THÈSE DE DOCTORAT DE
l'UPL, UNIVERSITÉ PARIS NANTERRE

Spécialité : **Mathématiques**

Ecole doctorale "Connaissance, Langage et Modélisation"

Présentée par

Luca Tamanini

Pour obtenir le grade de

DOCTEUR de l'UPL, UNIVERSITÉ PARIS NANTERRE

Analysis and geometry of RCD spaces via the Schrödinger problem

---

Sous la direction de :

Pr. Christian Léonard et Pr. Nicola Gigli

Devant le jury composé de :

M. Andrei Agrachev	Examinateur
M. Massimiliano Berti	Examinateur
M. Gianni Dal Maso	Examinateur
M. Ivan Gentil	Rapporteur
M. Nicola Gigli	Co-directeur de thèse
M. Christian Léonard	Directeur de thèse
M. Olivier Raimond	Examinateur
M. Giuseppe Savaré	Rapporteur

Il presente lavoro costituisce la tesi presentata da Luca Tamanini, sotto la direzione di ricerca del prof. Nicola Gigli (SISSA) e del prof. Christian Léonard (Université Paris Nanterre), al fine di ottenere l'attestato di ricerca post-universitaria di *Doctor Philosophiae in Matematica applicata* presso la SISSA. Ai sensi dell'art. 18, comma 3, dello Statuto della SISSA pubblicato sulla G.U. n. 62 del 15.03.2001, il predetto attestato è equipollente al titolo di Dottore di Ricerca in Matematica.

Trieste, anno accademico 2016/2017

Acknowledgements

I would like to warmly thank my advisors prof. Nicola Gigli and prof. Christian Léonard for their guidance, help and support during all of my Ph.D. studies and in particular the creation of this thesis. Their input has always been invaluable, truly appreciated and very inspiring, because of the different nature (geometric on the one hand and the probabilistic on the other one) of their approach towards the mathematical challenges that we faced.

I would also like to thank G. De Philippis, E. Pasqualetto and C. Rigoni for the frequent discussions in SISSA about analysis and geometry of metric measure spaces; S. Di Marino and G. Savaré for interesting comments about the dynamic representation of the entropic cost, which will hopefully lead to further developments; I. Gentil for useful information on the Bakry-Émery approach to the theory of curvature-dimension bounds and for the course he gave in Grenoble. I am also grateful to SISSA for the high level courses I had the opportunity to take.

Also S. Almi, C. Biolo, G. Bogo, P. Bonicatto, V. Crismale, G. Lazzaroni, I. Lucardesi, E. Marconi, M. Morandotti, L. Nardini, A. Olgiati, G. Orlando and G. Pitton, with whom I spent a delightful and unforgettable year (and a half), deserve a heartful thank you. On the French side, I warmly thank P. de Buyer, S. Doucet, E. Joly, N. Noiry and C. Tillier: following a presidential election has never been so serious and hilarious at the same time.

All the professors, with particular mention of A. Chambaz for the crucial help through French bureaucracy, researchers, post-docs, secretarial staff and other workers both in SISSA and Paris Nanterre are included in the acknowledgements.

Finally, I would like to extend my gratitude to my family, Arianna and all my friends for the continuous support and encouragement when more it was required.

The author is thankful to the *Labex MME-DII: Modèles Mathématiques et Économiques de la Dynamique, de l'Incertitude et des Interactions* for partially funding his Ph.D. grant and to the *Université Franco-Italienne/Università Franco-Italiana* for the financial support of the Vinci Programme.

Contents

Introduction

Dans les derniers dix ans, l'intéresse pour les espaces métriques mesurés est augmenté de manière considérable, comme témoigné par la littérature riche et florissante. Pour une liste non exhaustive des résultats les plus importants sur le sujet, le lecteur peut s'adresser à [87], [115], [116], [60], [6], [7], [58], [8], [105], [106], [61], [55], [57], [73], [10], [98], [26], [25]. Les points de départ de cette recherche si active et prospère ont été les travaux-phare [87] et [115], [116], qui ont lié pour la première fois des bornes inférieures sur la courbure de Ricci sur des espaces métriques mesurés à des propriétés de fonctionnelles de type 'entropie' en connexion avec la géométrie de l'espace de Wasserstein. Successivement ([6]) on s'est aperçu que l'analyse de Sobolev est aussi associée à la géométrie Wasserstein et donc, en construisant sur ces considérations, la définition originelle d'espace CD à la Lott-Sturm-Villani a evolué vers celle d'espace RCD ([7], [58]).

Un exemple de lien entre l'analyse de Sobolev et la géométrie de l'espace de Wasserstein est donné par l'énoncé suivant, montré dans [55] :

Theorem 0.0.1 (Formule de dérivation du premier ordre)**.** *Soient*

- $(\mathrm{X}, \mathsf{d}, \mathfrak{m})$ *un espace* $\mathsf{RCD}(K, \infty)$*, avec* $K \in \mathbb{R}$
- (μ_t) *une géodésique Wasserstein constituée par des mesures à support borné telles que* $\mu_t \leq C\mathfrak{m}$ *pour tout* $t \in [0,1]$ *et une certaine constante* $C > 0$
- $f \in W^{1,2}(\mathrm{X})$.

Alors la fonction

$$[0,1] \ni t \quad \mapsto \quad \int f \, \mathrm{d}\mu_t$$

est C^1 *et l'identité suivante est satisfaite*

$$\frac{\mathrm{d}}{\mathrm{d}t} \int f \, \mathrm{d}\mu_t|_{t=0} = -\int \mathrm{d}f(\nabla\varphi) \, \mathrm{d}\mu_0,$$

où φ *est un potentiel de Kantorovich localement lipschitzien de* μ_0 *à* μ_1.

Rappelons que sur un espace $\mathsf{RCD}(K, \infty)$ toute géodésique Wasserstein (μ_t) entre deux mesures à support et densité bornés a elle-même densité bornée de manière uniforme, c'est à dire $\mu_t \leq C\mathfrak{m}$ pour tout $t \in [0,1]$ et une certaine constante $C > 0$ ([106]), de sorte que le théorème implique l'existence de plusieurs fonctions C^1 sur les espaces RCD. Il est important de remarquer que la régularité C^1 - qui a été crucial dans [55] – n'est pas du tout triviale, même si la fonction f est supposée lipschitzienne. En plus, les énoncés concernant la régularité C^1 sont assez rares en géométrie métrique.

Le Théorème 0.0.1 peut être vu comme une version intégrée de la formule de base

$$\frac{\mathrm{d}}{\mathrm{d}t} f(\gamma_t)_{|t=0} = \mathrm{d}f(\gamma_0')$$

qui est valable dans le cadre lisse ; d'un point de vue strictement technique, la preuve de cette affirmation est liée au fait que la géodésique (μ_t) résout l'équation de continuité

$$\frac{\mathrm{d}}{\mathrm{d}t}\mu_t + \mathrm{div}(\nabla(-\varphi_t)\mu_t) = 0, \tag{0.0.1}$$

où les φ_t sont des choix convenables de potentiels de Kantorovich (voir aussi [59] dans cette direction).

Dans [56], Gigli a développé le calcul du deuxième ordre sur les espaces RCD, en définissant en particulier l'espace $H^{2,2}(\mathrm{X})$ et, pour toute fonction $f \in H^{2,2}(\mathrm{X})$ la Hessienne $\mathrm{Hess}(f)$, voir [56] et la section préliminaire de la thèse (Section 1.2). Il est ensuite naturel de se poser la question si une version 'intégrée' de la formule de dérivation du deuxième ordre

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2} f(\gamma_t)_{|t=0} = \mathrm{Hess}(f)(\gamma_0', \gamma_0') \qquad \text{pour toute géodésique } \gamma$$

est satisfaite dans ce contexte. Dans ce manuscrit, et plus précisément dans le Chapitre 7, on donne une réponse affirmative à cette question et le résultat les plus important que l'on établira est le suivant.

Theorem 0.0.2 (Formule de dérivation du deuxième ordre)**.** *Soient*

- $(\mathrm{X}, \mathsf{d}, \mathfrak{m})$ *un espace* $\mathsf{RCD}^*(K, N)$*, avec* $K \in \mathbb{R}$ *et* $N < \infty$
- (μ_t) *une géodésique Wasserstein telle que* $\mu_t \leq C\mathfrak{m}$ *et* μ_t *a support compact pour tout* $t \in [0,1]$ *et une certaine constante* $C > 0$
- $f \in H^{2,2}(\mathrm{X})$.

Alors la fonction

$$[0,1] \ni t \quad \mapsto \quad \int f \, \mathrm{d}\mu_t$$

est C^2 *et l'identité suivante est satisfaite*

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2} \int f \, \mathrm{d}\mu_t{}_{|t=0} = \int \mathrm{Hess}(f)(\nabla\varphi, \nabla\varphi) \, \mathrm{d}\mu_0, \tag{0.0.2}$$

où φ *est un potentiel de Kantorovich de* μ_0 *à* μ_1.

Le lecteur peut voir aussi le Théorème 7.1.2 pour une formulation alternative mais tout à fait équivalente du résultat. Ce résultat a été prouvé en premier lieu dans [63] sous l'hypothèse supplémentaire de compacité, mais comme l'on expliquera dans la suite cette supposition peut être ôtée par le biais de techniques de localisation. Par contre, l'hypothèse de dimension finie joue un rôle clé, parce que par exemple on invoque à plusieurs reprises l'inégalité de Li-Yau.

Avoir à disposition une telle formule de dérivation du deuxième ordre, c'est intéressant non seulement d'un point de vue purement théorique, mais aussi pour les applications dans

l'étude de la géométrie des espaces RCD. Par exemple, les preuves des théorèmes de splitting et 'from volume cone to metric cone' peuvent être simplifiées à l'aide de cette formule. En plus, un aspect de la théorie des espaces RCD qui n'est pas encore clair est le suivant : est-ce que leur dimension est constante ? Pour les espaces qui sont des 'Ricci limits' la réponse est connue et positive, grâce au résultat de Colding-Naber [32], qui utilise de manière cruciale la dérivation du deuxième ordre le long des géodésiques. Donc, notre résultat est nécessaire afin de répliquer la preuve de Colding et Naber dans le cadre non lisse (toutefois, il n'est pas suffisant : ils utilisent aussi des techniques de calcul avec les champs de Jacobi qui, pour l'instant, n'admettent pas encore des analogues non lisses).

Stratégie de la preuve

Le point de départ est une formule de dérivation du deuxième ordre obtenue dans [56], valable sous certaines hypothèse de régularité :

Theorem 0.0.3. *Soit (μ_t) une courbe W_2-absolument continue solution de l'équation de continuité*

$$\frac{\mathrm{d}}{\mathrm{d}t}\mu_t + \mathrm{div}(X_t\mu_t) = 0,$$

pour un certain champ de vecteurs $(X_t) \subset L^2(T\mathrm{X})$ au sens suivant : pour toute $f \in W^{1,2}(\mathrm{X})$ la fonction $t \mapsto \int f \,\mathrm{d}\mu_t$ est absolument continue et l'identité

$$\frac{\mathrm{d}}{\mathrm{d}t}\int f \,\mathrm{d}\mu_t = \int \langle \nabla f, X_t\rangle \,\mathrm{d}\mu_t$$

est vérifiée. Supposons que

(i) $t \mapsto X_t \in L^2(T\mathrm{X})$ *est absolument continue,*

(ii) $\sup_t\{\|X_t\|_{L^2} + \|X_t\|_{L^\infty} + \|\nabla X_t\|_{L^2}\} < +\infty$.

Alors pour toute $f \in H^{2,2}(\mathrm{X})$ la fonction $t \mapsto \int f \,\mathrm{d}\mu_t$ est $C^{1,1}$ et la formule

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}\int f\mathrm{d}\mu_t = \int \mathrm{Hess}(f)(X_t, X_t) + \left\langle \nabla f, \tfrac{\mathrm{d}}{\mathrm{d}t}X_t + \nabla_{X_t}X_t\right\rangle \mathrm{d}\mu_t \tag{0.0.3}$$

est satisfaite pour p.t. $t \in [0,1]$.

Si les champs de vecteurs X_t sont du type gradient, c'est à dire $X_t = \nabla\phi_t$ pour tout t et l''accélération' a_t est définie comme étant

$$\frac{\mathrm{d}}{\mathrm{d}t}\phi_t + \frac{|\nabla\phi_t|^2}{2} =: a_t$$

alors (0.0.3) devient

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}\int f\mathrm{d}\mu_t = \int \mathrm{Hess}(f)(\nabla\phi_t, \nabla\phi_t)\,\mathrm{d}\mu_t + \int \langle \nabla f, \nabla a_t\rangle \,\mathrm{d}\mu_t. \tag{0.0.4}$$

Dans le cas des géodésiques, les fonctions φ_t qui apparaissent dans (0.0.1) résolvent (en un sens qui ne sera pas précisé ici) l'équation de Hamilton-Jacobi

$$\frac{\mathrm{d}}{\mathrm{d}t}\varphi_t = \frac{|\nabla\varphi_t|^2}{2}, \tag{0.0.5}$$

et donc dans ce cas l'accélération a_t est identiquement nulle (observons le signe moins dans (0.0.1)). Pour cette raison, si les champs de vecteurs $(\nabla\varphi_t)$ satisfaisaient les conditions de régularité $(i),(ii)$ du dernier théorème, on serait facilement capable d'obtenir le Théorème 0.0.2. Pourtant ceci n'est pas le cas en général ; de manière informelle, le problème est lié au fait que pour les solutions de l'équation de Hamilton-Jacobi on n'a pas d'estimations du deuxième ordre suffisamment fortes.

Afin de montrer le Théorème 0.0.2 il est donc naturel de chercher des approximations 'lisses' appropriées des géodésiques, auxquelles on peut appliquer le Théorème 0.0.3 précédent et puis passer à la limite dans la formule (0.0.3). Étant donné que le manque de régularité des géodésiques Wasserstein est conséquence du manque de régularité des solutions de (0.0.5), en parallèle avec la théorie classique de l'approximation visqueuse pour l'équation de Hamilton-Jacobi, il y a une approche assez naturelle à essayer : résoudre, pour $\varepsilon > 0$, l'équation

$$\frac{\mathrm{d}}{\mathrm{d}t}\varphi_t^\varepsilon = \frac{|\nabla\varphi_t^\varepsilon|^2}{2} + \frac{\varepsilon}{2}\Delta\varphi_t^\varepsilon, \qquad \varphi_0^\varepsilon := \varphi,$$

où φ est un potentiel de Kantorovich donné et fixé, associé à la géodésique (μ_t), et ensuite résoudre

$$\frac{\mathrm{d}}{\mathrm{d}t}\mu_t^\varepsilon - \mathrm{div}(\nabla\varphi_t^\varepsilon\mu_t^\varepsilon) = 0, \qquad \mu_0^\varepsilon := \mu_0.$$

Ce projet peut effectivement être mis en place et, en suivant les idées présentées dans ce manuscrit, on peut montrer que si $(\mathrm{X}, \mathsf{d}, \mathfrak{m})$ est un espace $\mathsf{RCD}^*(K,N)$ et la géodésique (μ_t) est constituée par des mesures ayant densités uniformément bornées, alors lorsque $\varepsilon \downarrow 0$:

i) les courbes (μ_t^ε) convergent de façon W_2-uniforme à la géodésique (μ_t) et les mesures μ_t^ε ont densités uniformément bornées ;

ii) les fonctions φ_t^ε sont uniformément lipschitziennes et convergent de manière localement uniforme aussi bien que dans la topologie $W^{1,2}_{loc}$ vers l'unique solution visqueuse (φ_t) de (0.0.5) avec φ comme donnée initiale ; en particulier, l'équation de continuité (0.0.1) est satisfaite par la courbe limite.

Ces résultats de convergence sont basés sur les estimations de Hamilton pour le gradient et sur l'inégalité de Li-Yau et ils sont suffisants pour passer à la limite dans le terme avec la Hessienne dans (0.0.4). Pour ces courbes l'accélération est donnée par $a_t^\varepsilon = -\frac{\varepsilon}{2}\Delta\varphi_t^\varepsilon$ et donc il ne reste à prouver que la quantité

$$\varepsilon \int \langle \nabla f, \nabla\Delta\varphi_t^\varepsilon \rangle \, \mathrm{d}\mu_t^\varepsilon$$

disparaît à la limite en un certain sens. Cependant, il se trouve qu'il n'y a pas d'espoir d'obtenir cela à travers des techniques fondées sur les EDP. Le problème est dû au fait que ce type d'approximation visqueuse peut produire à la limite une courbe qui n'est pas une géodésique si φ n'est pas c-concave : en peu de mots, cela arrive dès qu'un choc se produit dans l'équation de Hamilton-Jacobi. Comme on ne peut pas espérer que la formule (0.0.2) est vraie pour des courbes qui ne sont pas géodésiques, on voit bien que très difficilement il est possible d'obtenir le résultat souhaité à travers des telles approximations visqueuses.

Pour cette raison on va utiliser une façon différente de se rapprocher aux géodésiques : le ralentissement des interpolations entropiques. Décrivons brièvement en quoi cela consiste, en se plaçant dans le plus familier cadre euclidien.

Fixons deux mesures de probabilité $\mu_0 = \rho_0 \mathcal{L}^d$, $\mu_1 = \rho_1 \mathcal{L}^d$ sur $\mathbb{R}^d$. Les équations fonctionnelles de Schrödinger sont les suivantes

$$\rho_0 = f\, \mathsf{h}_1 g \qquad\qquad \rho_1 = g\, \mathsf{h}_1 f, \tag{0.0.6}$$

les inconnues étant les fonctions boreliennes $f, g : \mathbb{R}^d \to [0, \infty)$, où $\mathsf{h}_t f$ est le flot de la chaleur démarrant en f et évalué à l'instant t. En grande généralité, ces équations admettent une solution, qui est unique à la transformtion triviale $(f, g) \mapsto (cf, g/c)$ près, pour une certaine constante $c > 0$. Cette solution peut être trouvée de la manière suivante : soit R la mesure sur $(\mathbb{R}^d)^2$ dont la densité par rapport à $\mathcal{L}^{2d}$ est donnée par le noyau de la chaleur $\mathsf{r}_t(x, y)$ à l'instant $t = 1$ et minimisons l'entropie de Boltzmann-Shannon $H(\boldsymbol{\gamma} \,|\, \mathsf{R})$ parmi tous les couplages $\boldsymbol{\gamma}$ de μ_0 à μ_1. L'équation d'Euler pour le minimiseur force celui-ci à avoir la forme $f \otimes g\, \mathsf{R}$ pour des fonctions boreliennes $f, g : \mathbb{R}^d \to [0, \infty)$ appropriées, où $f \otimes g(x, y) := f(x)g(y)$ (on montrera à nouveau ce résultat déjà connu dans la Proposition 4.1.5).

Une fois que l'on a trouvé la solution de (0.0.6) on peut l'utiliser conjointement avec le flot de la chaleur, afin d'interpoler de ρ_0 à ρ_1 en définissant

$$\rho_t := \mathsf{h}_t f\, \mathsf{h}_{1-t} g.$$

Ceci est appelé interpolation entropique. Maintenant on ralentit le flot de la chaleur : fixons $\varepsilon > 0$ et en imitant ce que l'on vient de présenter trouvons $f^\varepsilon, g^\varepsilon$ telles que

$$\rho_0 = f^\varepsilon\, \mathsf{h}_{\varepsilon/2} g^\varepsilon \qquad\qquad \rho_1 = g^\varepsilon\, \mathsf{h}_{\varepsilon/2} f^\varepsilon, \tag{0.0.7}$$

(le facteur 1/2 ne joue aucun rôle special, mais il est utile pour les calculs). Puis, définissons

$$\rho_t^\varepsilon := \mathsf{h}_{t\varepsilon/2} f^\varepsilon\, \mathsf{h}_{(1-t)\varepsilon/2} g^\varepsilon.$$

Ce qui est remarquable et non trivial, c'est que lorsque $\varepsilon \downarrow 0$ les courbes de mesures $(\rho_t^\varepsilon \mathcal{L}^d)$ convergent à la géodésique Wasserstein de μ_0 à μ_1.

Les premiers liens entre les équations de Schrödinger et le transport optimal ont été découverts par Mikami dans [91] pour le coût quadratique sur $\mathbb{R}^d$; ensuite, Mikami et Thieullen [97] ont montré que le lien persiste même pour des coûts plus généraux. L'énoncé qu'on vient de présenter au sujet de la convergence des interpolations entropiques à celles par déplacement a été prouvé par Léonard dans [79]. En fait, Léonard a travaillé dans un cadre beaucoup plus abstrait et général : comme c'est peut-être évident à en juger par la présentation, la construction des interpolations entropiques peut être mise en œuvre en grande généralité, parce qu'il suffit d'avoir un noyau de la chaleur. Il a aussi fourni une intuition élémentaire pour expliquer pourquoi cette convergence a lieu : l'idée essentielle est que si le noyau de la chaleur admet l'expansion asymptotique $\varepsilon \log \mathsf{r}_\varepsilon(x, y) \sim -\frac{\mathsf{d}^2(x,y)}{2}$ (au sens des grandes déviations), alors les fonctionnelles d'entropie rééchelonnées $\varepsilon H(\cdot \,|\, \mathsf{R}_\varepsilon)$ convergent à $\frac{1}{2} \int \mathsf{d}^2(x, y)\, \mathrm{d}\cdot$ (au sens de la Γ-convergence). Le lecteur est adressé à [81] pour une dissertation plus profonde de cet argument, pour des remarques historiques et tout renseignement complémentaire.

En partant de ces intuitions et résultats et travaillant dans le cadre des espaces $\mathsf{RCD}^*(K, N)$ on obtient des nouvelles informations sur la convergence des interpolations entropiques vers celles par déplacement. Afin d'énoncer nos résultats, il est plus pratique d'introduire les potentiels de Schrödinger $\varphi_t^\varepsilon, \psi_t^\varepsilon$ comme étant

$$\varphi_t^\varepsilon := \varepsilon \log \mathsf{h}_{t\varepsilon/2} f^\varepsilon \qquad\qquad \psi_t^\varepsilon := \varepsilon \log \mathsf{h}_{(1-t)\varepsilon/2} g^\varepsilon.$$

À la limite pour $\varepsilon \downarrow 0$, ils convergent vers des potentiels de Kantorovich 'forward' et 'backward' le long de la géodésique limite (μ_t), voir aussi en bas. Dans cette direction, il vaut la peine de noter que, alors que pour $\varepsilon > 0$ il y a un lien étroit entre les potentiels et les densités, car

$$\varphi_t^\varepsilon + \psi_t^\varepsilon = \varepsilon \log \rho_t^\varepsilon,$$

à la limite cela devient la célèbre (et plus faible) relation entre les potentiels 'forward' et 'backward' et les mesures (μ_t) :

$$\begin{aligned} \varphi_t + \psi_t &= 0 \qquad \text{on } \operatorname{supp}(\mu_t), \\ \varphi_t + \psi_t &\leq 0 \qquad \text{on X}, \end{aligned}$$

voir par exemple la Remark 7.37 dans [121] (en faisant attention aux signes différents). Par des calculs directs on peut verifier que $(\varphi_t^\varepsilon), (\psi_t^\varepsilon)$ résolvent les équations de Hamilton-Jacobi-Bellman

$$\frac{\mathrm{d}}{\mathrm{d}t}\varphi_t^\varepsilon = \frac{1}{2}|\nabla\varphi_t^\varepsilon|^2 + \frac{\varepsilon}{2}\Delta\varphi_t^\varepsilon \qquad\qquad -\frac{\mathrm{d}}{\mathrm{d}t}\psi_t^\varepsilon = \frac{1}{2}|\nabla\psi_t^\varepsilon|^2 + \frac{\varepsilon}{2}\Delta\psi_t^\varepsilon, \tag{0.0.8}$$

donc, en introduisant les fonctions

$$\vartheta_t^\varepsilon := \frac{\psi_t^\varepsilon - \varphi_t^\varepsilon}{2}$$

il n'est pas difficile de contrôler que

$$\frac{\mathrm{d}}{\mathrm{d}t}\rho_t^\varepsilon + \operatorname{div}(\nabla\vartheta_\mathrm{t}^\varepsilon\,\rho_\mathrm{t}^\varepsilon) = 0 \tag{0.0.9}$$

est satisfait et

$$\frac{\mathrm{d}}{\mathrm{d}t}\vartheta_t^\varepsilon + \frac{|\nabla\vartheta_t^\varepsilon|^2}{2} = a_t^\varepsilon, \qquad \text{where} \qquad a_t^\varepsilon := -\frac{\varepsilon^2}{8}\Big(2\Delta\log\rho_t^\varepsilon + |\nabla\log\rho_t^\varepsilon|^2\Big).$$

Ceci dit, nos résultats principaux concernant les interpolations entropiques peuvent être résumés de la façon suivante. Sous les hypothèses que l'espace métrique mesuré est $\mathsf{RCD}^*(K,N)$, $N < \infty$, et ρ_0, ρ_1 appartiennent à $L^\infty(\mathrm{X})$ on a :

- <u>Ordre zéro</u>
  - *borne* Pour quelque $C > 0$ on a $\rho_t^\varepsilon \leq C\mathfrak{m}$ pour tout $\varepsilon \in (0,1)$ et $t \in [0,1]$ (Proposition 5.2.2).
  - *convergence* Les courbes $(\rho_t^\varepsilon\mathfrak{m})$ convergent de manière W_2-uniforme vers l'unique géodésique Wasserstein (μ_t) de μ_0 à μ_1 (Propositions 6.1.1 and 6.2.2).
- <u>Premier ordre</u>
  - *borne* Pour tout $t \in (0,1]$ les fonctions $\{\varphi_t^\varepsilon\}_{\varepsilon\in(0,1)}$ sont localement uniformement lipschitziens (Proposition 5.2.3). Le même pour les ψ.
  - *convergence* Pour toute suite $\varepsilon_n \downarrow 0$ il existe une sous-suite – non réétiquetée – telle que pour tout $t \in (0,1]$ les fonctions φ_t^ε convergent de manière localement uniforme aussi bien que dans la topologie $W^{1,2}_{loc}(\mathrm{X})$ vers une fonction φ_t telle que $-t\varphi_t$ est un potential de Kantorovich de μ_t à μ_0 (voir les Propositions 6.1.1, 6.2.2 et 6.2.6 pour la formulation précise des résultats). De même pour les ψ.

- <u>Deuxième ordre</u> Pour tout $\delta \in (0, 1/2)$ on a

  – *borne*

$$\sup_{\varepsilon\in(0,1)} \iint_\delta^{1-\delta} \big(|\mathrm{Hess}(\vartheta_t^\varepsilon)|^2_{\mathsf{HS}} + \varepsilon^2 |\mathrm{Hess}(\log \rho_t^\varepsilon)|^2_{\mathsf{HS}}\big) \rho_t^\varepsilon \,\mathrm{d}t\,\mathrm{d}\mathfrak{m} < \infty, \tag{0.0.10}$$
$$\sup_{\varepsilon\in(0,1)} \iint_\delta^{1-\delta} \big(|\Delta \vartheta_t^\varepsilon|^2 + \varepsilon^2 |\Delta \log \rho_t^\varepsilon|^2\big) \rho_t^\varepsilon \,\mathrm{d}t\,\mathrm{d}\mathfrak{m} < \infty,$$

  (Lemme 5.3.4). Observons que, puisque en général le laplacien n'est pas la trace de la Hessienne, il n'y a pas de liens directs entre ces deux bornes.

  – *convergence* Pour toute fonction $h \in W^{1,2}(\mathrm{X})$ avec $\Delta h \in L^\infty(\mathrm{X})$ on a

$$\lim_{\varepsilon\downarrow 0} \iint_\delta^{1-\delta} \langle \nabla h, \nabla a_t^\varepsilon \rangle \, \rho_t^\varepsilon \,\mathrm{d}t\,\mathrm{d}\mathfrak{m} = 0, \tag{0.0.11}$$

  (Theorem 7.1.2).

À l'exception de la convergence $\rho_t^\varepsilon \mathfrak{m} \to \mu_t$, tous ces résultats sont nouveaux, même sur les variétés lisses compactes (même sur le tore plat). Les bornes d'ordre zéro et du premier ordre sont conséquences des équations de Hamilton-Jacobi-Bellman (0.0.8) satisfaites par les φ et les ψ et peuvent être obtenues à partir de l'estimation de Hamilton pour le gradient et de l'inégalité de Li-Yau. Les faits que la courbe limite est la géodésique Wasserstein entre μ_0 et μ_1 et que les potentiels limite sont des potentiels de Kantorovich, ce sont conséquences du fait que l'on peut passer à la limite dans l'équation de continuité (0.0.9) et que les potentiels limite satisfont l'équation de Hamilton-Jacobi. À cet égard, il est crucial que l'on se rapproche au même instant du potentiel 'forward' ψ et de celui 'backward' φ : le lecteur peut regarder la preuve de la Proposition 6.2.2 et se souvenir du fait que les approximations visqueuses 'simples' (non pas entropiques) peuvent bien converger vers des courbes qui ne sont pas géodésiques Wasserstein.

Ces résultat de convergence d'ordre zéro et un sont suffisants pour passer à la limite dans le terme avec la Hessienne dans (0.0.4).

Comme on l'a déjà remarqué avant, l'approximation visqueuse peut produire le même type de convergence. L'avantage principal et caractéristique provenant du fait que l'on s'appuie sur les interpolations entropiques apparaît dans le résultat de convergence du deuxième ordre (0.0.11), qui montre en quel sens le terme avec l'accélération dans (0.0.4) disparaît à la limite. Par conséquent, cela nous permettra de montrer le résultat principal, c'est à dire le Théorème 7.1.2. Dans cette direction, signalons de manière informelle que, étant l'équation des géodésiques une équation du deuxième ordre, quand on cherche une procédure d'approximation il est naturel de tâcher d'en trouver une qui produise une convergence jusqu'au deuxième ordre.

La propriété de convergence (0.0.11) est pour la plupart une conséquence – bien que non triviale – de la borne (0.0.10) (voir en particulier le Lemme 5.3.5 et la preuve du Théorème 7.1.2), donc concentrons-nous sur la manière pour obtenir (0.0.10). Le point de démarrage est une formule montrée par Léonard dans [76] ; il s'est aperçu de la connexion qui existe entre interpolations entropiques et bornes inférieures sur la courbure de Ricci : il a calculé explicitement la dérivée d'ordre deux de l'entropie le long des interpolations entropiques, en

obtenant

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2} H(\rho_t^\varepsilon \mathfrak{m} \,|\, \mathfrak{m}) = \int \big(\Gamma_2(\vartheta_t^\varepsilon) + \tfrac{\varepsilon^2}{4}\Gamma_2(\log \rho_t^\varepsilon)\big)\rho_t^\varepsilon \,\mathrm{d}\mathfrak{m}, \tag{0.0.12}$$

où Γ_2 est l'opérateur carré du champ itéré, défini comme étant

$$\Gamma_2(f) := \Delta \frac{|\nabla f|^2}{2} - \langle \nabla f, \nabla \Delta f \rangle$$

(dans le cadre des espaces RCD il faut faire attention dans le traitement de cet objet, parce qu'en général $\Gamma_2(f)$ est seulement une mesure, mais pour l'instant oublions cet aspect).

Donc, si l'on est sur une variété avec courbure de Ricci non négative, alors l'inégalité de Bochner

$$\Gamma_2(f) \geq |\mathrm{Hess}(f)|^2_{\mathsf{HS}} \tag{0.0.13}$$

assure que l'entropie est convexe le long des interpolations entropiques.

