y]^{\frac{p-1}{p}} \\ & \geq 2^{1/q-1} \sup_{-M \leq x \leq m(\mu)} \mu[-M, x]^{1/q} \hat{\nu}[x, y(x)]^{(p-1)/p} \\ & = H^o. \end{aligned}$$

This plus the boundary condition gives us the lower estimate of B_* . The proofs for the other assertions are similar. $\square$

Corollary 2.8 *The Hardy-type inequality (2.1) holds iff $H^o \vee H^\partial < \infty$.*

To generalize Theorem 2.6 to a general normed linear space $\mathbb{B}$, as in the study of the upper estimate $B_{\mathbb{B}}^*$, a natural way is starting from B_{g*} :

$$\sup_g B_{g*} = \sup_g \sup_{x < y} \left\{ (\mu_g[-M, x]^{\frac{1}{1-q}} + \mu_g[y, N]^{\frac{1}{1-q}})^{\frac{1-q}{q}} \hat{\nu}[x, y]^{\frac{p-1}{p}} \right\}.$$

Then it is not clear how to handle with this expression in terms of the norm $\mathbb{B}$. A crucial point here is that the measure μ appears in the last expression twice rather than a single term in the last section. The next result is an extension and improvement of the basic estimates given in [15, Theorem 2.2].

Theorem 2.9 *Let $\mu[-M, N] < \infty$, $\mu_{pp} = 0$, and $\mathcal{G}$ satisfy Hypotheses 1.10. Then*

(1) *for $p \in (1, 2]$, the optimal constant $A_{\mathbb{B}}$ in the inequality*

$$\| (f - \pi(f))^2 \|_{\mathbb{B}}^{1/2} \leq A_{\mathbb{B}} \|f'\|_{\nu, p}$$

satisfies $A_{\mathbb{B}} \leq k_{2,p} B_{\mathbb{B}}^*$, where the constant $B_{\mathbb{B}}^*$ is defined by

$$B_{\mathbb{B}}^{*- \frac{p}{p-1}} = \inf_{-M < x < y < N} \left\{ \|\mathbb{1}_{[-M, x]}\|_{\mathbb{B}}^{-\frac{p}{2(p-1)}} + \|\mathbb{1}_{[y, N]}\|_{\mathbb{B}}^{-\frac{p}{2(p-1)}} \right\} \hat{\nu}[x, y]^{-1}.$$

(2) *For $1 < p, q < \infty$, the optimal constant $A_{\mathbb{B}}$ in the inequality*

$$\| |f - \pi(f)|^q \|_{\mathbb{B}}^{1/q} \leq A_{\mathbb{B}} \|f'\|_{\nu, p}$$

satisfies $A_{\mathbb{B}} \geq B_{\mathbb{B}*}$, where

$$B_{\mathbb{B}*} = \sup_{-M < x < y < N} \gamma_{\mathbb{B}}(x, y; q) \hat{\nu}[x, y]^{\frac{p-1}{p}}$$

and

$$\gamma_{\mathbb{B}}(x, y; q) = \inf_{z \in (0, 1)} \left\| \mathbb{1}_{[-M, x]} z^q + \mathbb{1}_{[y, N]} (1 - z)^q \right\|_{\mathbb{B}}^{1/q}.$$

Proof The first assertion is a copy of Lemma 2.3. To prove the second assertion, we return to the construction used in the proof of Lemma 2.4. That is, we use the notation $\bar{\theta}$ and f introduced there. First, we have

$$\begin{aligned} \| |f - \pi(f)|^q \|_{\mathbb{B}} &= \| |f|^q \|_{\mathbb{B}} \\ &\geq \| |f|^q \mathbb{1}_{[-M, m]} + |f|^q \mathbb{1}_{[n, N]} \|_{\mathbb{B}} \\ &= \| \mathbb{1}_{[-M, m]} \hat{\nu}[m, \bar{\theta}]^q + \mathbb{1}_{[n, N]} \hat{\nu}[\bar{\theta}, n]^q \|_{\mathbb{B}} \\ &\geq \gamma_{\mathbb{B}}(m, n; q)^q \hat{\nu}[m, n]^q. \end{aligned}$$

Combining this with (2.4), we obtain

$$\frac{\| |f - \pi(f)|^q \|_{\mathbb{B}}^{1/q}}{\|f'\|_{\nu, p}} \geq \gamma_{\mathbb{B}}(m, n; q) \hat{\nu}[m, n]^{\frac{p-1}{p}}.$$

Now the required assertion follows by making supremum with respect to (x, y) with $x < y$. $\square$

Example 2.10 Let $\mu(dx) = \nu(dx) = e^{-bx} dx$ ($b > 0$) on $(0, \infty)$. Then the inequality (2.1) does not hold if $q > p$. When $q = p$, we have

$$B^* = B_* = H^\partial = \frac{1}{b}(p-1)^{1-\frac{1}{p}}.$$

The upper estimate $2B^*$ in Theorem 2.6 (1) is sharp in the case of $p = q = 2$, refer to [7, Example 5.3].

Proof We have

$$\mu(0, x) = \frac{1}{b}(1 - e^{-bx}), \quad \mu(y, \infty) = \frac{1}{b}e^{-by}, \quad y(x) = -\frac{1}{b} \log(1 - e^{-bx}), \quad m(\mu) = \frac{1}{b} \log 2.$$

Next,

$$h(x) = e^{\frac{bx}{p-1}}, \quad \hat{\nu}[x, y] = \frac{p-1}{b}(e^{\frac{by}{p-1}} - e^{\frac{bx}{p-1}}).$$

Thus,

$$H_{\mu, \nu}(x, y) = b^{-\frac{1}{q}} \left(\frac{p-1}{b} \right)^{1-\frac{1}{p}} [(1 - e^{-bx})^{\frac{1}{1-q}} + e^{-\frac{by}{1-q}}]^{\frac{1}{q}-1} (e^{\frac{by}{p-1}} - e^{\frac{bx}{p-1}})^{1-\frac{1}{p}}.$$

Hence

$$\begin{aligned} H^\partial &= \overline{\lim}_{y \rightarrow \infty} H_{\mu, \nu}(x, y) = b^{-\frac{1}{q}} \left(\frac{p-1}{b} \right)^{1-\frac{1}{p}} \lim_{y \rightarrow \infty} e^{by(\frac{1}{p}-\frac{1}{q})} \\ &= \begin{cases} b^{-1}(p-1)^{1-\frac{1}{p}}, & \text{if } q = p, \\ \infty, & \text{if } q > p. \end{cases} \end{aligned}$$

Thus, the inequality (2.1) does not hold if $q > p$. Next, assume that $q = p$. Then we have

$$\begin{aligned} H^o &= 2^{\frac{1}{p}-1} b^{-\frac{1}{p}} \left(\frac{p-1}{b} \right)^{1-\frac{1}{p}} \sup_{x \in (0, b^{-1} \log 2)} (1 - e^{-bx})^{\frac{1}{p}} ((1 - e^{-bx})^{-\frac{1}{p-1}} - e^{\frac{bx}{p-1}})^{1-\frac{1}{p}} \\ &= 2^{\frac{1}{p}-1} b^{-\frac{1}{p}} \left(\frac{p-1}{b} \right)^{1-\frac{1}{p}} \sup_{x \in (0, \log 2)} (1 - e^{-x})^{\frac{1}{p}} ((1 - e^{-x})^{-\frac{1}{p-1}} - e^{\frac{x}{p-1}})^{1-\frac{1}{p}}. \end{aligned}$$

To compute H^o , we observe that

$$\overline{\lim}_{x \rightarrow 0} (1 - e^{-x})^{\frac{1}{p}} ((1 - e^{-x})^{-\frac{1}{p-1}} - e^{\frac{x}{p-1}})^{1-\frac{1}{p}} = 1.$$

Because of this and the decreasing property of the function on the left-hand side, we obtain

$$H^o = 2^{\frac{1}{p}-1} b^{-\frac{1}{p}} \left(\frac{p-1}{b} \right)^{1-\frac{1}{p}} = 2^{\frac{1}{p}-1} \frac{1}{b} (p-1)^{1-\frac{1}{p}}.$$

Having H^∂ and H^o at hand, it is easy to compute B^* and B_* . $\square$

Example 2.11 Let $\mu(dx) = x^{-2} dx$ and $\nu(dx) = dx$ on $(1, \infty)$. Then the inequality (2.1) does not hold if $\frac{1}{p} + \frac{1}{q} < 1$. Otherwise, if $\frac{1}{p} + \frac{1}{q} = 1$, then $B^* = B_* = H^\partial = 1$. In particular,

when $p = q = 2$, our upper estimate $2B^*$ in Theorem 2.6(1) is exact since $A = 2$ (cf. [7, Example 5.4]). If $\frac{1}{p} + \frac{1}{q} > 1$, then $H^o \leq B_* \leq 2^{1-1/q} H^o$ and $2^{1/p-1/q} H^o \leq B^* \leq 2^{1-1/q} H^o$:

$$H^o = 2^{\frac{1}{q}-1} \left[\frac{2\beta}{\sqrt{\alpha^2 - 6\alpha\beta + \beta^2} + \alpha - \beta} + 1 \right]^\alpha \left[2 - \frac{\sqrt{\alpha^2 - 6\alpha\beta + \beta^2} + \alpha + \beta}{2\beta} \right]^\beta,$$

where $\alpha = 1 - \frac{1}{p} - \frac{1}{q} < 0$ and $\beta = 1 - \frac{1}{p} > 0$. In this case, $2^{1/p-1} \leq B^*/B_* \leq 2^{1-1/q}$. For fixed $p = 5/4$, when q varies over $[2, 4.25]$, the curves of H^o , B_* , B^* , and $2^{3/2-1/p} B^*$ are given in Figure 2. The ratio of the upper and lower bounds is increasing in q but no more than 2.

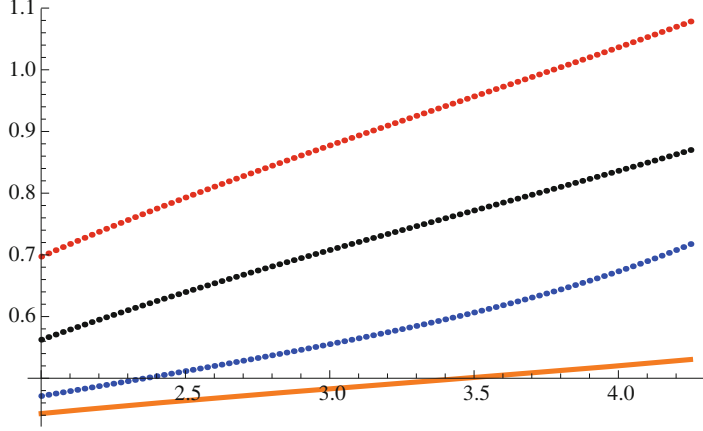


Figure 2 The curves from bottom to top are H^o , B_* , B^* , and $k_{2,p} B^*$, respectively

Proof Note that $h \equiv 1$. We have

$$\mu(1, x) = \int_1^x \frac{1}{z^2} = \frac{x-1}{x}, \quad \mu(y, \infty) = \int_y^\infty \frac{1}{z^2} = \frac{1}{y}, \quad y(x) = \frac{x}{x-1}, \quad m(\mu) = 2.$$

Then

$$H_{\mu,\nu}(x, y) = \left[\left(\frac{x-1}{x} \right)^{1/(1-q)} + y^{1/(q-1)} \right]^{1/q-1} (y-x)^{(p-1)/p}.$$

Thus,

$$H^\partial = \overline{\lim}_{y \rightarrow \infty} H_{\mu,\nu}(x, y) = \begin{cases} 1, & \text{if } \frac{1}{p} + \frac{1}{q} = 1, \\ \infty, & \text{if } \frac{1}{p} + \frac{1}{q} < 1, \\ 0, & \text{if } \frac{1}{p} + \frac{1}{q} > 1. \end{cases}$$

Hence the inequality does not hold if $\frac{1}{p} + \frac{1}{q} < 1$. When $p = 2$, this means that the inequality does not hold whenever $q > 2$. The assertion is known as a sharp result for Nash inequality (cf. [10, Table 8.2]). Next, we have

$$\begin{aligned} H^o &= 2^{1/q-1} \sup_{x \in (1, 2]} \left(\frac{x-1}{x} \right)^{1/q} \left(\frac{x}{x-1} - x \right)^{(p-1)/p} \\ &= 2^{1/q-1} \sup_{x \in (1, 2]} \left(1 + \frac{1}{x-1} \right)^{1-\frac{1}{p}-\frac{1}{q}} (2-x)^{1-\frac{1}{p}}. \end{aligned}$$

When $\frac{1}{p} + \frac{1}{q} = 1$, we have $H^\circ = 2^{1/q-1}$, and so $B^* = B_* = H^\partial = 1$. When $\frac{1}{p} + \frac{1}{q} > 1$, we have

$$B^* = \sup_{x < y} \left[\left(1 - \frac{1}{x} \right)^{\frac{p}{(1-p)q}} + y^{\frac{p}{(p-1)q}} \right]^{\frac{1}{p}-1} (y-x)^{1-\frac{1}{p}},$$

$$B_* = \sup_{x < y} \left[\left(1 - \frac{1}{x} \right)^{\frac{1}{1-q}} + y^{\frac{1}{q-1}} \right]^{\frac{1}{q}-1} (y-x)^{1-\frac{1}{p}}.$$

H° has an explicit expression as shown above. $\square$

3 Nash Inequality, Sobolev-type Inequality, and Logarithmic Sobolev Inequality

In this section, we study first the Nash inequality and its closely related Sobolev-type inequality, as a typical application of Theorem 2.6. Then we study the logarithmic Sobolev inequality by the similar method introduced in the paper.

3.1 Nash Inequality and Sobolev-type Inequality

Recall that the probability measure π is defined by $\mu/\mu[-M, N]$. Consider the Nash inequality:

$$\|f - \pi(f)\|_{\mu, 2}^{2+4/\gamma} \leq A_N \|f'\|_{\nu, 2}^2 \|f\|_{\mu, 1}^{4/\gamma}, \quad \gamma > 0. \quad (3.1)$$

It seems more symmetric to replace $\|f\|_{\mu, 1}$ by $\|f - \pi(f)\|_{\mu, 1}$ on the right-hand side of (3.1). Let us denote the latter one by (3.1)'. If (3.1) holds for every absolutely continuous $f \in L^2(\mu)$, then so does (3.1)' regarding $f - \pi(f)$ as a new f . Conversely, assume (3.1)'. Since $\|f - \pi(f)\|_{\mu, 1} \leq 2\|f\|_{\mu, 1}$, we certainly have (3.1). Hence (3.1) and (3.1)' are equivalent.

It is known (cf. [16], [10, §4.8 and §6.5] for related results and more references) that this inequality, when $\gamma > 2$, is equivalent to the Sobolev-type inequality

$$\|f - \pi(f)\|_{\mu, 2\gamma/(\gamma-2)}^2 \leq A_S \|f'\|_{\nu, 2}^2. \quad (3.2)$$

Now, as an application of Theorem 2.6, we have the following result.

Theorem 3.1 *Let $\mu[-M, N] < \infty$ and $\mu_{pp} = 0$. Then*

(1) *when $\gamma > 2$, the Nash inequality (3.1) or equivalently, the Sobolev-type inequality (3.2) holds iff $H^\circ \vee H^\partial < \infty$, where H° and H^∂ are defined by (2.7) and (2.8), respectively, with $p = 2$ and $q = 2\gamma/(\gamma - 2)$. Furthermore, we have*

$$B_{S*} \leq A_S \leq 4B_{S*}^*,$$

where

$$B_S^* = \sup_{x < y} [\mu(-M, x)^{\frac{2}{\gamma}-1} + \mu(y, N)^{\frac{2}{\gamma}-1}]^{-1} \hat{\nu}[x, y],$$

$$B_{S*} = \inf_{x < y} [\mu(-M, x)^{\frac{4}{2+\gamma}-1} + \mu(y, N)^{\frac{4}{2+\gamma}-1}]^{-\frac{2}{\gamma}-1} \hat{\nu}[x, y],$$

and $B_{S*} \leq B_S^* \leq 2^{2/\gamma} B_{S*}$.

(2) *Let $M, N = \infty$. If $\hat{\nu}(-\infty, \theta] \wedge \hat{\nu}[\theta, \infty) = \infty$,*

$$\overline{\lim}_{x \rightarrow -\infty} \mu(-\infty, x) \hat{\nu}(x, \theta) < \infty \quad \text{and} \quad \overline{\lim}_{y \rightarrow \infty} \mu(y, \infty) \hat{\nu}(\theta, y) < \infty,$$

$$\underline{\lim}_{x \rightarrow -\infty} \mu(-\infty, x) \hat{\nu}(x, \theta) > 0 \quad \text{and} \quad \underline{\lim}_{y \rightarrow \infty} \mu(y, \infty) \hat{\nu}(\theta, y) > 0$$

for some $\theta \in \mathbb{R}$, then the Nash inequality (3.1) does not hold whenever $\gamma \in (0, 2]$.

Proof (a) The assertion in Part (1) on the estimates of A_S is a straightforward consequence of Theorem 2.6 with $p = 2$, $q = 2\gamma/(\gamma - 2)$, and $(A_S, B_S^*, B_{S*}) = (A^2, B^{*2}, B_*^2)$. The criterion is a copy of Corollary 2.8.

(b) The restriction “ $\gamma > 2$ ” comes from the reduction of Nash inequality to the Sobolev-type one. This costs a smaller gap of the criterion for the Nash inequality, marked as (ε) in the last line on page 15 and the last sentence in Theorem 1.10 of [10], for instance. The restriction is recently removed in [17] in the discrete situation. Here we show that a direct proof is also possible under a technical condition.

To see that the Nash inequality (3.1) does not hold for $\gamma \in (1, 2]$, rewrite for a moment the inequality as

$$\text{Var}_\mu(f)^r \leq A_r \|f'\|_{\nu,2}^2, \quad f \in L^2(\mu), \quad \|f\|_{\mu,1} = 1, \quad (3.3)$$

where A_r denotes the optimal constant. When γ varies 0 to 2, r moves from ∞ to 2. By the splitting technique (replacing $\|f\|_{\mu,1}$ by $\|f - \pi(f)\|_{\mu,1}$ on the right-hand side of (3.1)), we may consider the half-space $(-M, N) = (0, \infty)$ only, and reduce (3.3) to

$$\|f\|_{\mu,2}^{2r} \leq C_r \|f'\|_{\nu,2}^2, \quad f(0) = 0, \quad f \in L^2(\mu), \quad \|f\|_{\mu,1} = 1,$$

Since $\|f\|_{\mu,2} \geq \|f\|_{\mu,1} = 1$, it is clear that the last inequality becomes stronger when r increases. Thus, it is sufficient to show that the inequality (3.3) does not hold when $r = 2$ (i.e., $\gamma = 2$).

To do so, fix a point $y > 0$ and let

$$f(x) = \hat{\nu}[0, x \wedge y], \quad x \geq 0.$$

Then

$$\begin{aligned} \|f\|_{\mu,1} &= \int_0^\infty \mu(dz) \hat{\nu}[0, z \wedge y] = \int_0^y h \mu(\cdot, \infty), \\ \|f\|_{\mu,2}^{2+4/\gamma} &\geq \mu(y, \infty)^{1+2/\gamma} \hat{\nu}[0, y]^{2+4/\gamma}, \\ \|f'\|_{\nu,2}^2 &= \hat{\nu}[0, y]. \end{aligned}$$

Now, we have

$$\begin{aligned} \frac{\|f\|_{\mu,2}^{2+4/\gamma}}{\|f'\|_{\nu,2}^2 \|f\|_{\mu,1}^{4/\gamma}} &\geq \frac{(\mu(y, \infty) \hat{\nu}[0, y]^2)^{1+2/\gamma}}{\hat{\nu}[0, y] [\int_0^y h \mu(\cdot, \infty)]^{4/\gamma}} \\ &= \frac{\mu(y, \infty)^{1+2/\gamma} \hat{\nu}[0, y]^{1+4/\gamma}}{[\int_0^y h \mu(\cdot, \infty)]^{4/\gamma}} \\ &= \frac{\mu(y, \infty)^{1-2/\gamma} \hat{\nu}[0, y]}{[\int_0^y h \mu(\cdot, \infty) / (\mu(y, \infty) \hat{\nu}[0, y]) + 1]^{4/\gamma}}. \end{aligned}$$

Since $\int_0^y h \mu(\cdot, \infty) > \mu(y, \infty) \hat{\nu}[0, y]$, when $\gamma = 2$, we need only to study the ratio

$$\hat{\nu}[0, y] \left[\frac{\mu(y, \infty) \hat{\nu}[0, y]}{\int_0^y h \mu(\cdot, \infty)} \right]^2.$$

By assumption, if $\int_0^\infty h \mu(\cdot, \infty) < \infty$, then the right-hand side goes to infinity as so does y by assumption again. This implies that $A_N = \infty$. Therefore the Nash inequality (3.1) does not hold at $\gamma \in (0, 2]$ in this case.

Next, if $\int_0^\infty h \mu(\cdot, \infty) = \infty$, then

$$\begin{aligned}
 \frac{\hat{\nu}[0, y]}{[\int_0^y h \mu(\cdot, \infty)]^2} &\sim \frac{h(y)}{h(y)\mu(y, \infty) \int_0^y h \mu(\cdot, \infty)} \quad (\text{by l'Hôpital's rule}) \\
 &= \frac{\hat{\nu}[0, y]}{\mu(y, \infty) \hat{\nu}[0, y] \int_0^y h \mu(\cdot, \infty)} \\
 &\sim \frac{\hat{\nu}[0, y]}{\int_0^y h \mu(\cdot, \infty)} \\
 &\sim \frac{h(y)}{h(y)\mu(y, \infty)} \\
 &= \frac{1}{\mu(y, \infty)} \rightarrow \infty \quad \text{as } y \rightarrow \infty.
 \end{aligned}$$

Therefore, we have arrived the required assertion again. $\square$

We remark that the condition

$$\overline{\lim}_{x \rightarrow -\infty} \mu(-\infty, x) \hat{\nu}(x, \theta) < \infty \quad \text{and} \quad \overline{\lim}_{y \rightarrow \infty} \mu(y, \infty) \hat{\nu}(\theta, y) < \infty$$

in Theorem 3.1 (2) means that the corresponding diffusion process is exponentially ergodic, otherwise, the Nash inequality can not hold since the latter inequality is stronger than the former ergodicity (cf. [10, Table 5.1 and Theorem 1.9]). Actually, by the cited results, once the transition probability of the process has a density, one can even assume a stronger condition that $\int_0^\infty h \mu(\cdot, \infty) < \infty$, then the proof above can be simplified.

The rough factor $4^{1+1/\gamma}$ ($\gamma > 2$) in Theorem 3.1 (1) is clearly smaller than the first factor 8 given by [10, Theorem 6.8]. The second factor given by the cited theorem is clearly less sharp than the first one and so is than what we have here.

3.2 Logarithmic Sobolev Inequality

We now turn to study the logarithmic Sobolev inequality

$$\text{Ent}_\pi(f^2) \leq A_{\text{LS}} \|f'\|_{\nu, 2}^2, \quad (3.4)$$

where

$$\text{Ent}_\pi(f) = \int_{-M}^N f \log \left(\frac{f}{\pi(f)} \right) d\pi \quad \text{for } f \geq 0.$$

Theorem 3.2 *Let $\mu[-M, N] < \infty$ and $\mu_{\text{pp}} = 0$. Then the optimal constant A_{LS} in (3.4) satisfies the following estimates*

$$B_* \leq A_{\text{LS}} \leq 4B^*,$$

where

$$\begin{aligned}
 B^{*-1} &= \inf_{x < y} \left\{ \hat{\nu}[x, y]^{-1} \left(\left[\pi[-M, x] \log \left(1 + \frac{e^2}{\pi[-M, x]} \right) \right]^{-1} \right. \right. \\
 &\quad \left. \left. + \left[\pi[y, N] \log \left(1 + \frac{e^2}{\pi[y, N]} \right) \right]^{-1} \right) \right\}, \\
 B_*^{-1} &= \inf_{\theta \in (-M, N)} \inf_{(x, y) \ni \theta} \left\{ \hat{\nu}[x, y]^{-1} \left(\left[\pi[-M, x] \log \left(1 + \frac{1 - \pi[-M, \theta]}{\pi[-M, x]} \right) \right]^{-1} \right. \right. \\
 &\quad \left. \left. + \left[\pi[y, N] \log \left(1 + \frac{1 - \pi[\theta, N]}{\pi[y, N]} \right) \right]^{-1} \right) \right\},
 \end{aligned}$$

or alternatively,

$$B_*^{-1} = \inf_{x < y} \left\{ \hat{\nu}[x, y]^{-1} \left(\left[\pi[-M, x] \log \left(1 + \frac{z^*(x, y)}{\pi[-M, x]} \right) \right]^{-1} + \left[\pi[y, N] \log \left(1 + \frac{1 - z^*(x, y)}{\pi[y, N]} \right) \right]^{-1} \right) \right\}, \quad (3.5)$$

where $z^*(x, y)$ is the unique solution to the equation

$$\begin{aligned} & \left[\pi[-M, x] \log \left(1 + \frac{z}{\pi[-M, x]} \right) \right]^2 \left(1 + \frac{z}{\pi[-M, x]} \right) \\ &= \left[\pi[y, N] \log \left(1 + \frac{1 - z}{\pi[y, N]} \right) \right]^2 \left(1 + \frac{1 - z}{\pi[y, N]} \right), \quad z \in (0, 1). \end{aligned} \quad (3.6)$$

In particular, we have

$$B_*^{-1} \leq \inf_{(x, y) \ni m(\pi)} \left\{ \hat{\nu}[x, y]^{-1} \left(\left[\pi[-M, x] \log \left(1 + \frac{1}{2\pi[-M, x]} \right) \right]^{-1} + \left[\pi[y, N] \log \left(1 + \frac{1}{2\pi[y, N]} \right) \right]^{-1} \right) \right\},$$

where $m(\pi)$ is the median of π .

Proof (a) Upper bound. Even though this theorem is not a consequence of Theorem 2.9, the idea of the proof for the upper bound is more or less the same as we used several times in the last two sections. Given $f \in L^2(\pi)$ with $\nu(f'^2) \in (0, \infty)$ and $\theta \in (-M, N)$, let

$$\tilde{f} = f - f(\theta), \quad \tilde{f}_- = \tilde{f} \mathbb{1}_{[-M, \theta]}, \quad \tilde{f}_+ = \tilde{f} \mathbb{1}_{(\theta, N]}.$$

The following facts were proved in the first part of [18, Proof of Theorem 3] for the specific $\theta = m(\pi)$, but the proof remains true for general θ :

$$\begin{aligned} \text{Ent}_\pi(f^2) &\leq \text{Ent}_\pi(\tilde{f}^2) + 2\pi(\tilde{f}^2) \quad (\text{by [10, Lemma 4.14]}) \\ &\leq \text{Ent}_\pi(\tilde{f}_-^2) + 2\pi(\tilde{f}_-^2) + \text{Ent}_\pi(\tilde{f}_+^2) + 2\pi(\tilde{f}_+^2) \end{aligned}$$

(since Ent_π is sub-additive: $\text{Ent}_\pi(f + g) \leq \text{Ent}_\pi(f) + \text{Ent}_\pi(g)$) and

$$\text{Ent}_\pi(\tilde{f}_\pm^2) + 2\pi(\tilde{f}_\pm^2) \leq 4B_\theta^\pm \nu(\tilde{f}_\pm'^2),$$

where

$$\begin{aligned} B_\theta^- &= \sup_{x < \theta} \pi[-M, x] \log \left(1 + \frac{e^2}{\pi[-M, x]} \right) \hat{\nu}[x, \theta], \\ B_\theta^+ &= \sup_{y > \theta} \pi[y, N] \log \left(1 + \frac{e^2}{\pi[y, N]} \right) \hat{\nu}[\theta, y]. \end{aligned}$$

Thus, we have

$$\begin{aligned} \frac{\text{Ent}_\pi(f^2)}{\nu(f'^2)} &\leq \frac{\text{Ent}_\pi(\tilde{f}^2) + 2\pi(\tilde{f}^2)}{\nu(f'^2)} \\ &\leq \frac{\text{Ent}_\pi(\tilde{f}_-^2) + 2\pi(\tilde{f}_-^2) + \text{Ent}_\pi(\tilde{f}_+^2) + 2\pi(\tilde{f}_+^2)}{\nu(f_-'^2) + \nu(f_+'^2)} \\ &\leq \frac{\text{Ent}_\pi(\tilde{f}_-^2) + 2\pi(\tilde{f}_-^2)}{\nu(f_-'^2)} \vee \frac{\text{Ent}_\pi(\tilde{f}_+^2) + 2\pi(\tilde{f}_+^2)}{\nu(f_+'^2)} \\ &\leq (4B_\theta^-) \vee (4B_\theta^+). \end{aligned}$$

The original proof for the upper estimate stopped here with $\theta = m(\pi)$. Because θ is arbitrary, we obtain

$$\frac{\text{Ent}_\pi(f^2)}{\nu(f'^2)} \leq 4 \inf_{\theta} (B_{\bar{\theta}}^- \vee B_{\bar{\theta}}^+).$$

Note that the right-hand side is independent of f . By choosing $\bar{\theta}$ such that $B_{\bar{\theta}}^- = B_{\bar{\theta}}^+$, it follows that

$$\frac{\text{Ent}_\pi(f^2)}{\nu(f'^2)} \leq 4B_{\bar{\theta}}^-.$$

Now, since f with $\nu(f'^2) \in (0, \infty)$ is arbitrary, we obtain

$$A_{\text{LS}} = \sup_{\nu(f'^2) \in (0, \infty)} \frac{\text{Ent}_\pi(f^2)}{\nu(f'^2)} \leq 4B_{\bar{\theta}}^-.$$

Next, as an application of Lemma 1.2, we have

$$\begin{aligned} B^* &= \sup_{x < y} \frac{\hat{\nu}[x, \bar{\theta}] + \hat{\nu}(\bar{\theta}, y)}{[\pi[-M, x] \log(1 + \frac{e^2}{\pi[-M, x]})]^{-1} + [\pi[y, N] \log(1 + \frac{e^2}{\pi[y, N]})]^{-1}} \\ &\geq B_{\bar{\theta}}^- \wedge B_{\bar{\theta}}^+ \\ &= B_{\bar{\theta}}^-. \end{aligned}$$

Combining the last two estimates together, we obtain the required upper bound. Once again, the unknown $\bar{\theta}$ disappears in the expression of B^* .

(b) Lower bound. We adopt a similar method as used in the proof of Lemma 2.4. Define

$$f(z) = -\mathbb{1}_{\{z \leq \theta\}} \hat{\nu}[x \vee z, \theta] + \mathbb{1}_{\{z > \theta\}} \hat{\nu}[\theta, y \wedge z], \quad z \in [-M, N],$$

where x, y, θ with $(x, y) \ni \theta$ are fixed. First, let us apply [18, Proof of Theorem 3] to this specific test function f .

$$\begin{aligned} \text{Ent}_\pi(f^2 \mathbb{1}_{\{z \leq \theta\}}) &= \sup \left\{ \int_{-M}^{\theta} f^2 g d\pi : \int_{-M}^{\theta} e^g d\pi \leq 1 \right\} \\ &\geq \sup \left\{ \int_{-M}^{\theta} f^2 g d\pi : g \geq 0 \text{ and } \int_{-M}^{\theta} e^g d\pi \leq 1 \right\} \\ &\geq \hat{\nu}[x, \theta]^2 \sup \left\{ \int_{-M}^{\theta} \mathbb{1}_{[-M, x]} g d\pi : g \geq 0 \text{ and } \int_{-M}^{\theta} e^g d\pi \leq 1 \right\}. \end{aligned}$$

Applying [18, Lemma 6] to the last supremum, it follows that

$$\begin{aligned} \text{Ent}_\pi(f^2 \mathbb{1}_{[-M, \theta]}) &\geq \hat{\nu}[x, \theta]^2 \varphi(x, \theta), \\ \varphi(x, \theta) &:= \pi[-M, x] \log \left(1 + \frac{1 - \pi[-M, \theta]}{\pi[-M, x]} \right). \end{aligned} \tag{3.7}$$

Symmetrically, we have

$$\begin{aligned} \text{Ent}_\pi(f^2 \mathbb{1}_{[\theta, N]}) &\geq \hat{\nu}[\theta, y]^2 \psi(\theta, y), \\ \psi(\theta, y) &:= \pi[y, N] \log \left(1 + \frac{1 - \pi[\theta, N]}{\pi[y, N]} \right). \end{aligned} \tag{3.8}$$

Next, by logarithmic Sobolev inequality,

$$A_{\text{LS}} \nu(f'^2 \mathbb{1}_{[-M, \theta]}) \geq \text{Ent}_\pi(f^2 \mathbb{1}_{[-M, \theta]}),$$

it follows that

$$A_{\text{LS}} \nu(f'^2 \mathbb{1}_{[-M, \theta]}) \geq \hat{\nu}[x, \theta]^2 \varphi(x, \theta).$$

Similarly,

$$A_{\text{LS}} \nu(f'^2 \mathbb{1}_{[\theta, N]}) \geq \hat{\nu}[\theta, y]^2 \psi(\theta, y).$$

We now arrive at the place different from the known proofs. Summing up the last two inequalities, it follows that

$$A_{\text{LS}} \nu(f'^2) \geq \hat{\nu}[x, \theta]^2 \varphi(x, \theta) + \hat{\nu}[\theta, y]^2 \psi(\theta, y).$$

Since one can replace $\nu(f'^2)$ by $\hat{\nu}(f'^2)$ in the original inequality, by definition of f , we have $\hat{\nu}(f'^2) = \hat{\nu}[x, y]$, and so

$$A_{\text{LS}} \geq \frac{1}{\hat{\nu}[x, y]} [\hat{\nu}[x, \theta]^2 \varphi(x, \theta) + \hat{\nu}[\theta, y]^2 \psi(\theta, y)].$$

Therefore, we have

$$A_{\text{LS}} \geq \hat{\nu}[x, y] \left[\left(\frac{\hat{\nu}[x, \theta]}{\hat{\nu}[x, y]} \right)^2 \varphi(x, \theta) + \left(1 - \frac{\hat{\nu}[x, \theta]}{\hat{\nu}[x, y]} \right)^2 \psi(\theta, y) \right].$$

Noting that the function $c_1 z^2 + c_2(1 - z)^2$ on $[0, 1]$ achieves its minimum $(c_1^{-1} + c_2^{-1})^{-1}$ at $z^* = c_2/(c_1 + c_2)$, it follows that

$$A_{\text{LS}} \geq \hat{\nu}[x, y] (\varphi(x, \theta)^{-1} + \psi(\theta, y)^{-1})^{-1}.$$

Since $(x, y) \ni \theta$ are arbitrary, we finally arrive at

$$A_{\text{LS}} \geq \sup_{\theta \in (-M, N)} \sup_{(x, y) \ni \theta} \hat{\nu}[x, y] (\varphi(x, \theta)^{-1} + \psi(\theta, y)^{-1})^{-1}. \quad (3.9)$$

This gives us the first version of B_* . Then the final assertion of the theorem follows by setting $\theta = m(\pi)$.

(c) We now prove the alternative assertion of B_* . Since supremums are exchangeable, one may rewrite B_* as

$$B_* = \sup_{x < y} \left\{ \hat{\nu}[x, y] \sup_{\theta \in (x, y)} (\varphi(x, \theta)^{-1} + \psi(\theta, y)^{-1})^{-1} \right\}.$$

Fix $x < y$ and make a change of the variable θ by $z = 1 - \pi[-M, \theta]$. Then $1 - \pi[\theta, N] = 1 - z$ and the functions $\varphi(x, \theta)$ and $\psi(\theta, y)$ become $\tilde{\varphi}(x, z)$ and $\tilde{\psi}(z, y)$, respectively. The supremum above should be achieved at the point for which the derivative in z of $(\tilde{\varphi}(x, z)^{-1} + \tilde{\psi}(z, y)^{-1})^{-1}$ vanishes. This leads to the unique solution $z^* = z^*(x, y)$ to equation (3.6). Then we obtain (3.5). $\square$

We mention that there is a large number of publications on the logarithmic Sobolev inequalities, in the one-dimensional case for instance, one may refer to [18, 19], [10, §4.6 and §6.6] for related results and more references. Generally speaking, Theorem 3.2 clearly improves [18, Theorem 3] (since $\alpha \vee \beta$ is replaced by $\alpha + \beta$), to which the factor of the upper and lower bounds is at most 16, the best one we have known up to now. In the special case that the measures π and $\hat{\nu}$ are symmetric with respect to $m(\pi)$, the computation of B^* and B_* can be reduced to half space. Then Theorem 3.2 coincides with [18, Theorem 3]. Besides, having Theorem 3.2 at

hand, it should be easy to introduce the corresponding H^o and H^∂ as we did in the previous sections.

The methods introduced in the paper should have more applications. For instance, in parallel to the proof of Theorem 3.2, we may have an improved version of the Sobolev inequality and the Latała–Oleszkiewicz inequality presented by [18, Theorems 11 and 13].

Acknowledgements The results of the paper have been presented in our seminar. Thanks are given to Y. H. Mao, F. Y. Wang, Y. H. Zhang, and the participants for their helpful comments and suggestions which lead to some improvements of the paper. Thanks are also given to a referee and an editor for correcting a number of typos in a previous version of the paper.

References

- [1] Kokilashvili, V., Meshki, A., Persson, L. E.: Weighted Norm Inequalities for Integral Transforms with Product Weights, Nova Sci. Publ., New York, 2010
- [2] Kufner, A., Maligranda, L., Persson, L. E.: The Hardy Inequality: About its History and Some Related Results, Vydavatel'sky Servis, 2007
- [3] Kufner, A., Persson, L. E.: Weighted Inequalities of Hardy-type, World Scientific, Singapore, 2003
- [4] Maz'ya, V.: Sobolev Spaces with Applications to Elliptic Partial Differential Equations (2nd Ed.), Springer, Berlin, 2011
- [5] Opic, B., Kufner, A.: Hardy-type Inequalities, Longman, New York, 1990
- [6] Chen, M. F.: Speed of stability for birth-death processes. *Front. Math. China*, **5**(3), 379–515 (2010)
- [7] Chen, M. F.: Basic estimates of stability rate for one-dimensional diffusions. Chapter 6 in “Probability Approximations and Beyond”, eds. A.D. Barbour, H.P. Chan and D. Siegmund, Lecture Notes in Statistics, **205**, 2012, 75–99
- [8] Chen, M. F., Wang, F. Y.: Cheeger's inequalities for general symmetric forms and existence criteria for spectral gap. *Ann. Prob.*, **28**(1), 235–257 (2000)
- [9] Chen, M. F.: Explicit bounds of the first eigenvalue. *Sci. China, Ser. A*, **43**(10), 1051–1059 (2000)
- [10] Chen, M. F.: Eigenvalues, Inequalities, and Ergodic Theory, Springer, London, 2005
- [11] Chen, M. F., Zhang, Y. H., Zhao, X. L.: Dual variational formulas for the first Dirichlet eigenvalue on half-line. *Sci. China*, **46**(6), 847–861 (2003)
- [12] Bennett, G.: Some elementary inequalities III. *Quart. J. Math. Oxford Ser. (2)*, **42**(166), 149–174 (1991)
- [13] Manakov, V. M.: On the best constant in weighted inequalities for Riemann–Liouville integrals. *Bull. London Math. Soc.*, **24**, 442–448 (1992)
- [14] Muckenhoupt, B.: Hardy's inequality with weights. *Studia Math.*, **XLIV**, 31–38 (1972)
- [15] Chen, M. F.: Variational formulas of Poincaré-type inequalities in Banach spaces of functions on the line. *Acta Mathematica Sinica, English Series*, **18**(3), 417–436 (2002)
- [16] Mao, Y. H.: Nash inequalities for Markov processes in dimension one. *Acta Mathematica Sinica, English Series*, **18**(1), 147–156 (2002)
- [17] Wang, J.: Criteria for functional inequalities for ergodic birth-death processes. *Acta Mathematica Sinica, English Series*, **28**(2), 357–370 (2012)
- [18] Barthe, F., Roberto, C.: Sobolev inequalities for probability measures on the real line. *Studia Math.*, **159**(3), 481–497 (2003)
- [19] Mao, Y. H.: The logarithmic Sobolev inequalities for birth-death process and diffusion process on the line. *Chin. J. Appl. Prob. Statist.*, **18**(1), 94–100 (2002)

Mixed Principal Eigenvalues in Dimension One

Mu-Fa CHEN, Ling-Di WANG, Yu-Hui ZHANG

School of Mathematics Sciences, Beijing Normal University, Laboratory of Mathematics and Complex Systems (Beijing Normal University), Ministry of Education, Beijing 100875, China

© Higher Education Press and Springer-Verlag Berlin Heidelberg 2012

Abstract This is one of a series of papers exploring the stability speed of one-dimensional stochastic processes. The present paper emphasizes on the principal eigenvalues of elliptic operators. The eigenvalue is just the best constant in the L^2 -Poincaré inequality and describes the decay rate of the corresponding diffusion process. We present some variational formulas for the mixed principal eigenvalues of the operators. As applications of these formulas, we obtain case by case explicit estimates, a criterion for positivity, and an approximating procedure for the eigenvalue.

Keywords Eigenvalue, variational formula, explicit estimates, positivity criterion, approximating procedure

MSC 60J60, 34L15

1 Introduction

This paper is a continuation of [5] in which the stability speed was carefully studied in the discrete situation (birth–death processes) and partially in the continuous one (diffusions). For a large part of the study, the description of the problem is equivalent to that of the Poincaré-type inequalities or the principal eigenvalue. On the last two topics, there are a great number of publications (cf. [4, 6] and references therein for the background and motivation of the study on these topics). However, to save the space here, most of the references are not repeated in this paper. Consider a finite interval $(0, D)$ for a moment. We are interested in some typical Sturm-Liouville eigenvalue problems. According to the Dirichlet (denoted by code “D”) and Neumann (denoted by code “N”) boundaries at the left- or right-endpoint, we have four cases of boundary condition: DD, ND, DN and NN. In the diffusion context, the DD- and NN-cases are largely handled in [1–5] and [7, 8]. The present paper is mainly devoted to the ND- and DN-cases. As will be seen in the next section, the classification for the boundaries is also meaningful when $D = \infty$.

Received March 6, 2012; accepted June 8, 2012

Corresponding author: Ling-Di Wang, E-mail: wanglingdi@mail.bnu.edu.cn

The paper is organized as follows. In the next section, we focus on the ND-case. First, we introduce several variational formulas for the eigenvalue. As a consequence, we obtain the basic estimates, a criterion for positivity, an approximating procedure, and improved estimates for the eigenvalue. As far as we know, most of these results, except Theorem 5, have not yet appeared in the literature. The proofs of them are sketched in Section 3. From [5; Section 10], we know that the DN-case and the ND-case are dual to each other. Thus, as a dual to the ND-case, it is natural to study the DN-case, to which Section 4 is devoted, partial results come from the duality but some of them are not and need direct proofs. The main extension to the earlier study is that here we do not assume the uniqueness of the processes, instead of which we adopt the maximal extension of the Dirichlet form or the maximal process. Finally, some supplement to [2, 3, 8] in the NN-case (i.e., the ergodic case.) is presented in Section 5. The complete proofs of the results presented in this paper are quite technical and long. However, a large part of them are parallel to [5] and so we omit mostly the “translation” from the discrete situation to the continuous one. Instead, we emphasize on the difference between them (Lemmas 8–14, for instance), and illustrate a little of the translation for the reader’s reference. We may leave the details to our homepage or publish them elsewhere.

The basic estimates are also studied in [9] in terms of H -transform. Some examples of the study are illustrated in [6; Section 5].

Here we discuss briefly about the problem on the whole line. First, we consider the ND-case. Then one may regard the whole line $\mathbb{R}$ as a limit of $[M, \infty)$ as M decreases to $-\infty$. Then the mixed eigenvalue problem on line is known by what we are studying in the paper. Next, consider the DD-case, one may split $\mathbb{R}$ into two parts: $(-\infty, 0)$ and $(0, \infty)$. The case with ND-boundaries on $(0, \infty)$ is studied in Sections 2 and 3. Besides, the case with DN-boundaries on $(-\infty, 0)$ is simply a reverse of the ND-case on $(0, \infty)$. Therefore, the behavior of the original operator on the whole line should be clear. However, there is an interesting point here. On $(0, \infty)$, we use the minimal Dirichlet form but on $(-\infty, 0)$ we adopt the maximal one. Thus, the domain of the original Dirichlet form on the whole line may be neither the maximal nor the minimal one. Therefore, it is essentially different from DD- or NN-cases on the whole line we have studied in [5, 6, 7].

To conclude this section, we mention that in a more general context, for the Poincaré-type inequalities, the DN-case was completed earlier (cf. [4; Chapter 6]), the basic estimates for the ND-case in the discrete situation was given by [5; Theorem 8.5], from which one can write down easily the continuous version.

2 ND-case

Define

$$\mathcal{C}[0, D] = \{f : f \text{ is continuous on } [0, D]\} \quad \text{and}$$

$$\mathcal{C}^k(0, D) = \{f : f \text{ has continuous derivatives of order } k \text{ on } (0, D)\}, \quad k \geq 1.$$

Here and in what follows, when $D = \infty$, the notation $\mathcal{C}[0, D]$ simply means

$\mathcal{C}[0, D)$. The convention should be clear in other cases and we will not mention time by time. Let

$$L = a(x) \frac{d^2}{dx^2} + b(x) \frac{d}{dx}$$

be an elliptic operator on an interval $(0, D)$ ($D \leq \infty$). Set

$$C(x) = \int_0^x \frac{b(u)}{a(u)} du.$$

Throughout this paper, we need the following hypothesis (which is trivial in the discrete situation):

The functions a, b are Borel measurable on $[0, D]$ and a is positive on $[0, D]$, b/a and e^C/a are locally integrable on $[0, D]$. (1)

Note that for continuous functions a and b , the hypothesis (1) is reduced to the condition $a > 0$ only. In this section, we consider the ND-boundaries only. More precisely, as usual, the Dirichlet boundary condition at D means that $g(D) = 0$ when $D < \infty$. When $D = \infty$, it is natural to take “ $\lim_{x \rightarrow \infty} g(x) = 0$ ” as a boundary condition. However, this is not pre-assumed but proved later (cf. Lemma 14 below). Therefore, the code “ND” is still meaningful even if $D = \infty$.

Throughout this section, we work on the following mixed principal eigenvalue:

$$\lambda_0 = \inf \{ D(f) : \mu(f^2) = 1, f \in \mathcal{C}_K[0, D], f(D) = 0 \text{ if } D < \infty \}, \quad (2)$$

where $\mu(f) = \int_0^D f d\mu$,

$$\begin{aligned} \mathcal{C}_K[0, D] &= \{ f : f \in \mathcal{C}^1(0, D) \cap \mathcal{C}[0, D], f \text{ has compact support} \}, \\ D(f) &= \int_0^D a f'^2 d\mu, \quad \mu(dx) = \frac{e^{C(x)}}{a(x)} dx. \end{aligned}$$

Besides μ , throughout the paper, we often use another measure:

$$\nu(dx) = e^{-C(x)} dx.$$

When $D < \infty$, λ_0 coincides with the minimal solution λ to the following eigenequation:

$$Lf = -\lambda f, \quad f'(0) = 0, \quad \text{and} \quad f(D) = 0 \text{ if } D < \infty.$$

To state our results, we need some notation. Define

$$\begin{aligned} I(f)(x) &= -\frac{e^{-C(x)}}{f'(x)} \int_0^x f d\mu \quad (\text{single integral form}), \\ II(f)(x) &= \frac{1}{f(x)} \int_{(x, D) \cap \text{supp}(f)} \nu(ds) \int_0^s f d\mu, \quad x \in \text{supp}(f) \quad (\text{double integral form}), \\ R(h)(x) &= -(ah^2 + bh + ah')(x) \quad (\text{differential form}). \end{aligned}$$

The domains of the three operators defined above are, respectively, as follows:

$$\begin{aligned}\mathcal{F}_I &= \{f : f \in \mathcal{C}^1(0, D) \cap \mathcal{C}[0, D], f|_{(0, D)} > 0, \text{ and } f'|_{(0, D)} < 0\}, \\ \mathcal{F}_{II} &= \{f : f \in \mathcal{C}[0, D] \text{ and } f|_{(0, D)} > 0\}, \\ \mathcal{H} &= \{h : h \in \mathcal{C}^1(0, D) \cap \mathcal{C}[0, D], h(0) = 0, h|_{(0, D)} < 0 \text{ if } \nu(0, D) < \infty, \\ &\quad \text{and } h|_{(0, D)} \leq 0 \text{ if } \nu(0, D) = \infty\}, \quad \nu(\alpha, \beta) := \int_{\alpha}^{\beta} d\nu.\end{aligned}$$

These sets are used for the lower estimates of λ_0 . For the upper bounds, some modifications are needed to avoid the non-integrability problem, as shown below:

$$\begin{aligned}\widetilde{\mathcal{F}}_I &= \{f : f \in \mathcal{C}^1(x_0, x_1) \cap \mathcal{C}[x_0, x_1], f'|_{(x_0, x_1)} < 0 \text{ for some } x_0, x_1 \in [0, D] \text{ with} \\ &\quad x_0 < x_1, \text{ and } f = f(\cdot \vee x_0)\mathbb{1}_{[0, x_1)}\}, \\ \widetilde{\mathcal{F}}_{II} &= \{f : \exists x_0 \in (0, D) \text{ such that } f = f\mathbb{1}_{[0, x_0)} \text{ and } f \in \mathcal{C}[0, x_0]\}, \\ \widetilde{\mathcal{H}} &= \left\{h : \exists x_0 \in (0, D) \text{ such that } h \in \mathcal{C}^1(0, x_0) \cap \mathcal{C}[0, x_0], h|_{(0, x_0)} < 0, \right. \\ &\quad \left. h|_{[x_0, D]} = 0, \text{ and } h(0) = 0, \sup_{(0, x_0)} (ah^2 + bh + ah') < 0\right\}.\end{aligned}$$

Here and in what follows, we adopt the usual convention $1/0 = \infty$. The superscript “ $\sim$ ” means modified. In the formulas of Theorem 1 below, “sup inf” are used for lower bounds of λ_0 , each test function f produces a lower bound $\inf_x I(f)(x)^{-1}$, and so this part is called variational formula for the lower estimate of λ_0 . Dually, the “inf sup” are used for upper estimates of λ_0 . Among them, the ones expressed by the operator R are easiest to compute in practice, and the ones expressed by II are hardest to compute but provide better estimates. Because of “inf sup”, a localizing procedure is used for the test functions to avoid $I(f) \equiv \infty$ for instance, which is removed out automatically for the “sup inf” part. Each part of Theorem 1 below plays a role in our study. Parts (1) and (2) are applied to Theorems 5 and 6, respectively. Part (3) is a comparison with Proposition 3, which is then used as a dual form of Theorem 17 (3).

Theorem 1. *Under hypothesis (1), the following variational formulas hold for λ_0 defined by (2).*

(1) *Single integral forms:*

$$\inf_{f \in \widetilde{\mathcal{F}}_I} \sup_{x \in (0, D)} I(f)(x)^{-1} = \lambda_0 = \sup_{f \in \mathcal{F}_I} \inf_{x \in (0, D)} I(f)(x)^{-1},$$

(2) *Double integral forms:*

$$\inf_{f \in \widetilde{\mathcal{F}}_{II}} \sup_{x \in \text{supp}(f)} II(f)(x)^{-1} = \lambda_0 = \sup_{f \in \mathcal{F}_{II}} \inf_{x \in (0, D)} II(f)(x)^{-1}.$$

Moreover, if $a, b \in \mathcal{C}[0, D]$, then we have additionally

(3) differential forms:

$$\inf_{h \in \widetilde{\mathcal{H}}} \sup_{x \in (0, D)} R(h)(x) = \lambda_0 = \sup_{h \in \mathcal{H}} \inf_{x \in (0, D)} R(h)(x).$$

Furthermore, the supremum on the right-hand side of the above three formulas can be attained.

The next result, similar to the discrete case, either extends the domain of λ_0 , or adds some additional sets of test functions for operators I and II , respectively. Besides, as an application of the lower variational formula (Theorem 1 (2)), we obtain the vanishing property of the eigenfunction (Lemma 14) which leads to the crucial part (1) of the proposition below. The vanishing property is the meaning of the Dirichlet boundary at $D = \infty$ as we expected. A more common description of λ_0 is given by Lemma 9 below.

Proposition 2. *Let hypothesis (1) hold. Then*

(1) *we have*

$$\begin{aligned} \lambda_0 &= \inf \left\{ D(f) : \mu(f^2) = 1, f \in \mathcal{C}^1(0, D) \cap \mathcal{C}[0, D] \text{ and } f(D) = 0 \right\} \\ &=: \tilde{\lambda}_0, \end{aligned}$$

where $f(D) = \lim_{x \rightarrow D} f(x)$ in the case of $D = \infty$.

(2) *Moreover, we have*

$$\inf_{f \in \widetilde{\mathcal{F}}'_I} \sup_{x \in (0, D)} I(f)(x)^{-1} = \lambda_0 = \sup_{f \in \mathcal{F}_I} \inf_{x \in (0, D)} II(f)(x)^{-1}, \quad (3)$$

$$\inf_{f \in \widetilde{\mathcal{F}}_{II} \cup \widetilde{\mathcal{F}}'_{II}} \sup_{x \in \text{supp}(f)} II(f)(x)^{-1} = \lambda_0 = \inf_{f \in \mathcal{F}_{II}} \sup_{x \in \text{supp}(f)} II(f)(x)^{-1}, \quad (4)$$

where

$$\begin{aligned} \widetilde{\mathcal{F}}'_I &= \{f : \exists x_0 \in (0, D) \text{ such that } f = f \mathbb{1}_{[0, x_0]}, f \in \mathcal{C}^1(0, x_0) \\ &\quad \cap \mathcal{C}[0, x_0], \text{ and } f'|_{(0, x_0)} < 0\}, \\ \widetilde{\mathcal{F}}_{II} &= \{f : f > 0, f \in \mathcal{C}[0, D], \text{ and } f II(f) \in L^2(\mu)\}. \end{aligned}$$

Besides, the supremum over $\{f \in \mathcal{F}_I\}$ in (3) can be attained.

The operator $\overline{R}$ defined below was first introduced in [8; Theorem 2.1] based on a probabilistic (coupling) technique. Different from R , it is a “bridge” in proving the duality of the ND- and DN-cases. It also leads to a different variational formula for λ_0 as follows.

Proposition 3. *Suppose that $a, b \in \mathcal{C}^1(0, D) \cap \mathcal{C}[0, D]$ and $a > 0$ on $(0, D)$. Set*

$$\overline{\mathcal{H}} = \{h : h(0) = 0, h \in \mathcal{C}^2(0, D) \cap \mathcal{C}[0, D], \text{ and } h|_{(0, D)} < 0\}$$

and define

$$\overline{R}(h)(x) = -\frac{(ah' + bh)'(x)}{h(x)}.$$

Then

(1) we have

$$\sup_{h \in \mathcal{H}} \inf_{x \in (0, D)} \overline{R}(h)(x) \geq \lambda_0$$

and the equality holds once $\mu(0, D) = \infty$.

(2) In general, we have

$$\lambda_0 = \sup_{h \in \mathcal{H}_*} \inf_{x \in (0, D)} \overline{R}(h)(x), \quad (5)$$

where

$$\mathcal{H}_* = \{h \in \mathcal{C}^2(0, D) \cap \mathcal{C}[0, D] : h(0) = 0 \text{ and } h < 0, h' < -a^{-1}bh \text{ on } (0, D)\}.$$

Moreover, the supremum in (5) can be attained.

Remark 4. (Comparison of R and $\overline{R}$) With $h = g'/g$, we have

$$-\frac{Lg}{g} = -(ah^2 + bh + ah') = R(h).$$

Next, with $h = g'$, we have

$$-\frac{(Lg)'}{g'} = -\frac{(ah' + bh)'}{h} = \overline{R}(h).$$

As an application of Theorem 1(1) to the test function $\nu(x, D)^\gamma$ with $\gamma = 1/2$ or 1 , we obtain the basic estimates and furthermore a criterion as follows.

Theorem 5. (Criterion and basic estimates) *Let hypothesis (1) hold. Then $\lambda_0 > 0$ iff*

$$\delta := \sup_{x \in (0, D)} \mu(0, x) \nu(x, D) < \infty, \quad \mu(\alpha, \beta) := \int_\alpha^\beta d\mu.$$

More precisely, we have

$$(4\delta)^{-1} \leq \lambda_0 \leq \delta^{-1}.$$

In particular, when $D = \infty$, we have $\lambda_0 = 0$ if $\nu(0, D) = \infty$, and $\lambda_0 > 0$ if $\int_0^\infty \mu(0, x) \nu(dx) < \infty$.

The next result is an application of Theorem 1(2), repeated with $f = f_n$, starting from the initial function f_1 , the test function just mentioned before Theorem 5. The result provides us a way to improve the basic estimates step by step. In view of the last criterion, for any improvement, one may assume that $\delta < \infty$.

Theorem 6. (Approximating procedure) *Let hypothesis (1) hold and assume that $\delta < \infty$. Set $\varphi(x) = \nu(x, D)$.*

(1) *Let $f_1 = \sqrt{\varphi}$, $f_n = f_{n-1}H(f_{n-1})$, and $\delta_n = \sup_{x \in (0, D)} H(f_n)(x)$ for $n \geq 1$. Then δ_n is decreasing in n and*

$$\lambda_0 \geq \delta_n^{-1} \geq (4\delta)^{-1}, \quad n \geq 1.$$

(2) *For fixed $x_0, x_1 \in [0, D]$ with $x_0 < x_1$, define*

$$f_1^{x_0, x_1} = \nu(\cdot \vee x_0, x_1) \mathbb{1}_{[0, x_1]}, \quad f_n^{x_0, x_1} = (f_{n-1}^{x_0, x_1} H(f_{n-1}^{x_0, x_1}))(\cdot \vee x_0) \mathbb{1}_{[0, x_1]}, \quad n \geq 1,$$

and let

$$\delta'_n = \sup_{x_0, x_1: x_0 < x_1} \inf_{x < x_1} H(f_n^{x_0, x_1})(x).$$

Then $\delta^{-1} \geq \delta_n'^{-1} \geq \lambda_0$ for $n \geq 1$.

(3) Define

$$\bar{\delta}_n = \sup_{x_0, x_1: x_0 < x_1} \frac{\|f_n^{x_0, x_1}\|}{D(f_n^{x_0, x_1})}, \quad n \geq 1.$$

Then $\bar{\delta}_n^{-1} \geq \lambda_0$, $\bar{\delta}_{n+1} \geq \delta'_n$ for $n \geq 1$, and $\bar{\delta}_1 = \delta'_1$.

The next result comes from the first step of the approximation above.

Corollary 7. (Improved estimates) *Let hypothesis (1) hold. For λ_0 , we have*

$$\delta^{-1} \geq \delta_1'^{-1} \geq \lambda_0 \geq \delta_1^{-1} \geq (4\delta)^{-1},$$

where

$$\begin{aligned} \delta_1 &= \sup_{x \in (0, D)} \frac{1}{\sqrt{\varphi(x)}} \int_0^D \sqrt{\varphi} \varphi(\cdot \vee x) d\mu \\ &= \sup_{x \in (0, D)} \left(\sqrt{\varphi(x)} \int_0^x \sqrt{\varphi} d\mu + \frac{1}{\sqrt{\varphi(x)}} \int_x^D \varphi^{3/2} d\mu \right), \end{aligned} \quad (6)$$

$$\delta_1' = \sup_{x \in (0, D)} \left(\mu(0, x) \varphi(x) + \frac{1}{\varphi(x)} \int_x^D \varphi^2 d\mu \right) \in [\delta, 2\delta]. \quad (7)$$

3 Partial proofs of the results in Section 2

Some preparations are needed to prove our main results. The first six lemmas below, except Lemma 9, are mainly devoted to describe the eigenfunction of λ_0 . The Lemmas are essential in our study. Note that their proofs are very different from the discrete situation. The first one below is taken from [10; Theorems 1.2.1 and 2.2.1].

- Lemma 8.** (1) *Let hypothesis (1) hold. Then, whenever g and g' are initially not vanished simultaneously, there exists uniquely a non-zero function $g \in \mathcal{C}^1[0, D]$ such that g' is absolutely continuous on each compact subinterval of $[0, D)$ and the eigenequation $Lg = -\lambda g$ holds almost everywhere.*
- (2) *Suppose additionally a and b are continuous on $[0, D]$. Then $g \in \mathcal{C}^2[0, D]$ and the eigenequation holds everywhere on $[0, D]$.*

In what follows, we call the function g given in part (1) of Lemma 8 *a.e. eigenfunction* of λ . Remember we need “a.e.” only in the case where g'' is used. Of course, we remove “a.e.” if the eigenequation holds everywhere.

The next result enables us to return to a more common description of the eigenvalue.

Lemma 9. *Let $\mathcal{A}[\alpha, \beta]$ be the set of all absolutely continuous functions on $[\alpha, \beta]$. Define*

$$\lambda_* = \inf \{ D(f) : f \in \mathcal{A}[0, D], \|f\| = 1, f \text{ has compact support and } f(D) = 0 \text{ if } D < \infty \}$$

Then $\lambda_0 = \lambda_*$.