Maintenant observons que si $f : [0,1] \to [0,\infty)$ est convexe, alors pour $t \in (0,1)$ la quantité $|f'(t)|$ peut être contrôlée en termes de $f(0), f(1)$ et t seulement. Pour cette raison, comme la valeur de $H(\rho_t^\varepsilon \mathfrak{m} \,|\, \mathfrak{m})$ aux instants $t = 0, 1$ est indépendante de $\varepsilon > 0$, on obtient la borne uniforme

$$\sup_{\varepsilon>0} \int_\delta^{1-\delta} \frac{\mathrm{d}^2}{\mathrm{d}t^2} H(\mu_t^\varepsilon \,|\, \mathfrak{m}) \,\mathrm{d}t = \sup_{\varepsilon>0} \left(\frac{\mathrm{d}}{\mathrm{d}t} H(\mu_t^\varepsilon \,|\, \mathfrak{m})_{|t=1-\delta} - \frac{\mathrm{d}}{\mathrm{d}t} H(\mu_t^\varepsilon \,|\, \mathfrak{m})_{|t=\delta} \right) < \infty$$

qui par (0.0.13) et (0.0.12) garantit la première borne dans (0.0.10). La deuxième est déduite de manière similaire, en utilisant l'inégalité de Bochner suivante

$$\Gamma_2(f) \geq \frac{(\Delta f)^2}{N}$$

à la place de (0.0.13).

Structure de la thèse

Bien qu'elle soit très importante, la formule de dérivation du deuxième ordre n'est pas le seul aspect saillant auquel on s'est intéressé et c'est pourquoi il vaut la peine d'ajouter quelques mots sur l'organisation de la thèse, les différents thèmes envisagés et tous les résultats sécondaires que l'on n'a pas encore mentionnés.

La thèse est divisée en trois parties, trois étant aussi les domaines mathématiques principaux qui interviennent de manière plus ou moins assidue dans le cours des pages suivantes : l'analyse, la géométrie et les probabilités. Bien que très liées entre elles, les trois parties sont caractérisées par la prédominance d'un domain sur les autres et cela est bien évident à en juger par les titres choisis pour chaque partie. Il vaut la peine de dire quelques mots sur les motivations derrière un tel choix et, surtout, sur l'organisation du manuscrit. Comme on l'a déjà remarqué avant, la motivation originelle qui est à la base du présent travail de recherche est essentiellement géométrique, mais il est évident que pour sa preuve l'analyse est un ingrédient incontournable. Pour cette raison dans la première partie, après une introduction préliminaire

ayant comme but celui de donner au lecteur tous les notions et résultats sur le transport optimal et la théorie des espaces RCD nécessaires pour la compréhension de l'œuvre, on s'occupe de l'approche basée sur des techniques EDP et suggéré toute à l'heure : il n'est pas surprenant de voir qu'un rôle clé est joué par le logarithme des solutions de l'équation de la chaleur (parce que cela donne une solution de l'équation de Hamilton-Jacobi-Bellman) et par la manière dont on peut l'estimer. Dans la deuxième partie on traite le problème de Schr »odinger et donc on aborde le thème des grandes déviations ; bien que le langage adopté soit analytique, parce qu'il s'adapte mieux au contexte des espaces RCD, le sujet est spécifiquement probabiliste dans l'esprit et nous donne les outils corrects pour approximer les géodésiques Wasserstein : interpolations entropiques et potentiels de Schrödinger. Finalement, dans la troisième partie on fusionne les résultats des chapitres précédents, ce qui aboutit à donner une réponse affirmative au problème de départ, c'est à dire :

> Est-ce que l'on peut établir la formule de dérivation du deuxième ordre le long des géodésiques ?

En plus, on examine les informations analytiques et géométriques cachées dans les interpolations entropiques et fournit des applications géométriques de la formule de dérivation du deuxième ordre. Ainsi, les trois domaines sur lesquels cette thèse est fondée résulteront liés entre eux de manière encore plus stricte et forte.

Maintenant on va détailler avec plus de précision les trois parties surmentionnés

Partie I.
La théorie du transport optimal tire son origine des travaux d'ingénieur de Gaspard Monge, dans le XVIIIe siècle, et à travers les décennies l'intérêt pour ce domaine ne s'est jamais épuisé ; par contre, il s'est de plus en plus enrichi, avec les contributions de Kantorovich dans les Années Quarante du XXe siècle et il a progressé jusqu'à nos jours. Par contre, beaucoup plus récent est l'étude des bornes synthétiques sur la courbure et la dimension, une branche des mathématiques qui s'intéresse à caractériser la structure géométrique des variétés riemanniennes en termes de géométrie métrique et théorie de la mesure, ce qui permet d'étendre les notions de (bornes sur la) courbure de Ricci à des cadres beaucoup plus généraux et abstraits que les variétés riemanniennes.

Comme on l'a déjà remarqué, les deux domaines sont liés entre eux et pour cette raison dans le Chapitre 1 le lecteur peut trouver un guide aux définitions et résultats incontournables pour la compréhension du manuscrit. Dans un *crescendo*, on passera de la structure différentielle du premier ordre dont les espaces polonais peuvent être munis à celle du deuxième ordre, qui par contre nécessite de la condition RCD, à travers des énoncés de plus en plus détaillés sur le transport optimal.

Le Chapitre 2 est par contre entièrement consacré à l'analyse sur les espaces RCD et un rôle clé y est joué par les estimations gaussiennes sur le noyau de la chaleur, qui interviennent à plusieurs reprises pour montrer les inégalités de Hamilton et de Li-Yau. Ces théorèmes étaient déjà connus, même sur les espaces RCD, mais on en donne la preuve dans le cas compact parce que presque la même stratégie du cas lisse peut être adopté ; en plus, on en déduit des résultats qui, eux, sont nouveaux et nécessaires pour la suite.

Partie II.
En 1931 Erwin Schrödinger proposa un nouveau problème d'interpolation qui, dès son apparition, montra des analogies frappantes avec la mécanique ondulatoire, qui venait d'être

découverte, et l'équation de Schrödinger ; des analogies beaucoup plus marquées et évidentes que celles encodées dans l'équation de Fokker-Planck et dans les études de Smoluchowski sur le mouvement brownien. Cependant, malgré les nombreux défis proposés par Schrödinger dans [110], ils ne devinrent pas aussi célèbres que l'équation nommée d'après lui ; au contraire, le problème d'interpolation fut presque entièrement oublié, même redécouvert plusieurs ans après, et ce n'est que récemment qu'un intérêt diffus pour la thématique a apparu. Cette deuxième partie veut donc être en premier lieu une présentation historique du problème, avec fréquents aperçus sur son interprétation physique, et en deuxième instance un guide d'utilisation pour les non probabilistes ; en fait, tout le langage probabiliste traditionnellement employé dans la description du problème est ici traduit dans un langage analytique, de sorte qu'un plus vaste public puisse en bénéficier.

Pour cette raison dans le Chapitre 3 on éclaircira les décennies écoulées entre 1931 et nos jours. En ce qui concerne l'histoire du problème, le point de départ est donné par la formulation originelle, la motivation et l'interprétation physiques du phénomène (voir [110], [111], l'étude [81] et les monographies [100], [102]). Souvent, on utilisera les mots de Schrödinger à cause de leur puissance enrichissante, éclairante et suggestive et, à travers un raisonnement informel enraciné dans la physique statistique, les équations fonctionnelles de Schrödinger se traduiront dans la formulation de Föllmer, où un problème de minimisation entropique apparaît. Comme remarque finale, plusieurs développements et applications importants sont rappelés.

Une discussion mathématique sur le problème de Schrödinger est au centre des Chapitre 4 et 5. Dans le premier, la version de Föllmer du problème, ébauchée de manière informelle précédemment, est déduite rigoureusement et énoncée proprement sur un espace polonais quelconque. En plus, les versions dynamique et duale du problème sont aussi introduites. Pour elles, les résultats classiques d'existence sont énoncés et les liens entre les trois formulations différentes sont abordés. En outre, bien que moins général que les résultats déjà connus dans la littérature, on fournit un théorème d'existence qui est partiellement nouveau ; la spécificité du résultat consiste en la caractérisation de l'unique minimiseur du problème de Schrödinger parmi tous les couplages associés aux contraints marginales. Dans la dernière partie du chapitre, en imitant certains résultats classiques de la théorie du transport optimal, comme par exemple la convexité du coût de transport, la propriété de restriction et la stabilité, on met en place une boîte à outils pour le problème de Schrödinger qui, à notre connaissance, est encore manquante.

Dans le Chapitre 5 on abandonne le cadre des espaces polonais pour se placer dans celui des espaces RCD, où les bornes sur la courbure et la dimension déterminent une connaissance plus profonde des interpolations entropiques. Pierres angulaires de ce chapitre, ce sont les suivantes : une borne uniforme pour les densités des interpolations entropiques (Proposition 5.2.2), la lipschitzianité localement uniforme des potentiels de Schrödinger leur associés (Proposition 5.2.3), des formules explicites pour les dérivées première et seconde de l'entropie le long des interpolations entropiques (Proposition 5.3.3) et, par conséquent, un contrôle L^2 à poids uniforme pour les Hessiennes de ces potentiels (Lemme 5.3.4). Tous ces résultats, sauf les formules des dérivées du premier et deuxième ordre, sont nouveaux même dans le cas lisse.

Partie III.

Après une première partie d'inspiration analytique et une deuxième d'esprit probabiliste, la troisième et dernière partie du manuscrit est dédiée à la preuve de la formule de dérivation du deuxième ordre et aux applications géométriques des outils que l'on a développés dans les chapitres précédents, notamment les inégalités de Hamilton et de Li-Yau pour le côté

analytique et tous les bornes (localement) uniformes au sujet des interpolations entropiques et des potentiels de Schrödinger leur associés pour le côté à priori probabiliste. En montrant de quelle manière les interpolations entropiques et les potentiels de Schrödinger convergent vers les géodésiques Wasserstein et les potentiels de Kantorovich respectivement, on bâtira un pont solide entre le problème de Schrödinger et le transport optimal qui nous permettra non seulement de franchir la cible principale de ce projet de recherche, mais encore de pousser plus loin les analogies entre les deux problèmes de minimisation, découvrant ainsi des liens inespérés, dont on va parler bientôt.

Le Chapitre 6 s'ouvre avec l'application directe et immédiate des Propositions 5.2.2 et 5.2.3 : grâce aux bornes uniformes y énoncées on a assez de compacité à disposition pour montrer l'existence d'une courbe limite de mesures et des potentiels limite. Ensuite, la question naturelle qui se présente est la suivante : qu'est-ce que l'on peut dire de ces trajectoires limite ? Est-ce que l'on peut les caractériser ? La réponse est affirmative et l'on voit que la courbe de mesures limite est une géodésique W_2 et donc elle est unique, étant uniques les géodésiques dans l'espace de Wasserstein bâti sur un espace RCD ; en ce qui concerne les potentiels, ils convergent vers les solutions visqueuses de l'équation de Hamilton-Jacobi, par analogie avec le cas lisse, ce qui donne un aperçu des possibles développements de la théorie de la viscosité sur les espaces métriques. Les résultats de convergence qu'on vient de présenter permettent d'obtenir, d'une manière qu'on peut dire 'entropique', des faits déjà connus (par exemple la (K, N)-convexité de l'entropie le long des interpolations par déplacement et l'inégalité HWI) sur les espaces RCD ; pourtant, il vaut la peine de donner ces application, parce qu'elles sont déduites avec peu d'effort, ce qui est une conséquence directe des bonnes propriétés de régularité des interpolations entropiques.

Dans le Chapitre 7, en s'appuyant sur tous les résultats précédents on donne la preuve du théorème principal, c'est à dire la formule de dérivation du deuxième ordre pour les géodésiques Wasserstein. Une première conséquence du résultat est une formule de dérivation du premier ordre pour des champs de vecteurs (voir Corollaire 7.1.3), mais d'autres applications surgissent dans le domaine de l'analyse géométrique : on expliquera comment la formule de dérivation du deuxième ordre intervient dans le théorème de splitting, en simplifiant la stratégie de preuve.

Introduction

In the last ten years there has been a great interest in the study of metric measure spaces with Ricci curvature bounded from below, as witnessed by the flourishing literature. For a non exhaustive list of the most important results on the subject, the reader can see for instance [87], [115], [116], [60], [6], [7], [58], [8], [105], [106], [61], [55], [57], [73], [10], [98], [26], [25]. The starting points of this research line have been the seminal papers [87] and [115], [116] which linked lower Ricci bounds on metric measure spaces to properties of entropy-like functionals in connection with W_2-geometry. Later ([6]) it emerged that also Sobolev calculus is linked to W_2-geometry and building on top of this the original definition of CD spaces by Lott-Sturm-Villani has evolved into that of RCD spaces ([7], [58]).

An example of link between Sobolev calculus and W_2-geometry is the following statement, proved in [55]:

Theorem 0.0.4 (First order differentiation formula)**.** *Let* $(\mathrm{X}, \mathsf{d}, \mathfrak{m})$ *be a* $\mathsf{RCD}(K, \infty)$ *space,* (μ_t) *a* W_2*-geodesic made of measures with bounded support and such that* $\mu_t \leq C\mathfrak{m}$ *for every* $t \in [0,1]$ *and some* $C > 0$*. Then for every* $f \in W^{1,2}(\mathrm{X})$ *the map*

$$[0,1] \ni t \quad \mapsto \quad \int f \, \mathrm{d}\mu_t$$

is C^1 *and we have*

$$\frac{\mathrm{d}}{\mathrm{d}t} \int f \, \mathrm{d}\mu_t|_{t=0} = - \int \mathrm{d}f(\nabla \varphi) \, \mathrm{d}\mu_0,$$

where φ *is any locally Lipschitz Kantorovich potential from* μ_0 *to* μ_1*.*

Recall that on $\mathsf{RCD}(K, \infty)$ spaces every W_2-geodesic (μ_t) between measures with bounded density and support is such that $\mu_t \leq C\mathfrak{m}$ for every $t \in [0,1]$ and some $C > 0$ ([106]), so that the theorem also says that we can find 'many' C^1 functions on RCD spaces. We remark that such C^1 regularity - which was crucial in [55] - is non-trivial even if the function f is assumed to be Lipschitz and that statements about C^1 smoothness are quite rare in metric geometry.

One might think at Theorem 0.0.4 as an 'integrated' version of the basic formula

$$\frac{\mathrm{d}}{\mathrm{d}t} f(\gamma_t)_{|t=0} = \mathrm{d}f(\gamma_0')$$

valid in the smooth framework; at the technical level the proof of the claim has to do with the fact that the geodesic (μ_t) solves the continuity equation

$$\frac{\mathrm{d}}{\mathrm{d}t} \mu_t + \mathrm{div}(\nabla(-\varphi_t)\mu_t) = 0, \tag{0.0.14}$$

where the φ_t's are appropriate choices of Kantorovich potentials (see also [59] in this direction).

In [56], Gigli developed a second-order calculus on RCD spaces, in particular defining the space $H^{2,2}(\mathrm{X})$ and for $f \in H^{2,2}(\mathrm{X})$ the Hessian $\mathrm{Hess}(f)$, see [56] and the preliminary section (Section 1.2). It is then natural to ask whether an 'integrated' version of the second order differentiation formula

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2} f(\gamma_t)_{|t=0} = \mathrm{Hess}(f)(\gamma_0', \gamma_0') \qquad \text{for } \gamma \text{ geodesic}$$

holds in this framework. In this manuscript, and more precisely in Chapter 7, we provide affirmative answer to this question, our main result being:

Theorem 0.0.5 (Second order differentiation formula)**.** *Let* $(\mathrm{X}, \mathsf{d}, \mathfrak{m})$ *be a* $\mathsf{RCD}^*(K, N)$ *space, with* $K \in \mathbb{R}$ *and* $N < \infty$*,* (μ_t) *a* W_2*-geodesic such that* $\mu_t \leq C\mathfrak{m}$ *and* μ_t *has compact support for every* $t \in [0, 1]$ *and some* $C > 0$ *and* $f \in H^{2,2}(\mathrm{X})$*.*

Then the function

$$[0, 1] \ni t \quad \mapsto \quad \int f \, \mathrm{d}\mu_t$$

is C^2 *and we have*

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2} \int f \, \mathrm{d}\mu_t{}_{|t=0} = \int \mathrm{Hess}(f)(\nabla\varphi, \nabla\varphi) \, \mathrm{d}\mu_0, \tag{0.0.15}$$

where φ *is any Kantorovich potential from* μ_0 *to* μ_1*.*

See also Theorem 7.1.2 for an alternative, but equivalent, formulation of the result. We wish to stress that based on the kind of arguments used in our proof, we do not believe the compactness assumption to be crucial (but being our proof based on global analysis, to remove it is not a trivial task, especially in the case $K < 0$), while on the other hand the finite dimensionality plays a key role (e.g. because we use the Li-Yau inequality).

Having at disposal such second order differentiation formula - perhaps without the restriction of working in compact spaces - is interesting not only at the theoretical level, but also for applications to the study of the geometry of RCD spaces. For instance, the proofs of both the splitting theorem and of the 'volume cone implies metric cone' in this setting can be greatly simplified by using such formula. Also, one aspect of the theory of RCD spaces which is not yet clear is whether they have constant dimension: for Ricci-limit spaces this is known to be true by a result of Colding-Naber [32] which uses second order derivatives along geodesics in a crucial way. Thus our result is necessary to replicate Colding-Naber argument in the non-smooth setting (but not sufficient: they also use a calculus with Jacobi fields which as of today does not have a non-smooth counterpart).

Strategy of the proof

Our starting point is a related second order differentiation formula obtained in [56], available under proper regularity assumptions:

Theorem 0.0.6. *Let* (μ_t) *be a* W_2*-absolutely continuous curve solving the continuity equation*

$$\frac{\mathrm{d}}{\mathrm{d}t}\mu_t + \mathrm{div}(X_t\mu_t) = 0,$$

for some vector fields $(X_t) \subset L^2(T\mathrm{X})$ *in the following sense: for every* $f \in W^{1,2}(\mathrm{X})$ *the map* $t \mapsto \int f \, \mathrm{d}\mu_t$ *is absolutely continuous and it holds*

$$\frac{\mathrm{d}}{\mathrm{d}t} \int f \, \mathrm{d}\mu_t = \int \langle \nabla f, X_t \rangle \, \mathrm{d}\mu_t.$$

Assume that

(i) $t \mapsto X_t \in L^2(T\mathrm{X})$ *is absolutely continuous,*

(ii) $\sup_t \{ \|X_t\|_{L^2} + \|X_t\|_{L^\infty} + \|\nabla X_t\|_{L^2} \} < +\infty$.

Then for $f \in H^{2,2}(\mathrm{X})$ *the map* $t \mapsto \int f \, \mathrm{d}\mu_t$ *is* $C^{1,1}$ *and the formula*

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2} \int f \mathrm{d}\mu_t = \int \mathrm{Hess}(f)(X_t, X_t) + \left\langle \nabla f, \tfrac{\mathrm{d}}{\mathrm{d}t} X_t + \nabla_{X_t} X_t \right\rangle \mathrm{d}\mu_t \tag{0.0.16}$$

holds for a.e. $t \in [0,1]$.

If the vector fields X_t are of gradient type, so that $X_t = \nabla \phi_t$ for every t and the 'acceleration' a_t is defined as

$$\frac{\mathrm{d}}{\mathrm{d}t} \phi_t + \frac{|\nabla \phi_t|^2}{2} =: a_t$$

then (0.0.16) reads as

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2} \int f \mathrm{d}\mu_t = \int \mathrm{Hess}(f)(\nabla \phi_t, \nabla \phi_t) \, \mathrm{d}\mu_t + \int \langle \nabla f, \nabla a_t \rangle \, \mathrm{d}\mu_t. \tag{0.0.17}$$

In the case of geodesics, the functions φ_t appearing in (0.0.14) solve (in a sense which we will not make precise here) the Hamilton-Jacobi equation

$$\frac{\mathrm{d}}{\mathrm{d}t} \varphi_t = \frac{|\nabla \varphi_t|^2}{2}, \tag{0.0.18}$$

thus in this case the acceleration a_t is identically 0 (notice the minus sign in (0.0.14)). Hence if the vector fields $(\nabla \varphi_t)$ satisfy the regularity requirements $(i), (ii)$ in the last theorem we would easily be able to establish Theorem 0.0.5. However in general this is not the case; informally speaking this has to do with the fact that for solutions of the Hamilton-Jacobi equations we do not have sufficiently strong second order estimates.

In order to establish Theorem 0.0.5 it is therefore natural to look for suitable 'smooth' approximation of geodesics for which we can apply Theorem 0.0.6 above and then pass to the limit in formula (0.0.16). Given that the lack of smoothness of W_2-geodesic is related to the lack of smoothness of solutions of (0.0.18), also in line with the classical theory of viscous approximation for the Hamilton-Jacobi equation there is a quite natural thing to try: solve, for $\varepsilon > 0$, the equation

$$\frac{\mathrm{d}}{\mathrm{d}t} \varphi_t^\varepsilon = \frac{|\nabla \varphi_t^\varepsilon|^2}{2} + \frac{\varepsilon}{2} \Delta \varphi_t^\varepsilon, \qquad \varphi_0^\varepsilon := \varphi,$$

where φ is a given, fixed, Kantorovich potential for the geodesic (μ_t), and then solve

$$\frac{\mathrm{d}}{\mathrm{d}t} \mu_t^\varepsilon - \mathrm{div}(\nabla \varphi_t^\varepsilon \mu_t^\varepsilon) = 0, \qquad \mu_0^\varepsilon := \mu_0.$$

This plan can actually be pursued and following the ideas in this paper one can show that if the space $(\mathrm{X}, \mathsf{d}, \mathfrak{m})$ is compact and $\mathsf{RCD}^*(K,N)$ and the geodesic (μ_t) is made of measures with equibounded densities, then as $\varepsilon \downarrow 0$:

i) the curves (μ_t^ε) W_2-uniformly converge to the geodesic (μ_t) and the measures μ_t^ε have equibounded densities.

ii) the functions φ_t^ε are equi-Lipschitz and converge both uniformly and in the $W^{1,2}$-topology to the only viscous solution (φ_t) of (0.0.18) with φ as initial datum; in particular the continuity equation (0.0.14) for the limit curve holds.

These convergence results are based on Hamilton's gradient estimates and the Li-Yau inequality and are sufficient to pass to the limit in the term with the Hessian in (0.0.17). For these curves the acceleration is given by $a_t^\varepsilon = -\frac{\varepsilon}{2}\Delta\varphi_t^\varepsilon$ and thus we are left to prove that the quantity

$$\varepsilon \int \langle \nabla f, \nabla \Delta \varphi_t^\varepsilon \rangle \, \mathrm{d}\mu_t^\varepsilon$$

goes to 0 in some sense. However, there appears to be *no hope of obtaining this by PDE estimates*. The problem is that this kind of viscous approximation can produce in the limit a curve which is not a geodesic if φ is not c-concave: shortly said, this happens as soon as a shock appears in Hamilton-Jacobi. Since there is no hope for formula (0.0.15) to be true for non-geodesics, we see that there is little chance of obtaining it via such viscous approximation.

We therefore use another way of approximating geodesics: the slowing down of entropic interpolations. Let us briefly describe what this is in the familiar Euclidean setting.

Fix two probability measures $\mu_0 = \rho_0 \mathcal{L}^d$, $\mu_1 = \rho_1 \mathcal{L}^d$ on $\mathbb{R}^d$. The Schrödinger functional equations are

$$\rho_0 = f\, \mathsf{h}_1 g \qquad\qquad \rho_1 = g\, \mathsf{h}_1 f, \tag{0.0.19}$$

the unknown being the Borel functions $f, g : \mathbb{R}^d \to [0, \infty)$, where $\mathsf{h}_t f$ is the heat flow starting at f evaluated at time t. It turns out that in great generality these equations admit a solution which is unique up to the trivial transformation $(f, g) \mapsto (cf, g/c)$ for some constant $c > 0$. Such solution can be found in the following way: let R be the measure on $(\mathbb{R}^d)^2$ whose density w.r.t. $\mathcal{L}^{2d}$ is given by the heat kernel $\mathsf{r}_t(x, y)$ at time $t = 1$ and minimize the Boltzmann-Shannon entropy $H(\boldsymbol{\gamma} \,|\, \mathsf{R})$ among all transport plans $\boldsymbol{\gamma}$ from μ_0 to μ_1. The Euler equation for the minimizer forces it to be of the form $f \otimes g\, \mathsf{R}$ for some Borel functions $f, g : \mathbb{R}^d \to [0, \infty)$, where $f \otimes g(x, y) := f(x)g(y)$ (we shall reprove this known result in Proposition 4.1.5). Then the fact that $f \otimes g\, \mathsf{R}$ is a transport plan from μ_0 to μ_1 is equivalent to (f, g) solving (0.0.19).

Once we have found the solution of (0.0.19) we can use it in conjunction with the heat flow to interpolate from ρ_0 to ρ_1 by defining

$$\rho_t := \mathsf{h}_t f\, \mathsf{h}_{1-t} g.$$

This is called entropic interpolation. Now we slow down the heat flow: fix $\varepsilon > 0$ and by mimicking the above find $f^\varepsilon, g^\varepsilon$ such that

$$\rho_0 = f^\varepsilon\, \mathsf{h}_{\varepsilon/2} g^\varepsilon \qquad\qquad \rho_1 = g^\varepsilon\, \mathsf{h}_{\varepsilon/2} f^\varepsilon, \tag{0.0.20}$$

(the factor 1/2 plays no special role, but is convenient in computations). Then define

$$\rho_t^\varepsilon := \mathsf{h}_{t\varepsilon/2} f^\varepsilon\, \mathsf{h}_{(1-t)\varepsilon/2} g^\varepsilon.$$

The remarkable and non-trivial fact here is that as $\varepsilon \downarrow 0$ the curves of measures $(\rho_t^\varepsilon \mathcal{L}^d)$ converge to the W_2-geodesic from μ_0 to μ_1.

The first connections between Schrödinger equations and optimal transport have been obtained by Mikami in [91] for the quadratic cost on $\mathbb{R}^d$; later Mikami-Thieullen [97] showed that a link persists even for more general cost functions. The statement we have just made about convergence of entropic interpolations to displacement ones has been proved by Léonard in [79]. Actually, Léonard worked in much higher generality: as it is perhaps clear from the presentation, the construction of entropic interpolation can be done in great generality, as only a heat kernel is needed. He also provided a basic intuition about why such convergence is in place: the basic idea is that if the heat kernel admits the asymptotic expansion $\varepsilon \log \mathsf{r}_\varepsilon(x,y) \sim -\frac{\mathsf{d}^2(x,y)}{2}$ (in the sense of Large Deviations), then the rescaled entropy functionals $\varepsilon H(\cdot \,|\, \mathsf{R}_\varepsilon)$ converge to $\frac{1}{2}\int \mathsf{d}^2(x,y)\,\mathrm{d}\cdot$ (in the sense of Γ-convergence). We refer to [81] for a deeper discussion of this topic, historical remarks and much more.

Starting from these intuitions and results, working in the setting of compact $\mathsf{RCD}^*(K,N)$ spaces we gain new information about the convergence of entropic interpolations to displacement ones. In order to state our results, it is convenient to introduce the Schrödinger potentials $\varphi_t^\varepsilon, \psi_t^\varepsilon$ as

$$\varphi_t^\varepsilon := \varepsilon \log \mathsf{h}_{t\varepsilon/2} f^\varepsilon \qquad\qquad \psi_t^\varepsilon := \varepsilon \log \mathsf{h}_{(1-t)\varepsilon/2} g^\varepsilon.$$

In the limit $\varepsilon \downarrow 0$ these will converge to forward and backward Kantorovich potentials along the limit geodesic (μ_t) (see below). In this direction, it is worth to notice that while for $\varepsilon > 0$ there is a tight link between potentials and densities, as we trivially have

$$\varphi_t^\varepsilon + \psi_t^\varepsilon = \varepsilon \log \rho_t^\varepsilon,$$

in the limit this becomes the well known (weaker) relation that is in place between forward/backward Kantorovich potentials and measures (μ_t):

$$\begin{aligned} \varphi_t + \psi_t &= 0 \qquad \text{on } \operatorname{supp}(\mu_t), \\ \varphi_t + \psi_t &\le 0 \qquad \text{on X}, \end{aligned}$$

see e.g. Remark 7.37 in [121] (paying attention to the different sign convention). By direct computation one can verify that $(\varphi_t^\varepsilon), (\psi_t^\varepsilon)$ solve the Hamilton-Jacobi-Bellman equations

$$\frac{\mathrm{d}}{\mathrm{d}t}\varphi_t^\varepsilon = \frac{1}{2}|\nabla \varphi_t^\varepsilon|^2 + \frac{\varepsilon}{2}\Delta \varphi_t^\varepsilon \qquad\qquad -\frac{\mathrm{d}}{\mathrm{d}t}\psi_t^\varepsilon = \frac{1}{2}|\nabla \psi_t^\varepsilon|^2 + \frac{\varepsilon}{2}\Delta \psi_t^\varepsilon, \tag{0.0.21}$$

thus introducing the functions

$$\vartheta_t^\varepsilon := \frac{\psi_t^\varepsilon - \varphi_t^\varepsilon}{2}$$

it is not hard to check that it holds

$$\frac{\mathrm{d}}{\mathrm{d}t}\rho_t^\varepsilon + \operatorname{div}(\nabla \vartheta_\mathrm{t}^\varepsilon\, \rho_\mathrm{t}^\varepsilon) = 0 \tag{0.0.22}$$

and

$$\frac{\mathrm{d}}{\mathrm{d}t}\vartheta_t^\varepsilon + \frac{|\nabla \vartheta_t^\varepsilon|^2}{2} = a_t^\varepsilon, \qquad \text{where} \qquad a_t^\varepsilon := -\frac{\varepsilon^2}{8}\Big(2\Delta \log \rho_t^\varepsilon + |\nabla \log \rho_t^\varepsilon|^2\Big).$$

With this said, our main results about entropic interpolations can be summarized as follows. Under the assumptions that the metric measure space is compact and $\mathsf{RCD}^*(K,N)$, $N < \infty$, and that ρ_0, ρ_1 belong to $L^\infty(\mathrm{X})$ we have:

- Zeroth order
  - *bound* For some $C > 0$ we have $\rho_t^\varepsilon \leq C\mathfrak{m}$ for every $\varepsilon \in (0,1)$ and $t \in [0,1]$ (Proposition 5.2.2).
  - *convergence* The curves $(\rho_t^\varepsilon \mathfrak{m})$ W_2-uniformly converge to the unique W_2-geodesic (μ_t) from μ_0 to μ_1 (Propositions 6.1.1 and 6.2.1).
- First order
  - *bound* For any $t \in (0,1]$ the functions $\{\varphi_t^\varepsilon\}_{\varepsilon\in(0,1)}$ are equi-Lipschitz (Proposition 5.2.3). Similarly for the ψ's.
  - *convergence* For every sequence $\varepsilon_n \downarrow 0$ there is a subsequence - not relabeled - such that for any $t \in (0,1]$ the functions φ_t^ε converge both uniformly and in $W^{1,2}(\mathrm{X})$ to a function φ_t such that $-t\varphi_t$ is a Kantorovich potential from μ_t to μ_0 (see Propositions 6.1.1, 6.2.1 and 6.2.5 for the precise formulation of the results). Similarly for the ψ's.
- Second order For every $\delta \in (0,1/2)$ we have
  - *bound*

$$\begin{aligned} &\sup_{\varepsilon\in(0,1)} \iint_\delta^{1-\delta} \big(|\mathrm{Hess}(\vartheta_t^\varepsilon)|_{\mathsf{HS}}^2 + \varepsilon^2 |\mathrm{Hess}(\log \rho_t^\varepsilon)|_{\mathsf{HS}}^2\big)\rho_t^\varepsilon \,\mathrm{d}t\,\mathrm{d}\mathfrak{m} < \infty, \\ &\sup_{\varepsilon\in(0,1)} \iint_\delta^{1-\delta} \big(|\Delta\vartheta_t^\varepsilon|^2 + \varepsilon^2 |\Delta\log \rho_t^\varepsilon|^2\big)\rho_t^\varepsilon \,\mathrm{d}t\,\mathrm{d}\mathfrak{m} < \infty, \end{aligned} \tag{0.0.23}$$

    (Lemma 5.3.4). Notice that since in general the Laplacian is not the trace of the Hessian, there is no direct link between these two bounds.
  - *convergence* For every function $h \in W^{1,2}(\mathrm{X})$ with $\Delta h \in L^\infty(\mathrm{X})$ it holds

$$\lim_{\varepsilon\downarrow 0} \iint_\delta^{1-\delta} \langle \nabla h, \nabla a_t^\varepsilon\rangle \,\rho_t^\varepsilon \,\mathrm{d}t\,\mathrm{d}\mathfrak{m} = 0, \tag{0.0.24}$$

    (Theorem 7.1.2).