Proof It is obvious that $\lambda_* \leq \lambda_0$. Next, let g be the a.e. eigenfunction of λ_* . Then, $g \in \mathcal{C}^1[0, D]$ by Lemma 8(1). By making inner product with g on the both sides of $Lg = -\lambda_* g$ with respect to μ , it follows that

$$-(e^C g g') \Big|_0^D + D(g) = \lambda_* \|g\|^2.$$

Since $g'(0) = 0$ and $(gg')(D) \leq 0$, we have $\lambda_* \geq D(g)/\|g\|^2$. Because $g \in \mathcal{C}^1[0, D]$, it is clear that $D(g)/\|g\|^2 \geq \lambda_0$. We have thus obtained that

$$\lambda_0 \leq \lambda_* \leq \lambda_0,$$

and so $\lambda_0 = \lambda_*$. There is a small gap in the proof above since in the case of $D = \infty$, the a.e. eigenfunction g may not belong to $L^2(\mu)$ and we have not yet proved that $(gg')(D) \leq 0$. However, one may avoid this by a standard approximating procedure,¹ using $[0, p_n]$ instead of $[0, D]$ with $p_n \uparrow D$ provided $D = \infty$ ²:

$$\begin{aligned} \lim_{n \rightarrow \infty} \lambda_0^{(0, p_n)} &= \lim_{n \rightarrow \infty} \inf \{ D(f) : f \in \mathcal{C}[0, p_n] \cap \mathcal{C}^1(0, p_n), \mu(f^2) = 1, f|_{[p_n, D]} = 0 \} \\ &= \inf \{ D(f) : \mu(f^2) = 1, f \in \mathcal{C}_K[0, D], f(D) = 0 \text{ if } D < \infty \} \\ &= \lambda_0. \quad \square \end{aligned}$$

¹If $D = \infty$ and $\lambda_* < \lambda_0$, then there would exist $p_n < \infty$ such that $\lambda_*^{(0, p_n)} < \lambda_0^{(0, p_n)}$ which is a contradiction with what we have just proved.

²By definition of λ_0 , for any $n > 0$, there exists $p_n \in (0, D)$ such that $f^{(n)} = f \mathbb{1}_{[0, p_n]} \in \mathcal{C}_K[0, D]$, $\mu(f^{(n)2}) = 1$ and $\lambda_0 \leq D(f^{(n)}) \leq \lambda_0 + 1/n$. So

$$\lambda_0 \leq \lambda_0^{(0, p_n)} \leq D(f^{(n)}) \leq \lambda_0 + \frac{1}{n},$$

and the required assertion holds by letting $n \rightarrow \infty$.

Remark 10. *Define*

$$\begin{aligned}\lambda_*^{(0,p_n)} &= \{D(f) : f \in \mathcal{A}[0, p_n], \mu(f^2) = 1, \text{ and } f|_{[p_n, D]} = 0\}, \\ \tilde{\lambda}_* &= \{D(f) : f \in \mathcal{A}[0, D], \mu(f^2) = 1, \text{ and } f(D) = 0\}\end{aligned}$$

Then we have $\lambda_* = \tilde{\lambda}_*$. Indeed, on the one hand, by definition, if $p_{n+1} \geq p_n$, then $\lambda_*^{(0,p_n)} \geq \lambda_*^{(0,p_{n+1})}$. We have thus obtained

$$\lim_{n \rightarrow \infty} \lambda_*^{(0,p_n)} \geq \lambda_*^{(0,D)} = \tilde{\lambda}_*.$$

On the other hand, by definition of $\tilde{\lambda}_*$, for any fixed $\varepsilon > 0$, there exists f satisfy $\|f\| = 1$, $f(D) = 0$, and $D(f) \leq \tilde{\lambda}_* + \varepsilon$. Let $p_n \uparrow D$ and $f^{(n)} = (f - f(p_n))\mathbb{1}_{[0,p_n]}$. Then we have

$$D(f_n) \uparrow D(f) \quad \text{as } n \rightarrow \infty.$$

Choose subsequence $\{n_m\}_{m \geq 1}$ if necessary such that

$$\overline{\lim}_{n \rightarrow \infty} \frac{D(f_n)}{\|f_n\|} = \lim_{m \rightarrow \infty} \frac{D(f_{n_m})}{\|f_{n_m}\|}.$$

By Fatou's lemma and the fact that $f(D) = 0$, we have

$$\underline{\lim}_{m \rightarrow \infty} \|f_{n_m}\| \geq \left\| \underline{\lim}_{m \rightarrow \infty} f_{n_m} \right\| = \|f\| = 1.$$

Therefore, we obtain

$$\overline{\lim}_{n \rightarrow \infty} \lambda_*^{(0,p_n)} \leq \overline{\lim}_{n \rightarrow \infty} \frac{D(f_n)}{\|f_n\|} = \lim_{m \rightarrow \infty} \frac{D(f_{n_m})}{\|f_{n_m}\|} \leq \frac{\lim_{m \rightarrow \infty} D(f_{n_m})}{\underline{\lim}_{m \rightarrow \infty} \|f_{n_m}\|} \leq D(f) \leq \tilde{\lambda}_* + \varepsilon.$$

Combining this with the assertion above, we get the required one. Moreover, we have

$$\tilde{\lambda}_0 \geq \tilde{\lambda}_* = \lambda_* = \lambda_0 \geq \tilde{\lambda}_0.$$

So $\tilde{\lambda}_0 = \lambda_0 = \lambda_* = \tilde{\lambda}_*$.

Clearly, because of hypothesis (1), we have $\lambda_0 > 0$ once $D < \infty$. The next result is a simple comparison. For given α, β ($\alpha < \beta$), denote by $\lambda_0^{(\alpha, \beta)}$ and $\lambda_1^{(\alpha, \beta)}$, respectively, the principal ND- and NN-eigenvalue (the latter is also called the first nontrivial eigenvalue or the spectral gap in the ergodic case). For simplicity, we use $\downarrow$ (resp. $\downarrow\downarrow, \uparrow, \uparrow\uparrow$) to denote decreasing (resp. strictly decreasing, increasing, strictly increasing).

Lemma 11. (1) For $p, q \in (0, D)$ with $p < q$, we have $\lambda_0^{(0,p)} > \lambda_0^{(0,q)}$. Furthermore, $\lambda_0^{(0,p_n)} \downarrow\downarrow \lambda_0$ as $p_n \uparrow\uparrow D$.
(2) For $p \in (0, D)$, we have $\lambda_1^{(0,p)} > \lambda_0^{(0,p)}$.

Proof (a) Let $g (\neq 0)$ be an a.e. eigenfunction of $\lambda_0^{(0,p)}$. Then $g'(0) = 0$, $g(p) = 0$, and $Lg = -\lambda_0^{(0,p)}g$ a.e. on $(0, p)$ by Lemma 8 (1). Moreover

$$\lambda_0^{(0,p)} = \frac{D_{0,p}(g)}{\|g\|_{L^2(0,p;\mu)}^2}, \quad D_{\alpha,\beta}(f) = \int_{\alpha}^{\beta} a f'^2 d\mu.$$

By Lemma 9, the proof of the first assertion in part (1) will be done once we choose a function $\tilde{g} \in \mathcal{A}[0, q]$ such that $\tilde{g}'(0) = 0$, $\tilde{g}(q) = 0$, and

$$\frac{D_{0,p}(g)}{\|g\|_{L^2(0,p;\mu)}^2} > \frac{D_{0,q}(\tilde{g})}{\|\tilde{g}\|_{L^2(0,q;\mu)}^2} \quad \left(\geq \lambda_0^{(0,q)} \right). \quad (8)$$

To do so, without loss of generality, assume that $g|_{(0,p)} > 0$ (this is a well-known property as a reverse of the DN-case for finite intervals, cf. [4; Theorem 3.7]). Then the required assertion follows for

$$\tilde{g}(x) = (g + \varepsilon)\mathbb{1}_{[0,p)}(x) + \frac{\varepsilon(x-q)}{(p-q)}\mathbb{1}_{[p,q]}(x), \quad x \in [0, q],$$

once ε is sufficiently small. Actually, by simple calculation, we have

$$\begin{aligned} D_{0,q}(\tilde{g}) &= D_{0,p}(g) + \frac{\varepsilon^2}{(p-q)^2} \int_p^q e^{C(x)} dx, \\ \|\tilde{g}\|_{L^2(0,q;\mu)}^2 &= \|g\|_{L^2(0,p;\mu)}^2 + \varepsilon \int_0^p (2g + \varepsilon) d\mu + \frac{\varepsilon^2}{(p-q)^2} \int_p^q (x-q)^2 d\mu. \end{aligned}$$

Thus, (8) holds if and only if

$$\begin{aligned} &\frac{\varepsilon}{(p-q)^2} \int_p^q e^C dx \|g\|_{L^2(0,p;\mu)}^2 \\ &< \left(\int_0^p (2g + \varepsilon) d\mu + \frac{\varepsilon}{(p-q)^2} \int_p^q (x-q)^2 d\mu \right) D_{0,p}(g). \end{aligned}$$

Since $\lambda_0^{(0,p)} = D_{0,p}(g)/\|g\|_{L^2(0,p;\mu)}^2$, it suffices to show that

$$\frac{\varepsilon}{(p-q)^2} \int_p^q e^C dx < \lambda_0^{(0,p)} \left(2 \int_0^p g d\mu \right),$$

which is obvious for sufficiently small ε .

The second assertion in part (1) has just been proved at the end of the last proof.

(b) Part (2) of the Lemma strengthens in the present situation a general result that $\lambda_1 \geq \lambda_0$ proved in [2; Proposition 3.2]. Let $g \neq \text{constant}$ be an a.e. eigenfunction of $\lambda_1^{(0,p)}$. Then $g'(0) = 0$, $g'(p) = 0$ and $Lg = -\lambda_1^{(0,p)}g$ a.e. on $(0, p)$ by Lemma 8 (1). Moreover

$$\lambda_1^{(0,p)} = \frac{D_{0,p}(g)}{\text{Var}_{(0,p)}(g)}, \quad \text{Var}_{(\alpha,\beta)}(f) = \int_{\alpha}^{\beta} f^2 d\mu - \frac{\mu_{\alpha,\beta}(f)^2}{\mu(\alpha,\beta)}.$$

Without loss of generality, assume that g is strictly increasing (cf. [5; Proposition 6.4]). Then we have

$$\tilde{g}(x) := g(p) - g(x) > 0 \quad \text{on } (0, p).$$

Thus, $\tilde{g}'(0) = 0$, $\tilde{g}(p) = 0$ and moreover ³

$$\lambda_1^{(0,p)} = \frac{D_{0,p}(\tilde{g})}{\text{Var}_{(0,p)}(\tilde{g})} = \frac{D_{0,p}(\tilde{g})}{\|\tilde{g}\|_{L^2(0,p;\mu)}^2 - \mu_{0,p}(\tilde{g})^2/\mu(0,p)} > \frac{D_{0,p}(\tilde{g})}{\|\tilde{g}\|_{L^2(0,p;\mu)}^2} \geq \lambda_0^{(0,p)}. \quad \square$$

Before moving further, let us mention a nice expression of L :

$$L = \frac{d}{d\mu} \frac{d}{d\nu}$$

which can be checked by a simple computation. Next, a large part of the results in the last section is related to the Poisson equation $Lg = -f$, a.e., from which we obtain

$$\frac{d}{d\nu}g(\beta) - \frac{d}{d\nu}g(\alpha) = - \int_{\alpha}^{\beta} f d\mu, \quad \alpha, \beta \in [0, D], \alpha < \beta. \quad (9)$$

Furthermore, if $g'(\alpha) = 0$, then we have

$$g(q) - g(p) = - \int_p^q \nu(d\beta) \int_{\alpha}^{\beta} f d\mu, \quad p, q \in [0, D], p < q. \quad (10)$$

Especially, because

$$\frac{d}{d\nu}g(0) = e^{C(0)}g'(0) = 0,$$

and (9), with $f = \lambda_0 g$, it follows that

$$\frac{d}{d\nu}g(s) = -\lambda_0 \int_0^s g d\mu, \quad s \in (0, D). \quad (11)$$

Lemmas 12 – 14 given below consist of the basis of the test functions used in the definitions of $\mathcal{F}_{\#}$ and $\mathcal{H}$.

Lemma 12. *Let g be a non-zero a.e. eigenfunction of $\lambda_0 > 0$. Then g is strictly monotone.*

Proof Because $\lambda_0 > 0$, g can not be a constant. We need only to prove that $g' \neq 0$ on $(0, D)$. Suppose that there is a $p \in (0, D)$ such that $g'(p) = 0$. Then, by the eigenequation restricted to $(0, p)$, we would have $\lambda_0 \geq \lambda_1^{(0,p)}$, where $\lambda_1^{(0,p)}$ is the minimal eigenvalue with Neumann boundaries at 0 and p . To see this, by (11), we have $\mu_{0,p}(g) = 0$ since $g'(0) = 0$ and $g'(p) = 0$. Here, it is

³Note that $D_{0,p}(\tilde{g}) = D_{0,p}(g)$ and $\text{Var}_{(0,p)}(\tilde{g}) = \text{Var}_{(0,p)}(g)$.

quite standard to prove the required assertion. By making inner product with g on the both sides of the eigenequation with respect to $\mu_{0,p}$, it follows that

$$-(e^C g g')|_0^p + D_{0,p}(g) = \lambda_0 \mu_{0,p}(g^2).$$

Again, because of $g'(0) = g'(p) = 0$, we obtain $\lambda_0 = D_{0,p}(g)/\mu_{0,p}(g^2)$. Hence,

$$\begin{aligned} \lambda_0 &= \frac{D_{0,p}(g)}{\mu_{0,p}(g^2)} \\ &= \frac{D_{0,p}(g)}{\text{Var}_{(0,p)}(g)} \quad (\text{since } \mu_{0,p}(g) = 0) \\ &\geq \inf \left\{ \frac{D_{0,p}(f)}{\text{Var}_{(0,p)}(f)} : f \in \mathcal{C}^1(0,p) \cap \mathcal{C}[0,p], f \in L^2(0,p;\mu), f \neq \text{constant} \right\} \\ &= \lambda_1^{(0,p)}. \end{aligned}$$

Now, by Lemma 11, we obtain

$$\lambda_0 \geq \lambda_1^{(0,p)} > \lambda_0^{(0,p)} > \lambda_0.$$

This is a contradiction. $\square$

Lemma 13. *The a.e. eigenfunction g of λ_0 is either positive or negative everywhere.*

Proof If $\lambda_0 = 0$, then g must be a constant and so the assertion is obvious. Now, let $\lambda_0 > 0$. By Lemma 12, without loss of generality, assume that $g'|_{(0,D)} < 0$ and $g(0) > 0$ ⁴. We need only to prove that $g \neq 0$ on $(0,D)$. If otherwise $g(p) = 0$ for some $p \in (0,D)$, then, since $\lambda_0^{(0,p)}$ is the minimal ND-eigenvalue on $(0,p)$, the eigenequation restricted to $(0,p)$ shows that⁵

$$\lambda_0 \geq \lambda_0^{(0,p)} > \lambda_0,$$

which is a contradiction. $\square$

Because of (11), we have $I(g)^{-1} \equiv \lambda_0$. This explains where the operator I comes from. Next, from (10), we have

$$g(x) - g(D) = \lambda_0 \int_x^D \nu(ds) \int_0^s g d\mu. \quad (12)$$

When $D < \infty$, since $g(D) = 0$ by our boundary condition, we obtain $II(g)^{-1} \equiv \lambda_0$. This explains the meaning of the operator II . To show that the last assertion holds even for $D = \infty$, it is necessary to prove that $g(\infty) = 0$. This is impossible if $\lambda_0 = 0$ since then g can be an arbitrary non-zero constant.

⁴About $g(0) > 0$. Since $Lg = -\lambda_0 g$ on $(0,D)$, by (11), we have $dg(s)/d\nu = -\lambda_0 \int_0^s g d\mu$, $s \in (0,D)$. So $\int_0^s g d\mu > 0$ by $g' < 0$. This implies that $g(0) > 0$. Otherwise, one would get a contradiction with $\int_0^s g d\mu > 0$ since $g' < 0$ and then $g \leq 0$.

⁵The proof of $\lambda_0 \geq \lambda_0^{(0,p)}$ is similar to the one of $\lambda_0 \geq \lambda_1^{(0,p)}$ given in the last proof.

Lemma 14. *Let $D = \infty$. If $\lambda_0 > 0$, then its a.e. eigenfunction g satisfies $g(\infty) = 0$.*

Proof Without loss of generality, by Lemmas 12 and 13, assume that $g'|_{(0,D)} < 0$ and $g|_{[0,D)} > 0$.

(a) By what we have just seen and the decreasing property of g , we have

$$\frac{g(x) - g(\infty)}{\lambda_0} = \int_x^\infty \nu(ds) \int_0^s g d\mu \geq g(\infty) \int_x^\infty \nu(ds) \int_0^s d\mu.$$

Thus, $g(\infty) = 0$ once $\int_0^\infty \nu(ds) \int_0^s d\mu = \infty$ (which is the uniqueness criterion for the semigroup or the nonexplosive criterion for the minimal process) since the left-hand side is finite.

(b) Otherwise, we have

$$M(x) := \int_x^\infty \nu(ds) \int_0^s d\mu < \infty, \quad x \in (0, D).$$

Let $f = g - g(\infty)$ and suppose that $g(\infty) > 0$. Then $f \in \mathcal{F}_H$ and moreover,

$$fH(f)(x) = \lambda_0^{-1}(g(x) - g(\infty)) - g(\infty)M(x) = \lambda_0^{-1}f(x) - g(\infty)M(x).$$

We arrive at

$$\sup_{x \in (0, \infty)} H(f)(x) = \frac{1}{\lambda_0} - g(\infty) \inf_{x \in (0, \infty)} \frac{M(x)}{f(x)}.$$

Since $f(\infty) = 0$ and $M(\infty) = 0$, by Cauchy's mean value theorem, we have

$$\begin{aligned} & \inf_{x \in (0, \infty)} \frac{M(x)}{f(x)} \\ & \geq \inf_{x \in (0, \infty)} \frac{M'(x)}{f'(x)} \\ & = \inf_{x \in (0, \infty)} -\frac{e^{-C(x)}}{g'(x)} \int_0^x \frac{e^{C(u)}}{a(u)} du \\ & \geq \inf_{x \in (0, \infty)} \left(-\frac{e^{-C(x)}}{g(0)g'(x)} \int_0^x g d\mu \right) \quad (\text{since } g' < 0 \text{ and } g > 0 \text{ on } (0, D)) \\ & = \inf_{x \in (0, \infty)} \frac{1}{g(0)} I(g)(x) \\ & = \frac{1}{\lambda_0 g(0)} \\ & > 0. \end{aligned}$$

Inserting this into the previous equation, it follows that $\lambda_0 < \inf_{x \in (0, \infty)} H(f)(x)^{-1}$. But $\inf_{x \in (0, \infty)} H(f)(x)^{-1} \leq \lambda_0$ is a part of Theorem 1 (2) and will be proved soon below, without using the properties of the a.e. eigenfunction g . We have thus obtained a contradiction. $\square$

From now on in this section, we assume that the a.e. eigenfunction (say g) satisfies $g > 0$ and $g' < 0$ on $(0, D)$, $g'(0) = 0$, and $g(D) = 0$ (recall that $g(D) = \lim_{x \rightarrow D} g(x)$ if $D = \infty$).

Proof of Theorem 1 and Proposition 2 Similar to the proof of [5; Theorem 2.4 and Proposition 2.5], we can prove the assertions by two circle arguments.

To prove the lower estimates, we adopt the following circle arguments: ⁶

$$\begin{aligned} \lambda_0 &\geq \tilde{\lambda}_0 \geq \sup_{f \in \mathcal{F}_I} \inf_{x \in (0, D)} I(f)(x)^{-1} \\ &= \sup_{f \in \mathcal{F}_I} \inf_{x \in (0, D)} II(f)(x)^{-1} = \sup_{f \in \mathcal{F}_{II}} \inf_{x \in (0, D)} II(f)(x)^{-1} \end{aligned} \quad (13)$$

$$\geq \sup_{h \in \mathcal{H}} \inf_{x \in (0, D)} R(h)(x) \geq \lambda_0. \quad (14)$$

For the upper estimation of λ_0 , we adopt the following circle arguments: ⁷

$$\lambda_0 \leq \inf_{f \in \mathcal{F}_{II} \cup \tilde{\mathcal{F}}'_{II}} \sup_{x \in \text{supp}(f)} II(f)(x)^{-1} = \inf_{f \in \tilde{\mathcal{F}}_{II}} \sup_{x \in \text{supp}(f)} II(f)(x)^{-1} \quad (15)$$

$$\begin{aligned} &= \inf_{f \in \tilde{\mathcal{F}}_I} \sup_{x \in \text{supp}(f)} II(f)(x)^{-1} = \inf_{f \in \tilde{\mathcal{F}}_I} \sup_{x \in (0, D)} I(f)(x)^{-1} \\ &= \inf_{f \in \tilde{\mathcal{F}}'_I} \sup_{x \in (0, D)} I(f)(x)^{-1} \end{aligned} \quad (16)$$

$$\leq \inf_{h \in \mathcal{H}} \sup_{x \in (0, D)} R(h)(x) \leq \lambda_0. \quad (17)$$

In fact, most of the proof here are parallel to those in the discrete case (see [5; Section 2]). Actually, one can follow the cited proofs with some changes illustrated here. For instance, to prove

$$\tilde{\lambda}_0 \geq \sup_{f \in \mathcal{F}_{II}} \inf_{x \in (0, D)} II(f)(x)^{-1},$$

following [5; Part I (a) of the proof of Theorem 2.4 and Proposition 2.5], let g (irrelated to the eigenfunction) be a test function of $\tilde{\lambda}_0$: $g \in \mathcal{C}^1(0, D) \cap \mathcal{C}[0, D]$, $g(D) = 0$, $g'(0) = 0$ and $\mu(g^2) = 1$. Then for every h with $h|_{(0, D)} > 0$, we have

$$\begin{aligned} 1 &= \mu(g^2) = \int_0^D \frac{e^{C(x)}}{a(x)} \left(\int_x^D g'(t) dt \right)^2 dx \\ &\leq \int_0^D \frac{e^{C(x)}}{a(x)} dx \int_x^D \frac{e^{C(t)}}{h(t)} g'(t)^2 dt \int_x^D \frac{h(s)}{e^{C(s)}} ds \quad (\text{by Cauchy-Schwarz's inequality}) \\ &= \int_0^D \frac{e^{C(t)}}{h(t)} g'(t)^2 dt \int_0^t \frac{e^{C(x)}}{a(x)} dx \int_x^D \frac{h(s)}{e^{C(s)}} ds \quad (\text{by Fubini's Theorem}) \\ &\leq D(g) \sup_{t \in (0, D)} \frac{1}{h(t)} \int_0^t \frac{e^{C(x)}}{a(x)} dx \int_x^D \frac{h(s)}{e^{C(s)}} ds \\ &=: D(g) \sup_{t \in (0, D)} H(t). \end{aligned}$$

⁶The details are given in Appendix A.1.

⁷The details are given in Appendix A.2.

For $f \in \mathcal{F}_H$ satisfying $\sup_{x \in (0, D)} H(f)(x) < \infty$, we specify

$$h(t) = \int_0^t a(s)^{-1} e^{C(s)} f(s) ds.$$

Then by Cauchy's mean value theorem, it follows that

$$\sup_{t \in (0, D)} H(t) \leq \sup_{x \in (0, D)} \frac{1}{f(x)} \int_x^D e^{-C(s)} ds \int_0^x \frac{e^{C(u)}}{a(u)} f(u) du = \sup_{x \in (0, D)} H(f)(x).$$

Hence,

$$\inf_{x \in (0, D)} H(f)(x)^{-1} \leq \inf_{t \in (0, D)} H(t)^{-1} \leq D(g).$$

Making infimum with respect to g , we obtain the required assertion. We have also completed the proof of Lemma 14.

As mentioned before Lemma 14, the operators I and II are all from the eigenequation. Here we show that so is the operator R . Rewrite the eigenequation as

$$-\frac{Lg}{g} = \lambda_0$$

which is meaningful since $g > 0$. To simplify the left-hand side, in the discrete case, one uses the ratio $g(x+1)/g(x)$. However, this is useless in the present continuous situation. What instead is using the function $h = g'/g$. Then

$$-\frac{Lg}{g} = -(ah^2 + bh + ah') = R(h).$$

The conditions $g > 0$ and $g' < 0$ on $(0, D)$ lead to the restraint $h|_{(0, D)} < 0$ in defining $\mathcal{H}$. Note that the inverse transform $h \rightarrow g$ is unique up to a positive constant:

$$g(x) = \exp \left[\int_0^x h(u) du \right].$$

The restraint allowing $h = 0$ in the definition of $\mathcal{H}$ is to include the degenerated case that $g' \equiv 0$ when $\lambda_0 = 0$ (then $D = \infty$ by hypothesis (1)). Clearly, the use of R is essentially the use of L . For this reason, we make the continuous condition on a and b once concerning with R . Because of this point, we need two additional terms in the circle arguments above: the right-hand side of (13) is not less than λ_0 and the right-hand side of (16) is no more than λ_0 . This is rather easy since for the a.e eigenfunction g , we have $I(g)^{-1} \equiv \lambda_0$ and $II(g)^{-1} \equiv \lambda_0$ by (11), (12), and Lemma 14. Actually, the required assertion was also contained in the corresponding proof of the discrete situation.

As another illustration of the proof when moving from the discrete case to the continuous one, we consider a proof for the upper estimates. For instance, we prove that

$$\lambda_0 \leq \inf_{f \in \tilde{\mathcal{F}}_H \cup \tilde{\mathcal{F}}'_H} \sup_{x \in \text{supp}(f)} H(f)(x)^{-1}.$$

Before moving to the details, let us mention that, for the upper estimates of λ_0 , we are actually using a comparison between λ_0 and $\lambda_0^{(0,x_0)}$. Thus, for the upper estimates of λ_0 , we indeed use the restriction on $[0, x_0]$ for the test functions, ignoring their behaviors out of $[0, x_0]$.

Given $f \in \widetilde{\mathcal{F}}_H$ with $f = f\mathbb{1}_{[0,x_0)}$ for some $x_0 \in (0, D)$, let

$$g = fH(f)\mathbb{1}_{\text{supp}(f)}.$$

Then $g \in L^2(\mu)$. Since

$$(e^C g')(x) = - \int_0^x f \, d\mu \quad \text{on } [0, x_0],$$

by the integration by parts formula, we have

$$\begin{aligned} D(g) &= \int_0^D e^{C(x)} g'(x)^2 dx = - \int_0^{x_0} (dg(x)) \int_0^x f \, d\mu \\ &= - \int_0^{x_0} f(t) \int_t^{x_0} dg(t) \mu(dt) \quad (\text{by Fubini's Theorem}) \\ &= \int_0^{x_0} f(t)(g(x_0) - g(t)) \mu(dt) \\ &= \int_0^{x_0} fg \, d\mu \quad (\text{since } g(x_0) = 0). \end{aligned}$$

Hence

$$D(g) \leq \int_0^{x_0} g^2 \, d\mu \sup_{(0,x_0)} \frac{f}{g} = \mu(g^2) \sup_{x \in (0,x_0)} H(f)(x)^{-1}.$$

Since $g \in L^2(\mu)$, it follows that

$$\lambda_0 \leq \frac{D(g)}{\mu(g^2)} \leq \sup_{\text{supp}(f)} H(f)^{-1} \quad (18)$$

for every $f \in \widetilde{\mathcal{F}}_H$. It remains to show that the same assertion holds for every $f \in \widetilde{\mathcal{F}}'_H$. Recall that in the proof above, the conclusion $g \in L^2(\mu)$ comes from the finiteness of x_0 . Otherwise, if $x_0 = D = \infty$, then $f \in \widetilde{\mathcal{F}}'_H$ means that the function $g = fH(f)$ is assumed to be in $L^2(\mu)$, and the proof above still works. So we obtain the required assertion.

Hopefully, we have explained enough the difference between the discrete and the continuous cases. Now, one may follow [5; Proof of Theorem 2.4 and Proposition 2.5] (quite long and technical) to complete the whole proof. $\square$

Before moving further, let us mention a fact about the localizing procedures used in Theorem 6 (2). Instead of the approximating to the infinite state space ($D = \infty$) by finite ones, it seems more natural to use the truncating procedure for the test function f : $f^{(n)} = f\mathbb{1}_{[0,x_n)}$ with $x_n \uparrow \infty$. The next result shows that such a procedure is not practical in general.

Remark 15. Assume that hypothesis (1) holds. Let $D = \infty$, let g be the eigenfunction of $\lambda_0 > 0$, define $g^{(n)} = g\mathbb{1}_{[0, x_n]}$ for some $x_n \in (0, \infty)$. Then

$$\inf_{x \in \text{supp}(g^{(n)})} II(g^{(n)})(x) = 0.$$

In particular, $\inf_{x \in \text{supp}(g^{(n)})} II(g^{(n)})(x)$ does not converge to λ_0 as $x_n \rightarrow \infty$.

Proof By definition of $g^{(n)}$, we have

$$\begin{aligned} \inf_{x \in \text{supp}(g^{(n)})} II(g^{(n)})(x) &= \inf_{x \in [0, x_n]} \frac{1}{g^{(n)}(x)} \int_x^{x_n} e^{-C(s)} ds \int_0^s g^{(n)} d\mu \\ &= \inf_{x \in [0, x_n]} \frac{1}{g(x)} \int_x^{x_n} e^{-C(s)} ds \int_0^s g d\mu \\ &= \inf_{x \in [0, x_n]} \frac{1}{g(x)} \int_x^{x_n} (-\lambda_0^{-1} g'(s)) ds \quad (\text{by (11)}) \\ &= \inf_{x \in [0, x_n]} \frac{1}{\lambda_0 g(x)} (g(x) - g(x_n)) \\ &= \inf_{x \in [0, x_n]} \frac{1}{\lambda_0} \left(1 - \frac{g(x_n)}{g(x)} \right) \\ &= 0 \quad (\text{since } g \in \mathcal{C}[0, D] \text{ and } g \downarrow\downarrow). \quad \square \end{aligned}$$

Proof of Proposition 3 (1) Let $g \in \mathcal{C}^1(0, D)$ with $g > 0$ and $g' < 0$ on $(0, D)$, and let

$$\bar{h}(x) = -e^{-C(x)} \int_0^x g d\mu.$$

Then ⁸ $\bar{h} \in \mathcal{H}_*$ and

$$\overline{R}(\bar{h})(x) = -\frac{(a\bar{h}' + b\bar{h})'(x)}{\bar{h}(x)} = \frac{g'(x)}{\bar{h}(x)} > 0.$$

This clearly implies that $\sup_{h \in \mathcal{H}_*} \inf_{x \in (0, D)} \overline{R}(h)(x) \geq 0$.

(2) Without loss of generality, assume that $\lambda_0 > 0$. Since $a, b \in \mathcal{C}^1(0, D)$, there exists an eigenfunction g such that ⁹ $\bar{h} := g' \in \mathcal{H}_*$ and

$$\overline{R}(\bar{h})(x) = \frac{-(Lg)'(x)}{g'(x)} \equiv \lambda_0.$$

⁸ $\bar{h}(0) = 0$, $\bar{h}' = -bh/a - g/a < -bh/a$. So $\bar{h} \in \mathcal{C}^2(0, D) \cap \mathcal{C}[0, D]$. Thus, $\bar{h} \in \mathcal{H}_*$. Moreover, $a\bar{h}'' + b\bar{h}' = -g < 0$.

⁹We have $g'(0) = 0$, $g \in \mathcal{C}^2[0, D]$, $g > 0$ and $g' < 0$ on $(0, D)$ by Lemmas 8, 12 and 13. Since $\bar{h} = g'$ and $a, b \in \mathcal{C}[0, D] \cap \mathcal{C}^1(0, D)$, we can see that $\bar{h}(0) = 0$ and

$$h' = g'' = -\frac{\lambda_0 g + bg'}{a} \in \mathcal{C}^1(0, D), \quad a\bar{h}' + b\bar{h} = ag'' + bg' = -\lambda_0 g < 0.$$

So $\bar{h} \in \mathcal{C}^2(0, D) \cap \mathcal{C}[0, D]$, $\bar{h}' < -a^{-1}b\bar{h}$ and then $\bar{h} \in \mathcal{H}_*$.

Thus

$$\sup_{h \in \mathcal{H}} \inf_{x \in (0, D)} \overline{R}(h)(x) \geq \sup_{h \in \mathcal{H}_*} \inf_{x \in (0, D)} \overline{R}(h)(x) \geq \lambda_0.$$

Now, one can complete the proof following that of the discrete case ([5; Proof of Proposition 2.7])¹⁰. $\square$

To prove Theorem 5, we need the following result.

Lemma 16. *Given two nonnegative, measurable, and locally integrable functions m and n on $[0, D]$, suppose that*

$$\int_0^D n(y) dy < \infty \quad \text{and} \quad c := \sup_{x \in (0, D)} \int_0^x m(y) dy \int_x^D n(y) dy < \infty.$$

Set $\psi(x) = \int_x^D n(y) dy$. Then for every $r \in (0, 1)$, we have

$$\int_0^x m(y) \psi^r(y) dy \leq \frac{c}{1-r} \psi^{r-1}(x), \quad x \in (0, D).$$

*Proof*¹¹ Let $M(x) = \int_0^x m(y) dy$. Noticing that $M'(x) = m(x)$ and $M\psi \leq c$, we obtain the assertion by using the integration by parts formula. $\square$

Proof of Theorem 5 To prove the lower estimate, without loss of generality, assume that $\delta < \infty$. Applying Lemma 16 to $m(x) = e^{C(x)}/a(x)$ and $n(x) = e^{-C(x)}$, we get

$$\int_0^x \varphi^r(y) \mu(dy) = \int_0^x \varphi^r(y) m(y) dy \leq \frac{\delta}{1-r} \varphi^{r-1}(x), \quad x \in (0, D).$$

Put $f = \varphi^r$. Then $f \in \mathcal{F}_I$ and $I(f)(x) \leq \delta/(r - r^2)$. Optimizing the inequality with respect to r , it follows that

$$I(f)(x) \leq \inf_{0 < r < 1} \frac{\delta}{(r - r^2)} = 4\delta. \quad (19)$$

We have thus proved the lower estimate.

For the upper estimate¹², we choose the test function as $f = \nu(x_0 \vee \cdot, x_1) \mathbb{1}_{[0, x_1)}$ for some $x_0, x_1 \in [0, D]$ with $x_0 < x_1$. Then, the assertion follows by using either the variational formula for upper estimate given by Theorem 1(1):

$$\lambda_0 \leq \inf_{f \in \mathcal{F}_I} \sup_{x \in (0, D)} I(f)(x)^{-1}$$

or the classical variational formula:

$$\lambda_0^{-1} = \lambda_*^{-1} \geq \sup_{x_0, x_1: x_0 < x_1} \frac{\|f_1^{x_0, x_1}\|}{D(f_1^{x_0, x_1})}$$

¹⁰The details are given in Appendix A.3.

¹¹The details are given in Appendix A.4.

¹²The details are given in Appendix A.5.

and then letting $x_1 \rightarrow D$.

At last, if $\nu(0, D) = \infty$, then we have $\nu(x, D) = \infty$ because of hypothesis (1). Furthermore, $\mu(0, x)\nu(x, D) = \infty$ for every $x \in (0, D)$. Therefore, $\delta = \infty$ and $\lambda_0 = 0$. If

$$\int_0^\infty \mu(0, x) \nu(dx) < \infty,$$

then for each $x \in (0, D)$, we have

$$\mu(0, x)\nu(x, D) = \int_x^D \mu(0, x)\nu(dt) < \int_x^\infty \mu(0, t)\nu(dt) < \int_0^\infty \mu(0, x) \nu(dx) < \infty.$$

Hence, $\delta < \infty$ and $\lambda_0 > 0$. $\square$

Proof of Theorem 6 and Corollary 7 ¹³ Simply follow [5; Proof of Theorem 3.2 and Corollary 3.3]. We mention that the proof of “ $\delta'_1 \leq 2\delta$ ” and the computation of δ'_1 are not easy. $\square$

4 DN-case

We now turn to study the DN-case. As Section 2, we use the same notation $\mathcal{C}[0, D]$, $\mathcal{C}^k(0, D)$ and the operator L . The main different point for the eigenequation $Lg = -\lambda_0 g$ is the boundary condition: $g(0) = 0$ and $g'(D) = 0$ if $D < \infty$. Now define

$$\lambda_0 = \inf \left\{ \frac{D(f)}{\mu(f^2)} : f \in \mathcal{C}^1(0, D) \cap \mathcal{C}[0, D], D(f) < \infty, f(0) = 0, f \neq 0 \right\}, \quad (20)$$

where

$$D(f) = \int_0^D a f'^2 d\mu, \quad \mu(dx) = \frac{e^{C(x)}}{a(x)} dx, \quad C(x) = \int_0^x \frac{b(u)}{a(u)} du.$$

Again, define $\nu(dx) = e^{-C(x)} dx$. Here, we have used the hypothesis (1). The restraint “ $D(f) < \infty$ ” in (20) is to avoid ∞/∞ since we allow $\mu(f^2) = \infty$. Then the restraint “ $f \neq 0$ ” is needed to avoid $0/0$. Note that the restriction on the set $\mathcal{C}_K$ for test functions disappears in (20). This means that the maximal Dirichlet form or the maximal process is used here, instead of the minimal one used in Section 2. In other words, we do not assume the uniqueness of the semigroup, which is different from what we studied earlier in [1–4] and [8]. The constant λ_0 defined above describes the optimal constant $C = \lambda_0^{-1}$ in the following *weighted Hardy inequality* :

$$\mu(f^2) \leq CD(f), \quad f(0) = 0.$$

(See [4; Section 5.2]). In other words, we are studying the weighted Hardy inequality in this section. To save the notation, we use the same notation λ_0 ,

¹³The details are given in Appendix A.6.

I , II , R and so on as before, each of them plays a similar role but may have different meaning in different context.

Before going to our main text, we note that in definition of λ_0 , one may replace $\mathcal{C}^1(0, D) \cap \mathcal{C}[0, D]$ by $\mathcal{A}[0, D]$ as shown by Lemma 9.

Now, we review some notation defined originally in [3, 8] and introduce some new ones as follows.

$$\begin{aligned} I(f)(x) &= \frac{e^{-C(x)}}{f'(x)} \int_x^D f \, d\mu \quad (\text{single integral form}), \\ II(f)(x) &= \frac{1}{f(x)} \int_0^x \nu(ds) \int_s^D f \, d\mu \quad (\text{double integral form}), \\ R(h)(x) &= -(ah^2 + bh + ah')(x) \quad (\text{differential form}). \end{aligned}$$

The domains of I , II and R , respectively, are as follows.

$$\begin{aligned} \mathcal{F}_I &= \{f : f \in \mathcal{C}^1(0, D) \cap \mathcal{C}[0, D], f(0) = 0, \text{ and } f'|_{(0, D)} > 0\}, \\ \mathcal{F}_{II} &= \{f : f \in \mathcal{C}[0, D], f(0) = 0, \text{ and } f|_{(0, D)} > 0\}, \\ \mathcal{H} &= \left\{h : h \in \mathcal{C}^1(0, D) \cap \mathcal{C}[0, D], h|_{(0, D)} > 0, \text{ and } \int_{0+} h(u) du = \infty\right\}, \end{aligned}$$

where $\int_{0+}$ means $\int_0^\varepsilon$ for sufficiently small $\varepsilon > 0$. These sets ¹⁴ are used for the estimates on lower bounds of λ_0 . For the upper bounds, we have the following domains:

$$\begin{aligned} \widetilde{\mathcal{F}}_I &= \{f : \exists x_0 \in (0, D), f \in \mathcal{C}^1(0, x_0) \cap \mathcal{C}[0, D], f(0) = 0, f = f(\cdot \wedge x_0), \\ &\quad \text{and } f'|_{(0, x_0)} > 0\}, \\ \widetilde{\mathcal{F}}_{II} &= \{f : \exists x_0 \in (0, D), f \in \mathcal{C}[0, x_0], f(0) = 0, f = f(\cdot \wedge x_0) \text{ and } f|_{(0, x_0)} > 0\}, \\ \widetilde{\mathcal{H}} &= \left\{h : \exists x_0 \in (0, D), h \in \mathcal{C}^1(0, x_0) \cap \mathcal{C}[0, D], h|_{(0, x_0)} > 0, \int_{0+} h(u) du = \infty, \right. \\ &\quad \left. h|_{[x_0, D]} = 0, \text{ and } \sup_{(0, x_0)} (ah^2 + bh + ah') < 0\right\}. \end{aligned}$$

Besides, we also need

$$\widetilde{\mathcal{F}}'_{II} = \{f : f > 0, f \in \mathcal{C}[0, D], \text{ and } fII(f) \in L^2(\mu)\}.$$

Under hypothesis (1), if $\mu(0, D) = \infty$, then λ_0 defined by (20) is trivial. Indeed, let

$$f = \mathbb{1}_{(\delta, D]} + h\mathbb{1}_{[0, \delta]},$$

where h is chosen such that $h(0) = 0$ and $f \in \mathcal{C}^1(0, D) \cap \mathcal{C}[0, D]$ (for example, $h(x) = -x^2 \cdot \delta^{-2} + 2x \cdot \delta^{-1}$). Then $D(f) \in (0, \infty)$ and $\mu(f^2) = \infty$. It follows that $\lambda_0 = 0$.

¹⁴About the set of $\mathcal{H}$: The functions g and h are one to one: $h = g'/g$ and conversely $g(x) = g(\varepsilon) \exp[\int_\varepsilon^x h(u) du]$, $x \in (0, D)$. So $g(0) = 0$ implies $h(0) = \infty$ and conversely $g(0) = 0$ is implied by $\int_{0+} h(u) du = \infty$.

Otherwise, $\mu(0, D) < \infty$. Then for every f with $\mu(f^2) = \infty$, by setting $f^{(x_0)} = f(\cdot \wedge x_0) \in L^2(\mu)$, we have

$$\infty > D(f^{(x_0)}) \uparrow D(f), \quad \infty > \mu(f^{(x_0)^2}) \rightarrow \mu(f^2) \quad \text{as } x_0 \rightarrow D.$$

In other words, for each non-square-integrable function f , both $\mu(f^2)$ and $D(f)$ can be approximated by a sequence of square-integrable ones. Hence, we can rewrite λ_0 as follows.

$$\lambda_0 = \inf \{D(f) : \mu(f^2) = 1, f(0) = 0, \text{ and } f \in \mathcal{C}^1(0, D) \cap \mathcal{C}[0, D]\}. \quad (21)$$

In this case, as will be seen soon but not obvious, we also have

$$\lambda_0 = \inf \{D(f) : \mu(f^2) = 1, f(0) = 0, f = f(\cdot \wedge x_0), f \in \mathcal{C}^1(0, x_0) \cap \mathcal{C}[0, x_0] \\ \text{for some } x_0 \in (0, D)\} =: \tilde{\lambda}_0.$$

Now we introduce our main results. Their relations are very much the same as that indicated in Section 2, except that the test function used in Theorem 18 is $\nu(0, x)^\gamma$ but not $\nu(x, D)^\gamma$ ($\gamma = 1/2$ or 1).

Theorem 17. ¹⁵ *Let hypothesis (1) hold. Assume that $\mu(0, D) < \infty$. Then λ_0 defined by (20) or (21) coincides with $\tilde{\lambda}_0$ and the following variational formulas hold.*

(1) *Single integral forms:*

$$\inf_{f \in \tilde{\mathcal{F}}_I} \sup_{x \in (0, D)} I(f)(x)^{-1} = \lambda_0 = \sup_{f \in \mathcal{F}_I} \inf_{x \in (0, D)} I(f)(x)^{-1}.$$

(2) *Double integral forms:*

$$\lambda_0 = \inf_{f \in \tilde{\mathcal{F}}_I} \sup_{x \in (0, D)} II(f)(x)^{-1} = \inf_{f \in \tilde{\mathcal{F}}_{II}} \sup_{x \in (0, D)} II(f)(x)^{-1} = \inf_{f \in \tilde{\mathcal{F}}_{II} \cup \tilde{\mathcal{F}}'_{II}} \sup_{x \in (0, D)} II(f)(x)^{-1} \\ \lambda_0 = \sup_{f \in \mathcal{F}_I} \inf_{x \in (0, D)} II(f)(x)^{-1} = \sup_{f \in \mathcal{F}_{II}} \inf_{x \in (0, D)} II(f)(x)^{-1}.$$

Moreover, if $a, b \in \mathcal{C}[0, D]$, then we also have

(3) *differential forms:*

$$\inf_{h \in \tilde{\mathcal{H}}} \sup_{x \in (0, D)} R(h)(x) = \lambda_0 = \sup_{h \in \mathcal{H}} \inf_{x \in (0, D)} R(h)(x).$$

Theorem 18. (Criterion and basic estimates) *Let hypothesis (1) hold. Then λ_0 defined by (20) (or equivalently, $\tilde{\lambda}_0$ provided $\mu(0, D) < \infty$) is positive if and only if*

$$\delta := \sup_{x \in (0, D)} \nu(0, x) \mu(x, D) < \infty.$$

¹⁵The details of the proof are given in Appendix B.1.

More precisely, we have

$$(4\delta)^{-1} \leq \lambda_0 \leq \delta^{-1}.$$

In particular, we have $\lambda_0 = 0$ if $\mu(0, D) = \infty$, and $\lambda_0 > 0$ if

$$D < \infty \quad \text{or} \quad \int_0^D \left(a(u)^{-1} e^{C(u)} + e^{-C(u)} \right) du < \infty.$$

Proof The result was proved in [2; Theorem 1.1] except the case that $\mu(0, D) = \infty$, which implies $\lambda_0 = 0$ ($\delta = \infty$) and so the assertion is trivial.

Theorem 19. ¹⁶ (Approximating procedure) *Let hypothesis (1) hold. Assume that $\mu(0, D) < \infty$ and $\delta < \infty$. Set $\varphi(x) = \nu(0, x)$ for $x \in (0, D)$.*

(1) *Define $f_1 = \sqrt{\varphi}$, $f_n = f_{n-1} II(f_{n-1})$, $n \geq 2$, and let*

$$\delta_n = \sup_{x \in (0, D)} II(f_n)(x), \quad n \geq 1.$$

Then δ_n is decreasing in n and

$$\lambda_0 \geq \delta_n^{-1} \geq (4\delta)^{-1}, \quad n \geq 1.$$

(2) *For fixed $x_0 \in (0, D)$, define*

$$f_1^{(x_0)} = \varphi(\cdot \wedge x_0), \quad f_n^{(x_0)} = (f_{n-1}^{(x_0)} II(f_{n-1}^{(x_0)}))(\cdot \wedge x_0), \quad n \geq 2,$$

and let

$$\delta'_n = \sup_{x_0 \in (0, D)} \inf_{x \in (0, D)} II(f_n^{(x_0)})(x).$$

Then δ'_n is increasing in n and

$$\delta^{-1} \geq \delta_n'^{-1} \geq \lambda_0, \quad n \geq 1.$$

Next, define

$$\bar{\delta}_n = \sup_{x_0 \in (0, D)} \frac{\|f_n^{(x_0)}\|}{D(f_n^{(x_0)})}, \quad n \geq 1.$$

Then $\bar{\delta}_n^{-1} \geq \lambda_0$, $\bar{\delta}_{n+1} \geq \delta'_n$ for every $n \geq 1$ and $\bar{\delta}_1 = \delta_1$.

Corollary 20. ¹⁷ (Improved estimates) *We have the following estimates:*

$$\delta^{-1} \geq \delta_1'^{-1} \geq \lambda_0 \geq \delta_1^{-1} \geq (4\delta)^{-1},$$

where

$$\begin{aligned} \delta_1 &= \sup_{x \in (0, D)} \frac{1}{\sqrt{\varphi(x)}} \int_0^D \varphi(x \wedge \cdot) \sqrt{\varphi} d\mu \\ &= \sup_{x \in (0, D)} \left(\frac{1}{\sqrt{\varphi(x)}} \int_0^x \varphi^{3/2} d\mu + \sqrt{\varphi(x)} \int_x^D \sqrt{\varphi} d\mu \right), \\ \delta_1' &= \sup_{x \in (0, D)} \frac{1}{\varphi(x)} \int_0^D \varphi(\cdot \wedge x)^2 d\mu \in [\delta, 2\delta]. \end{aligned}$$

¹⁶The details of the proof are given in Appendix B.2.

¹⁷The details of the proof are given in Appendix B.3.

Since the proofs of the results above are either known from [2, 3] or parallel to [5], here we make some remarks only.

Remark 21. (1) As mentioned in [5], the original proofs given in [2, 3] are still suitable to support the idea using the maximal Dirichlet form instead of the uniqueness assumption.

(2) As discussed in the last section, it is natural to extend a and b from continuous to measurable when using operators I and II only.

(3) About the duality. Recall that

$$L = \frac{d}{d\mu} \frac{d}{d\nu}.$$

The dual operator of L is simply defined as

$$L^* = \frac{d}{d\mu^*} \frac{d}{d\nu^*}, \quad \mu^* := \nu, \nu^* := \mu.$$

For the boundaries, simply exchange the names of Dirichlet and Neumann. The basic results for these operators are $\lambda_0(L) = \lambda_0(L^*)$ and $\delta = \delta^*$, where $\lambda_0(L)$ and δ are defined in Section 2, and $\lambda_0(L^*)$ and δ^* are defined in this section replacing L with L^* . The proof goes as follows.

(a) Reduce to finite D . By an approximating procedure we have used many times before, it suffices to prove the assertion for finite D . The point is that for $\lambda_0(L)$, one needs to consider only the test functions having compact support; for $\lambda_0(L^*)$, it suffices to consider the test function $f = f(\cdot \wedge x_0)$, where x_0 varies over $(0, D)$.

(b) By a standard smoothing procedure, one may assume that a and b are smooth.

(c) The identity of $\lambda_0(L)$ and $\lambda_0(L^*)$ is a combination of Proposition 3 (2) and Theorem 17 (3). The discrete case was given in [5; Section 5]. An alternative proof of this assertion was presented in [6] based on isospectrality. Note that in the last proof, the finiteness of D is crucial, otherwise, the domains of L and L^* are essential different unless the Dirichlet form corresponding to L^* is assumed to be regular.

(4) When $D < \infty$, one may simply reverse the variable to obtain one from the other between the ND- and DN- cases. In this sense, the identity $\lambda_0(L) = \lambda_0(L^*)$ stated in (3) is quite natural even though the duality is not a “reverse transform”. When $D = \infty$, these two cases are certainly different since the Dirichlet boundary at 0 is touchable but not the one at ∞ . We mention that the variational formulas and then the approximating procedure in this section are different from those deduced by the dual approach. It is interesting that in the discrete situation, the approximating procedure given by Theorem 19 is often less powerful than those given by Theorem 6 in terms of duality.

5 Supplement to NN-case

Everything is the same as those in the last section except the mixed eigenvalue λ_0 is replaced by

$$\lambda_1 = \inf \{D(f) : \mu(f) = 0, \mu(f^2) = 1, \text{ and } f \in \mathcal{C}^1(0, D) \cap \mathcal{C}[0, D]\}. \quad (22)$$

Let us repeat that throughout this section, we assume that hypothesis (1) holds and $\mu(0, D) < \infty$.

The supplement consists of three parts. The first one is using the maximal Dirichlet form instead of the uniqueness assumption of the semigroup. The second one is using the “a.e. eigenfunction” instead of “eigenfunction”. These two parts have already been studied in the previous sections. See also [8] for some supplement to the original paper. The third part is about the monotonicity of an approximating procedure which we are going to study below.

Define

$$\bar{f} = f - \pi(f), \quad f_1 = \sqrt{\varphi}, \quad f_n = \bar{f}_{n-1} II(\bar{f}_{n-1}), \quad \eta_n = \sup_{x \in (0, D)} I(\bar{f}_n)(x),$$

where $\pi = \mu/\mu(0, D)$. Here our main question is about the monotonicity of $\{\eta_n\}$. Unlike the sequences $\{\delta_n\}$ and $\{\delta'_n\}$ defined in Theorems 6 and 19, their monotonicity results from simply twice applications of Cauchy’s mean value theorem, the method does not work for the sequence $\{\eta_n\}$ since each $\bar{f}_n$ can be zero in $(0, D)$. We were unable to solve this problem for years until the appearance of the recent paper [5; Section 6], in which the problem was solved in the discrete context. Note that $\lambda_1 > 0$ if and only if

$$\delta := \sup_{x \in (0, D)} \nu(0, x) \mu(x, D) < \infty$$

by [2; Theorem 3.7], [5; Theorem 6.2], and Theorem 18.

Proposition 22. *Let hypothesis (1) hold and assume that $\delta < \infty$. Then the sequence $\{\eta_n\}$ defined above (i.e., $\{\eta''_n\}$ in [3; Theorem 1.4]) is non-decreasing.*

Proof (a) First, we show that $f_1 \in L^1(\mu)$. Recall that $\varphi(x) = \nu(0, x)$. Clearly, for arbitrarily fixed $x_0 \in (0, D)$, we have¹⁸

$$\mu(\sqrt{\varphi}) = \int_0^{x_0} \sqrt{\varphi} d\mu + \int_{x_0}^D \sqrt{\varphi} d\mu \leq \int_0^{x_0} \sqrt{\varphi} d\mu + \frac{2\delta}{\sqrt{\varphi(x_0)}} < \infty.$$

Hence $\sqrt{\varphi} \in L^1(\mu)$.

(b) Define two sequences $\{h_n\}$ and $\{\tilde{f}_n\}$ by the same recurrence $h_n = h_{n-1} II(h_{n-1})$ but different initial condition:

$$h_0 = 1, \quad \tilde{f}_1 = f_1 = \sqrt{\varphi}.$$

¹⁸By the integration by parts formula and $\varphi(x)\mu(x, D) \leq \delta$, we have $\int_{x_0}^D \sqrt{\varphi} d\mu \leq 2\delta/\sqrt{\varphi(x_0)} < \infty$.

We now study $\{\tilde{f}_n\}$ first. From [3; Theorem 1.2 (1)], we have known that $\tilde{f}_2 \leq 4\delta \tilde{f}_1$. Assume that $\tilde{f}_{n-1} \leq (4\delta)^{n-2} \tilde{f}_1$ for some $n \geq 3$. Then

$$\tilde{f}_n = \int_0^\cdot \nu(dy) \int_y^D \tilde{f}_{n-1} d\mu \leq (4\delta)^{n-2} \int_0^\cdot \nu(dy) \int_y^D \tilde{f}_1 d\mu = (4\delta)^{n-2} \tilde{f}_2 \leq (4\delta)^{n-1} \tilde{f}_1.$$

By induction, this estimate holds for $n \geq 2$. Hence $\tilde{f}_n \in L^1(\mu)$ for $n \geq 1$ by (a).

Next, we study the sequence $\{h_n\}$. Fix $x_0 \in (0, D)$. For $x > x_0$, we have

$$\begin{aligned} h_1(x) &= h_1(x_0) + \int_{x_0}^x \nu(dy) \mu(y, D) \\ &\leq h_1(x_0) + \frac{1}{\sqrt{\varphi(x_0)}} \tilde{f}_2(x) \\ &\leq h_1(x_0) + \frac{4\delta}{\sqrt{\varphi(x_0)}} \tilde{f}_1(x). \end{aligned}$$

By induction, it is not difficult to verify that

$$h_n(x) \leq \sum_{k=1}^n \frac{(4\delta)^k}{\sqrt{\varphi(x_0)}} h_1^{n-k}(x_0) \tilde{f}_1(x) + h_1^n(x_0).$$

Hence $h_n \in L^1(\mu)$ for $n \geq 1$.

(c) Now we look for the relationship between f_n and $\tilde{f}_n$. We begin with

$$f_1 = \tilde{f}_1 = \sqrt{\varphi}, \quad f_2 = \int_0^\cdot \nu(dy) \int_y^D \tilde{f}_1 d\mu = \tilde{f}_2 - \pi(f_1) h_1.$$

By induction, we have in general

$$f_n = \tilde{f}_n - \sum_{k=1}^{n-1} h_{n-k} \pi(f_k) \quad n \geq 2.$$

Thus $f_n \in L^1(\mu)$ for every $n \geq 1$ by (b).

(d) We now come to the central part of the proof: showing the monotonicity of η_n . By definition of f_n , we have

$$\eta_n = \sup_{x \in (0, D)} \frac{e^{-C(x)}}{\tilde{f}_n'(x)} \int_x^D \tilde{f}_n d\mu = \sup_{x \in (0, D)} \left(\int_x^D \tilde{f}_n d\mu \right) \left(\int_x^D \tilde{f}_{n-1} d\mu \right)^{-1}. \quad (23)$$

¹⁹Since $\sqrt{\varphi} \uparrow$, we have $\int_y^D \sqrt{\varphi} d\mu \geq \sqrt{\varphi} \mu(y, D)$. So

$$\int_{x_0}^x \nu(dy) \mu(y, D) \leq \int_{x_0}^x \frac{1}{\sqrt{\varphi(y)}} \int_y^D \sqrt{\varphi} d\mu \nu(dy) \leq \frac{1}{\sqrt{\varphi(x_0)}} \int_{x_0}^x \int_y^D \sqrt{\varphi} d\mu \nu(dy) \leq \frac{\tilde{f}_2(x)}{\sqrt{\varphi(x_0)}}$$

Thus, $\eta_n \leq \eta_{n-1}$ if and only if

$$\int_x^D (\bar{f}_n - \eta_{n-1} \bar{f}_{n-1}) d\mu \leq 0, \quad x \in [0, D).$$

That is,

$$\int_x^D (f_n - \eta_{n-1} f_{n-1}) d\mu \leq (\pi(f_n) - \eta_{n-1} \pi(f_{n-1})) \mu(x, D),$$

or equivalently,

$$S(x) := \frac{1}{\mu(x, D)} \int_x^D (\eta_{n-1} f_{n-1} - f_n) d\mu \geq \eta_{n-1} \pi(f_{n-1}) - \pi(f_n) = S(0). \quad (24)$$

This is our key observation and leads to the study on the monotonicity of S .

(e) In view of (24), we have reduced our proof to showing non-decreasing property of S . For this, it is enough to show that

$$\mu(y, D) \int_x^D (\eta_{n-1} f_{n-1} - f_n) d\mu \leq \mu(x, D) \int_y^D (\eta_{n-1} f_{n-1} - f_n) d\mu$$

for any $x, y \in [0, D)$ with $x < y$. By separating f_n and f_{n-1} , the last inequality is equivalent to the following one:

$$\eta_{n-1} \int_y^D \mu(du) \int_x^y (f_{n-1}(t) - f_{n-1}(u)) \mu(dt) \leq \int_y^D \mu(du) \int_x^y (f_n(t) - f_n(u)) \mu(dt). \quad (25)$$

To see this, it suffices to check that

$$f_n(u) - f_n(t) \leq \eta_{n-1} (f_{n-1}(u) - f_{n-1}(t)), \quad u \geq t.$$

To check the last inequality, consider $n \geq 3$ first. Then

$$\begin{aligned} f_n(u) - f_n(t) &= \int_t^u \nu(dy) \int_y^D \bar{f}_{n-1} d\mu \quad (\text{by definition of } f_n) \\ &\leq \eta_{n-1} \int_t^u \nu(dy) \int_y^D \bar{f}_{n-2} d\mu \quad (\text{by (23)}) \\ &= \eta_{n-1} (f_{n-1}(u) - f_{n-1}(t)) \quad (\text{by definition of } f_{n-1}), \quad u \geq t. \end{aligned}$$

It remains to check the required inequality for $n = 2$. By definition of η_1 , we have

$$\frac{e^{-C(y)}}{\bar{f}'_1(y)} \int_y^D \bar{f}_1 d\mu = I(\bar{f}_1)(y) \leq \eta_1.$$

It follows that

$$f_2(u) - f_2(t) = \int_t^u \nu(dy) \int_y^D \bar{f}_1 d\mu \leq \eta_1 \int_t^u \bar{f}'_1(y) dy \leq \eta_1 (f_1(u) - f_1(t)), \quad u \geq t.$$

We have thus completed the proof of the monotonicity of $\{\eta_n\}$ in the continuous context. $\square$

The monotonicity of $\{\eta_n\}$ means we can theoretically improve our lower estimates of λ_1 step by step. There is a similar result for the upper estimates but omitted here. It is regretted that the converges of $\{\eta_n^{-1}\}$ to λ_1 (as $n \rightarrow \infty$) remains open. All examples we have ever computed support the convergence.

Acknowledgements The work was supported in part by the National Natural Science Foundation of China (Grant No.11131003), the Specialized Research Fund for the Doctoral Program of Higher Education (Grant No. 20100003110005), the “985” project from the Ministry of Education in China and the Fundamental Research Funds for the Central Universities.