With the exception of the convergence $\rho_t^\varepsilon\mathfrak{m} \to \mu_t$, all these results are new even on compact smooth manifolds (in fact, even in the flat torus). The zeroth and first order bounds are both consequences of the Hamilton-Jacobi-Bellman equations (0.0.21) satisfied by the φ's and ψ's and can be obtained from Hamilton's gradient estimate and the Li-Yau inequality. The facts that the limit curve is the W_2-geodesic and that the limit potentials are Kantorovich potentials are consequence of the fact that we can pass to the limit in the continuity equation (0.0.22) and that the limit potentials satisfy the Hamilton-Jacobi equation. In this regard it is key that we approximate at the same time both the 'forward' potentials ψ and the 'backward' one φ: see the proof of Proposition 6.2.1 and recall that the simple viscous approximation may converge to curves which are not W_2-geodesics.

These zeroth and first order convergences are sufficient to pass to the limit in the term with the Hessian in (0.0.17).

As said, also the viscous approximation could produce the same kind of convergence. The crucial advantage of dealing with entropic interpolations is thus in the second order convergence result (0.0.24) which shows that the term with the acceleration in (0.0.17) vanishes in the limit and thus eventually allows us to prove our main result Theorem 0.0.5. In this direction, we informally point out that being the geodesic equation a second order one, in searching for an approximation procedure it is natural to look for one producing some sort of second order convergence.

The limiting property (0.0.24) is mostly a consequence - although perhaps non-trivial - of the bound (0.0.23) (see in particular Lemma 5.3.5 and the proof of Theorem 7.1.2), thus let us focus on how to get (0.0.23). The starting point here is a formula due to Léonard [76], who realized that there is a connection between entropic interpolation and lower Ricci bounds: he computed the second order derivative of the entropy along entropic interpolations obtaining

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2} H(\rho_t^\varepsilon \mathfrak{m} \,|\, \mathfrak{m}) = \int \big(\Gamma_2(\vartheta_t^\varepsilon) + \tfrac{\varepsilon^2}{4}\Gamma_2(\log \rho_t^\varepsilon)\big)\rho_t^\varepsilon \,\mathrm{d}\mathfrak{m}, \tag{0.0.25}$$

where Γ_2 is the 'iterated carré du champ' operator defined as

$$\Gamma_2(f) := \Delta \frac{|\nabla f|^2}{2} - \langle \nabla f, \nabla \Delta f \rangle$$

(in the setting of RCD spaces some care is needed when handling this object, because $\Gamma_2(f)$ is in general only a measure, but let us neglect this issue here).

Thus if, say, we are on a manifold with non-negative Ricci curvature, then the Bochner inequality

$$\Gamma_2(f) \geq |\mathrm{Hess}(f)|^2_{\mathsf{HS}} \tag{0.0.26}$$

grants that the entropy is convex along entropic interpolations.

Now notice that if $f : [0,1] \to [0,\infty)$ is convex, then for $t \in (0,1)$ the quantity $|f'(t)|$ can be bounded in terms of $f(0), f(1)$ and t only. Thus since the value of $H(\rho_t^\varepsilon \mathfrak{m} \,|\, \mathfrak{m})$ at $t = 0, 1$ is independent on $\varepsilon > 0$, we have the uniform bound

$$\sup_{\varepsilon>0} \int_\delta^{1-\delta} \frac{\mathrm{d}^2}{\mathrm{d}t^2} H(\mu_t^\varepsilon \,|\, \mathfrak{m}) \,\mathrm{d}t = \sup_{\varepsilon>0} \left(\frac{\mathrm{d}}{\mathrm{d}t} H(\mu_t^\varepsilon \,|\, \mathfrak{m})_{|t=1-\delta} - \frac{\mathrm{d}}{\mathrm{d}t} H(\mu_t^\varepsilon \,|\, \mathfrak{m})_{|t=\delta} \right) < \infty$$

which by (0.0.26) and (0.0.25) grants the first in (0.0.23). The second is obtained in a similar way using the Bochner inequality in the form

$$\Gamma_2(f) \geq \frac{(\Delta f)^2}{N}$$

in place of (0.0.26).

Structure of the thesis

Although very important, the second order differentiation formula is not the only prominent aspect we have been working on and for this reason it is worth spending some words on the organization of the thesis, on the different topics and all the secondary results that we have not mentioned yet.

The thesis is divided into three parts, as three are the main mathematical areas more or less involved: analysis, geometry and probability. Although deeply linked together, all the parts are characterized by the predominance of one domain over the others and this is made evident by the choice of the titles. It is worth spending some words on such a choice and, most of all, on the organization of the three parts. As already discussed above, the original motivation that gave rise to the present manuscript is geometric in spirit, but clearly relies on analysis for its proof. For this reason in the first part, after a preliminary introduction providing the reader with all the required notions and results on optimal transport and RCD spaces, we deal with the 'PDE' approach to the problem previously suggested: not surprisingly, a key role is played by the logarithm of solutions to the heat equation (since it is a solution to the Hamilton-Jacobi-Bellman equation) and how to estimate it. In the second part we move to the Schrödinger problem and thus to large deviations; even if an analytical language is adopted, because it better fits the RCD framework, the topic is probabilistic in spirit and provides us with the correct tools for approximating Wasserstein geodesics: entropic interpolations and Schrödinger potentials. Finally, in the third part we merge the results of the previous chapters and succeed to give a positive answer to our main problem, that is:

Can we establish the second order differentiation formula along geodesics?

Besides, we investigate the analytical and geometric information encoded in entropic interpolations and provide geometric applications of the second order differentiation formula, thus linking even more the three already cited domains this thesis is based on.

We now give a more accurate description of the three parts that we have just cited.

Part I. Optimal transport is rooted in Gaspard Monge's seminal works, thus it is a very longstanding theory that dates back to the 18th century and through the decades the interest towards it has never decreased; on the contrary, it has considerably grown thanks to the contribution of Kantorovich in the 1940s and still nowadays is a very active and flourishing research topic. On the contrary, the study of synthetic curvature-dimension bounds is much more recent. Main aim in this field is the characterization of the intrinsic geometric structure of smooth Riemannian manifolds in terms of metric geometry and measure theory via non-differential conditions, thus allowing to extend the notion of Ricci curvature (lower bounds) to abstract and non-smooth framework.

As already said, the two domains are deeply linked together and for this reason in Chapter 1 the reader can find all those definitions and results that are necessary for the understanding of the manuscript. In a *crescendo*, we will pass from the first order differential structure available on Polish spaces to the second order one, which instead requires the RCD condition; in doing so, we will also provide a plethora of optimal transport tools getting more and more sophisticated along with the refinement of the framework we will work within.

On the other hand, Chapter 2 is entirely devoted to analysis on RCD spaces, where a key role is played by the Gaussian estimated on the heat kernel: they come into play several times, especially in the proof of Hamilton's gradient estimate and Li-Yau inequality. These results were already known in the literature, even on RCD spaces, but we give the proof in the compact case because the strategy of the smooth case applies almost verbatim. In addition, we deduce from them some related estimates that on the contrary are new and needed for what follows.

Part II. In 1931 Erwin Schrödinger addressed a new interpolation problem which immediately showed shocking analogies with recently born wave mechanics and the Schrödinger equation, analogies that are much stronger than the ones encoded in the Fokker-Planck equation and in the studies of Smoluchowski on the Brownian motion. However, in spite of the several mathematical challenges proposed by Schrödinger in [110], they did not become as famous as the equation named after him; on the contrary, the interpolation problem was almost forgotten, even rediscovered many years after, and only recently a widespread interest for the topic has spotted. This part aims to be a historical presentation of the problem with physical insights and a user's guide for non-probabilists; in fact, all the probabilistic terminology commonly adopted in the literature is here translated into an analytic language.

For this reason in Chapter 3 we shed light on the decades passed between 1931 and the recent years. As regards the history of the problem, we start with the original formulation, the motivation and the physical interpretation lying behind (see [110], [111], the survey [81] and the monographs [100], [102]). We often quote Schrödinger's words, because of their suggestive and enlightening power, and moving from them and following a statistical physics approach we sketch how Föllmer's formulation as an entropy minimization problem can be deduced. As a concluding remark, various developments and relevant applications are recalled.

A mathematical dissertation about the Schrödinger problem is carried out in Chapter 4 and Chapter 5. In the former, Föllmer's version of the problem is established in a precise way on general Polish spaces as well as a dual and a dynamical formulation. For them we present the basic existence theorems and stress the relationship between static, dynamic and dual solutions; moreover, although less general than the results already known in the literature, we provide the reader with a partially new existence theorem, whose peculiarity is the characterization of the (unique) minimizer of the Schrödinger problem among all transport plans between the marginal constraints. In the last part of the section, mimicking some classical results of optimal transport theory (convexity of the cost, restriction property and stability) we build a toolbox for the Schrödinger problem that, to the best of our knowledge, is still missing in the literature.

In Chapter 5 we move to the RCD framework, where the bounds on both the curvature and the dimension entail a deeper knowledge of the so called entropic interpolations and of the behaviour of the entropy along them. Main achievements of this chapter are equiboundedness of the densities of the entropic interpolations (Proposition 5.2.2), local equi-Lipschitz continuity of the Schrödinger potentials associated to them (Proposition 5.2.3), explicit formulas for the first and second derivative of the entropy along them (Proposition 5.3.3) and, as a byproduct, a uniform weighted L^2 control of the Hessian of such potentials (Lemma 5.3.4). All these results, except for first and second derivative formulas, are new even in the smooth setting.

Part III. After the first part, where analysis was predominant, and the second one with probabilistic insights, in the the third and last part of the manuscript we focus our attention on the proof of the second order differentiation formula and on the geometric applications of the tools we developed throughout the previous chapters (in particular, Hamilton's gradient estimate and Li-Yau inequality on the analytical side and the locally uniform bounds on entropic interpolations and Schrödinger potentials as regards the 'probabilistic' side). By showing that entropic interpolations and Schrödinger potentials converge to Wasserstein geodesics and Kantorovich potentials respectively (making precise in which sense the convergence happens), we are able to build a solid bridge between Schrödinger problem and optimal transport; as a first

outcome, we overcome all the difficulties in proving the second order differentiation formula, thus giving a positive answer to the main question of this research project, and secondly we deeply investigate the analogies between the two minimization problems, discovering new connections.

The *ouverture* of Chapter 6 immediately provides the reader with an application of Propositions 5.2.2 and 5.2.3: thanks to the uniform bounds stated therein we have indeed enough compactness for showing the existence of a limit curve of measures and limit potentials. The question that naturally arises is then the following: what can we say about these limit trajectories? Can we characterize them? The answer is positive: we actually prove that the limit curve of measures is a Wasserstein geodesic, hence unique, being unique the geodesics in $(\mathscr{P}_2(\mathrm{X}), W_2)$ where X is a RCD space; about the potentials, they converge to the viscosity solutions of the Hamilton-Jacobi equation, in complete analogy with the smooth case: this could suggest some possible developments in the theory of viscosity solutions on metric spaces. The convergence results just stated have interesting byproducts, as they allow to deduce in an 'entropic' and thus alternative way some already known facts on RCD spaces, for instance the (K, N)-displacement convexity of the entropy and the HWI inequality; moreover, the proofs are light, as a direct consequence of the regularity features of entropic interpolations.

In Chapter 7, relying on all the previous results we finally show the validity of the second order differentiation formula for W_2-geodesics. A first consequence is a closely related first order differentiation formula for vector fields (see Corollary 7.1.3 for the precise statement), but other applications arise in geometric analysis: we shall explain where and how our main result can be used in the proof of the splitting theorem to simplify the strategy.

Basic notations

$(\mathrm{X}, \mathsf{d}, \mathfrak{m})$	metric measure space (from Chapter 5 always assumed to be $\mathsf{RCD}^*(K, N)$
$\mathrm{Adm}(\mu, \nu)$	set of admissible plans for the couple (μ, ν)
d	differential
Δ, $\boldsymbol{\Delta}$	Laplacian and measure-valued Laplacian
$\mathrm{Geo}(\mathrm{X})$	space of constant speed geodesics with values in the metric space (X, d)
$\boldsymbol{\Gamma}_2$	measure-valued iterated *carré du champ* operator
$\mathrm{OptGeo}(\mu, \nu)$	set of optimal plans for the couple (μ, ν) representing a geodesic
$\mathrm{Hess}(f)$	Hessian of f
h_t	heat semigroup in $L^2(\mathrm{X})$
$\mathsf{r}_t[x]$	heat kernel at time t starting in x
$\mathrm{Lip}(f)$	Lipschitz constant of f
$\mathrm{lip}(f)$	local Lipschitz constant of f
$\mathfrak{h}_t$	mollified heat semigroup
$\mathscr{B}(\mathrm{X})$	Borel σ-algebra over X
$\mathrm{Opt}(\mu, \nu)$	set of optimal plans for the couple (μ, ν)
$\overline{\mathbb{R}}$	the extended real line
$\mathscr{P}(\mathrm{X})$	space of Borel probability measures over X
$\mathscr{P}_2(\mathrm{X})$	space of Borel probability measures over X with finite second moment
supp	support of (a function or a measure)
$\mathrm{Test}(\mathrm{X})$	test functions over $(\mathrm{X}, \mathsf{d}, \mathfrak{m})$
$\mathrm{TestV}(\mathrm{X})$	test vector fields over $(\mathrm{X}, \mathsf{d}, \mathfrak{m})$

φ^c	c-transform of φ
$AC([0,1],\mathrm{X})$	space of absolutely continuous curves on $[0,1]$ with values in the metric space (X,d)
$AC^2([0,1],\mathrm{X})$	space of absolutely continuous curves on $[0,1]$ with values in the metric space (X,d) and 2-integrable metric derivative
$AC_{loc}([0,\infty),H)$	space of locally absolutely continuous curves on $[0,\infty)$ with values in the Hilbert space H
$C(\mathrm{X})$	space of continuos functions over X
$C_b(\mathrm{X})$	space of continuos and bounded functions over X
$D(\Phi)$	domain of the functional $\Phi:\mathrm{X}\to\mathbb{R}\cup\{+\infty\}$
E	Cheeger's energy functional
$L^2(T\mathrm{X})$, $L^2(T^*\mathrm{X})$	tangent and cotangent module
$L^2(T^{\otimes 2}\mathrm{X})$, $L^2((T^*)^{\otimes 2}\mathrm{X})$	tensor product of tangent/cotangent module with itself
$L^p(\mathrm{X})$, $L^p(\mathfrak{m})$, $L^p(\mathrm{X},\mathfrak{m})$	space of p-integrable functions on X w.r.t. $\mathfrak{m}$
Q_t	Hopf-Lax semigroup
$S^2(\mathrm{X})$	Sobolev class over $(\mathrm{X},\mathsf{d},\mathfrak{m})$
$W^{1,2}(\mathrm{X})$, $W^{2,2}(\mathrm{X})$	Sobolev spaces over $(\mathrm{X},\mathsf{d},\mathfrak{m})$
W_2	Wasserstein distance
$x_n\downarrow\bar{x}$	decreasing convergence of x_n to $\bar{x}$
$x_n\rightharpoonup\bar{x}$	weak convergence of x_n to $\bar{x}$
$x_n\to\bar{x}$	convergence of x_n to $\bar{x}$
$x_n\uparrow\bar{x}$	increasing convergence of x_n to $\bar{x}$
$\mathrm{LIP}(\mathrm{X})$	space of Lipschitz continuos functions over X

Part I

Analysis of RCD spaces

Chapter 1

Preliminaries

Aim of this chapter is twofold: on the one hand, for the reader's sake we introduce the basic metric, topological, measure-theoretic and differential concepts used in the thesis; on the other one, we recall some preliminary results involving these notions that will be extensively invoked throughout the manuscript.

In Section 1.1 we begin with purely metric definitions and we progressively enrich the structure of our framework, up to the notion of infinitesimal Hilbertianity. More precisely, we start with the well-understood concepts of test plan and minimal weak upper gradient until we reach the sophisticated machinery of $L^2(\mathfrak{m})$-normed modules, which is the key ingredient for the understanding of the first order differential structure of general metric measure spaces. We conclude the section with a quick overview of the Monge-Kantorovich optimal transport problem and a first look at absolutely continuous curves in the Wasserstein space, pointing out the connection with the continuity equation.

The (reduced) curvature-dimension condition is the main feature of Section 1.2, because the introduction of this purely synthetic notion in an infinitesimally Hilbertian space (thus leading to the definition of RCD^* space) enables a wide discussion on the geometric information encoded therein, on the regularization properties of the heat flow and on the second order differential structure, whose cornerstones are the Sobolev space $W^{2,2}(\mathrm{X})$ and the notion of Hessian. As a final remark, we provide some further results in optimal transport under the RCD^* assumption.

1.1 Infinitesimally Hilbertian spaces

1.1.1 First order calculus

Let us first present those notions that only need a metric structure to be defined. For this reason, let (X, d) be a metric space; by $\mathscr{P}(\mathrm{X})$ we denote the space of Borel probability measures on it and by $\mathscr{P}_2(\mathrm{X}) \subset \mathscr{P}(\mathrm{X})$ the subclass of those with finite second moment, i.e. $\mu \in \mathscr{P}_2(\mathrm{X})$ if and only if

$$\int \mathsf{d}^2(x, x_0)\mathrm{d}\mu(x) < \infty$$

for some, and thus any, $x_0 \in \mathrm{X}$. The first notion we recall is the following: if $\mu \in \mathscr{P}(\mathrm{X})$ and $T : \mathrm{X} \to \mathrm{Y}$ is a μ-measurable map taking values in the topological space Y, the push-forward measure $T_\#\mu \in \mathscr{P}(\mathrm{Y})$ is defined by $T_\#\mu(B) := \mu(T^{-1}(B))$ for all Borel set $B \subset \mathrm{Y}$. In many

occasions, the map T will be given by the projection from a product space on one of its factors or by the **evaluation map** $\mathrm{e}_t : C([0,1],(\mathrm{X},\mathsf{d})) \to \mathrm{X}$ defined by $\mathrm{e}_t(\gamma) := \gamma_t$, where $C([0,1],(\mathrm{X},\mathsf{d}))$, or simply $C([0,1],\mathrm{X})$, denotes the space of continuous curves with values on the metric space (X,d).

A curve $\gamma : [0,1] \to \mathrm{X}$ is said to be **absolutely continuous** provided there exists a function $f \in L^1(0,1)$ such that

$$\mathsf{d}(\gamma_t,\gamma_s) \le \int_s^t f(r)\mathrm{d}r \qquad \forall s,t \in [0,1],\ s<t$$

and the collection of such curves will be denoted by $AC([0,1],(\mathrm{X},\mathsf{d}))$ or simply $AC([0,1],\mathrm{X})$. The **metric speed** $t \mapsto |\dot\gamma_t| \in L^1(0,1)$ of an absolutely continuous curve γ is defined as the essential infimum of all the functions $f \in L^1(0,1)$ such that the inequality above holds but it can be equivalently seen as limit of incremental ratios, namely

$$|\dot\gamma_t| = \lim_{h\to 0} \frac{\mathsf{d}(\gamma_{t+h},\gamma_t)}{|h|} \qquad \text{for a.e. } t \in [0,1],$$

where the a.e. existence of the limit is part of the statement (see for instance Theorem 1.1.2 of [5] for a proof). In what follows we will write $\int_0^1 |\dot\gamma_t|^2\mathrm{d}t$ even if $\gamma \in C([0,1],\mathrm{X})$: in the case γ is not absolutely continuous, the integral is set equal to $+\infty$ by definition. In the case the metric speed $|\dot\gamma_t|$ belongs to $L^p(0,1)$ (resp. $L^p_{loc}(0,1)$), the curve γ is said p-absolutely continuous (resp. locally p-absolutely continuous) and the family of such curves will be denoted by $AC^2([0,1],\mathrm{X})$ (resp. $AC_{loc}([0,1],\mathrm{X})$).

A (constant speed) **geodesic** is a curve $\gamma : [0,1] \to \mathrm{X}$ such that

$$\mathsf{d}(\gamma_s,\gamma_t) = |t-s|\mathsf{d}(\gamma_0,\gamma_1) \qquad \forall s,t \in [0,1].$$

We will denote by $\mathrm{Geo}(\mathrm{X})$ the set made up by all of them and say that (X,d) is a **geodesic space** if for any $x_0,x_1 \in \mathrm{X}$ there exists a constant speed geodesic γ satisfying $\gamma_0 = x_0$ and $\gamma_1 = x_1$. A weaker property is the one of **length space**: for any $x_0,x_1 \in \mathrm{X}$ and $\varepsilon > 0$ there exists $\gamma \in AC([0,1],\mathrm{X})$ such that $\gamma_0 = x_0$, $\gamma_1 = x_1$ and

$$\ell(\gamma) := \int_0^1 |\dot\gamma_t|\mathrm{d}t \le \mathsf{d}(x_0,x_1) + \varepsilon.$$

Finally, recall that given $f : \mathrm{X} \to \mathbb{R}$ the **upper** and **lower slopes** $|D^+f|, |D^-f| : \mathrm{X} \to [0,\infty]$ are defined as 0 on isolated points and otherwise

$$|D^+f|(x) := \overline{\lim_{y\to x}} \frac{(f(y)-f(x))^+}{\mathsf{d}(x,y)} \qquad\qquad |D^-f|(x) := \overline{\lim_{y\to x}} \frac{(f(y)-f(x))^-}{\mathsf{d}(x,y)}.$$

Similarly, the **local Lipschitz constant** $\mathrm{lip}(f) : \mathrm{X} \to [0,\infty]$ is defined as 0 on isolated points and otherwise as

$$\mathrm{lip}f(x) := \max\{|D^+f|(x), |D^-f|(x)\} = \limsup_{y\to x} \frac{|f(x)-f(y)|}{d(x,y)}.$$

This allows us to single out some properties of the **Hopf-Lax semigroup** in metric spaces. For $f : \mathrm{X} \to \mathbb{R}\cup\{+\infty\}$ and $t > 0$ we define the function $Q_t f : \mathrm{X} \to \mathbb{R}\cup\{-\infty\}$ as

$$Q_t f(x) := \inf_{y\in\mathrm{X}} \frac{\mathsf{d}^2(x,y)}{2t} + f(y) \tag{1.1.1}$$

and set $t_* := \sup\{t > 0 : Q_t f(x) > -\infty \text{ for some } x \in \mathrm{X}\}$; it is worth saying that t_* does not actually depend on x, since if $Q_t f(x) > -\infty$, then $Q_s f(y) > -\infty$ for all $s \in (0,t)$ and all $y \in \mathrm{X}$. With this premise we have the following result, whose proof can be found in [6] (see Theorems 3.5 and 3.6 therein) and [85] :

Proposition 1.1.1. *Let* (X, d) *be a metric space and* $f : \mathrm{X} \to \mathbb{R} \cup \{+\infty\}$. *Then for all* $x \in \mathrm{X}$ *the following facts hold:*

(i) $Q_t f(x) \uparrow f(x)$ *as* $t \downarrow 0$*;*

(ii) $Q_t(Q_s f(x)) \geq Q_{t+s} f(x)$ *for all* $t, s \geq 0$*;*

(iii) the map $(0, t_*) \ni t \mapsto Q_t f(x)$ *is locally Lipschitz and*

$$\frac{\mathrm{d}}{\mathrm{d}t} Q_t f(x) + \frac{1}{2}\Big(\mathrm{lip} Q_t f(x)\Big)^2 = 0 \tag{1.1.2}$$

for all but countably many $t \in (0, t^*)$.

In the case we further assume (X, d) *to be geodesic, (ii) and (iv) are true with equality.*

As a next step, let us assume from now on (X, d) to be complete and separable and let us endow it with a Borel non-negative measure $\mathfrak{m}$ which is finite on bounded sets and with full support (this last assumption is not really needed, but allows to avoid some technicalities). For the forthcoming definitions of test plans, Sobolev class and of minimal weak upper gradient we draw from [6], but the interested reader can also refer to the previous works [28], [112] for alternative - but equivalent - definitions of Sobolev functions and to [66] for a wide dissertation on the topic.

A path measure $\boldsymbol{\pi} \in \mathscr{P}(C([0,1], \mathrm{X}))$ is a **test plan** if it has bounded compression and finite 2-action, namely there exists a constant $C = C(\boldsymbol{\pi}) > 0$ such that

$$(\mathrm{e}_t)_* \boldsymbol{\pi} \leq C\mathfrak{m} \qquad \forall t \in [0,1],$$
$$\iint_0^1 |\dot{\gamma}_t|^2 \mathrm{d}t \mathrm{d}\boldsymbol{\pi} < +\infty.$$

Recall that $\int_0^1 |\dot{\gamma}_t|^2 \mathrm{d}t := +\infty$ if γ is not absolutely continuous; thus, a fortiori any test plan must be concentrated on $AC^2([0,1], \mathrm{X})$.

The **Sobolev class** $S^2(\mathrm{X}, \mathsf{d}, \mathfrak{m})$, also denoted by $S^2(\mathrm{X})$ when no ambiguity arises, is defined as the space of all Borel functions $f : \mathrm{X} \to \mathbb{R}$ for which there exists a non-negative function $G \in L^2(\mathfrak{m})$, called weak upper gradient, such that

$$\int |f(\gamma_1) - f(\gamma_0)| \mathrm{d}\boldsymbol{\pi}(\gamma) \leq \iint_0^1 G(\gamma_t) |\dot{\gamma}_t| \mathrm{d}t \mathrm{d}\boldsymbol{\pi}(\gamma) \tag{1.1.3}$$

for all test plan $\boldsymbol{\pi}$. It turns out that there exists a minimal G in the $\mathfrak{m}$-a.e. sense satisfying the inequality above: such minimal function will be called **minimal weak upper gradient** (although defined in duality with distance, so that it should rather be called minimal weak upper differential) and denoted by $|Df|$. It is not difficult to see that (1.1.3) implies a localized version of the same inequality: for all $0 \leq s \leq t \leq 1$ it holds

$$\int |f(\gamma_t) - f(\gamma_s)| \mathrm{d}\boldsymbol{\pi}(\gamma) \leq \iint_s^t G(\gamma_r) |\dot{\gamma}_r| \mathrm{d}r \mathrm{d}\boldsymbol{\pi}(\gamma). \tag{1.1.4}$$

Moreover, minimal weak upper gradients enjoy the following important locality property:

$$|Df| = |Dg| \qquad \mathfrak{m}\text{-a.e. on } \{f = g\} \qquad\qquad \forall f, g \in S^2(\mathrm{X}).$$

The **Sobolev space** $W^{1,2}(\mathrm{X})$ is then defined as $L^2(\mathrm{X}) \cap S^2(\mathrm{X})$ and it turns out to be a Banach space when endowed with the norm

$$\|f\|^2_{W^{1,2}(\mathrm{X})} := \|f\|^2_{L^2(\mathrm{X})} + \||Df|\|^2_{L^2(\mathrm{X})}.$$

However in general $W^{1,2}(\mathrm{X})$ is not Hilbert and this motivates a further step.

Along with [58], the space $(\mathrm{X}, \mathsf{d}, \mathfrak{m})$ is said **infinitesimally Hilbertian** provided $W^{1,2}(\mathrm{X})$ is a Hilbert space. Introducing the **Cheeger energy** as the convex and lower-semicontinuous functional $E : L^2(\mathrm{X}) \to [0, \infty]$ given by

$$E(f) := \begin{cases} \dfrac{1}{2} \displaystyle\int |Df|^2 \, \mathrm{d}\mathfrak{m} & \text{for } f \in W^{1,2}(\mathrm{X}) \\ +\infty & \text{otherwise} \end{cases}$$

it is well known that infinitesimal Hilbertianity is equivalent to the fact that E is a Dirichlet form. Its infinitesimal generator Δ, which is a closed self-adjoint linear operator on $L^2(\mathrm{X})$, is called **Laplacian** on $(\mathrm{X}, \mathsf{d}, \mathfrak{m})$ and its domain denoted by $D(\Delta) \subset W^{1,2}(\mathrm{X})$. Relying on the classical theory of gradient flows of convex functions on Hilbert spaces, whose main results can be found in [5] with detailed bibliographical references or in [6], we deduce existence and uniqueness of a 1-parameter semigroup (h_t) of continuous operators from $L^2(\mathrm{X})$ to itself such that for any $f \in L^2(\mathrm{X})$ the curve $t \mapsto \mathsf{h}_t f \in L^2(\mathrm{X})$ is continuous on $[0, \infty)$, locally absolutely continuous on $(0, \infty)$ and the only solution of

$$\frac{\mathrm{d}}{\mathrm{d}t} \mathsf{h}_t f = \Delta \mathsf{h}_t f \quad \text{a.e. } t > 0, \qquad \mathsf{h}_t f \overset{L^2}{\to} f \text{ as } t \downarrow 0$$

where it is part of the statement the fact that $\mathsf{h}_t f \in D(\Delta)$ for every $f \in L^2(\mathrm{X})$ and $t > 0$. Such 1-parameter semigroup is called **heat flow**.

Beside this notion of L^2-valued Laplacian, we shall also need that of measure-valued Laplacian ([58]). A function $f \in W^{1,2}(\mathrm{X})$ is said to have **measure-valued Laplacian**, and in this case we write $f \in D(\boldsymbol{\Delta})$, provided there exists a Borel (signed) measure μ whose total variation is finite on bounded sets and such that

$$\int g \, \mathrm{d}\mu = -\int \langle \nabla g, \nabla f \rangle \, \mathrm{d}\mathfrak{m}, \qquad \forall g \text{ Lipschitz with bounded support.} \tag{1.1.5}$$

In this case μ is unique and denoted $\boldsymbol{\Delta} f$. This notion is compatible with the previous one in the sense that

$$f \in D(\boldsymbol{\Delta}),\ \boldsymbol{\Delta} f \ll \mathfrak{m} \text{ and } \frac{\mathrm{d}\boldsymbol{\Delta} f}{\mathrm{d}\mathfrak{m}} \in L^2(\mathfrak{m}) \qquad \Leftrightarrow \qquad f \in D(\Delta) \text{ and in this case } \Delta f = \frac{\mathrm{d}\boldsymbol{\Delta} f}{\mathrm{d}\mathfrak{m}}.$$

Finally, let us provide a quick overview about tangent and cotangent modules and some of the differential objects that, more or less naturally, come along with them; all the forthcoming definitions and observations are taken from [56], unless otherwise specified.

Recall that a $L^2(\mathrm{X})$-normed $L^\infty(\mathrm{X})$-module, or simply an $L^2(\mathrm{X})$-normed module, is a Banach space $(\mathscr{M}, \|\cdot\|_{\mathscr{M}})$ endowed with a bilinear map

$$\begin{array}{rcl} L^\infty(\mathrm{X}) \times \mathscr{M} & \to & \mathscr{M} \\ (f, v) & \mapsto & f \cdot v \end{array}$$

called multiplication by $L^\infty(\mathrm{X})$ functions, and a map $|\cdot| : \mathscr{M} \to L^2(\mathrm{X})$ with non-negative values, called pointwise $L^2(\mathrm{X})$-norm, such that:

(i) for every $v \in \mathscr{M}$ and $f, g \in L^\infty(\mathrm{X})$

$$f \cdot (g \cdot v) = (fg) \cdot v, \qquad 1 \cdot v = v, \qquad \|f \cdot v\|_{\mathscr{M}} \leq \|f\|_{L^\infty(\mathrm{X})} \|v\|_{\mathscr{M}}$$

where 1 denotes the function identically equal to 1;

(ii) for every $v \in \mathscr{M}$ and $f \in L^\infty(\mathcal{X}, \mathfrak{m})$ it holds

$$\||v|\|_{L^2(\mathrm{X})} = \|v\|_{\mathscr{M}}, \qquad\qquad |f \cdot v| = |f|\,|v| \quad \mathfrak{m}\text{-a.e.}$$

An isomorphism between two $L^2(\mathrm{X})$-normed modules is a linear bijection which preserves the norm, the product with $L^\infty(\mathrm{X})$ functions and the pointwise norm.

When $(\mathscr{M}, \|\cdot\|_{\mathscr{M}})$ is a Hilbert space, we shall say that $\mathscr{M}$ as $L^2(\mathrm{X})$-normed module is a Hilbert module. The dual module $\mathscr{M}^*$ of $\mathscr{M}$ is defined as $\mathrm{Hom}(\mathscr{M}, L^1(\mathrm{X}))$ and it is an $L^2(\mathrm{X})$-normed module too. With these premises, we can state the following existence theorem, thus defining the cotangent module.

Theorem 1.1.2 (Definition of cotangent module)**.** *There exists a unique couple $(L^2(T^*\mathrm{X}), \mathrm{d})$, where $L^2(T^*\mathrm{X})$ is an $L^2(\mathrm{X})$-normed module and $\mathrm{d} : S^2(\mathrm{X}) \to L^2(T^*\mathrm{X})$ is a linear and continuous map, such that:*

(i) $|\mathrm{d}f| = |Df|$ $\mathfrak{m}$-a.e. for every $f \in S^2(\mathrm{X})$;

(ii) the space $\{\mathrm{d}f : f \in S^2(\mathrm{X})\}$ generates $L^2(T^*\mathrm{X})$ in the sense of modules.

Uniqueness is meant up to unique isomorphism, namely if $(\mathscr{M}, \mathrm{d}')$ satisfies the same properties, then there is a unique isomorphism $\Phi : L^2(T^*\mathrm{X}) \to \mathscr{M}$ such that $\Phi(\mathrm{d}f) = \mathrm{d}'f$ for all $f \in S^2(\mathrm{X})$.

The $L^2(\mathrm{X})$-normed module $L^2(T^*\mathrm{X})$ is called **cotangent module** and d is the **differential**, which is a closed operator when seen as unbounded operator on $L^2(\mathrm{X})$ (see the forthcoming Lemma 1.1.3). The **tangent module** $L^2(T\mathrm{X})$ is defined as the dual of $L^2(T^*\mathrm{X})$. The elements of $L^2(T^*\mathrm{X})$ will be called cotangent vector fields or 1-forms, while the elements of $L^2(T\mathrm{X})$ vector fields. If $(\mathrm{X}, \mathsf{d}, \mathfrak{m})$ is infinitesimally Hilbertian, as it is now and shall always be the case, the cotangent and tangent modules are canonically isomorphic both as Hilbert modules and as Hilbert spaces via the Riesz theorem. Thus, if we denote by $\langle \cdot\,, \cdot \rangle$ the scalar product associated to the pointwise norm of $L^2(T\mathrm{X})$, we can introduce the (musical) isomorphisms $\flat : L^2(T\mathrm{X}) \to L^2(T^*\mathrm{X})$ and $\sharp : L^2(T^*\mathrm{X}) \to L^2(T\mathrm{X})$ as

$$V^\flat(W) := \langle V, W \rangle, \qquad\qquad \langle \omega^\sharp, W \rangle := \omega(W) \tag{1.1.6}$$

$\mathfrak{m}$-a.e. for every $V, W \in L^2(T\mathrm{X})$ and $\omega \in L^2(T^*\mathrm{X})$. This allows to define the **gradient** of $f \in S^2(\mathrm{X})$ as $\nabla f := (\mathrm{d}f)^\sharp$; equivalently, it can be defined as the unique $W \in L^2(T\mathrm{X})$ such

that $\mathrm{d}f(W) = |W|^2 = |\mathrm{d}f|^2$ $\mathfrak{m}$-a.e., this definition being in line with the one used in Finsler geometry (see e.g. [16]).

The **divergence** of a vector field is defined as (minus) the adjoint of the differential, i.e. we say that $W \in L^2(T\mathrm{X})$ has a divergence, and write $W \in D(\mathrm{div})$, provided there is a function $g \in L^2(\mathrm{X})$ such that

$$\int fg \,\mathrm{d}\mathfrak{m} = -\int \mathrm{d}f(W)\,\mathrm{d}\mathfrak{m} \qquad \forall f \in W^{1,2}(\mathrm{X}).$$

In this case g is unique and is denoted $\mathrm{div}(W)$. It can also be verified that

$$f \in D(\Delta) \text{ if and only if } \nabla f \in D(\mathrm{div}) \text{ and in this case } \Delta f = \mathrm{div}(\nabla f),$$

in accordance with the smooth case and this immediately implies that $f \in D(\Delta)$ and $h = \Delta f$ if and only if

$$\int gh \,\mathrm{d}\mathfrak{m} = -\int \langle \nabla g, \nabla f\rangle \,\mathrm{d}\mathfrak{m}, \qquad \forall g \in W^{1,2}(\mathrm{X}).$$

We conclude this first introductory part with a lemma collecting all the calculus rules for differential, gradient, divergence and Laplacian which we shall use extensively throughout the whole thesis without further notice.

Lemma 1.1.3 (Calculus rules 1)**.** *Let* $(\mathrm{X}, \mathsf{d}, \mathfrak{m})$ *be an infinitesimally Hilbertian space. Then:*

(i) (Differential and gradient) d *satisfies the following calculus rules:*

$$\begin{aligned}
|\mathrm{d}f| &= |Df| \quad \mathfrak{m}\text{-a.e.} && \forall f \in S^2(\mathrm{X})\\
\mathrm{d}f &= \mathrm{d}g \quad \mathfrak{m}\text{-a.e. on } \{f = g\} && \forall f, g \in S^2(\mathrm{X})\\
\mathrm{d}(\varphi \circ f) &= \varphi' \circ f \,\mathrm{d}f && \forall f \in S^2(\mathrm{X}),\ \varphi : \mathbb{R} \to \mathbb{R} \text{ Lipschitz}\\
\mathrm{d}(fg) &= g\,\mathrm{d}f + f\,\mathrm{d}g && \forall f, g \in L^\infty \cap S^2(\mathrm{X})
\end{aligned}$$

where it is part of the properties the fact that $\varphi \circ f, fg \in S^2(\mathrm{X})$ *for* φ, f, g *as above; analogous statements hold for the gradient.*

(ii) (Divergence) for all $f \in W^{1,2}(\mathrm{X}), W \in D(\mathrm{div})$ *such that* $|f|, |W| \in L^\infty(\mathrm{X})$ *it holds*

$$\mathrm{div}(fW) = \mathrm{d}f(W) + f\mathrm{div}(W),$$

where it is part of the statement that $fW \in D(\mathrm{div})$ *for* f, W *as above.*

(iii) (Laplacian) Δ *enjoys the chain and Leibniz rules:*

$$\Delta(\varphi \circ f) = \varphi'' \circ f |\mathrm{d}f|^2 + \varphi' \circ f \Delta f \tag{1.1.7a}$$

$$\Delta(fg) = g\Delta f + f\Delta g + 2\,\langle \nabla f, \nabla g\rangle \tag{1.1.7b}$$

where in the first equality we assume that $f \in D(\Delta), \varphi \in C^2(\mathbb{R})$ *are such that* $f, |\mathrm{d}f| \in L^\infty(\mathrm{X})$ *and* $\varphi', \varphi'' \in L^\infty(\mathbb{R})$ *and in the second that* $f, g \in D(\Delta) \cap L^\infty(\mathrm{X})$ *and* $|\mathrm{d}f|, |\mathrm{d}g| \in L^\infty(\mathrm{X})$ *and it is part of the claims that* $\varphi \circ f, fg$ *are in* $D(\Delta)$.

proof See [56] for properties (i), (ii) and (1.1.7b). As regards (1.1.7a), the proof is in the same spirit of the Leibniz rule for Δ, but for sake of completeness we prefer to present it, since not present in [56]. First notice that the assumptions on f, φ grant that for any $h \in W^{1,2}(\mathrm{X})$ we have $\varphi \circ f, \varphi' \circ f, h\varphi' \circ f \in W^{1,2}(\mathrm{X})$ and also that the right-hand side of (1.1.7a) belongs to $L^2(\mathrm{X})$. Hence, using the integration by parts characterization of the Laplacian and the chain and Leibniz rules for gradients just stated we see that

$$\begin{aligned}
-\int \langle \nabla h, \nabla(\varphi \circ f)\rangle \mathrm{d}\mathfrak{m} &= -\int \langle \nabla h, \varphi' \circ f \nabla f\rangle \mathrm{d}\mathfrak{m} \\
&= -\int \Big(\langle \nabla(h\varphi' \circ f), \nabla f\rangle + h\langle \nabla(\varphi' \circ f), \nabla f\rangle \Big) \mathrm{d}\mathfrak{m} \\
&= \int h\Big(\varphi' \circ f \Delta f + \varphi'' \circ f |\nabla f|^2\Big) \mathrm{d}\mathfrak{m}
\end{aligned}$$

whence the conclusion. □

It is worth saying that the properties of the differential actually hold even if the space is not assumed infinitesimally Hilbertian.

1.1.2 A first glance at optimal transport

Following the standard terminology (see for instance [3] or [121]), on a complete and separable metric space (X, d) the **Monge-Kantorovich** or **optimal transport problem** with quadratic cost and marginal constraints $\mu_0, \mu_1 \in \mathscr{P}_2(\mathrm{X})$ is the minimization problem

$$\inf \int_{\mathrm{X}^2} \mathsf{d}^2(x, y) \mathrm{d}\boldsymbol{\gamma}(x, y) \tag{1.1.8}$$

the infimum being taken among all $\boldsymbol{\gamma} \in \mathscr{P}(\mathrm{X}^2)$ such that $\pi^0_* \boldsymbol{\gamma} = \mu_0$ and $\pi^1_* \boldsymbol{\gamma} = \mu_1$, where $\pi^0, \pi^1 : \mathrm{X}^2 \to \mathrm{X}$ are the canonical projections. The set of all such $\boldsymbol{\gamma}$ will be denoted by $\mathrm{Adm}(\mu_0, \mu_1)$ and its elements will be called admissible plans or, more frequently, **transport plans** for the couple (μ_0, μ_1). It is well known that the infimum is actually attained: the Wasserstein distance W_2 between μ_0 and μ_1 is then defined as the square root of (1.1.8), any minimizer is called **optimal** and their collection is denoted by $\mathrm{Opt}(\mu_0, \mu_1)$. The set $\mathscr{P}_2(\mathrm{X})$ equipped with W_2 turns out to be a complete and separable metric space and if (X, d) is a length (resp. geodesic) space, then so is $(\mathscr{P}_2(\mathrm{X}), W_2)$.

Passing to the dual formulation of (1.1.8), given by

$$\sup \Big\{ \int \varphi \, \mathrm{d}\mu_0 + \int \psi \, \mathrm{d}\mu_1 \Big\},$$

where the supremum runs over all functions $\varphi \in L^1(\mathrm{X}, \mu_0)$ and $\psi \in L^1(\mathrm{X}, \mu_1)$, we recall that the c-transform $\varphi^c : X \to \mathbb{R} \cup \{-\infty\}$ of a function $\varphi : \mathrm{X} \to \mathbb{R} \cup \{-\infty\}$ is defined as

$$\varphi^c(x) := \inf_{y \in X} \Big\{ \frac{\mathsf{d}^2(x, y)}{2} - \varphi(y) \Big\}$$

and that φ is said to be **c-concave** provided $\varphi = \psi^c$ for some ψ. With this premise, given $\mu_0, \mu_1 \in \mathscr{P}_2(\mathrm{X})$ a function $\varphi : \mathrm{X} \to \mathbb{R} \cup \{-\infty\}$ is called **Kantorovich potential** from μ_0 to μ_1 provided it is c-concave and

$$\int \varphi \, \mathrm{d}\mu_0 + \int \varphi^c \, \mathrm{d}\mu_1 = \frac{1}{2} W_2^2(\mu_0, \mu_1),$$

which implies in particular that the supremum in the dual problem is attained by the couple (φ, φ^c). By the classical theory of optimal transport, existence for potentials is granted under very mild assumptions (which in particular hold for the quadratic cost), but it is worth recalling that on general complete and separable metric spaces (X, d) we have that for $\mu_0, \mu_1 \in \mathscr{P}_2(\mathrm{X})$ with bounded support there exists a Kantorovich potential from μ_0 to μ_1 which is Lipschitz and bounded.

This can be obtained starting from an arbitrary Kantorovich potential ψ and then defining

$$\varphi(x) := \min\Big\{C, \inf_{y\in X} \frac{\mathsf{d}^2(x,y)}{2} - \psi^c(y)\Big\}$$

for C sufficiently big.

Looking back at (1.1.8), given $\mu_0, \mu_1 \in \mathscr{P}_2(\mathrm{X})$ it is well known that the optimal transport problem admits a dynamical version, namely

$$\inf \iint_0^1 |\dot\gamma_t|^2 \mathrm{d}\boldsymbol{\pi}(\gamma) \tag{1.1.9}$$

where the infimum is taken among all $\boldsymbol{\pi} \in \mathscr{P}(C([0,1],\mathrm{X}))$ such that $(\mathrm{e}_0)_*\boldsymbol{\pi} = \mu_0, (\mathrm{e}_1)_*\boldsymbol{\pi} = \mu_1$. Moreover, if a dynamical minimizer $\boldsymbol{\pi}$ exists, then its marginal flow $((\mathrm{e}_t)_*\boldsymbol{\pi})$ is a Wasserstein geodesic in $(\mathscr{P}_2(\mathrm{X}), W_2)$ and $(\mathrm{e}_0, \mathrm{e}_1)_*\boldsymbol{\pi} \in \mathrm{Opt}(\mu_0, \mu_1)$. These considerations suggest that a helpful approach to (1.1.8) may come from a better understanding of W_2-geodesics and, more generally, W_2-absolutely continuous curves.

In this direction, it is well known that on $\mathbb{R}^d$, curves of measures which are W_2-absolutely continuous are in correspondence with appropriate solutions of the **continuity equation**, as proved in Sections 8.1 and 8.2 of [5]. It has been later shown in [59] that the same connection holds on arbitrary metric measure spaces $(\mathrm{X}, \mathsf{d}, \mathfrak{m})$, provided the measures are controlled by $C\mathfrak{m}$ for some $C > 0$; the formulation of such result and the good notion of solution of the continuity equation which we shall need are the following, where Lisini's superposition result [84] is taken into account too.

Theorem 1.1.4 (Continuity equation and W_2-AC curves)**.** *Let* $(\mathrm{X}, \mathsf{d}, \mathfrak{m})$ *be infinitesimally Hilbertian,* $(\mu_t) \subset \mathscr{P}_2(\mathrm{X})$ *be weakly continuous and* $t \mapsto X_t \in L^0(T\mathrm{X})$ *be a family of vector fields, possibly defined only for a.e.* $t \in [0,1]$*. Assume that the map* $t \mapsto \int |X_t|^2 \mathrm{d}\mu_t$ *is Borel and:*

$$\mu_t \le C\mathfrak{m} \qquad \forall t \in [0,1] \text{ for some } C > 0 \tag{1.1.10a}$$

$$\int_0^1 \int |X_t|^2 \,\mathrm{d}\mu_t \,\mathrm{d}t < \infty \tag{1.1.10b}$$

and that the continuity equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\mu_t + \mathrm{div}(X_t\mu_t) = 0,$$

is satisfied in the following sense: for any $f \in W^{1,2}(\mathrm{X})$ *the map* $[0,1] \ni t \mapsto \int f \,\mathrm{d}\mu_t$ *is absolutely continuous and it holds*

$$\frac{\mathrm{d}}{\mathrm{d}t}\int f \,\mathrm{d}\mu_t = \int \mathrm{d}f(X_t)\,\mathrm{d}\mu_t \qquad \text{a.e. } t.$$

Then $(\mu_t) \in AC([0,1],(\mathscr{P}_2(\mathrm{X}),W_2))$ *and there exists a* lifting $\boldsymbol{\pi}$ *of it, i.e. a measure* $\boldsymbol{\pi} \in \mathscr{P}(C([0,1],\mathrm{X}))$ *such that* $(\mathrm{e}_t)_*\boldsymbol{\pi} = \mu_t$ *for every* $t \in [0,1]$*, in such a way that*

$$|\dot{\mu}_t|^2 = \int |X_t|^2 \,\mathrm{d}\mu_t = \int |\dot{\gamma}_t|^2 \,\mathrm{d}\boldsymbol{\pi}(\gamma) \qquad \text{a.e. } t \in [0,1].$$

The correspondence between W_2-absolutely continuous curves and solutions of the continuity equation also allows to recover the following version of the Benamou-Brenier formula (see [56]).

Theorem 1.1.5. *Let* $(\mathrm{X},\mathsf{d},\mathfrak{m})$ *be infinitesimally Hilbertian and* $\mu,\nu \in \mathscr{P}_2(\mathrm{X})$ *be such that there exists a* W_2*-geodesic* (μ_t) *connecting them such that* $\mu_t \le C\mathfrak{m}$ *for all* $t \in [0,1]$ *for some* $C > 0$*. Then*

$$W_2^2(\mu,\nu) = \min \int_0^1 \int |X_t|^2 \mathrm{d}\mu_t \mathrm{d}t,$$

where the minimum is taken among all solutions (μ_t, X_t) *of the continuity equation such that* $\mu_0 = \mu$ *and* $\mu_1 = \nu$*.*

For a more exhaustive picture of the matter, including the relevant case of Wasserstein geodesics, further assumptions on $(\mathrm{X},\mathsf{d},\mathfrak{m})$ have to be required and this is going to be realized in the next section.

1.2 The RCD condition

Let us begin with the very definition of $\mathsf{CD}(K,\infty)$ space, independently proposed by Sturm in [115], [116] and by Lott-Villani in [87]. Such definition has been the starting point of an extremely rich and fruitful investigation, as already said in the introduction, but rephrasing Bernard of Chartres' words this has been possible by standing on the shoulders of seminal (and completely different) papers appeared before.

On the one hand, motivated by the study of hypercontractivity for diffusion processes, in [13] and [14] D. Bakry and V. Émery formulated a curvature-dimension condition in terms of the iterated *carré du champ* operator associated to a given Dirichlet form and this approach boosted in an impressive way the study of functional inequalities (such as Li-Yau-Harnack and logarithmic Sobolev ones) and concentration estimates, especially in the infinite-dimensional case (see also [75] and [15] for the subsequent developments and further references).

On the other hand, inspired by the works of F. Otto and C. Villani [104] and of D. Cordero-Erausquin, R. McCann and M. Schmuckenschläger [33], in [122] K.-T. Sturm and M. K. von Renesse were able to characterize Ricci lower bounds in several different ways, e.g. contraction and gradient estimates for the heat flow, for the gradient flow of the relative entropy and for couplings of Brownian motions; however, their most famous characterization involved K-displacement convexity of the relative entropy functional and thus optimal transport: the $\mathsf{CD}(K,\infty)$ condition à la Lott-Sturm-Villani precisely relies on this convexity requirement and reads as follows.

Definition 1.2.1. *A complete and separable metric measure space* $(\mathrm{X},\mathsf{d},\mathfrak{m})$ *is a* $\mathsf{CD}(K,\infty)$ *space or has Ricci curvature bounded below by* K *with* $K \in \mathbb{R}$*, if for any* $\mu_0,\mu_1 \in D(H(\cdot\,|\,\mathfrak{m}))$ *there exists a constant speed* W_2*-geodesic* (ν_t) *such that* $\nu_0 = \mu_0$*,* $\nu_1 = \mu_1$ *and*

$$H(\nu_t\,|\,\mathfrak{m}) \le (1-t)H(\mu_0\,|\,\mathfrak{m}) + tH(\mu_1\,|\,\mathfrak{m}) - \frac{K}{2}t(1-t)W_2^2(\mu_0,\mu_1)$$

for all $t \in [0,1]$.

After a fruitful investigation of the heat flow in non-smooth setting and its role in bridging Wasserstein geometry and contraction estimates (see [53], [60] and [6]), the relationship between the optimal transport-oriented approach à la Lott-Sturm-Villani and the Bakry-Émery one was made clear in [8], where equivalence is established under the assumption of infinitesimal Hilbertianity.

In order to take also the dimension into account, we adopt the $\mathsf{CD}^*(K,N)$ condition, due to K. Bacher and K.-T. Sturm [12]. To this aim, let us first define the Rényi entropy functional and the distortion coefficients. For $N \geq 1$ and for $\mu \in \mathscr{P}(\mathrm{X})$ with $\mu \ll \mathfrak{m}$, the former is given by

$$H_N(\mu \,|\, \mathfrak{m}) := -\int_{\mathcal{X}} \left(\frac{\mathrm{d}\mu}{\mathrm{d}\mathfrak{m}}\right)^{1-1/N} \mathrm{d}\mathfrak{m}. \tag{1.2.1}$$

For $\theta \geq 0$ and $t \in [0,1]$ we then introduce

$$\mathfrak{S}_k(\theta) := \begin{cases} \dfrac{\sin(\sqrt{k}\theta)}{\sqrt{k}\theta} & \text{if } k > 0 \\ 1 & \text{if } k = 0 \\ \dfrac{\sinh(\sqrt{-k}\theta)}{\sqrt{-k}\theta} & \text{if } k < 0 \end{cases}$$

and set the latter equal to

$$\sigma^{(t)}_{K,N}(\theta) := \begin{cases} +\infty & \text{if } K\theta^2 \geq N\pi^2 \\ t\dfrac{\mathfrak{S}_{K/N}(t\theta)}{\mathfrak{S}_{K/N}(\theta)} & \text{otherwise} \end{cases}.$$

Hence, given $K \in \mathbb{R}$ and $N \in [1,\infty)$, a complete and separable metric measure space $(\mathrm{X}, \mathsf{d}, \mathfrak{m})$ is said to be a $\mathsf{CD}^*(K,N)$ **space** or satisfies the **reduced curvature-dimension condition** $\mathsf{CD}^*(K,N)$ if, for every $\mu_0, \mu_1 \in \mathscr{P}_2(\mathrm{X})$ absolutely continuous w.r.t. $\mathfrak{m}$, there exist $\boldsymbol{\gamma} \in \mathrm{Opt}(\mu_0,\mu_1)$ and a constant speed W_2-geodesic (ν_t) such that $\nu_0 = \mu_0$, $\nu_1 = \mu_1$ and

$$H_{N'}(\nu_t \,|\, \mathfrak{m}) \leq -\int_{\mathrm{X}^2} \Big(\sigma^{(1-t)}_{K,N'}(\mathsf{d}(x_0,x_1))\rho_0^{-1/N'}(x_0) + \sigma^{(t)}_{K,N'}(\mathsf{d}(x_0,x_1))\rho_1^{-1/N'}(x_1)\Big)\mathrm{d}\boldsymbol{\gamma}(x_0,x_1)$$

for all $t \in [0,1]$ and all $N' \geq N$, where ρ_0, ρ_1 are the Radon-Nikodym derivatives of μ_0, μ_1 w.r.t. $\mathfrak{m}$ respectively. Let us point out that

$$\mathsf{CD}^*(K,N) \qquad \Rightarrow \qquad \mathsf{CD}(K,\infty).$$

Combining this condition and the $\mathsf{CD}(K,\infty)$ one with infinitesimal Hilbertianity, the notions of $\mathsf{RCD}^*(K,N)$ and $\mathsf{RCD}(K,\infty)$ space follow.

Definition 1.2.2. *Given $K \in \mathbb{R}$ and $N \in [1,\infty)$, a complete and separable metric measure space $(\mathrm{X}, \mathsf{d}, \mathfrak{m})$ is said to be a $\mathsf{RCD}^*(K,N)$ space (resp. $\mathsf{RCD}(K,\infty)$ space) provided it is an infinitesimally Hilbertian $\mathsf{CD}^*(K,N)$ space (resp. $CD(K,\infty)$ space).*

This definition was originally proposed in the case $N = \infty$ in [7] and [4] (for finite reference measures in the former and for σ-finite ones in the latter) and in the finite-dimensional case in [58], after having noticed that not only lower Ricci bounds but also Sobolev calculus is linked to W_2-geometry thanks to the understanding of the crucial role played by the heat flow, as done in the already cited works [53], [60] and [6]. Truth to be told, in [58] the notion of $\mathsf{RCD}(K, N)$ space, and not the one of $\mathsf{RCD}^*(K, N)$, was introduced:

$$\mathsf{RCD}(K, N) := \mathsf{CD}(K, N) + \text{infinitesimally Hilbertian}$$

thus relying on Lott-Sturm-Villani's $\mathsf{CD}(K, N)$ condition. However, in this theory there is an important and longstanding open problem: does a $\mathsf{CD}(K, N)$ space satisfy the globalization (or local-to-global) property? Up to technical assumptions, the answer was known to be affirmative for the cases $N = \infty$ and $K = 0$, as shown in [115] and [121] respectively. The recent advancements of Cavalletti and Milman (see [25]) suggest that this is actually true on every essentially non-branching geodesic $\mathsf{CD}(K, N)$ space for $K \in \mathbb{R}$ and $N \in (1, \infty)$; the case $N = 1$ is excluded because of critical pathologies. Nevertheless, in [25] the reference measure $\mathfrak{m}$ is assumed to have finite mass and although this hypothesis seems to be only technical, it has not been removed yet. On the other hand, the $\mathsf{CD}^*(K, N)$ condition has the local-to-global property (it was exactly for this reason that Bacher and Sturm proposed it as an alternative to CD) and this explains the success of this subsequent definition. Thus, throughout the whole manuscript the $\mathsf{CD}^*(K, N)$ condition will be more common than other synthetic curvature-dimension conditions.

One of the main advantages of the RCD condition w.r.t. the CD one is the fact that it generalizes Riemannian geometry ruling out Finsler-like structures and this enables to get the stability results already cited in the Introduction as well as heat kernel bounds and Laplacian estimates.

As regards the geometric features of finite-dimensional $\mathsf{RCD}^*(K, N)$ spaces, we first recall the **Bishop-Gromov inequality** (see [115], [116]), since it actually holds true on finite-dimensional $\mathsf{CD}^*(K, N)$ spaces: for any $x \in \mathrm{supp}(\mathfrak{m})$ and

- for any $0 < r \leq R < \infty$ if $K \leq 0$
- for any $0 < r \leq R \leq \sqrt{\frac{N-1}{K}}\pi$ if $K > 0$ (because of the Bonnet-Myers maximal diameter theorem, see [116])

it holds

$$\frac{\mathfrak{m}(B_r(x))}{\mathfrak{m}(B_R(x))} \geq \begin{cases} \dfrac{\int_0^r \sin(t\sqrt{K/(N-1)})^{N-1}\mathrm{d}t}{\int_0^R \sin(t\sqrt{K/(N-1)})^{N-1}\mathrm{d}t} & \text{if } K > 0 \\ \left(\dfrac{r}{R}\right)^N & \text{if } K = 0 \\ \dfrac{\int_0^r \sinh(t\sqrt{-K/(N-1)})^{N-1}\mathrm{d}t}{\int_0^R \sinh(t\sqrt{-K/(N-1)})^{N-1}\mathrm{d}t} & \text{if } K < 0 \end{cases} . \tag{1.2.2}$$

It is also worth recalling its spherical version, which reads as

$$\frac{\mathfrak{s}_r(x)}{\mathfrak{s}_R(x)} \geq \begin{cases} \left(\dfrac{\sin(r\sqrt{K/(N-1)})}{\sin(R\sqrt{K/(N-1)})}\right)^{N-1} & \text{if } K>0 \\ \left(\dfrac{r}{R}\right)^{N-1} & \text{if } K=0 \\ \left(\dfrac{\sinh(r\sqrt{-K/(N-1)})}{\sinh(R\sqrt{-K/(N-1)})}\right)^{N-1} & \text{if } K<0 \end{cases} \tag{1.2.3}$$

where

$$\mathfrak{s}_r(x) := \limsup_{\delta\downarrow 0} \frac{1}{\delta}\mathfrak{m}(\overline{B_{r+\delta}(x)} \setminus B_r(x)).$$

Such result has a couple of interesting consequences. First of all, it implies that $\mathfrak{m}$ is uniformly locally doubling with an explicit expression for the local doubling costant, i.e. for all $x \in \mathrm{X}$ and $r > 0$ it holds

$$\mathfrak{m}(B_{2r}(x)) \leq 2^N \cosh\left(2\sqrt{\frac{-K}{N-1}}r\right)^{N-1}\mathfrak{m}(B_r(x)); \tag{1.2.4}$$

in the case $K \geq 0$, $\mathfrak{m}$ is doubling with doubling constant given by 2^N, that is for all $x \in \mathrm{X}$ and $r > 0$ we have

$$\mathfrak{m}(B_{2r}(x)) \leq 2^N \mathfrak{m}(B_r(x)).$$

Analogous to (1.2.4) is the following volume growth condition: for all $x \in \mathrm{X}$ there exists a positive constant C depending on K, N, x only such that

$$\mathfrak{m}(B_r(x)) \leq Ce^{Cr}, \qquad \forall r > 0 \tag{1.2.5}$$

and this ensures that if $\mathfrak{m}$ is Radon (as it is always assumed), then it is finite on every bounded set.

In the case of $\mathsf{RCD}(K,\infty)$ spaces all these properties are no longer true, but a volume growth control (worse than (1.2.5), of course) is still available, as proved in [115] (see Theorem 4.24 therein); namely, for all $x \in \mathrm{X}$ there exists a positive constant C depending on K, x only such that

$$\mathfrak{m}(B_r(x)) \leq Ce^{Cr^2}, \qquad \forall r > 0 \tag{1.2.6}$$

and thus $\mathfrak{m}$ is still finite on bounded sets, provided it is Radon. It also implies that

$$\int e^{-M\mathsf{d}^2(\cdot,\bar{x})}\mathrm{d}\mathfrak{m} < +\infty \tag{1.2.7}$$

for some $\bar{x} \in \mathrm{X}$ and $M > 0$ sufficiently large and this is crucial for the Boltzmann-Shannon entropy to be well defined, as we will see in Chapter 5. In the case of finite-dimensional $\mathsf{RCD}^*(K,N)$ spaces, where (1.2.5) holds, $\mathsf{d}^2(\cdot,\bar{x})$ can be replaced by $\mathsf{d}(\cdot,\bar{x})$ or M can be any positive constant.

1.2.1 Second order calculus

After this first bit of information, let us review the analytical properties of RCD spaces and sketch how a second order differential structure can be introduced over them.