References

1. Chen M.F. Analytic proof of dual variational formula for the first eigenvalue in dimension one. *Sci in China, A*, 1999, 42(8): 805–815
2. Chen M.F. Explicit bounds of the first eigenvalue. *Sci Chin Ser, A*, 2000, 43(10): 1051–1059
3. Chen M.F. Variational formulas and approximation theorems for the first eigenvalue in dimension one. *Sci in China, A*, 2001, 44(4): 409–418
4. Chen M.F. *Eigenvalues, Inequalities, and Ergodic theory*. New York: Springer, 2005
5. Chen M.F. Speed of stability for birth-death process. *Front Math China*, 2010, 5(3): 379–516
6. Chen M.F. Basic estimates and stability rate for one-dimensional diffusion. Chapter 6 in “Probability Approximations and Beyond”, *Lecture Notes in Statistics* 205, 2012, 75–99
7. Chen M.F. Lower bounds of principal eigenvalue in dimension one, *Front. Math. China* 7:4, 645–668, 2012.
8. Chen M.F., Wang F.Y. Estimation of spectral gap for elliptic operators. *Trans Amer Math Soc*, 1997, 349(3): 1239–1267. See also book [4] in for some supplement: http://math.bnu.edu.cn/~chenmf/main_eng.htm
9. Wang J. First eigenvalue of one-dimensional diffusion processes. *Elect Comm Probab*, 2009, 14: 232–244
10. Zettl A. *Sturm-Liouville Theory*. AMS, 2005

Appendix A Complement of the proofs in Section 3

A.1 Complementary proof of the two circle arguments: lower estimates

Let us review the circle arguments for the lower estimates first.

$$\begin{aligned}\lambda_0 &\geq \tilde{\lambda}_0 \geq \sup_{f \in \mathcal{F}_I} \inf_{x \in (0,D)} I(f)(x)^{-1} \\ &= \sup_{f \in \mathcal{F}_I} \inf_{x \in (0,D)} II(f)(x)^{-1} = \sup_{f \in \mathcal{F}_{II}} \inf_{x \in (0,D)} II(f)(x)^{-1} \\ &\geq \sup_{h \in \mathcal{H}} \inf_{x \in (0,D)} R(h)(x) \geq \lambda_0,\end{aligned}$$

where

$$\tilde{\lambda}_0 = \inf \{ D(f) : \mu(f^2) = 1, f \in \mathcal{C}^1(0,D) \cap \mathcal{C}[0,D], f'(0) = 0, \text{ and } f(D) = 0 \}.$$

We prove the circle arguments through the following (a)-(e) steps.

(a) Prove that $\lambda_0 \geq \tilde{\lambda}_0 \geq \sup_{f \in \mathcal{F}_{II}} \inf_{x \in (0,D)} II(f)(x)^{-1}$.

The first assertion is obvious by definitions of λ_0 and $\tilde{\lambda}_0$. The second one is proved in the main text.

(b) Prove that

$$\sup_{f \in \mathcal{F}_I} \inf_{x \in (0,D)} I(f)(x)^{-1} = \sup_{f \in \mathcal{F}_I} \inf_{x \in (0,D)} II(f)(x)^{-1} = \sup_{f \in \mathcal{F}_{II}} \inf_{x \in (0,D)} II(f)(x)^{-1}.$$

For $f \in \mathcal{F}_I$, without loss of generality, assume that $\sup_{x \in (0,D)} I(f)(x) < \infty$. By using Cauchy's mean value theorem, we have

$$\begin{aligned}\sup_{x \in (0,D)} II(f)(x) &= \sup_{x \in (0,D)} \frac{1}{f(x)} \int_x^D \nu(ds) \int_0^s f d\mu \\ &\leq \sup_{x \in (0,D)} - \left(\int_x^D e^{-C(s)} ds \int_0^s f d\mu \right) \left(\int_x^D f'(t) dt \right)^{-1} \\ &\leq \sup_{t \in (0,D)} - \frac{e^{-C(t)}}{f'(t)} \int_0^t f d\mu \\ &= \sup_{x \in (0,D)} I(f)(x) < \infty.\end{aligned}$$

Making infimum with respect to $f \in \mathcal{F}_I$, we have

$$\inf_{f \in \mathcal{F}_I} \sup_{x \in (0,D)} II(f)(x) \leq \inf_{f \in \mathcal{F}_I} \sup_{x \in (0,D)} I(f)(x).$$

Since $\mathcal{F}_I \subset \mathcal{F}_{II}$, the left-hand side is bounded below by $\inf_{f \in \mathcal{F}_{II}} \sup_{x \in (0,D)} II(f)(x)$.

Hence

$$\sup_{f \in \mathcal{F}_I} \inf_{x \in (0,D)} I(f)(x)^{-1} \leq \sup_{f \in \mathcal{F}_I} \inf_{x \in (0,D)} II(f)(x)^{-1} \leq \sup_{f \in \mathcal{F}_{II}} \inf_{x \in (0,D)} II(f)(x)^{-1}.$$

To obtain the equality signs, it suffices to show

$$\sup_{f \in \mathcal{F}_H} \inf_{x \in (0, D)} II(f)(x)^{-1} \leq \sup_{f \in \mathcal{F}_I} \inf_{x \in (0, D)} I(f)(x)^{-1}.$$

To do so, let $f \in \mathcal{F}_H$. Without loss of generality, assume that $\inf_{x \in (0, D)} II(f)(x)^{-1} > 0$. Then $f \in \mathcal{C}[0, D]$ and $f > 0$. Put

$$g(x) = fII(f)(x) = \int_x^D \nu(ds) \int_0^s f d\mu.$$

Then $g \in \mathcal{F}_I$ and

$$-g'(s)e^{C(s)} = \int_0^s f d\mu \geq \int_0^s g d\mu \inf_{x \in (0, D)} \frac{f(x)}{g(x)} \quad \text{for } s \in (0, D).$$

That is $I(g)(s)^{-1} \geq \inf_{x \in (0, D)} II(f)(x)^{-1}$. Making infimum with respect to $s \in (0, D)$, we obtain

$$\inf_{s \in (0, D)} I(g)(s)^{-1} \geq \inf_{x \in (0, D)} II(f)(x)^{-1}.$$

The assertion now follows by making supremum with respect to $g \in \mathcal{F}_I$ on the both sides of the inequality first and then with respect to $f \in \mathcal{F}_H$.

A different way to prove the equalities here and in (a), without using the continuity of a and b , is to show that

$$\sup_{f \in \mathcal{F}_I} \inf_{x \in (0, D)} I(f)(x)^{-1} \geq \lambda_0.$$

By the comments below Lemma 13 and (11), we have seen that $\lambda_0 = I(g)(x)^{-1}$ for $x \in (0, D)$. In view of Lemmas 12 and 13, it follows that $g \in \mathcal{F}_I$ and

$$\lambda_0 = \inf_{x \in (0, D)} I(g)(x)^{-1} \leq \sup_{f \in \mathcal{F}_I} \inf_{x \in (0, D)} I(f)(x)^{-1}.$$

In the following two steps, assume that $a, b \in \mathcal{C}[0, D]$.

(c) We show that $\sup_{f \in \mathcal{F}_H} \inf_{x \in (0, D)} II(f)(x)^{-1} \geq \sup_{h \in \mathcal{H}} \inf_{x \in (0, D)} R(h)(x)$.

To this end, recall that for each $h \in \mathcal{H}$ with $h = g'/g$ (see Remark 4), we have

$$R(h) = -(ah^2 + bh + ah') = -\frac{Lg}{g}.$$

Before moving further, we prove that if $R(h) > 0$ for a positive g with $g'(0) = 0$ and $h = g'/g$, then g must be strictly decreasing. In fact, we have $R(h) = -Lg/g > 0$. Let $f = gR(h) > 0$. Then $Lg = -f$. Moreover,

$$g'(x) = -e^{C(x)} \int_0^x f d\mu$$

since $g'(0) = 0$. So $g' < 0$ on $(0, D)$.

Now, we return to our main assertion. It suffices to show that

$$\inf_{x \in (0, D)} R(h)(x) \leq \sup_{f \in \mathcal{F}_H} \inf_{x \in (0, D)} II(f)(x)^{-1}, \quad h \in \mathcal{H}.$$

Without loss of generality, assume that $\inf_{x \in (0, D)} R(h)(x) > 0$. Then

$$R(h)(x) > 0 \quad \text{for } x \in (0, D).$$

Let $f = -(ag'' + bg') = gR(h)$ (g is the function given above). Then $Lg = -f$, $f > 0$ and $f \in \mathcal{C}[0, D]$ since $a, b \in \mathcal{C}[0, D]$. Since $Lg = -f$ and $g'(0) = 0$, we obtain

$$g(x) - g(D) = \int_x^D \nu(ds) \int_0^s f d\mu = f(x)II(f)(x)$$

by (10). That is $g(x) \geq f(x)II(f)(x)$ since $g(D) \geq 0$. So

$$R(h)(x)^{-1} = \frac{g(x)}{f(x)} \geq II(f)(x) \quad \text{for } x \in (0, D).$$

Furthermore,

$$\inf_{x \in (0, D)} R(h)(x) \leq \inf_{x \in (0, D)} II(f)(x)^{-1} \leq \sup_{f \in \mathcal{F}_H} \inf_{x \in (0, D)} II(f)(x)^{-1},$$

and the assertion follows since $h \in \mathcal{H}$ is arbitrary.

(d) Prove that $\sup_{h \in \mathcal{H}} \inf_{x \in (0, D)} R(h)(x) \geq \lambda_0$.

Let $f \in L^1(\mu)$, $g = fII(f)$, and $\bar{h} = g'/g$ on $[0, D]$. Then

$$\bar{h} \in \mathcal{H}, \quad Lg = -f, \quad \text{and} \quad R(h) = \frac{-Lg}{g} = \frac{f}{g} > 0.$$

Thus, $\sup_{h \in \mathcal{H}} \inf_{x \in (0, D)} R(h)(x) \geq 0$. Without loss of generality, assume that $\lambda_0 > 0$. Since $a, b \in \mathcal{C}[0, D]$, by Lemma 8 (2), there exists an eigenfunction g such that $Lg = -\lambda_0 g$. Furthermore,

$$g'(0) = 0, \quad g|_{(0, D)} > 0, \quad g'|_{(0, D)} < 0 \quad \text{and} \quad g \in \mathcal{C}^1(0, D) \cap \mathcal{C}[0, D].$$

Let $h = g'/g$. Then $h \in \mathcal{C}^1(0, D) \cap \mathcal{C}[0, D]$, $h(0) = 0$, $h \in \mathcal{H}$, and

$$R(h)(x) = -\frac{Lg(x)}{g(x)} = \lambda_0 \quad \text{for } x \in (0, D).$$

So the assertion follows immediately.

(e) We now prove that the supremum in the first circle arguments can be attained. The case that $\lambda = 0$ is easier since

$$0 = \lambda_0 \geq \inf_{x \in (0, D)} II(f)(x)^{-1} \geq 0 \quad \text{and} \quad 0 = \lambda_0 \geq \inf_{x \in (0, D)} I(f)(x)^{-1} \geq 0$$

for every f in their corresponding domains, as an application of the first circle arguments. Similarly, the conclusion holds for R as seen from proof (d): noting in the degenerated case that $\nu(0, D) = \infty$, we have $\lambda_0 = 0$ (which is a simple consequence of definition (2), see also the proof of Theorem 5 given below) and then $h = 0$ since the eigenfunction is constant in the case.

Next, we consider the case that $\lambda_0 > 0$. Let g be its eigenfunction. For R the supremum is attained at $h = g'/g$ as seen from the last paragraph of proof (d). For the operator I and II , we have already seen that $I(g) \equiv II(g) \equiv \lambda_0^{-1}$ according to Lemma 14 and the remarks below Lemma 13. $\square$

A.2 Complementary proof of the two circle arguments: upper estimates

For the upper estimation of λ_0 , we review and show the circle arguments in the following.

$$\begin{aligned} \lambda_0 &\leq \inf_{f \in \widetilde{\mathcal{F}}_{II} \cup \widetilde{\mathcal{F}}'_{II}} \sup_{x \in \text{supp}(f)} II(f)(x)^{-1} = \inf_{f \in \widetilde{\mathcal{F}}_{II}} \sup_{x \in \text{supp}(f)} II(f)(x)^{-1} \\ &= \inf_{f \in \widetilde{\mathcal{F}}_I} \sup_{x \in \text{supp}(f)} II(f)(x)^{-1} = \inf_{f \in \widetilde{\mathcal{F}}_I} \sup_{x \in (0, D)} I(f)(x)^{-1} \\ &= \inf_{f \in \widetilde{\mathcal{F}}'_I} \sup_{x \in (0, D)} I(f)(x)^{-1} \\ &\leq \inf_{h \in \mathcal{H}} \sup_{x \in (0, D)} R(h)(x) \leq \lambda_0. \end{aligned}$$

(f) The assertion that $\lambda_0 \leq \inf_{f \in \widetilde{\mathcal{F}}_{II} \cup \widetilde{\mathcal{F}}'_{II}} \sup_{x \in \text{supp}(f)} II(f)(x)^{-1}$ is proved in our main text.

(g) Prove that

$$\inf_{f \in \widetilde{\mathcal{F}}_{II}} \sup_{x \in \text{supp}(f)} II(f)(x)^{-1} = \inf_{f \in \widetilde{\mathcal{F}}_I} \sup_{x \in \text{supp}(f)} II(f)(x)^{-1} = \inf_{f \in \widetilde{\mathcal{F}}_I} \sup_{x \in (0, D)} I(f)(x)^{-1}.$$

Let $f \in \widetilde{\mathcal{F}}_I$, there exists $x_0, x_1 \in [0, D)$ such that $f = f(\cdot \vee x_0) \mathbb{1}_{[0, x_1)} \in \mathcal{C}^1(x_0, x_1) \cap \mathcal{C}[x_0, x_1]$. By Cauchy's mean value theorem, we have

$$\begin{aligned} \inf_{x \in \text{supp}(f)} II(f)(x) &= \inf_{x \in [x_0, x_1)} \frac{1}{f(x)} \int_x^{x_1} e^{-C(t)} \int_0^t f d\mu dt \\ &\geq \inf_{x \in [x_0, x_1)} \frac{1}{-f'(x)} e^{-C(x)} \int_0^x f d\mu \\ &= \inf_{x \in (0, D)} I(f)(x). \end{aligned}$$

So the assertion that

$$\inf_{f \in \widetilde{\mathcal{F}}_{II}} \sup_{x \in \text{supp}(f)} II(f)(x)^{-1} \leq \inf_{f \in \widetilde{\mathcal{F}}_I} \sup_{x \in \text{supp}(f)} II(f)(x)^{-1} \leq \inf_{f \in \widetilde{\mathcal{F}}_I} \sup_{x \in (0, D)} I(f)(x)^{-1}$$

follows by $\widetilde{\mathcal{F}}_I \subset \widetilde{\mathcal{F}}_{II}$.

There are two choices to prove the equalities. The first choice is proving the assertion that

$$\inf_{f \in \widetilde{\mathcal{F}}_I} \sup_{x \in (0, D)} I(f)(x)^{-1} \leq \inf_{f \in \widetilde{\mathcal{F}}_{II}} \sup_{x \in \text{supp}(f)} II(f)(x)^{-1}.$$

For $f \in \widetilde{\mathcal{F}}_{II}$, $\exists x_0 \in (0, D)$ such that $f = f \mathbb{1}_{[0, x_0]}$ and $f \in \mathcal{C}[0, x_0]$. Let $g = f II(f) \mathbb{1}_{\text{supp}(f)}$. Then

$$g(x) = \int_x^{x_0} \nu(ds) \int_0^s f d\mu \mathbb{1}_{[0, x_0]}(x), \quad g \in \widetilde{\mathcal{F}}'_1 \subseteq \widetilde{\mathcal{F}}_I,$$

and $Lg = -f$ on $[0, x_0]$ by simple calculation. Since $g'(0) = 0$, replacing $[0, D]$ with $[0, x_0]$ in (9), we have

$$-e^{C(x)} g'(x) = \int_0^x f d\mu \leq \int_0^x g d\mu \sup_{t \in (0, x_0)} \frac{f(t)}{g(t)}, \quad x < x_0.$$

Hence,

$$-e^{C(x)} g'(x) \left(\int_0^x g d\mu \right)^{-1} \leq \sup_{x \in \text{supp}(f)} II(f)(x)^{-1}, \quad x < x_0.$$

Making supremum with respect to $x \in (0, x_0)$, we have

$$\sup_{x \in (0, D)} I(g)(x)^{-1} = \sup_{x \in (0, x_0)} I(g)(x)^{-1} \leq \sup_{x \in \text{supp}(f)} II(f)(x)^{-1}.$$

The assertion now follows by making infimum with respect to $g \in \widetilde{\mathcal{F}}_I$ first, then with respect to $f \in \widetilde{\mathcal{F}}_{II}$.

The second method for the identity is making a small circle below.
Since

$$\begin{aligned} \lambda_0 &\leq \inf_{f \in \widetilde{\mathcal{F}}_{II}} \sup_{x \in \text{supp}(f)} II(f)(x)^{-1} \leq \inf_{f \in \widetilde{\mathcal{F}}_I} \sup_{x \in \text{supp}(f)} II(f)(x)^{-1} \\ &\leq \inf_{f \in \widetilde{\mathcal{F}}_I} \sup_{x \in (0, D)} I(f)(x)^{-1} \leq \inf_{f \in \widetilde{\mathcal{F}}'_I} \sup_{x \in (0, D)} I(f)(x)^{-1}, \end{aligned}$$

it suffices to show $\inf_{f \in \widetilde{\mathcal{F}}'_I} \sup_{x \in (0, D)} I(f)(x)^{-1} \leq \lambda_0$.

To see this, we introduce an approximating procedure. Recall that

$$\lambda_0^{(0, p)} = \inf \{ D(f) : \mu(f^2) = 1, f \in \mathcal{C}^1(0, p) \cap \mathcal{C}[0, p], f|_{[p, D]} = 0 \}.$$

Let $p_n \in (0, D)$, $p_n \uparrow D$. Then $\lambda_0^{(0, p_n)} \downarrow \lambda_0$ by Lemma 11 (1), where $\lambda_0^{(0, p_n)}$ is the first eigenvalue of the Dirichlet form $(D, \mathcal{D}(D))$ restricted to $(0, p_n)$ with ND-boundaries. Now, let g be the eigenfunction of $\lambda_0^{(0, p_n)} > 0$. Extend g to

the whole space by setting $g = g\mathbb{1}_{[0,p_n]}$. By using Lemmas 12 and 13, it follows that $g \in \widetilde{\mathcal{F}}_I'$. Furthermore,

$$\lambda_0^{(0,p_n)} = \sup_{x \in (0,p_n)} I(g)(x)^{-1} = \sup_{x \in (0,D)} I(g)(x)^{-1} \geq \inf_{g \in \widetilde{\mathcal{F}}_I'} \sup_{x \in (0,D)} I(g)(x)^{-1}.$$

The assertion now follows by letting $n \rightarrow \infty$ because of $\tilde{\lambda}_0 = \lambda_0$.

In the following two steps, we assume that $a, b \in \mathcal{C}[0, D]$.

(h) Prove that

$$\inf_{f \in \widetilde{\mathcal{F}}_{II}} \sup_{x \in (0,D)} II(f)(x)^{-1} \leq \inf_{h \in \widetilde{\mathcal{H}}} \sup_{x \in (0,D)} R(h)(x).$$

First, for $h \in \widetilde{\mathcal{H}}$, $\exists x_0 \in (0, D)$ such that $R(h) > 0$ on $(0, x_0)$. We use $R(g)|_{[0,x_0]}$ instead of $R(h)|_{[0,x_0]}$ as in (c). Hence,

$$R(g)(x) = \begin{cases} -(Lg/g)(x), & x < x_0; \\ 0, & x \geq x_0. \end{cases}$$

Secondly, we turn to the main assertion.

Let $f = gR(g)$. Then

$$f = f\mathbb{1}_{[0,x_0]}, \quad Lg = -f \quad \text{on } (0, x_0),$$

and $f \in \widetilde{\mathcal{F}}_{II}$ since $a, b \in \mathcal{C}[0, D]$. Noting that $g'(0) = 0$ and $g(x_0) = 0$, by (10), we have

$$g(y) = \int_y^{x_0} \nu(dx) \int_0^x f d\mu = f(y)II(f)(y) \quad \text{for } y < x_0.$$

So

$$R(h)(y) = \frac{f(y)}{g(y)} = II(f)(y)^{-1} \quad \text{for } y < x_0.$$

Making supremum with respect to $y \in (0, x_0)$, we have

$$\sup_{y \in (0,x_0)} R(h)(y) = \sup_{y \in (0,x_0)} II(f)(y)^{-1},$$

and the assertion follows immediately by making infimum with respect to $f \in \widetilde{\mathcal{F}}_{II}$ firstly and then making infimum with respect to $h \in \widetilde{\mathcal{H}}$.

(i) Prove that $\inf_{h \in \widetilde{\mathcal{H}}} \sup_{x \in (0,D)} R(h)(x) \leq \lambda_0$.

When $D < \infty$, since $a, b \in \mathcal{C}[0, D]$, there is an eigenfunction g satisfying

$$h := \frac{g'}{g} \in \widetilde{\mathcal{H}}, \quad R(h) = -\frac{Lg}{g} = \lambda_0.$$

Indeed, since $a, b \in \mathcal{C}[0, D]$, we have $g \in \mathcal{C}^2[0, D]$, $g'(0) = 0$, $g(D) = 0$, and $g' < 0$ on $(0, D)$. Hence, $h(0) = 0$, $h(D) = 0$, $h < 0$ on $(0, D)$, and $h \in \mathcal{C}^1(0, D) \cap \mathcal{C}[0, D]$. Moreover,

$$R(h)(x) = -(Lg/g)(x) = \lambda_0 > 0.$$

So the assertion holds for $D < \infty$.

When $D = \infty$, let $p_n \uparrow \infty$. For fixed p_n , as the last part of (g), denote by g the eigenfunction of $\lambda_0^{(0, p_n)} > 0$, i.e.

$$Lg(x) = -\lambda_0^{(0, p_n)} g(x), \quad x \in (0, p_n).$$

Since $a, b \in \mathcal{C}[0, D]$, we have

$$g \downarrow\downarrow \text{ on } (0, p_n), \quad g'(0) = 0, \quad g(p_n) = 0, \quad g \in \mathcal{C}^2[0, p_n].$$

by Lemmas 12, 13 and 14.

Let $\bar{h}_n(x) = g'(x)\mathbb{1}_{[0, p_n)}(x)/g(x)$. Then $\bar{h}_n \in \widetilde{\mathcal{H}}$ and

$$\begin{aligned} \lambda_0^{(0, p_n)} &= \sup_{x \in (0, p_n)} R(g)(x) \\ &\geq \inf_{h \in \widetilde{\mathcal{H}}, \text{supp}(h) = (0, p_n)} \sup_{x \in (0, p_n)} R(h)(x) \\ &\geq \inf_{h \in \widetilde{\mathcal{H}}, \text{supp}(h) = (0, p_n)} \sup_{x \in (0, D)} R(h)(x) \\ &\geq \inf_{h \in \widetilde{\mathcal{H}}} \sup_{x \in (0, D)} R(h)(x). \end{aligned}$$

The assertion now follows by letting $n \rightarrow \infty$. $\square$

A.3 Proof of Proposition 3

The proof consists of the following four parts.

(a) The assertion that

$$\sup_{h \in \overline{\mathcal{H}}} \inf_{x \in (0, D)} \bar{R}(h)(x) \geq \sup_{h \in \mathcal{H}_*} \inf_{x \in (0, D)} \bar{R}(h)(x) \geq \lambda_0 \geq 0$$

is proved in our main text.

(b) Prove that $\lambda_0 = \sup_{h \in \overline{\mathcal{H}}} \inf_{x \in (0, D)} \bar{R}(h)(x)$ whenever $\mu(0, D) = \infty$. From (a), it suffices to show that

$$\lambda_0 \geq \sup_{h \in \overline{\mathcal{H}}} \inf_{x \in (0, D)} \bar{R}(h)(x),$$

or equivalently

$$\lambda_0 \geq \inf_{x \in (0, D)} \bar{R}(h)(x) \quad \text{for every } h \in \overline{\mathcal{H}}.$$

In view of (a), without loss of generality, assume that $\inf_{x \in (0, D)} \bar{R}(h)(x) > 0$ for a given $h \in \mathcal{H}$. Let $f = -(ah' + bh)$. Since $h < 0$ on $(0, D)$ and $h(0) = 0$, we have

$$f' = h\bar{R}(h) < 0 \quad \text{and} \quad f'(0) = 0.$$

Thus, $f \in \mathcal{C}^1(0, D) \cap \mathcal{C}[0, D]$. It follows that $f \in \mathcal{F}_I$ once we show that $f > 0$ on $(0, D)$.

For this, fix $x \in (0, D)$. By integration formula by parts and $h(0) = 0$, we obtain

$$\int_0^x f d\mu = -e^{C(x)}h(x) > 0. \quad (26)$$

Since $f \searrow$, if $f(x_0) \leq 0$ for some $x_0 \in (0, D)$, then

$$\int_{x_0}^x f d\mu \leq f(x_0)\mu(x_0, x), \quad x > x_0,$$

and the right-hand side of the above inequality converges to $-\infty$ as $x \rightarrow D$.

By (26), we obtain

$$0 < \int_0^x f d\mu = \int_0^{x_0} f d\mu + \int_{x_0}^x f d\mu \rightarrow -\infty \quad \text{as } x \rightarrow D,$$

which is a contradiction. So $f > 0$ on $(0, D)$.

Because of (26), we have

$$\bar{R}(h)(x) = \frac{f'(x)}{h(x)} = \left(-\frac{e^{-C(x)}}{f'(x)} \int_0^x f d\mu \right)^{-1} = I(f)(x)^{-1}.$$

Making infimum with respect to $x \in (0, D)$ first and then making supremum with respect to $f \in \mathcal{F}_I$, we obtain the assertion by the variational formulas for lower bounds in Theorem 1 (1).

(c) Prove that $\lambda_0 = \sup_{h \in \mathcal{H}_*} \inf_{x \in (0, D)} \bar{R}(h)(x)$.

It suffices to prove that

$$\lambda_0 \geq \sup_{h \in \mathcal{H}_*} \inf_{x \in (0, D)} \bar{R}(h)(x).$$

The main body in proof (b) is to prove that the function f defined there is positive, this is automatic due to the definition of $\mathcal{H}_*$. So the proof (b) can be applied to $h \in \mathcal{H}_*$ directly. Hence, $\lambda_0 \geq \sup_{h \in \mathcal{H}_*} \inf_{x \in (0, D)} \bar{R}(h)(x)$ and then the equality $\lambda_0 = \sup_{h \in \mathcal{H}_*} \inf_{x \in (0, D)} \bar{R}(h)(x)$ follows. $\square$

A.4 Proof of Lemma 16

Define $M(x) = \int_0^x m(y)dy$. Using integration by parts formula, we have

$$\begin{aligned}
 \int_0^x m(y)\psi^r(y)dy &= \int_0^x \psi^r(y)dM(y) \\
 &= \psi^r(x)M(x) - r \int_0^x \psi^{r-1}(y)\psi'(y)M(y)dy \quad (\text{by } M(0) = 0, \psi(0) < \infty) \\
 &\leq c\psi^{r-1}(x) - cr \int_0^x \psi^{r-2}(y)\psi'(y)dy \quad (\text{since } M\psi \leq c, \psi' = -n \leq 0) \\
 &= c\psi^{r-1}(x) - \frac{cr}{r-1}(\psi^{r-1}(x) - \psi^{r-1}(0)) \\
 &= \frac{c}{1-r}\psi^{r-1}(x) - \frac{cr}{1-r}\psi^{r-1}(0) \leq \frac{c}{1-r}\psi^{r-1}(x). \quad \square
 \end{aligned}$$

A.5 Proof of Theorem 5

Firstly, the assertion $\lambda_0 \geq (4\delta)^{-1}$ is proved in the main text.

Now, we show that $\lambda_0 \leq \delta^{-1}$. Let $x_0, x_1 \in [0, D)$ with $x_0 < x_1$. Set $f = \nu(x_0 \vee \cdot, x_1)\mathbb{1}_{[0, x_1]}$. Then $f \in \widetilde{\mathcal{F}}_I$, $f' = -e^{-C}$ on (x_0, x_1) , and

$$I(f)(x) = \begin{cases} \int_{x_0}^x f d\mu + \nu(x_0, x_1)\mu(0, x_0), & x \in (x_0, x_1); \\ \infty \text{ (by convention, } 1/0 = \infty), & \text{otherwise.} \end{cases}$$

Thus, $I(f)(x)$ achieves its minimum at $x = x_0+$, and

$$\inf_{x \in (0, D)} I(f)(x) = \inf_{x \in (x_0, x_1)} I(f)(x) = \nu(x_0, x_1)\mu(0, x_0) \rightarrow \mu(0, x_0)\nu(x_0, D)$$

as $x_1 \rightarrow D$. Hence,

$$\lambda_0^{-1} \geq \sup_{f \in \widetilde{\mathcal{F}}_I} \inf_{x \in (0, D)} I(f)(x) \geq \sup_{x_0 \in (0, D)} \mu(0, x_0)\nu(x_0, D) = \delta.$$

Another method to show that $\lambda_0 \leq \delta^{-1}$ is using the classical variational formula:

$$\lambda_*^{-1} \geq \sup_{x_0, x_1: x_0 < x_1} \frac{\|f^{x_0, x_1}\|}{D(f^{x_0, x_1})}.$$

By simple calculation, we have

$$\begin{aligned}
 \|f^{x_0, x_1}\| &= \int_0^D f^{x_0, x_1 2} d\mu = \nu^2(x_0, x_1)\mu(0, x_0) + \int_{x_0}^{x_1} \nu^2(t, x_1) d\mu, \\
 D(f^{x_0, x_1}) &= \nu(x_0, x_1).
 \end{aligned}$$

Thus,

$$\lambda_0^{-1} = \lambda_*^{-1} \geq \sup_{x_0, x_1: x_0 < x_1} \frac{\|f^{x_0, x_1}\|}{D(f^{x_0, x_1})} \geq \sup_{x_0, x_1: x_0 < x_1} \nu(x_0, x_1)\mu(0, x_0) = \delta$$

by Lemma 9 and Proposition 2(1). $\square$

A.6 Proof of Theorem 6 and Corollary 7

We prove the assertions through the following six steps.

(a) By Cauchy's mean value theorem and (19), we have

$$\delta_1 = \sup_{x \in (0, D)} II(f_1)(x) \leq \sup_{x \in (0, D)} I(f_1)(x) \leq 4\delta,$$

and

$$\delta_{n+1} = \sup_{x \in (0, D)} II(f_{n+1})(x) = \sup_{x \in (0, D)} \frac{f_{n+2}(x)}{f_{n+1}(x)} \leq \sup_{x \in (0, D)} \frac{f_{n+1}(x)}{f_n(x)} = \delta_n,$$

which means the monotonicity of δ_n with respect to n .

Notice that

$$f_1(x) = \sqrt{\varphi(x)} > 0, \quad f_1'(x) = -\frac{e^{-C(x)}}{2\sqrt{\varphi(x)}}.$$

We have $f_1 \in \mathcal{C}[0, D]$ and $f_1 \in \mathcal{F}_H$. Moreover, by induction, we have $f_n \in \mathcal{F}_H$ for $n \geq 1$. Therefore,

$$\lambda_0 \geq \sup_{f \in \mathcal{F}_H} \inf_{x \in (0, D)} II(f)(x)^{-1} \geq \delta_n^{-1} \quad \text{for } n \geq 1.$$

(b) By Cauchy's mean value theorem, we have

$$\begin{aligned} \inf_{x < x_1} II(f_1^{x_0, x_1})(x) &= \inf_{x \in [x_0, x_1)} \frac{1}{f_1^{x_0, x_1}(x)} \int_x^{x_1} e^{-C(s)} ds \int_0^s f_1^{x_0, x_1} d\mu \\ &\geq \inf_{x \in [x_0, x_1)} \int_0^x f_1^{x_0, x_1} d\mu \\ &= \inf_{x \in [x_0, x_1)} \int_0^x \int_{t \vee x_0}^{x_1} d\nu \mu(dt) \\ &= \int_0^{x_0} d\mu \int_{x_0}^{x_1} d\nu \\ &= \mu(0, x_0) \nu(x_0, x_1). \end{aligned}$$

Thus,

$$\sup_{x_0, x_1: x_0 < x_1} \inf_{x < x_1} II(f_1^{x_0, x_1})(x) \geq \sup_{x_0, x_1: x_0 < x_1} \mu(0, x_0) \nu(x_0, x_1) = \sup_{x \in (0, D)} \mu(0, x) \nu(x, D),$$

which is just $\delta'_1 \geq \delta$.

Meanwhile, for the same reason, we obtain

$$\inf_{x < x_1} II(f_{n+1}^{x_0, x_1})(x) = \inf_{x < x_1} \frac{f_{n+2}^{x_0, x_1}(x)}{f_{n+1}^{x_0, x_1}(x)} \geq \inf_{x < x_1} \frac{f_{n+1}^{x_0, x_1}(x)}{f_n^{x_0, x_1}(x)} = \inf_{x < x_1} II(f_n^{x_0, x_1})(x).$$

which implies that

$$\delta'_{n+1} = \sup_{x_0, x_1: x_0 < x_1} \inf_{x < x_1} II(f_{n+1}^{x_0, x_1})(x) \geq \sup_{x_0, x_1: x_0 < x_1} \inf_{x < x_1} II(f_n^{x_0, x_1})(x) = \delta'_n,$$

i.e. δ'_n is increasing in n .

(c) Noticing that

$$f_1^{x_0, x_1} = \nu(x_0 \vee \cdot, x_1) \mathbb{1}_{[0, x_1)} \quad \text{and} \quad (f_1^{x_0, x_1})'(x) = -e^{-C(x)} \quad \text{on } (x_0, x_1),$$

we have $f_1^{x_0, x_1} \in \mathcal{C}^1(x_0, x_1) \cap \mathcal{C}[x_0, x_1]$ and further $f_1^{x_0, x_1} \in \widetilde{\mathcal{F}}_I \subset \widetilde{\mathcal{F}}_{II}$. It is easy to verify that $f_n^{x_0, x_1} \in \widetilde{\mathcal{F}}_I \subset \widetilde{\mathcal{F}}_{II}$ for $n \geq 1$ by induction. From the variational formula for upper bounds, we obtain the following inequalities below immediately.

$$\lambda_0 \leq \inf_{f \in \widetilde{\mathcal{F}}_I} \sup_{x \in \text{supp}(f)} II(f)(x)^{-1} \leq \inf_{x_0, x_1: x_0 < x_1} \sup_{x < x_1} II(f_n^{x_0, x_1})(x)^{-1} = \delta_n'^{-1}, \quad n \geq 1.$$

(d) Since $f_n^{x_0, x_1}$ By definition of λ_* , it is obvious that

$$(\bar{\delta}_n)^{-1} \geq \lambda_* = \lambda_0$$

by Lemma 9 and Proposition 2 (1). Next, let $f = f_n^{x_0, x_1}$. Replacing $f II(f)(\cdot \wedge x_0) \mathbb{1}_{\text{supp}(f)} = f_{n+1}^{x_0, x_1}$ with g in (18), by definition of λ_* , we obtain that $\bar{\delta}_{n+1} \geq \delta'_n$.

(e) The computation of δ_1 is simple, Here we only compute $\bar{\delta}_1$ and δ'_1 in details. Firstly, for any $x \in [x_0, x_1)$, we have

$$\begin{aligned} f_2^{x_0, x_1}(x) &= \int_x^{x_1} \nu(ds) \int_0^s f_1^{x_0, x_1} d\mu = \int_0^{x_1} f_1^{x_0, x_1}(t) \mu(dt) \int_{t \vee x}^{x_1} d\nu \\ &= \int_0^{x_1} \mu(dt) \int_{t \vee x_0}^{x_1} d\nu \int_{t \vee x}^{x_1} d\nu \\ &= \int_0^{x_0} \mu(dt) \int_{x_0}^{x_1} d\nu \int_x^{x_1} d\nu + \int_{x_0}^x \mu(dt) \int_t^{x_1} d\nu \int_x^{x_1} d\nu \\ &\quad + \int_x^{x_1} \mu(dt) \int_t^{x_1} d\nu \int_t^{x_1} d\nu \\ &= \nu(x_0, x_1) \nu(x, x_1) \mu(0, x_0) + \nu(x, x_1) \int_{x_0}^x \nu(t, x_1) \mu(dt) \\ &\quad + \int_x^{x_1} \nu(t, x_1)^2 \mu(dt) \\ &=: \left(\nu(x_0, x_1) \mu(0, x_0) + H_1(x) \right) \nu(x, x_1), \end{aligned}$$

where

$$H_1(x) = \int_{x_0}^x \nu(t, x_1) \mu(dt) + \frac{1}{\nu(x, x_1)} \int_x^{x_1} \nu^2(t, x_1) \mu(dt), \quad x \in [x_0, x_1).$$

Noticing that $f_1^{x_0, x_1}(x) = \nu(x, x_1)$ for every $x \in [x_0, x_1)$, we have

$$\inf_{x_0 \leq x < x_1} \frac{f_2^{x_0, x_1}(x)}{f_1^{x_0, x_1}(x)} = \nu(x_0, x_1) \mu(0, x_0) + \inf_{x_0 \leq x < x_1} H_1(x) = \nu(x_0, x_1) \mu(0, x_0) + H_1(x_0).$$

In the last equality, we have used the fact that H_1 is non-decreasing on $[x_0, x_1]$. Indeed, fix $x, y \in [x_0, x_1]$ with $x < y$. Since $\nu(t, x_1)$ is decreasing in $t \in (x_0, x_1)$, we have

$$\frac{1}{\nu(x, x_1)} \int_x^y \nu^2(t, x_1) \mu(dt) \leq \int_x^y \nu(t, x_1) \mu(dt) \quad \text{and} \quad \frac{1}{\nu(y, x_1)} - \frac{1}{\nu(x, x_1)} > 0.$$

Moreover,

$$\begin{aligned} \frac{1}{\nu(x, x_1)} \int_x^y \nu^2(t, x_1) \mu(dt) &\leq \int_x^y \nu(t, x) \mu(dt) \\ &\quad + \left(\frac{1}{\nu(y, x_1)} - \frac{1}{\nu(x, x_1)} \right) \int_y^{x_1} \nu^2(t, x_1) \mu(dt), \end{aligned} \quad (27)$$

which implies that $H_1(x) \leq H_1(y)$. So H_1 is non-decreasing on $[x_0, x_1]$. Therefore,

$$\begin{aligned} \delta'_1 &= \sup_{x_0, x_1: x_0 < x_1} \inf_{x < x_1} II(f_1^{x_0, x_1})(x) = \sup_{x_0, x_1: x_0 < x_1} \inf_{x_0 \leq x < x_1} \frac{f_2^{x_0, x_1}(x)}{f_1^{x_0, x_1}(x)} \\ &= \sup_{x_0, x_1: x_0 < x_1} \left(\nu(x_0, x_1) \mu(0, x_0) + \inf_{x_0 \leq x < x_1} H_1(x) \right) \\ &= \sup_{x_0, x_1: x_0 < x_1} \left(\nu(x_0, x_1) \mu(0, x_0) + \frac{1}{\nu(x_0, x_1)} \int_{x_0}^{x_1} \nu^2(t, x_1) \mu(dt) \right) \end{aligned} \quad (28)$$

$$= \sup_{x_0 \in (0, D)} \left(\nu(x_0, D) \mu(0, x_0) + \frac{1}{\nu(x_0, D)} \int_{x_0}^D \nu^2(t, D) \mu(dt) \right). \quad (29)$$

In (29), we have used the fact that

$$H_2(x) := \nu(x_0, x) \mu(0, x_0) + \frac{1}{\nu(x_0, x)} \int_{x_0}^x \nu^2(t, x) \mu(dt), \quad x > x_0$$

is non-decreasing in x . In fact, for $x_0 < x < y$, $H_2(x) \leq H_2(y)$ if and only if

$$\nu(x, y) \mu(0, x_0) + \frac{1}{\nu(x_0, y)} \int_x^y \nu^2(t, y) \mu(dt) + \int_{x_0}^x \left(\frac{\nu^2(t, y)}{\nu(x_0, y)} - \frac{\nu^2(t, x)}{\nu(x_0, x)} \right) \mu(dt) \geq 0. \quad (30)$$

For $t \in [x_0, x]$, we have

$$\frac{\nu^2(t, y)}{\nu^2(t, x)} = \left(1 + \frac{\nu(x, y)}{\nu(t, x)} \right)^2 \geq 1 + \frac{\nu(x, y)}{\nu(t, x)} \geq 1 + \frac{\nu(x, y)}{\nu(x_0, x)} = \frac{\nu(x_0, y)}{\nu(x_0, x)},$$

and

$$\frac{\nu^2(t, y)}{\nu(x_0, y)} \geq \frac{\nu^2(t, x)}{\nu(x_0, x)},$$

So the inequality (30) follows, which implies that H_2 is non-decreasing in x .

Now, we compute $\bar{\delta}_1$. Noticing that

$$\begin{aligned}\|f_1^{x_0, x_1}\|^2 &= \int_0^{x_0} \nu^2(x_0, x_1) \mu(dx) + \int_{x_0}^{x_1} \nu^2(x, x_1) \mu(dx) \\ &= \nu^2(x_0, x_1) \mu(0, x_0) + \int_{x_0}^{x_1} \nu^2(x, x_1) \mu(dx), \\ D(f_1^{x_0, x_1}) &= \int_{x_0}^{x_1} e^{C(s)} ((f_1^{x_0, x_1})'(s))^2 ds = \int_{x_0}^{x_1} e^{-C(s)} ds = \nu(x_0, x_1).\end{aligned}$$

we obtain

$$\begin{aligned}\bar{\delta}_1 &= \sup_{x_0, x_1: x_0 < x_1} \frac{\|f_1^{x_0, x_1}\|^2}{D(f_1^{x_0, x_1})} \\ &= \sup_{x_0, x_1: x_0 < x_1} \left(\nu(x_0, x_1) \mu(0, x_0) + \frac{1}{\nu(x_0, x_1)} \int_{x_0}^{x_1} \nu^2(x, x_1) \mu(dx) \right).\end{aligned}$$

Comparing this with the expression of δ'_1 in (28), we obtain $\bar{\delta}_1 = \delta'_1$.

(f) At last, we show that $\delta'_1 \leq 2\delta$. Without loss of generality, assume that $\delta < \infty$. Using the integration by parts formula, we have

$$\begin{aligned}\int_0^x \nu^2(s \vee x_0, D) \mu(ds) &= \nu^2(x, D) \mu(0, x) + 2 \int_{x_0}^x \mu(0, s) \nu(s, D) e^{-C(s)} ds \\ &\leq \delta \nu(x, D) + 2\delta \int_{x_0}^x e^{-C(s)} ds, \quad x \geq x_0.\end{aligned}$$

By letting $x \rightarrow D$, we obtain

$$\int_0^D \nu^2(s \vee x_0, D) \mu(ds) \leq 2\delta \nu(x_0, D),$$

or equivalently,

$$\nu(x_0, D) \mu(0, x_0) + \frac{1}{\nu(x_0, D)} \int_{x_0}^D \nu^2(s, D) \mu(ds) \leq 2\delta, \quad x_0 \in (0, D).$$

Making supremum with respect to $x_0 \in (0, D)$ on the both sides of the inequality, the assertion that $\delta'_1 \leq 2\delta$ follows from (29) immediately. $\square$

Appendix B Complement of the proofs in section 4

B.1 Proof of theorem 17

Similar to the ND situation, we adopt two circle arguments follows.

$$\tilde{\lambda}_0 \geq \lambda_0 \quad (31)$$

$$\begin{aligned} &\geq \sup_{f \in \tilde{\mathcal{F}}_H} \inf_{x \in (0,D)} II(f)(x)^{-1} = \sup_{f \in \tilde{\mathcal{F}}_I} \inf_{x \in (0,D)} II(f)(x)^{-1} \\ &= \sup_{f \in \tilde{\mathcal{F}}_I} \inf_{x \in (0,D)} I(f)(x)^{-1} \end{aligned} \quad (32)$$

$$\geq \sup_{h \in \mathcal{H}} \inf_{x \in (0,D)} R(h)(x) \quad (33)$$

$$\geq \tilde{\lambda}_0. \quad (34)$$

and

$$\lambda_0 \leq \inf_{f \in \tilde{\mathcal{F}}_H \cup \tilde{\mathcal{F}}'_H} \sup_{x \in (0,D)} II(f)(x)^{-1} \quad (35)$$

$$\begin{aligned} &\leq \inf_{f \in \tilde{\mathcal{F}}_H} \sup_{x \in (0,D)} II(f)(x)^{-1} = \inf_{f \in \tilde{\mathcal{F}}_I} \sup_{x \in (0,D)} II(f)(x)^{-1} \\ &= \inf_{f \in \tilde{\mathcal{F}}_I} \sup_{x \in (0,D)} I(f)(x)^{-1} \end{aligned} \quad (36)$$

$$\leq \inf_{h \in \tilde{\mathcal{H}}} \sup_{x \in (0,D)} R(h)(x) \quad (37)$$

$$\leq \lambda_0. \quad (38)$$

The assertions below are proved in [3; Theorem 1.1] and [4; Chapter 6].

$$\lambda_0 \geq \sup_{f \in \tilde{\mathcal{F}}_H} \inf_{x \in (0,D)} II(f)(x)^{-1} = \sup_{f \in \tilde{\mathcal{F}}_I} \inf_{x \in (0,D)} I(f)(x)^{-1} = \sup_{f \in \tilde{\mathcal{F}}_I} \inf_{x \in (0,D)} II(f)(x)^{-1}; \quad (39)$$

$$\lambda_0 \leq \inf_{f \in \tilde{\mathcal{F}}_H} \sup_{x \in (0,D)} II(f)(x)^{-1} = \inf_{f \in \tilde{\mathcal{F}}_I} \sup_{x \in (0,D)} I(f)(x)^{-1} = \inf_{f \in \tilde{\mathcal{F}}_I} \sup_{x \in (0,D)} II(f)(x)^{-1}. \quad (40)$$

Actually, from [3, 4], it is known that

$$\lambda_0 \geq \sup_{f \in \tilde{\mathcal{F}}_H} \inf_{x \in (0,D)} II(f)(x)^{-1} = \sup_{f \in \tilde{\mathcal{F}}_I} \inf_{x \in (0,D)} I(f)(x)^{-1},$$

and

$$\lambda_0 \leq \inf_{f \in \tilde{\mathcal{F}}_H} \sup_{x \in (0,D)} II(f)(x)^{-1} = \inf_{f \in \tilde{\mathcal{F}}_I} \sup_{x \in (0,D)} I(f)(x)^{-1}.$$

Thus, (39) holds since $\tilde{\mathcal{F}}_I \subseteq \tilde{\mathcal{F}}_H$ and

$$\sup_{x \in (0,D)} II(f)(x) \leq \sup_{x \in (0,D)} I(f)(x)$$

by Cauchy's mean value theorem. (40) holds for the similar reason:

$$\tilde{\mathcal{F}}_I \subseteq \tilde{\mathcal{F}}_H \quad \text{and} \quad \inf_{x \in (0,D)} II(f)(x) \geq \inf_{x \in (0,D)} I(f)(x).$$

In particular, we have known (31), (32) and (36) since the inequalities in (31) and (36) are obvious. It remains to prove (33)–(35), (37) and (38).

We now begin to work on the additional part of the proof under the assumption that $a, b \in \mathcal{C}$ except proof (c) below.

(a) Prove that

$$\sup_{f \in \mathcal{F}_H} \inf_{x \in (0, D)} H(f)(x)^{-1} \geq \sup_{h \in \mathcal{H}} \inf_{x \in (0, D)} R(h)(x).$$

Given $h \in \mathcal{H}$, let

$$g(x) = g(\varepsilon) \exp \left[\int_{\varepsilon}^x h(u) du \right], \quad x \in (0, D)$$

for a fixed $\varepsilon > 0$. Then $g \in \mathcal{C}^2(0, D) \cap \mathcal{C}[0, D]$,

$$g(0) = 0, \quad g' > 0, \quad \text{and} \quad h = \frac{g'}{g} \quad \text{on } (0, D).$$

Furthermore,

$$R(h) = -(ah^2 + bh + ah') = -\frac{Lg}{g}.$$

To show that

$$\inf_{x \in (0, D)} R(h)(x) \leq \sup_{f \in \mathcal{F}_H} \inf_{x \in (0, D)} H(f)(x)^{-1} \quad \text{for every } h \in \mathcal{H},$$

without loss of generality, assume that $\inf_{x \in (0, D)} R(h)(x) > 0$. This ensures

$$f := -(ag'' + bg') = gR(h) > 0.$$

Then $f > 0$ and $f \in \mathcal{C}[0, D]$ since $a, b \in \mathcal{C}[0, D]$. Since $f = -(ag'' + bg') = Lg$ and $g'(D) \geq 0$, by (9), we have

$$g'(s) \geq g'(s) - g'(D) = e^{-C(s)} \int_s^D f d\mu, \quad s \in (0, D).$$

Moreover, we obtain

$$g(x) \geq \int_0^x \nu(ds) \int_s^D f d\mu = f(x) H(f)(x), \quad x \in (0, D)$$

since $g(0) = 0$. Thus,

$$R(h)(x)^{-1} \geq \frac{g(x)}{f(x)} \geq H(f)(x), \quad x \in (0, D).$$

Therefore,

$$\inf_{x \in (0, D)} R(h)(x) \leq \inf_{x \in (0, D)} II(f)(x)^{-1} \leq \sup_{f \in \widetilde{\mathcal{F}}_H} \inf_{x \in (0, D)} II(f)(x)^{-1}.$$

The assertion follows since h is arbitrary.

(b) Prove that $\sup_{h \in \mathcal{H}} \inf_{x \in (0, D)} R(h)(x) \geq \tilde{\lambda}_0$.

First, we show that

$$\sup_{h \in \mathcal{H}} \inf_{x \in (0, D)} R(h)(x) \geq 0.$$

For a given positive $f \in L^1(\mu)$, let $g = fII(f)$. Then

$$g'(x) = e^{-C(x)} \int_x^D f d\mu > 0.$$

Let $h = g'/g$. By simple calculation, we get $-f = ag'' + bg'$ and

$$R(h) = -(ah^2 + bh + ah') = -\frac{Lg}{g} = \frac{f}{g} > 0 \quad \text{on } (0, D).$$

This implies $\inf_{x \in (0, D)} R(h)(x) \geq 0$ and the required assertion follows.

When $\lambda_0 > 0$, It was proved in [1; Theorem 2.2] and [4; Proof (d) of Theorem 3.7] (also mentioned in the proofs of [3; Theorem 1.2]) that the eigenfunction of λ_0 is strictly increasing. Even though $\tilde{\lambda}_0$ could be formally bigger than λ_0 , the same proofs still work for the eigenfunction g of $\tilde{\lambda}_0$ since the constructed function g used there satisfies $g = g(\cdot \wedge x_0)$ for some $x_0 \in (0, D)$. Hence, there exists an eigenfunction g such that

$$Lg = -\tilde{\lambda}_0 g, \quad g(0) = 0, \quad \text{and} \quad g \in \mathcal{C}^2(0, D) \cap \mathcal{C}[0, D].$$

Let $h = g'/g$. Then $h \in \mathcal{C}^1(0, D) \cap \mathcal{C}[0, D]$, $h \in \mathcal{H}$ and

$$R(h)(x) = -\frac{Lg(x)}{g(x)} = \tilde{\lambda} \quad \text{for } x \in (0, D).$$

So the assertion follows.

(c) Prove that $\lambda_0 \leq \inf_{f \in \widetilde{\mathcal{F}}_H \cup \widetilde{\mathcal{F}}_H'} \sup_{x \in (0, D)} II(f)(x)^{-1}$.

When $D = \infty$, this is almost done in the original proof of [3; Theorem 1.1] except that one requires an additional condition $g \in L^2(\mu)$, provided $x_0 = \infty$ is allowed. This is the reason why the set $\widetilde{\mathcal{F}}_H'$ is added. Anyhow the proof is similar to that of Theorem 1 presented in Section 3.

(d) Prove that

$$\inf_{f \in \widetilde{\mathcal{F}}_H} \sup_{x \in (0, D)} II(f)(x)^{-1} \leq \inf_{h \in \mathcal{H}} \sup_{x \in (0, D)} R(h)(x).$$

First, for $h \in \widetilde{\mathcal{H}}$, $\exists x_0 \in (0, D)$ such that $h|_{(0, x_0)} > 0$, $\int_{0+} h(u) du = \infty$, and $h|_{[x_0, D]} = 0$. Similar to the proof (b) above, given g , we change the form of $R(h)$ on $(0, x_0)$. Thus,

$$R(h) = \begin{cases} -(ag'' + bg')/g, & \text{on } (0, x_0); \\ 0, & \text{otherwise.} \end{cases}$$

Next, for $h \in \widetilde{\mathcal{H}}$, let

$$f(x) = [gR(h)](x) = -ae^{-C}(e^C g')'(x) \quad \text{for } x \leq x_0$$

$f(x) = f(x_0)$ for $x > x_0$. Then $f \in \widetilde{\mathcal{F}}_H$ since $a, b \in \mathcal{C}[0, D]$, and

$$e^{C(x)} g'(x) = \int_x^{x_0} f d\mu + e^{C(x_0)} g'(x_0) \quad \text{for } x \leq x_0,$$

$$e^{C(x_0)} g'(x_0) = \int_{x_0}^D f d\mu.$$

Moreover, we have $g'(x) = e^{-C(x)} \int_x^D f d\mu$, which implies that

$$g(x) = \int_0^x \nu(ds) \int_s^D f d\mu$$

and

$$R(h)(x)^{-1} = \frac{g(x)}{f(x)} \leq H(f)(x) \quad \text{for } x \in (0, x_0).$$

Therefore, we get

$$\begin{aligned} \sup_{x \in (0, x_0)} R(h)(x) &\geq \sup_{x \in (0, x_0)} H(f)(x)^{-1} \\ &\geq \inf_{f \in \widetilde{\mathcal{F}}_H, f=f \cdot \wedge_{x_0}} \sup_{x \in (0, x_0)} H(f)(x)^{-1} \\ &\geq \inf_{f \in \widetilde{\mathcal{F}}_H} \sup_{x \in (0, D)} H(f)(x)^{-1}. \end{aligned}$$

Furthermore, we obtain

$$\inf_{f \in \widetilde{\mathcal{F}}_H} \sup_{x \in (0, D)} H(f)(x)^{-1} \leq \inf_{h \in \widetilde{\mathcal{H}}} \sup_{x \in (0, D)} R(h)(x).$$

(e) Prove that $\inf_{h \in \widetilde{\mathcal{H}}} \sup_{x \in (0, D)} R(h)(x) \leq \lambda_0$.

Recall the definition of λ_0 :

$$\begin{aligned} \lambda_0 = \inf \{ D(f) : \mu(f^2) = 1, f(0) = 0, f = f(\cdot \wedge x_0), f \in \mathcal{C}^1(0, x_0) \cap \mathcal{C}[0, x_0] \\ \text{for some } x_0 \in (0, D) \} =: \tilde{\lambda}_0. \end{aligned}$$

Let $p_n \uparrow D$ and denote by $\lambda_0^{(0, p_n)}$ the corresponding eigenvalue determined by $L|_{(0, p_n)}$ (The same as the proof of [3; Theorem 1.2]. Then $\lambda_0^{(0, p_n)} \downarrow \lambda_0$ by using the proof of [8; Lemma 5.1]. $\square$

B.2 Proof of Theorem 19

(a) We remark that the sequence $\{f_n^{(x_0)}\}_{n \in \mathbb{N}}$ is clearly contained in $\widetilde{\mathcal{F}}_I$. But the modified sequence used in [3; Theorem 1.2]:

$$\widetilde{f}_1^{(x_0)} = \varphi(\cdot \wedge x_0), \quad \widetilde{f}_n^{(x_0)} = \widetilde{f}_{n-1}^{(x_0)}(\cdot \wedge x_0) II(\widetilde{f}_{n-1}^{(x_0)}(\cdot \wedge x_0)), \quad n \geq 2,$$

is usually not contained in $\widetilde{\mathcal{F}}_{II}$. However,

$$\begin{aligned} \delta'_n &= \sup_{x_0 \in (0, D)} \inf_{x \in (0, D)} II(f_n^{(x_0)})(x) \\ &= \sup_{x_0 \in (0, D)} \inf_{x \in (0, x_0)} II(f_n^{(x_0)})(x) \\ &= \sup_{x_0 \in (0, D)} \inf_{x \in (0, x_0)} II(\widetilde{f}_n^{(x_0)}(\cdot \wedge x_0))(x) \\ &= \sup_{x_0 \in (0, D)} \inf_{x \in (0, D)} II(\widetilde{f}_n^{(x_0)}(\cdot \wedge x_0))(x). \end{aligned}$$

Here in the last step we have used the convention that $1/0 = \infty$. Hence, these two sequences produce the same $\{\delta'_n\}$. The assertions about δ_n and δ'_n were proved in [3; Theorem 1.2].

(b) Prove that $\bar{\delta}_{n+1} \geq \delta'_n$ and $\bar{\delta}_n^{-1} \geq \lambda_0$ ($n \geq 1$).

The assertion of $\bar{\delta}_n^{-1} \geq \lambda_0$ is obvious since every function in $\{f_n^{(x_0)} : n \geq 1\}$ is a test function of $\tilde{\lambda}_0$ and $\tilde{\lambda}_0 = \lambda_0$. Similar to the ND case, it is easy to see that $\bar{\delta}_{n+1} \geq \delta'_n$, which is a consequence of the proof of [3; Theorem 1.1].

Indeed, when proving

$$\lambda_0 \leq \inf_{f \in \widetilde{\mathcal{F}}_{II}} \sup_{x \in (0, D)} II(f)(x)^{-1}$$

(i.e., ξ'_0 there), we know that $g = [f II(f)](\cdot \wedge x_0)$ satisfies

$$D(g)/\mu(g^2) \leq \sup_{x \in (0, D)} II(f)(x)^{-1}.$$

By the relation between $f_n^{(x_0)}$ and $f_{n+1}^{(x_0)}$, we have

$$\frac{\|f_{n+1}^{(x_0)}\|}{D(f_{n+1}^{(x_0)})} \geq \inf_{x \in (0, D)} II(f_n^{(x_0)})(x),$$

which implies that

$$\bar{\delta}_{n+1} = \sup_{x_0 \in (0, D)} \frac{\|f_{n+1}^{(x_0)}\|}{D(f_{n+1}^{(x_0)})} \geq \delta'_n.$$

The assertion that $\delta'_1 = \bar{\delta}_1$ is proved in the appendix B.3 below. $\square$

B.3 Proof of Corollary 20

The degenerated case that $\mu(0, D) = \infty$ is trivial since $\lambda_0 = 0$ and

$$\delta = \delta_1 = \delta'_1 = \infty.$$

The main assertion of Corollary 20 is a consequence of Theorem 19. Here, we compute δ_1 , δ'_1 and prove that $\delta'_1 \in [\delta, 2\delta]$. Compute δ_1 first.