If $(\mathrm{X}, \mathsf{d}, \mathfrak{m})$ is a $\mathsf{RCD}(K, \infty)$ space, it has been proved in [7] and [4] that there exists the **heat kernel**, namely a function

$$(0,\infty) \times \mathrm{X}^2 \ni (t,x,y) \quad \mapsto \quad \mathsf{r}_t[x](y) \in (0,\infty) \tag{1.2.8}$$

which is symmetric, satisfies the Chapman-Kolmogorov formula and provides an integral representation of the heat flow, i.e.

$$\mathsf{r}_t[x](y) = \mathsf{r}_t[y](x) \qquad \text{for } \mathfrak{m}\otimes\mathfrak{m}\text{-a.e. } (x,y) \in \mathrm{X}^2,\ \forall t>0, \tag{1.2.9a}$$

$$\mathsf{r}_{t+s}[x](y) = \int \mathsf{r}_t[x](z)\mathsf{r}_s[z](y)\mathrm{d}\mathfrak{m}(z) \qquad \text{for } \mathfrak{m}\otimes\mathfrak{m}\text{-a.e. } (x,y) \in \mathrm{X}^2,\ \forall t,s\geq 0, \tag{1.2.9b}$$

$$\mathsf{h}_t f(x) = \int f(y)\mathsf{r}_t[x](y)\,\mathrm{d}\mathfrak{m}(y) \qquad \forall t>0,\ \forall f\in L^2(\mathrm{X}). \tag{1.2.9c}$$

For every $x \in \mathrm{X}$ and $t > 0$, $\mathsf{r}_t[x]$ is a probability density and thus (1.2.9c) can be used to extend the heat flow to $L^1(\mathrm{X})$ and shows that the flow is **mass preserving** and satisfies the **maximum principle**, i.e.

$$f \leq c \quad \mathfrak{m}\text{-a.e.} \qquad \Rightarrow \qquad \mathsf{h}_t f \leq c \quad \mathfrak{m}\text{-a.e.},\ \forall t>0. \tag{1.2.10}$$

One of the advantages of a finite-dimensionality assumption is that for $\mathsf{RCD}^*(K,N)$ spaces with $N < \infty$ it has been recently proved that (see [71], where the approach of [113], [114] is adapted to the RCD setting) the heat kernel satisfies **Gaussian estimates**, i.e. for every $\delta > 0$ there are positive constants $C_1 = C_1(K,N,\delta)$ and $C_2 = C_2(K,N,\delta)$ such that for every $x, y \in \mathrm{X}$ and $t > 0$ it holds

$$\frac{1}{C_1\mathfrak{m}(B_{\sqrt{t}}(y))}\exp\Big(-\frac{\mathsf{d}^2(x,y)}{(4-\delta)t} - C_2 t\Big) \leq \mathsf{r}_t[x](y) \leq \frac{C_1}{\mathfrak{m}(B_{\sqrt{t}}(y))}\exp\Big(-\frac{\mathsf{d}^2(x,y)}{(4+\delta)t} + C_2 t\Big). \tag{1.2.11}$$

Passing to the regularization features of the heat semigroup, if $(\mathrm{X}, \mathsf{d}, \mathfrak{m})$ is a $\mathsf{RCD}(K,\infty)$ space then it is well known that h_t is a contraction from $L^2(\mathrm{X})$ into itself and a bounded operator from $L^2(\mathrm{X})$ into $W^{1,2}(\mathrm{X})$; furthermore, the following a priori estimates hold true for every $f \in L^2(\mathrm{X})$ and $t > 0$:

$$E(\mathsf{h}_t f) \leq \frac{1}{4t}\|f\|^2_{L^2(\mathrm{X})}, \tag{1.2.12a}$$

$$\|\Delta \mathsf{h}_t f\|^2_{L^2(\mathrm{X})} \leq \frac{1}{2t^2}\|f\|^2_{L^2(\mathrm{X})}. \tag{1.2.12b}$$

The proof can be found for instance in [56], following the arguments in Section 3.4.4.

A further important property is the **Bakry-Émery contraction estimate** (see [7] and [60]):

$$|\mathrm{d}\mathsf{h}_t f|^2 \leq e^{-2Kt}\mathsf{h}_t(|\mathrm{d}f|^2) \qquad \forall f \in W^{1,2}(\mathrm{X}),\ t \geq 0. \tag{1.2.13}$$

We also recall that $\mathsf{RCD}(K,\infty)$ spaces have the **Sobolev-to-Lipschitz** property, as pointed out in [7] and [55], i.e.

$$f \in W^{1,2}(\mathrm{X}),\ |\mathrm{d}f| \in L^\infty(\mathrm{X}) \qquad \Rightarrow \qquad \exists \tilde{f} = f\ \mathfrak{m}\text{-a.e. with } \mathrm{Lip}(\tilde{f}) \leq \||\mathrm{d}f|\|_{L^\infty}, \tag{1.2.14}$$

and thus we shall typically identify Sobolev functions with bounded differentials with their Lipschitz representative; in particular this will be the case for functions in Test(X), whose definition is going to be immediately provided. Still from [7] we know that

$$f \in L^2 \cap L^\infty(\mathrm{X}),\ t > 0 \qquad \Rightarrow \qquad \mathsf{h}_t(f) \in \mathrm{Test}(\mathrm{X}). \tag{1.2.15}$$

Since already appeared twice, let us introduce the vector space Test(X) of 'test functions' on $\mathsf{RCD}(K,\infty)$ spaces: relying on the results of [109], this can be done as follows

$$\mathrm{Test}(\mathrm{X}) := \Big\{ f \in D(\Delta) \cap L^\infty(\mathrm{X}) \,:\, |\nabla f| \in L^\infty(\mathrm{X}),\, \Delta f \in W^{1,2}(\mathrm{X}) \Big\}$$

and notice that this is an algebra dense in $W^{1,2}(\mathrm{X})$. We shall also make use of the vector space

$$\mathrm{Test}^\infty(\mathrm{X}) := \Big\{ f \in D(\Delta) \cap L^\infty(\mathrm{X}) \ :\ |\nabla f| \in L^\infty(\mathrm{X}),\ \Delta f \in L^\infty \cap W^{1,2}(\mathrm{X}) \Big\}$$

which is an algebra dense in $W^{1,2}(\mathrm{X})$ too. The fact that $\mathrm{Test}(\mathrm{X}), \mathrm{Test}^\infty(\mathrm{X})$ are algebras is based on the following property, proved in [109] (see Lemma 3.2 therein)

$$\begin{aligned} f \in \mathrm{Test}(\mathrm{X}) \qquad \Rightarrow \qquad & |\mathrm{d}f|^2 \in W^{1,2}(\mathrm{X}) \quad \text{with} \\ & \int |\mathrm{d}(|\mathrm{d}f|^2)|^2 \,\mathrm{d}\mathfrak{m} \leq \||\mathrm{d}f|\|_{L^\infty}^2 \Big(\||\mathrm{d}f|\|_{L^2} \||\mathrm{d}\Delta f|\|_{L^2} + |K| \||\mathrm{d}f|\|_{L^2}^2 \Big) \end{aligned} \tag{1.2.16}$$

and actually a further regularity property of test functions, stated in Lemma 3.2 of [109] too, is that

$$f \in \mathrm{Test}(\mathrm{X}) \qquad \Rightarrow \qquad |\mathrm{d}f|^2 \in D(\boldsymbol{\Delta}), \tag{1.2.17}$$

so that it is possible to introduce the **measure-valued Γ_2 operator** as

$$\boldsymbol{\Gamma}_2(f) := \boldsymbol{\Delta} \frac{|\mathrm{d}f|^2}{2} - \langle \nabla f, \nabla \Delta f \rangle\, \mathfrak{m} \qquad \forall f \in \mathrm{Test}(\mathrm{X}).$$

By construction, the assignment $f \mapsto \boldsymbol{\Gamma}_2(f)$ is a quadratic form.

The existence of the space of test functions and the language of $L^2(\mathrm{X})$-normed $L^\infty(\mathrm{X})$-modules allow to introduce the space $W^{2,2}(\mathrm{X})$ and $W^{1,2}_C(T\mathrm{X})$, following the procedure depicted in [56], and thus the notions of Hessian and covariant derivative. We first consider the **tensor product** $L^2((T^*)^{\otimes 2}\mathrm{X})$ of the cotangent module $L^2(T^*\mathrm{X})$ with itself. Roughly speaking, this is built in the following way: consider the algebraic tensor product, namely the space of formal finite sums of objects of the kind $\omega_1 \otimes \omega_2$ with $\omega_1, \omega_2 \in L^2(T^*\mathrm{X})$ having the bilinearity property and satisfying $f(\omega_1 \otimes \omega_2) = (f\omega_1) \otimes \omega_2 = \omega_1 \otimes (f\omega_2)$ for all $f \in L^\infty(\mathrm{X})$; endow it with the bilinear form : defined by

$$(\omega_1 \otimes \omega_2) : (\tilde{\omega}_1 \otimes \tilde{\omega}_2) := \langle \omega_1, \tilde{\omega}_1 \rangle \langle \omega_2, \tilde{\omega}_2 \rangle, \qquad \forall \omega_1, \omega_2, \tilde{\omega}_1, \tilde{\omega}_2 \in L^2(T^*\mathrm{X}),$$

where $\langle \cdot\,,\,\cdot \rangle$ is the scalar product of $L^2(T^*\mathrm{X})$, and taking values in $L^0(\mathrm{X})$ equipped with the topology of $\mathfrak{m}$-a.e. convergence. Then the tensor product $L^2((T^*)^{\otimes 2}\mathrm{X})$ is defined as the completion of the space of A's belonging to the algebraic tensor product such that $A : A \in L^1(\mathrm{X})$ equipped with the norm

$$\|A\|_{L^2((T^*)^{\otimes 2}\mathrm{X})} := \sqrt{\int A : A \,\mathrm{d}\mathfrak{m}}.$$

Introducing the pointwise norm $|\cdot|_{\mathsf{HS}} : L^2((T^*)^{\otimes 2}\mathrm{X}) \to L^0(\mathrm{X})$ as $|A|_{\mathsf{HS}} := \sqrt{A : A}$, denoted in this way to remind that in the smooth case it coincides with the Hilbert-Schmidt one, it is not difficult to see that $L^2((T^*)^{\otimes 2}\mathrm{X})$ has a canonical structure of $L^2(\mathrm{X})$-normed module. In the same way it is possible to build the tensor product of $L^2(T\mathrm{X})$ with itself, denoted by $L^2(T^{\otimes 2}\mathrm{X})$.

After this premise, we say that a function $f \in W^{1,2}(\mathrm{X})$ belongs to $W^{2,2}(\mathrm{X})$ provided there exists $A \in L^2((T^*)^{\otimes 2}\mathrm{X})$ symmetric, i.e. such that $A(W_1, W_2) = A(W_2, W_1)$ $\mathfrak{m}$-a.e. for every $W_1, W_2 \in L^2(T\mathrm{X})$, for which it holds

$$\int hA(\nabla g, \nabla g)\,\mathrm{d}\mathfrak{m} = \int -\langle \nabla f, \nabla g\rangle \,\mathrm{div}(h\nabla g) - h\langle \nabla f, \nabla \frac{|\nabla g|^2}{2}\rangle \,\mathrm{d}\mathfrak{m} \qquad \forall g, h \in \mathrm{Test}(\mathrm{X}).$$

In this case A is unique, called **Hessian** of f and denoted by $\mathrm{Hess}(f)$. The space $W^{2,2}(\mathrm{X})$ endowed with the norm

$$\|f\|^2_{W^{2,2}(\mathrm{X})} := \|f\|^2_{L^2(\mathrm{X})} + \|\mathrm{d}f\|^2_{L^2(T^*\mathrm{X})} + \|\mathrm{Hess}(f)\|^2_{L^2((T^*)^{\otimes 2}\mathrm{X})}$$

is a separable Hilbert space which contains $\mathrm{Test}(\mathrm{X})$ and in particular is dense in $W^{1,2}(\mathrm{X})$. It is proved in [56] that $D(\Delta) \subset W^{2,2}(\mathrm{X})$ with

$$\int |\mathrm{Hess}(f)|^2_{\mathsf{HS}}\,\mathrm{d}\mathfrak{m} \le \int (\Delta f)^2 - K|\nabla f|^2\,\mathrm{d}\mathfrak{m} \qquad \forall f \in D(\Delta). \tag{1.2.18}$$

However, it is unknown whether $D(\Delta)$ is dense in $W^{2,2}(\mathrm{X})$: for this reason we define the space $H^{2,2}(\mathrm{X})$ as its $W^{2,2}(\mathrm{X})$-closure. The Hessian is a local operator, in a sense which is more subtle than the one already discussed for the differential; indeed, for given $f, g \in W^{2,2}(\mathrm{X})$ we have

$$\mathrm{Hess}(f) = \mathrm{Hess}(g), \qquad \mathfrak{m}\text{-a.e. on the interior of } \{f = g\} \tag{1.2.19}$$

whereas if $f, g \in H^{2,2}(\mathrm{X})$, then

$$\mathrm{Hess}(f) = \mathrm{Hess}(g), \qquad \mathfrak{m}\text{-a.e. on } \{f = g\} \tag{1.2.20}$$

holds.

On the other hand, a vector $W \in L^2(T\mathrm{X})$ belongs to $W^{1,2}_C(T\mathrm{X})$ if there exists $T \in L^2(T^{\otimes 2}\mathrm{X})$ such that

$$\int hT : (\nabla g_1 \otimes \nabla g_2)\,\mathrm{d}\mathfrak{m} = \int -\langle W, \nabla g_2\rangle \mathrm{div}(h\nabla g_1) - h\mathrm{Hess}(g_2)(W, \nabla g_1)\,\mathrm{d}\mathfrak{m}$$
$$\forall g_1, g_2, h \in \mathrm{Test}(\mathrm{X}).$$

In this case T is unique, called **covariant derivative** of W and denoted by ∇W. The space $W^{1,2}_C(T\mathrm{X})$ endowed with the norm

$$\|W\|^2_{W^{1,2}_C(T\mathrm{X})} := \|W\|^2_{L^2(T\mathrm{X})} + \|\nabla W\|^2_{L^2(T^{\otimes 2}\mathrm{X})}$$

is a separable Hilbert space which contains the class of 'test vector fields' TestV(X), given by

$$\mathrm{TestV(X)} := \Big\{ \sum_{i=1}^n g_i \nabla f_i \,:\, n \in \mathbb{N},\, f_i, g_i \in \mathrm{Test(X)}\, i = 1, ..., n \Big\}.$$

By the properties of Test(X) it is not difficult to see that TestV(X) is dense in $L^2(T\mathrm{X})$, so that $W^{1,2}_C(T\mathrm{X})$ is dense in $L^2(T\mathrm{X})$ as well; in addition, $\mathrm{TestV(X)} \subset L^1 \cap L^\infty(\mathrm{X})$ and $\mathrm{TestV(X)} \subset D(\mathrm{div})$. However, it is not known yet whether TestV(X) is dense in $W^{1,2}_C(T\mathrm{X})$ or not: this motivates the introduction of the space $H^{1,2}_C(T\mathrm{X})$ as the $W^{1,2}_C(T\mathrm{X})$-closure of TestV(X).

As for cotangent and tangent modules, $L^2((T^*)^{\otimes 2}\mathrm{X})$ and $L^2(T^{\otimes 2}\mathrm{X})$ are canonically isomorphic both as Hilbert modules and as Hilbert spaces via the Riesz theorem. Thus, with a little abuse of notation, we introduce the (musical) isomorphisms $\flat : L^2(T^{\otimes 2}\mathrm{X}) \to L^2((T^*)^{\otimes 2}\mathrm{X})$ and $\sharp : L^2((T^*)^{\otimes 2}\mathrm{X}) \to L^2(T^{\otimes 2}\mathrm{X})$ as

$$T^\flat(S) := T : S, \qquad A^\sharp : T := A(T)$$

$\mathfrak{m}$-a.e. for every $T, S \in L^2(T^{\otimes 2}\mathrm{X})$ and $A \in L^2((T^*)^{\otimes 2}\mathrm{X})$.

As regards the calculus tools directly linked to the notions we have just introduced, they are collected in the following lemma. The proof can be found in [56] (Proposition 3.3.22 and Theorem 3.4.2 therein).

Lemma 1.2.3 (Calculus rules 2)**.** *Let* $(\mathrm{X}, \mathsf{d}, \mathfrak{m})$ *be a* $\mathsf{RCD}^*(K, N)$ *with* $K \in \mathbb{R}$ *and* $N \in [1, \infty)$*. Then:*

(i) (Leibniz rule) for all $f, g \in \mathrm{Test(X)}$ *it holds*

$$\mathrm{d}\,\langle \nabla f, \nabla g\rangle = \mathrm{Hess}(f)(\nabla g, \cdot) + \mathrm{Hess}(g)(\nabla f, \cdot) \qquad \mathfrak{m}\text{-a.e.} \tag{1.2.21}$$

(ii) for all $W = \sum_i g_i \nabla f_i \in \mathrm{TestV(X)}$ *we have*

$$\nabla W = \sum_i \nabla g_i \otimes \nabla f_i + g_i(\mathrm{Hess}(f_i))^\sharp. \tag{1.2.22}$$

The second order structure of RCD spaces also allows to talk about the **Bochner inequality**: on $\mathsf{RCD}(K, \infty)$ spaces it takes the form of an inequality between measures ([56] - see also the previous contributions [109], [118]):

$$\mathbf{\Gamma}_2(f) \geq \big(|\mathrm{Hess}(f)|^2_{\mathsf{HS}} + K|\mathrm{d}f|^2\big)\mathfrak{m} \qquad \forall f \in \mathrm{Test(X)}, \tag{1.2.23}$$

and if the space is $\mathsf{RCD}^*(K, N)$ for some finite N it also holds ([8]):

$$\mathbf{\Gamma}_2(f) \geq \Big(\frac{(\Delta f)^2}{N} + K|\mathrm{d}f|^2\Big)\mathfrak{m} \qquad \forall f \in \mathrm{Test(X)}. \tag{1.2.24}$$

Notice that since the Laplacian is in general not the trace of the Hessian, the former does not trivially imply the latter (in connection to this, see [65]). Furthermore (1.2.24) is the RCD analogue of the Bakry-Émery curvature-dimension condition with dimension term introduced in [14]; the relationship between the theory of $\mathsf{CD}^*(K,N)$ and $\mathsf{RCD}^*(K,N)$ spaces on the one hand and the Γ-calculus approach on the other one is fully described in [45] and [10].

We conclude the section recalling the notion of Regular Lagrangian Flow, introduced by Ambrosio-Trevisan in [11] as the generalization to RCD spaces of the analogous concept existing on $\mathbb{R}^d$ as proposed by Ambrosio in [1]:

Definition 1.2.4 (Regular Lagrangian Flow)**.** *Given $(v_t) \in L^1([0,1], L^2(T\mathrm{X}))$, the function $F : [0,1] \times \mathrm{X} \to \mathrm{X}$ is a Regular Lagrangian Flow for (v_t) provided:*

i) $[0,1] \ni t \mapsto F_t(x)$ is continuous for every $x \in \mathrm{X}$

ii) for every $f \in \mathrm{Test}^\infty(\mathrm{X})$ and $\mathfrak{m}$-a.e. x the map $t \mapsto f(F_t(x))$ belongs to $W^{1,1}([0,1])$ and

$$\frac{\mathrm{d}}{\mathrm{d}t} f(F_t(x)) = \mathrm{d}f(v_t)(F_t(x)) \qquad \text{a.e. } t \in [0,1].$$

iii) it holds

$$(F_t)_*\mathfrak{m} \le C\mathfrak{m} \qquad \forall t \in [0,1]$$

for some constant $C > 0$.

In [11] the authors prove that under suitable assumptions on the v_t's, Regular Lagrangian Flows exist and are unique. We shall use the following formulation of their result (weaker than the one provided in [11]):

Theorem 1.2.5. *Let $(\mathrm{X}, \mathsf{d}, \mathfrak{m})$ be a $\mathsf{RCD}(K,\infty)$ space and $(v_t) \in L^1([0,1], L^2(T\mathrm{X}))$ be such that $v_t \in D(\mathrm{div})$ for a.e. t and*

$$\mathrm{div}(v_t) \in L^1([0,1], L^2(\mathrm{X})) \qquad (\mathrm{div}(v_t))^- \in L^1([0,1], L^\infty(\mathrm{X})).$$

Then there exists a unique, up to $\mathfrak{m}$-a.e. equality, Regular Lagrangian Flow F for (v_t).

For such flow, the quantitative bound

$$(F_t)_*\mathfrak{m} \le \exp\Big(\int_0^1 \|(\mathrm{div}(v_t))^-\|_{L^\infty(\mathrm{X})}\,\mathrm{d}t\Big)\mathfrak{m} \tag{1.2.25}$$

holds for every $t \in [0,1]$ and for $\mathfrak{m}$-a.e. x the curve $t \mapsto F_t(x)$ is absolutely continuous and its metric speed $\mathrm{ms}_t(F_\cdot(x))$ at time t satisfies

$$\mathrm{ms}_t(F_\cdot(x)) = |v_t|(F_t(x)) \qquad \text{a.e. } t \in [0,1]. \tag{1.2.26}$$

To be precise, (1.2.26) is not explicitly stated in [11]; its proof is anyway not hard and can be obtained, for instance, following the arguments in [56].

1.2.2 A second glance at optimal transport

With this said, we can come back to optimal transport theory and resume the discussion about W_2-absolutely continuous curves and the continuity equation. As we are going to notice immediately, the curvature-dimension assumption has a deep impact on the structure of Wasserstein geodesics. We first recall the following version of Brenier-McCann theorem on RCD spaces ((i) comes from [54] and [106], (ii) from [7] and [58], (iii) from [6] and (iv) from [62]).

Theorem 1.2.6. *Let $(\mathrm{X}, \mathsf{d}, \mathfrak{m})$ be a $\mathsf{RCD}(K, \infty)$ space and $\mu_0, \mu_1 \in \mathscr{P}_2(\mathrm{X})$ with bounded support and such that $\mu_0, \mu_1 \leq C\mathfrak{m}$ for some $C > 0$. Also, let φ be a Kantorovich potential for the couple (μ_0, μ_1) which is locally Lipschitz on a neighbourhood of $\mathrm{supp}(\mu_0)$. Then:*

(i) There exists a unique geodesic (μ_t) from μ_0 to μ_1, it satifies

$$\mu_t \leq C'\mathfrak{m} \qquad \forall t \in [0,1] \text{ for some } C' > 0 \tag{1.2.27}$$

and there is a unique lifting $\boldsymbol{\pi}$ *of it, i.e. a unique measure $\boldsymbol{\pi} \in \mathscr{P}(C([0,1], \mathrm{X}))$ such that $(\mathrm{e}_t)_* \boldsymbol{\pi} = \mu_t$ for every $t \in [0,1]$ and*

$$\iint_0^1 |\dot{\gamma}_t|^2 \, \mathrm{d}t \, \mathrm{d}\boldsymbol{\pi}(\gamma) = W_2^2(\mu_0, \mu_1).$$

(ii) For every $f \in W^{1,2}(\mathrm{X})$ the map $t \mapsto \int f \, \mathrm{d}\mu_t$ is differentiable at $t = 0$ and

$$\frac{\mathrm{d}}{\mathrm{d}t} \int f \, \mathrm{d}\mu_t|_{t=0} = -\int \mathrm{d}f(\nabla\varphi) \, \mathrm{d}\mu_0.$$

(iii) The identity

$$|\mathrm{d}\varphi|(\gamma_0) = |D^+\varphi|(\gamma_0) = \mathsf{d}(\gamma_0, \gamma_1)$$

holds for $\boldsymbol{\pi}$-a.e. γ.

(iv) If the space is $\mathsf{RCD}^*(K, N)$ for some $N < \infty$, then $(i), (ii), (iii)$ holds with μ_1 only assumed to be with bounded support, with the caveat that (1.2.27) holds in the form: for every $\delta \in (0, 1/2)$ there is $C_\delta > 0$ so that $\mu_t \leq C'_\delta \mathfrak{m}$ for every $t \in [0, 1-\delta]$. In addition, for every $x \in \mathrm{X}$ the following holds: for $\mathfrak{m}$-a.e. y there is a unique geodesic connecting y to x.

A property related to the above is the fact that although the Kantorovich potentials are not uniquely determined by the initial and final measures, their gradients are. This is expressed by the following result, which also says that if we sit in the intermediate point of a geodesic and move to one extreme or the other, then the two corresponding velocities are one the opposite of the other (see Lemma 5.8 and Lemma 5.9 in [55] for the proof):

Lemma 1.2.7. *Let $(\mathrm{X}, \mathsf{d}, \mathfrak{m})$ be a $\mathsf{RCD}(K, \infty)$ space with $K \in \mathbb{R}$ and $(\mu_t) \subset \mathscr{P}_2(\mathrm{X})$ a W_2-geodesic such that $\mu_t \leq C\mathfrak{m}$ for every $t \in [0,1]$ for some $C > 0$. For $t \in [0,1]$ let $\phi_t, \phi'_t : \mathrm{X} \to \mathbb{R}$ be locally Lipschitz functions such that for some $s, s' \neq t$ the functions $-(s-t)\phi_t$ and $-(s'-t)\phi'_t$ are Kantorovich potentials from μ_t to μ_s and from μ_t to $\mu_{s'}$ respectively.*

Then

$$\nabla\phi_t = \nabla\phi'_t \qquad \mu_t\text{-a.e..}$$

On RCD spaces, W_2-geodesics made of measures with bounded density also have the weak continuity property of the densities expressed by the following lemma. The proof follows by a simple argument involving Young's measures and the continuity of the entropy along a geodesic (see Corollary 5.7 in [55]):

Lemma 1.2.8. *Let* $(\mathrm{X}, \mathsf{d}, \mathfrak{m})$ *be a* $\mathsf{RCD}(K, \infty)$ *space with* $K \in \mathbb{R}$ *and* $(\mu_t) \subset \mathscr{P}_2(X)$ *a* W_2*-geodesic such that* $\mu_t \leq C\mathfrak{m}$ *for every* $t \in [0,1]$ *for some* $C > 0$*. Let* ρ_t *be the density of* μ_t*.*

Then for any $t \in [0,1]$ *and any sequence* $(t_n)_{n\in\mathbb{N}} \subset [0,1]$ *converging to* t *there exists a subsequence* $(t_{n_k})_{k\in\mathbb{N}}$ *such that*

$$\rho_{t_{n_k}} \to \rho_t, \quad \mathfrak{m}\text{-a.e.}$$

as $k \to \infty$*.*

Chapter 2

Old estimates in a new setting

In this chapter we carry on the dissertation of the regularizing properties of the heat flow on RCD spaces, focusing the attention on gradient and Laplacian estimates for $\log \mathsf{h}_t u$. The former, also known as Hamilton's gradient estimate and singled out in [64] for compact Riemannian manifolds and later extended in [74] to non-compact ones, says that for any $u_0 \in L^p \cap L^\infty(\mathrm{X})$ positive with $p \in [1, \infty)$ it holds

$$t|\nabla \log u_t|^2 \leq (1 + 2K^- t) \log \left(\frac{\|u_0\|_{L^\infty(\mathfrak{m})}}{u_t} \right), \qquad \mathfrak{m}\text{-a.e.}$$

for all $t > 0$, where $u_t := \mathsf{h}_t u_0$, K is a lower bound on the Ricci curvature and $K^- := \max\{0, -K\}$. The latter was proved in the seminal paper [83] by P. Li and S.-T. Yau, whence its name, for Riemannian manifolds with non-negative Ricci curvature, where it reads as

$$\Delta \log u_t \geq -\frac{N}{2t}, \qquad \mathfrak{m}\text{-a.e.} \tag{2.0.1}$$

for all $t > 0$, N being the dimension of the manifold, or by trivial manipulations

$$|\nabla \log u_t|^2 - \frac{\Delta u_t}{u_t} \leq \frac{N}{2t}, \qquad \mathfrak{m}\text{-a.e.}$$

During the last years many efforts have been done in order to generalize such inequalities to the metric framework and the results are very recent. As we learnt while working on [63], Hamilton's gradient estimate has already been proved on proper $\mathsf{RCD}(K, \infty)$ spaces by Jiang and Zhang in [72] but, as we will later point out, the assumption for the space to be proper can be removed. On the other hand, the Li-Yau inequality as stated above is known on $\mathsf{RCD}^*(0, N)$ spaces from [51] and [70]; in the same papers, the case of negative curvature is treated too, thus establishing the following Baudoin-Garofalo inequality

$$|\nabla \log u_t|^2 \leq e^{-2Kt/3} \frac{\Delta u_t}{u_t} + \frac{NK}{3} \frac{e^{-4Kt/3}}{1 - e^{-2Kt/3}}, \qquad \mathfrak{m}\text{-a.e.}$$

Section 2.1 is based on [63] and entirely devoted to old and new results for compact RCD spaces. We present the proof of Hamilton's gradient estimate, although already known by [72] on proper $\mathsf{RCD}(K, \infty)$ spaces, because the original proof of [64] can be adapted almost verbatim and the additional compactness assumption allows to avoid the technicalities of the

non-compact case. In this direction, we also prove a bound which seems new in the non-smooth context, namely a uniform bound on $|\nabla \log \mathsf{h}_t u|$ in the special case $|\nabla \log u| \in L^\infty(\mathrm{X})$, see Proposition 2.1.5. We conclude with a version of the Li-Yau inequality because, even if K is possibly negative, under the compactness hypothesis we get an estimate closer to (2.0.1) rather than to Baudoin-Garofalo inequality.

In Section 2.2 we extend the differential notions of the previous chapter to locally integrable objects and we show that on $\mathsf{RCD}(K,\infty)$ spaces there exist 'good' cut-off functions; this result is not new in the literature, but we point out some bounds on the cut-off functions in terms of the relative position of the sets they divide that have not been explicitly written yet. As a consequence, we prove that $\log \mathsf{h}_t u$ is locally well behaved.

Finally, in Section 2.3 we start from the already cited Hamilton's gradient estimate and Li-Yau inequality and modify them to make them fit to our purposes, stressing also the differences w.r.t. the compact case. In particular, on finite-dimensional RCD spaces we remove the dependence on the L^∞ norm of the initial datum present on the right-hand side of Hamilton's estimate.

2.1 Comments on the compact case

Throughout the whole section $(\mathrm{X}, \mathsf{d}, \mathfrak{m})$ will be a compact RCD space equipped with a probability measure (this is not restrictive, because by the Bishop-Gromov inequality $\mathfrak{m}$ has to be finite). This allows to define a further (and slightly different) class of test functions as follows

$$\mathrm{Test}^\infty_{>0}(\mathrm{X}) := \Big\{ f \in \mathrm{Test}^\infty(\mathrm{X}) \ : \ f \geq c \ \mathfrak{m}\text{-a.e. for some } c > 0 \Big\}$$

and combining the Gaussian estimates on compact $\mathsf{RCD}^*(K,N)$ spaces, $N < \infty$, with the results in [109] we see that

$$\begin{aligned} f \in L^1(\mathrm{X}),\ t > 0 \qquad &\Rightarrow \qquad \mathsf{h}_t(f) \in \mathrm{Test}^\infty(\mathrm{X}), \\ f \in L^1(\mathrm{X}),\ f \geq 0,\ \int f \,\mathrm{d}\mathfrak{m} > 0,\ t > 0 \qquad &\Rightarrow \qquad \mathsf{h}_t(f) \in \mathrm{Test}^\infty_{>0}(\mathrm{X}). \end{aligned} \tag{2.1.1}$$

2.1.1 Comparison principles

The proofs of Hamilton's gradient estimate and of the Li-Yau inequality are based on the following two comparison principles, valid in general infinitesimally Hilbertian spaces $(\mathrm{Y}, \mathsf{d}_\mathrm{Y}, \mathfrak{m}_\mathrm{Y})$.