Since

$$\begin{aligned} \int_0^x \nu(dt) \int_t^D \sqrt{\varphi} d\mu &= \int_0^x \nu(dt) \int_t^x \sqrt{\varphi} d\mu + \int_0^x \nu(dt) \int_x^D \sqrt{\varphi} d\mu \\ &= \int_0^x \sqrt{\varphi(s)} \mu(ds) \int_0^s d\nu + \varphi(x) \int_x^D \sqrt{\varphi} d\mu \\ &= \int_0^x \sqrt{\varphi} \varphi d\mu + \varphi(x) \int_x^D \sqrt{\varphi} d\mu \\ &= \int_0^D \sqrt{\varphi(s)} \varphi(s \wedge x) \mu(ds), \end{aligned}$$

we have

$$\begin{aligned} \delta_1 &= \sup_{x \in (0, D)} II(\sqrt{\varphi})(x) \\ &= \sup_{x \in (0, D)} \left(\frac{1}{\sqrt{\varphi(x)}} \int_0^x \sqrt{\varphi} \varphi d\mu + \sqrt{\varphi(x)} \int_x^D \sqrt{\varphi} d\mu \right) \\ &= \sup_{x \in (0, D)} \frac{1}{\sqrt{\varphi(x)}} \int_0^D \sqrt{\varphi(s)} \varphi(s \wedge x) \mu(ds). \end{aligned}$$

Now, we compute δ'_1 . Note that

$$II(f_1^{(x_0)})(x) = \frac{1}{\varphi(x \wedge x_0)} \int_0^x e^{-C(t)} dt \int_t^D \varphi(s \wedge x_0) \mu(ds).$$

The right-hand side is clearly increasing in x for $x \geq x_0$ and decreasing for $x \leq x_0$. Hence, $II(f_1^{(x_0)})$ achieves its minimum at $x = x_0$. By exchanging the order of the integrals, its minimum is equal to

$$\frac{1}{\varphi(x_0)} \int_0^D \varphi^2(s \wedge x_0) \mu(ds).$$

So

$$\delta'_1 = \sup_{x_0 \in (0, D)} \inf_{x \in (0, D)} II(f_1^{(x_0)})(x) = \sup_{x_0 \in (0, D)} \frac{1}{\varphi(x_0)} \int_0^D \varphi^2(s \wedge x_0) \mu(ds).$$

Next, following the proof in the discrete case [5], we have

$$D(f_1^{(x_0)}) = \int_0^D e^{C(x)} [\varphi'(x_0 \wedge x)]^2 dx = \int_0^{x_0} e^{C(x)} (e^{-C(x)})^2 dx = \varphi(x_0),$$

and

$$\mu(f_1^{(x_0)}) = \int_0^D \varphi^2(x_0 \wedge x) \mu(dx).$$

Thus,

$$\bar{\delta}_1 = \sup_{x_0 \in (0, D)} \frac{\mu[(f_1^{(x_0)})^2]}{D(f_1^{(x_0)})} = \delta'_1.$$

At last, we prove that $\delta'_1 \in [\delta, 2\delta]$. Following the corresponding proof in the discrete case [5; Corollary 4.4], we have

$$\delta'_1 = \sup_{x_0 \in (0, D)} \frac{1}{\varphi(x_0)} \int_0^D \varphi^2(s \wedge x_0) \mu(ds) \geq \sup_{x \in (0, D)} \varphi(x) \mu(x, D) = \delta.$$

On the other hand, using the integration by parts formula, for $x < x_0$, we have

$$\int_x^{x_0} \varphi^2(s) \mu(ds) = -\varphi^2(s) \mu(s, D) \Big|_x^{x_0} + 2 \int_x^{x_0} \varphi(s) \varphi'(s) \mu(s, D) ds.$$

So

$$\begin{aligned} \int_x^D \varphi^2(s \wedge x_0) \mu(ds) &= \int_x^{x_0} \varphi^2(s) \mu(ds) + \varphi^2(x_0) \mu(x_0, D) \\ &= \varphi^2(x) \mu(x, D) + 2 \int_x^{x_0} \varphi(s) \varphi'(s) \mu(s, D) ds \\ &\leq \delta \varphi(x) + 2\delta \int_x^{x_0} e^{-C(s)} ds \\ &\rightarrow 2\delta \varphi(x_0) \quad \text{as } x \rightarrow 0. \end{aligned}$$

Thus $\int_0^D \varphi^2(s \wedge x_0) \mu(ds) / \varphi(x_0) \leq 2\delta$ and the assertion $\delta'_1 \leq 2\delta$ follows immediately. $\square$

The optimal constant in Hardy-type inequalities

Mu-Fa Chen

(Beijing Normal University, Beijing 100875, China)

March 19, 2013

Abstract

To estimate the optimal constant in Hardy-type inequalities, some variational formulas and approximating procedures are introduced. The known basic estimates are improved considerably. The results are illustrated by typical examples. It is shown that the sharp factor is meaningful for each finite interval and a classical sharp model is re-examined.

1 Introduction

For given two Borel measures μ and ν on an interval $(-M, N)$ ($M, N \leq \infty$), the Hardy-type inequality says that the $L^q(\mu)$ -norm of each absolutely continuous function f is controlled from above by the $L^p(\nu)$ -norm of its derivative f' up to a constant factor A :

$$\left(\int_{-M}^N |f|^q d\mu \right)^{1/q} \leq A \left(\int_{-M}^N |f'|^p d\nu \right)^{1/p}, \quad (1)$$

i.e., $\|f\|_{\mu,q} \leq A \|f'\|_{\nu,p}$, $p, q \in (1, \infty)$.

The inequalities have been well studied in the past decades, cf. [15, 10, 7]. In particular, with the constraint that f vanishes at $M = 0$, the following basic estimates for the optimal constant A in (1) are well known:

$$B \leq A \leq k_{q,p} B, \quad (2)$$

where B is a quantity described by M, N, μ, ν, p and q , and $k_{q,p} \in [1, 2]$ is a constant factor (cf. (3), (22) and (23) below. See also Appendix for more details). The goal of this paper is to show that there is still a room for improvements of (2). Such a qualitative study is valuable since the optimal

2000 *Mathematics Subject Classifications.* 26D10, 60J60, 34L15.

Key words and phrases. Hardy-type inequality, optimal constant, variational formulas, approximating procedure.

constant A describes the speed of some type of stability (cf. [6; Chapter 6] and references therein). We begin our story with an example which is typical in the sense that it is the only one, except the special case that $p = q = 2$, we have known so far for having the exact constant A for all $p, q \in (1, \infty)$ (see also Example 2.5 and Proposition 4.5 for additional information).

There are mainly four different types of boundary conditions in (1). We concentrate in the paper only on the one vanishing at $-M$. The other cases will be handled subsequently. In particular, the results of this paper are extended by [12] to the discrete context.

Example 1.1 Let $(-M, N) = (0, 1)$ and $d\mu = d\nu = dx$. Then

(1) the optimal constant A in (1) is

$$A = \frac{p^{\frac{1}{q}} q^{1-\frac{1}{p}} (pq + p - q)^{\frac{1}{p}-\frac{1}{q}}}{(p-1)^{\frac{1}{p}} B\left(\frac{1}{q}, 1 - \frac{1}{p}\right)}, \quad p, q \in (1, \infty),$$

where $B(\alpha, \beta)$ is the Beta function

$$B(\alpha, \beta) = \int_0^1 s^{\alpha-1} (1-s)^{\beta-1} ds = \frac{1}{\alpha} \int_0^1 (1-t^{\frac{1}{\alpha}})^{\beta-1} dt \quad (\text{change variable } t = s^\alpha).$$

In particular, if $q = p$, then

$$A = \frac{p}{\pi(p-1)^{1/p}} \sin \frac{\pi}{p}.$$

More particularly, $A = 2/\pi$ if $q = p = 2$.

(2) *Basic estimates.* The constants used for the basic estimates in (2) for $q \geq p$ are as follows:

$$B = \frac{p^{\frac{1}{q}} ((p-1)q)^{1-\frac{1}{p}}}{(pq + p - q)^{1-\frac{1}{p}+\frac{1}{q}}},$$

$$k_{q,p} = \left[\frac{\Gamma\left(\frac{pq}{q-p}\right)}{\Gamma\left(\frac{q}{q-p}\right)\Gamma\left(\frac{p(q-1)}{q-p}\right)} \right]^{1/p-1/q} = \left[\frac{q-p}{pB\left(\frac{p}{q-p}, \frac{p(q-1)}{q-p}\right)} \right]^{1/p-1/q}, \quad q > p, \quad (3)$$

cf. [7; Example 1.12] and references therein. In particular, if $q = p$, then

$$B = (1/p)^{1/p} (1/p^*)^{1/p^*}, \quad k_{p,p} = p^{1/p} p^{*1/p^*},$$

where p^* is the conjugate of p : $1/p + 1/p^* = 1$.

(3) *Improvements.* As the first step of our approximating procedures introduced in the paper, we have new upper and lower bounds δ_1 and $\bar{\delta}_1$, respectively. Besides, we also have another upper estimate A^* . More precisely, the basic estimates in (2) are improved in the paper by

$$B \leq \bar{\delta}_1 \leq A \leq A^* \leq \delta_1 \leq k_{q,p} B, \quad (4)$$

where

$$\bar{\delta}_1 = \frac{p^{1/q}((p-1)(q+1))^{1-1/p}}{(pq+p-q)^{1-1/p+1/q}}, \quad (5)$$

$$A^* = \left(\frac{p^*}{q}\right)^{1/q} \left(\frac{p^*+q}{\pi p^*} \sin \frac{\pi p^*}{p^*+q}\right)^{1/p^*+1/q}, \quad (6)$$

$$\delta_1 = \frac{1}{(q\gamma^*/p^*+1)^{1/q}} \left\{ \sup_{x \in (0,1)} \frac{1}{x\gamma^*} \int_0^x (1-y^{q\gamma^*/p^*+1})^{p^*/q} dy \right\}^{1/p^*}, \quad \gamma^* := \frac{q}{p^*+q}. \quad (7)$$

In the last formula, the function under $\sup_{x \in (0,1)}$ is unimodal on $(0,1)$, its integral term is indeed an incomplete Beta function:

$$B(x, \alpha, \beta) = \int_0^x s^{\alpha-1} (1-s)^{\beta-1} ds = \frac{1}{\alpha} \int_0^{x^\alpha} (1-t^{1/\alpha})^{\beta-1} dt.$$

Note that $\bar{\delta}_1$ is very much the same as B : the factor $q+1$ in $\bar{\delta}_1$ is replaced by q in B . Besides, $A^* = A$ if $q = p$. Except these facts, the comparison of the quantities in (4) are non-trivial, as shown by Figures 1–4.

(4) *Figures.* First, consider the case that $q = p$. Figure 1 shows the basic estimates of the optimal constant A .

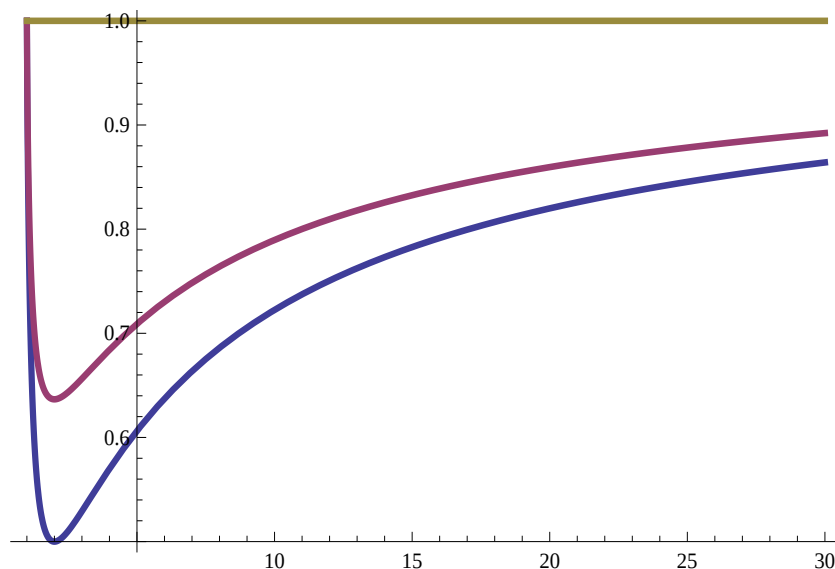


Figure 1 The middle curve is the exact value of A . The top straight line and the bottom curve consist of the basic estimates of A .

Our improved upper bound δ_1 and lower one $\bar{\delta}_1$ are added to Figure 1, as shown in Figure 2.

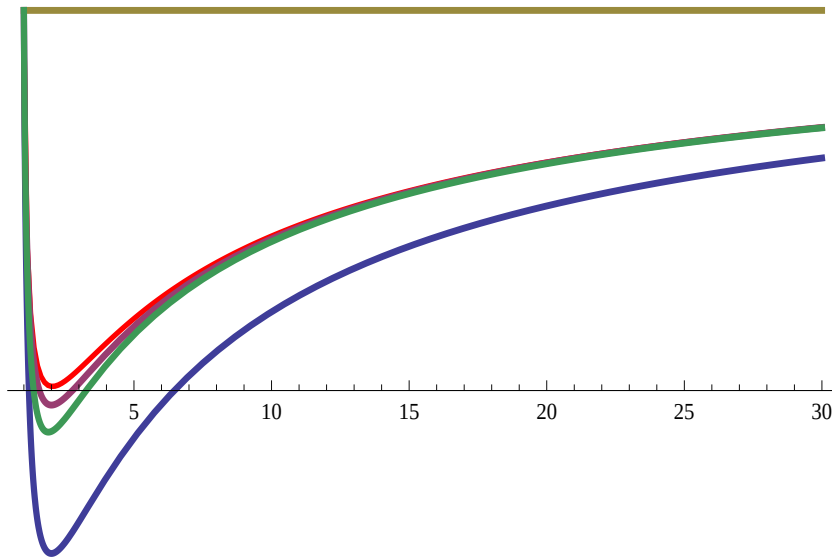


Figure 2 The new bounds δ_1 and $\bar{\delta}_1$ are almost overlapped with the exact value A except in a small neighborhood of $p = 2$, δ_1 is a little bigger, and $\bar{\delta}_1$ is a little smaller than A .

Next, consider the case that $q > p$. For convenience, we rewrite q as $p + r$, where r varies over $(0, 15)$. The six quantities in (4) are shown in Figures 3 and 4 according to $p = 2$ and $p = 5$, respectively.

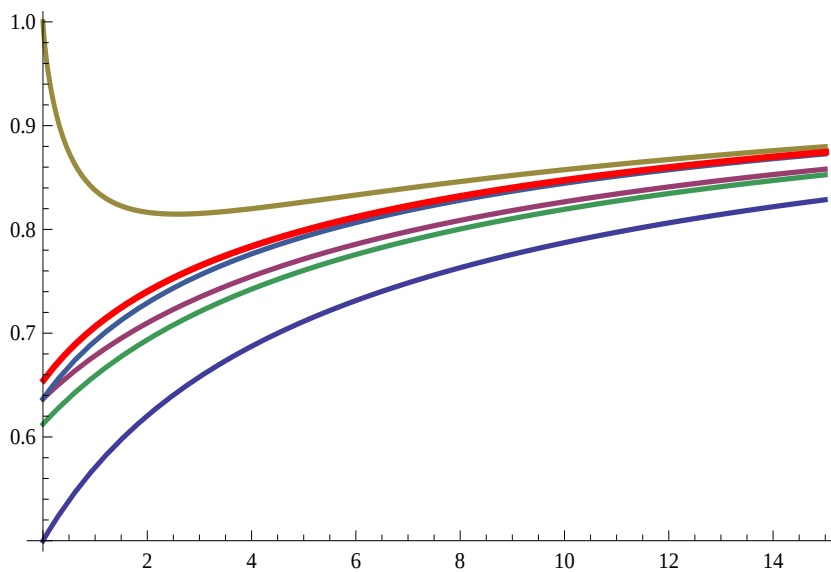


Figure 3 The six curves from top to bottom are $k_{q,p}B$, δ_1 , A^* , A (the exact value), $\bar{\delta}_1$, and B , respectively, for $p = 2$, $q = p + r$, and $r \in (0, 15)$.

In view of Figures 1 and 2, it is clear that the six curves should be closer for larger p .

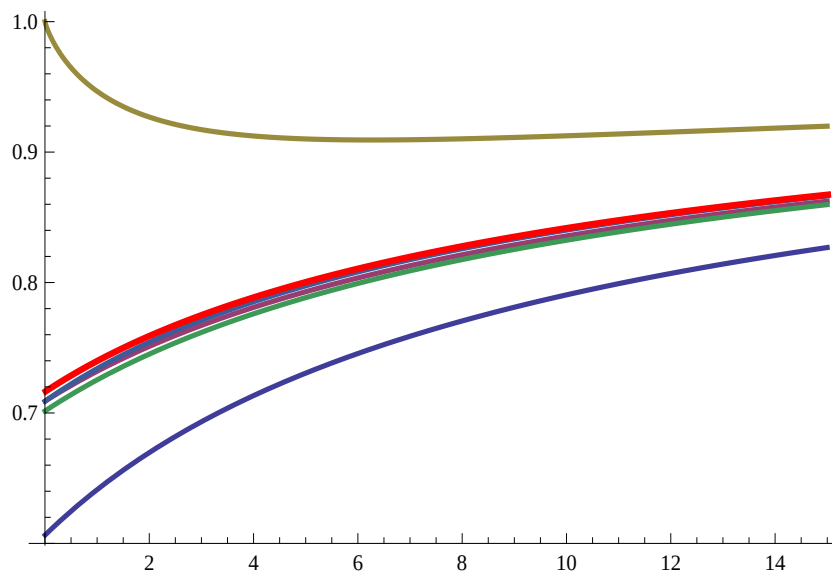


Figure 4 The only change of this figure from the last one is replacing $p = 2$ by $p = 5$. Certainly, the six curves are located in the same order. Except the basic estimates, the other four curves are almost overlapped.

Figures 1–4 illustrate the effectiveness of our improvements. It is surprising and unexpected that the new estimates can be so closed to the exact value. The general results are presented in the next section. Their proofs are given in Section 3. In Appendix, we will come back to study the basic estimates (2) and the optimal factor $k_{q,p}$.

To conclude this section, we make some historical remarks on Example 1.1. The optimal constant given in the example was presented in [16; page 357] with optimizer but without details. The detailed proofs were presented in [9] and [1]. We mention that the boundary condition used in the cited papers are vanishing at both endpoints. This is using the Dirichlet boundaries at two endpoints. The result is the same if we replace Dirichlet boundaries with Neumann ones (cf. [7]). However, as mentioned before, we consider only the Dirichlet boundary at the left-endpoint in this paper. Then the optimal constant here is a double of those given in the cited papers. For more recent progress on p -Laplacian (which is an alternative description of the Hardy-type inequality in the case of $q = p$), one may refer to the book [11] and references therein. Actually, in this case, the story is now quite complete. The new progress will be published elsewhere in [8].

2 Main results

From now on, for simplicity, we fix $(-M, N) = (0, D)$, $D \leq +\infty$. Set

$$\begin{aligned}\mathcal{A}[0, D] &= \{f : f \text{ is absolutely continuous in } [0, D] \text{ (or } [0, D) \text{ if } D = \infty)\}, \\ \mathcal{A}_0[0, D] &= \{f \in \mathcal{A}[0, D] : f(0) = 0\}.\end{aligned}$$

Then the optimal constant A in (1) is described by the following classical variational formula

$$A = \sup_{f \in \mathcal{A}_0[0,D], \|f'\|_{\nu,p} \in (0,\infty)} \frac{\|f\|_{\mu,q}}{\|f'\|_{\nu,p}} \quad (8)$$

To state our results, we need some notation. Denote by v the density of the absolutely continuous part of ν with respect to the Lebesgue measure dx and let

$$\hat{v} = v^{-\frac{1}{p-1}} = v^{1-p^*}.$$

For upper estimates, define two operators II^* and I^* :

$$II^*(f)(x) = \frac{1}{f(x)} \int_0^x dy \hat{v}(y) \left(\int_y^D f^{q/p^*} d\mu \right)^{p^*/q}, \quad x \in (0, D), \quad (9)$$

$$I^*(f)(x) = \frac{\hat{v}}{f'}(x) \left(\int_x^D f^{q/p^*} d\mu \right)^{p^*/q}, \quad x \in (0, D), \quad (10)$$

with domains

$$\mathcal{F}_{II} = \{f : f(0) = 0, f > 0 \text{ on } (0, D)\}, \quad (11)$$

$$\mathcal{F}_I = \{f : f(0) = 0, f' > 0 \text{ on } (0, D)\}, \quad (12)$$

respectively. For lower estimates, we need different operators:

$$II(f)(x) = \frac{1}{f(x)} \int_0^x dy \hat{v}(y) \left(\int_y^D f^{q-1} d\mu \right)^{p^*-1}, \quad x \in (0, D), \quad (13)$$

$$I(f)(x) = \frac{\hat{v}}{f'}(x) \left(\int_x^D f^{q-1} d\mu \right)^{p^*-1}, \quad x \in (0, D). \quad (14)$$

When $q = p$, we have $II = II^*$ and $I = I^*$. To avoid the non-integrability problem, the domain of II and I have to be modified from $\mathcal{F}_{II}$ and $\mathcal{F}_I$:

$$\begin{aligned} \widetilde{\mathcal{F}}_{II} = \{f \in \mathcal{F}_{II} : \exists x_0 \in (0, D] \text{ such that } f = f(\cdot \wedge x_0) \text{ and moreover} \\ f II(f) \in L^q(\mu) \text{ if } x_0 = D\}, \end{aligned} \quad (15)$$

$$\begin{aligned} \widetilde{\mathcal{F}}_I = \{f : f(0) = 0, \exists x_0 \in (0, D] \text{ such that } f = f(\cdot \wedge x_0), f' > 0 \text{ on } (0, x_0), \\ \text{and moreover } f II(f) \in L^q(\mu) \text{ if } x_0 = D\}, \end{aligned} \quad (16)$$

where $\alpha \wedge \beta = \min\{\alpha, \beta\}$ and similarly, $\alpha \vee \beta = \max\{\alpha, \beta\}$. Thus, the operators we actually use for the lower estimates are modified from II and I as follows: when $f = f(\cdot \wedge x_0)$, set

$$II^\sim(f)(x) = II(f)(x \wedge x_0) = \frac{1}{f(x)} \int_0^{x \wedge x_0} dy \hat{v}(y) \left(\int_y^D f^{q-1} d\mu \right)^{p^*-1}, \quad (17)$$

$$I^\gamma(f)(x) = I(f)(x \wedge x_0) = \frac{\hat{v}}{f'}(x \wedge x_0) \left(\int_{x \wedge x_0}^D f^{q-1} d\mu \right)^{p^*-1}, \quad x \in (0, D). \quad (18)$$

Here we adopt the usual convention that $1/0 = \infty$.

We can now state our variational result.

Theorem 2.1 For the optimal constant A in (1), we have

(1) upper estimate:

$$A \leq \inf_{f \in \mathcal{F}_H} \left[\sup_{x \in (0, D)} II^*(f)(x) \right]^{1/p^*} = \inf_{f \in \mathcal{F}_I} \left[\sup_{x \in (0, D)} I^*(f)(x) \right]^{1/p^*} \quad (19)$$

for $q \geq p$, and

(2) lower estimate:

$$A \geq \sup_{f \in \mathcal{F}_H} \|f II^\gamma(f)\|_{\mu, q}^{1-q/p} \left(\inf_{x \in (0, D)} II^\gamma(f)(x) \right)^{(q-1)/p} \quad (20)$$

for $p, q \in (1, \infty)$. In particular, when $q = p$, we have additionally that

$$\sup_{f \in \mathcal{F}_H} \left(\inf_{x \in (0, D)} II^\gamma(f)(x) \right)^{1/p^*} = \sup_{f \in \mathcal{F}_I} \left(\inf_{x \in (0, D)} I^\gamma(f)(x) \right)^{1/p^*}. \quad (21)$$

Recall that for general $q \geq p$, $p, q \in (1, \infty)$, the basic estimates read as follows

$$B \leq A \leq k_{q,p} B \quad (22)$$

where $k_{q,p}$ is given in (3) and

$$B = \sup_{x \in (0, D)} \hat{v}(0, x)^{1/p^*} \mu(x, D)^{1/q}, \quad (23)$$

here $\hat{v}(\alpha, \beta) = \int_\alpha^\beta \hat{v}$ as usual, and similar for $\mu(\alpha, \beta)$ (cf. [10; pages 45–47] and Appendix below).

As an application of Theorem 2.1, we have the following approximating procedures.

Theorem 2.2 (1) Let $q \geq p$, $p, q \in (1, \infty)$,

$$f_1(x) = \hat{v}(0, x)^{\gamma^*}, \quad \gamma^* = \frac{q}{p^* + q}, \quad (24)$$

and define

$$f_{n+1}(x) = f_n II(f_n) = \int_0^x dy \hat{v}(y) \left(\int_y^D f_n^{q/p^*} d\mu \right)^{p^*/q}, \quad n \geq 1, \quad (25)$$

$$\delta_n = \begin{cases} \left(\sup_{x \in (0, D)} \frac{f_{n+1}(x)}{f_n(x)} \right)^{1/p^*}, & n = 1 \text{ or } n \geq 2 \text{ but } \delta_1 < \infty \\ \infty, & n \geq 2 \text{ and } \delta_1 = \infty. \end{cases} \quad (26)$$

Then we have $A \leq \delta_n$ for all $n \geq 1$ and $\{\delta_n\}$ is decreasing in n .

(2) Let $p, q \in (1, \infty)$,

$$f_1^{(x_0)}(x) = \hat{\nu}(0, x \wedge x_0), \quad f_{n+1}^{(x_0)} = f_n^{(x_0)} \Pi(f_n^{(x_0)}), \quad n \geq 1 \quad (27)$$

and define

$$\tilde{\delta}_n = \sup_{x_0 \in (0, D]} \|f_n^{(x_0)} \Pi(f_n^{(x_0)})\|_{\mu, q}^{1-q/p} \left(\inf_{x \in (0, D)} \Pi(f_n^{(x_0)})(x) \right)^{(q-1)/p}, \quad (28)$$

$$\bar{\delta}_n = \sup_{x_0 \in (0, D]} \frac{\|f_n^{(x_0)}\|_{\mu, q}}{\|f_n^{(x_0)}\|_{v, p}}, \quad n \geq 1. \quad (29)$$

Then we have $A \geq \tilde{\delta}_n \vee \bar{\delta}_n$ for all $n \geq 1$.

Actually, in view of Corollary 2.3 below, we have $\delta_1 < \infty$ iff $B < \infty$. When $q = p$, it is known from [8] that $\{\tilde{\delta}_n\}_{n \geq 1}$ is increasing in n and $\tilde{\delta}_1 \geq B$.

We can now summarize the first step of our approximating procedures as follows.

Corollary 2.3 For $q \geq p > 1$, we have

$$B \leq \bar{\delta}_1 \vee \tilde{\delta}_1 \leq A \leq \delta_1 \wedge (k_{q,p} B) \leq \delta_1 \leq \tilde{k}_{q,p} B, \quad (30)$$

where

$$\tilde{k}_{q,p} = \left(1 + \frac{q}{p^*}\right)^{1/q} \left(1 + \frac{p^*}{q}\right)^{1/p^*} \quad (\geq k_{q,p} \text{ if } q \geq p).$$

More precisely, let $\varphi(x) = \hat{\nu}(0, x)$. Then we have

$$\delta_1 = \left\{ \sup_{x \in (0, D)} \frac{1}{\varphi(x)^{\gamma^*}} \int_0^x dy \hat{\nu}(y) \left(\int_y^D \varphi^{q\gamma^*/p^*} d\mu \right)^{p^*/q} \right\}^{1/p^*} \leq \tilde{k}_{q,p} B, \quad (31)$$

where $\gamma^* = \frac{q}{p^*+q}$. Next, we have

$$\bar{\delta}_1 = \left\{ \sup_{x \in (0, D)} \left[\frac{1}{\varphi(x)^{q/p}} \int_0^x \varphi^q d\mu + \varphi(x)^{q/p^*} \mu(x, D) \right] \right\}^{1/q} \geq B, \quad (32)$$

$$\tilde{\delta}_1 = \sup_{x_0 \in (0, D)} \|f_2^{(x_0)}\|_{\mu, q}^{1-q/p} \left[\frac{f_2^{(x_0)}(x_0)}{\varphi(x_0)} \right]^{(q-1)/p}, \quad (33)$$

where

$$f_2^{(x_0)}(x) = \int_0^{x \wedge x_0} dy \hat{\nu}(y) \left[\int_y^{x_0} \varphi^{q-1} d\mu + \varphi(x_0)^{q-1} \mu(x_0, D) \right]^{p^*-1}, \quad x \in [0, D].$$

It is known that $\tilde{k}_{p,p} = \lim_{q \downarrow p} k_{q,p}$. When $q > p$, we have $\tilde{k}_{q,p} > k_{q,p}$. Their small differences are shown by Figure 5. Their ratios have similar shape as in Figure 5 and are located in the interval $[1, 1.23)$ with maximum 1.2274 at $(p, q) \approx (2, 2 + 2.5758)$. Besides,

$$\sup_{q \geq p > 1} \tilde{k}_{q,p} = \sup_{q \geq p > 1} k_{q,p} = \sup_{p > 1} \tilde{k}_{p,p} = \tilde{k}_{2,2} = 2.$$

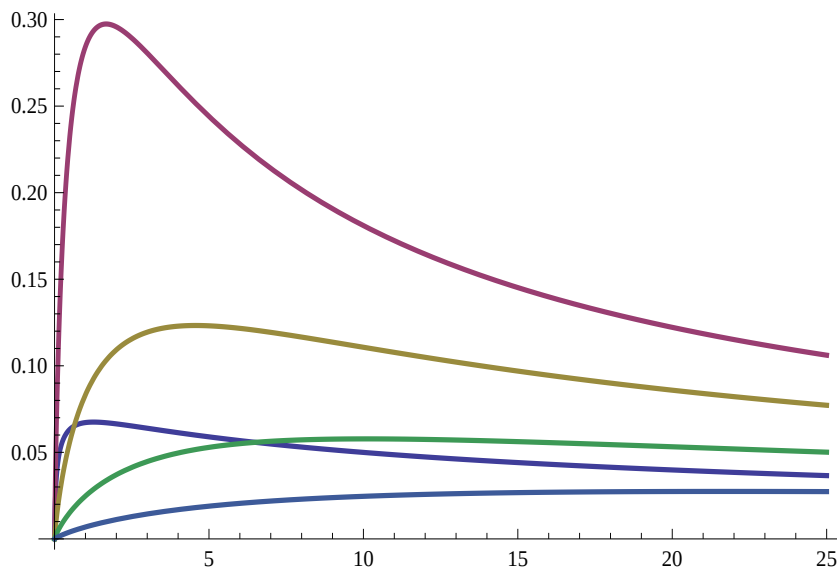


Figure 5 The difference $\tilde{k}_{q,p} - k_{q,p}$ for $q = p + x$ ($p = 1.1, 2, 5, 10, 20$) and x varies over $(0.0001, 25)$. When $p = 1.1$, the curve is special, located at lower level and intersects with two others. The remaining curves from top to bottom correspond to $p = 2, 5, 10$, and 20 , respectively.

Thus, our upper bound $\tilde{k}_{q,p}B$ given in (30) is a little bigger than the basic one (22). As illustrated by Figures 1–4, δ_1 improves $k_{q,p}B$ (not only $\tilde{k}_{q,p}B$) remarkably. However, the proof for the sharp factor $k_{q,p}$ (when $q > p$) becomes much more technical (cf. [10; pages 45–47] for historical remarks and [3]. See also Example 2.5 below). Therefore, we prove only the upper bound given in (31), as ones often do so [15; Theorem 1.14]. Actually, one often regards (22) replacing $k_{q,p}$ by $\tilde{k}_{q,p}$ as “basic estimates”, due to the reasons just mentioned above.

Among δ_1 , $\bar{\delta}_1$, and $\tilde{\delta}_1$ in the corollary, the most complicated one is $\tilde{\delta}_1$. It is not simple even for the simplest Example 1.1:

$$\tilde{\delta}_1 = \sup_{z \in (0,1]} \left(\frac{f_2^{(z)}(z)}{z} \right)^{(q-1)/p} \|f_2^{(z)}\|_{\mu,q}^{1-q/p},$$

$$f_2^{(z)}(x) = \int_0^{x \wedge z} dy \left[\frac{1}{q} (z^q - y^q) + z^{q-1}(1-z) \right]^{p^*-1}, \quad x \in [0, 1].$$

The main contribution of the sequence $\{\tilde{\delta}_n\}$ is, when $q = p$, its increasing property which then implies that $\{\tilde{\delta}_n\}$ is closer and closer, step by step, to A . Therefore, the sequence $\{\bar{\delta}_n\}$ possesses the same property since $\bar{\delta}_{n+1} \geq \tilde{\delta}_n$ by [8]. However, there is no direct proof for the increasing property of the sequence $\{\bar{\delta}_n\}$ even though it is believed to be true. From [8], it is also known that in the particular case of $q = p$, we have $\bar{\delta}_1 \geq \tilde{\delta}_1$ if $p \geq 2$, $\bar{\delta}_1 \leq \tilde{\delta}_1$ if $p \in (1, 2]$, and $\bar{\delta}_1 = \tilde{\delta}_1$ if $q = p = 2$. Thus, only in a small region of (p, q) , $\tilde{\delta}_1$ can be better

than $\bar{\delta}_1$. For instance, setting $p = 1.1$ in our Example 1.1, then only for those $q \in [1.1, 1.55]$, one has $\bar{\delta}_1 \leq \tilde{\delta}_1$. Next, let $p = 2$, then we have $\bar{\delta}_1 > \tilde{\delta}_1$ once $q > p$. For this reason, unlike the case of $q = p$, here we do not pay much attention to study the sequence $\{\tilde{\delta}_n\}$ in the case of $q > p$.

Having Corollary 2.3 at hand, it is not difficult to compute δ_1 and $\bar{\delta}_1$ given in Example 1.1. To obtain the constant A^* there, we need more work.

Remark 2.4 We are now going to describe the upper estimate (19) in a different way. First, when $q = p$, we can rewrite Π^* as $\Pi_{\mu,p}^{\nu,p}$:

$$\Pi_{\mu,p}^{\nu,p}(f)(x) = \frac{1}{f(x)} \int_0^x dy \hat{v}(y) \left(\int_y^D f^{p-1} d\mu \right)^{p^*-1}.$$

At the same time, we rewrite A in (1) as $A_{\mu,p}^{\nu,p}$. In this case, in view of the first inequality of (19), we have obtained

$$A_{\mu,p}^{\nu,p} \leq \inf_{f \in \mathcal{F}_\Pi} \left[\sup_{x \in (x,D)} \Pi_{\mu,p}^{\nu,p}(f)(x) \right]^{1/p^*}.$$

Actually, by [8; Theorem 2.1], the equality sign here holds:

$$A_{\mu,p}^{\nu,p} = \inf_{f \in \mathcal{F}_\Pi} \left[\sup_{x \in (x,D)} \Pi_{\mu,p}^{\nu,p}(f)(x) \right]^{1/p^*}. \quad (34)$$

Next, for general p and q , we may use the similar notation $\Pi_{\mu,q}^{\nu,p}$ and $A_{\mu,q}^{\nu,p}$. When $q \geq p$, noting that corresponding to $\tilde{p} = q/p^* + 1$ and $\tilde{v} = v^{q/p}$, we have

$$\hat{v} = \tilde{v}^{-\frac{1}{\tilde{p}-1}} = \left(v^{\frac{q}{p}}\right)^{-\frac{1}{q/p^*+1-1}} = v^{-\frac{1}{p-1}} = \hat{v},$$

It follows that $\Pi^* = \Pi_{\mu,\tilde{p}}^{\nu^{q/p},\tilde{p}}$, here $\nu^{q/p}$ denotes for a moment the measure when the density of ν is replaced by its power of q/p . By using the first inequality of (19) again, we have

$$\begin{aligned} A_{\mu,q}^{\nu,p} &\leq \inf_{f \in \mathcal{F}_\Pi} \left[\sup_{x \in (x,D)} \Pi_{\mu,\tilde{p}}^{\nu^{q/p},\tilde{p}}(f)(x) \right]^{1/p^*} \\ &= \left\{ \inf_{f \in \mathcal{F}_\Pi} \left[\sup_{x \in (x,D)} \Pi_{\mu,\tilde{p}}^{\nu^{q/p},\tilde{p}}(f)(x) \right]^{q/(q+p^*)} \right\}^{1/p^*+1/q} \end{aligned}$$

since the conjugate of $\tilde{p} = 1 + q/p^*$ is $1 + p^*/q$. By (34), we have thus obtain the following estimate

$$A_{\mu,q}^{\nu,p} \leq \left[A_{\mu,\tilde{p}}^{\nu^{q/p},\tilde{p}} \right]^{1/p^*+1/q}. \quad (35)$$

In other words, when $q \neq p$, we are estimating the optimal constant $A_{\mu,q}^{\nu,p}$ of a mapping $L^p(\nu) \rightarrow L^q(\mu)$ by the one of $L^{\tilde{p}}(\tilde{\nu}) \rightarrow L^{\tilde{p}}(\mu)$. When $q = p$, the

right-hand side of (35) coincides with its left-hand side and so (35) becomes an equality. In Example 1.1, the upper bound A^* denotes the right-hand side of (35). Note that without assuming (34), by part (1) of Theorem 2.1, the estimate (35) is the best one we can expected. This indicates a limitation of (19) since Figure 3 shows that there is a small difference between the two sides of (35) (see also Example 2.5 below). In contract to part (1) of Theorem 2.1, part (2) of the theorem can be sharp at least when there is a solution to the Euler-Lagrange equation (or “eigenequation”):

$$(vg'^{p-1})' + ug^{q-1} = 0, \quad g, g' > 0 \text{ on } (0, D).$$

We conclude this section by looking an extremal example to which there is no room for improving the upper estimate in (22). Refer to Lemma 4.4 and Proposition 4.5 in Appendix for more general results.

Example 2.5 Let $q > p > 1$, $D = \infty$, $\mu(dx) = x^{-q/p^*-1}dx$, and $\nu(dx) = dx$. Then the optimal constant in the Hardy-type inequality is

$$A = \left(\frac{p^*}{q}\right)^{1/q} \left[\frac{\Gamma\left(\frac{pq}{q-p}\right)}{\Gamma\left(\frac{q}{q-p}\right)\Gamma\left(\frac{p(q-1)}{q-p}\right)} \right]^{1/p-1/q} = \left(\frac{p^*}{q}\right)^{1/q} k_{q,p}$$

which can be attained by a simple optimizer f having derivative

$$f'(x) = \frac{\alpha}{(\beta x^\gamma + 1)^{(\gamma+1)/\gamma}}, \quad \alpha, \beta > 0, \gamma = \frac{q}{p} - 1.$$

Refer to [3] or Appendix for details. Since

$$B = \sup_{x>0} \hat{\nu}(0, x)^{1/p^*} \mu(x, \infty)^{1/q} = \sup_{x>0} x^{1/p^*} \left[\int_x^\infty y^{-q/p^*-1} \right]^{1/q} = \left(\frac{p^*}{q}\right)^{1/q},$$

the upper bound of the basic estimates in (22) is sharp. Actually, this is where the optimal factor $k_{q,p}$ comes from.

Even though there is now nothing more to do about the upper estimate of A , to understand what happened in such an extremal situation, we compute δ_n . Because $\varphi(x) = x$, $f_1 = \varphi^{\gamma^*}$, where $\gamma^* = \frac{q}{p^*+q}$, and

$$\begin{aligned} \int_y^D \varphi^{q\gamma^*/p^*} d\mu &= \int_y^\infty z^{q\gamma^*/p^*-q/p^*-1} = \frac{p^*}{q(1-\gamma^*)} y^{q(\gamma^*-1)/p^*}, \\ \int_0^x dy \left[\int_y^D \varphi^{q\gamma^*/p^*} d\mu \right]^{p^*/q} &= \left[\frac{p^*}{q(1-\gamma^*)} \right]^{p^*/q} \int_0^x y^{\gamma^*-1} dy = \frac{1}{\gamma^*} \left[\frac{p^*}{q(1-\gamma^*)} \right]^{p^*/q} x^{\gamma^*}. \end{aligned}$$

we have

$$f_2(x) = \int_0^x dy \left[\int_y^\infty f_1^{q/p^*} d\mu \right]^{p^*/q} = \frac{1}{\gamma^*} \left[\frac{p^*}{q(1-\gamma^*)} \right]^{p^*/q} f_1(x) =: C f_1(x).$$

By induction, it follows that $f_{n+1} = C^n f_1$ and hence

$$\delta_n = \left(\sup_x \frac{f_{n+1}(x)}{f_n(x)} \right)^{1/p^*} = C^{1/p^*} = \left(1 + \frac{p^*}{q} \right)^{1/p^* + 1/q}, \quad n \geq 1.$$

It is now easy to check that $\delta_n = \tilde{k}_{q,p} B (\geq k_{q,p} B)$ for all $n \geq 1$. Thus, no improvement of the upper bound $\tilde{k}_{q,p} B$ can be made by our approach. This is not surprising since δ_1 is already a sharp estimate of the right-hand side of (35). To see this, let $q \downarrow p$, we get

$$A_{\mu,q}^{\nu,p} \rightarrow A_{\mu,p}^{\nu,p} = \left(\frac{p^*}{p} \right)^{1/p} k_{p,p} = \left(\frac{p^*}{p} \right)^{1/p} \tilde{k}_{p,p} = p^*.$$

(Actually, when $q = p$, we come back to the original Hardy inequality, its optimal constant is well known to be p^* .) Then replacing p with $\tilde{p} = q/p^* + 1$, noting that $\tilde{p}^* = p^*/q + 1$, we obtain the optimal constant

$$A_{\mu,\tilde{p}}^{\nu q/p, \tilde{p}} = p^*/q + 1$$

appearing on the right-hand side of (35) which is clearly equal to δ_n . For general $q > p$, $\delta_n \equiv \tilde{k}_{q,p} B$ is actually bigger than, and so can not improve $k_{q,p} B$.

Next, we compute $\bar{\delta}_1$. Because

$$\begin{aligned} \frac{1}{\varphi(x)^{q/p}} \int_0^x \varphi^q d\mu + \varphi(x)^{q/p^*} \mu(x, D) &= \frac{1}{x^{q/p}} \int_0^x y^{q/p-1} + x^{q/p^*} \int_x^\infty y^{-q/p^*-1} \\ &= \frac{p}{q} + \frac{p^*}{q} x^{q/p}, \end{aligned}$$

we have by Corollary 2.3,

$$\bar{\delta}_1 = \left[\frac{pp^*}{q} \right]^{1/q}$$

which is clearly bigger than B : $\bar{\delta}_1/B = p^{1/q} > 1$, and hence improves the lower bound of the basic estimates in (22). Since $\bar{\delta}_1$ is not sharp, the lower bound can be usually improved step by step using the sequence $\{\bar{\delta}_n\}$. By Corollary 2.3, for general $q > p > 1$, the ratio $\delta_1/\bar{\delta}_1$ is controlled by $\tilde{k}_{q,p}$. However, from our experience we do have (without proof) that

$$\sup_{q>p>1} \delta_1/\bar{\delta}_1 \leq \sqrt{2} < 2 = \sup_{q>p>1} k_{q,p}.$$

In this sense, the ratio of the estimates in (22) is improved.

Some illustrations of $A (= k_{q,p} B)$ and its lower bound $\bar{\delta}_1$ are given in Figures 6 and 7. From which, one sees that our estimates are still effective even in such an extremal situation.

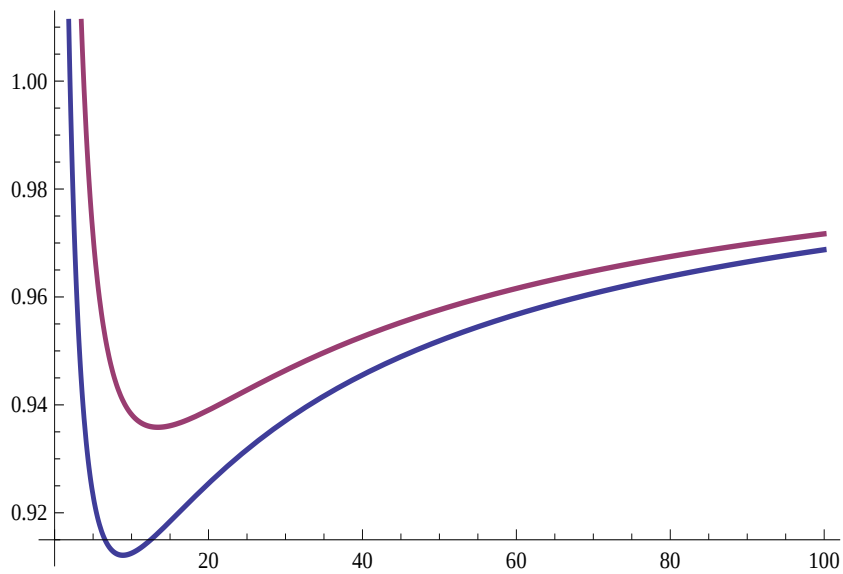


Figure 6 The constant $A = k_{q,p}B$ and its lower bound $\bar{\delta}_1$ in the case of $p = 2$, $q = p + r$. $r \in [0, 100]$.

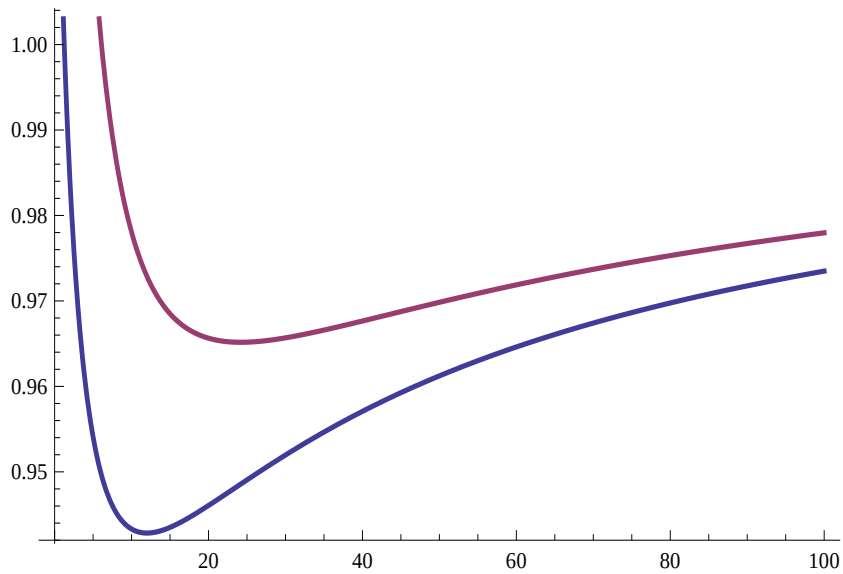


Figure 7 Everything is same as in the last figure except $p = 2$ is replaced by $p = 5$.

3 Proofs

It is now standard (cf. the explanation in the paragraph above [7; (9)]) that to prove the main results stated in the last section, one may assume that μ

has a density u if necessary. Similarly, one can assume that ν has a density v . Besides, one can also assume some integrability for $\hat{\nu}$ by an approximating procedure if necessary in the proofs below.

Proof of Theorem 2.1. (a) First, we prove (19). Let g satisfy $\|g\|_{\mu,q} = 1$ and $g(0) = 0$. Then for each positive h , by a good use of the Hölder inequality, we have

$$\begin{aligned} 1 &= \int_0^D g^q d\mu = \int_0^D \mu(dx) \left(\int_0^x g' \right)^q = \int_0^D \mu(dx) \left(\int_0^x g' v^{1/p} h^{-1/p} v^{-1/p} h \right)^q \\ &\leq \int_0^D \mu(dx) \left(\int_0^x g'^p v h^{-p} \right)^{q/p} \left(\int_0^x \hat{v} h^{p^*} \right)^{q/p^*} \quad (\text{since } p > 1). \end{aligned} \quad (36)$$

Here and in what follow, the Lebesgue measure dx is omitted. To separate out the term $\int_0^D v g'^p$, we need an exchange of the order of integration. When $q = p$, this is not a problem: one simply uses Fubini's theorem and nothing is lost. However, when $q \neq p$, this is not trivial. Fortunately, for $q \geq p$, one can apply the Hölder-Minkowski inequality:

$$\left\{ \int_{E_1} \mu(dx) \left[\int_{E_2} f(x,y) \nu(dy) \right]^r \right\}^{1/r} \leq \int_{E_2} \nu(dy) \left[\int_{E_1} f(x,y)^r \mu(dx) \right]^{1/r},$$

$\mu, \nu : \sigma\text{-finite measures}, r \in [1, \infty), f \geq 0.$

Applying this inequality to $r = q/p$, $E_1 = E_2 = [0, D]$, $\nu(dy) = (g'^p v h^{-p})(y) dy$, and

$$f(x,y) = \mathbb{1}_{[0,x]}(y) \left(\int_0^x \hat{v} h^{p^*} \right)^{p/p^*},$$

it follows that the right-hand of (36) is controlled by

$$\left\{ \int_0^D dy (g'^p v h^{-p})(y) \left[\int_y^D \mu(dx) \left(\int_0^x \hat{v} h^{p^*} \right)^{q/p^*} \right]^{p/q} \right\}^{q/p}.$$

Note that here we have only " $\leq$ " rather than " $=$ ". Now, making a power $1/q$, we get

$$\begin{aligned} 1 &\leq \left\{ \int_0^D dy (g'^p v h^{-p})(y) \left[\int_y^D \mu(dx) \left(\int_0^x \hat{v} h^{p^*} \right)^{q/p^*} \right]^{p/q} \right\}^{1/p} \\ &\leq \left(\int_0^D g'^p v \right)^{1/p} \left\{ \sup_{y \in (0,D)} \frac{1}{h(y)} \left[\int_y^D \mu(dx) \left(\int_0^x \hat{v} h^{p^*} \right)^{q/p^*} \right]^{1/q} \right\}. \end{aligned}$$

Replacing h by $h^{1/q}$, it follows that

$$1 \leq \left(\int_0^D g'^p v \right)^{1/p} \left\{ \sup_{y \in (0,D)} \frac{1}{h(y)} \int_y^D \mu(dx) \left(\int_0^x \hat{v} h^{p^*/q} \right)^{q/p^*} \right\}^{1/q}. \quad (37)$$

To move further, we need an extension of the mean value theorem for integrals.

Lemma 3.1 Let $g > 0$ on (α, β) and $\int_{\alpha}^{\beta} g d\mu < \infty$. Suppose that the integral $\int_{\alpha}^{\beta} f d\mu$ exists (may be $+\infty$). Then

$$\sup_{x \in (\alpha, \beta)} \frac{\int_x^{\beta} f d\mu}{\int_x^{\beta} g d\mu} \leq \sup_{x \in (\alpha, \beta)} \frac{f}{g}(x) \quad \text{and dually} \quad \inf_{x \in (\alpha, \beta)} \frac{\int_x^{\beta} f d\mu}{\int_x^{\beta} g d\mu} \geq \inf_{x \in (\alpha, \beta)} \frac{f}{g}(x).$$

Proof. Set $\xi = \sup_{x \in (\alpha, \beta)} (f/g)(x)$. Without loss of generality, assume that $\xi < \infty$. Otherwise, the first assertion is trivial. By assumptions, $g > 0$ and moreover

$$f \leq \xi g \quad \text{on} \quad (\alpha, \beta).$$

Making integration over the interval (x, β) , it follows that

$$\int_x^{\beta} f d\mu \leq \xi \int_x^{\beta} g d\mu, \quad x \in (\alpha, \beta).$$

The first assertion then follows since $\int_{\alpha}^{\beta} g d\mu \in (0, \infty)$. Dually, we can prove the second assertion. $\square$

We now come back to the proof of the inequality in (19). Actually, we prove a (formally) stronger conclusion. Let

$$\mathcal{F}_H^* = \{f : f(0) \geq 0, f > 0 \text{ on } (0, D)\} \quad [\supset \mathcal{F}_H].$$

For a given $f \in \mathcal{F}_H^*$, without loss of generality, assume that $\sup_{x \in (0, D)} H^*(f)(x) < \infty$. Otherwise, the upper bound we are going to prove is trivial. Let $h(x) = \int_x^D f^{q/p^*} d\mu$. As an application of Lemma 3.1, since $h < \infty$, we have

$$\begin{aligned} & \sup_{x \in (0, D)} \frac{1}{h(x)} \int_x^D \mu(dy) \left(\int_0^y \hat{v} h^{p^*/q} \right)^{q/p^*} \\ & \leq \left\{ \sup_{x \in (0, D)} \frac{1}{f(x)} \int_0^x dy \hat{v}(y) \left[\int_y^D f^{q/p^*} d\mu \right]^{p^*/q} \right\}^{q/p^*}. \end{aligned}$$

Inserting this into (37) and making supremum with respect to g , it follows that

$$A \leq \left[\sup_{x \in (0, D)} H^*(f)(x) \right]^{1/p^*}$$

and then

$$A \leq \inf_{f \in \mathcal{F}_H^*} \left[\sup_{x \in (0, D)} H^*(f)(x) \right]^{1/p^*} \leq \inf_{f \in \mathcal{F}_H} \left[\sup_{x \in (0, D)} H^*(f)(x) \right]^{1/p^*}.$$

This gives us the first inequality in (19). Furthermore, applying Lemma 3.1 again, we obtain

$$A \leq \inf_{f \in \mathcal{F}_H} \left[\sup_{x \in (0, D)} H^*(f)(x) \right]^{1/p^*} \leq \inf_{f \in \mathcal{F}_I} \left[\sup_{x \in (0, D)} I^*(f)(x) \right]^{1/p^*}.$$

Now, for a given $f \in \mathcal{F}_H$ with $\sup_x H^*(f) < \infty$, let $g = fH^*(f)$. Then $g \in \mathcal{F}_I$ and

$$g'(x) = \hat{v}(x) \left[\int_x^D f^{q/p^*} d\mu \right]^{p^*/q} \geq \hat{v}(x) \left[\int_x^D g^{q/p^*} d\mu \right]^{p^*/q} \left[\inf_x H^*(f)^{-1} \right].$$

That is,

$$\sup_x H^*(f)(x) \geq \frac{\hat{v}(x)}{g'(x)} \left[\int_x^D g^{q/p^*} d\mu \right]^{p^*/q},$$

and then

$$\sup_x H^*(f)(x) \geq \sup_x I^*(g)(x).$$

On both sides, making successively, power $1/p^*$, infimum with respect to $g \in \mathcal{F}_I$, and then infimum with respect to $f \in \mathcal{F}_H$, we obtain

$$\inf_{f \in \mathcal{F}_H} \left[\sup_x H^*(f)(x) \right]^{1/p^*} \geq \inf_{g \in \mathcal{F}_I} \left[\sup_x I^*(g)(x) \right]^{1/p^*}.$$

Therefore, the equality in (19) holds.

(b) Next, we prove (20). Given $f \in \widetilde{\mathcal{F}}_H$, define $g_0 = [fH(f)](\cdot \wedge x_0)$. Then

$$\int_0^D v g_0'^p = \int_0^D v g_0'^{p-1} dg_0 = (v g_0 g_0'^{p-1})(x_0-) - \int_0^{x_0} g_0 (v g_0'^{p-1})'.$$

By definition of g_0 , we have

$$(v g_0 g_0'^{p-1})(x_0-) = g_0(x_0) \int_{x_0}^D f^{q-1} d\mu, \quad (v g_0'^{p-1})' = -f^{q-1} d\mu.$$

Hence we have

$$\int_0^D v g_0'^p = \int_0^D g_0 f^{q-1} d\mu \leq \left(\sup_{x \in (0, D)} \frac{f}{g_0} \right)^{q-1} \int_0^D g_0^q d\mu.$$

That is,

$$\|g'\|_{\nu, p} \leq \left(\sup_{x \in (0, D)} \frac{f}{g_0} \right)^{(q-1)/p} \|g_0\|_{\mu, q}^{q/p}.$$

In other words,

$$\frac{\|g_0'\|_{\nu, p}}{\|g_0\|_{\mu, q}} \leq \left(\sup_x \frac{f}{g_0} \right)^{(q-1)/p} \|g_0\|_{\mu, q}^{q/p-1}.$$

We have thus obtain

$$A \geq \sup_{f \in \mathcal{F}_H} \|fH(f)\|_{\mu, q}^{1-q/p} \left(\inf_{x \in (0, D)} H(f)(x) \right)^{(q-1)/p}.$$

This proves the first assertion of part (2). Then the second one follows by using the proof similar to the last part of proof (a). $\square$

Proof of Theorem 2.2. The approximating sequences $\{\delta_n\}$ and $\{\tilde{\delta}_n\}$ are simply successive application of Theorem 2.1. The sequence $\{\bar{\delta}_n\}$ is a direct application of (8). The monotonicity of δ_n in n is obtained by using Lemma 3.1 twice. $\square$

To prove the (basic) upper bound given in Corollary 2.3, we need the following result.

Lemma 3.2 Let $\varphi > 0$ on $(0, D)$ and

$$B := \sup_{x \in (0, D)} \varphi(x)^{1/p^*} \mu(x, D)^{1/q} < \infty.$$

Then for each $\gamma \in (0, 1)$, we have

$$\left(\int_x^D \varphi^{\gamma q/p^*} d\mu \right)^{1/q} \leq \frac{B}{(1-\gamma)^{1/q}} \varphi^{(\gamma-1)/p^*}.$$

Proof. For a function $h \in C[0, D] \cap C^1(0, D)$ with $h(0) = 0$, write

$$\int_x^D h d\mu = - \int_x^D h(y) dM(y), \quad M(y) := \mu(y, D).$$

Applying [4; Proof of Lemma 1.2] to $c = B^q$ with a change of φ by φ^{q/p^*} , it follows that

$$\int_x^D \varphi^{\gamma q/p^*} d\mu \leq \frac{B^q}{1-\gamma} \varphi^{q(\gamma-1)/p^*}.$$

The required assertion now follows immediately. $\square$

Proof of Corollary 2.3. The main assertion as well as the formula of δ_1 are obtained by Theorem 2.2 directly, except the estimates involving B and the formulas of $\bar{\delta}_1$ and $\tilde{\delta}_1$. The inequality involving B in the middle is based on (22).

(a) To prove the upper bound given in (31), we specify φ used in Lemma 3.2: $\varphi(x) = \hat{\nu}(0, x)$, and set $f = \varphi^\gamma$. Then

$$\frac{\hat{\nu}}{f'} = \frac{1}{\gamma \varphi^{\gamma-1}}.$$

By Lemma 3.2, we have

$$[I^*(f)(x)]^{1/p^*} \leq B \gamma^{-1/p^*} (1-\gamma)^{-1/q}.$$

Optimizing the right-hand side with respect to γ , the minimum

$$\left(1 + \frac{q}{p^*}\right)^{1/q} \left(1 + \frac{p^*}{q}\right)^{1/p^*} \quad [= \tilde{k}_{q,p}]$$

of $\gamma^{-1/p^*}(1-\gamma)^{-1/q}$ is attained at

$$\gamma^* = \frac{q}{p^* + q}.$$

We have finally arrived at $\delta_1 \leq \tilde{k}_{q,p}B$ by using the equality in (19) with the specific $f = \varphi^{\gamma^*}$. From the proof, the main reason why δ_1 can improve $\tilde{k}_{q,p}B$ is clear: δ_1 is defined by using the operator II^* , but its upper bound $\tilde{k}_{q,p}B$ is deduced from the operator I^* . Usually, there is a gap between $\sup_x II^*(f)$ and $\sup_x I^*(f)$ for a fixed f .

(b) To compute $\bar{\delta}_1$, recall our test function

$$\varphi(x) = \varphi^{(x_0)}(x) = \int_0^{x \wedge x_0} \hat{v}.$$

The reason to choose this function is the following observation:

$$\int_0^D v \varphi'^p = \int_0^{x_0} v^{p(1-p^*)} v = \int_0^{x_0} \hat{v} = \varphi(x_0).$$

Next, because of

$$\int_0^D \varphi^q d\mu = \int_0^{x_0} \varphi^q d\mu + \varphi(x_0)^q \mu(x_0, D),$$

it follows that

$$\frac{\|\varphi\|_{\mu,q}}{\|\varphi'\|_{v,p}} = \left[\frac{1}{\varphi(x_0)^{q/p}} \int_0^{x_0} \varphi^q d\mu + \varphi(x_0)^{q/p^*} \mu(x_0, D) \right]^{1/q}.$$

Making supremum with respect to x_0 , we obtain $\bar{\delta}_1$.

The proof of $\bar{\delta}_1 \geq B$ is rather easy. Simply ignore the first term in the sum in (32). The improvement of $\bar{\delta}_1$ from B is obvious.

(c) To compute $\bar{\delta}_1$, recall that

$$\begin{aligned} f_1^{(x_0)}(x) &= \hat{v}(0, x \wedge x_0), \\ f_2^{(x_0)}(x) &= \int_0^{x \wedge x_0} dy \hat{v}(y) \left[\int_y^{x_0} \varphi^{q-1} d\mu + \varphi(x_0)^{q-1} \mu(x_0, D) \right]^{p^*-1}, \quad x \in [0, D]. \end{aligned}$$

For simplicity, in what follows, we ignore the superscript (x_0) in $f_1^{(x_0)}$ and $f_2^{(x_0)}$. Clearly, we have

$$\inf_{x \in (0, D)} f_2(x)/f_1(x) = \inf_{x \in (0, x_0)} f_2(x)/f_1(x)$$

by the convention that $1/0 = \infty$. Next, we show that the derivative of f_2/f_1 is non-positive on $(0, x_0)$, that is

$$\varphi(x) \left[\int_x^D \varphi(\cdot \wedge x_0)^{q-1} d\mu \right]^{p^*-1} - \int_0^x dy \hat{v}(y) \left[\int_y^D \varphi(\cdot \wedge x_0)^{q-1} d\mu \right]^{p^*-1} \leq 0$$

on $(0, x_0)$. This is obvious since for each $h \geq 0$,

$$\int_0^x dy \hat{v}(y) \left[\int_y^D h d\mu \right]^{p^*-1} \geq \left(\int_0^x \hat{v} \right) \left[\int_x^D h d\mu \right]^{p^*-1} = \varphi(x) \left[\int_x^D h d\mu \right]^{p^*-1}.$$

Hence we indeed have

$$\inf_{x \in (0, D)} f_2(x)/f_1(x) = f_2(x_0)/f_1(x_0).$$

We have thus obtained $\tilde{\delta}_1$ as stated in the corollary. $\square$

4 Appendix. The inequalities on finite intervals and the sharp factor

As far as we know, the basic estimates (2) with universal optimal constant $k_{q,p}$ was proved only for the half-line (cf. [13]). In this appendix, we show that the estimates with the same factor $k_{q,p}$ actually hold for every finite interval. The study on this problem also provides us a chance to examine how to obtain (2). The main result of this section is Theorem 4.6. We begin with our study on three comparison results for the optimal constants and their basic upper estimates in different intervals. The first one is a comparison for the optimal constants only.