To formulate the result we need to introduce the dual of $W^{1,2}(\mathrm{Y})$, which we shall denote $W^{-1,2}(\mathrm{Y})$. As usual, the fact that $W^{1,2}(\mathrm{Y})$ embeds in $L^2(\mathrm{Y})$ with dense image allows to see $L^2(\mathrm{Y})$ as a dense subset of $W^{-1,2}(\mathrm{Y})$, where $f \in L^2(\mathrm{Y})$ is identified with the mapping $W^{1,2}(\mathrm{Y}) \ni g \mapsto \int fg \,\mathrm{d}\mathfrak{m}_\mathrm{Y}$.

Notice also that even in this generality, a regularization via the heat flow shows that $D(\Delta)$ is dense in $W^{1,2}(\mathrm{Y})$ and, with the use of the maximum principle (1.2.10), that non-negative functions in $D(\Delta)$ are $W^{1,2}$-dense in the space of non-negative functions in $W^{1,2}$.

Proposition 2.1.1. *Let $(\mathrm{Y}, \mathsf{d}_\mathrm{Y}, \mathfrak{m}_\mathrm{Y})$ be an infinitesimally Hilbertian space. Then the following two comparison principles hold:*

(i) let $(F_t), (G_t) \in AC_{loc}([0,\infty), L^2(\mathrm{Y}))$ be respectively a weak super- and weak sub- solution of the heat equation, i.e. such that for all $h \in D(\Delta)$ non-negative and a.e. $t > 0$ it holds

$$\frac{\mathrm{d}}{\mathrm{d}t}\int hF_t \mathrm{d}\mathfrak{m}_{\mathrm{Y}} \geq \int \Delta h F_t \mathrm{d}\mathfrak{m}_{\mathrm{Y}}, \qquad \frac{\mathrm{d}}{\mathrm{d}t}\int hG_t \mathrm{d}\mathfrak{m}_{\mathrm{Y}} \leq \int \Delta h G_t \mathrm{d}\mathfrak{m}_{\mathrm{Y}}.$$

Assume that $F_0 \geq G_0$ $\mathfrak{m}$-a.e. Then $F_t \geq G_t$ $\mathfrak{m}$-a.e. for every $t > 0$.

(ii) Let $a_0, a_1, a_2 \in \mathbb{R}$ and $(v_t) \in L^1_{loc}([0,\infty), W^{1,2}(\mathrm{Y}))$ with $v_t \in D(\Delta)$ for a.e. t and $\|\Delta v_t\|_{L^\infty} \in L^1_{loc}([0,\infty))$ and let $(F_t), (G_t) \in L^\infty_{loc}([0,\infty), L^\infty(\mathrm{Y})) \cap L^\infty_{loc}([0,\infty), W^{1,2}(\mathrm{Y})) \cap AC_{loc}([0,\infty), W^{-1,2}(\mathrm{Y}))$ be respectively a weak super- and weak sub- solution of

$$\frac{\mathrm{d}}{\mathrm{d}t}u_t = \Delta u_t + a_0 u_t^2 + a_1 u_t + \langle \nabla u_t, \nabla v_t\rangle + a_2 \tag{2.1.2}$$

in the following sense: for all $h \in D(\Delta)$ non-negative and a.e. $t > 0$ it holds

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\int hF_t \mathrm{d}\mathfrak{m}_{\mathrm{Y}} &\geq \int \Delta h F_t \mathrm{d}\mathfrak{m}_{\mathrm{Y}} + \int h\Big(a_0 F_t^2 + a_1 F_t + \langle \nabla F_t, \nabla v_t\rangle + a_2\Big)\mathrm{d}\mathfrak{m}_{\mathrm{Y}},\\ \frac{\mathrm{d}}{\mathrm{d}t}\int hG_t \mathrm{d}\mathfrak{m}_{\mathrm{Y}} &\leq \int \Delta h G_t \mathrm{d}\mathfrak{m}_{\mathrm{Y}} + \int h\Big(a_0 G_t^2 + a_1 G_t + \langle \nabla G_t, \nabla v_t\rangle + a_2\Big)\mathrm{d}\mathfrak{m}_{\mathrm{Y}}.\end{aligned}$$

Assume that $F_0 \geq G_0$ $\mathfrak{m}_{\mathrm{Y}}$-a.e. Then $F_t \geq G_t$ $\mathfrak{m}_{\mathrm{Y}}$-a.e. for every $t > 0$.

proof
(i) By linearity it is not restrictive to assume $G_t \equiv 0$ for all $t \geq 0$. Fix $\varepsilon > 0$, notice that $t \mapsto \mathsf{h}_\varepsilon F_t$ belongs to $AC_{loc}([0,\infty), L^2(\mathrm{Y}))$ with values in $D(\Delta)$. Then pick $h \in D(\Delta)$ non-negative, notice that $\mathsf{h}_\varepsilon h$ is non-negative as well to get

$$\int h\frac{\mathrm{d}}{\mathrm{d}t}\mathsf{h}_\varepsilon F_t \mathrm{d}\mathfrak{m}_{\mathrm{Y}} = \frac{\mathrm{d}}{\mathrm{d}t}\int h\mathsf{h}_\varepsilon F_t \mathrm{d}\mathfrak{m}_Y = \frac{\mathrm{d}}{\mathrm{d}t}\int (\mathsf{h}_\varepsilon h) F_t \mathrm{d}\mathfrak{m}_{\mathrm{Y}} \geq \int \Delta \mathsf{h}_\varepsilon h\, F_t \mathrm{d}\mathfrak{m}_{\mathrm{Y}} = \int h\Delta \mathsf{h}_\varepsilon F_t \mathrm{d}\mathfrak{m}_{\mathrm{Y}}.$$

Since this is true for all $h \in D(\Delta)$ non-negative and, by what we said before, this class of functions is L^2-dense in the set of non-negative L^2-functions, we deduce that for a.e. $t > 0$ it holds

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathsf{h}_\varepsilon F_t \geq \Delta \mathsf{h}_\varepsilon F_t, \quad \mathfrak{m}_Y\text{-a.e.}. \tag{2.1.3}$$

Now notice that being $F_0 \geq 0$, by the maximum principle (1.2.10) we see that $\mathsf{h}_\varepsilon F_0 \geq 0$ too and we claim that from this fact and (2.1.3) it follows that $\mathsf{h}_\varepsilon(F_t) \geq 0$ for every $t \geq 0$. Thus let us consider

$$\Phi(t) := \frac{1}{2}\int |\phi(\mathsf{h}_\varepsilon F_t)|^2 \mathrm{d}\mathfrak{m}_{\mathrm{Y}},$$

where $\phi(z) := z^- = \max\{0, -z\}$. Observe that $\Phi \in AC_{loc}([0,\infty))$, that $\Phi(0) = 0$ and compute

$$\begin{aligned}\Phi'(t) &= \int \phi(\mathsf{h}_\varepsilon F_t)\frac{\mathrm{d}}{\mathrm{d}t}\phi(\mathsf{h}_\varepsilon F_t)\,\mathrm{d}\mathfrak{m}_{\mathrm{Y}} = \int \phi'(\mathsf{h}_\varepsilon F_t)\,\phi(\mathsf{h}_\varepsilon F_t)\frac{\mathrm{d}}{\mathrm{d}t}\mathsf{h}_\varepsilon F_t\,\mathrm{d}\mathfrak{m}_{\mathrm{Y}}\\ &= -\int \phi(\mathsf{h}_\varepsilon F_t)\frac{\mathrm{d}}{\mathrm{d}t}\mathsf{h}_\varepsilon F_t\,\mathrm{d}\mathfrak{m}_{\mathrm{Y}}\end{aligned}$$

and therefore taking (2.1.3) into account we see that

$$\begin{aligned}\Phi'(t) &\le -\int \phi(\mathsf{h}_\varepsilon F_t)\Delta \mathsf{h}_\varepsilon F_t \,\mathrm{d}\mathfrak{m}_{\mathrm{Y}} = \int \langle \nabla\phi(\mathsf{h}_\varepsilon F_t), \nabla \mathsf{h}_\varepsilon F_t\rangle \,\mathrm{d}\mathfrak{m}_{\mathrm{Y}}\\ &= -\int |\nabla\phi(\mathsf{h}_\varepsilon F_t)|^2 \,\mathrm{d}\mathfrak{m}_{\mathrm{Y}} \le 0.\end{aligned}$$

Thus $\Phi(t)=0$ for every $t\ge 0$, i.e. $\mathsf{h}_\varepsilon F_t \ge 0$ for all $t\ge 0$. Letting $\varepsilon\downarrow 0$ we conclude.
(ii) Since $(F_t)\in L^\infty_{loc}([0,\infty), W^{1,2}(\mathrm{Y}))$, the fact that it is a supersolution of (2.1.2) can be written as

$$\frac{\mathrm{d}}{\mathrm{d}t}\int hF_t \mathrm{d}\mathfrak{m}_{\mathrm{Y}} \ge -\int \langle \nabla h, \nabla F_t\rangle \,\mathrm{d}\mathfrak{m}_{\mathrm{Y}} + \int h\Big(a_0F_t^2 + a_1F_t + \langle \nabla F_t, \nabla v_t\rangle + a_2\Big)\mathrm{d}\mathfrak{m}_{\mathrm{Y}} \tag{2.1.4}$$

for every $h\in D(\Delta)$ non-negative. Recalling that the class of such functions is $W^{1,2}$-dense in the one of non-negative $W^{1,2}$ functions, passing through the integral formulation - in time - of (2.1.4) it is immediate to see that (2.1.4) also holds for any $h\in W^{1,2}(\mathrm{Y})$ non-negative. Using the fact that $W^{-1,2}(\mathrm{Y})$ has the Radon-Nikodym property (because it is Hilbert) we see that (F_t) seen as curve with values in $W^{-1,2}(\mathrm{Y})$ must be differentiable at a.e. t and it is then clear that for any point of differentiability t, the inequality (2.1.4) holds for any $h\in W^{1,2}(\mathrm{Y})$ non-negative, i.e. that the set of t's for which (2.1.4) holds is independent on h. The analogous property holds for (G_t).

Now we apply Lemma 2.1.2 below to $h_t := G_t - F_t$ to get that $\Phi(t) := \frac{1}{2}\int |(G_t-F_t)^+|^2 \,\mathrm{d}\mathfrak{m}_Y$ is absolutely continuous and

$$\Phi'(t) = \int (G_t-F_t)^+ \frac{\mathrm{d}}{\mathrm{d}t}(G_t-F_t)\,\mathrm{d}\mathfrak{m}_Y,$$

where the right hand side is intended as the coupling of $\frac{\mathrm{d}}{\mathrm{d}t}(G_t-F_t)\in W^{-1,2}(\mathrm{Y})$ and the function $(G_t-F_t)^+\in W^{1,2}(\mathrm{Y})$. Fix t which is a differentiability point of both (F_t) and (G_t), pick $h:=(G_t-F_t)^+$ in (2.1.4) and in the analogous inequality for (G_t) to obtain

$$\begin{aligned}\Phi'(t) \le \int &-\big\langle \nabla((G_t-F_t)^+), \nabla(G_t-F_t)\big\rangle\\ &+ (G_t-F_t)^+\Big(a_0(G_t^2-F_t^2) + a_1(G_t-F_t) + \langle \nabla(G_t-F_t), \nabla v_t\rangle\Big)\,\mathrm{d}\mathfrak{m}_{\mathrm{Y}}\end{aligned}$$

and since $\langle \nabla h^+, \nabla h\rangle = |\nabla h^+|^2$ and $h^+\nabla h = \frac{1}{2}\nabla(h^+)^2$ for any $h\in W^{1,2}$, we have

$$\begin{aligned}\Phi'(t) &\le \int -|\nabla((G_t-F_t)^+)|^2 + |(G_t-F_t)^+|^2\big(a_0(G_t+F_t) + a_1 - \tfrac{1}{2}\Delta v_t\big)\,\mathrm{d}\mathfrak{m}_{\mathrm{Y}}\\ &\le 2\Phi(t)\big(|a_0|\|G_t+F_t\|_{L^\infty} + |a_1| + \tfrac{1}{2}\|\Delta v_t\|_{L^\infty}\big).\end{aligned}$$

Since the assumption $F_0\ge G_0$ gives $\Phi(0)=0$, by Gronwall's lemma we conclude that $\Phi(t)=0$ for any $t\ge 0$, which is the thesis. □

Lemma 2.1.2. *Let $(h_t)\in L^\infty_{loc}([0,\infty), W^{1,2}(\mathrm{Y}))\cap AC_{loc}([0,\infty), W^{-1,2}(\mathrm{Y}))$.*

Then $t\mapsto \frac{1}{2}\int |(h_t)^+|^2\,\mathrm{d}\mathfrak{m}_{\mathrm{Y}}$ is locally absolutely continuous on $[0,\infty)$ and it holds

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{1}{2}\int |(h_t)^+|^2\,\mathrm{d}\mathfrak{m}_{\mathrm{Y}} = \int (h_t)^+\frac{\mathrm{d}}{\mathrm{d}t}h_t\,\mathrm{d}\mathfrak{m}_{\mathrm{Y}}, \qquad \text{a.e. } t, \tag{2.1.5}$$

where the integral in the right hand side is intended as the coupling of $(h_t)^+\in W^{1,2}(\mathrm{Y})$ with $\frac{\mathrm{d}}{\mathrm{d}t}h_t\in W^{-1,2}(\mathrm{Y})$.

proof If $(h_t) \in AC_{loc}([0,\infty), L^2(\mathrm{Y}))$, the claim follows easily with the same computations done in (2.1.4). The general case follows by approximation via the heat flow. Fix $\varepsilon > 0$ and notice that the fact that h_ε is a contraction in $W^{1,2}$ and a bounded operator from L^2 to $W^{1,2}$ yield the inequalities

$$\|\mathsf{h}_\varepsilon f\|_{L^2} = \sup_{\|g\|_{L^2}\le 1} \int \mathsf{h}_\varepsilon f\, g\, \mathrm{d}\mathfrak{m}_\mathrm{Y} \le \sup_{\|g\|_{L^2}\le 1} \|\mathsf{h}_\varepsilon g\|_{W^{1,2}} \|f\|_{W^{-1,2}} \le C_\varepsilon \|f\|_{W^{-1,2}}$$
$$\|\mathsf{h}_\varepsilon f\|_{W^{-1,2}} = \sup_{\|g\|_{W^{1,2}}\le 1} \int \mathsf{h}_\varepsilon f\, g\, \mathrm{d}\mathfrak{m}_\mathrm{Y} \le \sup_{\|g\|_{W^{1,2}}\le 1} \|\mathsf{h}_\varepsilon g\|_{W^{1,2}} \|f\|_{W^{-1,2}} \le \|f\|_{W^{-1,2}},$$

for all $f \in L^2$, which together with the density of L^2 in $W^{-1,2}$ ensures that h_ε can be uniquely extended to a linear bounded operator from $W^{-1,2}$ to L^2 which is also a contraction when seen with values in $W^{-1,2}$. It is then clear that $\mathsf{h}_\varepsilon f \to f$ in $W^{-1,2}$ as $\varepsilon \downarrow 0$ for any $f \in W^{-1,2}$. It follows that for (h_t) as in the assumption, $(\mathsf{h}_\varepsilon h_t) \in AC_{loc}([0,\infty), L^2(\mathrm{Y}))$, so that by what previously said the thesis holds for such curve and writing the identity (2.1.5) in integral form we have

$$\frac{1}{2}\int |(\mathsf{h}_\varepsilon h_{t_1})^+|^2 - |(\mathsf{h}_\varepsilon h_{t_0})^+|^2\, \mathrm{d}\mathfrak{m}_\mathrm{Y} = \int_{t_0}^{t_1} \int \big(\mathsf{h}_\varepsilon h_t\big)^+ \mathsf{h}_\varepsilon\Big(\frac{\mathrm{d}}{\mathrm{d}t} h_t\Big)\, \mathrm{d}\mathfrak{m}_\mathrm{Y}\, \mathrm{d}t \qquad \forall 0 \le t_0 \le t_1.$$

Letting $\varepsilon \downarrow 0$, using the continuity at $\varepsilon = 0$ of h_ε seen as operator on all the spaces $W^{1,2}, L^2, W^{-1,2}$ and the continuity of $h \mapsto h^+$ as map from $W^{1,2}$ with the strong topology to $W^{1,2}$ with the weak one (which follows from the continuity of the same operator in L^2 together with the fact that it decreases the $W^{1,2}$ norm), we obtain

$$\frac{1}{2}\int |h_{t_1}^+|^2 - |h_{t_0}^+|^2\, \mathrm{d}\mathfrak{m}_\mathrm{Y} = \int_{t_0}^{t_1} \int (h_t)^+ \frac{\mathrm{d}}{\mathrm{d}t} h_t\, \mathrm{d}\mathfrak{m}_\mathrm{Y}\, \mathrm{d}t \qquad \forall 0 \le t_0 \le t_1. \tag{2.1.6}$$

Now the bound

$$\Big| \int_{t_0}^{t_1} \int (h_t)^+ \frac{\mathrm{d}}{\mathrm{d}t} h_t\, \mathrm{d}\mathfrak{m}_\mathrm{Y}\, \mathrm{d}t \Big| \le \|(h_t)\|_{L^\infty([t_0,t_1],W^{1,2})} \int_{t_0}^{t_1} \Big\| \frac{\mathrm{d}}{\mathrm{d}t} h_t \Big\|_{W^{-1,2}} \mathrm{d}t$$

grants the local absolute continuity of $t \mapsto \frac{1}{2}\int |h_t^+|^2\, \mathrm{d}\mathfrak{m}_\mathrm{Y}$ and the conclusion follows by differentiating (2.1.6). □

2.1.2 Hamilton's gradient estimates and related inequalities

We start proving Hamilton's gradient estimate on compact $\mathsf{RCD}(K,\infty)$ spaces, with a proof which closely follows the original one in [64]. As already said, in fact the same result is known to be true - from [72] - on the more general class of $\mathsf{RCD}(K,\infty)$ spaces, but given that the compactness assumption slightly simplifies the argument, we provide the proof. A key tool for the forthcoming computations is the following:

$$\varphi \circ f \in \mathrm{Test}^\infty(\mathrm{X}) \quad \forall f \in \mathrm{Test}^\infty(\mathrm{X}),\ \varphi : \mathbb{R} \to \mathbb{R} \text{ which is } C^\infty \text{ on the image of } f \tag{2.1.7}$$

(see [109]). We shall extensively make use of it without further notice in the case $\varphi(z) := \log(z)$ and f has bounded image.

Proposition 2.1.3. *Let* $(\mathrm{X}, \mathsf{d}, \mathfrak{m})$ *be a compact* $\mathsf{RCD}(K, \infty)$ *space with* $K \in \mathbb{R}$ *and let* $u_0 \in L^\infty(\mathfrak{m})$ *be such that* $u_0 \geq c$ *for some positive constant* c*. Put* $u_t := \mathsf{h}_t u_0$ *for all* $t > 0$*. Then*

$$t|\nabla \log u_t|^2 \leq (1 + 2K^- t) \log\left(\frac{\|u_0\|_{L^\infty(\mathfrak{m})}}{u_t}\right), \quad \mathfrak{m}\text{-a.e.} \tag{2.1.8}$$

for all $t > 0$*, where* $K^- := \max\{0, -K\}$*.*

proof Let us assume for the moment that $u_0 \in \mathrm{Test}^\infty_{>0}(\mathrm{X})$. Set $M := \|u_0\|_{L^\infty(\mathfrak{m})}$ and define for $t \geq 0$

$$v_t := \varphi_t \frac{|\nabla u_t|^2}{u_t} - u_t \log \frac{M}{u_t}, \qquad \text{with} \quad \varphi_t := \frac{t}{1 + 2K^- t}.$$

Notice that by the maximum principle (1.2.10) we know that $c \leq u_t \leq M$ for all $t \geq 0$, thus the definition of v_t is well posed.

Our thesis is equivalent to the fact that $v_t \leq 0$ and we shall prove this via the comparison principle for the heat flow stated in point (i) of Proposition 2.1.1. The fact that $(u_t) \in AC_{loc}([0, \infty), W^{1,2}(\mathrm{X}))$ and - by the maximum principle (1.2.10) and the Bakry-Émery inequality (1.2.13) - that $(\log(u_t)), (|\nabla u_t|) \in L^\infty_{loc}([0, \infty), L^\infty(\mathrm{X}))$ grant that $(v_t) \in AC_{loc}([0, \infty), L^2(\mathrm{X}))$. Since by construction we have $v_0 \leq 0$, we are left to prove that for any $h \in D(\Delta)$ non-negative it holds

$$\int h \frac{\mathrm{d}}{\mathrm{d}t} v_t \, \mathrm{d}\mathfrak{m} \leq \int v_t \Delta h \mathrm{d}\mathfrak{m} \qquad \text{a.e. } t.$$

We have $u_t \in D(\Delta)$ and, by (1.2.17), that $|\nabla u_t|^2 \in D(\mathbf{\Delta})$ for any $t \geq 0$, thus since as said $0 < c \leq u_t \leq M$ for all $t \geq 0$, we deduce that $v_t \in D(\mathbf{\Delta})$ for any $t \geq 0$. Hence our thesis can be rewritten as

$$\Big(\frac{\mathrm{d}}{\mathrm{d}t} v_t\Big)\mathfrak{m} \leq \mathbf{\Delta} v_t \qquad \text{a.e. } t.$$

The conclusion now follows by direct computation. We have

$$\frac{\mathrm{d}}{\mathrm{d}t} v_t = \varphi_t' \frac{|\nabla u_t|^2}{u_t} + \varphi_t \Big(\frac{2}{u_t} \langle \nabla u_t, \nabla \Delta u_t \rangle - \Delta u_t \frac{|\nabla u_t|^2}{u_t^2}\Big) - \Delta u_t \log \frac{M}{u_t} + \Delta u_t \tag{2.1.9}$$

and

$$\Delta\Big(u_t \log \frac{M}{u_t}\Big) = (\Delta u_t) \log \frac{M}{u_t} - \Delta u_t - \frac{|\nabla u_t|^2}{u_t}. \tag{2.1.10}$$

Moreover

$$\mathbf{\Delta} \frac{|\nabla u_t|^2}{u_t} = \frac{1}{u_t} \mathbf{\Delta} |\nabla u_t|^2 + \Big(|\nabla u_t|^2 \Delta(u_t^{-1}) + 2 \left\langle \nabla |\nabla u_t|^2, \nabla(u_t^{-1}) \right\rangle\Big)\mathfrak{m}$$

so that using the Bochner inequality (1.2.23) we obtain

$$\begin{aligned} \mathbf{\Delta} \frac{|\nabla u_t|^2}{u_t} \geq \Big(&\frac{2}{u_t} |\mathrm{Hess}(u_t)|^2_{\mathsf{HS}} + \frac{2}{u_t} \langle \nabla u_t, \nabla \Delta u_t \rangle + \frac{2K}{u_t} |\nabla u_t|^2 \\ &- \Delta u_t \frac{|\nabla u_t|^2}{u_t^2} + 2 \frac{|\nabla u_t|^4}{u_t^3} - \frac{2}{u_t^2} \left\langle \nabla u_t, \nabla |\nabla u_t|^2 \right\rangle\Big)\mathfrak{m}. \end{aligned} \tag{2.1.11}$$

Putting together (2.1.9), (2.1.10) and (2.1.11) and using the identity

$$\Big|\mathrm{Hess}(u_t) - \frac{\nabla u_t \otimes \nabla u_t}{u_t}\Big|^2_{\mathsf{HS}} = |\mathrm{Hess}(u_t)|^2_{\mathsf{HS}} + \frac{|\nabla u_t|^4}{u_t^2} - \frac{\big\langle \nabla u_t, \nabla |\nabla u_t|^2\big\rangle}{u_t}$$

we obtain

$$\Big(\frac{\mathrm{d}}{\mathrm{d}t} v_t\Big)\mathfrak{m} - \boldsymbol{\Delta} v_t \le \Big(\frac{|\nabla u_t|^2}{u_t}\big(\varphi_t' - 2K\varphi - 1\big) - \frac{2}{u_t}\Big|\mathrm{Hess}(u_t) - \frac{\nabla u_t \otimes \nabla u_t}{u_t}\Big|^2_{\mathsf{HS}}\Big)\mathfrak{m}$$

and the conclusion follows noticing that by the definition of φ_t we have

$$\varphi_t' - 2K\varphi_t - 1 \le 0 \qquad \forall t \ge 0.$$

For the general case, recall that by (1.2.15) and our assumption on u_0 we have that $u_\varepsilon \in \mathrm{Test}^\infty_{>0}(\mathrm{X})$ for every $\varepsilon > 0$ and notice that what we have just proved grants that

$$t|\nabla \log u_{t+\varepsilon}|^2 \le (1 + 2K^- t)\log\Big(\frac{\|u_\varepsilon\|_{L^\infty}}{u_{t+\varepsilon}}\Big), \qquad \mathfrak{m}\text{-a.e.}, \quad \forall t \ge 0.$$

By the maximum principle (1.2.10) we have that $\|u_\varepsilon\|_{L^\infty} \le \|u_0\|_{L^\infty}$, then the conclusion easily follows letting $\varepsilon \downarrow 0$ and using the continuity of $\varepsilon \mapsto u_\varepsilon, |\nabla u_\varepsilon| \in L^2(\mathfrak{m})$. □

In the compact finite-dimensional case, thanks to the Gaussian estimates for the heat kernel we can now easily obtain a bound independent on the L^∞ norm of the initial datum present in inequality (2.1.8):

Theorem 2.1.4. *Let $(\mathrm{X}, \mathsf{d}, \mathfrak{m})$ be a compact $\mathsf{RCD}^*(K, N)$ space with $K \in \mathbb{R}$ and $N \in [1, \infty)$. Then there is a constant C depending on K, N and $D := \mathrm{diam}(\mathrm{X})$ only such that for any $u_0 \in L^1(\mathrm{X})$ non-negative and not identically 0 the inequality*

$$|\nabla \log(u_t)|^2 \le C\Big(1 + \frac{1}{t^2}\Big), \qquad \mathfrak{m}\text{-a.e.} \tag{2.1.12}$$

holds for all $t > 0$, where $u_t := \mathsf{h}_t u$. In particular, for every $\delta > 0$ there is a constant $C_\delta > 0$ depending on K, N, D, δ only such that

$$\sup_{\varepsilon \in (0,1)} \varepsilon \|\nabla \log(u_{\varepsilon t})\|_{L^\infty} \le C_\delta \qquad \forall t \ge \delta. \tag{2.1.13}$$

proof Recall the representation formula (1.2.9c):

$$u_t(x) = \int u(y) r_t[y](x)\,\mathrm{d}\mathfrak{m}(y) \qquad \forall x \in \mathrm{X}$$

and that for the transition probability densities $\mathsf{r}_t[y](x)$ we have the Gaussian estimates (1.2.11)

$$\frac{C_0}{\mathfrak{m}(B_{\sqrt{t}}(y))} e^{-C_1 \frac{D^2}{t}} \le \mathsf{r}_t[y](x) \le \frac{C_2}{\mathfrak{m}(B_{\sqrt{t}}(y))} \qquad \forall x, y \in \mathrm{X},$$

for appropriate constants C_0, C_1, C_2 depending only on K, N. Therefore we have

$$\|u_t\|_{L^\infty} = \sup_x u_t(x) \le C_2 \int \frac{u(y)}{\mathfrak{m}(B_{\sqrt{t}}(y))}\,\mathrm{d}\mathfrak{m}(y),$$

$$\inf_x u_{2t}(x) \ge C_0 e^{-C_1 \frac{D^2}{t}} \int \frac{u(y)}{\mathfrak{m}(B_{\sqrt{2t}}(y))}\,\mathrm{d}\mathfrak{m}(y) > 0.$$

By the Bishop-Gromov inequality we know that for some constant $C_3 > 0$ it holds

$$\mathfrak{m}(B_{\sqrt{2t}}(y)) \leq C_3 \mathfrak{m}(B_{\sqrt{t}}(y)) \qquad \forall y \in \mathrm{X},\ t > 0,$$

hence the above yields

$$\frac{\|u_t\|_{L^\infty}}{u_{2t}(x)} \leq \frac{C_2 C_3}{C_0} e^{C_1 \frac{D^2}{t}} \qquad \forall x \in \mathrm{X},\ t > 0.$$

We now apply Proposition 2.1.3 with u_t in place of u_0 (notice that the assumptions are fulfilled) to get

$$t|\nabla \log(u_{2t})|^2 \leq (1 + 2K^- t) \log\Big(\frac{\|u_t\|_{L^\infty}}{u_{2t}}\Big) \leq (1 + 2K^- t)\Big(\log\big(\frac{C_2 C_3}{C_0}\big) + C_1 \frac{D^2}{t}\Big) \qquad \mathfrak{m}\text{-a.e.},$$

which is (equivalent to) the bound (2.1.12). The last statement is now obvious. □

In inequality (2.1.12), the right hand side blows-up at $t = 0$ and thus it gives no control for small t's. In the next simple proposition we show that if the initial datum is good enough, then we have a control for all t's:

Proposition 2.1.5. *Let* $(\mathrm{X}, \mathsf{d}, \mathfrak{m})$ *be a compact* $\mathsf{RCD}(K, \infty)$ *space with* $K \in \mathbb{R}$ *and let* $u_0 : \mathrm{X} \to (0, \infty)$ *be such that* $\log u_0$ *is Lipschitz. Put* $u_t := \mathsf{h}_t u_0$ *for all* $t > 0$. *Then*

$$|\nabla \log u_t| \leq e^{-Kt} \||\nabla \log u_0|\|_{L^\infty} \qquad \mathfrak{m}\text{-a.e.}.$$

proof Assume for a moment that $u_0 \in \mathrm{Test}^\infty_{>0}(\mathrm{X})$ and put $\varphi_t := \log u_t \in \mathrm{Test}^\infty(\mathrm{X})$ so that, also recalling the calculus rules stated in the preliminary section, we have $(\varphi_t) \in AC_{loc}([0, \infty), L^2(\mathrm{X}))$ and

$$\frac{\mathrm{d}}{\mathrm{d}t}\varphi_t = |\nabla \varphi_t|^2 + \Delta \varphi_t. \tag{2.1.14}$$

By the maximum principle (1.2.10) we know that $u_t(x) \in [c, C]$ for any t, x, for some $[c, C] \subset (0, \infty)$ and from this fact and the chain rule for the differential and Laplacian it easily follows that $(\Delta \varphi_t) \in L^\infty_{loc}([0, \infty), W^{1,2}(\mathrm{X}))$ and $(|\nabla \varphi_t|) \in L^\infty_{loc}([0, \infty), L^\infty(\mathrm{X}))$. Hence taking (1.2.16) into account we see that $|\nabla \varphi_t|^2 \in L^\infty_{loc}([0, \infty), W^{1,2}(\mathrm{X}))$ as well. Therefore from (2.1.14) we deduce that $(\varphi_t) \in AC_{loc}([0, \infty), W^{1,2}(\mathrm{X}))$ so that putting

$$F_t := |\nabla \varphi_t|^2,$$

we have that (F_t) satisfies the regularity assumptions needed in point (ii) of Proposition 2.1.1 (notice that trivially $AC_{loc}([0, \infty), L^2(\mathrm{X})) \subset AC_{loc}([0, \infty), W^{-1,2}(\mathrm{X}))$). Moreover, from (2.1.14) we get

$$\frac{\mathrm{d}}{\mathrm{d}t} F_t = 2\,\langle \nabla \varphi_t, \nabla F_t \rangle + 2\,\langle \nabla \varphi_t, \nabla \Delta \varphi_t \rangle$$

and therefore from the Bochner inequality (1.2.23) written for φ_t - neglecting the term with the Hessian - we see that for any $h \in \mathrm{Test}^\infty(\mathrm{X})$ it holds

$$\frac{\mathrm{d}}{\mathrm{d}t} \int h F_t \,\mathrm{d}\mathfrak{m} \leq \int \Delta h F_t + 2h\Big(\langle \nabla \varphi_t, \nabla F_t \rangle - K F_t\Big)\,\mathrm{d}\mathfrak{m},$$

showing that (F_t) is a weak subsolution of (2.1.2) with

$$a_0 = 0 \qquad a_1 = -2K \qquad a_2 = 0 \qquad v_t = 2\varphi_t.$$

On the other hand, the function

$$G_t(x) := e^{-2Kt}\|F_0\|_{L^\infty}$$

is a solution of (2.1.2) and $F_0 \le G_0$ $\mathfrak{m}$-a.e. Since from the chain rule for the Laplacian and the maximum principle (1.2.10) we have $\Delta\varphi_t \in L^\infty_{loc}([0,\infty), L^\infty(\mathrm{X}))$, we see that we are in position to apply point (ii) of Proposition 2.1.1 and deduce that $F_t \le G_t$ $\mathfrak{m}$-a.e. for every $t > 0$, which is the thesis.