Lemma 4.1 Let A_D be the optimal constant in the Hardy-type inequality on the interval $(0, D)$. Then we have $A_D \uparrow A_{D'}$ as $D \uparrow D' \leq \infty$. Here we use the same notation (μ, ν) to denote the Borel measures on $[0, D']$ and their restriction to $[0, D]$. In particular, if the inequality holds on $(0, D')$, then it also holds with the same constant $A_{D'}$ on $(0, D)$ for every $D < D'$.

Proof. (a) Extending f from $[0, D]$ to $[0, D']$ by setting $f = f(\cdot \wedge D)$, it follows that

$$\left[\int_0^D |f|^q d\mu \right]^{1/q} \leq \left[\int_0^{D'} |f|^q d\mu \right]^{1/q} \leq A_{D'} \left[\int_0^{D'} |f'|^p d\nu \right]^{1/p} = A_{D'} \left[\int_0^D |f'|^p d\nu \right]^{1/p}.$$

The last assertion of the lemma is now obvious. We have thus proved the monotonicity: $A_D \leq A_{D'}$ whenever $D \leq D'$.

(b) To prove the convergence in the first assertion, consider first the simplest case that $\mu[0, D'] = \infty$. Then $D' = \infty$ since μ is Borel. Clearly, we have $B_{D'} = \infty$ and so is $A_{D'}$ by our basic estimates. Besides, restricting to $[0, n]$, we have

$$A_n \geq B_n = \sup_{x \in [0, n]} \mu[x, n]^{1/q} \hat{\nu}[0, x]^{1/p^*} \geq \mu[1, n]^{1/q} \hat{\nu}[0, 1]^{1/p^*} \rightarrow \infty \text{ as } n \rightarrow \infty,$$

hence the convergence in the first assertion holds in this case.

(c) Let $\mu[0, D'] < \infty$ and $A_{D'} < \infty$. Then for every f satisfying $\|f'\|_{\nu, p} \in (0, \infty)$, we have

$$\frac{\|f\mathbb{1}_{[0, D]}\|_{\mu, q}}{\|f'\mathbb{1}_{[0, D]}\|_{\nu, p}} \rightarrow \frac{\|f\|_{\mu, q}}{\|f'\|_{\nu, p}} \leq A_{D'} \quad \text{as } D \uparrow D'.$$

Since $A_{D'} < \infty$, for every $\varepsilon > 0$, we can choose first $f = f_\varepsilon$ such that $\|f'\|_{\nu, p} \in (0, \infty)$ and

$$A_{D'} \leq \frac{\|f\|_{\mu, q}}{\|f'\|_{\nu, p}} + \varepsilon,$$

then we can choose D closed to D' such that

$$\frac{\|f\|_{\mu, q}}{\|f'\|_{\nu, p}} \leq \frac{\|f\mathbb{1}_{[0, D]}\|_{\mu, q}}{\|f'\mathbb{1}_{[0, D]}\|_{\nu, p}} + \varepsilon.$$

Therefore, we obtain

$$A_D \leq A_{D'} \leq \frac{\|f\mathbb{1}_{[0, D]}\|_{\mu, q}}{\|f'\mathbb{1}_{[0, D]}\|_{\nu, p}} + 2\varepsilon \leq A_D + 2\varepsilon.$$

From this, we conclude that the convergence also holds in the present case.

(d) Finally, the proof in the case that $\mu[0, D'] < \infty$ but $A_{D'} = \infty$ is in parallel to the proof (c). $\square$

The next result is a comparison of the factor in the basic estimates for different intervals.

Lemma 4.2 Let $A_D(\mu, \nu)$ and $B_D(\mu, \nu)$ denote the constants A and B , respectively, given in the basic estimates (2) for the inequality on interval $[0, D]$ with measures μ and ν . Next, let $D < D' \leq \infty$ and (μ', ν') be an extension of (μ, ν) to $[0, D']$: $\mu'|_{[0, D]} = \mu$, $\nu'|_{[0, D]} = \nu$, and moreover $\mu'|_{(0, D')} = 0$.

- (1) Suppose that $A_{D'}(\mu', \nu') \leq k B_{D'}(\mu', \nu')$ for a universal constant k , then we have $A_D(\mu, \nu) \leq k B_D(\mu, \nu)$.
- (2) In particular, if the inequality in part (1) holds for arbitrary (resp. absolutely continuous) pair (μ', ν') , then so does the conclusion for arbitrary (resp. absolutely continuous) pair (μ, ν) .

Proof. Clearly, we need only to prove the first assertion. Then the second one follows immediately. As in the last proof, extend f from $[0, D]$ to $[0, D']$

by setting $f = f(\cdot \wedge D)$. Then we have

$$\begin{aligned}
 \left[\int_0^D |f|^q d\mu \right]^{1/q} &= \left[\int_0^{D'} |f|^q d\mu' \right]^{1/q} \quad (\text{since } \mu'|_{(D, D')} = 0) \\
 &\leq A_{D'}(\mu', \nu') \left[\int_0^{D'} |f'|^p d\nu' \right]^{1/p} \quad (\text{by definition of } A_{D'}(\mu', \nu')) \\
 &\leq kB_{D'}(\mu', \nu') \left[\int_0^{D'} |f'|^p d\nu' \right]^{1/p} \quad (\text{by assumption}) \\
 &= kB_{D'}(\mu', \nu') \left[\int_0^D |f'|^p d\nu \right]^{1/p} \quad (\text{since } f'|_{(D, D')} = 0).
 \end{aligned}$$

Because

$$\begin{aligned}
 B_{D'}(\mu', \nu') &= \sup_{x \in (0, D')} \hat{\nu}'(0, x)^{1/p^*} \mu'(x, D')^{1/q} \\
 &= \sup_{x \in (0, D)} \hat{\nu}'(0, x)^{1/p^*} \mu'(x, D)^{1/q} \quad (\text{since } \mu'|_{(D, D')} = 0) \\
 &= \sup_{x \in (0, D)} \hat{\nu}(0, x)^{1/p^*} \mu(x, D)^{1/q} \quad (\text{since } \mu'|_{[0, D]} = \mu \text{ and } \nu'|_{[0, D]} = \nu) \\
 &= B_D(\mu, \nu),
 \end{aligned}$$

it follows that

$$\left[\int_0^D |f|^q d\mu \right]^{1/q} \leq kB_D(\mu, \nu) \left[\int_0^D |f'|^p d\nu \right]^{1/p}.$$

Hence $A_D(\mu, \nu) \leq kB_D(\mu, \nu)$ as required. $\square$

The next result is somehow a refinement of Lemma 4.1, but in an opposite way: from local sub-intervals to the whole interval. It provides us an approximating procedure for unbounded interval.

Lemma 4.3 Given Borel measures μ^D and ν^D on $[0, D]$, extend them to $[0, D')$, $D < D' \leq \infty$, as follows:

$$\tilde{\mu}^D = \begin{cases} \mu^D & \text{on } [0, D], \\ 0 & \text{on } (D, D'); \end{cases}$$

$$\tilde{\nu}^{D, \#} = \begin{cases} \nu^D & \text{on } [0, D], \\ \# & \text{on } (D, D'), \end{cases}$$

where $\#$ is an arbitrary Borel measure. Then we have $A_D = A(\tilde{\mu}^D, \tilde{\nu}^{D, \#})$ and $B_D = B(\tilde{\mu}^D, \tilde{\nu}^{D, \#})$.

Proof. Following the proof of Lemma 4.2, it is easy to check that $B_D = B(\tilde{\mu}^D, \tilde{\nu}^{D,\#})$. Next, applying the inequality

$$\|f\|_{L^q(\tilde{\mu}^D)} \leq A(\tilde{\mu}^D, \tilde{\nu}^{D,\#}) \|f'\|_{L^p(\nu^{D,\#})},$$

to $f^D = f(\cdot \wedge D)$, we obtain

$$\|f^D\|_{L^q(\mu^D)} \leq A(\tilde{\mu}^D, \tilde{\nu}^{D,\#}) \|(f^D)'\|_{L^p(\nu^D)}.$$

Because f is arbitrary and so is f^D , this implies that $A_D \leq A(\tilde{\mu}^D, \tilde{\nu}^{D,\#})$. Conversely, for every function f on $(0, D')$ with $f(0) = 0$, we have

$$\begin{aligned} \left[\int_0^{D'} |f|^q d\tilde{\mu}^D \right]^{1/q} &= \left[\int_0^D |f|^q d\mu^D \right]^{1/q} \\ &\leq A_D \left[\int_0^D |f'|^p d\nu^D \right]^{1/p} \quad (\text{by definition of } A_D) \\ &\leq A_D \left[\int_0^{D'} |f'|^p d\tilde{\nu}^{D,\#} \right]^{1/p} \quad (\text{since } D < D'). \end{aligned}$$

This implies that $A(\tilde{\mu}^D, \tilde{\nu}^{D,\#}) \leq A_D$ and then the equality holds. $\square$

Lemma 4.4 (Bliss, 1930) Let $q > p$ ($p, q \in (1, \infty)$), $\nu(dx) = dx$, and $\mu(dx) = x^{-q/p^*-1}dx$ on $[0, D]$. Then we have $A \leq k_{q,p} (p^*/q)^{1/q}$ with equality sign holds provided $D = \infty$.

Proof. The case that $D = \infty$ was proved in Bliss' original paper [3]. Then by Lemma 4.1, the conclusion also holds for finite D . $\square$

The next result is a generalization of Bliss's lemma. It says that the basic upper estimate in (2) is sharp for a large class of (μ, ν) .

Proposition 4.5 Let $q > p$ ($p, q \in (1, \infty)$), $\nu(dx) = v(x)dx$, and define $\hat{v}(x) = v(x)^{1/(1-p)}$. Then the Hardy-type inequality holds on $[0, D]$ with $\mu(dx) := u(x)dx$,

$$0 \leq u(x) \leq -B_1^q \frac{d}{dx} \left(\int_0^x \hat{v} \right)^{-q/p^*},$$

where $B_1 \in (0, \infty)$ is a constant. Moreover, its optimal constant A_D satisfies $A_D \leq k_{q,p} B_1$. In particular, when $D = \infty$,

$$\hat{v}(0, \infty) = \infty \quad \text{and} \quad \sup_{x \in (0, D)} \left[\int_0^x \hat{v} \right]^{1/p^*} \left[\int_x^\infty u \right]^{1/q} = B_1, \quad (38)$$

the upper bound is sharp with $B_1 = B$ defined by (23).

Proof. Throughout the proof, we restrict ourselves to the special case that

$$u(x) = -B_1^q \frac{d}{dx} \left(\int_0^x \hat{v} \right)^{-q/p^*} > 0.$$

The general case stated in the proposition then follows immediately. In this situation, the last assertion of the proposition is due to [13; Theorem 1]. Actually, the essential part of the proof is in the special case that $B_1 = 1$. The use of B_1 indicates an additional freedom for the choice of u , even in the present non-linear situation.

(a) By definition of u , we have

$$\int_x^D u = B_1^q \left[\left(\int_0^x \hat{v} \right)^{-q/p^*} - \left(\int_0^D \hat{v} \right)^{-q/p^*} \right] \leq B_1^q \left(\int_0^x \hat{v} \right)^{-q/p^*}.$$

Note that here the equality sign holds iff so does the first condition in (38). Hence

$$B = \sup_{x \in (0, D)} \left[\int_0^x \hat{v} \right]^{1/p^*} \left[\int_x^D u \right]^{1/q} \leq B_1 < \infty,$$

and $B_1 = B$ once the first condition in (38) holds. Then the second condition in (38) is automatic in the present special case.

(b) Define

$$s(x) = \int_0^x \hat{v}, \quad (39)$$

$$\varphi(s(x)) = f(x) \hat{v}(x)^{-1} \quad (\varphi = \varphi_f). \quad (40)$$

Since $s(x)$ is increasing in x , its inverse function s^{-1} is well-defined. Then the last equation can be rewritten as

$$\varphi(s) = f(s^{-1}(s)) \hat{v}(s^{-1}(s))^{-1}.$$

Because

$$f(x) dx = f(x) \hat{v}(x)^{-1} ds(x) = \varphi(s) ds, \quad (41)$$

we have by (41),

$$Hf(x) := \int_0^x f = \int_0^{s(x)} \varphi = H\varphi(s(x)).$$

Next, because of definition of u and s ,

$$\begin{aligned} u(x) &= \frac{q}{p^*} B_1^q \left(\int_0^x \hat{v} \right)^{-q/p^*-1} \hat{v}(x) = \frac{q}{p^*} B_1^q s(x)^{-q/p^*-1} \hat{v}(x), \\ u(x) dx &= \frac{q}{p^*} B_1^q s(x)^{-q/p^*-1} ds(x) \quad (\text{cf. (39)}), \end{aligned}$$

we obtain

$$\begin{aligned} \left[\int_0^D (Hf(x))^q u(x) dx \right]^{1/q} &= B_1 \left(\frac{q}{p^*} \right)^{1/q} \left[\int_0^{s(D)} (H\varphi(s))^q s^{-q/p^*-1} ds \right]^{1/q} \\ &\leq k_{q,p} B_1 \left[\int_0^{s(D)} \varphi(s)^p ds \right]^{1/p} \end{aligned}$$

by Bliss' lemma and Lemma 4.1. The equality sign holds once $s(D) = \infty$, i.e. (38) holds. Since by (40),

$$f(x)^p v(x) = \varphi(s(x))^p \hat{v}(x)^p v(x) = \varphi(s(x))^p \hat{v}(x),$$

and then by (39),

$$f(x)^p v(x) dx = \varphi(s(x))^p ds(x),$$

we have

$$\left[\int_0^{s(D)} \varphi(s)^p ds \right]^{1/p} = \left[\int_0^D f(x)^p v(x) dx \right]^{1/p}.$$

Therefore, we have proved that

$$\left[\int_0^D (Hf(x))^q u(x) dx \right]^{1/q} \leq k_{q,p} B_1 \left[\int_0^D f(x)^p v(x) dx \right]^{1/p}.$$

This leads to the conclusion that $A_D \leq k_{q,p} B_1$ as required. Again, the equality sign holds under (38). $\square$

We can now state the main result in this section. When $D = \infty$, it is just [2; Theorem 8]. If additionally (38) holds, then it is [13; Theorem 2].

Theorem 4.6 Let $q > p$ ($p, q \in (1, \infty)$) and $D \leq \infty$. Then the basic estimates in (2) hold for given μ and ν .

Proof. The lower estimate in (2) is shown in the proof of (30) (Corollary 2.3). Our main task is to prove the upper estimate in (2).

Without loss of generality, assume that $\mu(dx) = u(x)dx$ and $\nu(dx) = v(x)dx$ on $[0, D]$ (cf. [14]), and moreover $B < \infty$. Next, by part (2) of Lemma 4.2, it suffices to prove the case that $D = \infty$. Note that

$$\begin{aligned} \int_0^\infty (Hf(x))^q u(x) dx &= \int_0^\infty \left(\int_0^x d(Hf(t))^q \right) u(x) dx \\ &= \int_0^\infty \left(\int_t^\infty u(x) dx \right) d(Hf(t))^q \quad (\text{by Fubini's theorem}) \\ &\leq B^q \int_0^\infty \left(\int_0^t \hat{v}(x) dx \right)^{-q/p^*} d(Hf(t))^q \\ &\quad (\text{by definition of } B) \end{aligned}$$

$$= B^q \int_0^\infty s(t)^{-q/p^*} d(Hf(t))^q \quad (\text{by (39)}). \quad (42)$$

Next, note that

$$\begin{aligned} \int_t^\infty s(x)^{-q/p^*-1} ds(x) &= -\frac{p^*}{q} s(x)^{-q/p^*} \Big|_{x=t}^\infty \\ &= -\frac{p^*}{q} s(\infty)^{-q/p^*} + \frac{p^*}{q} s(t)^{-q/p^*} \\ &= \frac{p^*}{q} s(t)^{-q/p^*} \quad \text{if (38) holds.} \end{aligned}$$

That is,

$$s(t)^{-q/p^*} = \frac{q}{p^*} \int_t^\infty s(x)^{-q/p^*-1} ds(x) \quad \text{if (38) holds.}$$

Combining this with (42), under (38), we obtain

$$\begin{aligned} \int_0^\infty (Hf(x))^q u(x) dx &\leq \frac{q}{p^*} B^q \int_0^\infty \left[\int_t^\infty s(x)^{-q/p^*-1} ds(x) \right] d(Hf(t))^q \\ &= \frac{q}{p^*} B^q \int_0^\infty \left[\int_0^x d(Hf(t))^q \right] s(x)^{-q/p^*-1} ds(x) \\ &= \frac{q}{p^*} B^q \int_0^\infty (Hf(x))^q s(x)^{-q/p^*-1} ds(x) \\ &= \frac{q}{p^*} B^q \int_0^\infty (H\varphi(s))^q s^{-q/p^*-1} ds \quad (\text{by (41) and (38)}). \end{aligned}$$

Therefore,

$$\left[\int_0^\infty (Hf(x))^q u(x) dx \right]^{1/q} \leq \left[\frac{q}{p^*} \right]^{1/q} B \left[\int_0^\infty (H\varphi(s))^q s^{-q/p^*-1} ds \right]^{1/q}.$$

By Bliss's Lemma, the right-hand side is controlled by

$$k_{q,p} B \left[\int_0^\infty \varphi(s)^p ds \right]^{1/p} = k_{q,p} B \left[\int_0^\infty f(x)^p v(x) dx \right]^{1/p}.$$

We have thus proved the required assertion under (38).

To remove condition (38) used in the proof above, we use Lemma 4.3. For given μ and ν , we define naturally μ^N and ν^N to be the restriction of μ and ν on $[0, N]$. Then we clearly have $\tilde{\nu}^{N,\nu} = \nu$, respectively. We have already proved that

$$A(\tilde{\mu}^N, \tilde{\nu}^{N,dx}) \leq k_{q,p} B(\tilde{\mu}^N, \tilde{\nu}^{N,dx})$$

since $\tilde{\nu}^{N,dx}$ satisfies condition (38). By Lemma 4.3, we get

$$A(\tilde{\mu}^N, \nu) = A_N = A(\tilde{\mu}^N, \tilde{\nu}^{N,dx}), \quad B(\tilde{\mu}^N, \nu) = B_N = B(\tilde{\mu}^N, \tilde{\nu}^{N,dx}),$$

and then

$$A(\tilde{\mu}^N, \nu) \leq k_{q,p} B(\tilde{\mu}^N, \nu).$$

The assertion now follows by letting $N \rightarrow \infty$. $\square$

We conclude the Appendix by a discussion on the eigenequation corresponding to the Hardy-type inequality,

Proposition 4.7 Again, let $\mu(dx) = u(x)dx$ and $\nu(dx) = v(x)dx$. When $q \neq p$, the eigenequation for the Hardy-type inequality becomes

$$(vg'^{p-1})' = -ug^{q-1}, \quad g, g' > 0 \quad \text{a.e. on } (-M, N)$$

and with boundary condition $(vg'^{p-1}g)|_{-M}^N = 0$ once $M, N < \infty$. Actually, the eigenequation is equivalent to the following assertion:

$$\frac{ug^{q-1}(x)}{(vg'^{p-1})'(x)} =: -\eta \quad \text{is independent of a.e. } x \text{ on } (-M, N).$$

If the boundary condition holds, then the optimal constant A is given by

$$A = \frac{\|g\|_{L^q(\mu)}}{\|g'\|_{L^p(\nu)}} = \eta^{1/q} \left[\int_{-M}^N vg'^p \right]^{1/q-1/p}.$$

Proof. The first assertion comes from the Euler-Lagrange equation in variational methods. Roughly speaking, the idea goes as follows. Let g be the minimizer of the inequality and let $h \in \mathcal{C}_0^\infty(-M, N)$. Define

$$F(\varepsilon) = \left(\int_{-M}^N (g + \varepsilon h)^p v \right)^{1/p} \left(\int_{-M}^N (g + \varepsilon h)^q u \right)^{-1/q}, \quad \varepsilon > 0.$$

Then, it is easy to check that $\frac{d}{d\varepsilon} F(\varepsilon) = 0$ iff

$$Q \int_{-M}^N g'^{p-1} h' v = P \int_{-M}^N g^{q-1} u h,$$

where

$$P = \int_{-M}^N g'^p v, \quad Q = \int_{-M}^N g^q u.$$

Using the integration by parts formula, we obtain

$$\int_{-M}^N \left[\frac{Q}{P} (g'^{p-1} v)' + g^{q-1} u \right] h = 0.$$

Since h is arbitrary, this gives us

$$\frac{Q}{P} (g'^{p-1} v)' + g^{q-1} u = 0, \quad \text{a.e.}$$

Here a key is the inhomogenous, one may replace g by ξg if necessary for some constant ξ , so that the coefficient Q/P can be set to be one. This gives us the first assertion and then one leads to the equivalent assertion.

Multiplying by g on both sides of the eigenequation, and using the integral by parts formula, it follows that

$$(vg'^{p-1}g)|_{-M}^N - \int_{-M}^N vg'^p = -\eta^{-1} \int_{-M}^N ug^q.$$

By boundary condition, we obtain

$$\int_{-M}^N ug^q = \eta \int_{-M}^N vg'^p.$$

Hence

$$\|g\|_{L^q(\mu)} = \eta^{1/q} \|g'\|_{L^p(\nu)}^{p/q}$$

which is the last assertion. $\square$

To apply the last result to Example 2.5, let

$$g(x) = \frac{\alpha x}{(1 + \beta x^\gamma)^{1/\gamma}} > 0, \quad \alpha, \beta > 0, \quad \gamma = \frac{q}{p} - 1 > 0.$$

Then

$$g'(x) = \frac{\alpha}{(\beta x^\gamma + 1)^{(\gamma+1)/\gamma}} > 0.$$

Clearly, we have $(gg'^{p-1})(0) = 0$. Next, since $g(x) \sim 1$ and $g'(x) \sim x^{-q/p}$ as $x \rightarrow \infty$, we also have $\lim_{x \rightarrow \infty} (gg'^{p-1})(x) = 0$. Some computations show that

$$\begin{aligned} (g'^{p-1})'(x) &= -\frac{\alpha^{p-1}\beta(p-1)(\gamma+1)x^{\gamma-1}}{(1+\beta x^\gamma)^{p+(p-1)/\gamma}} \\ &= -\frac{\alpha^{p-1}\beta(p-1)(\gamma+1)x^{\gamma-1}}{(1+\beta x^\gamma)^{(q-1)/\gamma}} \quad \left(p + \frac{p-1}{\gamma} = \frac{q-1}{\gamma}\right), \\ \frac{g(x)^{q-1}}{x^{\gamma p + p - \gamma}} &= \frac{\alpha^{q-1}x^{\gamma - p(1+\gamma) + q-1}}{(1+\beta x^\gamma)^{(q-1)/\gamma}} \\ &= \frac{\alpha^{q-1}x^{\gamma-1}}{(1+\beta x^\gamma)^{(q-1)/\gamma}} \quad (q = p(1+\gamma)). \end{aligned}$$

Hence

$$\frac{g(x)^{q-1}}{x^{\gamma p + p - \gamma}} \bigg/ (g'^{p-1})'(x) = -\frac{\alpha^{q-p}}{\beta(p-1)(\gamma+1)} =: -\eta.$$

[The right-hand side is independent of x for all β , for simplicity, one may simply set $\beta = 1$. Then one can also set $\alpha = 1$ in computing A . This observation

may simplify the computation blow.] Set $t = \beta x^\gamma$, then

$$\begin{aligned} dt &= \beta \gamma x^{\gamma-1} dx, & x &= \left(\frac{t}{\beta}\right)^{1/\gamma}. \\ g'(x)^p dx &= \frac{\alpha^p}{(1 + \beta x^\gamma)^{p(1+\gamma)/\gamma}} dx \\ &= \frac{\alpha^p x^{1-\gamma}}{\beta \gamma (1 + \beta x^\gamma)^{p(1+\gamma)/\gamma}} dt \\ &= \frac{\alpha^p}{\beta^{1/\gamma} \gamma} \frac{t^{1/\gamma-1}}{(1+t)^{p(1+\gamma)/\gamma}} dt. \end{aligned}$$

Because

$$B(x, y) = \int_0^\infty \frac{t^{x-1}}{(1+t)^{x+y}} dt,$$

we obtain

$$\int_0^\infty g'(x)^p dx = \frac{\alpha^p}{\beta^{1/\gamma} \gamma} \int_0^\infty \frac{t^{1/\gamma-1}}{(1+t)^{p(1+\gamma)/\gamma}} dt = \frac{p \alpha^p \beta^{p/(p-q)}}{q-p} B\left(\frac{p}{q-p}, \frac{p(q-1)}{q-p}\right).$$

Therefore, we arrive at

$$\begin{aligned} &\left[\int_0^\infty \frac{g(x)^q}{x^{\gamma p + p - \gamma}} dx \right]^{1/q} \bigg/ \left[\int_0^\infty g'(x)^p dx \right]^{1/p} \\ &= \left[\frac{\alpha^{q-p}}{\beta(p-1)(\gamma+1)} \right]^{1/q} \left[\frac{p \alpha^p \beta^{p/(p-q)}}{q-p} B\left(\frac{p}{q-p}, \frac{p(q-1)}{q-p}\right) \right]^{1/q-1/p} \\ &= \left[\frac{p^*}{q} \right]^{1/q} \left[\frac{q-p}{p B\left(\frac{p}{q-p}, \frac{q(p-1)}{q-p}\right)} \right]^{1/p-1/q} \end{aligned}$$

which is the exact upper bound given in Example 2.5 or Lemma 4.4.

Acknowledgments. The paper was started when the author visited Taiwan in October, 2012. The invitation and financial support from Inst. Math. Acad. Sin. (hosts: Chii-Ruey Hwang and Shuenn-Jyi Sheu) are acknowledged. The author thanks the very warm hospitality made by C.R. Hwang, S.J. Sheu, Tzuu-Shuh Chiang, Yun-Shyong Chow and their wives. The author also thanks the following professors for their warm hospitality and the financial support from their universities: Lung-Chi Chen at Fu Jen Catholic University, Hong-Kun Xu, Mong-Na Lo Huang, and Mei-Hui Guo at National Sun Yat-sen University, Yuh-Jia Lee at National University of Kaohsiung, Shuenn-Jyi Sheu at National Central University, Yuan-Chung Sheu at National Chiao Tung University, Tien-Chung Hu at National Tsing Hua University. In the past years, the author has obtained a lot of help from Y.S. Chow, as well as the librarians at the Institute of Mathematics for providing some old papers. Without their help, this paper would not exist.

Research is supported in part by the National Natural Science Foundation of China (No. 11131003), the “985” project from the Ministry of Education in China,

and the Project Funded by the Priority Academic Program Development of Jiangsu Higher Education Institutions.

The careful corrections to an earlier version of the paper by an unknown referee is also acknowledged.

References

- [1] Agapito, J.C.C. Paredes, L.I. Rey, R.M. and Sy, P.W. (2002). On the asymptotic behaviors of the positive solution of $\Delta_p u + |u|^{p-1}u = 0$, Taiwan J. Math. 6:4, 555–563.
- [2] Bennett, G. (1991). Some elementary inequalities III, Quart. J. Math. Oxford Ser. (2) 42, 149–174.
- [3] Bliss, G.A. (1930). An integral inequality, Journ. L.M.S. 5, 40–46.
- [4] Chen, M.F. (2000). Explicit bounds of the first eigenvalue, Sci. China (A) 43:10, 1051–1059.
- [5] Chen, M.F. (2004). From Markov Chains to Non-Equilibrium Particle Systems, World Scientific, Singapore, Second Edition.
- [6] Chen, M.F. (2005). Eigenvalues, Inequalities, and Ergodic Theory, Springer, London.
- [7] Chen, M.F. (2013). Bilateral Hardy-type inequalities, Acta Math. Sin. Eng. Ser. 29:1, 1–32.
- [8] Chen, M.F., Wang, L.D. and Zhang, Y.H. (2015). Mixed eigenvalues of p-Laplacian, Front. Math. China 10:2, 249–274.
- [9] Drábek, P. and Manásevich, R. (1999). On the closed solution to some nonhomogeneous eigenvalue problems with p-Laplacian, Differential Integral Equations 12, 773–788.
- [10] Kufner, A., Maligranda, L. and Persson, L.E. (2007). The Hardy Inequality: About its History and Some Related Results, Vydavatelský Servis, Pilsen.
- [11] Lang, J. and Edmunds, D. (2011). Eigenvalues, Embeddings and Generalised Trigonometric Functions, LNM 2016, Springer, Heidelberg.
- [12] Liao, Z.W. (2014). Discrete Hardy-type inequalities, to appear in Advanced Non-linear Studies.
- [13] Manakov, V.M. (1992). On the best constant in weighted inequalities for Riemann-Liouville integrals, Bull. London Math. Soc. 24, 442–448.
- [14] Muckenhoupt, B. (1972). Hardy's inequality with weights, Studia Math. XLIV, 31–38.
- [15] Opic, B. and Kufner, A. (1990). Hardy-type Inequalities, Longman, New York.
- [16] Talenti, G. (1976). Best constant in Sobolev inequality, Ann. Mat. Pura ed Appl. 110:1, 353–372.

School of Mathematical Sciences, Beijing Normal University, Laboratory of Mathematics and Complex Systems (Beijing Normal University), Ministry of Education, Beijing 100875, The People's Republic of China.

E-mail: mfchen@bnu.edu.cn

Home page: http://math.bnu.edu.cn/~chenmf/main_eng.htm

Progress on Hardy-type Inequalities

Mu-Fa Chen

(Beijing Normal University, Beijing 100875, China)

October 4, 2013

Abstract

This paper surveys some of our recent progress on Hardy-type inequalities which consist of a well-known topic in Harmonic Analysis. In the first section, we recall the original probabilistic motivation dealing with the stability speed in terms of the L^2 -theory. A crucial application of a result by Fukushima and Uemura (2003) is included. In the second section, the non-linear case (a general Hardy-type inequality) is handled with a direct and analytic proof. In the last section, it is illustrated that the basic estimates presented in the first two sections can still be improved considerably.

This paper mainly concerns with the following Hardy-type inequality

$$\left(\int_{-M}^N |f - \pi(f)|^q d\mu \right)^{1/q} \leq A \left(\int_{-M}^N |f'|^p d\nu \right)^{1/p}, \quad (1)$$

where $p, q \in (1, \infty)$, μ and ν are Borel measures on an interval $[M, N]$ ($M, N \leq \infty$). Here, we assume that $\mu[-M, N] < \infty$ and define a probability measure $\pi = (\mu[-M, N])^{-1}\mu$. Then $\pi(f) := \int f d\pi$. The functions f are assumed to be absolutely continuous on $(-M, N)$ and belong to $L^q(\mu)$. For simplicity, we may also write the inequality as

$$\|f - \pi(f)\|_{L^q(\mu)} \leq A \|f'\|_{L^p(\nu)}.$$

To save our notation, assume the constant A to be optimal. The linear case that $p = q = 2$ is discussed in the next section, where a result by Fukushima and Uemura [14] plays an important role. The general case is studied in Section 2. In the last section, we show the possibility for improving further the basic estimates of the optimal constant.

2000 *Mathematics Subject Classifications.* 26D10, 60J60, 34L15.

Key words and phrases. Hardy-type inequality, stability speed, optimal constant, basic estimate.

1 Linear case: $p = q = 2$.

Let us recall the original probabilistic problem. Throughout this section, we fix $p = q = 2$. Consider a second-order elliptic operator on $(-M, N)$:

$$L = a(x) \frac{d^2}{dx^2} + b(x) \frac{d}{dx}, \quad a(x) > 0 \text{ on } (-M, N).$$

Then the two measures used in inequality (1) are as follows

$$\mu(dx) = \frac{e^{C(x)}}{a(x)} dx, \quad \nu(dx) = e^{C(x)} dx, \quad C(x) := \int_{\theta}^x \frac{b}{a}, \quad (2)$$

here in the last integral and in what follows, the Lebesgue measure dx is omitted, $\theta \in (-M, N)$ is a reference point. Denote by $\{P_t\}_{t \geq 0}$ the (maximal) semigroup generated by L on $L^2(\mu)$. Here “maximal” means the Dirichlet form having the maximal domain which we learnt earlier from Fukushima [12] (cf. [1; §6.7]). We are interested in the stability speed, for instance, the L^2 -exponential convergence rate ε :

$$\|P_t f - \pi(f)\|_{L^2(\mu)} \leq \|f - \pi(f)\|_{L^2(\mu)} e^{-\varepsilon t}, \quad t \geq 0.$$

Then, it turns out that the largest rate $\varepsilon_{\max}$ coincides with A^{-1} given in (1) (cf. [1; Theorem 9.1]).

We can state one of our recent results as follows.

Theorem 1.1 [4; Theorem 10.2] Let $a > 0$, a and b be continuous on $[-M, N]$ (or $(-M, N]$ if $M = \infty$, for instance). Assume that $\mu(-M, N) < \infty$. Then for the optimal constant A , we have the basic estimates: $\kappa \leq A \leq 2\kappa$, where

$$\kappa^{-2} = \inf_{-M < x < y < N} [\mu(-M, x)^{-1} + \mu(y, N)^{-1}] \hat{\nu}(x, y)^{-1},$$

$$\mu(x, y) = \int_x^y d\mu, \text{ and } \hat{\nu}(dx) = e^{-C(x)} dx.$$

The continuity assumption on a and b is not essential and will be removed in the next section. To understand the proof of this theorem, assume that $M, N < \infty$. Then the general case can be done by an approximating procedure. The optimal constant actually describes an eigenvalue $\lambda_1 (= A^{-2})$ defined by

$$\lambda_1 = \inf \left\{ \|f'\|_{L^2(\nu)}^2 : \pi(f) = 0, \|f\|_{L^2(\mu)} = 1 \right\}.$$

Let g be the eigenfunction corresponding to λ_1 :

$$Lg = -\lambda_1 g, \quad g \neq 0.$$

To ignore a constant (say $\pi(g)$) from g , making derivative on both sides of the equation and replacing g' by f , we get

$$L_S f = -\lambda_1 f,$$

where

$$L_S = a(x) \frac{d^2}{dx^2} + (a'(x) + b(x)) \frac{d}{dx} + b'(x)$$

which is a Schrödinger operator. Because $g'(-M) = 0 = g'(N)$, the boundary condition for L_S becomes $f(-M) = 0 = f(N)$. This leads to the principal eigenvalue of L_S :

$$\lambda_S = \sup_{f \in \mathcal{F}} \inf_{x \in (-M, N)} \frac{-L_S f}{f}(x),$$

where

$$\mathcal{F} = \{f \in \mathcal{C}^1[-M, N] \cap \mathcal{C}^2(-M, N) : f(-M) = f(N) = 0, f|_{(-M, N)} > 0\}.$$

Note that the zero-point of g in the original eigenequation for L is located inside of the interval $(-M, N)$, not explicitly known; the zero-points for the eigenvalue λ_S are located at the boundaries $-M$ and N only. This is the advantage of L_S . However, there is an extra term b' in operator L_S which costs some trouble as usual. To avoid this, we rewrite L as

$$L = \frac{d}{d\mu} \frac{d}{d\hat{\nu}}.$$

Then we define a dual operator of L as follows.

$$L^* = \frac{d}{d\mu^*} \frac{d}{d\hat{\nu}^*} := \frac{d}{d\hat{\nu}} \frac{d}{d\mu},$$

i.e. an exchange of μ and $\hat{\nu}$. More explicitly,

$$L^* = a(x) \frac{d^2}{dx^2} + (a'(x) - b(x)) \frac{d}{dx}.$$

Next, define

$$\lambda_0^* = \sup_{f^* \in \mathcal{F}} \inf_{x \in (-M, N)} \frac{-L^* f^*}{f^*}(x).$$

In view of the next result which is crucial in proving Theorem 1.1, it is clear that we have thus removed the extra term b' in the operator L_S .

Proposition 1.2 We have $A^{-2} = \lambda_1 = \lambda_S = \lambda_0^*$.

Here, in proving $\lambda_1 = \lambda_S$, we have used a mathematical tool — the coupling technique (cf. [8] or [2]). We have also used another tool — dual technique in proving $\lambda_S = \lambda_0^*$. It is interesting that they are the main tools used in the study on interacting particle systems (cf. [1] and references therein). To obtain the basic estimates listed in Theorem 1.1, we need one more mathematical tool — the capacitary method. The next result is taken from Fukushima & Uemura [14] and [2, 3], see also [13].

Theorem 1.3 For a regular transient Dirichlet form $(D, \mathcal{D}(D))$ with locally compact state space $(E, \mathcal{E})$, the optimal constant $A_{\mathbb{B}}$ in the Poincaré-type inequality

$$\|f^2\|_{\mathbb{B}} \leq A_{\mathbb{B}}^2 D(f), \quad f \in \mathcal{C}_K^\infty(E)$$

satisfies $B_{\mathbb{B}} \leq A_{\mathbb{B}} \leq 2B_{\mathbb{B}}$, where $\|\cdot\|_{\mathbb{B}}$ is the norm in a normed linear space $\mathbb{B}$ and

$$B_{\mathbb{B}}^2 = \sup_{\text{compact } K} \text{Cap}(K)^{-1} \|\mathbb{1}_K\|_{\mathbb{B}}.$$

The space $\mathbb{B}$ can be very general, for instance $L^p(\mu)$ ($p \geq 1$) or the Orlicz spaces. In the present context, $D(f) = \int_{-M}^N f'^2 e^C = \|f'\|_{L^2(\nu)}^2$, $\mathcal{D}(D)$ is the closure of $\mathcal{C}_K^\infty(-M, N)$ with respect to the norm $\|\cdot\|_D$: $\|f\|_D^2 = \|f\|^2 + D(f)$, and

$$\text{Cap}(K) = \inf \{D(f) : f \in \mathcal{C}_K^\infty(-M, N), f|_K \geq 1\}.$$

Note that we have the universal factor 2 here and the isoperimetric constant $B_{\mathbb{B}}$ has a very compact form. We now need to compute the capacity only. The problem is that the capacity is usually not explicitly computable. For instance, at the moment, we do not know how to compute it for Schrödinger operators even for the elliptic operators having killings. very luckily we are able to compute the capacity for the one-dimensional elliptic operators. The result has a simple expression:

$$B_{\mathbb{B}}^2 = \sup_{-M < x < y < N} [\hat{\nu}(-M, x)^{-1} + \hat{\nu}(y, N)^{-1}]^{-1} \|\mathbb{1}_{(x, y)}\|_{\mathbb{B}}.$$

It looks strange to have double inverse here. So, making inverse in both sides, we get

$$B_{\mathbb{B}}^{-2} = \inf_{-M < x < y < N} [\hat{\nu}(-M, x)^{-1} + \hat{\nu}(y, N)^{-1}] \|\mathbb{1}_{(x, y)}\|_{\mathbb{B}}^{-1}.$$

Applying this result to $\mathbb{B} = L^1(\mu)$, we obtain the solution to the case having double Dirichlet boundaries: $\lambda_0 = A_{L^1(\mu)}^{-2}$ and

$$\kappa_0^{-2} = B_{L^1(\mu)}^{-2} = \inf_{-M < x < y < N} [\hat{\nu}(-M, x)^{-1} + \hat{\nu}(y, N)^{-1}] \mu(x, y)^{-1}.$$

Applying the last result to the dual process, we have not only

$$(\kappa_0^*)^2 \leq \lambda_1 = \lambda_S = \lambda_0^* \leq 4(\kappa_0^*)^2$$

but also

$$\begin{aligned} (\kappa_0^*)^{-2} &= \inf_{x < y} [\hat{\nu}^*(-M, x)^{-1} + \hat{\nu}^*(y, N)^{-1}] \mu^*(x, y)^{-1} \\ &= \inf_{x < y} [\mu(-M, x)^{-1} + \mu(y, N)^{-1}] \hat{\nu}(x, y)^{-1} \\ &= \kappa^{-2}. \end{aligned}$$

We have thus arrived at the assertion of Theorem 1.1. Refer to [4; §10] and [5] for more details.

To conclude this section, we remark that the use of the capacity is natural in the higher dimensions, since in which the boundary may be very complicated. However, it seems unnecessary to use it in the present one-dimensional situation. This leads to a direct proof of Theorem 1.1, given in the next section, without using the three mathematical tools just mentioned above.

2 Non-linear case

We now return to our general inequality (1). First, we need a measure $\hat{\nu}$, as in the last section, deduced from ν . Let $\nu^\#$ be the absolutely continuous part of ν with respect to the Lebesgue measure. Then, define

$$\hat{\nu}(dx) = \hat{\nu}_p(dx) = \left(\frac{d\nu^\#}{dx} \right)^{-1/(p-1)} dx, \quad p > 1.$$

Next, we need a universal factor

$$k_{q,p} = \left[\frac{q-p}{pB\left(\frac{p}{q-p}, \frac{p(q-1)}{q-p}\right)} \right]^{1/p-1/q} \leq 2 \quad \text{if } q \geq p,$$

where $B(\alpha, \beta)$ is the Beta function. In particular (as the limit of $q \downarrow p$),

$$k_{p,p} = p^{1/p} (p^*)^{1/p^*},$$

where p^* is the conjugate of $p \in (1, \infty)$: $1/p + 1/p^* = 1$.

Theorem 2.1 [6; Theorem 2.6] Let $\mu(-M, N) < \infty$. Then the optimal constant A in the Hardy-type inequality (1) satisfies

- (1) the upper estimate $A \leq k_{2,p} B^*$ for $1 < p \leq 2 \leq q < \infty$ once the pure point part of μ (denoted by μ_{pp}) vanishes, and
- (2) the lower estimate $A \geq B_*$ for $1 < p, q < \infty$, where

$$B^* = \sup_{x \leq y} \frac{\hat{\nu}(x, y)^{(p-1)/p}}{\left\{ \mu(-M, x)^{\frac{p}{q(1-p)}} + \mu(y, N)^{\frac{p}{q(1-p)}} \right\}^{(p-1)/p}},$$

$$B_* = \sup_{x \leq y} \frac{\hat{\nu}(x, y)^{(p-1)/p}}{\left\{ \mu(-M, x)^{\frac{1}{1-q}} + \mu(y, N)^{\frac{1}{1-q}} \right\}^{(q-1)/q}}.$$

Moreover, $B_* \leq B^* \leq 2^{1/p-1/q} B_*$ once $q \geq p$.

Here are some remarks on Theorem 2.1.

- (a) The isoperimetric constants B^* and B_* are expressed explicitly in measures μ and $\hat{\nu}$.
- (b) The boundaries $-M$ and N symmetric in the formulas of B^* and B_* .
- (c) Even though $B^* \geq B_*$ in general, but the rough ratio $k_{q,p}2^{1/p-1/q}$ of the upper and lower bounds is still ≤ 2 .
- (d) When $q = p$, we have

$$B^* = B_* = \sup_{x \leq y} \frac{\hat{\nu}(x, y)^{(p-1)/p}}{\left\{ \mu(-M, x)^{\frac{1}{1-p}} + \mu(y, N)^{\frac{1}{1-p}} \right\}^{(p-1)/p}}.$$

- (e) Ignoring the $\mu(-M, x)$ -term in the expression of B^* or B_* , we obtain

$$B^+ = \sup_y \hat{\nu}(-M, y]^{1/p^*} \mu(y, N)^{1/q}.$$

Similarly, Ignoring the $\mu(y, N)$ -term in the expression of B^* or B_* , we obtain

$$B^- = \sup_x \hat{\nu}(x, N)^{1/p^*} \mu(-M, x)^{1/q}.$$

We have thus returned to one of the main results in the study of Hardy-type inequalities.

Theorem 2.2 (1920—1992) For the Hardy-type inequalities

$$\|f\|_{L^q(\mu)} \leq A^+ \|f'\|_{L^p(\nu)}, \quad f(-M) = 0$$

and

$$\|f\|_{L^q(\mu)} \leq A^- \|f'\|_{L^p(\nu)}, \quad f(N) = 0,$$

we have the basic estimates $B^\pm \leq A^\pm \leq k_{q,p} B^\pm$. Moreover, the factor $k_{q,p}$ is sharp.

There is a long history about the development of Theorem 2.2. The reader is urged to refer to [15] and [6] for a long list of references including five books.

Having the experience in proving Theorem 1.1 and known the history of Theorem 2.2, it is hardly imaginable how to find a direct proof of Theorem 1.1, or even much more general Theorem 2.2, without using capacity. To have a test, let us introduce the proof of a hard part — the upper estimate of Theorem 2.2.

The idea is starting from Theorem 2.2. For this, we split the interval $(-M, N)$ into two parts: $(-M, \theta)$ and (θ, N) ,



Denote by A_θ^- the optimal constant on the left subinterval $(-M, \theta)$ and by A_θ^+ the one on the right subinterval (θ, N) with the same boundary condition $f(\theta) = 0$. Then, we can rewrite Theorem 2.2 as follows.

Known Theorem Let $q \geq p$. Then we have $B_\theta^\pm \leq A_\theta^\pm \leq k_{q,p} B_\theta^\pm$, where

$$B_\theta^- = \sup_{x < \theta} \hat{\nu}(x, \theta)^{1/p^*} \mu(-M, x)^{1/q}, \quad B_\theta^+ = \sup_{y > \theta} \hat{\nu}(\theta, y)^{1/p^*} \mu(y, N)^{1/q}.$$

Proof of the upper estimate: $k_{2,p} B^* \geq A$.

Rewrite B^* as

$$\begin{aligned} B^* &= \sup_{x \leq y} \frac{\hat{\nu}(x, y)^{(p-1)/p}}{\left\{ \mu(-M, x)^{\frac{p}{q(1-p)}} + \mu(y, N)^{\frac{p}{q(1-p)}} \right\}^{(p-1)/p}} \\ &= \left\{ \sup_{x \leq y} \frac{\hat{\nu}(x, y)}{\mu(-M, x)^{\frac{p}{q(1-p)}} + \mu(y, N)^{\frac{p}{q(1-p)}}} \right\}^{(p-1)/p} \\ &=: \left\{ \sup_{x \leq y} \frac{\hat{\nu}(x, y)}{\varphi(x) + \psi(y)} \right\}^{1/p^*}. \end{aligned}$$

By proportional property, we have

$$\frac{\hat{\nu}(x, y)}{\varphi(x) + \psi(y)} = \frac{\hat{\nu}(x, \theta) + \hat{\nu}(\theta, y)}{\varphi(x) + \psi(y)} \geq \left\{ \frac{\hat{\nu}(x, \theta)}{\varphi(x)} \wedge \frac{\hat{\nu}(\theta, y)}{\psi(y)} \right\}, \quad \theta \in (x, y).$$

Here $a \wedge b = \min\{a, b\}$ and similarly $a \vee b = \max\{a, b\}$. Hence (omit what in $\{\dots\}$)

$$\frac{\hat{\nu}(x, y)}{\varphi(x) + \psi(y)} \geq \sup_{\theta \in (x, y)} \{\dots\}.$$

Then

$$\begin{aligned} \sup_{x \leq y} \frac{\hat{\nu}(x, y)}{\varphi(x) + \psi(y)} &\geq \sup_{x \leq y} \sup_{\theta \in (x, y)} \{\dots\} \\ &= \sup_{\theta} \sup_{(x, y) \ni \theta} \{\dots\} \\ &= \sup_{\theta} \left\{ \left[\sup_{x \leq \theta} \frac{\hat{\nu}(x, \theta)}{\varphi(x)} \right] \wedge \left[\sup_{y \geq \theta} \frac{\hat{\nu}(\theta, y)}{\psi(y)} \right] \right\}. \end{aligned}$$

Making power $1/p^*$ on both sides, by definition of $B_\theta^\pm$, we obtain

$$B^* \geq \sup_{\theta} (B_\theta^- \wedge B_\theta^+).$$

Here a problem appears: we need $\vee$ rather than $\wedge$ on the right-hand side. To overcome this, we assume that $\mu_{pp} = 0$. Then, there exists $\bar{\theta} \in (-M, N)$ such that $B_{\bar{\theta}}^- = B_{\bar{\theta}}^+$. Therefore

$$B^* \geq \sup_{\theta} (B_\theta^- \wedge B_\theta^+) \geq B_{\bar{\theta}}^- = B_{\bar{\theta}}^- \vee B_{\bar{\theta}}^+$$

and then

$$\begin{aligned} k_{q,p}B^* &\geq (k_{q,p}B_{\bar{\theta}}^-) \vee (k_{q,p}B_{\bar{\theta}}^+). \\ &\geq A_{\bar{\theta}}^- \vee A_{\bar{\theta}}^+ \quad (\text{Known Theorem}) \\ &\geq \inf_{\bar{\theta}} (A_{\bar{\theta}}^- \vee A_{\bar{\theta}}^+) \\ &\geq A \quad (\text{splitting technique}). \end{aligned}$$

Here each step holds for all $q \geq p$ except the last one. In which, some additional work is required, due to the appearance of $f - \pi(f)$ rather than f only. We prove the conclusion first for $q = 2 \geq p$ and then extend it to $q \geq 2$, even to a large class of normed linear space $\mathbb{B}$, as used in Theorem 1.3, using a known lifting procedure (cf. [2; §6.3]). Note that in the proof above, we use a bridge θ to combine the known results on two subintervals together. But then remove it, otherwise, $\bar{\theta}$ for instance, may not be computable. Nevertheless, it should be understandable that the present analytic proof does not use the coupling, duality, or capacitary techniques.

Actually, much more topics are studied in [6]: the bilateral Dirichlet boundaries, logarithmic Sobolev inequalities, Nash inequalities, and so on.

3 Improvements of the basic estimates

Note that for two given numbers having smaller ratio, their difference can be still quite big. Hence, it is meaningful to improve the basic estimates introduced in the last two sections. In this section, we show the possibility in doing so by a simplest example: $\mu = dx$ and $\nu = dx$ on $(0, 1)$. We need to consider the following Hardy-type inequality

$$\|f\|_{L^q(\mu)} \leq A \|f'\|_{L^p(\nu)}, \quad f(0) = 0 \quad (3)$$

only since the other cases (the ergodic case in particular) can be reduced to this one by symmetry. The basic estimates for the optimal constant A in (3) are given in Theorem 2.2.

Example 3.1 Let $\mu = dx$ and $\nu = dx$ on $(0, 1)$. Then the optimal constant A in (3) is given as follows.

(1) When $p = q = 2$, we have $A = 2/\pi$.

(2) When $p = q \in (1, \infty)$, we have

$$A = \frac{p}{\pi(p-1)^{1/p}} \sin \frac{\pi}{p}.$$

(3) For general $p, q \in (1, \infty)$, we have

$$A = \frac{p^{\frac{1}{q}} q^{1-\frac{1}{p}} (pq + p - q)^{\frac{1}{p} - \frac{1}{q}}}{(p-1)^{\frac{1}{p}} B\left(\frac{1}{q}, 1 - \frac{1}{p}\right)}.$$

This simplest example already shows that it is nontrivial from the special case $p = q = 2$ to the general one.

Proposition 3.2 [7] For the optimal constant A in the last example, we have the following improved estimates:

$$B \leq \bar{\delta}_1 \leq A \leq A^* \leq \delta_1 \leq k_{q,p}B,$$

where

$$B = \frac{p^{1/q}((p-1)q)^{1-1/p}}{(pq+p-q)^{1-1/p+1/q}},$$

$$\bar{\delta}_1 = \frac{p^{1/q}((p-1)(q+1))^{1-1/p}}{(pq+p-q)^{1-1/p+1/q}},$$

$$A^* = \left[\frac{p^*}{q} \right]^{\frac{1}{q}} \left[\frac{p^*+q}{\pi p^*} \sin \frac{\pi p^*}{p^*+q} \right]^{\frac{1}{p^*} + \frac{1}{q}} \quad \boxed{= A \text{ if } q = p}$$

$$\delta_1 = \frac{1}{(q\gamma^*/p^* + 1)^{\frac{1}{q}}} \left[\sup_{x \in (0,1)} \frac{1}{x^{\gamma^*}} \int_0^x (1 - y^{q/\gamma^* p^* + 1})^{\frac{p^*}{q}} dy \right]^{\frac{1}{p^*}}, \quad \gamma^* := \frac{q}{p^* + q}$$

The results in Proposition 3.2 are shown by Figures 1–4.

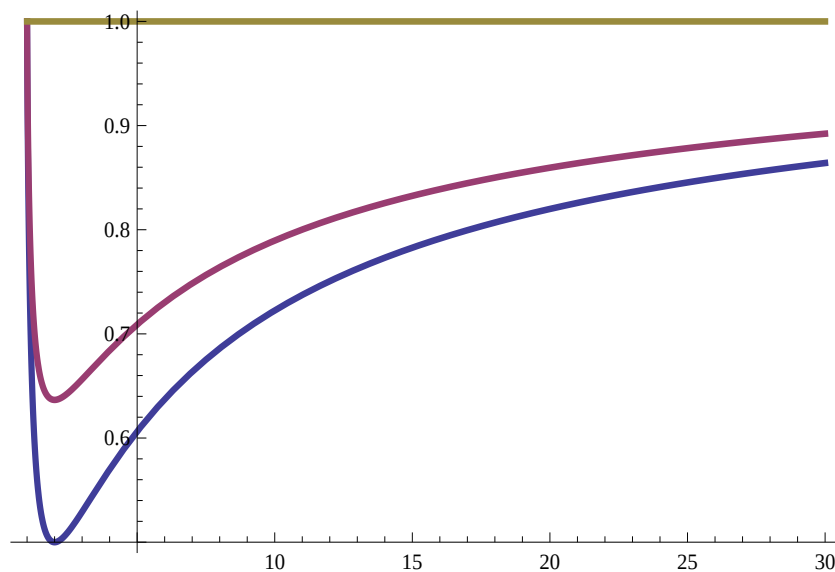


Figure 1 The basic estimates of A : $p = q \in (1, 30)$.

Note that in the case that $p = q$, we have $A^* = A$. So there are five curves only in Figure 2.

The improvements are surprisingly effective. Note that a suitable convex mean of the new upper and lower bounds should provides a quite precise approx-

ximation of A . However, the convex means of the basic estimates do not have this property. When $p = q$, much more refined results can be found from [9, 10, 11].

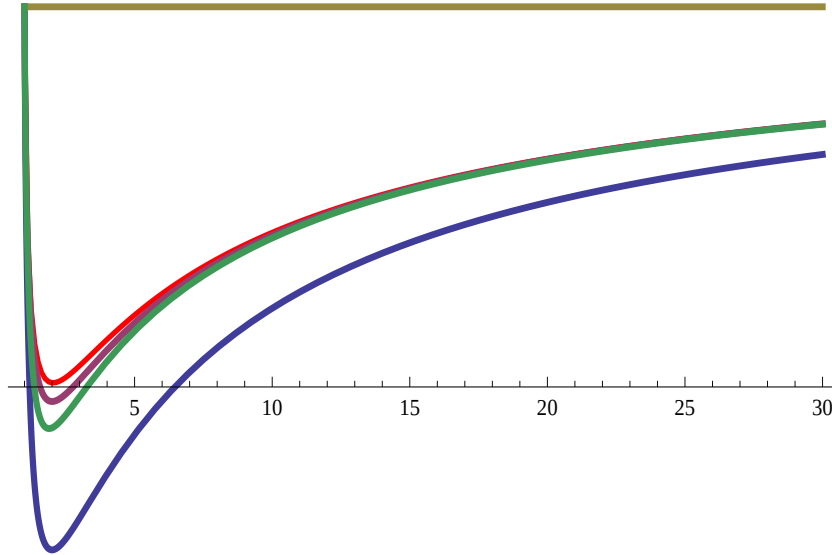


Figure 2 The basic estimates of A and their improvements: $p = q \in (1, 30)$.

Next, since $q \geq p$, we may write $q = p + r$ for some $r \geq 0$. In Figures 3 and 4, there are six curves, three of them are upper estimates and two of them are lower ones. The third curve from the bottom is the exact one; the top and the bottom curves consist of the basic estimates of the exact one. The other three curves are the improvements of the basic estimates.

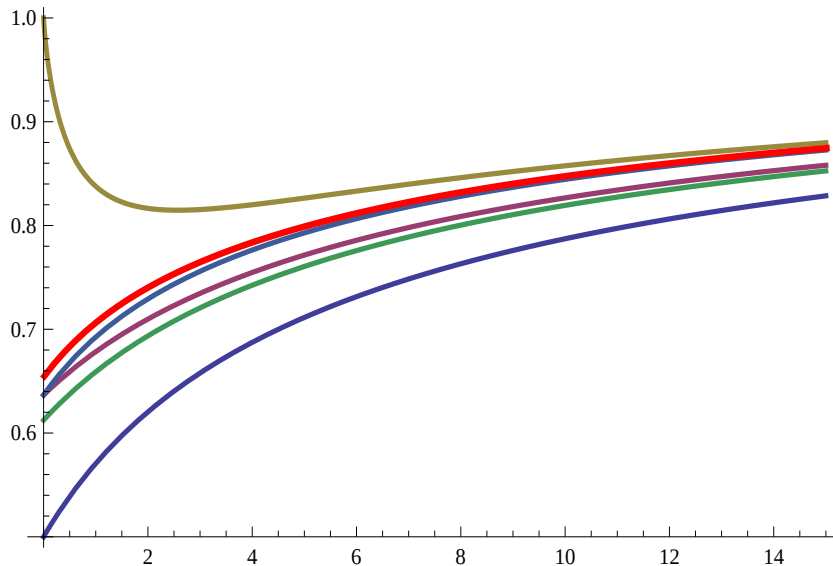


Figure 3 The basic estimates and their improvements: $p = 2$, $r \in (0, 15)$.

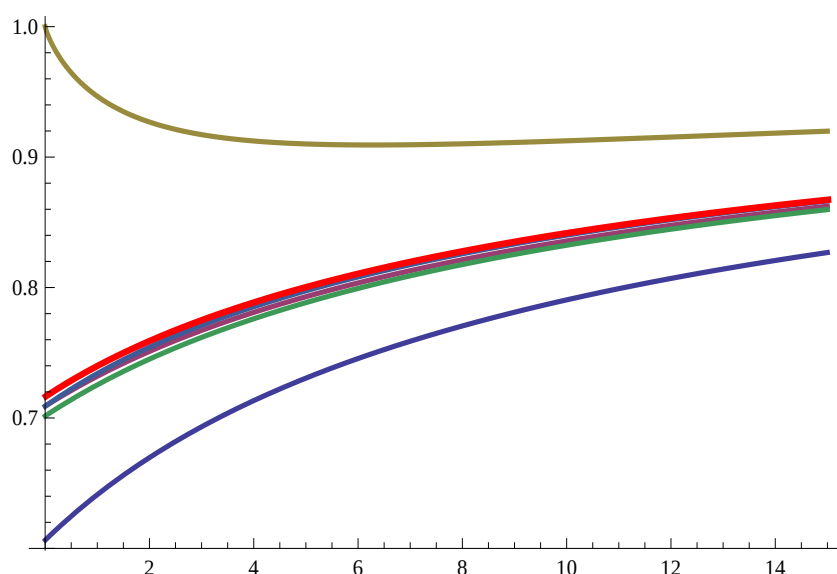


Figure 4 The basic estimates and their improvements: $p = 5$, $r \in (0, 15)$.

Note that in Figure 4, the new upper bounds and lower bounds are almost overlapped with A . In general, they are closer when p or $q \geq p$ is larger.

Finally, we mention that the main results in this note: Theorems 1.1 and 2.1, and Proposition 3.2 are new addition to the context of Hardy-type inequalities.

Acknowledgments. This short note introduces an example to illustrate the great help we have obtained from Professor Masatoshi Fukushima in the past decades. It is a nice chance to express our deepest thanks to him. Research supported in part by the National Natural Science Foundation of China (No. 11131003), and by the “985” project from the Ministry of Education in China.

References

- [1] Chen, M.F. (2004). *From Markov Chains To Non-Equilibrium Particle Systems*, 2nd ed. World Sci., Singapore.
- [2] Chen, M.F. (2005a). *Eigenvalues, Inequalities, and Ergodic Theory*. Springer, London.
- [3] Chen, M.F. (2005b). *Capacitary criteria for Poincaré-type inequalities*. Potential Analysis 23(4), 303–322.
- [4] Chen, M.F. (2010). *Speed of stability for birth–death processes*. Front. Math. China 5(3), 379–515.
- [5] Chen, M.F. (2012). *Basic estimates of stability rate for one-dimensional diffusions*. Chapter 6 in “Probability Approximations and Beyond”, 75–99, Lecture Notes in Statistics 205, eds. A.D. Barbour, H.P. Chan and D. Siegmund.
- [6] Chen, M.F. (2013a). *Bilateral Hardy-type inequalities*. Acta Math. Sin. Eng. Ser. 2013, 29:1, 1–32.

- [7] Chen, M.F. (2013b). *The optimal constant in Hardy-type inequalities*, in preparation.
- [8] Chen, M.F. and Wang, F.Y. (1997). Estimation of spectral gap for elliptic operators. *Trans. Amer. Math. Soc.* 349(3), 1239–1267.
- [9] Chen M.F., Wang L.D., Zhang Y.H. (2013a) *Mixed principal eigenvalues in dimension one*. *Front. Math. China*, 2013, 8(2): 317-343.
- [10] Chen, M.F., Wang, L.D., Zhang, Y.H. (2013b) *Mixed eigenvalues of discrete p -Laplacian*, preprint.
- [11] Chen, M.F., Wang, L.D., Zhang, Y.H. (2013c) *Mixed Eigenvalues of p -Laplacian*, preprint.
- [12] Fukushima, M. (1980). *Dirichlet Forms and Markov Processes*. North Holland Math. 23.
- [13] Fukushima, M., Oshima, Y. and Takeda, M. (2011). *Dirichlet Forms and Symmetric Markov Processes*, 2nd ed. Walter de Gruyter.
- [14] Fukushima, M. and Uemura, T. (2003). *Capacitary bounds of measures and ultracontractivity of time changed processes*. *J. Math. Pure et Appliquées* 82(5), 553–572.
- [15] Kufner, A., Maligranda, L. and Persson, L.E. (2007). *The Hardy Inequality: About its History and Some Related Results*. Vydavatelsky Servis.

School of Mathematical Sciences, Beijing Normal University,
Laboratory of Mathematics and Complex Systems (Beijing Normal University),
Ministry of Education, Beijing 100875, The People's Republic of China.

E-mail: mfchen@bnu.edu.cn

Home page: http://math.bnu.edu.cn/~chenmf/main_eng.htm

Mixed Eigenvalues of Discrete p -Laplacian

Mu-Fa CHEN¹, Ling-Di WANG^{1,2}, Yu-Hui ZHANG¹

¹ School of Mathematical Sciences, Beijing Normal University, Beijing 100875, China

² School of Mathematics and Information Sciences, Henan University, Kaifeng, Henan 475004, China

© Higher Education Press and Springer-Verlag Berlin Heidelberg 2013

Abstract This paper deals with the principal eigenvalue of discrete p -Laplacian on the set of nonnegative integers. Alternatively, it is studying the optimal constant of a class of weighted Hardy inequalities. The main goal is the quantitative estimates of the eigenvalue. The paper begins with the case having reflecting boundary at origin and absorbing boundary at infinity. Several variational formulas are presented in different formulation: the difference form, the single summation form and the double summation form. As their applications, some explicit lower and upper estimates, a criterion for positivity (which was known years ago), as well as an approximating procedure for the eigenvalue are obtained. Similarly, the dual case having absorbing boundary at origin and reflecting boundary at infinity is also studied. Two examples are presented at the end of Section 2 to illustrate the value of the investigation.

Keywords Discrete p -Laplacian, mixed eigenvalue, variational formula, explicit estimate, positivity criterion, approximating procedure

MSC 60J60, 34L15

1 Introduction

In the past years, we have been interested in various aspects of stability speed, such as exponentially ergodic rate, exponential decay rate, algebraic convergence speed, exponential convergence speeds. The convergence speeds are often described by principal eigenvalues or the optimal constants in different types of inequalities. Having a great effort on the L^2 -case (refer to [1–4] and reference therein), we now come to a more general setup, studying the nonlinear p -Laplacian, especially on the discrete space $E := \{0, 1, \dots, N\}$ ($N \leq \infty$) in this paper. This is a typical topic in harmonic analysis (cf. [8]). The method adopted in this paper is analytic rather than probabilistic. Let us presume that $N < \infty$ for a moment. Following the classification given in [4, 6], where $p = 2$ was treated, we have four types of boundary conditions: DD, DN, ND, and NN, according to Dirichlet (code ‘D’) or Neumann (code ‘N’) boundary at each of the endpoints. For instance, the Neumann condition at the right endpoint means that $f_{N+1} = f_N$. For Dirichlet condition, it means that $f_{N+1} = 0$. In the

Received August 13, 2013; accepted March 3, 2014

Corresponding author: Ling-Di WANG, E-mail: wanglingdi@mail.bnu.edu.cn

continuous context, the DN-case was partially studied in [7] by Jin and Mao. For the NN- and DD-cases, one may refer to [5]. Based on [4, 6], here, we study the mixed eigenvalues (i.e., the ND- and DN-cases) of discrete p -Laplacian. Certainly, the above classification of the boundaries for p -Laplacian remains meaningful even if $N = \infty$.

The paper is organized as follows. In Section 2, we study the ND-case. First, we introduce three groups of variational formulas for the eigenvalue. As a consequence, we obtain the basic estimates (i.e., the ratio of the upper and the lower bounds is a constant) of the eigenvalue. Furthermore, an approximating procedure and some improved estimates are presented. Except the basic estimates, when $p \neq 2$, the other results seem to be new. To illustrate the power of our main results, two examples are included at the end of Section 2. Usually, the nonlinear case (here it means $p \neq 2$) is much harder than the linear one ($p = 2$). We are lucky in the present situation since most of ideas developed in [4] are still suitable in the present general setup. This saves us a lot of spaces. Thus, we do not need to publish all details, but emphasize some key points and the difference to [4]. The sketched proofs are presented in Section 3. In Section 4, the corresponding results for the DN-case are presented.

2 ND-case

Throughout the paper, denoted by $\mathcal{C}_K$ the set of functions having compact support. In this section, let $E = \{i : 0 \leq i < N + 1\}$ ($N \leq \infty$). The discrete p -Laplacian is defined as follows:

$$\Omega_p f(k) = \nu_k |f_k - f_{k+1}|^{p-2} (f_{k+1} - f_k) - \nu_{k-1} |f_{k-1} - f_k|^{p-2} (f_k - f_{k-1}), \quad p > 1,$$

where $\{\nu_k : k \in E\}$ is a positive sequence with boundary condition $\nu_{-1} = 0$ (and $f_{-1} = f_0$). Alternatively, we may rewrite Ω_p as

$$\Omega_p f(k) = \nu_k |f_k - f_{k+1}|^{p-1} \operatorname{sgn}(f_{k+1} - f_k) - \nu_{k-1} |f_{k-1} - f_k|^{p-1} \operatorname{sgn}(f_k - f_{k-1}),$$

especially when $p \in (1, 2)$. Then we have the following discrete version of the p -Laplacian eigenvalue problem with ND-boundary conditions:

$$\text{'Eigenequation':} \quad \Omega_p g(k) = -\lambda \mu_k |g_k|^{p-2} g_k, \quad k \in E; \quad (1)$$

$$\text{ND-boundary conditions:} \quad 0 \neq g_0 = g_{-1} \text{ and } g_{N+1} = 0 \text{ if } N < \infty. \quad (2)$$

If (λ, g) is a solution to the eigenvalue problem, then λ is called an "eigenvalue" and g is its eigenfunction. Especially, when $p = 2$, the first (or principal) eigenvalue corresponds to the exponential decay rate for birth-death process on half line, where $\{\mu_k\}$ is just the invariant measure of the birth-death process and $\{\nu_k\}$ is a quantity related to the recurrence criterion of the process ([4; Sections 2 and 3]).

Define

$$D_p(f) = \sum_{k \in E} \nu_k |f_k - f_{k+1}|^p, \quad p \geq 1, \quad f \in \mathcal{C}_K,$$

and the ordinary inner product

$$(f, g) = \sum_{k \in E} f_k g_k.$$

Then we have

$$D_p(f) = (-\Omega_p f, f).$$

Actually,

$$\begin{aligned} (-\Omega_p f, f) &= \sum_{k=0}^N \nu_k f_k |f_k - f_{k+1}|^{p-2} (f_k - f_{k+1}) \\ &\quad + \sum_{k=0}^N \nu_{k-1} f_k |f_k - f_{k-1}|^{p-2} (f_k - f_{k-1}). \end{aligned}$$

Since $\nu_{-1} = 0$, one may rewrite the second term as $\sum_{k=1}^N$ and then as $\sum_{k=0}^{N-1}$ by a change of the index. Combining the resulting sum with the first one, we get

$$\begin{aligned} (-\Omega_p f, f) &= \sum_{k=0}^{N-1} \nu_k |f_k - f_{k+1}|^{p-2} (f_k - f_{k+1})^2 + \nu_N |f_N|^p \\ &= \sum_{k \in E} \nu_k |f_k - f_{k+1}|^p \quad (\text{since } f_{N+1} = 0). \end{aligned}$$

In this section, we are interested in the principal eigenvalue defined by the following classical variational formula:

$$\lambda_p = \inf \{ D_p(f) : \mu(|f|^p) = 1, f \in \mathcal{C}_K \}, \quad (3)$$

where $\mu(f) = \sum_{k \in E} \mu_k f_k$. We mention that the Neumann boundary at left endpoint is described by $f_0 = f_{-1}$ or $\nu_{-1} = 0$. The Dirichlet boundary condition at right endpoint is described by $f_{N+1} = 0$ if $N < \infty$. Actually, the condition also holds even if $N = \infty$ ($f_N := \lim_{i \rightarrow N} f_i$ provided $N = \infty$) as will be proved in Proposition 3.4 below. Formula (3) can be rewritten as the following weighted Hardy inequality:

$$\mu(|f|^p) \leq A D_p(f), \quad f \in \mathcal{C}_K$$

with optimal constant $A = \lambda_p^{-1}$. This explains the relationship between the p -Laplacian eigenvalue and the Hardy's inequality. Throughout this paper, we concentrate on $p \in (1, \infty)$ since the degenerated cases that $p = 1$ or ∞ are often easier (cf. [11; Lemmas 5.4, 5.6 on Page 49 and 56, respectively]).

2.1 Main results

To state our main results, we need some notations. For $p > 1$, let p^* be its conjugate number (i.e., $1/p + 1/p^* = 1$). Define $\hat{\nu}_j = \nu_j^{1-p^*}$ and three operators

which are parallel to those introduced in [4], as follows:

$$I_i(f) = \frac{1}{\nu_i(f_i - f_{i+1})^{p-1}} \sum_{j=0}^i \mu_j f_j^{p-1} \quad (\text{single summation form}),$$

$$II_i(f) = \frac{1}{f_i^{p-1}} \left[\sum_{j \in \text{supp}(f) \cap [i, N]} \hat{\nu}_j \left(\sum_{k=0}^j \mu_k f_k^{p-1} \right)^{p^*-1} \right]^{p-1}$$

(double summation form),

$$R_i(w) = \mu_i^{-1} [\nu_i(1 - w_i)^{p-1} - \nu_{i-1}(w_{i-1}^{-1} - 1)^{p-1}] \quad (\text{difference form}).$$

We make a convention that $w_{-1} > 0$ is free and $w_N = 0$ if $N < \infty$. For the lower estimates to be studied below, their domains are defined, respectively, as follows:

$$\mathcal{F}_I = \{f : f > 0 \text{ and } f \text{ is strictly decreasing}\},$$

$$\mathcal{F}_{II} = \{f : f > 0 \text{ on } E\},$$

$$\mathcal{W} = \left\{ w : w_i \in (0, 1) \text{ if } \sum_{j \in E} \hat{\nu}_j < \infty \text{ and } w_i \in (0, 1] \text{ if } \sum_{j \in E} \hat{\nu}_j = \infty \right\}.$$

For the upper estimates, some modifications are needed to avoid the non-summable problem:

$$\begin{aligned} \widetilde{\mathcal{F}}_I &= \{f : f \text{ is strictly decreasing on some } [n, m], 0 \leq n < m < N + 1, \\ &\quad f \cdot \nu_n \mathbb{1}_{\leq m}\}, \end{aligned}$$

$$\widetilde{\mathcal{F}}_{II} = \{f : f_i > 0 \text{ up to some } m \in [1, N + 1) \text{ and then vanishes}\},$$

$$\begin{aligned} \widetilde{\mathcal{W}} &= \left\{ w : \exists m \in [1, N + 1) \text{ such that } w_i > 0 \text{ up to } m - 1, w_m = 0, \right. \\ &\quad \left. w_i < 1 - (\nu_{i-1}/\nu_i)^{p-1} (w_{i-1}^{-1} - 1) \text{ for } i = 0, 1, \dots, m \right\}. \end{aligned}$$

In some extent, these functions are imitated of eigenfunction corresponding to λ_p . Each part of Theorem 2.1 below plays a different role in our study. Operator I is used to deduce the basic estimates (Theorem 2.3) and operator II is a tool to produce our approximating procedure (Theorem 2.4). In comparing with these two operators, the operator R is easier in the computation. Noting that for each $f \in \mathcal{F}_I$, the term $\inf_{i \in E} I_i(f)^{-1}$ given in part (1) below is a lower bound of λ_p , it indicates that the formulas on the right-hand side of each term in Theorem 2.1 are mainly used for the lower estimates. Similarly, the formulas on the left-hand side are used for the upper estimates.

Theorem 2.1 For λ_p ($p > 1$), we have

(1) single summation forms:

$$\inf_{f \in \widetilde{\mathcal{F}}_I} \sup_{i \in E} I_i(f)^{-1} = \lambda_p = \sup_{f \in \mathcal{F}_I} \inf_{i \in E} I_i(f)^{-1}.$$

(2) *Double summation forms:*

$$\inf_{f \in \widetilde{\mathcal{F}}_{II}} \sup_{i \in \text{supp}(f)} II_i(f)^{-1} = \lambda_p = \sup_{f \in \mathcal{F}_{II}} \inf_{i \in E} II_i(f)^{-1}.$$

(3) *Difference forms:*

$$\inf_{w \in \widetilde{\mathcal{W}}} \sup_{i \in E} R_i(w) = \lambda_p = \sup_{w \in \mathcal{W}} \inf_{i \in E} R_i(w).$$

Moreover, the supremum on the right-hand sides of the three above formulas can be attained.

The next proposition adds some additional sets of test functions for operators I and II . For simplicity, in what follows, we use $\downarrow$ (resp. $\downarrow\downarrow$) to denote decreasing (resp. strictly decreasing). In parallel, we also use the notation $\uparrow$ and $\uparrow\uparrow$.

Proposition 2.2 *For λ_p ($p > 1$), we have*

$$\begin{aligned} \inf_{f \in \widetilde{\mathcal{F}}_I} \sup_{i \in \text{supp}(f)} II_i(f)^{-1} &= \lambda_p = \sup_{f \in \mathcal{F}_I} \inf_{i \in E} II_i(f)^{-1}, \\ \lambda_p &= \inf_{f \in \widetilde{\mathcal{F}}'_I \cup \widetilde{\mathcal{F}}_{II}} \sup_{i \in \text{supp}(f)} II_i(f)^{-1} = \inf_{f \in \widetilde{\mathcal{F}}'_I} \sup_{i \in E} I_i(f)^{-1}, \end{aligned}$$

where

$$\begin{aligned} \widetilde{\mathcal{F}}'_I &= \{f : f \downarrow\downarrow, f \text{ is positive up to some } m \in [1, N+1), \text{ then vanishes}\} \subset \widetilde{\mathcal{F}}_I, \\ \widetilde{\mathcal{F}}_{II} &= \{f : f > 0 \text{ and } f II(f)^{p^*-1} \in L^p(\mu)\}. \end{aligned}$$

Throughout the paper, we write $\tilde{\mu}[m, n] = \sum_{j=m}^n \tilde{\mu}_j$ for a measure $\tilde{\mu}$ and define $k(p) = pp^{*p-1}$ (Figure 1). Next, define

$$\sigma_p = \sup_{n \in E} (\mu[0, n] \nu[n, N]^{p-1}).$$

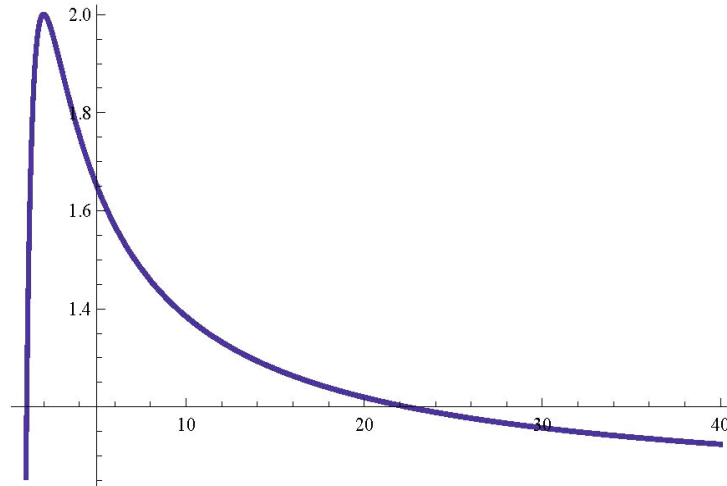


Figure 1 Function $p \rightarrow k(p)^{1/p}$ is unimodal with maximum 2 at $p = 2$.

Applying $f = \hat{\nu}[\cdot, N]^{r(p-1)}$ ($r = 1/2$ or 1) to Theorem 2.1 (1), we obtain the basic estimates given in Theorem 2.3 below. This result was known in 1990's (cf. [8; Page 58, Theorem 7]). See also [10].

Theorem 2.3 (Basic estimates) *For $p > 1$, we have $\lambda_p > 0$ if and only if $\sigma_p < \infty$. More precisely,*

$$(k(p)\sigma_p)^{-1} \leq \lambda_p \leq \sigma_p^{-1}.$$

In particular, when $N = \infty$, we have

$$\lambda_p = 0 \text{ if } \hat{\nu}[1, \infty) = \infty \text{ and } \lambda_p > 0 \text{ if } \sum_{k=0}^{\infty} \hat{\nu}_k \mu[0, k]^{p^*-1} < \infty.$$

As an application of variational formulas in Theorem 2.1 (2), we obtain an approximating procedure in the next theorem. This approach can improve the above basic estimates step by step. Noticing that λ_p is trivial once $\sigma_p = \infty$ by Theorem 2.3, we may assume that $\sigma_p < \infty$ in the study on the approximating procedure.

Theorem 2.4 (Approximating procedure) *Assume that $\sigma_p < \infty$. Let $p > 1$.*

(1) *When $\hat{\nu}[0, N] < \infty$, define*

$$f_1 = \hat{\nu}[\cdot, N]^{1/p^*}, \quad f_n = f_{n-1} II(f_{n-1})^{p^*-1} \quad (n \geq 2),$$

and $\delta_n = \sup_{i \in E} II_i(f_n)$. Otherwise, define $\delta_n = \infty$. Then δ_n is decreasing in n (denote its limit by δ_∞) and

$$\lambda_p \geq \delta_\infty^{-1} \geq \dots \geq \delta_1^{-1} \geq (k(p)\sigma_p)^{-1}.$$

(2) *For fixed $\ell, m \in E$, $\ell < m$, define*

$$f_1^{(\ell, m)} = \hat{\nu}[\cdot \vee \ell, m] \mathbb{1}_{\leq m}, \quad f_n^{(\ell, m)} = f_{n-1}^{(\ell, m)} (II(f_{n-1}^{(\ell, m)}))^{p^*-1} \mathbb{1}_{\leq m}, \quad n \geq 2,$$

where $\mathbb{1}_{\leq m}$ is the indicator of the set $\{0, 1, \dots, m\}$ and then define

$$\delta'_n = \sup_{\ell, m: \ell < m} \min_{i \leq m} II_i(f_n^{(\ell, m)}).$$

Then δ'_n is increasing in n (denote its limit by δ'_∞) and

$$\sigma_p^{-1} \geq \delta_1'^{-1} \geq \dots \geq \delta_\infty'^{-1} \geq \lambda_p.$$

Next, define

$$\bar{\delta}_n = \sup_{\ell, m: \ell < m} \frac{\mu(|f_n^{(\ell, m)}|^p)}{D_p(f_n^{(\ell, m)})}, \quad n \geq 1.$$

Then $\bar{\delta}_n^{-1} \geq \lambda_p$ and $\bar{\delta}_{n+1} \geq \delta'_n$ for $n \geq 1$.

The next result is a consequence of Theorem 2.4.

Corollary 2.5 (Improved estimates) *For $p > 1$, we have*

$$\sigma_p^{-1} \geq \delta_1'^{-1} \geq \lambda_p \geq \delta_1^{-1} \geq (k(p)\sigma_p)^{-1},$$

where

$$\delta_1 = \sup_{i \in E} \left[\frac{1}{\hat{\nu}[i, N]^{1/p^*}} \sum_{j=i}^N \hat{\nu}_j \left(\sum_{k=0}^j \mu_k \hat{\nu}[k, N]^{(p-1)/p^*} \right)^{p^*-1} \right]^{p-1},$$

$$\delta_1' = \sup_{\ell \in E} \frac{1}{\hat{\nu}[\ell, N]^{p-1}} \left[\sum_{j=\ell}^N \hat{\nu}_j \left(\sum_{k=0}^j \mu_k \hat{\nu}[k \vee \ell, N]^{p-1} \right)^{p^*-1} \right]^{p-1}.$$

Moreover,

$$\bar{\delta}_1 = \sup_{m \in E} \frac{1}{\hat{\nu}[m, N]} \sum_{j=0}^N \mu_j \hat{\nu}[j \vee m, N]^p \in [\sigma_p, p\sigma_p],$$

and $\bar{\delta}_1 \leq \delta_1'$ for $1 < p \leq 2$, $\bar{\delta}_1 \geq \delta_1'$ for $p \geq 2$.

An remarkable point of Corollary 2.5 is its last assertion which is comparable with the known result that $\bar{\delta}_1 = \delta_1'$ when $p = 2$ (cf. [4; Theorem 3.2]). This indicates that some additional work is necessary for general p than the specific one $p = 2$.

2.2 Examples

In the worst case that $p = 2$ (cf. Figure 1), the ratio $k(p)^{1/p}$ of the upper and lower estimates is no more than 2 which can be improved (no more than $\sqrt{2}$) by the improved estimates as shown by a large number of examples (cf. [4]). The same conclusion should also be true for general p as shown by two examples below. Actually, the effectiveness of the improved bounds δ_1 and $\bar{\delta}_1$ shown by the examples is quite unexpected.

Example 2.6 Assume that $E = \{0, 1, \dots, N\}$, $a > 0$, and $r > 1$. Let $\mu_k = r^k$, $\nu_k = ar^{k+1}$ for $k \in E$. Then

$$\sigma_p = \frac{r}{a(r-1)(r^{p^*-1}-1)^{p-1}},$$

$$\delta_1 = \frac{1}{ar(r^{1/p}-1)} \sup_{i \in E} \left\{ \sum_{j=i}^N \left(r^{i-j+[(j-i+1)/p]} - r^{i-j-(i/p)} \right)^{p^*-1} \right\}^{p-1},$$

$$\bar{\delta}_1 = \frac{r^{p^*}-1}{a(r^{p^*-1}-1)^p(r-1)},$$

$$\delta_1' = \frac{1}{ar} \sup_{\ell \in E} \left\{ \sum_{j=\ell}^N \left(\frac{r^{\ell+1}-1}{r^j(r-1)} + (j-\ell)r^{\ell-j} \right)^{p^*-1} \right\}^{p-1}.$$

The improved estimates given in Corollary 2.5 are shown in Figure 2 below.

The ratio between δ_1 and δ'_1 (or $\bar{\delta}_1$) is obvious smaller than the basic estimates $k(p)$ obtained in Theorem 2.3. When $p = 2$, $\bar{\delta}_1 = \delta'_1$ which is known as just mentioned.

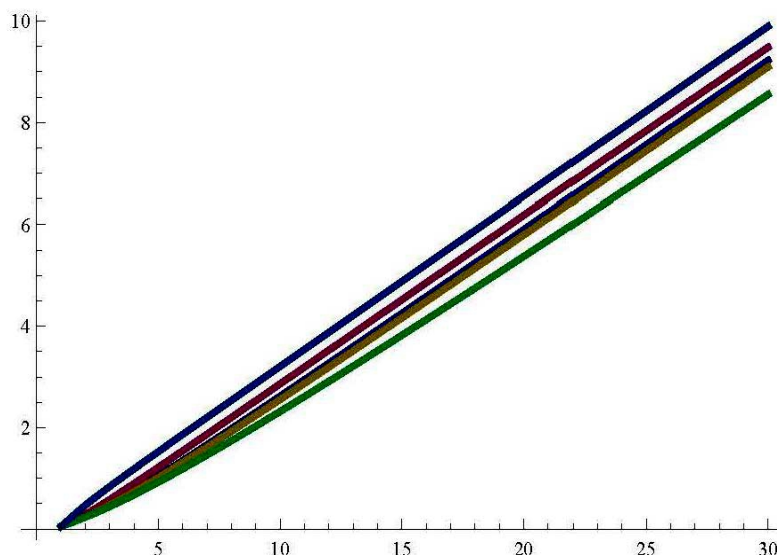


Figure 2 Let $N = 80$, $a = 1$, $r = 20$ and let p vary over $(1.001, 30.001)$ avoiding the singularity at $p = 1$. Viewing from the right-hand side, the curves from top to bottom are $(k(p)\sigma_p)^{1/p}$, $\delta_1^{1/p}$, $\bar{\delta}_1^{1/p}$, $\delta'_1^{1/p}$ and $\sigma_p^{1/p}$, respectively.

Note that the lower bounds $\bar{\delta}_1^{1/p}$ and $\delta'_1^{1/p}$ of $\lambda_p^{-1/p}$ are nearly overlapped.

Example 2.7 Assume that $E = \{0, 1, \dots, N\}$, $N < \infty$. Let $\mu_k = 1$ and $\nu_k = 1$ for $k \in E$. Then

$$\sigma_p = \sup_{n \in E} [(n+1)(N-n+1)^{p-1}],$$

$$\delta_1 = \sup_{i \in E} \left[\frac{1}{(N-i+1)^{(p-1)/p}} \sum_{j=i}^N \left(\sum_{k=0}^j (N-k+1)^{(p-1)/p^*} \right)^{p^*-1} \right]^{p-1},$$

$$\bar{\delta}_1 = \sup_{m \in E} \left(m(N-m+1)^{p-1} + \frac{1}{N-m+1} \sum_{j=m}^N (N-j+1)^p \right),$$

$$\delta'_1 = \sup_{\ell \in E} \left\{ \frac{1}{N-\ell+1} \sum_{j=\ell}^N \left[\ell(N-\ell+1)^{p-1} + \sum_{k=\ell}^j (N-k+1)^{p-1} \right]^{p^*-1} \right\}^{p-1}.$$

Surprisingly, the improved estimates δ_1 , δ'_1 and $\bar{\delta}_1$ are nearly overlapped as shown in Figure 3.

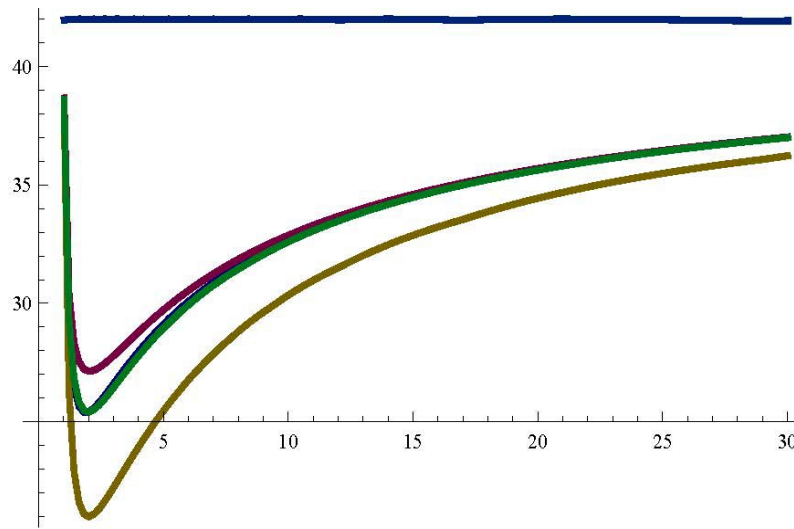


Figure 3 Let $N = 40$ and let p vary over $(1.0175, 30.0175)$ avoiding the singularity at $p = 1$. Viewing from the right-hand side, the curves from top to bottom are again $(k(p)\sigma_p)^{1/p}$, $\delta_1^{1/p}$, $\bar{\delta}_1^{1/p}$, $\delta_1'^{1/p}$, and $\sigma_p^{1/p}$, respectively. Note that $\bar{\delta}_1^{1/p}$ and $\delta_1'^{1/p}$ (lower bounds of $\lambda_p^{-1/p}$), as well as $\delta_1^{1/p}$ (upper bound) are nearly overlapped, except in a small neighborhood of $p = 2$.

For this example, the exact λ_p is unknown except that $\lambda_p = \sin^2 \frac{\pi}{2(N+2)}$ when $p = 2$.

3 Proofs of the main results in Section 2

This section is organized as follows. Some preparations are collected in Subsection 3.1. The preparations may not be needed completely for our proofs here, but they are useful for the study in a more general setup. The proofs of the main results are presented in Subsection 3.2.

3.1 Some preparations

A large part of the results stated in Section 2 depend on the properties of the eigenfunction g of λ_p . The goal of this subsection is studying these properties.

Define an operator

$$\bar{\Omega}_p f(k) = \Omega_p f(k) - \mu_k d_k |f_k|^{p-2} f_k, \quad k \in E, \quad p > 1,$$

where $\{d_k\}_{k \in E}$ is a fixed nonnegative sequence. Then there is an extended equation of (1):

$$\bar{\Omega}_p f(k) = -\bar{\lambda} \mu_k |f_k|^{p-2} f_k, \quad k \in E, \quad (4)$$

which coincides with equation (1) for $\bar{\lambda} = \lambda$ if $d_k = 0$ for every $k \in E$.

Proposition 3.1 *Define*

$$\bar{D}_p(f) = D_p(f) + \sum_{k \in E} d_k \mu_k |f_k|^p, \quad f \in \mathcal{C}_K.$$

Let

$$\bar{\lambda} = \inf \{ \bar{D}_p(f) : \mu(|f|^p) = 1, f \in \mathcal{C}_K \text{ and } f_{N+1} = 0 \text{ if } N < \infty \}. \quad (5)$$

Then the solution, say g , to equation (4) with ND-boundaries is either positive or negative. In particular, the assertion holds for the eigenfunction of λ_p .

Proof Since $g_{-1} = g_0$, by making summation from 0 to $i \in E$ with respect to k on both sides of (4), we get

$$\nu_i |g_i - g_{i+1}|^{p-2} (g_i - g_{i+1}) = \sum_{k=0}^i (\bar{\lambda} - d_k) \mu_k |g_k|^{p-2} g_k, \quad i \in E. \quad (6)$$

If $\bar{\lambda} = 0$, then the assertion is obvious by (6) and induction. If $\bar{\lambda} > 0$, then $g_0 \neq 0$ (otherwise $g \equiv 0$). Without loss of generality, assume that $g_0 = 1$ (if not, replace g with g/g_0). Suppose that there exists k_0 , $1 \leq k_0 < N$ such that $g_i > 0$ for $i < k_0$ and $g_{k_0} \leq 0$. Let

$$f_i = g_i \mathbb{1}_{i < k_0} + \varepsilon \mathbb{1}_{i = k_0}$$

for some $0 < \varepsilon < g_{k_0-1}$. Then f belongs to the setting defining $\bar{\lambda}$ (cf. (5)). Since $g_{k_0} \leq 0 < \varepsilon < g_{k_0-1}$ and $|\varepsilon - g_{k_0-1}| < |g_{k_0} - g_{k_0-1}|$, we have

$$\begin{aligned} \bar{\Omega}_p f(k_0-1) &= \bar{\Omega}_p g(k_0-1) + \nu_{k_0-1} |g_{k_0-1} - g_{k_0}|^{p-2} (g_{k_0-1} - g_{k_0}) \\ &\quad - \nu_{k_0-1} |\varepsilon - g_{k_0-1}|^{p-2} (g_{k_0-1} - \varepsilon) \\ &\geq \bar{\Omega}_p g(k_0-1) - \nu_{k_0-1} |\varepsilon - g_{k_0-1}|^{p-2} [(g_{k_0-1} - \varepsilon) - (g_{k_0-1} - g_{k_0})] \\ &= \bar{\Omega}_p g(k_0-1) - \nu_{k_0-1} (g_{k_0-1} - \varepsilon)^{p-2} (g_{k_0} - \varepsilon) \\ &> \bar{\Omega}_p g(k_0-1), \\ \bar{\Omega}_p f(k_0) &= -\nu_{k_0} \varepsilon^{p-1} + \nu_{k_0-1} |g_{k_0-1} - \varepsilon|^{p-2} (g_{k_0-1} - \varepsilon) - \mu_{k_0} d_{k_0} \varepsilon^{p-1}. \end{aligned}$$

Hence,

$$\begin{aligned} \bar{D}_p(f) &= (-\bar{\Omega}_p f, f) \\ &= - \sum_{i=0}^{k_0-2} f_i \bar{\Omega}_p f(i) - f_{k_0-1} \bar{\Omega}_p f(k_0-1) - \varepsilon \bar{\Omega}_p f(k_0) \\ &< - \sum_{i=0}^{k_0-2} g_i \bar{\Omega}_p g(i) - g_{k_0-1} \bar{\Omega}_p g(k_0-1) - \varepsilon \bar{\Omega}_p g(k_0) \\ &= \bar{\lambda} \sum_{i=0}^{k_0-1} \mu_i |g_i|^{p-2} g_i^2 + \varepsilon [(\nu_{k_0} + \mu_{k_0} d_{k_0}) \varepsilon^{p-1} - \nu_{k_0-1} (g_{k_0-1} - \varepsilon)^{p-1}]. \end{aligned}$$

In the second equality, we have used the fact that

$$\nu_i |g_i - g_{i+1}|^{p-2} (g_i - g_{i+1}) g_i = \sum_{k=0}^i (\bar{\lambda} - d_k) \mu_k |g_k|^p - \sum_{k=0}^{i-1} \nu_k |g_k - g_{k+1}|^p, \quad i \in E, \quad (7)$$

which can be obtained from (4), by a computation similar to that of $(-\Omega_p f, f)$ given above (3). Noticing that

$$\mu(|f|^p) = \sum_{i=0}^{k_0-1} \mu_i |g_i|^{p-2} g_i^2 + \mu_{k_0} \varepsilon^p$$

and

$$\nu_{k_0} + \mu_{k_0} d_{k_0} - \bar{\lambda} \mu_{k_0} < \nu_{k_0-1} \left(\frac{g_{k_0-1}}{\varepsilon} - 1 \right)^{p-1}$$

for small enough ε , we obtain a contradiction to (5):

$$\frac{\bar{D}_p(f)}{\mu(|f|^p)} < \bar{\lambda} \leq \frac{\bar{D}_p(f)}{\mu(|f|^p)}.$$

This proves the first assertion and then the second one is obvious. $\square$

Before moving further, we introduce an equation which is somehow more general than eigenequation:

$$\text{Poisson equation:} \quad \Omega_p g(k) = -\mu_k |f_k|^{p-2} f_k. \quad (8)$$

By putting $f = \lambda_p^{p^*-1} g$, we return to eigenequation. From (8), for $i, j \in E$ with $i < j$ we obtain

$$\nu_j |g_j - g_{j+1}|^{p-2} (g_j - g_{j+1}) - \nu_{i-1} |g_{i-1} - g_i|^{p-2} (g_{i-1} - g_i) = \sum_{k=i}^j \mu_k |f_k|^{p-2} f_k. \quad (9)$$

Moreover, if g is positive and decreasing, then

$$g_n - g_{N+1} = \sum_{j=n}^N \left(\frac{1}{\nu_j} \sum_{k=0}^j \mu_k |f_k|^{p-2} f_k \right)^{p^*-1}, \quad n \in E. \quad (10)$$

Besides Proposition 3.1, two more propositions are needed. One describes the monotonicity of the eigenfunction presented in the next proposition, and the other one is about the vanishing property to be presented in Proposition 3.4.

Proposition 3.2 *Assume that (λ_p, g) is a solution to (1) with ND-boundaries and $\lambda_p > 0$. Then the eigenfunction g is strictly monotone. Furthermore,*

$$\frac{1}{\lambda_p} = \left[\frac{1}{g_n - g_{N+1}} \sum_{k=n}^N \left(\frac{1}{\nu_k} \sum_{i=0}^k \mu_i g_i^{p-1} \right)^{p^*-1} \right]^{p-1}, \quad n \in E. \quad (11)$$

Proof Without loss of generality, assume that $g_0 = 1$. The first assertion follows by letting $i = 0$ and $f = \lambda_p^{p^*-1}g$ in (9). Moreover, it is clear that g is strictly decreasing. Formula (11) then follows from (10) by letting $f = \lambda_p^{p^*-1}g$. $\square$

As mentioned above, with $f = \lambda_p^{p^*-1}g$, (9) and (10) are simple variants of eigenequation (1). However, for general test function f , the left-hand side of the function g defined by (10) may be far away from the eigenfunction of λ_p . Nevertheless, we regard the resulting function (assuming $g_{N+1} = 0$) as a mimic of the eigenfunction. This explains where the operator II comes from: it is regarded as an approximation of λ_p^{-1} since $II(\lambda_p^{p^*-1}g) \equiv \lambda_p^{-1}$. Next, write $II_i(f)$ as u_i/v_i . Then $I_i(f)$ is defined by

$$I_i(f) = \frac{u_i - u_{i+1}}{v_i - v_{i+1}}.$$

In other words, the operator I comes from II in the use of proportional property. The operator R also comes from the eigenequation by setting $w_i = g_{i+1}/g_i$.

Remark 3.3 Define

$$\tilde{\lambda}_p = \inf \{D_p(f) : \mu(|f|^p) = 1 \text{ and } f_{N+1} = 0\}. \quad (12)$$

- (1) It is easy to check that the assertions in Propositions 3.1 and 3.2 also hold for $\tilde{\lambda}_p$ defined by (12).
- (2) Define

$$\lambda_p^{(n)} = \inf \{D_p(f) : \mu(|f|^p) = 1, f = f_{\leq n}\}.$$

We have $\lambda_p = \tilde{\lambda}_p$ and $\lambda_p^{(n)} \downarrow \lambda_p$ as $n \uparrow N$. Indeed, it is clear that

$$\lambda_p^{(n)} \geq \lambda_p \geq \tilde{\lambda}_p.$$

By definition of $\tilde{\lambda}_p$, for any fixed $\varepsilon > 0$, there exists $\bar{f} \in L^p(\mu)$ such that $\bar{f}_{N+1} = 0$ and

$$\frac{D_p(\bar{f})}{\mu(|\bar{f}|^p)} \leq \tilde{\lambda}_p + \varepsilon.$$

Define $f^{(n)} = \bar{f} \mathbb{1}_{[0,n]}$. Then $f^{(n)} \in L^p(\mu)$ and

$$D_p(f^{(n)}) \uparrow D_p(\bar{f}), \quad \mu(|f^{(n)}|^p) \uparrow \mu(|\bar{f}|^p) \quad \text{as } n \rightarrow N.$$

Next, for large enough n , we have

$$\lambda_p \leq \lambda_p^{(n)} \leq \frac{D_p(f^{(n)})}{\mu(|f^{(n)}|^p)} \leq \frac{D_p(\bar{f})}{\mu(|\bar{f}|^p)} + \varepsilon \leq \tilde{\lambda}_p + 2\varepsilon \leq \lambda_p + 2\varepsilon.$$

By letting $n \rightarrow N$ first and then $\varepsilon \downarrow 0$, it follows that $\lambda_p = \tilde{\lambda}_p$. Actually, we have $\lambda_p^{(n)} \downarrow \lambda_p$.

- (3) The test functions in (12) are described by $f_{N+1} = 0$, which can be seen as the imitations of eigenfunction, so the assertion in item (2) above also implies the vanishing property of eigenfunction to some extent.

From now on, in this section, without loss of generality, we assume that **the eigenfunction g corresponding to λ_p (or $\tilde{\lambda}_p$) is nonnegative and strictly decreasing**. The monotone property is important to our study. For example, it guarantees the meaning of the operator I defined in Section 2 since the denominator is $f_i - f_{i+1}$ there.

3.2 Sketch proof of the main results in ND-case

Since a large part of proofs are analogies of the case that $p = 2$, we need only to show some keys and some difference between the general p and the specific $p = 2$.

Proof of Theorem 2.1 We adopt the following circle arguments for the lower estimates:

$$\begin{aligned}
 \lambda_p &\geq \sup_{f \in \mathcal{F}_H} \inf_{i \in E} II_i(f)^{-1} \\
 &= \sup_{f \in \mathcal{F}_I} \inf_{i \in E} II_i(f)^{-1} \\
 &= \sup_{f \in \mathcal{F}_I} \inf_{i \in E} I_i(f)^{-1} \\
 &\geq \sup_{w \in \mathcal{W}} \inf_{i \in E} R_i(w) \\
 &\geq \lambda_p.
 \end{aligned}$$

Step 1 Prove that

$$\lambda_p \geq \sup_{f \in \mathcal{F}_H} \inf_{i \in E} II_i(f)^{-1}.$$

Clearly, we have $\lambda_p \geq \tilde{\lambda}_p$ by Remark 3.3 (2). Let $\{h_k : k \in E\}$ be a positive sequence and let g satisfy $\mu(|g|^p) = 1$ and $g_{N+1} = 0$. Then

$$\begin{aligned}
 1 &= \mu(|g|^p) \\
 &\leq \sum_{i=0}^N \mu_i \left(\sum_{k=i}^N |g_k - g_{k+1}| \right)^p \quad (\text{since } g_{N+1} = 0) \\
 &= \sum_{i=0}^N \mu_i \left(\sum_{k=i}^N |g_k - g_{k+1}| \left(\frac{\nu_k}{h_k} \right)^{1/p} \left(\frac{h_k}{\nu_k} \right)^{1/p} \right)^p \\
 &\leq \sum_{i=0}^N \mu_i \left[\left(\sum_{k=i}^N |g_k - g_{k+1}|^p \frac{\nu_k}{h_k} \right)^{1/p} \left(\sum_{j=i}^N \left(\frac{h_j}{\nu_j} \right)^{p^*/p} \right)^{1/p^*} \right]^p \\
 &\quad \quad \quad (\text{by Hölder's inequality}) \\
 &= \sum_{k=0}^N \frac{\nu_k}{h_k} |g_k - g_{k+1}|^p \sum_{i=0}^k \mu_i \left(\sum_{j=i}^N \hat{\nu}_j h_j^{p^*-1} \right)^{p-1} \\
 &\quad \quad \quad (\text{by exchanging the order of the sums}) \\
 &\leq D_p(g) \sup_{k \in E} H_k,
 \end{aligned}$$

where

$$H_k = \frac{1}{h_k} \sum_{i=0}^k \mu_i \left(\sum_{j=i}^N \hat{\nu}_j h_j^{p^*-1} \right)^{p-1}.$$

For every $f \in \mathcal{F}_H$ with $\sup_{i \in E} II_i(f) < \infty$, let

$$h_k = \sum_{j=0}^k \mu_j f_j^{p-1}.$$

Then

$$\sup_{k \in E} H_k \leq \sup_{k \in E} II_k(f)$$

by the proportional property. Hence,

$$D_p(g) \geq \inf_{k \in E} II_k(f)^{-1}$$

for every g with $\mu(|g|^p) = 1$, $g_{N+1} = 0$, and $f \in \mathcal{F}_H$. By making the supremum with respect to $f \in \mathcal{F}_H$ first and then the infimum with respect to g with $g_{N+1} = 0$ and $\mu(|g|^p) = 1$, it follows

$$\lambda_p \geq \tilde{\lambda}_p \geq \sup_{f \in \mathcal{F}_H} \inf_{i \in E} II_i(f)^{-1}.$$

Step 2 Prove that

$$\sup_{f \in \mathcal{F}_H} \inf_{i \in E} II_i(f)^{-1} = \sup_{f \in \mathcal{F}_I} \inf_{i \in E} II_i(f)^{-1} = \sup_{f \in \mathcal{F}_I} \inf_{i \in E} I_i(f)^{-1}.$$

Using the proportional property, on the one hand, for any fixed $f \in \mathcal{F}_I$, we have

$$\begin{aligned} \sup_{i \in E} II_i(f) &= \sup_{i \in E} \left[\frac{1}{f_i} \sum_{j=i}^N \hat{\nu}_j \left(\sum_{k=0}^j \mu_k f_k^{p-1} \right)^{p^*-1} \right]^{p-1} \\ &= \sup_{i \in E} \left\{ \left[\sum_{j=i}^N \hat{\nu}_j \left(\sum_{k=0}^j \mu_k f_k^{p-1} \right)^{p^*-1} \right] / \left[\sum_{j=i}^N (f_j - f_{j+1}) + f_{N+1} \right] \right\}^{p-1} \\ &\leq \sup_{j \in E} \frac{1}{\nu_j (f_j - f_{j+1})^{p-1}} \sum_{k=0}^j \mu_k f_k^{p-1} \quad (\text{since } \hat{\nu}_j = \nu_j^{1-p^*}) \\ &= \sup_{j \in E} I_j(f). \end{aligned}$$

Since $\mathcal{F}_I \subset \mathcal{F}_{II}$, by making the infimum with respect to $f \in \mathcal{F}_I$ on both sides of the inequality above, we have

$$\sup_{f \in \mathcal{F}_{II}} \inf_{i \in E} II_i(f)^{-1} \geq \sup_{f \in \mathcal{F}_I} \inf_{i \in E} II_i(f)^{-1} \geq \sup_{f \in \mathcal{F}_I} \inf_{i \in E} I_i(f)^{-1}.$$

On the other hand, for any fixed $f \in \mathcal{F}_{II}$, let

$$g = f II(f)^{p^*-1} \in \mathcal{F}_I.$$

Similar to the proof above, we have

$$\sup_{i \in E} I_i(g) \leq \sup_{i \in E} II_i(f)$$

Therefore

$$\sup_{f \in \mathcal{F}_I} \inf_{i \in E} I_i(f)^{-1} \geq \sup_{f \in \mathcal{F}_{II}} \inf_{i \in E} II_i(f)^{-1}$$

and the required assertion holds.

Step 3 Prove that

$$\sup_{f \in \mathcal{F}_{II}} \inf_{i \in E} II_i(f)^{-1} \geq \sup_{w \in \mathcal{W}} \inf_{i \in E} R_i(w).$$

First, we change the form of $R_i(w)$. Given $w \in \mathcal{W}$, let $u_{i+1} = w_i w_{i-1} \cdots w_0$ for $i \geq 0$, $u_0 = 1$. Then u is positive, strictly decreasing and $w_i = u_{i+1}/u_i$. By a simple rearrangement, we get

$$R_i(w) = \frac{1}{\mu_i} \left[\nu_i \left(1 - \frac{u_{i+1}}{u_i} \right)^{p-1} - \nu_{i-1} \left(\frac{u_{i-1}}{u_i} - 1 \right)^{p-1} \right] = -\frac{1}{\mu_i u_i^{p-1}} \Omega_p u(i).$$

Next, we prove the main assertion. Without loss of generality, assume that $\inf_{i \in E} R_i(w) > 0$. Let u be the function constructed above and let $f =$

$uR(w)^{p^*-1} > 0$. Then $\Omega_p u = -\mu f^{p-1}$. Since $\nu_{-1} = 0$, u is decreasing and $u > 0$, by (10), we have

$$u_i - u_{N+1} = \sum_{k=i}^N \left(\frac{1}{\nu_k} \sum_{j=0}^k \mu_j f_j^{p-1} \right)^{p^*-1}.$$

Hence,

$$R_i(w)^{1-p^*} = \frac{u_i}{f_i} \geq \frac{1}{f_i} \sum_{k=i}^N \left(\frac{1}{\nu_k} \sum_{j=0}^k \mu_j f_j^{p-1} \right)^{p^*-1} = II_i(f)^{p^*-1}.$$

Then

$$\sup_{f \in \mathcal{F}_{II}} \inf_{i \in E} II_i(f)^{-1} \geq \inf_{i \in E} R_i(w)$$

holds for every $w \in \mathcal{W}$ and the assertion follows immediately.

Step 4 Prove that

$$\sup_{w \in \mathcal{W}} \inf_{i \in E} R_i(w) \geq \lambda_p.$$

If $\sum_{i \in E} \hat{\nu}_i < \infty$, then choose f to be a positive function satisfying $h = f II(f)^{p^*-1} < \infty$. We have

$$h_i = \sum_{k=i}^N \hat{\nu}_k \left(\sum_{j=0}^k \mu_j f_j^{p-1} \right)^{p^*-1}, \quad h \downarrow, \quad h_i - h_{i+1} = \hat{\nu}_i \left(\sum_{j=0}^i \mu_j f_j^{p-1} \right)^{p^*-1}.$$

Let $\bar{w}_i = h_{i+1}/h_i$ for $i \in E$. By a simple calculation, we obtain

$$R_i(\bar{w}) = \frac{-\Omega_p h(i)}{\mu_i h_i^{p-1}} = \frac{f_i^{p-1}}{h_i^{p-1}} > 0.$$

If $\sum_{i \in E} \hat{\nu}_i = \infty$, then set $\bar{w} \equiv 1$. We have $R_i(\bar{w}) = 0$. So

$$\sup_{w \in \mathcal{W}} \inf_{i \in E} R_i(w) \geq 0.$$

Without loss of generality, assume that $\lambda_p > 0$. By Proposition 3.1, the eigenfunction g of λ_p is positive and strictly decreasing. Let $\bar{w}_i = g_{i+1}/g_i \in \mathcal{W}$. Then the assertion follows from the fact that $R_i(\bar{w}) = \lambda_p$ for every $i \in E$.

Step 5 We prove that the supremum in the circle arguments can be attained.

As an application of the circle arguments before Step 1, the assertion is easy in the case of $\lambda_p = 0$ since

$$0 = \lambda_p \geq \inf_{i \in E} II_i(f)^{-1} \geq 0, \quad 0 = \lambda_p \geq \inf_{i \in E} I_i(f)^{-1} \geq 0$$

for every f in the set defining λ_p and

$$\lambda_p \geq \sup_{w \in \mathcal{W}} \inf_{i \in E} R_i(w) \geq \inf_{i \in E} R_i(\bar{w}) \geq 0$$

for $\bar{w}$ used in Step 4 above. In the case that $\lambda_p > 0$ with eigenfunction g satisfying $g_0 = 1$, let $\bar{w}_i = g_{i+1}/g_i$. Then $R_i(\bar{w}) = \lambda_p$ as seen from the remarks after Proposition 3.2 and $I_i(g) = \lambda_p$ by letting $f = \lambda_p^{p^*-1}g$ in (9). Moreover, we have $II_i(g) = \lambda_p$ for $i \in E$ by letting $f = \lambda_p^{p^*-1}g$ in (10) whenever $g_{N+1} = 0$.

It remains to rule out the probability that $g_{N+1} > 0$. The Proposition 3.4 below, which is proved by the variational formulas verified in Step 1, gives us the positive answer.

Proposition 3.4 *Assume that $\lambda_p > 0$ and $p > 1$. Let g be an eigenfunction corresponding to λ_p . Then*

$$g_{N+1} := \lim_{i \rightarrow N+1} g_i = 0.$$

Proof Let $f = g - g_{N+1}$. Then $f \in \mathcal{F}_H$. By (11), we have

$$\lambda_p^{1-p^*} f_i = \sum_{j=i}^N \hat{\nu}_j \left(\sum_{k=0}^j \mu_k g_k^{p-1} \right)^{p^*-1}.$$

Denote

$$M_i = \sum_{j=i}^N \hat{\nu}_j \left(\sum_{k=0}^j \mu_k \right)^{p^*-1}.$$

If $M_i = \infty$, then

$$\lambda_p^{1-p^*} f_i = \sum_{j=i}^N \hat{\nu}_j \left(\sum_{k=0}^j \mu_k g_k^{p-1} \right)^{p^*-1} > M_i g_{N+1}.$$

There is a contradiction once $g_{N+1} \neq 0$. If $M_i < \infty$, then

$$\begin{aligned} \sup_{i \in E} II_i(f) &= \sup_{i \in E} \frac{1}{(g_i - g_{N+1})^{p-1}} \left[\sum_{j=i}^N \hat{\nu}_j \left(\sum_{k=0}^j \mu_k (g_k - g_{N+1})^{p-1} \right)^{p^*-1} \right]^{p-1} \\ &= \sup_{i \in E} \frac{1}{\lambda_p} \left\{ \sum_{j=i}^N \hat{\nu}_j \left[\sum_{k=0}^j \mu_k (g_k - g_{N+1})^{p-1} \right]^{p^*-1} \right. \\ &\quad \left. \Bigg/ \sum_{j=i}^N \hat{\nu}_j \left[\sum_{k=0}^j \mu_k g_k^{p-1} \right]^{p^*-1} \right\}^{p-1} \quad (\text{by (11)}) \\ &\leq \frac{1}{\lambda_p} \sup_{k \in E} \left(1 - \frac{g_{N+1}}{g_k} \right)^{p-1} \quad (\text{by the proportional property}) \\ &= \frac{1}{\lambda_p} \left(1 - \frac{g_{N+1}}{g_0} \right)^{p-1} \quad (\text{since } g \Downarrow) \end{aligned}$$

If $g_{N+1} > 0$, then by the variational formula for lower estimates proved in Step 1 above, we have

$$\lambda_p^{-1} \leq \inf_{f \in \mathcal{F}_H} \sup_{i \in E} II_i(f) \leq \sup_{i \in E} II_i(f) < \lambda_p^{-1},$$

which is a contradiction. So we must have $g_{N+1} = 0$. $\square$

By now, we have finished the proof for the lower estimates. From this proposition, we see that the vanishing property of eigenfunction holds naturally. So the classification also holds for $N = \infty$. Combining with (10), the vanishing property also further explains where the operator II comes from. Then, we come back to the main proof of Theorem 2.1.

For the upper estimates, we adopt the following circle arguments.

$$\begin{aligned} \lambda_p &\leq \inf_{f \in \widetilde{\mathcal{F}}'_H \cup \widetilde{\mathcal{F}}_H} \sup_{i \in \text{supp}(f)} II_i(f)^{-1} \\ &\leq \inf_{f \in \widetilde{\mathcal{F}}_H} \sup_{i \in \text{supp}(f)} II_i(f)^{-1} \\ &= \inf_{f \in \widetilde{\mathcal{F}}_I} \sup_{i \in \text{supp}(f)} II_i(f)^{-1} \\ &= \inf_{f \in \widetilde{\mathcal{F}}_I} \sup_{i \in E} I_i(f)^{-1} \\ &\leq \inf_{f \in \widetilde{\mathcal{F}}'_I} \sup_{i \in E} I_i(f)^{-1} \\ &\leq \inf_{w \in \widetilde{\mathcal{W}}} \sup_{i \in E} R_i(w) \\ &\leq \lambda_p. \end{aligned}$$

Since the proofs are parallel to that of the lower bounds part, we ignore most of the details here and only mention a technique when proving

$$\inf_{f \in \widetilde{\mathcal{F}}_H} \sup_{i \in \text{supp}(f)} II_i(f)^{-1} = \inf_{f \in \widetilde{\mathcal{F}}_I} \sup_{i \in \text{supp}(f)} II_i(f)^{-1} = \inf_{f \in \widetilde{\mathcal{F}}_I} \sup_{i \in E} I_i(f)^{-1}.$$

To see this, we adopt a small circle arguments below:

$$\lambda_p \leq \inf_{f \in \widetilde{\mathcal{F}}_H} \sup_{i \in \text{supp}(f)} II_i(f)^{-1} \leq \inf_{f \in \widetilde{\mathcal{F}}_I} \sup_{i \in \text{supp}(f)} II_i(f)^{-1} \leq \inf_{f \in \widetilde{\mathcal{F}}_I} \sup_{i \in E} I_i(f)^{-1} \leq \lambda_p.$$

The technique is about an approximating procedure, which is used to prove the last inequality above. Recall that

$$\lambda_p^{(m)} = \inf \{ D_p(f) : \mu(|f|^p) = 1, f = f \mathbf{1}_{\leq m} \}$$

and $\lambda_p^{(m)} \downarrow \lambda_p$ as $m \uparrow N$ (see Remark 3.3). Let g be an eigenfunction of $\lambda_p^{(m)} > 0$ with $g_0 = 1$. Then $\{g_i\}_{i=0}^m$ is strictly decreasing and $g_{m+1} = 0$ by letting E be

$E^{(m)} := [0, m] \cap E$ in Proposition 3.2. Extend g to E with $g_i = 0$ for $i \geq m+1$, we have

$$g \in \widetilde{\mathcal{F}}_I', \quad \text{supp}(g) = \{0, 1, \dots, m\} \quad \text{and} \quad \lambda_p^{(m)} = I_k(g)^{-1} \quad \text{for } k \leq m.$$

Hence,

$$\lambda_p^{(m)} = \sup_{i \leq m} I_i(g)^{-1} \geq \inf_{f \in \widetilde{\mathcal{F}}_I', \text{supp}(f) = E^{(m)}} \sup_{k \in E} I_k(f)^{-1} \geq \inf_{f \in \widetilde{\mathcal{F}}_I'} \sup_{k \in E} I_k(f)^{-1}.$$

Since $\widetilde{\mathcal{F}}_I' \subset \widetilde{\mathcal{F}}_I$, the right-hand side of the formula above is bounded below by $\inf_{f \in \widetilde{\mathcal{F}}_I} \sup_{k \in E} I_k(f)^{-1}$. So the required assertion follows by letting $m \rightarrow N$. $\square$

Instead of the approximating with finite state space used in the proof of the upper bound above, it seems more natural to use the truncating procedure for the “eigenfunction” g . However, the next result, which is easy to check by (9) and Proposition 3.2, shows that the procedure is not practical in general.

Remark 3.5 Let (λ_p, g) be a non-trivial solution to eigenequation (1) and (2) with $\lambda_p > 0$. Define $g^{(m)} = g \mathbb{1}_{\leq m}$. Then

$$\min_{i \in \text{supp}(g^{(m)})} II_i(g^{(m)}) = (1 - g_{m+1}/g_m)^{p-1} / \lambda_p.$$

In particular, the sequence $\{\min_{i \in \text{supp}(g^{(m)})} II_i(g^{(m)})\}_{m \geq 1}$ may not converge to λ_p^{-1} as $m \uparrow \infty$.

Proof The proof is simply an application of $f = \lambda_p^{1/(p-1)} g$ to (9), based on Proposition 3.2. $\square$

For simplicity, we write $\varphi_i = \hat{\nu}[i, N]^{p-1}$ in the proofs of Theorems 2.3, 2.4 and Corollary 2.5 below.

Proof of Theorem 2.3 (a) First, we prove that $\lambda_p \geq (k(p)\sigma_p)^{-1}$. Without loss of generality, assume that $\varphi_0 < \infty$ (otherwise $\sigma_p = \infty$). Let $f = \varphi^{1/p} < \infty$. Using the summation by parts formula, we have

$$\begin{aligned} \sum_{j=0}^i \mu_j f_j^{p-1} &= \mu[0, i] \varphi_i^{1/p^*} + \sum_{j=0}^{i-1} \mu[0, j] \left(\varphi_j^{1/p^*} - \varphi_{j+1}^{1/p^*} \right) \\ &\leq \sigma_p \left[\varphi_i^{-1/p} + \sum_{j=0}^{i-1} \frac{1}{\varphi_j} \left(\varphi_j^{1/p^*} - \varphi_{j+1}^{1/p^*} \right) \right] \\ &\leq p \sigma_p \varphi_i^{-1/p} \end{aligned}$$

In the last inequality, we have used the fact that

$$\sum_{j=0}^{i-1} \frac{1}{\varphi_j} \left(\varphi_j^{1/p^*} - \varphi_{j+1}^{1/p^*} \right) \leq (p-1) \varphi_i^{-1/p}.$$

To see this, since $\varphi_0 > 0$, it suffices to show that

$$\varphi_j^{1/p^*} - \varphi_{j+1}^{1/p^*} \leq (p-1)\varphi_j \left(\varphi_{j+1}^{-1/p} - \varphi_j^{-1/p} \right).$$

Multiplying $\varphi_{j+1}^{1/p}$ on both sides, this is equivalent to

$$p\varphi_j^{1/p^*} \varphi_{j+1}^{1/p} \leq (p-1)\varphi_j + \varphi_{j+1}^{1/p^*} \varphi_{j+1}^{1/p},$$

which is now obvious by Young's inequality:

$$\varphi_j^{1/p^*} \varphi_{j+1}^{1/p} \leq \frac{1}{p^*} \left(\varphi_j^{1/p^*} \right)^{p^*} + \frac{1}{p} \left(\varphi_{j+1}^{1/p} \right)^p.$$

Since

$$\frac{1}{\nu_i} = \hat{\nu}_i^{p-1} = \left(\varphi_i^{p^*-1} - \varphi_{i+1}^{p^*-1} \right)^{p-1}, \quad \varphi_i^{1/p(p-1)} \varphi_{i+1}^{1/p} \leq \frac{1}{p} \varphi_i^{p^*-1} + \frac{1}{p^*} \varphi_{i+1}^{p^*-1},$$

we have

$$\begin{aligned} I_i(f) &= \frac{1}{\nu_i (\varphi_i^{1/p} - \varphi_{i+1}^{1/p})^{p-1}} \sum_{j=0}^i \mu_j \varphi_j^{1/p^*} \\ &\leq \frac{p\sigma_p \varphi_i^{-1/p}}{\nu_i (\varphi_i^{1/p} - \varphi_{i+1}^{1/p})^{p-1}} \\ &= p\sigma_p \left[\left(\varphi_i^{p^*-1} - \varphi_{i+1}^{p^*-1} \right) / \left(\varphi_i^{p^*-1} - \varphi_i^{1/p(p-1)} \varphi_{i+1}^{1/p} \right) \right]^{p-1} \\ &\leq pp^{*p-1} \sigma_p. \end{aligned} \tag{13}$$

Then the required assertion follows by Theorem 2.1 (1).