For the case of general u_0 as in the assumptions, we put $u_0^\varepsilon := e^{\mathsf{h}_\varepsilon(\log(u_0))}$ and notice that by the Bakry-Émery estimate (1.2.13), it holds

$$\overline{\lim_{\varepsilon\downarrow 0}}\, \||\nabla \log u_0^\varepsilon|\|_{L^\infty} \le \||\nabla \log u_0|\|_{L^\infty}.$$

Then put $\varphi_t^\varepsilon := \log \mathsf{h}_t u_0^\varepsilon$ and notice that this last inequality together with what previously proved grants that

$$\overline{\lim_{\varepsilon\downarrow 0}}\, \||\nabla \varphi_t^\varepsilon|\|_{L^\infty} \le e^{-Kt}\||\nabla \log u_0|\|_{L^\infty}.$$

Conclude noticing that $\varphi_t^\varepsilon \to \log u_t$ $\mathfrak{m}$-a.e. as $\varepsilon \downarrow 0$ and use the closure of the differential. □

2.1.3 A Li-Yau type inequality

We now prove a version of Li-Yau inequality valid on general compact $\mathsf{RCD}^*(K,N)$ spaces, where K is possibly negative: the bound (2.1.15) that we obtain is not sharp (as it is seen by letting $K \uparrow 0$ in the estimate (2.1.16) provided in the proof) but we decide to present it because the idea of the proof is simple and the compactness assumption

Theorem 2.1.6. *Let* $(\mathrm{X}, \mathsf{d}, \mathfrak{m})$ *be a compact* $\mathsf{RCD}^*(K,N)$ *space with* $K \in \mathbb{R}$ *and* $N \in [1,\infty)$. *Then for every* $\delta > 0$ *there exists a constant* $C_\delta > 0$ *depending on* $K, N,$ $\mathrm{Diam}(\mathrm{X})$ *and* δ *only such that the following holds.*

For any $u_0 \in L^1(\mathrm{X})$ *non-negative and non-zero and* $\varepsilon \in (0,1)$ *it holds*

$$\varepsilon\Delta \log(\mathsf{h}_{\varepsilon t}(u_0)) \ge -C_\delta \qquad \forall t \ge \delta. \tag{2.1.15}$$

proof We can, and will, assume $K < 0$. Let C be the constant given by Theorem 2.3.2 (which only depends on K, N and $\mathrm{Diam}(\mathrm{X})$) and put

$$\alpha(t) := -KC\Big(1 + \frac{4}{t^2}\Big) > 0.$$

We shall prove that for u_0 as in the assumptions we have

$$\Delta \log u_t \ge -\sqrt{N\alpha(t)}\coth\Big(\sqrt{\frac{\alpha(t)}{N}}t\Big) \qquad \forall t > 0. \tag{2.1.16}$$

From this the thesis easily follows as the function $\phi(t,\varepsilon) := \varepsilon\sqrt{N\alpha(\varepsilon t)}\coth\Big(\sqrt{\frac{\alpha(\varepsilon t)}{N}}\varepsilon t\Big)$ is decreasing in t - as seen by direct computation - so that (2.1.15) follows from (2.1.16) and

$$\lim_{\varepsilon\downarrow 0}\phi(\delta,\varepsilon) = \sqrt{\frac{-4KCN}{\delta^2}}\coth\Big(\sqrt{\frac{-4KC}{N}}\Big) < +\infty.$$

Thus fix u_0 as in the statement and notice that $u_t \in \mathrm{Test}^\infty_{>0}(\mathrm{X})$ for every $t > 0$, so that $f_t := \log u_t \in \mathrm{Test}^\infty(\mathrm{X})$ for every $t > 0$. Arguing as in the proof of Proposition 2.1.5 we see that $(f_t) \in AC_{loc}((0,\infty), W^{1,2}(\mathrm{X}))$ with

$$\frac{\mathrm{d}}{\mathrm{d}t}f_t = \Delta f_t + |\nabla f_t|^2, \quad \text{for a.e. } t > 0. \tag{2.1.17}$$

Let $\eta > 0$ to be fixed later and put $F_t := \Delta f_{t+\eta}$. From the chain rules for the gradient and Laplacian it is readily verified that $(F_t) \in L^\infty_{loc}([0,\infty), L^\infty(\mathrm{X})) \cap L^\infty_{loc}([0,\infty), W^{1,2}(\mathrm{X}))$

Now, as in the proof of Lemma 2.1.2, the trivial estimate

$$\|\Delta f\|_{W^{-1,2}} = \sup_{\|g\|_{W^{1,2}}=1}\int g\Delta f\,\mathrm{d}\mathfrak{m} = \sup_{\|g\|_{W^{1,2}}=1} -\int \langle\nabla g,\nabla f\rangle\,\mathrm{d}\mathfrak{m} \le \|f\|_{W^{1,2}}$$

grants that $\Delta : D(\Delta) \to L^2(\mathrm{X})$ can be uniquely extended to a linear bounded functional, still denoted by Δ, from $W^{1,2}(\mathrm{X})$ to $W^{-1,2}(\mathrm{X})$. It is then clear that $(F_t) \in AC_{loc}([0,\infty), W^{-1,2}(\mathrm{X}))$.

We want to show that (F_t) is a weak supersolution of (2.1.2) for an appropriate choice of the parameters and to this aim we fix $h \in \mathrm{Test}^\infty_+(\mathrm{X})$ and notice that

$$\frac{\mathrm{d}}{\mathrm{d}t}\int hF_t\,\mathrm{d}\mathfrak{m} = \frac{\mathrm{d}}{\mathrm{d}t}\int \Delta h f_{t+\eta}\,\mathrm{d}\mathfrak{m} \overset{(2.1.17)}{=} \int \Delta h(F_t + |\nabla f_{t+\eta}|^2)\,\mathrm{d}\mathfrak{m}.$$

Using first the Bochner inequality (1.2.24) and then the gradient estimate (2.3.2) we obtain

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\int hF_t\,\mathrm{d}\mathfrak{m} &\ge \int \Delta hF_t + h\Big(2\langle\nabla f_{t+\eta},\nabla F_t\rangle + \frac{2}{N}F_t^2 + 2K|\nabla f_{t+\eta}|^2\Big)\,\mathrm{d}\mathfrak{m}\\ &\ge \int \Delta hF_t + h\Big(2\langle\nabla f_{t+\eta},\nabla F_t\rangle + \frac{2}{N}F_t^2 + 2KC\big(1+\frac{1}{\eta^2}\big)\Big)\,\mathrm{d}\mathfrak{m},\end{aligned}$$

thus indeed (F_t) is a weak supersolution of (2.1.2) for

$$a_0 := \frac{2}{N} \qquad a_1 := 0 \qquad a_2(\eta) := 2KC\big(1+\frac{1}{\eta^2}\big) \qquad v_t := 2f_{t+\eta}.$$

Noticing that $\alpha_2(\eta) < 0$, it is trivial to check that the function

$$y_t := -\sqrt{-\frac{a_2(\eta)N}{2}}\coth\Big(\sqrt{-\frac{2a_2(\eta)}{N}}(t+t_0)\Big)$$

is the only solution of

$$y_t' = \frac{2}{N}y_t^2 + a_2(\eta)$$

with $y_0 = -\sqrt{-\frac{a_2(\eta)N}{2}}\coth\Big(\sqrt{-\frac{2a_2(\eta)}{N}}t_0\Big)$. Now recall that $F_0 = \Delta f_\eta \in L^\infty$, so that choosing $t_0 > 0$ sufficiently small we have that $F_0 \ge y_0$ $\mathfrak{m}$-a.e.

Defining $G_t(x) := y_t$ it is then clear that (G_t) is a weak (sub)solution of (2.1.2), and since $F_0 \geq G_0$ holds $\mathfrak{m}$-a.e. and, as already argued in the proof of Proposition 2.1.5, $\Delta v_t \in L^\infty_{loc}([0,\infty), L^\infty(\mathrm{X}))$, Proposition 2.1.1 grants that for any $t > 0$ it holds $F_t \geq G_t$ $\mathfrak{m}$-a.e., that is:

$$\Delta \log(u_{t+\eta}) \geq -\sqrt{-\frac{\alpha_2(\eta)N}{2}} \coth\Big(\sqrt{-\frac{2\alpha_2(\eta)}{N}}(t+t_0)\Big) \geq -\sqrt{-\frac{\alpha_2(\eta)N}{2}} \coth\Big(\sqrt{-\frac{2\alpha_2(\eta)}{N}}t\Big).$$

Picking $\eta := t$ we obtain (an equivalent version of) (2.1.16). □

2.2 Local calculus

As noticed and widely used in the previous section, when $(\mathrm{X}, \mathsf{d}, \mathfrak{m})$ is a compact $\mathsf{RCD}^*(K,N)$ and $u \in L^1(\mathrm{X})$ is a non-zero non-negative function, by the regularization properties of the heat semigroup it follows that $\mathsf{h}_t u \in \mathrm{Test}(\mathrm{X})$ and $\mathsf{h}_t u \geq c$ for some constant $c > 0$, so that $\log \mathsf{h}_t u \in \mathrm{Test}(\mathrm{X})$ as well.

In the non-compact case, the lower bound on $\mathsf{h}_t u$ is no longer true and thus in general $\log \mathsf{h}_t u \notin L^2(\mathrm{X})$. As a consequence, in order to give a meaning to Hamilton's gradient estimates and Li-Yau inequality one needs to introduce local Sobolev spaces and related differential operators; this will allow us also to investigate those local regularity features that fail to be true on a global scale. A possible approach to the problem consists in generalizing the machinery of $L^2(\mathrm{X})$-normed modules introduced in Section 1.1 through the definition of $L^0(\mathrm{X})$-normed modules (for a detailed presentation, the reader is addressed to [56]).

A $L^0(\mathrm{X})$-normed module is a complete topological space $(\mathscr{M}, \tau)$ endowed with a bilinear map

$$\begin{array}{rcl} L^0(\mathrm{X}) \times \mathscr{M} & \to & \mathscr{M} \\ (f, v) & \mapsto & f \cdot v \end{array}$$

called multiplication by $L^0(\mathrm{X})$ functions, and a map $|\cdot| : \mathscr{M} \to L^0(\mathrm{X})$ with non-negative values, called pointwise norm, such that:

(i) for every $v \in \mathscr{M}$ and $f, g \in L^0(\mathrm{X})$

$$f \cdot (g \cdot v) = (fg) \cdot v, \qquad 1 \cdot v = v$$

where 1 denotes the function identically equal to 1;

(ii) for every $v \in \mathscr{M}$ and $f \in L^0(\mathrm{X})$ it holds $|f \cdot v| = |f|\,|v|$ $\mathfrak{m}$-a.e.

(iii) for some Borel partition (E_i) of X into sets of finite $\mathfrak{m}$-measure, $\mathscr{M}$ is complete w.r.t. the distance

$$\mathsf{d}_0(v,w) := \sum_{i\in\mathbb{N}} \frac{1}{2^i \mathfrak{m}(E_i)} \int_{E_i} \min\{1, |v-w|\} \mathrm{d}\mathfrak{m}$$

and τ is the topology induced by the distance.

An isomorphism between two $L^0(\mathrm{X})$-normed modules is a linear homeomorphism which preserves the pointwise norm and the product with $L^0(\mathrm{X})$ functions.

It is not difficult to see that the choice of the partition (E_i) in (iii) affects the distance, but not the completeness of $\mathscr{M}$ nor the topology τ. Moreover, if $\mathscr{M}$ is a $L^2(\mathrm{X})$-normed module, then there exist a unique $L^0(\mathrm{X})$-normed module $\mathscr{M}^0$ and a unique map $\iota : \mathscr{M} \to \mathscr{M}^0$ which is linear, preserves the pointwise norm and has dense image. Here uniqueness is intended up to unique isomorphism, in the following sense: if $\tilde{\mathscr{M}}^0$ and $\tilde{\iota}$ have the same properties, then there exists a unique isomorphism $\Phi : \mathscr{M} \to \tilde{\mathscr{M}}^0$ such that $\tilde{\iota} = \Phi \circ \iota$.

For our discussion it is also helpful to recall that on $\mathsf{RCD}^*(K, N)$ spaces very regular cut-off functions can be built. Although already known in the literature (see [9]), we give the full proof of the following result, because the estimates (2.2.1) will be widely exploited in Chapter 5 and Chapter 6 and have not been explicitly pointed out in [9].

Lemma 2.2.1. *Let* $(\mathrm{X}, \mathsf{d}, \mathfrak{m})$ *be a* $\mathsf{RCD}(K, \infty)$ *space with* $K \in \mathbb{R}$ *endowed with a Borel non-negative measure* $\mathfrak{m}$ *which is finite on bounded sets.*

Then for all $B \subset B' \subset \mathrm{X}$ *with* B *compact and* B' *open and relatively compact there exists a function* $\chi : \mathrm{X} \to \mathbb{R}$ *satisfying:*

(i) $0 \le \chi \le 1$, $\chi \equiv 1$ *on a neighbourhood of* B *and* $\mathrm{supp}(\chi) \subset B'$*;*

(ii) $\chi \in \mathrm{Test}^\infty(\mathrm{X})$.

Moreover, the following estimates hold

$$\||\nabla \chi|\|_{L^\infty(\mathfrak{m})} \le C \qquad \|\Delta \chi\|_{L^\infty(\mathfrak{m})} \le C' \tag{2.2.1}$$

where the constants $C, C' > 0$ *only depend on* K *and the distance between* B *and* $\mathrm{X} \setminus B'$.

proof Let $\eta \in \mathrm{Lip}(\mathbb{R})$ with bounded support and introduce the following notation:

$$\mathsf{d}(x, B) := \inf_{y \in B} \mathsf{d}(x, y) \qquad B^r := \{x \in \mathrm{X} \,:\, \mathsf{d}(x, B) \le r\}.$$

If we define $\xi := \eta \circ \mathsf{d}(\cdot, B)$, then it is easy to see that ξ has bounded support and $\xi \in W^{1,2}(\mathrm{X})$ with $|\nabla \xi| \le |\eta'(\mathsf{d}(\cdot, B))|$ $\mathfrak{m}$-a.e. Thus, since $\varepsilon := \inf_{x \notin B'} \mathsf{d}(x, B) > 0$, we can choose η in such a way that $0 \le \xi \le 1$, $\xi \equiv 1$ on $B^{\varepsilon/3}$ and $\mathrm{supp}(\xi) \subset B^{2\varepsilon/3}$, i.e. (i) holds.

In order to gain further regularity and build a function satisfying also (ii), let us first introduce the mollified heat flow $\mathfrak{h}_t$ as

$$\mathfrak{h}_t f := \frac{1}{t} \int_0^\infty \mathsf{h}_r f \kappa(r/t) \mathrm{d}r, \tag{2.2.2}$$

where $\kappa \in C_c^\infty(0, \infty)$ is a probability density, and observe that if $f \in L^2 \cap L^\infty(\mathfrak{m})$ then $\mathfrak{h}_t f \in \mathrm{Test}^\infty(\mathrm{X})$: indeed, by (1.2.15) and the a priori estimates (1.2.12a) and (1.2.12b) we see that $\mathfrak{h}_t f \in \mathrm{Test}(\mathrm{X})$ for every $t > 0$ and its Laplacian is explicitly given by

$$\begin{aligned} \Delta \mathfrak{h}_t f &= \frac{1}{t} \int_0^\infty \Delta \mathsf{h}_r f \kappa(r/t) \mathrm{d}r = \frac{1}{t} \int_0^\infty \Big(\frac{\mathrm{d}}{\mathrm{d}t} \mathsf{h}_r f \Big) \kappa(r/t) \mathrm{d}r \\ &= -\frac{1}{t^2} \int_0^\infty \mathsf{h}_r f \kappa'(r/t) \mathrm{d}r = -\frac{1}{t} \int_0^\infty \mathsf{h}_{st} f \kappa'(s) \mathrm{d}s \end{aligned}$$

whence, by the maximum principle,

$$\|\Delta \mathfrak{h}_t f\|_{L^\infty(\mathfrak{m})} \leq \frac{1}{t}\|f\|_{L^\infty(\mathfrak{m})}\|\kappa'\|_{L^1(0,\infty)}. \tag{2.2.3}$$

An explicit bound can be found for $|\nabla \mathfrak{h}_t f|$ too, provided $|\nabla f| \in L^\infty(\mathfrak{m})$: by the Bakry-Émery contraction estimate (1.2.13) we have

$$|\nabla \mathfrak{h}_t f| \leq \frac{1}{t}\int_0^\infty |\nabla \mathsf{h}_r f|\kappa(r/t)\mathrm{d}r = \int_0^\infty |\nabla \mathsf{h}_{st} f|\kappa(s)\mathrm{d}s \leq \int_0^\infty e^{-Kst}\sqrt{\mathsf{h}_{st}(|\nabla f|^2)}\kappa(s)\mathrm{d}s$$

whence, again by the maximum principle,

$$\||\nabla \mathfrak{h}_t f|\|_{L^\infty(\mathfrak{m})} \leq \||\nabla f|\|_{L^\infty(\mathfrak{m})}\int_0^\infty e^{-Kst}\kappa(s)\mathrm{d}s \leq C\||\nabla f|\|_{L^\infty(\mathfrak{m})}. \tag{2.2.4}$$

Now let us regularize ξ by means of the mollified heat flow, setting $\xi_t := \mathfrak{h}_t\xi$. The maximum principle, (2.2.4) and the Sobolev-to-Lipschitz property imply that $\{\xi_t\}_{t\in[0,T]}$ is a uniformly Lipschitz and uniformly bounded family for any $T > 0$. By the Ascoli-Arzelà theorem together with the fact that $\xi_t \to \xi$ in $L^2(\mathfrak{m})$ as $t \downarrow 0$ we deduce that $\xi_t \to \xi$ uniformly. As a consequence, again by the maximum principle there exists $\delta = \delta(\varepsilon) > 0$ sufficiently small so that

$$\frac{3}{4} \leq \xi_\delta \leq 1 \text{ on } B^{\varepsilon/3} \qquad \text{and} \qquad 0 \leq \xi_\delta \leq \frac{1}{4} \text{ on } \mathrm{X}\setminus B^{2\varepsilon/3}.$$

Then pick $\zeta \in C^2([0,1],[0,1])$ such that $\zeta([0,1/4]) = \{0\}$ and $\zeta([3/4,1]) = \{1\}$ and set $\chi := \zeta \circ \xi_\delta$. By the chain rule for the gradient and the Laplacian, we finally deduce that χ satisfies both (i) and (ii).

The estimates (2.2.1) follow from (2.2.3) and (2.2.4), noticing that the choice of κ and ζ is completely arbitrary, whereas η and δ only depend on ε. □

With this said, $L^2_{loc}(\mathrm{X})$ is defined as the space of functions $f \in L^0(\mathrm{X})$ such that for all compact set $\Omega \subset \mathrm{X}$ there exists a function $g \in L^2(\mathrm{X})$ such that $f = g$ $\mathfrak{m}$-a.e. in Ω. Analogously, it is possible to introduce $L^p_{loc}(\mathrm{X})$ for all $p \in [1,\infty]$. The local Sobolev class $S^2_{loc}(\mathrm{X})$ is then defined as

$$S^2_{loc}(\mathrm{X}) := \{f \in L^0(\mathrm{X}) : \forall \Omega \subset\subset \mathrm{X}\ \exists g \in S^2(\mathrm{X}) \text{ s.t. } f = g\ \mathfrak{m}\text{-a.e. in } \Omega\} \tag{2.2.5}$$

and the local minimal weak upper gradient of a function $f \in S^2_{loc}(\mathrm{X})$ is denoted by $|Df|$, omitting the locality feature, and defined for all $\Omega \subset\subset \mathrm{X}$ as

$$|Df| := |Dg| \qquad \mathfrak{m}\text{-a.e. in } \Omega$$

where g is as in (2.2.5). The definition does depend neither on Ω nor on the choice of g associated to it by locality of the minimal weak upper gradient; the definition of the local Sobolev space $W^{1,2}(\mathrm{X})$ follows naturally as $L^2_{loc} \cap S^2_{loc}(\mathrm{X})$.

As already seen in Section 1.1 for the global analogue, the notion of local minimal weak upper gradient enables the definition of the local differential through the following result.

Theorem 2.2.2 (Definition of $L^0(T^*\mathrm{X})$)**.** *There exists a unique couple* $(L^0(T^*\mathrm{X}), \mathrm{d})$*, where* $L^0(T^*\mathrm{X})$ *is a* $L^0(\mathrm{X})$*-normed module and* $\mathrm{d} : S^2_{loc}(\mathrm{X}) \to L^0(T^*\mathrm{X})$ *is a linear map, such that:*

(i) $|\mathrm{d}f| = |Df|$ $\mathfrak{m}$*-a.e. for every* $f \in S^2_{loc}(\mathrm{X})$*;*

(ii) the space $\{\mathrm{d}f \,:\, f \in S^2_{loc}(\mathrm{X})\}$ *generates* $L^0(T^*\mathrm{X})$ *in the sense of modules, namely* $L^0(\mathrm{X})$*-linear combinations of objects of the form* $\mathrm{d}f$ *are dense in* $L^0(T^*\mathrm{X})$*.*

Uniqueness is meant up to unique isomorphism, namely if $(\mathscr{M}, \mathrm{d}')$ *satisfies the same properties, then there is a unique isomorphism* $\Phi : L^0(T^*\mathrm{X}) \to \mathscr{M}$ *such that* $\Phi(\mathrm{d}f) = \mathrm{d}'f$ *for all* $f \in S^2_{loc}(\mathrm{X})$*.*

It is worth saying that $(L^0(T^*\mathrm{X}), \mathrm{d})$ can be fully identified with the $L^0(\mathrm{X})$-completion of $L^2(T^*\mathrm{X})$ in the sense that, if we denote by d' the differential associated to $L^2(T^*\mathrm{X})$ to avoid ambiguity, there is a unique linear map $\iota : L^2(T^*\mathrm{X}) \to L^0(T^*\mathrm{X})$ sending $\mathrm{d}'f$ to $\mathrm{d}f$, preserving the pointwise norm and with dense image. For this reason, we shall use the same notation for both the differentials of $L^2(T^*\mathrm{X})$ and $L^0(T^*\mathrm{X})$.

The space of vector fields $L^0(T\mathrm{X})$ can be now defined in two equivalent ways: as the dual of $L^0(T^*\mathrm{X})$ as $L^0(\mathrm{X})$-normed module or as the $L^0(\mathrm{X})$-completion of $L^2(T\mathrm{X})$. Then $L^2_{loc}(T^*\mathrm{X}) \subset L^0(T^*\mathrm{X})$ (resp. $L^2_{loc}(T\mathrm{X}) \subset L^0(T\mathrm{X})$) is defined as the collection of the 1-forms ω such that $|\omega| \in L^2_{loc}(\mathrm{X})$ (resp. the vector fields W such that $|W| \in L^2_{loc}(\mathrm{X})$); with this definition, we can observe that in $(L^0(T^*\mathrm{X}), \mathrm{d})$ the differential d actually takes values in $L^2_{loc}(T^*\mathrm{X})$. Furhermore, since $L^0(T^*\mathrm{X})$ and $L^0(T\mathrm{X})$ are canonically isomorphic via the (musical) isomorphisms

$$\flat : L^0(T\mathrm{X}) \to L^0(T^*\mathrm{X}) \qquad\qquad \sharp : L^0(T^*\mathrm{X}) \to L^0(T\mathrm{X}),$$

extensions of the ones introduced in (1.1.6), we can define the gradient of $f \in W^{1,2}_{loc}(\mathrm{X})$ as $\nabla f := (\mathrm{d}f)^\sharp \in L^2_{loc}(T\mathrm{X})$.

Remark 2.2.3. By a simple cut-off argument, one can show that $f \in S^2_{loc}(\mathrm{X})$ if and only if $\chi f \in S^2(\mathrm{X})$ for every Lipschitz function χ with bounded support. An analogous statement holds for $W^{1,2}_{loc}(\mathrm{X})$. ■

As regards the divergence, we say that $W \in L^2_{loc}(T\mathrm{X})$ has divergence in $L^2_{loc}(\mathrm{X})$, and write $W \in D(\mathrm{div}_{loc})$, provided there exists $h \in L^2_{loc}(\mathrm{X})$ such that

$$\int \phi h \mathrm{d}\mathfrak{m} = -\int \mathrm{d}\phi(W)\,\mathrm{d}\mathfrak{m}, \qquad \forall \phi \text{ Lipschitz with bounded support.}$$

In this case ϕ is unique and denoted by $\mathrm{div}(W)$. Arguing in the same manner, we say that a function $f \in W^{1,2}_{loc}(\mathrm{X})$ has Laplacian in $L^2_{loc}(\mathrm{X})$, and write $f \in D(\Delta_{loc})$, if there exists $g \in L^2_{loc}(\mathrm{X})$ such that

$$\int \phi g \mathrm{d}\mathfrak{m} = -\int \langle \nabla\phi, \nabla f\rangle \mathrm{d}\mathfrak{m}, \qquad \forall \phi \text{ Lipschitz with bounded support}$$

and in this case, since g is unique, we set $\Delta f := g$. In addition, a function $f \in W^{1,2}_{loc}(\mathrm{X})$ belongs to the domain of the (local) measure-valued Laplacian, and we write $f \in D(\mathbf{\Delta}_{loc})$, if there exists a set function $\mu : \mathscr{B}_{bd}(\mathrm{X}) \to \overline{\mathbb{R}}$, $\mathscr{B}_{bd}(\mathrm{X})$ being the family of bounded Borel subsets of X, which is a Borel (signed) measure with finite total variation when restricted to an element of $\mathscr{B}_{bd}(\mathrm{X})$ and such that (1.1.5) holds. Finally, let us introduce the local Sobolev space $W^{2,2}_{loc}(\mathrm{X})$ and the vector space of 'local test functions' $\mathrm{Test}_{loc}(\mathrm{X})$ as

$$W^{2,2}_{loc}(\mathrm{X}) := \{f \in L^2_{loc}(\mathrm{X}) \,:\, \forall \Omega \subset\subset \mathrm{X}\ \exists g \in W^{2,2}(\mathrm{X}) \text{ s.t. } f = g \ \mathfrak{m}\text{-a.e. in } \Omega\}$$
$$\mathrm{Test}_{loc}(\mathrm{X}) := \{f \in L^2_{loc}(\mathrm{X}) \,:\, \forall \Omega \subset\subset \mathrm{X}\ \exists g \in \mathrm{Test}(\mathrm{X}) \text{ s.t. } f = g \ \mathfrak{m}\text{-a.e. in } \Omega\}$$

and, for a function $f \in W^{2,2}_{loc}(\mathrm{X})$, denote the (local) Hessian of f as the global one and define it by locality, namely $\mathrm{Hess}(f) := \mathrm{Hess}(g)$ $\mathfrak{m}$-a.e. in Ω provided Ω has non-empty interior (because of the locality property (1.2.19)) where g is as in the definition above.

It is easy to see that the local definitions we have just provided are consistent with the global ones, in the sense that

- if $f \in W^{1,2}(\mathrm{X})$, then its local minimal weak upper gradient coincides $\mathfrak{m}$-a.e. with the global one;
- if $f \in W^{1,2}_{loc}(\mathrm{X})$ with $f, |Df| \in L^2(\mathrm{X})$, then $f \in W^{1,2}(\mathrm{X})$ and the local minimal weak upper radient turns out to be the global one;

thus motivating the same notation. Analogous statements hold for the divergence, the L^2-valued Laplacian, the measure-valued one and the Hessian. It is also worth saying that these local objects enjoy the same calculus rules of the corresponding global ones, up to slight adaptations; we collect them below.

Lemma 2.2.4 (Local calculus rules)**.** *Let* $(\mathrm{X}, \mathsf{d}, \mathfrak{m})$ *be a* $\mathsf{RCD}^*(K, N)$ *with* $K \in \mathbb{R}$ *and* $N \in [1, \infty)$*. Then:*

(i) (Differential and gradient) d *satisfies the following calculus rules:*

$$\begin{aligned} |\mathrm{d}f| &= |Df| \quad \mathfrak{m}\text{-a.e.} && \forall f \in S^2_{loc}(\mathrm{X}) \\ \mathrm{d}f &= \mathrm{d}g \quad \mathfrak{m}\text{-a.e. on } \{f = g\} && \forall f, g \in S^2_{loc}(\mathrm{X}) \\ \mathrm{d}(fg) &= g\,\mathrm{d}f + f\,\mathrm{d}g && \forall f, g \in L^\infty_{loc} \cap S^2_{loc}(\mathrm{X}) \end{aligned}$$

and for all $f \in S^2_{loc}(\mathrm{X})$ *and* $\varphi : \mathbb{R} \to \mathbb{R}$ *such that for all* $K \subset\subset \mathrm{X}$ *there exists* $I_K \subset\subset \mathbb{R}$ *in such a way that* $\mathscr{L}^1(f(K) \setminus I_K) = 0$ *and* $f_{|_{I_K}}$ *is Lipschitz it holds*

$$\mathrm{d}(\varphi \circ f) = \varphi' \circ f,$$

where it is part of the properties the fact that $\varphi \circ f, fg \in S^2_{loc}(\mathrm{X})$ *for* φ, f, g *as above; analogous statements hold for the gradient.*

(ii) (Divergence) for all $f \in W^{1,2}_{loc}(\mathrm{X}), W \in D(\mathrm{div}_{loc})$ *such that* $|f|, |W| \in L^\infty_{loc}(\mathrm{X})$ *it holds*

$$\mathrm{div}(fW) = \mathrm{d}f(W) + f\mathrm{div}(W),$$

where it is part of the statement that $fW \in D(\mathrm{div}_{loc})$ *for* f, W *as above.*

(iii) (Laplacian) Δ *enjoys the chain and Leibniz rules:*

$$\Delta(\varphi \circ f) = \varphi'' \circ f |\mathrm{d}f|^2 + \varphi' \circ f \Delta f \tag{2.2.6a}$$

$$\Delta(fg) = g\Delta f + f\Delta g + 2\,\langle \nabla f, \nabla g \rangle \tag{2.2.6b}$$

where in the first equality we assume that $f \in D(\Delta_{loc}), \varphi : \mathbb{R} \to \mathbb{R}$ *are such that* $f, |\mathrm{d}f| \in L^\infty_{loc}(\mathrm{X})$ *and* $\varphi', \varphi'' \in L^\infty(\mathbb{R})$ *and in the second that* $f, g \in D(\Delta_{loc}) \cap L^\infty_{loc}(\mathrm{X})$ *and* $|\mathrm{d}f|, |\mathrm{d}g| \in L^\infty_{loc}(\mathrm{X})$ *and it is part of the claims that* $\varphi \circ f, fg$ *are in* $D(\Delta_{loc})$*.*

(iv) (Hessian) for all $f, g \in \mathrm{Test}_{loc}(\mathrm{X})$ *it holds*

$$\mathrm{Hess}(f) = \mathrm{Hess}(g), \qquad \mathfrak{m}\text{-a.e. } on\ \{f = g\}$$

and the following Leibniz rule is satisfied

$$\mathrm{d}\,\langle \nabla f, \nabla g\rangle = \mathrm{Hess}(f)(\nabla g, \cdot) + \mathrm{Hess}(g)(\nabla f, \cdot) \qquad \mathfrak{m}\text{-a.e.} \tag{2.2.7}$$

proof All the locality properties are straightforward consequences of the definitions and of the analogous properties valid for the global objects. The Leibniz rules are also easy to obtain; for (2.2.7) it is sufficient to observe that $\mathrm{Test}(\mathrm{X}) \subset W^{2,2}(\mathrm{X})$ entails $\mathrm{Test}_{loc}(\mathrm{X}) \subset W^{2,2}_{loc}(\mathrm{X})$, so that $\mathrm{Hess}(f)$ is locally well defined for any $f \in \mathrm{Test}_{loc}(\mathrm{X})$ and the locality of the Hessian together with (1.2.21) give the conclusion.