(b) Next, we prove that $\lambda_p \geq \sigma_p^{-1}$. Let $f = \hat{\nu}[\cdot \vee n, m] \mathbb{1}_{\leq m}$ for some $m, n \in E$ with $n < m$. Then $f \in \widetilde{\mathcal{F}}_I$ and $f_i - f_{i+1} = \hat{\nu}_i \mathbb{1}_{n \leq i \leq m}$. By convention $1/0 = \infty$, we have

$$I_i(f) = \left(\sum_{k=0}^n \mu_k \hat{\nu}[n, m]^{p-1} + \sum_{k=n+1}^i \mu_k \hat{\nu}[k, m]^{p-1} \right) \mathbb{1}_{[n, m]} + \infty \mathbb{1}_{[n, m]^c}.$$

So

$$\begin{aligned} \lambda_p^{-1} &= \sup_{f \in \widetilde{\mathcal{F}}_I} \inf_{i \in E} I_i(f) \\ &\geq \inf_{i \in E} I_i(f) \\ &= \inf_{n \leq i \leq m} I_i(f) \\ &= \inf_{n \leq i \leq m} \left(\sum_{k=0}^n \mu_k \hat{\nu}[n, m]^{p-1} + \sum_{k=n+1}^i \mu_k \hat{\nu}[k, m]^{p-1} \right) \\ &= \sum_{k=0}^n \mu_k \hat{\nu}[n, m]^{p-1}, \quad m > n. \end{aligned}$$

The assertion that $\lambda_p^{-1} \geq \sigma_p$ follows by letting $m \rightarrow N$.

(c) At last, if $\hat{\nu}[1, \infty) = \infty$, then $\lambda_p = 0$ is obvious. If

$$\sum_{k=1}^{\infty} \hat{\nu}_k \mu[0, k]^{p^*-1} < \infty,$$

then

$$\varphi_n \mu[0, n] = \left(\sum_{k=n}^{\infty} \hat{\nu}_k \mu[0, k]^{p^*-1} \right)^{p-1} \leq \left(\sum_{k=1}^{\infty} \hat{\nu}_k \mu[0, k]^{p^*-1} \right)^{p-1} < \infty.$$

So $\sigma_p = \infty$ and $\lambda_p = 0$. $\square$

Proof of Theorem 2.4 By definitions of $\{\bar{\delta}_n\}$ and $\{\delta'_n\}$, using the proportional property, it is not hard to prove most of the results except that $\bar{\delta}_{n+1} \geq \delta'_n$ ($n \geq 1$). Put $g = f_{n+1}^{(\ell, m)}$ and $f = f_n^{(\ell, m)}$. Then $g = f \Pi(f)^{p^*-1} \mathbb{1}_{\leq m}$. By a simple calculation, we have

$$(g_i - g_{i+1})^{p-1} = \frac{1}{\nu_i} \sum_{k=0}^i \mu_k f_k^{p-1}, \quad i \leq m.$$

Inserting this term into $D_p(g)$, we obtain

$$D_p(g) = \sum_{i=0}^m \nu_i (g_i - g_{i+1})^{p-1} (g_i - g_{i+1}) = \sum_{i=0}^m \sum_{k=0}^i \mu_k f_k^{p-1} (g_i - g_{i+1}).$$

Noticing $g_{m+1} = 0$ and exchanging the order of the sums, we obtain

$$D_p(g) = \sum_{k=0}^m \mu_k f_k^{p-1} \sum_{i=k}^m (g_i - g_{i+1}) = \sum_{k=0}^m \mu_k f_k^{p-1} g_k \leq \sum_{k=0}^m \mu_k g_k^p \max_{0 \leq i \leq m} \frac{f_k^{p-1}}{g_k^{p-1}},$$

i.e.,

$$D_p(g) \leq \mu(|g|^p) \sup_{0 \leq i \leq m} \Pi_i(f)^{-1}.$$

So the required assertion follows by definitions of $\bar{\delta}_{n+1}$ and δ'_n . $\square$

Most of the results in Corollary 2.5 can be obtained from Theorem 2.4 directly. Here, we study only those assertions concerning δ'_1 and $\bar{\delta}_1$.

Proof of Corollary 2.5 (a) We compute δ'_1 first. Since $p > 1$ and

$$\frac{1}{\hat{\nu}[i, m]} \sum_{j=i}^m \hat{\nu}_j \left(\sum_{k=0}^j \mu_k \hat{\nu}[k \vee \ell, m]^{p-1} \right)^{p^*-1}$$

is increasing in $i \in [\ell, m]$ (not hard to check), we have

$$\min_{i \leq m} \Pi_i(f_1^{(\ell, m)}) = \left[\frac{1}{\hat{\nu}[\ell, m]} \sum_{j=\ell}^m \hat{\nu}_j \left(\sum_{k=0}^j \mu_k \hat{\nu}[k \vee \ell, m]^{p-1} \right)^{p^*-1} \right]^{p-1}.$$

We claim that

$$\delta'_1 = \sup_{\ell \in E} \frac{1}{\varphi_\ell} \left[\sum_{j=\ell}^N \hat{\nu}_j \left(\sum_{k=0}^j \mu_k \varphi_{k \vee \ell} \right)^{p^*-1} \right]^{p-1}$$

because

$$\frac{1}{\hat{\nu}[\ell, m]^{p-1}} \left[\sum_{j=\ell}^m \hat{\nu}_j \left(\sum_{k=0}^j \mu_k \hat{\nu}[k \vee \ell, m]^{p-1} \right)^{p^*-1} \right]^{p-1}$$

is increasing in m ($m > \ell$). To see this, it suffices to show that

$$\begin{aligned} & \frac{1}{\hat{\nu}[\ell, m+1]} \sum_{j=\ell}^{m+1} \hat{\nu}_j \left(\sum_{k=0}^j \mu_k \hat{\nu}[k \vee \ell, m+1]^{p-1} \right)^{p^*-1} \\ & \geq \frac{1}{\hat{\nu}[\ell, m]} \sum_{j=\ell}^m \hat{\nu}_j \left(\sum_{k=0}^j \mu_k \hat{\nu}[k \vee \ell, m]^{p-1} \right)^{p^*-1}. \end{aligned}$$

Equivalently,

$$\sum_{j=\ell}^{m+1} \hat{\nu}_j \left(\sum_{k=0}^j \mu_k \frac{\hat{\nu}[k \vee \ell, m+1]^{p-1}}{\hat{\nu}[\ell, m+1]^{p-1}} \right)^{p^*-1} \geq \sum_{j=\ell}^m \hat{\nu}_j \left(\sum_{k=0}^j \mu_k \frac{\hat{\nu}[k \vee \ell, m]^{p-1}}{\hat{\nu}[\ell, m]^{p-1}} \right)^{p^*-1}.$$

It suffices to show that

$$\frac{\hat{\nu}[k \vee \ell, m+1]}{\hat{\nu}[\ell, m+1]} \geq \frac{\hat{\nu}[k \vee \ell, m]}{\hat{\nu}[\ell, m]}, \quad k \in E.$$

When $k \leq \ell$, the required assertion is obvious. When $k > \ell$, the inequality is just

$$\frac{\hat{\nu}[k, m+1]}{\hat{\nu}[k, m]} \geq \frac{\hat{\nu}[\ell, m+1]}{\hat{\nu}[\ell, m]}.$$

Noticing that $\hat{\nu}[i, m+1] = \hat{\nu}_{m+1} + \hat{\nu}[i, m]$ for any fixed $i \leq m$ and $\hat{\nu}[k, m] < \hat{\nu}[\ell, m]$ for $k > \ell$, we have

$$\frac{\hat{\nu}[k, m+1]}{\hat{\nu}[k, m]} = 1 + \frac{\hat{\nu}_{m+1}}{\hat{\nu}[k, m]} > 1 + \frac{\hat{\nu}_{m+1}}{\hat{\nu}[\ell, m]} = \frac{\hat{\nu}[\ell, m+1]}{\hat{\nu}[\ell, m]},$$

and then the required monotone property follows.

(b) Computing $\bar{\delta}_1$. Since

$$\mu(|f_1^{(\ell, m)}|^p) = \sum_{j=0}^m \mu_j \hat{\nu}[\ell \vee j, m]^p = \mu[0, \ell] \hat{\nu}[\ell, m]^p + \sum_{j=\ell+1}^m \mu_j \hat{\nu}[j, m]^p,$$

and

$$\begin{aligned} D_p(f_1^{(\ell,m)}) &= \sum_{j=0}^m \nu_j \left(f_1^{(\ell,m)}(j) - f_1^{(\ell,m)}(j+1) \right)^p \\ &= \sum_{j=\ell}^m \nu_j \hat{\nu}_j^p \\ &= \hat{\nu}[\ell, m] \quad \left(\text{since } \hat{\nu}_k = \nu_k^{1-p^*} \right), \end{aligned}$$

we have

$$\frac{\mu(|f_1^{(\ell,m)}|^p)}{D_p(f_1^{(\ell,m)})} = \hat{\nu}[\ell, m]^{p-1} \mu[0, \ell] + \frac{1}{\hat{\nu}[\ell, m]} \sum_{k=\ell+1}^m \mu_k \hat{\nu}[k, m]^p.$$

So

$$\bar{\delta}_1 = \sup_{\ell, m \in E: \ell < m} \left(\hat{\nu}[\ell, m]^{p-1} \mu[0, \ell] + \frac{1}{\hat{\nu}[\ell, m]} \sum_{k=\ell+1}^m \mu_k \hat{\nu}[k, m]^p \right).$$

The assertion on $\bar{\delta}_1$ follows immediately once we show that

$$\hat{\nu}[\ell, m]^{p-1} \mu[0, \ell] + \frac{1}{\hat{\nu}[\ell, m]} \sum_{k=\ell+1}^m \mu_k \hat{\nu}[k, m]^p$$

is increasing in m ($\ell < m$). To see this, it suffices to show that

$$\frac{1}{\hat{\nu}[\ell, m]} \sum_{k=\ell+1}^m \mu_k \hat{\nu}[k, m]^p \leq \frac{1}{\hat{\nu}[\ell, m+1]} \sum_{k=\ell+1}^{m+1} \mu_k \hat{\nu}[k, m+1]^p,$$

or equivalently,

$$\frac{\mu_{m+1}}{\hat{\nu}[\ell, m+1]} \hat{\nu}_{m+1}^p + \sum_{k=\ell+1}^m \mu_k \left(\frac{\hat{\nu}[k, m+1]^p}{\hat{\nu}[\ell, m+1]} - \frac{\hat{\nu}[k, m]^p}{\hat{\nu}[\ell, m]} \right) \geq 0.$$

Since $p > 1$ and $k > \ell$, we have

$$\left(\frac{\hat{\nu}[k, m+1]}{\hat{\nu}[k, m]} \right)^p > \frac{\hat{\nu}[k, m+1]}{\hat{\nu}[k, m]} = 1 + \frac{\hat{\nu}_{m+1}}{\hat{\nu}[k, m]} > 1 + \frac{\hat{\nu}_{m+1}}{\hat{\nu}[\ell, m]} = \frac{\hat{\nu}[\ell, m+1]}{\hat{\nu}[\ell, m]}.$$

So the required assertion holds.

(c) We compare $\bar{\delta}_1$ with σ_p and δ'_1 .

For the convenience of comparison of $\bar{\delta}_1$ with σ_p and δ'_1 , we rewrite $\bar{\delta}_1$ as follows.

$$\bar{\delta}_1 = \sup_{\ell \in E} \left(\varphi_\ell \mu[0, \ell] + \frac{1}{\varphi_\ell^{p^*-1}} \sum_{k=\ell+1}^N \mu_k \varphi_k^{p(p^*-1)} \right).$$

By definition of σ_p , it is clear that $\bar{\delta}_1 \geq \sigma_p$. To compare $\bar{\delta}_1$ with δ'_1 , we further change the form of $\bar{\delta}_1$. By definition of φ , we have

$$\begin{aligned} \sum_{j=0}^N \mu_j \varphi_{j \vee m}^{p^*} &= \sum_{j=0}^{m-1} \mu_j \varphi_m \sum_{k=m}^N \hat{\nu}_k + \sum_{j=m}^N \mu_j \varphi_j \sum_{k=j}^N \hat{\nu}_k \\ &= \sum_{k=m}^N \hat{\nu}_k \sum_{j=0}^{m-1} \mu_j \varphi_m + \sum_{k=m}^N \hat{\nu}_k \sum_{j=m}^k \mu_j \varphi_j. \end{aligned}$$

So

$$\bar{\delta}_1 = \sup_{\ell \in E} \left(\frac{1}{\varphi_\ell^{p^*-1}} \sum_{k=0}^N \mu_k \varphi_{k \vee \ell}^{p^*} \right) = \sup_{m \in E} \frac{1}{\varphi_m^{p^*-1}} \sum_{k=m}^N \hat{\nu}_k \sum_{j=0}^k \mu_j \varphi_{m \vee j}.$$

Denote $a_\ell(k) = \hat{\nu}_k / \varphi_\ell^{p^*-1}$. Then $\sum_{k=\ell}^N a_\ell(k) = 1$ (i.e., $\{a_\ell(k) : k = \ell, \dots, N\}$ is a probability measure on $\{\ell, \ell+1, \dots, N\}$). By the increasing property of the moments $\mathbb{E}(|X|^s)^{1/s}$ in $s > 0$, it follows that

$$\begin{aligned} \delta'_1 &= \sup_{\ell \in E} \left[\sum_{j=\ell}^N a_\ell(j) \left(\sum_{k=0}^j \mu_k \varphi_{k \vee \ell} \right)^{p^*-1} \right]^{p-1} \\ &\geq \sup_{\ell \in E} \sum_{j=\ell}^N a_\ell(j) \sum_{k=0}^j \mu_k \varphi_{k \vee \ell} \quad (\text{if } p^* - 1 > 1) \\ &= \bar{\delta}_1. \end{aligned}$$

Hence, $\bar{\delta}_1 \leq \delta'_1$ for $1 < p \leq 2$. Otherwise, $\bar{\delta}_1 \leq \delta'_1$ for $p \geq 2$.

(d) At last, we prove that $\bar{\delta}_1 \leq p\sigma_p$. Using the summation by parts formula, we have

$$\sum_{j=0}^N \mu_j \varphi_{j \vee \ell}^{p^*} = \sum_{j=\ell}^N \left(\varphi_j^{p^*} - \varphi_{j+1}^{p^*} \right) \mu[0, j].$$

Hence,

$$\begin{aligned} \frac{1}{\varphi_m^{p^*-1}} \sum_{j=0}^N \mu_j \varphi_{j \vee m}^{p^*} &= \frac{1}{\varphi_m^{p^*-1}} \sum_{j=m}^N \left(\varphi_j^{p^*} - \varphi_{j+1}^{p^*} \right) \mu[0, j] \\ &\leq \sigma_p \frac{1}{\varphi_m^{p^*-1}} \sum_{j=m}^N \frac{1}{\varphi_j} \left(\varphi_j^{p^*} - \varphi_{j+1}^{p^*} \right) \\ &\leq \sigma_p \sum_{j=m}^N \frac{1}{\varphi_j} \left(\varphi_j^{p^*} - \varphi_{j+1}^{p^*} \right) \bigg/ \sum_{j=m}^N \left(\varphi_j^{p^*-1} - \varphi_{j+1}^{p^*-1} \right) \\ &\quad (\text{since } \varphi_{N+1} = 0). \end{aligned}$$

By Young's inequality, we have

$$\varphi_j \varphi_{j+1}^{p^*-1} \leq \frac{1}{p^*} \varphi_j^{p^*} + \frac{1}{p} \varphi_{j+1}^{p^*}.$$

Combining this inequality with the proportional property, we obtain

$$\begin{aligned} \frac{1}{\varphi_m^{p^*-1}} \sum_{j=0}^N \mu_j \varphi_{j \vee m}^{p^*} &\leq \sigma_p \sup_{j \in E} \frac{\varphi_j^{p^*} - \varphi_{j+1}^{p^*}}{\varphi_j (\varphi_j^{p^*-1} - \varphi_{j+1}^{p^*-1})} \\ &= \sigma_p \sup_{j \in E} \frac{\varphi_j^{p^*} - \varphi_{j+1}^{p^*}}{\varphi_j^{p^*} - \varphi_j \varphi_{j+1}^{p^*-1}} \\ &\leq p \sigma_p. \end{aligned}$$

So the assertion holds. $\square$

4 DN-case

In this section, we use the same notations as the last section because they play the same role. However, they have different meaning in different sections. Set $E = \{i \in \mathbb{N} : 1 \leq i < N+1\}$. Let $\{\mu_i\}_{i \in E}$ and $\{\nu_i\}_{i \in E}$ be two positive sequences. Similar to the ND-case, we have the discrete version of p -Laplacian eigenvalue problem with DN-boundaries:

$$\text{'Eigenequation': } \Omega_p g(k) = -\lambda \mu_k |g_k|^{p-2} g_k, \quad k \in E;$$

$$\text{DN-boundary conditions: } g_0 = 0 \text{ and } g_{N+1} = g_N \text{ if } N < \infty,$$

where

$$\Omega_p f(k) = \nu_{k+1} |f_{k+1} - f_k|^{p-2} (f_{k+1} - f_k) - \nu_k |f_k - f_{k-1}|^{p-2} (f_k - f_{k-1}), \quad p > 1,$$

$\nu_{N+1} := 0$ if $N < \infty$ and $\nu_{N+1} := \lim_{i \rightarrow \infty} \nu_i$ if $N = \infty$. Let λ_p denote the first eigenvalue. Then

$$\lambda_p = \inf \{D_p(f)/\mu(|f|^p) : f \neq 0, D_p(f) < \infty\}. \quad (14)$$

where

$$\mu(f) = \sum_{k \in E} \mu_k f_k \leq \infty, \quad D_p(f) = \sum_{k \in E} \nu_k |f_k - f_{k-1}|^p, \quad f_0 = 0.$$

The constant λ_p describes the optimal constant $A = \lambda_p^{-1}$ in the following *weighted Hardy inequality*:

$$\mu(|f|^p) \leq A D_p(f), \quad f(0) = 0,$$

or equivalently,

$$\|f\|_{L^p(\mu)} \leq A^{1/p} \|\partial^- f\|_{L^p(\nu)}, \quad f(0) = 0,$$

where $\partial^- f(k) = f_{k-1} - f_k$. In other words, we are studying again the weighted Hardy inequality in this section. In view of this, by a duality [9; Page 13], the optimal constant $\lambda_p^{-1/p}$ in the last inequality coincides with $\lambda_{p^*}^{-1/p^*}$ which is the optimal constant in the inequality

$$\|f\|_{L^{p^*}(\nu^{1-p^*})} \leq A'^{1/p^*} \|\partial^+ f\|_{L^{p^*}(\mu^{1-p^*})}, \quad f_{N+1} = 0,$$

where $\partial^+ f(k) = f_{k+1} - f_k$, studied in Section 2. However, due to the difference of boundaries in these two cases, the variational formulas and the approximating procedure are different (cf. [4]). Therefore, it is worthy to present some details here. Similar notations as Section 2 are defined as follows. Define $\hat{\nu}_j = \nu_j^{1-p^*}$ for $j \in E$, and

$$\begin{aligned} I_i(f) &= \frac{1}{\nu_i(f_i - f_{i-1})^{p-1}} \sum_{j=i}^N \mu_j f_j^{p-1} \quad (\text{single summation form}), \\ II_i(f) &= \frac{1}{f_i^{p-1}} \left[\sum_{j=1}^i \hat{\nu}_j \left(\sum_{k=j}^N \mu_k f_k^{p-1} \right)^{p^*-1} \right]^{p-1} \quad (\text{double summation form}), \\ R_i(w) &= \mu_i^{-1} [\nu_i(1 - w_{i-1}^{-1})^{p-1} - \nu_{i+1}(w_i - 1)^{p-1}], \quad w_0 := \infty \quad (\text{difference form}). \end{aligned}$$

For the lower bounds, the domains of the operators are defined respectively as follows:

$$\begin{aligned} \mathcal{F}_I &= \{f : f > 0 \text{ and is strictly increasing on } E\}, \\ \mathcal{F}_{II} &= \{f : f > 0 \text{ on } E\}, \\ \mathcal{W} &= \{w : w_i > 1 \text{ for } i \in E\}. \end{aligned}$$

Note that the test function given in $\mathcal{F}_I$ is different from that given in Section 2. Again, this is due to the property of eigenfunction (which is proved in Proposition 4.6 later). For the upper bounds, we need modify these sets as follows:

$$\begin{aligned} \widetilde{\mathcal{F}}_I &= \{f : \exists m \in E \text{ such that } f \text{ is strictly increasing on } \{1, \dots, m\} \text{ and } f = f_{\wedge m}\}, \\ \widetilde{\mathcal{F}}_{II} &= \{f : f = f_{\wedge m} > 0 \text{ for some } m \in E\}, \\ \widetilde{\mathcal{W}} &= \bigcup_{m \in E} \left\{ w : 1 < w_i < 1 + \nu_i^{p^*-1} (1 - w_{i-1}^{-1}) \nu_{i+1}^{1-p^*} \text{ for } 1 \leq i \leq m-1 \right. \\ &\quad \left. \text{and } w_i = 1 \text{ for } i \geq m \right\}. \end{aligned}$$

Define $\tilde{R}$ acting on $\widetilde{\mathcal{W}}$ as a modified form of R by replacing μ_m with $\tilde{\mu}_m := \sum_{k=m}^N \mu_k$ in $R_i(w)$ for the same m in $\widetilde{\mathcal{W}}$. The change of μ_m is due to the Neumann boundary at right endpoint. Note that if $w_i = 1$ for every $i \geq m$, then

$$\tilde{R}_i(w) = R_i(w) = 0, \quad i > m.$$

Besides, we also need the following set:

$$\widetilde{\mathcal{F}}_H' = \{f : f > 0 \text{ on } E \text{ and } f II(f)^{p^*-1} \in L^p(\mu)\}.$$

If $\sum_{i \in E} \mu_i = \infty$, let $f_i = 1$ for $i \in E$ and $f_0 = 0$. Then

$$D_p(f) = \sum_{k=1}^N \nu_k |f_k - f_{k-1}|^p = \nu_1 < \infty \quad \text{and} \quad \mu(|f|^p) = \infty.$$

So $\lambda_p = 0$ by (14). If $\sum_{i \in E} \mu_i < \infty$, then as we will prove later (Lemma 4.5) that λ_p coincides with

$$\lambda_p^{[1]} := \inf \{D_p(f) : \mu(|f|^p) = 1\}.$$

Actually, the later is also coincides with

$$\lambda_p^{[1]} = \inf \{D_p(f) : \mu(|f|^p) = 1, f_i = f_{i \wedge m} \text{ for some } m \in E\}.$$

Now we introduce the main results, many of which are parallel to that in Section 2. However, the exchange of boundary conditions ‘D’ and ‘N’ makes many difference. For example, the results related to $\widetilde{R}$, the definition of σ_p (see Theorem 4.2 later), and so on.

Theorem 4.1 *Assume that $p > 1$ and $\sum_{k=1}^N \mu_i < \infty$. Then the following variational formulas holds for λ_p (equivalently, $\lambda_p^{[1]}$ or $\lambda_p^{[2]}$).*

(1) *Single summation forms:*

$$\inf_{f \in \widetilde{\mathcal{F}}_I} \sup_{i \in E} I_i(f)^{-1} = \lambda_p = \sup_{f \in \widetilde{\mathcal{F}}_I} \inf_{i \in E} I_i(f)^{-1}.$$

(2) *Double summation forms:*

$$\begin{aligned} \lambda_p &= \inf_{f \in \widetilde{\mathcal{F}}_H} \sup_{i \in E} II_i(f)^{-1} = \inf_{f \in \widetilde{\mathcal{F}}_I} \sup_{i \in E} II_i(f)^{-1} = \inf_{f \in \widetilde{\mathcal{F}}_H \cup \widetilde{\mathcal{F}}_I} \sup_{i \in E} II_i(f)^{-1}, \\ \lambda_p &= \sup_{f \in \mathcal{F}_H} \inf_{i \in E} II_i(f)^{-1} = \sup_{f \in \mathcal{F}_I} \inf_{i \in E} II_i(f)^{-1}. \end{aligned}$$

(3) *Difference forms:*

$$\inf_{w \in \widetilde{\mathcal{W}}} \sup_{i \in E} \widetilde{R}_i(w) = \lambda_p = \sup_{w \in \mathcal{W}} \inf_{i \in E} R_i(w).$$

As an application of the variational formulas in Theorem 4.1 (1), we have the following theorem. This result was known in 1990’s (cf. [8; Page 58, Theorem 7] plus the duality technique, cf. [9; Page 13]). See also [10]. It can be regarded as a dual of Theorem 2.3.

Theorem 4.2 (Basic estimates) *For $p > 1$, we have λ_p (or equivalently, $\lambda_p^{[1]}$ or $\lambda_p^{[2]}$ provided $\sum_{k \in E} \mu_k < \infty$) is positive if and only if $\sigma_p < \infty$, where*

$$\sigma_p = \sup_{n \in E} (\mu[n, N] \hat{\nu}[1, n]^{p-1}).$$

More precisely,

$$(k(p)\sigma_p)^{-1} \leq \lambda_p \leq \sigma_p^{-1},$$

where $k(p) = pp^{*p-1}$. In particular, we have $\lambda_p = 0$ if $\sum_{i \in E} \mu_i = \infty$ and $\lambda_p > 0$ if $N < \infty$, or $\sum_{k=1}^{\infty} \mu[k, N]^{p^*-1} \nu_k < \infty$, or $\sum_{k=1}^{\infty} (\mu_k + \hat{\nu}_k) < \infty$.

The next result is an application of the variational formulas in Theorem 4.1 (2). It is interesting that the result is not a direct dual of Theorem 2.4.

Theorem 4.3 (Approximating procedure) *Assume that $p > 1$, $\sum_{k \in E} \mu_k < \infty$ and $\sigma_p < \infty$. Then the following assertions hold.*

(1) Define

$$f_1 = \hat{\nu}[1, \cdot]^{1/p^*}, \quad f_n = f_{n-1} II(f_{n-1})^{p^*-1} \quad (n \geq 2), \quad \delta_n = \sup_{i \in E} II_i(f_n).$$

Then δ_n is decreasing in n and

$$\lambda_p \geq \delta_{\infty}^{-1} \geq \dots \geq \delta_1^{-1} \geq (k(p)\sigma_p)^{-1}.$$

(2) For fixed $m \in E$, define

$$f_1^{(m)} = \hat{\nu}[1, \cdot \wedge m], \quad f_n^{(m)} = f_{n-1}^{(m)} II(f_{n-1}^{(m)})(\cdot \wedge m)^{p^*-1}, \quad n \geq 2,$$

and

$$\delta'_n = \sup_{m \in E} \inf_{i \in E} II_i(f_n^{(m)}).$$

Then δ'_n is increasing in n and

$$\sigma_p^{-1} \geq \delta_1'^{-1} \geq \dots \geq \delta_{\infty}'^{-1} \geq \lambda_p.$$

Next, define

$$\bar{\delta}_n = \sup_{m \in E} \frac{\mu(f_n^{(m)p})}{D_p(f_n^{(m)})}, \quad n \in E.$$

Then $\bar{\delta}_n \geq \lambda_p$ and $\bar{\delta}_{n+1} \geq \delta'_n$ for every $n \geq 1$.

Corollary 4.4 (Improved estimates) *Assume that $\sum_{k \in E} \mu_k < \infty$. For $p > 1$, we have*

$$\sigma_p^{-1} \geq \delta_1'^{-1} \geq \lambda_p^{-1} \geq \delta_1^{-1} \geq (k(p)\sigma_p)^{-1},$$

where

$$\delta_1 = \sup_{i \in E} \left[\hat{\nu}[1, i]^{-1/p^*} \sum_{j=1}^i \hat{\nu}_j \left(\sum_{k=j}^N \mu_k \hat{\nu}[1, k]^{(p-1)/p^*} \right)^{p^*-1} \right]^{p-1},$$

$$\delta'_1 = \sup_{m \in E} \frac{1}{\hat{\nu}[1, m]^{p-1}} \left[\sum_{j=1}^m \hat{\nu}_j \left(\sum_{k=j}^N \mu_k \hat{\nu}[1, k \wedge m]^{p-1} \right)^{p^*-1} \right]^{p-1}.$$

Moreover,

$$\bar{\delta}_1 = \sup_{m \in E} \frac{1}{\hat{\nu}[1, m]} \sum_{j=1}^N \mu_j \hat{\nu}[1, j \wedge m]^p \in [\sigma_p, p\sigma_p],$$

and $\bar{\delta}_1 \geq \delta'_1$ for $p \geq 2$, $\bar{\delta}_1 \leq \delta'_1$ for $1 < p \leq 2$.

When $p = 2$, the result that $\delta'_1 = \bar{\delta}_1$ is also known (see [4; Theorem 4.3]).

4.1 Partial proofs of main results

Before moving to the proofs of the main results, we give some more descriptions of λ_p . Define

$$\lambda_p^{(m)} = \inf \{ D_p(f) : \mu(|f|^p) = 1, f_i = f_{i \wedge m}, i \in E \}.$$

Let

$$\tilde{D}_p(f) = \sum_{i=1}^m \tilde{\nu}_i |f_i - f_{i-1}|^p, \quad \tilde{\mu}(f) = \sum_{i=1}^m \tilde{\mu}_i |f_i|^p,$$

where $\tilde{\nu}$ and $\tilde{\mu}$ are defined as follows:

$$\tilde{\nu}_i = \nu_i \quad \text{for } i \leq m; \quad \tilde{\mu}_i = \mu_i \quad \text{for } i \leq m-1, \quad \tilde{\mu}_m = \sum_{k=m}^N \mu_k.$$

For $f = f_{\cdot \wedge m}$, we have

$$\tilde{D}_p(f) = D_p(f), \quad \tilde{\mu}(|f|^p) = \mu(|f|^p).$$

So $\lambda_p^{(m)}$ is the first eigenvalue of the local Dirichlet form $(\tilde{D}, \mathcal{D}(\tilde{D}))$ on $E^{(m)} := \{1, 2, \dots, m\}$ with reflecting (Neumann) boundary at $m+1$ and absorbing (Dirichlet) boundary at 0. Furthermore, we have the following fact.

Lemma 4.5 *If $\sum_{i \in E} \mu_i < \infty$, then $\lambda_p = \lambda_p^{[1]} = \lambda_p^{[2]}$. Moreover, $\lambda_p^{(m)} \downarrow \lambda_p^{[2]}$ as $m \rightarrow N$.*

Proof Since each f with $\mu(|f|^p) = \infty$ can be approximated by $f_i^{(m)} = f_{i \wedge m}$ ($m \in E$) with respect to norm $\|\cdot\|^p = D_p(\cdot) + \mu(|\cdot|^p)$, it is clear that $\lambda_p = \lambda_p^{[1]}$.

We now prove that $\lambda_p^{[1]} = \lambda_p^{[2]}$. It is clear that $\lambda_p^{[2]} \geq \lambda_p^{[1]}$ since $\sum_{k \in E} \mu_k < \infty$. For any fixed $\varepsilon > 0$, there exists $f \in L^p(\mu)$ such that

$$D_p(f)/\mu(|f|^p) \leq \lambda_p^{[1]} + \varepsilon.$$

Let $f^{(n)} = f \cdot \chi_n$. Then

$$D_p(f^{(n)}) \rightarrow D_p(f), \quad \mu(|f^{(n)}|^p) \rightarrow \mu(|f|^p) \quad \text{as } n \rightarrow \infty.$$

By definitions of $\lambda_p^{(n)}$ and $\lambda_p^{[2]}$, for large enough $n \in E$, we have

$$\lambda_p^{[2]} \leq \lambda_p^{(n)} \leq \frac{D_p(f^{(n)})}{\mu(|f^{(n)}|^p)} \leq \frac{D_p(f)}{\mu(|f|^p)} + \varepsilon \leq \lambda_p^{[1]} + 2\varepsilon \leq \lambda_p^{[2]} + 2\varepsilon.$$

Hence, $\lambda_p^{[1]} = \lambda_p^{[2]}$ and $\lambda_p^{(n)} \downarrow \lambda_p^{[2]}$.

Proof of Theorem 4.1 In parallel to the ND-case, we also adopt two circle arguments to prove the theorem. For instance, the circle argument below is adopted for the upper estimates:

$$\begin{aligned} \lambda_p &\leq \inf_{f \in \widetilde{\mathcal{F}}_I \cup \widetilde{\mathcal{F}}_E} \sup_{i \in E} II_i(f)^{-1} \\ &\leq \inf_{f \in \widetilde{\mathcal{F}}_E} \sup_{i \in E} II_i(f)^{-1} \\ &= \inf_{f \in \widetilde{\mathcal{F}}_I} \sup_{i \in E} II_i(f)^{-1} \\ &= \inf_{f \in \widetilde{\mathcal{F}}_I} \sup_{i \in E} I_i(f)^{-1} \\ &\leq \inf_{w \in \widetilde{\mathcal{W}}} \sup_{i \in E} \widetilde{R}_i(w) \\ &\leq \lambda_p. \end{aligned}$$

The proofs are similar to the ND-case. Here, we present the proofs of the assertions related to the operator $\widetilde{R}$ in the above circle argument, which are obvious different from that in Section 2 due to the boundary conditions.

(1) We first prove that

$$\inf_{f \in \widetilde{\mathcal{F}}_E} \sup_{i \in E} II_i(f)^{-1} \leq \inf_{w \in \widetilde{\mathcal{W}}} \sup_{i \in E} \widetilde{R}_i(w).$$

For $w \in \widetilde{\mathcal{W}}$, it is easy to check that $\widetilde{R}_i(w) > 0$. Let $u_0 = 0$ and $u_i = u_{i \wedge m} > 0$ for $i \in E$ such that $w_i (= u_{i+1}/u_i) \in \widetilde{\mathcal{W}}$. Then u_i is strictly increasing on $[0, m]$. Put

$$f_i^{p-1} = \begin{cases} \mu_i^{-1} [\nu_i(u_i - u_{i-1})^{p-1} - \nu_{i+1}(u_{i+1} - u_i)^{p-1}], & i \leq m-1, \\ \widetilde{\mu}_m^{-1} \nu_m(u_m - u_{m-1})^{p-1}, & i \geq m. \end{cases}$$

We have $f \in \widetilde{\mathcal{F}}_\Pi$ and

$$\Omega_p u(k) = -\mu_k f_k^{p-1}, \quad k \leq m-1,$$

$$\nu_m(u_m - u_{m-1})^{p-1} = \sum_{j=m}^N \mu_j f_m^{p-1} = \sum_{j=m}^N \mu_j f_j^{p-1}.$$

By a simple reorganization and making summation from 1 to i ($\leq m$) with respect to k , we obtain

$$\sum_{k=1}^i \hat{\nu}_k \left(\sum_{j=k}^N \mu_j f_j^{p-1} \right)^{p^*-1} = u_i - u_0 = u_i, \quad i \leq m.$$

Therefore,

$$\tilde{R}_i(w)^{p^*-1} = \frac{f_i}{u_i} = f_i \left[\sum_{k=1}^i \hat{\nu}_k \left(\sum_{j=k}^N \mu_j f_j^{p-1} \right)^{p^*-1} \right]^{-1} = \Pi_i(f)^{1-p^*}, \quad i \leq m,$$

and then

$$\sup_{i \in E} \tilde{R}_i(w) \geq \max_{i \leq m} \tilde{R}_i(w) \geq \sup_{i \in E} \Pi_i(f)^{-1} \geq \inf_{f \in \widetilde{\mathcal{F}}_\Pi} \sup_{i \in \text{supp}(f)} \Pi_i(f)^{-1}.$$

(2) We prove that

$$\inf_{w \in \widetilde{\mathcal{W}}} \sup_{i \in E} \tilde{R}_i(w) \leq \lambda_p.$$

Let g denote the solution to the equation

$$-\Omega_p g(i) = \lambda_p^{(m)} \tilde{\mu}_i |g_i|^{p-2} g_i, \quad g_0 = 0, \quad g_{m+1} := g_m, \quad i \in E^{(m)} := \{0, 1, \dots, m\}.$$

Without loss of generality, assume $g_1 > 0$. Then g is strictly increasing (by Proposition 4.6 below) and

$$-\nu_{i+1}(g_{i+1} - g_i)^{p-1} + \nu_i(g_i - g_{i-1})^{p-1} = \lambda_p^{(m)} \mu_i g_i^{p-1}, \quad i \leq m-1.$$

$$\nu_m(g_m - g_{m-1})^{p-1} = \lambda_p^{(m)} \tilde{\mu}_m g_m^{p-1}.$$

That is,

$$\nu_i \left(1 - \frac{g_{i-1}}{g_i} \right)^{p-1} - \nu_{i+1} \left(\frac{g_{i+1}}{g_i} - 1 \right)^{p-1} = \lambda_p^{(m)} \mu_i, \quad i \leq m-1;$$

$$\nu_m \left(1 - \frac{g_{m-1}}{g_m} \right)^{p-1} = \lambda_p^{(m)} \tilde{\mu}_m.$$

Let $\bar{w}_i = g_{i+1}/g_i$ for $i \leq m-1$ and $\bar{w}_i = 1$ for $i \geq m$. Then $\bar{w} \in \widetilde{\mathcal{W}}$ and $\tilde{R}_i(\bar{w}) = \lambda_p^{(m)}$ for every $i \leq m$. Therefore,

$$\begin{aligned} \lambda_p^{(m)} &= \max_{0 \leq i \leq m} \tilde{R}_i(\bar{w}) \\ &\geq \inf_{w \in \widetilde{\mathcal{W}}: w_i = w_{i \wedge m}} \sup_{0 \leq i \leq m} \tilde{R}_i(w) \\ &\geq \inf_{w \in \widetilde{\mathcal{W}}: w_i = w_{i \wedge n} \text{ for some } n \in E} \sup_{i \in E} \tilde{R}_i(w) \\ &\geq \inf_{w \in \widetilde{\mathcal{W}}} \sup_{i \in E} \tilde{R}_i(w). \end{aligned}$$

We obtain the required assertion by letting $m \rightarrow N$. $\square$

Noticing the difference between the ND- and the DN-cases, one may finish the proofs of other theorems in this section without much difficulties following Section 2 or [4; Section 4]. So we ignore the details here but present one proposition below, which is essential to our study and is used to verify the last inequalities related to R or $\tilde{R}$ in the two circle arguments. The proposition, whose proof is independent of the other assertions in this paper, provides the basis for imitating the eigenfunction to construct the corresponding test functions of the operators.

Proposition 4.6 *Assume that g is a nontrivial solution to p -Laplacian problem with DN-boundary conditions. Then g is monotone. Moreover, g is increasing provided $g_1 > 0$.*

Proof The proof is parallel to that in [3; Proposition 3.4]. We give the skeleton of the proof. The proof is quite easy in the case of $\lambda_p = 0$. For the case that $\lambda_p > 0$, suppose that there exists $n \in E$ such that $g_0 < g_1 < \cdots < g_n \geq g_{n+1}$. Then define $\bar{g}_i = g_{i \wedge n}$. By a simple calculation and (14), we obtain

$$\lambda_p \leq \frac{D_p(\bar{g})}{\mu(|\bar{g}|^p)} = \frac{\lambda_p \sum_{k=1}^{n-1} \mu_k |g_k|^p + \nu_n |g_n - g_{n-1}|^{p-2} (g_n - g_{n-1}) g_n}{\sum_{k=1}^{n-1} \mu_k |g_k|^p + |g_n|^p \sum_{k=n}^N \mu_k} < \lambda_p.$$

In the last inequality, we have used the following fact:

$$\nu_n |g_n - g_{n-1}|^{p-2} (g_n - g_{n-1}) g_n \leq -g(n) \Omega_p g(n) = \lambda_p \mu_n |g_n|^p < \lambda_p |g_n|^p \sum_{k=n}^N \mu_k$$

for $n < N$. Therefore, there is a contradiction and so the required assertion holds. $\square$

Acknowledgements The authors thank Professors Yonghua Mao and Yutao Ma for their helpful comments and suggestions. The work was supported in part by NSFC (Grant No.11131003), SRFDP (Grant No. 20100003110005), the “985” project from the Ministry of Education in

China, and the Fundamental Research Funds for the Central Universities.

References

1. Chen M.F. Variational formulas and approximation theorems for the first eigenvalue in dimension one. *Sci. in China (A)* 2001, 31(1): 28-36 (Chinese Edition); 44(4): 409-418 (English Edition).
2. Chen M.F. Explicit bounds of the first eigenvalue. *Sci. in China (A)*, 2000, 43(10): 1051-1059.
3. Chen M.F. *Eigenvalues, Inequalities, and Ergodic theory*. New York, Springer, 2005.
4. Chen M.F. Speed of stability for birth-death process. *Front. Math. China*, 2010, 5(3): 379-516.
5. Chen M.F. Bilateral Hardy-type inequalities. *Acta Math. Sinica*, 2013, 29 (1): 1-32.
6. Chen M.F., Wang L.D., Zhang Y.H. Mixed principal eigenvalues in dimension one. *Front. Math. China*, 2013, 8(2): 317-343.
7. Jin H.Y, Mao Y.H. Estimation of the optimal constants in the L_p -Poincaré inequalities on the half Line. *Acta Math. Sinica (Chinese Series)*, 2012, 55(1): 169-178.
8. Kufner A., Maligranda L, Persson L.E. *The Hardy Inequality: About its History and Some Related Results*. Pilson, 2007.
9. Kufner A, Persson L.E. *Weighted Inequalities of Hardy Type*. World Sci. 2003.
10. Mao Y.H. Nash inequalities for markov processes in dimension one. *Acta Math. Sinica*, 2002, 18(1): 147-156.
11. Opic B., Kufner A. *Hardy Type Inequalities*. Longman Scientific and Technical, 1990.

School of Mathematical Sciences, Beijing Normal University;

Laboratory of Mathematics and Complex Systems (Beijing Normal University),

Ministry of Education;

Beijing 100875, China

Emails: M.F. Chen mfchen@bnu.edu.cn Y.H. Zhang zhangyh@bnu.edu.cn

Mixed Eigenvalues of p -Laplacian

Mu-Fa CHEN¹, Ling-Di WANG^{1,2}, Yu-Hui ZHANG¹

¹ Beijing Normal University, Beijing 100875, China

² Henan University, Kaifeng, Henan 475004, China

© Higher Education Press and Springer-Verlag Berlin Heidelberg 2013

Abstract The mixed principal eigenvalue of p -Laplacian (equivalently, the optimal constant of weighted Hardy inequality in L^p space) is studied in this paper. Several variational formulas for the eigenvalue are presented. As applications of the formulas, a criterion for the positivity of the eigenvalue is obtained. Furthermore, an approximating procedure and some explicit estimates are presented case by case. An example is included to illustrate the power of the results of the paper.

Keywords p -Laplacian, Hardy inequality in L^p space, mixed boundaries, explicit estimates, eigenvalue, approximating procedure

MSC 60J60, 34L15

1 Introduction

As a natural extension of Laplacian from linear to nonlinear, p -Laplacian plays a typical role in mathematics, especially in nonlinear analysis. Refer to [1, 10] for recent progresses on this subject. Motivated by the study on stability speed, we come to this topic, see [2, 3] and references therein. The present paper is a continuation of [5] in which the estimates of the mixed principal eigenvalue for discrete p -Laplacian were carefully studied. This paper deals with the same problem but for continuous p -Laplacian, its principal eigenvalue is equivalent to the optimal constant in the weighted Hardy inequality. Even though the discrete case is often harder than the continuous one, the latter has its own difficulty. For instance, the existence of the eigenfunction is rather hard in the nonlinear context, but it is not a problem in the discrete situation. Similar to the case of $p = 2$ ([3, 4]), there are four types of boundaries: Neumann (denoted by code “N”) or Dirichlet (denoted by code “D”) boundary at the left- or right-endpoint of the half line $[0, D]$. In [7], Jin and Mao studied a class of weighted Hardy inequality and presented two variational formulas in the DN-case. Here, we study ND-case carefully and add some results to [7]. The DD- and NN-cases

Received August 13, 2013; accepted March 3, 2014

Corresponding author: Ling-Di WANG, E-mail:wanglingdi@henu.edu.cn

will be handled elsewhere. Comparing with our previous study, here the general weights are allowed.

The paper is organized as follows. In the next section, restricted in the ND-case, we introduce the main results: variational formulas and the basic estimates for the optimal constant (cf. [8, 9]). As an application, we improve the basic estimates step by step through an approximating procedure. To illustrate the power of the results, an example is included. The sketched proofs of the results in Section 2 are presented in Section 3. For another mixed case: DN-case studied in [7], some complementary are presented in Section 4.

2 ND-case

Let μ, ν be two positive Borel measures on $[0, D]$, $D \leq \infty$ (replace $[0, D]$ by $[0, D)$ if $D = \infty$), $d\mu = u(x)dx$ and $d\nu = v(x)dx$. Next, let

$$L_p f = (v|f'|^{p-2}f')', \quad p > 1.$$

Then the eigenvalue problem with ND-boundary conditions reads:

$$\begin{cases} \text{Eigenequation :} & L_p g(x) = -\lambda u(x)|g|^{p-2}g(x); \\ \text{ND-boundaries :} & g'(0) = 0, \quad g(D) = 0 \text{ if } D < \infty. \end{cases} \quad (1)$$

If (λ, g) is a solution to the eigenvalue problem above, $g \neq 0$, then we call λ an 'eigenvalue' and g is an 'eigenfunction' of λ . When $p = 2$, the operator L_p defined above returns to the diffusion operator defined in [4]: $u^{-1}(vf')'$, where $u(x)dx$ is the invariant measure of the diffusion process and v is a Borel measurable function related to its recurrence criterion. For $\alpha \leq \beta$, define

$$\mathcal{C}[\alpha, \beta] = \{f : f \text{ is continuous on } [\alpha, \beta]\},$$

$$\mathcal{C}^k(\alpha, \beta) = \{f : f \text{ has continuous derivatives of order } k \text{ on } (\alpha, \beta)\}, \quad k \geq 1,$$

and

$$\mu_{\alpha, \beta}(f) = \int_{\alpha}^{\beta} f d\mu, \quad D_p^{\alpha, \beta}(f) = \int_{\alpha}^{\beta} |f'|^p d\nu.$$

Similarly, one may define $\mathcal{C}(\alpha, \beta)$. In this section, we study the first eigenvalue (the minimal one), denoted by λ_p , described by the following classical variational formula:

$$\lambda_p = \inf \{D_p(f) : f \in \mathcal{C}_K[0, D], \mu(|f|^p) = 1, f(D) = 0 \text{ if } D < \infty\}, \quad (2)$$

where $\mu(f) = \mu_{0, D}(f)$, $D_p(f) = D_p^{0, D}(f)$ and

$$\mathcal{C}_K[\alpha, \beta] = \{f \in \mathcal{C}[\alpha, \beta] : v^{p^*-1}f' \in \mathcal{C}(\alpha, \beta) \text{ and } f \text{ has compact support}\},$$

with p^* the conjugate number of p (i.e., $p^{-1} + p^{*-1} = 1$). When $p = 2$, it reduces to the linear case studied in [4]. Thus, the aim of the paper is extending the results in linear case ($p = 2$) to nonlinear one. Set

$$\mathcal{A}[\alpha, \beta] = \{f : f \text{ is absolutely continuous on } [\alpha, \beta]\}.$$

As will be proved soon (see Lemmas 3.3 and 3.4), we can rewrite λ_p as

$$\tilde{\lambda}_{*,p} := \inf \{D_p(f) : \mu(|f|^p) = 1, f \in \mathcal{A}[0, D], f(D) = 0\}. \quad (3)$$

By making inner product with g on both sides of eigenequation (1) with respect to the Lebesgue measure over (α, β) , we obtain

$$\lambda \mu_{\alpha,\beta}(|g|^p) = D_p^{\alpha,\beta}(g) - (v|g'|^{p-2}g'g)|_{\alpha}^{\beta}.$$

Moreover, since $g'(0) = 0$, we have

$$\lambda \mu(|g|^p) = D_p(g) - (v|g'|^{p-2}g'g)(D),$$

where, throughout this paper, $f(D) := \lim_{x \rightarrow D} f(x)$ provided $D = \infty$. Hence, with

$$\mathcal{D}(D_p) = \{f : f \in \mathcal{A}[0, D], D_p(f) < \infty\},$$

$A := \lambda_p^{-1}$ is the optimal constant of the following *weighted Hardy inequality*:

$$\text{Hardy inequality : } \mu(|f|^p) \leq AD_p(f), \quad f \in \mathcal{D}(D_p);$$

$$\text{Boundary condition : } f(D) = 0.$$

Note that the boundary condition “ $f'(0) = 0$ ” is unnecessary in the inequality.

Throughout this paper, we concentrate on $p \in (1, \infty)$ since the degenerated cases that either $p = 1$ or ∞ are often easier to handle (cf. [11; Lemmas 5.4 and 5.6 on pages 49 and 56, respectively]).

Main notation and results

For $p > 1$, let p^* be its conjugate number. Define $\hat{v}(x) = v^{1-p^*}(x)$ and $\hat{\nu}(dx) = \hat{v}(x)dx$. We use the following hypothesis throughout the paper:

$$u, \hat{v} \text{ are locally integrable with respect to the Lebesgue measure on } [0, D],$$

without mentioned time by time.

Our main operators are defined as follows.

$$I(f)(x) = -\frac{1}{(vf'|f|^{p-2})(x)} \int_0^x f^{p-1} d\mu \quad (\text{single integral form}),$$

$$II(f)(x) = \frac{1}{f^{p-1}(x)} \left[\int_{(x,D) \cap \text{supp}(f)} \hat{v}(s) \left(\int_0^s f^{p-1} d\mu \right)^{p^*-1} ds \right]^{p-1} \quad (\text{double integral form}),$$

$$R(h)(x) = u(x)^{-1} \left[-|h|^{p-2} (v'h + (p-1)(h^2 + h')v) \right](x) \quad (\text{differential form}).$$

These operators have domains, respectively, as follows.

$$\begin{aligned}\mathcal{F}_I &= \{f \in \mathcal{C}[0, D] : v^{p^*-1}f' \in \mathcal{C}(0, D), f|_{(0,D)} > 0, f'|_{(0,D)} < 0\}, \\ \mathcal{F}_{II} &= \{f : f \in \mathcal{C}[0, D], f|_{(0,D)} > 0\}, \\ \mathcal{H} &= \{h : h \in \mathcal{C}^1(0, D) \cap \mathcal{C}[0, D], h(0) = 0, h|_{(0,D)} < 0 \text{ if } \hat{\nu}(0, D) < \infty, \\ &\quad \text{and } h|_{(0,D)} \leq 0 \text{ if } \hat{\nu}(0, D) = \infty\},\end{aligned}$$

where $\nu(\alpha, \beta) = \int_\alpha^\beta d\nu$ for a measure ν . To avoid the non-integrability problem, some modifications of these sets are needed for studying the upper estimates.

$$\begin{aligned}\widetilde{\mathcal{F}}_I &= \{f \in \mathcal{C}[x_0, x_1] : v^{p^*-1}f' \in \mathcal{C}(x_0, x_1), f'|_{(x_0, x_1)} < 0 \text{ for some} \\ &\quad x_0, x_1 \in (0, D) \text{ with } x_0 < x_1, \text{ and } f = f(\cdot \vee x_0)\mathbb{1}_{[0, x_1]}\}, \\ \widetilde{\mathcal{F}}_{II} &= \{f : f = f\mathbb{1}_{[0, x_0]} \text{ for some } x_0 \in (0, D) \text{ and } f \in \mathcal{C}[0, x_0]\}, \\ \widetilde{\mathcal{H}} &= \left\{h : \exists x_0 \in (0, D) \text{ such that } h \in \mathcal{C}[0, x_0] \cap \mathcal{C}^1(0, x_0), h|_{(0, x_0)} < 0, \right. \\ &\quad \left. h|_{[x_0, D]} = 0, h(0) = 0, \text{ and } \sup_{(0, x_0)} (v'h + (p-1)(h^2 + h')v) < 0\right\},\end{aligned}$$

In Theorem 2.1 below, for each $f \in \mathcal{F}_I$, $\inf_{x \in (0, D)} I(f)(x)^{-1}$ produces a lower bound of λ_p . So the part having “sup inf” in each of the formulas is used for the lower estimates of λ_p . Dually, the part having “inf sup” is used for the upper estimates. These formulas deduce the basic estimates in Theorem 2.3 and the approximating procedure in Theorem 2.4.

Theorem 2.1 (Variational formulas) *For $p > 1$, we have*

(1) *single integral forms:*

$$\inf_{f \in \widetilde{\mathcal{F}}_I} \sup_{x \in (0, D)} I(f)(x)^{-1} = \lambda_p = \sup_{f \in \mathcal{F}_I} \inf_{x \in (0, D)} I(f)(x)^{-1},$$

(2) *double integral forms:*

$$\inf_{f \in \widetilde{\mathcal{F}}_{II}} \sup_{x \in \text{supp}(f)} II(f)(x)^{-1} = \lambda_p = \sup_{f \in \mathcal{F}_{II}} \inf_{x \in (0, D)} II(f)(x)^{-1}.$$

Moreover, if u and v' are continuous, then we have additionally

(3) *differential forms:*

$$\inf_{h \in \widetilde{\mathcal{H}}} \sup_{x \in (0, D)} R(h)(x) = \lambda_p = \sup_{h \in \mathcal{H}} \inf_{x \in (0, D)} R(h)(x).$$

Furthermore, the supremum on the right-hand side of the above three formulas can be attained.

The following proposition adds some additional sets of functions for operators I and II . It then provides alternative descriptions of the lower and upper estimates of λ_p .

Proposition 2.2 *For $p > 1$, we have*

$$\lambda_p = \sup_{f \in \mathcal{F}_I} \inf_{x \in (0, D)} II(f)(x)^{-1};$$

$$\lambda_p = \inf_{f \in \widetilde{\mathcal{F}}_{II} \cup \widetilde{\mathcal{F}}'_{II}} \sup_{x \in \text{supp}(f)} II(f)(x)^{-1} = \inf_{f \in \widetilde{\mathcal{F}}_I} \sup_{x \in \text{supp}(f)} II(f)(x)^{-1} = \inf_{f \in \widetilde{\mathcal{F}}'_I} \sup_{x \in (0, D)} I(f)(x)^{-1},$$

where

$$\widetilde{\mathcal{F}}'_{II} = \{f : f \in \mathcal{C}[0, D] \text{ and } fII(f) \in L^p(\mu)\},$$

$$\begin{aligned} \widetilde{\mathcal{F}}'_I = \{f : \exists x_0 \in (0, D), f = f\mathbb{1}_{[0, x_0)} \in \mathcal{C}[0, x_0], f'|_{(0, x_0)} < 0, \\ \text{and } v^{p^*-1}f' \in \mathcal{C}(0, x_0)\}. \end{aligned}$$

Define $k(p) = pp^{*p-1}$ for $p > 1$ and

$$\sigma_p = \sup_{x \in (0, D)} \mu(0, x) \hat{v}(x, D)^{p-1}.$$

As applications of the variational formulas in Theorem 2.1 (1), we have the following basic estimates known in [11].

Theorem 2.3 (Criterion and basic estimates) *For $p > 1$, the eigenvalue $\lambda_p > 0$ if and only if $\sigma_p < \infty$. Moreover, the following basic estimates hold:*

$$(k(p)\sigma_p)^{-1} \leq \lambda_p \leq \sigma_p^{-1},$$

In particular, we have $\lambda_p = 0$ if $\hat{v}(0, D) = \infty$ and $\lambda_p > 0$ if

$$\int_0^D \mu(0, s)^{p^*-1} \hat{v}(s) ds < \infty.$$

The approximating procedure below is an application of variational formulas in Theorem 2.1 (2). The main idea is an iteration, its first step produces Corollary 2.5 below. Noticing that λ_p is trivial once $\sigma_p = \infty$, we may assume that $\sigma_p < \infty$ for further study on the estimates of λ_p .

Theorem 2.4 (Approximating procedure) *Assume that $\sigma_p < \infty$.*

- (1) *Let $f_1 = \hat{v}(\cdot, D)^{1/p^*}$, $f_{n+1} = f_n II(f_n)^{p^*-1}$ and $\delta_n = \sup_{x \in (0, D)} II(f_n)(x)$ for $n \geq 1$. Then δ_n is decreasing and*

$$\lambda_p \geq \delta_n^{-1} \geq (k(p)\sigma_p)^{-1}.$$

(2) For fixed $x_0, x_1 \in (0, D)$ with $x_0 < x_1$, define

$$f_1^{x_0, x_1} = \hat{\nu}(\cdot \vee x_0, x_1) \mathbb{1}_{[0, x_1)}, \quad f_n^{x_0, x_1} = f_{n-1}^{x_0, x_1} H(f_{n-1}^{x_0, x_1})^{p^*-1} \mathbb{1}_{[0, x_1)},$$

and

$$\delta'_n = \sup_{x_0, x_1: x_0 < x_1} \inf_{x < x_1} H(f_n^{x_0, x_1})(x), \quad \text{for } n \geq 1.$$

Then δ'_n is increasing and

$$\sigma_p^{-1} \geq \delta_n'^{-1} \geq \lambda_p.$$

Next, define

$$\bar{\delta}_n = \sup_{x_0 < x_1} \frac{\|f_n^{x_0, x_1}\|_p^p}{D_p(f_n^{x_0, x_1})}, \quad n \geq 1.$$

Then $\bar{\delta}_n^{-1} \geq \lambda_p$ and $\bar{\delta}_{n+1} \geq \delta'_n$ for $n \geq 1$.

The following Corollary 2.5 can be obtained directly from Theorem 2.4. It provides us some improved and explicit estimates of the eigenvalue (see Example 2.6 below).

Corollary 2.5 (Improved estimates) Assume that $\sigma_p < \infty$. Then

$$\sigma_p^{-1} \geq \delta_1'^{-1} \geq \lambda_p \geq \delta_1^{-1} \geq (k(p)\sigma_p)^{-1},$$

where

$$\delta_1 = \sup_{x \in (0, D)} \left[\frac{1}{\hat{\nu}(x, D)^{1/p^*}} \int_x^D \hat{\nu}(s) \left(\int_0^s \hat{\nu}(t, D)^{(p-1)/p^*} \mu(dt) \right)^{p^*-1} ds \right]^{p-1};$$

$$\delta_1' = \sup_{x \in (0, D)} \frac{1}{\hat{\nu}(x, D)^{p-1}} \left[\int_x^D \hat{\nu}(s) \left(\int_0^s \hat{\nu}(t \vee x, D)^{p-1} \mu(dt) \right)^{p^*-1} ds \right]^{p-1}.$$

Moreover,

$$\bar{\delta}_1 = \sup_{x \in (0, D)} \left[\mu(0, x) \hat{\nu}(x, D)^{p-1} + \frac{1}{\hat{\nu}(x, D)} \int_x^D \hat{\nu}(t, D)^p \mu(dt) \right] \in [\sigma_p, p\sigma_p],$$

and $\bar{\delta}_1 \leq \delta_1'$ for $1 < p \leq 2$, $\bar{\delta}_1 \geq \delta_1'$ for $p \geq 2$.

When $p = 2$, the assertion that $\bar{\delta}_1 = \delta_1'$ was proved in [4; Theorem 3]. To illustrate the results above, we present an example as follows.

Example 2.6 Let $d\mu = d\nu = dx$ on $(0, 1)$. In the ND-case, the eigenvalue λ_p is

$$\lambda_p^{1/p} = \frac{\pi(p-1)^{1/p}}{p} \sin^{-1} \frac{\pi}{p}. \quad (4)$$

For the basic estimates, we have

$$\sigma_p^{1/p} = \left(\frac{1}{p}\right)^{1/p} \left(\frac{1}{p^*}\right)^{1/p^*}.$$

Furthermore, we have

$$\bar{\delta}_1^{1/p} = p^{1/p-2} (p^2 - 1)^{1-1/p},$$

$$\delta_1^{1/p} = \frac{1}{(p + 1/p - 1)^{1/p}} \left\{ \sup_{x \in (0,1)} \frac{1}{(1-x)^{1/p^*}} \int_0^{1-x} (1 - z^{p+1/p-1})^{p^*-1} dz \right\}^{1/p^*}.$$

The exact value $\lambda_p^{1/p}$ and its basic estimates are shown in Figure 1. Then, the

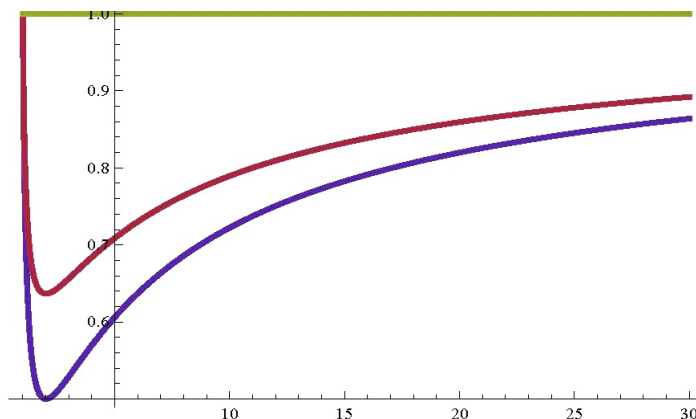


Figure 1 The middle curve is the exact value of $\lambda_p^{1/p}$. The top straight line and the bottom curve are the basic estimates of $\lambda_p^{1/p}$.

Figure 1 The middle curve is the exact value of $\lambda_p^{1/p}$. The top straight line and the bottom curve are the basic estimates of $\lambda_p^{1/p}$.

improved upper bound $\delta_1^{1/p}$ and lower one $\bar{\delta}_1^{1/p}$ are added to Figure 1, as shown in Figure 2. It is quite surprising and unexpected that both of $\delta_1^{1/p}$ and $\bar{\delta}_1^{1/p}$ are almost overlapped with the exact value $\lambda_p^{1/p}$ except in a small neighborhood of $p = 2$, where $\delta_1^{1/p}$ is a little bigger and $\bar{\delta}_1^{1/p}$ is a little smaller than $\lambda_p^{1/p}$. Here $\delta_1'^{1/p}$ is ignored since it improves $\bar{\delta}_1^{1/p}$ only a little bit for $p \in (1, 2)$.

3 Proofs of the main results

Some preparations for the proofs are collected in Subsection 3.1. They may not be used completely in the proofs but are helpful to understand the idea in this paper and may be useful in other cases. The proofs of the main results are presented in Subsection 3.2. For simplicity, we let $\uparrow$ (resp. $\Uparrow$, $\downarrow$, $\Downarrow$) denote

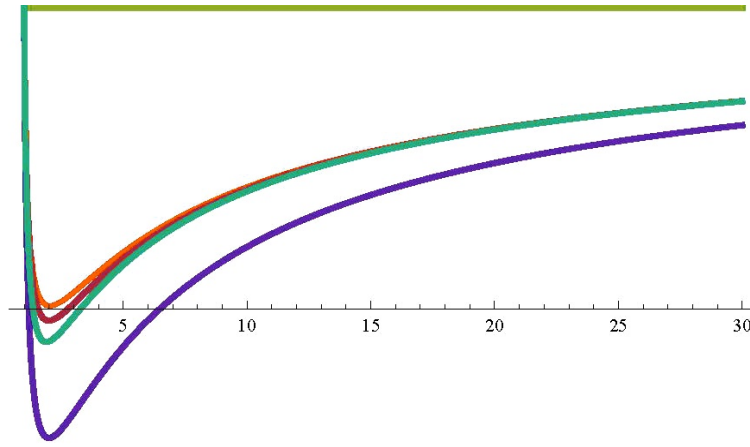


Figure 2 The improved bounds $\delta_1^{1/p}$ and $\bar{\delta}_1^{1/p}$ are added to Figure 1.

Figure 2 The improved bounds $\delta_1^{1/p}$ and $\bar{\delta}_1^{1/p}$ are added to Figure 1.

increasing (resp. strictly increasing, decreasing, strictly decreasing) throughout this paper.

3.1 Preparations

The next lemma is taken from [1; Theorem 1.1 on page 170](see [13] for its original idea). Combining with the following Remark 3.2, Lemmas 3.3 and 3.4, it guarantees the existence of the solution (λ_p, g) to the eigenvalue problem.

Lemma 3.1 (Existence and Uniqueness)

- (1) Suppose that u and v are locally integrable on $[0, D] \subseteq \mathbb{R}$ (or $[0, D) \subseteq \mathbb{R}$ provided $D = \infty$) and $v > 0$. Given constants A and B , for each fixed λ , there is uniquely a solution g such that $g(0) = A$, $g'(0) = B$ and the eigenequation (1) holds almost everywhere. Moreover, $v^{p^*-1}g'$ is absolutely continuous.
- (2) Suppose additionally u and v are continuous. Then $g \in \mathcal{C}^2[0, D]$ and the eigenequation holds everywhere on $[0, D]$.

If the eigenequation (1) holds (almost) everywhere for (λ_p, g) , then g is called an (a.e.) eigenfunction of λ_p .

Remark 3.2 (1) One may also refer to [6; Lemma 2.1] for the existence of solution to eigenvalue problem with ND boundary conditions provided $D < \infty$. When $D = \infty$, the Dirichlet boundary at D means $g(D) = 0$, which is proved by Proposition 3.7 below.

(2) By [11; Theorem 4.1, Theorem 4.7], we see that the eigenequation in (1) has solutions if and only if the following equation has solutions:

$$(|g'|^{p-2}g')'(x) = -\lambda\tilde{u}(x)|g|^{p-2}g(x)$$

where $\tilde{u}$ is related to v and u in the eigenequation. Hence the weight function v in the eigenequation is not a sensitive or key quantity to the existence of solution to the eigenequation and can be seen as a constant.

Define $\mathcal{A}_K[0, D] = \{f : f \in \mathcal{A}[0, D], f \text{ has compact support}\}$ and

$$\lambda_{*,p} = \inf\{D_p(f) : f \in \mathcal{A}_K[0, D], \|f\|_p = 1, \text{ and } f(D) = 0 \text{ if } D < \infty\}, \quad (5)$$

$$\tilde{\lambda}_p = \inf\{D_p(f) : v^{p^*-1}f' \in \mathcal{C}(0, D), f \in \mathcal{C}[0, D], \|f\|_p = 1, f(D) = 0\}, \quad (6)$$

where $\|\cdot\|_p$ means the norm in $L^p(\mu)$ space. The following quantities are also useful for us. Set $\alpha \in (0, D)$ and define

$$\lambda_{*,p}^{(0,\alpha)} = \inf\{D_p(f) : \mu(|f|^p) = 1, f \in \mathcal{A}[0, \alpha] \text{ and } f|_{[\alpha,D]} = 0\}.$$

$$\lambda_p^{(0,\alpha)} = \inf\{D_p(f) : f \in \mathcal{C}[0, \alpha], v^{p^*-1}f' \in \mathcal{C}(0, \alpha), \mu(|f|^p) = 1, f|_{[\alpha,D]} = 0\}.$$

The following three Lemmas describe in a refined way the first eigenvalue and lead to, step by step, the conclusion that

$$\tilde{\lambda}_p = \lambda_p = \lambda_{*,p} = \tilde{\lambda}_{*,p}.$$

Lemma 3.3 *We have $\lambda_p = \lambda_{*,p}$.*

Proof It is obvious that $\lambda_p \geq \lambda_{*,p}$. Next, let g be the a.e. eigenfunction of $\lambda_{*,p}$. Then $g \in \mathcal{C}[0, D]$ and $v^{p^*-1}g' \in \mathcal{C}(0, D)$ by Lemma 3.1. Since $L_p g = -\lambda_{*,p}|g|^{p-2}g$, by the arguments after formula (3), we have

$$-(vg|g'|^{p-2}g')\big|_0^D + D_p(g) = \lambda_{*,p}\|g\|_p^p.$$

Since $g'(0) = 0$ and $(gg')(D) \leq 0$, we have $\lambda_{*,p} \geq D_p(g)/\|g\|_p^p$. Because $g \in \mathcal{C}_K[0, D]$, it is clear that $D_p(g)/\|g\|_p^p \geq \lambda_p$. We have thus obtained that

$$\lambda_p \leq \lambda_{*,p} \leq \lambda_p,$$

and so $\lambda_p = \lambda_{*,p}$. There is a small gap in the proof above since in the case of $D = \infty$, the a.e. eigenfunction g may not belong to $L^p(\mu)$ and we have not yet proved that $(gg')(D) \leq 0$. However, one may avoid this by a standard approximating procedure, using $[0, \alpha_n]$ instead of $[0, D]$ with $\alpha_n \uparrow D$ provided $D = \infty$:

$$\lim_{n \rightarrow \infty} \lambda_p^{(0,\alpha_n)} = \lim_{n \rightarrow \infty} \inf\{D_p(f) : \mu(|f|^p) = 1, f \in \mathcal{C}[0, \alpha_n], v^{p^*-1}f' \in \mathcal{C}(0, \alpha_n), f|_{[\alpha_n,D]} = 0\} = \lambda_p.$$

Similarly, $\lambda_{*,p}^{(0,\alpha_n)} \rightarrow \lambda_{*,p}$ as $n \rightarrow \infty$. $\square$

Lemma 3.4 *For $\tilde{\lambda}_{*,p}$ defined in (3), we have $\tilde{\lambda}_{*,p} = \lambda_{*,p}$. Furthermore, $\tilde{\lambda}_p = \lambda_p = \lambda_{*,p} = \tilde{\lambda}_{*,p}$.*

Proof On one hand, by definition, if $\beta_{n+1} \geq \beta_n$, then $\lambda_{*,p}^{(0,\beta_n)} \geq \lambda_{*,p}^{(0,\beta_{n+1})}$. We have thus obtained

$$\lim_{n \rightarrow \infty} \lambda_{*,p}^{(0,\beta_n)} \geq \lambda_{*,p}^{(0,D)} = \tilde{\lambda}_{*,p}.$$

On the other hand, by definition of $\tilde{\lambda}_{*,p}$, for any fixed $\varepsilon > 0$, there exists f satisfying $\|f\|_p = 1$, $f(D) = 0$, and $D_p(f) \leq \tilde{\lambda}_{*,p} + \varepsilon$. Let $\beta_n \uparrow D$ and $f_n = (f - f(\beta_n))\mathbb{1}_{[0,\beta_n]}$. Then $D_p(f_n) \uparrow D_p(f)$ as $n \uparrow \infty$. Choose subsequence $\{n_m\}_{m \geq 1}$ if necessary such that

$$\overline{\lim}_{n \rightarrow \infty} \frac{D_p(f_n)}{\|f_n\|_p^p} = \lim_{m \rightarrow \infty} \frac{D_p(f_{n_m})}{\|f_{n_m}\|_p^p}.$$

By Fatou's lemma and the fact that $f(D) = 0$, we have

$$\underline{\lim}_{m \rightarrow \infty} \|f_{n_m}\|_p^p \geq \left\| \underline{\lim}_{m \rightarrow \infty} f_{n_m} \right\|_p^p = \|f\|_p^p = 1.$$

Therefore, we obtain

$$\begin{aligned} \overline{\lim}_{n \rightarrow \infty} \lambda_{*,p}^{(0,\beta_n)} &\leq \overline{\lim}_{n \rightarrow \infty} \frac{D_p(f_n)}{\|f_n\|_p^p} = \lim_{m \rightarrow \infty} \frac{D_p(f_{n_m})}{\|f_{n_m}\|_p^p} \leq \frac{\lim_{m \rightarrow \infty} D_p(f_{n_m})}{\underline{\lim}_{m \rightarrow \infty} \|f_{n_m}\|_p^p} \leq D_p(f) \\ &\leq \tilde{\lambda}_{*,p} + \varepsilon. \end{aligned}$$

Since $\lim_{n \rightarrow \infty} \lambda_{*,p}^{(0,\beta_n)} = \lambda_{*,p}$, we get $\tilde{\lambda}_{*,p} = \lambda_{*,p}$. Moreover,

$$\tilde{\lambda}_p \geq \tilde{\lambda}_{*,p} = \lambda_{*,p} = \lambda_p \geq \tilde{\lambda}_p$$

and the required assertion holds. $\square$

The following lemma, which serves for Lemma 3.6, presents us that $\{\lambda_{*,p}^{(0,\alpha)}\}$ is strictly decreasing with respect to α .

Lemma 3.5 For $\alpha, \beta \in (0, D)$ with $\alpha < \beta$, we have $\lambda_{*,p}^{(0,\alpha)} > \lambda_{*,p}^{(0,\beta)}$. Furthermore, $\lambda_{*,p}^{(0,\beta_n)} \searrow \lambda_{*,p}$ as $\beta_n \uparrow D$.

Proof Let $g (\neq 0)$ be an a.e. eigenfunction of $\lambda_{*,p}^{(0,\alpha)}$. Then $g'(0) = 0$, $g(\alpha) = 0$, and $L_p g = -\lambda_{*,p}^{(0,\alpha)} |g|^{p-2} g$ on $(0, \alpha)$. Moreover,

$$\lambda_{*,p}^{(0,\alpha)} = \frac{D_p^{0,\alpha}(g)}{\|g\|_{L^p(0,\alpha;\mu)}^p}, \quad D_p^{\alpha,\beta}(f) = \int_{\alpha}^{\beta} |f'|^p d\nu$$

(see arguments after formula (3)). By the proof of Lemma 3.3, the proof of the first assertion will be done once we choose a function $\tilde{g} \in \mathcal{A}[0, \beta]$ such that $\tilde{g}'(0) = 0$, $\tilde{g}(\beta) = 0$, and

$$\frac{D_p^{0,\alpha}(g)}{\|g\|_{L^p(0,\alpha;\mu)}^p} > \frac{D_p^{0,\beta}(\tilde{g})}{\|\tilde{g}\|_{L^p(0,\beta;\mu)}^p} \quad \left(\geq \lambda_{*,p}^{(0,\beta)} \right). \quad (7)$$

To do so, without loss of generality, assume that $g|_{(0,\alpha)} > 0$ (see [12; Lemma 2.4]). Then the required assertion follows for

$$\tilde{g}(x) = (g + \varepsilon)\mathbb{1}_{[0,\alpha)}(x) + \frac{\varepsilon(\beta - x)}{(\beta - \alpha)}\mathbb{1}_{[\alpha,\beta]}(x), \quad x \in [0, \beta],$$

once ε is sufficiently small. Actually, by simple calculation, we have

$$D_p^{0,\beta}(\tilde{g}) = D_p^{0,\alpha}(g) + \frac{\varepsilon^p}{(\beta - \alpha)^p}\nu(\alpha, \beta),$$

$$\|\tilde{g}\|_{L^p(0,\beta;\mu)}^p = \|g\|_{L^p(0,\alpha;\mu)}^p + \int_0^\alpha (|g + \varepsilon|^p - |g|^p) d\mu + \int_\alpha^\beta \frac{\varepsilon^p(\beta - x)^p}{(\beta - \alpha)^p} \mu(dx).$$

Since

$$\lambda_{*,p}^{(0,\alpha)} = D_p^{0,\alpha}(g) / \|g\|_{L^p(0,\alpha;\mu)}^p,$$

inequality (7) holds if and only if

$$\frac{\varepsilon^p \nu(\alpha, \beta)}{(\beta - \alpha)^p} < \left(\int_0^\alpha (|g + \varepsilon|^p - |g|^p) d\mu + \frac{\varepsilon^p}{(\beta - \alpha)^p} \int_\alpha^\beta (\beta - x)^p \mu(dx) \right) \lambda_{*,p}^{(0,\alpha)}.$$

It suffices to show that

$$\frac{\varepsilon^{p-1}}{(\beta - \alpha)^p} \nu(\alpha, \beta) < \lambda_{*,p}^{(0,\alpha)} \left(\int_0^\alpha \frac{|g(x) + \varepsilon|^p - |g(x)|^p}{\varepsilon} \mu(dx) \right).$$

By letting $\varepsilon \rightarrow 0$, the right-hand side is equal to

$$\lambda_{*,p}^{(0,\alpha)} \int_0^\alpha p g^{p-1} d\mu,$$

which is positive. So the required inequality is obvious for sufficiently small ε and the first assertion holds. The second assertion was proved at the end of the proofs of Lemma 3.4. $\square$

The following Lemma is about the eigenfunction of λ_p , which is the basis of the test functions used for the corresponding operators.

Lemma 3.6 *Let g be the first eigenfunction of eigenvalue problem (1). Then both g and g' do not change sign. Moreover, if $g > 0$, then $g' < 0$.*

Proof If there exists $\alpha \in (0, D)$ such that $g(\alpha) = 0$, then $\lambda_{*,p}^{(0,\alpha)} \leq \lambda_{*,p}$ by the minimum property of $\lambda_{*,p}^{(0,\alpha)}$. However, by Lemma 3.5, we get $\lambda_{*,p}^{(0,\alpha)} \Downarrow \lambda_{*,p}$ as $\alpha \Uparrow D$. This is a contradiction. So g does not change its sign. Next, consider g' . By [12; Lemma 2.3], if there exists $x \in (0, D)$ such that $g'(x) = 0$, then $\exists x_0 \in (0, x)$ such that $g(x_0) = 0$, which is impossible by the strictly decreasing property of $\lambda_{*,p}^{(0,\alpha)}$ with respect to α . So the assertion holds. $\square$

Before moving on, we introduce a general equation, non-linear ‘Poisson equation’ as follows:

$$L_p g(x) = -u(x)|f|^{p-2}f(x), \quad x \in (0, D). \quad (8)$$

Integration by parts yields that for $x, y \in (0, D)$ with $x < y$,

$$v(x)|g'|^{p-2}g'(x) - v(y)|g'|^{p-2}g'(y) = \int_x^y |f|^{p-2}f d\mu. \quad (9)$$

By replacing f with $\lambda^{p^*-1}g$, it is not hard to understand where the operator I comes from. Moreover, if g is positive and decreasing, $g'(0) = 0$, then

$$g(y) - g(D) = \int_y^D \hat{v}(x) \left(\int_0^x |f|^{p-2}f d\mu \right)^{p^*-1} dx, \quad y \in (0, D). \quad (10)$$

By replacing f with $\lambda^{p^*-1}g$, it is easy to see where the operator II comes from, provided $g(D) = 0$ (which is affirmative by Proposition 3.7 below). Finally, assume that (λ_p, g) is a solution to (1). Then $\lambda_p = -L_p g / (|g|^{p-2}g u)$. Hence, by letting $h = g'/g$, we deduce the operator R from the eigenequation.

3.2 Proof of the main results

Proof of Theorem 2.1 and Proposition 2.2 We adopt the circle arguments below to prove the lower estimates:

$$\begin{aligned} \lambda_p &\geq \tilde{\lambda}_p \geq \sup_{f \in \mathcal{F}_{II}} \inf_{x \in (0, D)} II(f)(x)^{-1} = \sup_{f \in \mathcal{F}_I} \inf_{x \in (0, D)} II(f)(x)^{-1} \\ &= \sup_{f \in \mathcal{F}_I} \inf_{x \in (0, D)} I(f)(x)^{-1} \\ &\geq \sup_{h \in \mathcal{H}} \inf_{x \in (0, D)} R(h)(x) \\ &\geq \lambda_p. \end{aligned}$$

Step 1 Prove that $\lambda_p \geq \tilde{\lambda}_p \geq \sup_{f \in \mathcal{F}_{II}} \inf_{x \in (0, D)} II(f)(x)^{-1}$.

It suffices to show the second inequality. For each fixed $h > 0$ and $g \in \mathcal{C}[0, D]$ with $\|g\|_p = 1$, $g(D) = 0$ and $v^{p^*-1}g' \in \mathcal{C}(0, D)$, we have

$$\begin{aligned} \int_0^D |g|^p d\mu &= \int_0^D \left| \int_x^D g'(t) \left(\frac{v(t)}{h(t)} \right)^{1/p} \left(\frac{h(t)}{v(t)} \right)^{1/p} dt \right|^p \mu(dx) \\ &\leq \int_0^D \int_x^D \frac{v(t)}{h(t)} |g'(t)|^p dt \left[\int_x^D \left(\frac{h(s)}{v(s)} \right)^{p^*-1} ds \right]^{p-1} \mu(dx) \\ &\quad \text{(by Hölder's inequality)} \\ &= \int_0^D \frac{v(t)}{h(t)} |g'(t)|^p dt \int_0^t \left[\int_x^D \left(\frac{h(s)}{v(s)} \right)^{p^*-1} ds \right]^{p-1} \mu(dx) \\ &\quad \text{(by Fubini's Theorem)} \\ &\leq D_p(g) \sup_{t \in (0, D)} H(t), \end{aligned}$$

where

$$H(t) = \frac{1}{h(t)} \int_0^t \left[\int_x^D \left(\frac{h(s)}{v(s)} \right)^{p^*-1} ds \right]^{p-1} \mu(dx).$$

For $f \in \mathcal{F}_H$ with $\sup_{x \in (0,D)} H(f)(x) < \infty$, let

$$h(t) = \int_0^t f^{p-1}(s) u(s) ds.$$

Then $h' = f^{p-1}u$. By Cauchy's mean-value theorem, we have

$$\sup_{x \in (0,D)} H(x) \leq \sup_{x \in (0,D)} H(f)(x).$$

Thus $\lambda_p \geq \inf_{x \in (0,D)} H(f)(x)^{-1}$. The assertion then follows by making the supremum with respect to $f \in \mathcal{F}_H$.

Step 2 Prove that

$$\sup_{f \in \mathcal{F}_H} \inf_{x \in (0,D)} H(f)(x)^{-1} = \sup_{f \in \mathcal{F}_I} \inf_{x \in (0,D)} H(f)(x)^{-1} = \sup_{f \in \mathcal{F}_I} \inf_{x \in (0,D)} I(f)(x)^{-1}.$$

(a) We prove the part ' $\geq$ '. Since $\mathcal{F}_I \subset \mathcal{F}_H$, it suffices to show that

$$\sup_{f \in \mathcal{F}_I} \inf_{x \in (0,D)} H(f)(x)^{-1} \geq \sup_{f \in \mathcal{F}_I} \inf_{x \in (0,D)} I(f)(x)^{-1}$$

for $f \in \mathcal{F}_I$ with $\sup_{x \in (0,D)} I(f) < \infty$. Since $f(D) \geq 0$, by replacing f in the denominator of $H(f)$ with $-\int_x^D f'(s) ds$ and using Cauchy's mean-value theorem, we have

$$\sup_{x \in (0,D)} H(f)(x) \leq \sup_{x \in (0,D)} I(f)(x) < \infty.$$

So the assertion holds by making the supremum with respect to $f \in \mathcal{F}_I$.

(b) To prove the equality, it suffices to show that

$$\sup_{f \in \mathcal{F}_I} \inf_{x \in (0,D)} I(f)(x)^{-1} \geq \sup_{f \in \mathcal{F}_H} \inf_{x \in (0,D)} H(f)(x)^{-1}.$$

For $f \in \mathcal{F}_H$, without loss of generality, assume that $\inf_{x \in (0,D)} H(f)(x)^{-1} > 0$. Let $g = f[H(f)]^{p^*-1}$. Then $g \in \mathcal{F}_I$. Moreover,

$$v(x)(-g'(x))^{p-1} = \int_0^x f^{p-1} d\mu \geq \int_0^x g^{p-1} d\mu \inf_{t \in (0,x)} \frac{f^{p-1}(t)}{g^{p-1}(t)},$$

i.e.,

$$I(g)(x)^{-1} \geq \inf_{x \in (0,D)} H(f)(x)^{-1}.$$

Hence,

$$\sup_{f \in \mathcal{F}_I} \inf_{x \in (0, D)} I(f)(x)^{-1} \geq \inf_{x \in (0, D)} I(g)(x)^{-1} \geq \inf_{x \in (0, D)} II(f)(x)^{-1}$$

and the assertion holds since $f \in \mathcal{F}_{II}$ is arbitrary.

Then there is another method to prove the equality: prove that

$$\sup_{f \in \mathcal{F}_I} \inf_{x \in (0, D)} I(f)(x)^{-1} \geq \lambda_p.$$

Let g be an a.e. eigenfunction corresponding to λ_p . Then g is positive and strictly decreasing. It is easy to check that $g \in \mathcal{F}_I$. By (9), we have

$$\lambda_p = \inf_{x \in (0, D)} I(g)(x)^{-1} \leq \sup_{f \in \mathcal{F}_I} \inf_{x \in (0, D)} I(f)(x)^{-1}.$$

Step 3 When u and v' are continuous, we prove that

$$\sup_{f \in \mathcal{F}_{II}} \inf_{x \in (0, D)} II(f)(x)^{-1} \geq \sup_{h \in \mathcal{H}} \inf_{x \in (0, D)} R(h)(x).$$

First, we change the form of $R(h)$. Let g with $g(D) = 0$ be a positive function on $[0, D)$ such that $h = g'/g$ (see the arguments after Lemma 3.6). Then

$$R(h) = -u^{-1} \{ |h|^{p-2} [v'h + (p-1)v(h^2 + h')] \} = -\frac{1}{ug^{p-1}} L_p g.$$

Now, we turn to our main text. It suffices to show that

$$\sup_{f \in \mathcal{F}_{II}} \inf_{x \in (0, D)} II(f)(x)^{-1} \geq \inf_{x \in (0, D)} R(h)(x) \quad \text{for every } h \in \mathcal{H}.$$

Without loss of generality, assume that $\inf_{x \in (0, D)} R(h)(x) > 0$, which implies $R(h) > 0$ on $(0, D)$. Let $f = g(R(h))^{p^*-1}$ (g is the function just specified). Since u, v' are continuous, we have $f \in \mathcal{F}_{II}$ and

$$u(x)f^{p-1}(x) = -L_p g(x), \quad x \in (0, D).$$

Moreover, by (10), we have

$$g(y) - g(D) = \int_y^D \hat{v}(x) \left(\int_0^x f^{p-1} d\mu \right)^{p^*-1} dx.$$

So $g^{p-1}/f^{p-1} \geq II(f)$ on $(0, D)$ and

$$\inf_{(0, D)} R(h) = \inf_{(0, D)} f^{p-1}/g^{p-1} \leq \inf_{(0, D)} II(f)^{-1} \leq \sup_{f \in \mathcal{F}_{II}} \inf_{x \in (0, D)} II(f)(x)^{-1}.$$

Hence, the required assertion holds.

Step 4 Prove that $\sup_{h \in \mathcal{H}} \inf_{x \in (0, D)} R(h)(x) \geq \lambda_p$ when u and v' are continuous.

Noticing that

$$\hat{v}(x, D) \left(\int_0^x f^{p-1} d\mu \right)^{p^*-1} \leq f H(f)(x)^{p^*-1} \leq \hat{v}(x, D) \left(\int_0^D f^{p-1} d\mu \right)^{p^*-1},$$

If $\hat{v}(0, D) < \infty$, then choose $f \in L^{p-1}(\mu)$ to be a positive function such that $g = f H(f)^{p^*-1} < \infty$. Set $\bar{h} = g'/g$. Then $\bar{h} \in \mathcal{H}$ since u and v' are continuous. Moreover, $L_p g = -u f^{p-1}$ and

$$R(\bar{h}) = -\frac{1}{u g^{p-1}} L_p g = \frac{f^{p-1}}{g^{p-1}} > 0.$$

If $\hat{v}(0, D) = \infty$, then set $\bar{h} = 0$. So $R(\bar{h}) = 0$. In other words, we always have

$$\sup_{h \in \mathcal{H}} \inf_{x \in (0, D)} R(h)(x) \geq 0.$$

Without loss of generality, assume that $\lambda_p > 0$ and g is an eigenfunction of λ_p , i.e.,

$$L_p g = -\lambda_p u |g|^{p-2} g.$$

Let $h = g'/g \in \mathcal{H}$. Then $R(h) = \lambda_p$ and the assertion holds.

Step 5 Prove that the supremum in the lower estimates can be attained.
Since

$$0 = \lambda_p \geq \inf_{x \in (0, D)} H(f)(x)^{-1} \geq 0, \quad 0 = \lambda_p \geq \inf_{x \in (0, D)} I(f)(x)^{-1} \geq 0$$

for every f in the set defining λ_p , the assertion is clear for the case that $\lambda_p = 0$. Similarly, the conclusion holds for operator R as seen from the preceding proof in Step 4. For the case that $\lambda_p > 0$, assume that g is an eigenfunction corresponding to λ_p . Let $\bar{h} = g'/g \in \mathcal{H}$. Then $R(\bar{h}) = \lambda_p$, $I(g)^{-1} \equiv \lambda_p$ by letting $f = \lambda_p^{p^*-1} g$ in (9) and $H(g)^{-1} \equiv \lambda_p$ by letting $f = \lambda_p^{p^*-1} g$ in (10) whenever $g(D) = 0$.

Now, it remains to show that the vanishing property of eigenfunction at D , which is proved in the following proposition by using the variational formula proved in Step 1 above.

Proposition 3.7 *Let g be an a.e. eigenfunction of $\lambda_p > 0$. Then $g(D) = 0$.*

Proof Let $f = g - g(D)$. Then $f \in \mathcal{F}_H$. By (10), we have

$$f(x) = \lambda_p^{p^*-1} \int_x^D \hat{v}(t) \left(\int_0^t g^{p-1} d\mu \right)^{p^*-1} dt.$$

We prove the proposition by dividing it into two cases. Denoted by

$$M(x) = \int_x^D \hat{v}(t) \left(\int_0^t d\mu \right)^{p^*-1} dt.$$

(a) If $M(x) = \infty$, then $f(x) = g(x) - g(D) < \infty$ and

$$\lambda_p^{1-p^*} f(x) = \int_x^D \hat{v}(t) \left(\int_0^t g^{p-1} d\mu \right)^{p^*-1} dt > g(D)M(x) = \infty$$

once $g(D) \neq 0$. So there is a contradiction.

(b) If $M(x) < \infty$, then

$$fH(f)(x)^{p^*-1} = \int_x^D \hat{v}(t) \left(\int_0^t (g - g(D))^{p-1} d\mu \right)^{p^*-1} dt < g(0)M(0) < \infty.$$

Replacing f in the denominator of $H(f)$ with this term and using Cauchy's mean-value theorem twice, we have

$$\sup_{(0,D)} H(f) \leq \frac{1}{\lambda_p} \sup_{(0,D)} \frac{f^{p-1}}{g^{p-1}} = \frac{1}{\lambda_p} \sup_{x \in (0,D)} \left(1 - \frac{g(D)}{g(x)} \right)^{p-1} = \frac{1}{\lambda_p} \left(1 - \frac{g(D)}{g(0)} \right)^{p-1}.$$

The last equality comes from the fact that $g \searrow$. If $g(D) > 0$, then

$$\lambda_p^{-1} \leq \inf_{f \in \mathcal{F}_H} \sup_{x \in (0,D)} H(f)(x) \leq \sup_{x \in (0,D)} H(f)(x) < \lambda_p^{-1},$$

which is a contradiction. Therefore, we must have $g(D) = 0$. $\square$

By now, we have finished the proof of the lower estimates of λ_p . Dually, one can prove the upper estimates without too much difficulty. We ignore the details here.

The following lemma or its variants have been used many times before (cf., [3; Proof of Theorem 3.1], [2; page 97], or [7], and the earlier publications therein). It is essentially an application of the integration by parts formula, and is a key to the proof of Theorem 2.3.

Lemma 3.8 *Assume that m and n are two non-negative locally integrable functions. For $p > 1$, define*

$$S(x) = \left(\int_x^D n(y) dy \right)^{p-1}, \quad M(x) = \int_0^x m(y) dy$$

and $c_0 = \sup_{x \in (0,D)} S(x)M(x) < \infty$. Then

$$\int_0^x m(y) S(y)^{p^*r/p} dy \leq \frac{c_0}{1 - p^*r/p} S(x)^{(p^*r/p)-1}, \quad r \in (0, p/p^*).$$

Proof of Theorem 2.3 First, we prove that $\lambda_p \geq (k(p)\sigma_p)^{-1}$. Fixing $r \in (0, p/p^*)$, let $f(x) = \hat{\nu}(x, D)^{p^*r/p}$. Applying $m(x) = u(x)$, $n(x) = \hat{\nu}(x)$ to Lemma 3.8, we have $M(x) = \mu(0, x)$, $S(x) = \hat{\nu}(x, D)^{p-1}$, $c_0 = \sigma_p$ and

$$\int_0^x \hat{\nu}(y, D)^r \mu(dy) \leq \frac{\sigma_p}{1 - p^*r/p} \hat{\nu}(x, D)^{r-(p/p^*)}.$$

Since

$$|f'|^{p-2} f' = - \left(\frac{p^*r}{p} \hat{\nu}(\cdot, D)^{(p^*r/p)-1} \hat{\nu}(\cdot) \right)^{p-1},$$

we have

$$\sup_{x \in (0, D)} I(f)(x) \leq \frac{[p/(p^*r)]^{p-1}}{1 - p^*r/p} \sigma_p. \quad (11)$$

By Theorem 2.1 (1), (11), and an optimization with respect to $r \in (0, p/p^*)$, we obtain

$$\lambda_p^{-1} \leq \left(\sup_{f \in \mathcal{F}_I} \inf_{x \in (0, D)} I(f)(x)^{-1} \right)^{-1} \leq pp^{*p-1} \sigma_p = k(p) \sigma_p.$$

Now we prove that $\lambda_p \leq \sigma_p^{-1}$. For fixed $x_0, x_1 \in (0, D)$ with $x_0 < x_1$, let $f(x) = \hat{\nu}(x \vee x_0, D) \mathbb{1}_{[0, x_1)}(x)$. Then

$$I(f)(x) = \hat{\nu}(x_0, D)^{p-1} \mu(0, x_0) + \int_{x_0}^x \hat{\nu}(t, D)^{p-1} \mu(dt), \quad x \in (x_0, x_1)$$

and $I(f)(x) = \infty$ on $[0, x_0] \cup [x_1, D]$ by convention $1/0 = \infty$. Combining with Theorem 2.1 (1), we have

$$\lambda_p^{-1} \geq \inf_{x < x_1} I(f)(x) = \hat{\nu}(x_0, D)^{p-1} \mu(0, x_0), \quad x_0 < x_1.$$

Thereby the assertion that $\lambda_p \leq \sigma_p^{-1}$ follows by letting $x_1 \rightarrow D$. Since

$$\mu(0, x)^{p^*-1} \hat{\nu}(x, D) \leq \int_x^D \mu(0, s)^{p^*-1} \hat{\nu}(s) ds \leq \int_0^D \mu(0, s)^{p^*-1} \hat{\nu}(s) ds,$$

the assertions hold. $\square$

From the proof above, it is easy to understand why we choose the test function as $f = \hat{\nu}[\cdot, D]^{1/p^*}$ in [5; Proof of Theorem 2.3 (a)] in the discrete case.

Proof of Theorem 2.4 Using Cauchy's mean-value theorem and definitions of δ'_n , δ_n , $\bar{\delta}_n$ and λ_p , it is not hard to show the most of the results except that $\bar{\delta}_{n+1} \geq \delta'_n$. Put $f = f_n^{x_0, x_1}$ and $g = f_{n+1}^{x_0, x_1}$. Then $g = fH(f)^{p^*-1}$. By simple calculation, we have

$$D_p(g) = \int_0^{x_1} |g'|^{p-1} |g'| (x) v(x) dx = \int_0^{x_1} v(x)^{-1} \int_0^x f^{p-1} d\mu |g'(x)| v(x) dx$$

Exchanging the order of the integrals, we have

$$\begin{aligned} D_p(g) &= - \int_0^{x_1} f^{p-1}(t) \mu(dt) \int_t^{x_1} g'(x) dx \quad (\text{by Fubini's Theorem}) \\ &\leq \int_0^{x_1} f^{p-1}(t) g(t) \mu(dt) \quad (\text{since } g(x_1) \geq 0) \\ &\leq \int_0^{x_1} g^p d\mu \sup_{t \in (0, x_1)} \left(\frac{f(t)}{g(t)} \right)^{p-1} \\ &\leq \mu(|g|^p) \sup_{x \in (0, x_1)} II(f)(x)^{-1}. \end{aligned}$$

So the required assertion holds. $\square$

Proof of Corollary 2.5 (a) The calculation of δ_1 is simple. We compute δ'_1 first. Consider the term $\inf_{x < x_1} II(f_1^{x_0, x_1})(x)$. By calculation, we obtain that for $x \in (x_0, x_1)$, the numerator of $(II(f_1^{x_0, x_1})(x)^{p^*-1})'|_x$ equals

$$\hat{v}(x) \left[\int_x^{x_1} \hat{v}(s) \left(\int_0^s (f_1^{x_0, x_1})^{p-1} d\mu \right)^{p^*-1} ds - \hat{v}(x, x_1) \left(\int_0^x (f_1^{x_0, x_1})^{p-1} d\mu \right)^{p^*-1} \right]$$

which is obviously non-negative. So

$$II(f_1^{x_0, x_1})(x) = \left[\frac{1}{\hat{v}(x, x_1)} \int_x^{x_1} \hat{v}(s) \left(\int_0^s \hat{v}(t \vee x_0, x_1)^{p-1} \mu(dt) \right)^{p^*-1} ds \right]^{p-1}$$

is increasing in $x \in (x_0, x_1)$. Hence,

$$\begin{aligned} \delta'_1 &= \sup_{x_0 < x_1} \left[\frac{1}{\hat{v}(x_0, x_1)} \int_{x_0}^{x_1} \hat{v}(s) \left(\int_0^s \hat{v}(t \vee x_0, x_1)^{p-1} \mu(dt) \right)^{p^*-1} ds \right]^{p-1} \\ &= \sup_{x_0 \in (0, D)} \frac{1}{\hat{v}(x_0, D)^{p-1}} \left[\int_{x_0}^D \hat{v}(s) \left(\int_0^s \hat{v}(t \vee x_0, D)^{p-1} \mu(dt) \right)^{p^*-1} ds \right]^{p-1}. \end{aligned}$$

In the last equality, we have used the fact that $II(f_1^{x_0, x_1})(x_0)$ is increasing in $x_1 \in [x_0, D]$. Indeed, let

$$N_k(s, y) = \int_{x_0}^s \hat{v}(t, y)^k \mu(dt), \quad f(s, y) = \hat{v}(s) N_{p-1}(s, y)^{p^*-1}.$$

Then

$$\begin{aligned} II(f_1^{(x_0, y)})(x_0)^{p^*-1} &= \frac{1}{\hat{v}(x_0, y)} \left[\int_{x_0}^y f(s, y) ds + \int_{x_0}^y \hat{v}(s) ds \int_0^{x_0} \hat{v}(x_0, y)^{p-1} \mu(dt) \right] \\ &= \frac{1}{\hat{v}(x_0, y)} \int_{x_0}^y f(s, y) ds + \mu(0, x_0) \hat{v}(x_0, y)^{p-1} \\ &=: H_1(y) + H_2(y), \end{aligned}$$

and

$$\begin{aligned}\frac{\partial}{\partial y} N_{p-1}(s, y) &= \int_{x_0}^s (p-1) \hat{\nu}(t, y)^{p-2} \hat{\nu}(y) \mu(dt) \\ &= (p-1) \hat{\nu}(y) N_{p-2}(s, y); \\ \frac{\partial}{\partial y} f(s, y) &= (p^* - 1) \hat{\nu}(s) N_{p-1}(s, y)^{p^*-2} \frac{\partial}{\partial y} N_{p-1}(s, y); \\ \frac{\partial}{\partial y} \int_{x_0}^y f(s, y) ds &= \int_{x_0}^y \frac{\partial}{\partial y} f(s, y) ds + f(y, y).\end{aligned}$$

Hence, the numerator of dH_1/dy equals

$$\begin{aligned}& \left(\frac{\partial}{\partial y} \int_{x_0}^y f(s, y) ds \right) \hat{\nu}(x_0, y) - \hat{\nu}(y) \int_{x_0}^y f(s, y) ds \\ &= \hat{\nu}(x_0, y) \hat{\nu}(y) \int_{x_0}^y \hat{\nu}(s) N_{p-1}(s, y)^{p^*-2} N_{p-2}(s, y) ds \\ & \quad + \hat{\nu}(x_0, y) f(y, y) - \hat{\nu}(y) \int_{x_0}^y f(s, y) ds \\ &= \hat{\nu}(y) \left(\hat{\nu}(x_0, y) \int_{x_0}^y \hat{\nu}(s) N_{p-1}(s, y)^{p^*-2} N_{p-2}(s, y) ds \right. \\ & \quad \left. - \int_{x_0}^y \hat{\nu}(s) N_{p-1}(s, y)^{p^*-1} ds \right) + \hat{\nu}(x_0, y) f(y, y).\end{aligned}$$

Since $\hat{\nu}(x_0, y) N_{p-2}(s, y) - N_{p-1}(s, y) \geq 0$ for $s \in [x_0, y]$, we see that dH_1/dy is positive. It is obvious that dH_2/dy is positive. So $II(f_1^{x_0, y})(x_0)$ is increasing in y and the required assertion holds.

(b) Compute $\bar{\delta}_1$. By definition of $\bar{\delta}_1$, we have

$$\begin{aligned}\|f_1^{x_0, x_1}\|_p^p &= \int_0^{x_1} \left(\int_{x_0 \vee x}^{x_1} \hat{\nu}(s) ds \right)^p \mu(dx) \\ &= \mu(0, x_0) \hat{\nu}(x_0, x_1)^p + \int_{x_0}^{x_1} \left(\int_x^{x_1} \hat{\nu}(t) dt \right)^p \mu(dx), \\ D_p(f_1^{x_0, x_1}) &= \int_{x_0}^{x_1} \hat{\nu}(t)^p v(t) dt = \hat{\nu}(x_0, x_1).\end{aligned}$$

Hence,

$$\begin{aligned}\bar{\delta}_1 &= \sup_{x_0 < x_1} \left(\mu(0, x_0) \hat{\nu}(x_0, x_1)^{p-1} + \frac{1}{\hat{\nu}(x_0, x_1)} \int_{x_0}^{x_1} \hat{\nu}(s, x_1)^p \mu(ds) \right) \\ &= \sup_{x_0 \in (0, D)} \left(\mu(0, x_0) \hat{\nu}(x_0, D)^{p-1} + \frac{1}{\hat{\nu}(x_0, D)} \int_{x_0}^D \hat{\nu}(s, D)^p \mu(ds) \right)\end{aligned}$$

In the second equality, we have used the fact that:

$$\mu(0, x_0) \hat{\nu}(x_0, x_1)^{p-1} + \frac{1}{\hat{\nu}(x_0, x_1)} \int_{x_0}^{x_1} \hat{\nu}(s, x_1)^p \mu(ds) \quad \uparrow \quad \text{in } x_1.$$

Indeed, it suffices to show that

$$\frac{1}{\hat{\nu}(x_0, x)} \int_{x_0}^x \hat{\nu}(s, x)^p \mu(ds) \leq \frac{1}{\hat{\nu}(x_0, y)} \int_{x_0}^y \hat{\nu}(s, y)^p \mu(ds), \quad x_0 \leq x < y,$$

which is equivalent to

$$\frac{1}{\hat{\nu}(x_0, y)} \int_x^y \hat{\nu}(s, y)^p \mu(ds) + \int_{x_0}^x \frac{\hat{\nu}(s, y)^p}{\hat{\nu}(x_0, y)} - \frac{\hat{\nu}(s, x)^p}{\hat{\nu}(x_0, x)} \mu(ds) \geq 0.$$

Since $p > 1$ and $\hat{\nu}(t, x) \leq \hat{\nu}(x_0, x)$ for $x \geq t \geq x_0$, we have

$$\frac{\hat{\nu}(t, y)^p}{\hat{\nu}(t, x)^p} = \left[\frac{\hat{\nu}(t, x) + \hat{\nu}(x, y)}{\hat{\nu}(t, x)} \right]^p \geq 1 + \frac{\hat{\nu}(x, y)}{\hat{\nu}(t, x)} \geq 1 + \frac{\hat{\nu}(x, y)}{\hat{\nu}(x_0, x)} = \frac{\hat{\nu}(x_0, y)}{\hat{\nu}(x_0, x)}$$

for $t \geq x_0$ and the required assertion holds.

(c) Comparing δ'_1 and $\bar{\delta}_1$. It is easy to see that

$$\begin{aligned} \int_x^D \hat{\nu}(\cdot, D)^p d\mu &= \int_x^D \hat{\nu}(t, D)^{p-1} \int_t^D \hat{\nu}(s) ds \mu(dt) \\ &= \int_x^D \hat{\nu}(s) \int_x^s \hat{\nu}(t, D)^{p-1} \mu(dt) ds; \\ \mu(0, x) \hat{\nu}(x, D)^p &= \int_x^D \hat{\nu}(s) \int_0^x \hat{\nu}(x, D)^{p-1} \mu(dt) ds. \end{aligned}$$

Let $a_x(s) = \hat{\nu}(s)/\hat{\nu}(x, D)$ for $s \in (x, D)$. Noticing that a_x is a probability on (x, D) , by the increasing property of moments $\mathbb{E}(|X|^s)^{1/s}$ in $s > 0$ and combining the preceding assertions (a) and (b), we have

$$\begin{aligned} \bar{\delta}_1 &= \sup_{x \in (0, D)} \int_x^D a_x(s) \int_0^s \hat{\nu}(t \vee x, D)^{p-1} \mu(dt) ds \\ &\leq \sup_{x \in (0, D)} \left[\int_x^D a_x(s) \left(\int_0^s \hat{\nu}(t \vee x, D)^{p-1} \mu(dt) \right)^{p^*-1} ds \right]^{p-1} \quad (\text{if } p^* - 1 > 1) \\ &= \delta'_1. \end{aligned}$$

Similarly, if $p^* - 1 < 1$ (i.e., $p > 2$), then $\bar{\delta}_1 \geq \delta'_1$.

(d) Prove that $\bar{\delta}_1 \leq p\sigma_p$. Using the integration by parts formula, we have

$$\begin{aligned} \int_{x_0}^x \hat{\nu}(y, D)^p \mu(dy) &= \hat{\nu}(y, D)^p \mu(0, y) \Big|_{x_0}^x + p \int_{x_0}^x \hat{\nu}(y, D)^{p-1} \hat{\nu}(y) \mu(0, y) dy \\ &\leq \sigma_p \hat{\nu}(x, D) - \hat{\nu}(x_0, D)^p \mu(0, x_0) + p\sigma_p \int_{x_0}^x \hat{\nu}(y) dy \end{aligned}$$

Since $\hat{v}(x, D) < \infty$, letting $x \rightarrow D$, we have

$$\begin{aligned}\bar{\delta}_1 &= \sup_{x_0 \in (0, D)} \left(\mu(0, x_0) \hat{v}(x_0, D)^{p-1} + \frac{1}{\hat{v}(x_0, D)} \int_{x_0}^D \hat{v}(\cdot, D)^p d\mu \right) \\ &\leq \sup_{x_0 \in (0, D)} \left[\mu(0, x_0) \hat{v}(x_0, D)^{p-1} + \right. \\ &\quad \left. \frac{1}{\hat{v}(x_0, D)} \left(-\hat{v}(x_0, D)^p \mu(0, x_0) + p\sigma_p \int_{x_0}^D \hat{v}(y) dy \right) \right] \\ &= p\sigma_p,\end{aligned}$$

and the required assertion holds. $\square$

4 DN-case

From now on, we concern on p -Laplacian eigenvalue with DN-boundaries. We use the same notation as the previous ND-case since they play the similar role but have different meaning in different context. Let $D \leq \infty$, $p > 1$. The p -Laplacian eigenvalue problem with DN-boundary conditions is

$$\begin{cases} \text{Eigenequation : } & L_p g(x) = -\lambda u(x) |g|^{p-2} g(x); \\ \text{DN-boundaries : } & g(0) = 0, \quad g'(D) = 0 \text{ if } D < \infty \end{cases} \quad (12)$$

The first eigenvalue λ_p has the following classical variational formula:

$$\lambda_p = \inf \left\{ \frac{D_p(f)}{\mu(|f|^p)} : f(0)=0, f \neq 0, f \in \mathcal{C}[0, D], v^{p^*-1} f' \in \mathcal{C}(0, D), D_p(f) < \infty \right\}. \quad (13)$$

Correspondingly, we are also estimating the optimal constant $A := \lambda_p^{-1}$ in the *weighted Hardy inequality*:

$$\mu(|f|^p) \leq AD_p(f), \quad f(0) = 0, \quad f \in \mathcal{D}(D_p).$$

For $p > 1$, define $\hat{v} = v^{1-p^*}$ and $\hat{v}(dx) = \hat{v}(x)dx$. We use the following operators:

$$\begin{aligned} I(f)(x) &= \frac{1}{(vf'|f'|^{p-2})(x)} \int_x^D f^{p-1} d\mu \quad (\text{single integral form}) \\ II(f)(x) &= \frac{1}{f^{p-1}(x)} \left[\int_0^x \hat{v}(s) \left(\int_s^D f^{p-1} d\mu \right)^{p^*-1} ds \right]^{p-1} \\ &\quad (\text{double integral form}) \\ R(h)(x) &= -u^{-1} \{ |h|^{p-2} [v'h + (p-1)v(h^2 + h')] \} (x) \quad (\text{differential form}). \end{aligned}$$

The three operators above have domains, respectively, as follows.

$$\begin{aligned}\mathcal{F}_I &= \{f \in \mathcal{C}[0, D] : v^{p^*-1}f' \in \mathcal{C}(0, D), f(0) = 0 \text{ and } f'|_{(0, D)} > 0\}, \\ \mathcal{F}_{II} &= \{f : f \in \mathcal{C}[0, D], f(0) = 0 \text{ and } f|_{(0, D)} > 0\}, \\ \mathcal{H} &= \left\{h : h \in \mathcal{C}^1(0, D) \cap \mathcal{C}[0, D], h|_{(0, D)} > 0 \text{ and } \int_{0+} h(u)du = \infty\right\},\end{aligned}$$

where $\int_{0+}$ means $\int_0^\varepsilon$ for sufficiently small $\varepsilon > 0$. Some modifications are needed when studying the upper estimates.

$$\begin{aligned}\widetilde{\mathcal{F}}_I &= \{f \in \mathcal{C}[0, x_0] : f(0) = 0, v^{p^*-1}f' \in \mathcal{C}(0, x_0), f'|_{(0, x_0)} > 0 \text{ for some} \\ &\quad x_0 \in (0, D), \text{ and } f = f(\cdot \wedge x_0)\}, \\ \widetilde{\mathcal{F}}_{II} &= \{f : f(0) = 0, \exists x_0 \in (0, D) \text{ such that } f = f(\cdot \wedge x_0) > 0 \text{ and } f \in \mathcal{C}[0, x_0]\}, \\ \widetilde{\mathcal{H}} &= \left\{h : \exists x_0 \in (0, D) \text{ such that } h \in \mathcal{C}[0, x_0] \cap \mathcal{C}^1(0, x_0), h|_{(0, x_0)} > 0, \right. \\ &\quad \left. h|_{[x_0, D]} = 0, \int_{0+} h(u)du = \infty, \text{ and } \sup_{(0, x_0)} [v'h + (p-1)(h^2 + h')v] < 0\right\}.\end{aligned}$$

When $D = \infty$, replace $[0, D]$ and $(0, D]$ with $[0, D)$ and $(0, D)$, respectively. Besides, we also need the following notation:

$$\widetilde{\mathcal{F}}'_{II} = \{f : f(0) = 0, f \in \mathcal{C}[0, D] \text{ and } fH(f) \in L^p(\mu)\}.$$

If $\mu(0, D) = \infty$, then λ_p defined by (13) is trivial. Indeed, let

$$f = \mathbb{1}_{(\delta, D]} + h\mathbb{1}_{[0, \delta]},$$

where h is chosen such that $h(0) = 0$ and $f \in \mathcal{C}^1(0, D) \cap \mathcal{C}[0, D]$ (for example, $h(x) = -x^2 \cdot \delta^{-2} + 2x \cdot \delta^{-1}$). Then $D_p(f) \in (0, \infty)$ and $\mu(|f|^p) = \infty$. It follows that $\lambda_p = 0$.

Otherwise, $\mu(0, D) < \infty$. Then for every f with $\mu(|f|^p) = \infty$, by setting $f^{(x_0)} = f(\cdot \wedge x_0) \in L^p(\mu)$, we have

$$\infty > D(f^{(x_0)}) \rightarrow D(f), \quad \infty > \mu(|f^{(x_0)}|^p) \rightarrow \mu(|f|^p) \quad \text{as } x_0 \rightarrow D.$$

In other words, for $f \notin L^p(\mu)$, both $\mu(|f|^p)$ and $D_p(f)$ can be approximated by a sequence of functions belonging to $L^p(\mu)$. Hence, we can rewrite λ_p as follows.

$$\lambda_p = \inf \{D_p(f) : \mu(|f|^p) = 1, f(0) = 0, \text{ and } f \in \mathcal{C}^1(0, D) \cap \mathcal{C}[0, D]\}. \quad (14)$$

In this case, we also have

$$\begin{aligned}\lambda_p &= \inf \{D_p(f) : \mu(|f|^p) = 1, f(0) = 0, f = f(\cdot \wedge x_0), \\ &\quad f \in \mathcal{C}^1(0, x_0) \cap \mathcal{C}[0, x_0] \text{ for some } x_0 \in (0, D)\}.\end{aligned}$$

We are now ready to state the main results in the present context.

Theorem 4.1 Assume that $\mu(0, D) < \infty$. For $p > 1$, the following variational formulas hold for λ_p defined by (14) (equivalently, (13)).

(1) *Single integral forms:*

$$\inf_{f \in \widetilde{\mathcal{F}}_I} \sup_{x \in (0, D)} I(f)(x)^{-1} = \lambda_p = \sup_{f \in \mathcal{F}_I} \inf_{x \in (0, D)} I(f)(x)^{-1},$$

(2) *Double integral forms:*

$$\begin{aligned} \lambda_p &= \inf_{f \in \widetilde{\mathcal{F}}_I} \sup_{x \in (0, D)} II(f)(x)^{-1} = \inf_{f \in \widetilde{\mathcal{F}}_{II} \cup \widetilde{\mathcal{F}}'_{II}} \sup_{x \in (0, D)} II(f)(x)^{-1}, \\ \lambda_p &= \sup_{f \in \mathcal{F}_I} \inf_{x \in (0, D)} II(f)(x)^{-1} = \sup_{f \in \mathcal{F}_{II}} \inf_{x \in (0, D)} II(f)(x)^{-1}. \end{aligned}$$

Moreover, if u and v' are continuous, then we have additionally

(3) *differential forms:*

$$\inf_{h \in \widetilde{\mathcal{H}}} \sup_{x \in (0, D)} R(h)(x) = \lambda_p = \sup_{h \in \mathcal{H}} \inf_{x \in (0, D)} R(h)(x).$$

Define $k(p) = pp^{*p-1}$ and

$$\sigma_p = \sup_{x \in (0, D)} \mu(x, D) \hat{\nu}(0, x)^{p-1}.$$

As an application of the variational formulas in Theorem 4.1 (1), we have the following theorem which was also known in 1990's (cf. [11; Lemmas 3.2 and 3.4 on pages 22 and 25, respectively]).

Theorem 4.2 (Criterion and basic estimates) For $p > 1$, $\lambda_p > 0$ if and only if $\sigma_p < \infty$. Moreover,

$$(k(p)\sigma_p)^{-1} \leq \lambda_p \leq \sigma_p^{-1}.$$

In particular, we have $\lambda_p = 0$ if $\mu(0, D) = \infty$ and $\lambda_p > 0$ if

$$\int_0^D \mu(s, D)^{p^*-1} \hat{\nu}(ds) < \infty.$$

The next result is an application of the variational formulas in Theorem 4.1 (2).

Theorem 4.3 (Approximating procedure) Assume that $\mu(0, D) < \infty$ and $\sigma_p < \infty$.

(1) Let $f_1 = \hat{\nu}(0, \cdot)^{1/p^*}$, $f_{n+1} = f_n II(f_n)^{p^*-1}$ and $\delta_n = \sup_{x \in (0, D)} II(f_n)(x)$ for $n \geq 1$. Then δ_n is decreasing in n and

$$\lambda_p \geq \delta_n^{-1} \geq (k(p)\sigma_p)^{-1}.$$

(2) For fixed $x_0 \in (0, D)$, let

$$f_1^{(x_0)} = \hat{\nu}(0, \cdot \wedge x_0), \quad f_n^{(x_0)} = f_{n-1}^{(x_0)} \Pi(f_{n-1}^{(x_0)})(\cdot \wedge x_0)^{p^*-1}$$

and $\delta'_n = \sup_{x_0 \in (0, D)} \inf_{x \in (0, D)} \Pi(f_n^{(x_0)})(x)$ for $n \geq 1$. Then δ'_n is increasing in n and

$$\sigma_p^{-1} \geq \delta'_n{}^{-1} \geq \lambda_p.$$

Moreover, define

$$\bar{\delta}_n = \sup_{x_0 \in (0, D)} \frac{\|f_n^{(x_0)}\|_p^p}{D_p(f_n^{(x_0)})}, \quad n \geq 1.$$

Then $\bar{\delta}_n^{-1} \geq \lambda_p$ and $\bar{\delta}_{n+1} \geq \delta'_n$ for $n \geq 1$.

Most of the result in Corollary 4.4 below can be obtained directly from Theorem 4.3.

Corollary 4.4 (Improved estimates) Assume that $\mu(0, D) < \infty$ and $\lambda_p > 0$. We have

$$\sigma_p^{-1} \geq \delta'_1{}^{-1} \geq \lambda_p \geq \delta_1^{-1} \geq (k(p)\sigma_p)^{-1},$$

where

$$\delta_1 = \sup_{x \in (0, D)} \left[\frac{1}{\hat{\nu}(0, x)^{1/p^*}} \int_0^x \hat{\nu}(s) \left(\int_s^D \hat{\nu}(0, t)^{p/p^*2} \mu(dt) \right)^{p^*-1} ds \right]^{p-1};$$

$$\delta'_1 = \sup_{x \in (0, D)} \frac{1}{\hat{\nu}(0, x)^{p-1}} \left[\int_0^x \hat{\nu}(s) \left(\int_s^D \hat{\nu}(0, t \wedge x)^{p-1} \mu(dt) \right)^{p^*-1} ds \right]^{p-1}.$$

Moreover,

$$\bar{\delta}_1 = \sup_{x \in (0, D)} \left(\mu(x, D) \hat{\nu}(0, x)^{p-1} + \frac{1}{\hat{\nu}(0, x)} \int_0^x \hat{\nu}(0, t)^p \mu(dt) \right) \in [\sigma_p, p\sigma_p],$$

and $\bar{\delta}_1 \geq \delta_1$ for $p \geq 2$ and $\bar{\delta}_1 \leq \delta'_1$ for $1 < p \leq 2$.

When $p = 2$, the equality $\delta_1 = \bar{\delta}_1$ was proved in [4; Theorem 6].

Most of the results in this section are parallel to that in Section 2. One may follow Section 3 or [4, 7] to complete the proofs without too many difficulties. The details are omitted here. Instead, we prove some properties of the eigenfunction g , which are used in choosing the test functions for the operators.

Lemma 4.5 Let (λ_p, g) be a solution to (12), $g \neq 0$. Then g' does not change sign, and so does g .

Proof First, the solution provided by Lemma 3.1 is trivial: $g = 0$, if the given constants A and B are zero. Because we are in the situation that $g(0) = 0$, we can assume that $g'(0) \neq 0$. Next, we prove that g' does not change sign by seeking a contradiction. If there exists $x_0 \in (0, D)$ such that $g'(x_0) = 0$, then $g(x_0) \neq 0$ by [12; Lemma 2.3]. Let $\bar{g} = g\mathbb{1}_{[0, x_0]} + g(x_0)\mathbb{1}_{(x_0, D]}$. By simple calculation, we obtain

$$D_p(\bar{g}) = (-L_p \bar{g}, \bar{g})_\mu = \lambda_p \mu_{0, x_0}(|g|^p).$$

So

$$\lambda_p \leq \frac{D_p(\bar{g})}{\mu(|\bar{g}|^p)} = \frac{\lambda_p \mu_{0, x_0}(|g|^p)}{\mu_{0, x_0}(|g|^p) + \mu(x_0, D)|g(x_0)|^p} < \lambda_p,$$

which is a contradiction. Therefore g' does not change sign. Since $g(0) = 0$, the second assertion holds naturally.

Acknowledgements The work is supported in part by NSFC (Grant No.11131003), S-RFDP (Grant No. 20100003110005), the “985” project from the Ministry of Education in China and the Fundamental Research Funds for the Central Universities. The authors also thank Professor Yong-Hua Mao for his helpful comments and suggestions.

References

1. Cañada A. I., Drábek P., Fonda, A. Handbook of differential equations: Ordinary differential equations, Vol. 1. North Holland: Elsevier, 2004, 161–357.
2. Chen M.F. Eigenvalues, Inequalities, and Ergodic Theory. New York: Springer, 2005.
3. Chen M.F. Speed of stability for birth-death process. Front. Math. China, 2010, 5(3): 379–516.
4. Chen M.F., Wang, L.D., Zhang Y.H. Mixed principal eigenvalues in dimension one. Front. Math. China, 2013, 8(2): 317–343.
5. Chen M.F., Wang L.D., Zhang Y.H. Mixed eigenvalues of discrete p -Laplacian. preprint.
6. Drábek P., Kufner, A. Discreteness and simplicity of the spectrum of a quasilinear Sturm-Liouville-type problem on an infinite interval. Proc Amer Math Soc, 2005, 134(1):235–242.
7. Jin H.Y., Mao Y.H. Estimation of the Optimal Constants in the L^p -Poincaré inequalities on the Half Line of L^p -poncaré inequality on half line. Acta Math Sin (Chinese Series), 2012, 55(1): 169–178.
8. Kufner A., Maligranda L., Persson L. The prehistory of the Hardy inequality. Amer Math Mon, 2006, 113(8), 715–732.
9. Kufner A., Maligranda L., Persson L. The Hardy inequality: About its history and some related results. Plisen 2007.
10. Lane J., Edmunds D. Eigenvalues, embeddings and generalised trigonometric functions. Lecture Notes in Math, vol. 2016, 2011.
11. Opic B., Kufner A. Hardy Type Inequalities. Longman Scientific and Technical, 1990.
12. Pinasco J. Comparison of eigenvalues for the p -Laplacian with integral inequalities. Appl Math & Comp 2006, 182, 1399–1404.
13. Ważewski T. Sur un principe topologique de l'examen de l'allure asymptotique des intégrales des équations différentielles ordinaires. Ann. Soc. Polon. Math. 1947, 20, 279–313.