As regards the chain rule, let us prove the one for the differential; the other one follows along the same lines. Let $f \in S^2_{loc}(\mathrm{X})$, $K \subset \mathrm{X}$ be a compact set, f_K a Sobolev function coinciding with f $\mathfrak{m}$-a.e. on K and φ_K a Lipschitz extension of $\varphi|_{I_K}$ to $\mathbb{R}$. Then $\varphi_K \circ f_K \in S^2(\mathrm{X})$ by Lemma 1.1.3, whence $\mathrm{d}(\varphi_K \circ f_K) = \varphi'_K \circ f_K\, \mathrm{d}f_K$, and $\varphi_K \circ f_K = \varphi \circ f$ $\mathfrak{m}$-a.e. in K; by locality of the differential and the fact that $\mathrm{d}f = 0$ $\mathfrak{m}$-a.e. on $f^{-1}(N)$ for any $\mathscr{L}^1$-negligible set $N \subset \mathbb{R}$ (thus getting rid of the points where φ and φ_K may be different) the thesis follows. □

Let us stress that $\mathrm{Test}(\mathrm{X}) \subset W^{2,2}(\mathrm{X})$ entails $\mathrm{Test}_{loc}(\mathrm{X}) \subset W^{2,2}_{loc}(\mathrm{X})$. Hence, taking also (1.2.16) and (1.2.17) into account, for any $f \in \mathrm{Test}_{loc}(\mathrm{X})$ we already know that $\mathrm{Hess}(f)$ is locally well defined and $|\nabla f|^2 \in D(\mathbf{\Delta}_{loc}) \cap W^{1,2}_{loc}(\mathrm{X})$. However, there is a flipside: given $f \in L^2_{loc}(\mathrm{X})$, how can we check if $f \in \mathrm{Test}_{loc}(\mathrm{X})$? Far from investigating the question in its full generality, as already anticipated we are mostly interested in the case $f = \log \mathsf{h}_t u$ and to answer the question in this particular case we will rely on Lemma 2.2.1, which implies that

$$\mathrm{Test}_{loc}(\mathrm{X}) = \Big\{ f \in D(\Delta_{loc}) \cap L^\infty_{loc}(\mathrm{X}) \,:\, |\nabla f| \in L^\infty_{loc}(\mathrm{X}),\, \Delta f \in W^{1,2}_{loc}(\mathrm{X}) \Big\}. \tag{2.2.8}$$

The '$\subset$' inclusion is obvious, while for the opposite one if f belongs to the set on the right-hand side of (2.2.8) and $\Omega \subset \mathrm{X}$ is a compact set, it is sufficient to take a cut-off function $\chi \in \mathrm{Test}^\infty(\mathrm{X})$ with bounded support and $\chi \equiv 1$ on Ω: then $\chi f \in \mathrm{Test}(\mathrm{X})$ and $\chi f \equiv f$ on Ω. This follows by Remark 2.2.3, the fact that χf still belongs to the set on the right-hand side of (2.2.8) and the calculus rules of Lemma 2.2.4.

The existence of a regular extension is not a trivial task. Thus, the advantage of (2.2.8) is clear: it characterizes $\mathrm{Test}_{loc}(\mathrm{X})$ in terms of $L^\infty_{loc}(\mathrm{X}), W^{1,2}_{loc}(\mathrm{X}), D(\Delta_{loc})$ and checking whether a function belongs to these spaces requires a distributional or cut-off approach, which is easier.

The fact that $\log \mathsf{h}_t u \in \mathrm{Test}_{loc}(\mathrm{X})$ is now a matter of direct computation.

Proposition 2.2.5. *Let* $(\mathrm{X}, \mathsf{d}, \mathfrak{m})$ *be a* $\mathsf{RCD}^*(K, N)$ *space with* $K \in \mathbb{R}$ *and* $N \in [1, \infty)$ *endowed with a Borel non-negative measure* $\mathfrak{m}$ *and let* $u_0 \in L^2 \cap L^\infty(\mathrm{X})$ *be non-negative not identically zero. Put* $u_t := \mathsf{h}_t u_0$ *for all* $t > 0$. *Then* $\log u_t \in \mathrm{Test}_{loc}(\mathrm{X})$.

proof Fix $t > 0$ and notice that by (1.2.15) $u_t \in \mathrm{Test}(\mathrm{X})$ and by (1.2.11) u_t is locally away from 0. Taking into account the fact that log is smooth on $(0, \infty)$, this information implies that $\log u_t$ is well defined, belongs to $L^2_{loc} \cap L^\infty_{loc}(\mathrm{X})$ and by the chain rule for the differential and the Laplacian stated in Lemma 2.2.4 $\log u_t \in W^{1,2}_{loc}(\mathrm{X}) \cap D(\Delta_{loc})$ with

$$\nabla \log u_t = \frac{\nabla u_t}{u_t}, \qquad \Delta \log u_t = \frac{\Delta u_t}{u_t} - \frac{|\nabla u_t|^2}{u_t^2}. \tag{2.2.9}$$

The first expression, $u_t \in \mathrm{Test}(\mathrm{X})$ and the fact that u_t is locally bounded away from 0 also entail that $|\nabla \log u_t| \in L^\infty_{loc}(\mathrm{X})$. Looking at the second expression, taking into account the local chain rule (which grant that $1/u_t, 1/u_t^2 \in W^{1,2}_{loc}(\mathrm{X})$) and the Leibniz one, we deduce that $\Delta \log u_t \in W^{1,2}_{loc}(\mathrm{X})$ and it is easy to check that

$$|\nabla \Delta \log u_t| \leq \frac{|\nabla \Delta u_t|}{u_t} + \frac{\Delta u_t |\nabla u_t|}{u_t^2} + \frac{2|\nabla u_t|}{u_t^2} |\mathrm{Hess}(u_t)|_{\mathsf{HS}} + \frac{2|\nabla u_t|^3}{u_t^3}. \tag{2.2.10}$$

By (2.2.8) this is sufficient to conclude. □

2.3 Hamilton and Li-Yau estimates

We start recalling Hamilton's gradient estimate on $\mathsf{RCD}(K,\infty)$ spaces, which is known to be true from [72]. In this work Jiang and Zhang proved the inequality on the class of proper $\mathsf{RCD}(K,\infty)$ spaces, but we would like to point out that it is not actually needed for $(\mathrm{X},\mathsf{d},\mathfrak{m})$ to be proper. In fact, in Section 2 of [72] the authors never use explicitly the fact that X is proper, but refer to [58], [4], [7] and [109]: in the last three papers $\mathfrak{m}$ is just a non-negative Radon measure, while only in the first one $\mathfrak{m}$ is assumed to be finite on bounded sets (see Proposition 4.24) and this is always the case on $\mathsf{RCD}(K,\infty)$ spaces because of (1.2.6). Hence we are going to state Hamilton's gradient estimate in a framework which is slightly better than the one of [72].

Proposition 2.3.1. *Let $(\mathrm{X},\mathsf{d},\mathfrak{m})$ be a $\mathsf{RCD}(K,\infty)$ space with $K \in \mathbb{R}$ endowed with a Borel non-negative measure $\mathfrak{m}$ and let $u_0 \in L^p \cap L^\infty(\mathfrak{m})$ be positive with $p \in [1,\infty)$. Put $u_t := \mathsf{h}_t u_0$ for all $t > 0$. Then*

$$t|\nabla \log u_t|^2 \leq (1 + 2K^- t) \log\left(\frac{\|u_0\|_{L^\infty(\mathfrak{m})}}{u_t}\right), \quad \mathfrak{m}\text{-a.e.} \tag{2.3.1}$$

for all $t > 0$, where $K^- := \max\{0, -K\}$.

In the finite-dimensional case, thanks to the Gaussian estimates for the heat kernel we can easily obtain a bound independent of the L^∞ norm of the initial datum (provided it has bounded support) present in inequality (2.3.1):

Theorem 2.3.2. *Let $(\mathrm{X},\mathsf{d},\mathfrak{m})$ be a $\mathsf{RCD}^*(K,N)$ space with $K \leq 0$ and $N \in [1,\infty)$ endowed with a Borel non-negative measure $\mathfrak{m}$ which is finite on bounded sets. Then there is a constant C depending on K, N only such that for any $u_0 \in L^1(\mathfrak{m})$ non-negative, not identically 0 and with bounded support the inequality*

$$|\nabla \log(u_t)|^2 \leq C\left(1 + \frac{1}{t}\right)\left(1 + t + \frac{D_0^2(x)}{t}\right), \quad \mathfrak{m}\text{-a.e.} \tag{2.3.2}$$

holds for all $t > 0$, where $u_t := \mathsf{h}_t u$ and

$$D_0(x) := \sup_{y \in \mathrm{supp}(u_0)} \mathsf{d}(x,y).$$

In particular, for every $0 < \delta \leq T < \infty$ and $\bar{x} \in \mathrm{X}$ there is a constant $C_{\delta,T} > 0$ depending on $K, N, \delta, T, \bar{x}$ and the diameter of $\mathrm{supp}(u_0)$ such that for every $\varepsilon \in (0,1)$ it holds

$$\varepsilon|\nabla \log(u_{\varepsilon t})| \leq C_{\delta,T}\big(1 + \mathsf{d}(\cdot,\bar{x})\big) \qquad \forall t \in [\delta, T]. \tag{2.3.3}$$

proof As a first remark, by the compactness of $\mathrm{supp}(u_0)$ we see that D_0 is always finite and thus the inequality (2.3.2) is meaningful. Recall the representation formula (1.2.9c)

$$u_t(x) = \int u_0(y) r_t[y](x)\,\mathrm{d}\mathfrak{m}(y) = \int_{\mathrm{supp}(u_0)} u_0(y) r_t[y](x)\,\mathrm{d}\mathfrak{m}(y) \qquad \forall x \in \mathrm{X}$$

and that for the transition probability densities $r_t[y](x)$ we have the Gaussian estimates (1.2.11), which can be simplified as

$$\frac{C_0}{\mathfrak{m}(B_{\sqrt{t}}(y))} \exp\Big(-\frac{\mathsf{d}^2(x,y)}{3t} - C_2 t\Big) \le \mathsf{r}_t[x](y) \le \frac{C_1}{\mathfrak{m}(B_{\sqrt{t}}(y))} e^{C_2 t} \qquad \forall x, y \in \mathrm{X},$$

for appropriate constants C_0, C_1, C_2 depending only on K, N. Therefore, we have

$$\|u_t\|_{L^\infty} = \sup_x u_t(x) \le C_1 e^{C_2 t} \int_{\mathrm{supp}(u_0)} \frac{u(y)}{\mathfrak{m}(B_{\sqrt{t}}(y))}\,\mathrm{d}\mathfrak{m}(y),$$

$$\inf_x u_{2t}(x) \ge C_0 e^{-2C_2 t} e^{-\frac{D_0^2(x)}{t}} \int_{\mathrm{supp}(u_0)} \frac{u(y)}{\mathfrak{m}(B_{\sqrt{2t}}(y))}\,\mathrm{d}\mathfrak{m}(y) > 0.$$

By the uniformly local doubling condition (1.2.4) we know that it holds

$$\mathfrak{m}(B_{\sqrt{2t}}(y)) \le \mathfrak{m}(B_{\sqrt{t}}(y)) C_3 e^{C_4\sqrt{t}} \qquad \forall y \in \mathrm{X},\ t > 0,$$

where C_3, C_4 only depend on K, N. As a consequence, the above yields

$$\frac{\|u_t\|_{L^\infty}}{u_{2t}(x)} \le C_5 e^{3C_2 t + C_4\sqrt{t} + \frac{D_0^2(x)}{t}} \qquad \forall x \in \mathrm{X},\ t > 0.$$

We now apply Proposition 2.3.1 with u_t in place of u_0 (notice that the assumptions are fulfilled) to get

$$\begin{aligned} t|\nabla \log(u_{2t})|^2 &\le (1 + 2K^- t) \log\Big(\frac{\|u_t\|_{L^\infty}}{u_{2t}}\Big) \\ &\le (1 + 2K^- t)\Big(\log C_5 + 3C_2 t + C_4\sqrt{t} + \frac{D_0^2(x)}{t}\Big) \qquad \mathfrak{m}\text{-a.e.}, \end{aligned}$$

which is (equivalent to) the bound (2.3.2), because D_0 is uniformly bounded away from 0; more precisely, since $\mathrm{supp}(u_0)$ is compact, we can say that $D_0 \ge \mathrm{diam}(\mathrm{supp}(u_0))/2 > 0$. Indeed, if this were not the case, then there would exist a sufficiently small constant $\eta > 0$ such that $\mathsf{d}(x,y) \le \mathrm{diam}(\mathrm{supp}(u_0))/2 - \eta$ for all $y \in \mathrm{supp}(u_0)$, whence $\mathsf{d}(y,y') \le \mathrm{diam}(\mathrm{supp}(u_0)) - 2\eta$ for all $y, y' \in \mathrm{supp}(u_0)$ by trinagle inequality and thus $\mathrm{diam}(\mathrm{supp}(u_0)) \le \mathrm{diam}(\mathrm{supp}(u_0)) - 2\eta$. The last statement is now obvious, noticing that

$$D_0(x) \le D_0(\bar{x}) + \mathsf{d}(x, \bar{x}) \tag{2.3.4}$$

for any $\bar{x} \in \mathrm{supp}(u_0)$. □

This result is the extension of Theorem 2.1.4 to the non-compact case, but while (2.1.12) provides a uniform (in space) control on $|\nabla \log(u_t)|$ which improves as $t \to \infty$, (2.3.2) is only a local boundedness condition, both in space and in time. The same difference exists between (2.1.13) and (2.3.3). However, since we are only be interested in small-time behaviour, (2.3.3) will be enough for our purposes.

A further result that we shall need soon is the Baudoin-Garofalo inequality (see [51] for the case of finite mass and [70] for the general one).

Theorem 2.3.3 (Baudoin-Garofalo inequality)**.** *Let* $(\mathrm{X}, \mathsf{d}, \mathfrak{m})$ *be a* $\mathsf{RCD}^*(K, N)$ *space with* $K \in \mathbb{R}$ *and* $N \in [1, \infty)$ *endowed with a non-negative Radon measure* $\mathfrak{m}$ *and let* $u_0 \in L^p(\mathfrak{m})$ *for some* $p \in [1, \infty)$ *be non-negative. Put* $u_t := \mathsf{h}_t u_0$ *for all* $t > 0$*. Then*

$$|\nabla \log u_t|^2 \leq e^{-2Kt/3} \frac{\Delta u_t}{u_t} + \frac{NK}{3} \frac{e^{-4Kt/3}}{1 - e^{-2Kt/3}}, \qquad \mathfrak{m}\text{-a.e.} \tag{2.3.5}$$

for all $t > 0$*, where* $\frac{NK}{3} \frac{e^{-4Kt/3}}{1-e^{-2Kt/3}}$ *is understood as* $\frac{N}{2t}$ *when* $K = 0$*.*

It is worth stressing that in the case $K = 0$ (2.3.5) reduces to the well known Li-Yau inequality

$$\Delta \log u_t, \qquad \forall t > 0,\ \mathfrak{m}\text{-a.e.}$$

which thus reads as in the smooth setting. The theorem above allows us to prove a version of the Li-Yau inequality valid on general $\mathsf{RCD}^*(K, N)$ spaces, where K is possibly negative. The bound (2.3.6) that we obtain is valid only locally in time but sufficient for our needs since it makes explicit the dependence on the distance.

Theorem 2.3.4. *Let* $(\mathrm{X}, \mathsf{d}, \mathfrak{m})$ *be a* $\mathsf{RCD}^*(K, N)$ *space with* $K \leq 0$ *and* $N \in [1, \infty)$ *endowed with a non-negative Radon measure* $\mathfrak{m}$*. Then for every* $0 < \delta \leq T < \infty$ *and* $\bar{x} \in \mathrm{X}$ *there exists a constant* $C_{\delta,T} > 0$ *depending on* $K, N, \delta, T, \bar{x}$ *and the diameter of* $\mathrm{supp}(u_0)$ *such that the following holds.*

For any $u_0 \in L^1(\mathfrak{m})$ *non-negative, not identically zero and with bounded support and for any* $\varepsilon \in (0, 1)$ *it holds*

$$\varepsilon \Delta \log(\mathsf{h}_{\varepsilon t}(u_0)) \geq -C_{\delta,T}\big(1 + \mathsf{d}^2(\cdot, \bar{x})\big) \qquad \forall t \in [\delta, T]. \tag{2.3.6}$$

proof As a first step, rewrite the Baudoin-Garofalo inequality (2.3.5) as

$$e^{-2Kt/3}\left(\frac{\Delta u_t}{u_t} - |\nabla \log u_t|^2\right) \geq (1 - e^{-2Kt/3})|\nabla \log u_t|^2 - \frac{NK}{3} \frac{e^{-4Kt/3}}{1 - e^{-2Kt/3}}.$$

Then, dividing both sides by $e^{-2Kt/3}$, using Hamilton's gradient estimate (2.3.2) and the fact that $\log u_t \in \mathrm{Test}_{loc}(\mathrm{X})$ (Proposition 2.2.5), which allows us to use all the calculus rules we need and thus to say that

$$\Delta \log u_t = \frac{\Delta u_t}{u_t} - |\nabla \log u_t|^2,$$

the inequality above becomes

$$\Delta \log u_t \geq C(e^{2Kt/3} - 1)\left(1 + \frac{1}{t}\right)\left(1 + t + \frac{D_0^2}{t}\right) - \frac{NK}{3} \frac{e^{-2Kt/3}}{1 - e^{-2Kt/3}},$$

whence by trivial manipulations

$$\Delta \log u_t \geq -\frac{2K^- C}{3}(1 + t)\left(1 + t + \frac{D_0^2}{t}\right) - \frac{N}{2t} e^{2K^- t/3}.$$

From (2.3.4) the conclusion is now immediate. □

As remarked for Theorem 2.3.2, also (2.3.6) is a generalization of (2.1.15) to the non-compact case with the 'disadvantage' of being only a local estimate (both in time and space); such disadvantage is just a consequence of (2.3.2), because if we could use (2.1.12) instead of (2.3.2) in the proof above, then the outcome would be a true Li-Yau type inequality: a space-independent lower bound on $\Delta \log u_t$ which improves as $t \to \infty$.

Part II

The Schrödinger problem

Chapter 3

From Schrödinger equation to entropy minimization

In 1931 Erwin Schrödinger addressed a new interpolation problem which immediately showed shocking analogies with recently born wave mechanics and the Schrödinger equation, analogies that are much stronger than the ones encoded in the Fokker-Planck equation and are in line with the studies of Smoluchowski on the Brownian motion and on rare events in diffusive particles systems. However, in spite of the several mathematical challenges proposed by Schrödinger in [110], they did not become as famous as the equation named after him; on the contrary, the interpolation problem was almost forgotten, even rediscovered many years after, and only recently a widespread interest for the topic has spotted. This chapter aims to be a historical presentation of the problem with several physical insights.

For this reason we shed light on the decades passed between 1931 and the recent years. As regards the history of the problem, we start with the original formulation, the motivation and the physical interpretation lying behind (see [110], [111], the survey [81] and the monographs [100], [102], whence most of the considerations are taken from). We often quote Schrödinger's words, because of their suggestive and enlightening power, and moving from them and following a statistical physics approach we sketch how Föllmer's formulation as an entropy minimization problem can be deduced. As a concluding remark, various developments and relevant applications are recalled.

3.1 A bridge between quantum mechanics and diffusion processes

In the origin it was the wave equation. Indeed, it was 1926 when Schrödinger invented what he called wave equation and became later known as (linear) Schrödinger equation, namely the following linear PDE

$$i\frac{\partial\psi}{\partial t} + \frac{1}{2}\Delta\psi + i\mathbf{b}(t,x)\cdot\nabla\psi - V(t,x)\psi = 0$$

which describes the non-relativistic evolution of a single particle in an electric field with potential energy V, as it is nowadays commonly understood. However the physical meaning lying behind the solutions of such equation had been remaining mysterious for a long time and had been triggering many people, before satisfactory interpretations appeared, one of them

proposed by Max Born (see [21]). He proposed a statistical interpretation of wave functions, observing that

$$|\psi(t,x)|^2 = \psi(t,x)\overline{\psi}(t,x)$$

can be regarded as the probability density describing the position of something, e.g. an electron, where ψ is a solution to the Schrödinger equation and $\overline{\psi}$ its complex conjugate. This observation was coherent with experiments, explained them well and thus became a pillar of the Copenhagen interpretation of quantum mechanics. Yet, physicists like Einstein [44] and Planck were still quite suspicious towards such a perspective; famous is Einstein's remark about quantum mechanics in a letter to Born, dated 4 December 1926:

> Die Quantenmechanik ist sehr achtung-gebietend. Aber eine innere Stimme sagt mir, daß das doch nicht der wahre Jakob ist. Die Theorie liefert viel, aber dem Geheimnis des Alten bringt sie uns kaum näher. Jedenfalls bin ich überzeugt, daß der nicht würfelt.[1][2]

Schrödinger himself struggled to find a satisfactory physical meaning, but in a different way: from these efforts (although they did not give a positive answer to the problem) and his peculiar formulation of Brownian motion the Schrödinger problem arose.

In order to explain the results of the German physicist, first and foremost one has to observe that the Schrödinger equation is a diffusion equation, as witnessed by the Laplacian, but unlike the heat equation it is also a wave equation because of the imaginary unit appearing in front of the time derivative, whence the name chosen by Schrödinger. Thus no particle notion should be encoded in it, but with Born's statistical interpretation this came along and this contributed to the growth and strengthening of the wave-particle duality. However, together with the particle notion, a particle theory for quantum mechanics should be developed, whence the following question:

> Which particle should be attached to a solution of the Schrödinger equation?

Since in the quantum world position and velocity can not be equally well determined because of Heisenberg's uncertainty principle, no classical particle driven by Newton's law can be considered. On the contrary, the irregularity feature of Brownian motion trajectories (Hölder continuity but no absolute continuity) exactly leads to an interpretation of Brownian particles as objects with position but no velocity. For this reason the first attempt of a particle model for quantum mechanics was addressed towards the Brownian motion. As already mentioned above, there is a formal resemblance between the Schrödinger equation and a diffusion one, because the couple $(\psi, \overline{\psi})$ given by the wave function and its complex conjugate solve the following system of Schrödinger equations

$$i\frac{\partial\psi}{\partial t} + \frac{1}{2}\Delta\psi + i\mathbf{b}(t,x)\cdot\nabla\psi - V(t,x)\psi = 0 \tag{3.1.1a}$$

$$-i\frac{\partial\overline{\psi}}{\partial t} + \frac{1}{2}\Delta\overline{\psi} + i\mathbf{b}(t,x)\cdot\nabla\overline{\psi} - V(t,x)\overline{\psi} = 0 \tag{3.1.1b}$$

[1] Quantum mechanics is certainly imposing. But an inner voice tells me that it is not yet the real thing. The theory says a lot, but does not really bring us any closer to the secret of the 'old one'. I, at any rate, am convinced that He is not playing at dice.

[2] Born, M.; Born, M. E. H. & Einstein, A. (1971). The Born–Einstein Letters: Correspondence between Albert Einstein and Max and Hedwig Born from 1916 to 1955, with commentaries by Max Born. I. Born, trans. London, UK: Macmillan. ISBN 978-0-8027-0326-2

while a diffusive system can be described via the following couple of PDEs

$$\frac{\partial u}{\partial t} + \frac{1}{2}\Delta u + \mathbf{b}(t,x)\cdot\nabla u - c(t,x)u = 0 \tag{3.1.2a}$$

$$-\frac{\partial v}{\partial t} + \frac{1}{2}\Delta v - \mathbf{b}(t,x)\cdot\nabla v - c(t,x)v = 0 \tag{3.1.2b}$$

where c represents a creation and killing term, and the appearance of the two systems is very similar. However the resemblance is very superficial, because mathematical properties of solutions to one system can not be transferred to solutions to the other one; for instance, a group of unitary operators is naturally associated to (3.1.1a) and (3.1.1b), while diffusion equations induce semigroups of non-negative operators. Furthermore, from a physical point of view, wave equations describe essentially reversible phenomena, while phenomena described by diffusion equations have an irreversible nature.

Nevertheless, can Schrödinger equations be put within the framework of diffusion processes and the other way round? Is there an alternative way to get a particle theory for quantum mechanics via Brownian motions? Although it does not fully answer these questions, a problem arising in classical physics with much stronger analogies to quantum mechanics was proposed by Schrödinger in 1931 in [110] (a French translation of the original German paper appeared the year after, see [111]).

Let us present it by quoting Schrödinger himself [111], who in turn cites Eddington's point of view on Born's statistical interpretation:

> Il s'agit d'un problème classique: problème de probabilités dans la théorie du mouvement brownien. Mais en fin de compte, il *ressortira* une analogie avec la mécanique ondulatoire, qui fut si frappante pour moi lorsque je l'eus trouvée, qu'il m'est difficile de la croire purement accidentelle.
>
> À titre d'introduction, je voudrais vous citer une remarque que j'ai trouvée dans les "Gifford lectures" de A. S. Eddington (Cambridge, 1928, p. 216 et sqq). Eddington, en parlant de l'interprétation de la mécanique ondulatoire, fait dans une note au bas de la page la remarque suivante:[3]
>
> "The whole interpretation is very obscure, but it seems to depend on whether you are considering the probability *after you know what has happened* or the probability for the purposes of prediction. The $\psi\overline{\psi}$ is obtained by introducing two symmetrical systems of ψ waves travelling in opposite directions in time; one of these must presumably correspond to probable inference from what is known (or is stated) to have been the condition at a later time."

Eddington's remark is intriguing, because one could be tempted to read between the lines that information from the future is available. However this is impossible and a possible way to explain that relies on Schrödinger's work [110]. For sake of simplicity and in line with [110],

[3]This is a classical problem: a probability problem in the Brownian motion theory. But eventually, an analogy with wave mechanics will come out, which was so shocking for me when I discovered it that it is difficult for me to believe that it is purely accidental.

As an introduction, I would like to quote a remark that I found in the "Gifford lectures" of A. S. Eddington (Cambridge, 1928, p. 216 et sqq). Speaking about the interpretation of wave mechanics, Eddington points out the following remark in a footnote:

Einleitung. Wenn für ein diffundierendes oder in Brownscher Bewegung begriffenes Teilchen die Aufenthaltswahrscheinlichkeit im Abszissenbereich $(x;\ x+dx)$ zur Zeit t_0

$$w(x, t_0)\,dx$$

gegeben ist,

$$w(x, t_0) = w_0(x),$$

so ist sie für $t > t_0$ diejenige Lösung $w(x, t)$ der Diffusionsgleichung

$$D\,\frac{\partial^2 w}{\partial x^2} = \frac{\partial w}{\partial t}, \qquad (1)$$

welche für $t = t_0$ der vorgegebenen Funktion $w_0(x)$ gleich wird. — Über Probleme dieser Art, mit vielen möglichen Komplikationen und Variationen, die durch spezielle Versuchsanordnungen und Beobachtungsmethoden nahegelegt waren, gibt es eine umfangreiche Literatur, wobei das System, um das es sich handelt, gar nicht ein diffundierendes Teilchen zu sein braucht, sondern beispielsweise die Elektrometernadel bei der K. W. F. Kohlrauschischen Anordnung zur Messung der Schweidlerschen Schwankungen, und an Stelle der Gleichung (1) ihre Verallgemeinerung tritt, die sogenannte Fokker-Plancksche partielle Differentialgleichung für das betreffende, irgendwelchen Zufallseinflüssen ausgesetzte System[1].

Solche Systeme geben nun zu einer Klasse von Wahrscheinlichkeitsproblemen Anlaß, die bisher keine oder wenig Beachtung gefunden hat und die schon rein mathematisch dadurch von Interesse ist, daß die Antwort nicht durch eine Lösung der Fokkerschen Gleichung geliefert wird, sondern, wie sich zeigen wird, durch das Produkt der Lösungen zweier adjungierter Gleichungen, wobei nicht der einzelnen Lösung, sondern dem Produkt zeitliche Grenzbedingungen auferlegt sind. Physikalisch besteht eine enge Verwandtschaft mit dem interessanten Problemenkreis, den M. von Smoluchowski[2] aufgerollt hat in seinen schönen letzten Arbeiten über Erwartungszeiten und Wiederkehrzeiten sehr unwahrscheinlicher Zustände in Systemen diffundierender Par-

[1] A. Fokker, Ann. d. Phys. **43**, 812, 1914; M. Planck, diese Berichte 10. Mai 1917.

[2] M. von Smoluchowski, Bull. Akad. Cracovie A, S. 418, 1913; Göttinger Vorträge (bei Teubner 1914) S. 89ff.; Sitz.-Ber. d. Wien. Akad. d. Wiss. 2a, **123**, 2381, 1914; **124**, 263, 1915; **124**, 339, 1915; Physik. ZS. **16**, 321, 1915; Ann. d. Phys. **48**, 1103, 1915.

Sitzungsber. phys.-math. Kl. 1931. (2)

— 413 —

Figure 3.1: First page of [110]

assume that $\mathbf{b} \equiv 0$ and $c \equiv 0$ in (3.1.2a) (and in (3.1.2b) too), so that (3.1.2a) reduces to the heat equation: this corresponds to the evolution of a single particle, whose motion is driven only by molecular shocks with other particles. Now suppose that

> ... wir hätten das Teilchen zur Zeit t_0 bei x_0, zur Zeit t_1 bei x_1 angetroffen [...] Ein Hilfsbeobachter hat die Lage des Teilchens zur Zeit t beobachtet, jedoch ohne uns sein Ergebnis mitzuteilen. Die Frage lautet dann: welche Wahrscheinlichkeitsschlüsse können wir aus unseren zwei Beobachtungen auf die Zwischenbeobachtung unseres Helfers ziehen?[4]

[4]... at time t_0 we find the particle in x_0, at time t_1 in x_1 [...] An assistant observer has monitored the particle

In a slightly different way, we can imagine that we do not exactly know the particle position at times t_0 and t_1. Instead, we are given an initial probability distribution ρ_0 as well as a final one ρ_1. How can we deduce the distribution $\rho_t = \rho(\cdot, t)$ for the position at intermediate times? With just one boundary datum, the problem would trivialize: indeed, if only ρ_0 is assigned then ρ_t is the solution of the (forward) heat equation with $\rho(\cdot, 0) = \rho_0$ while in the case only ρ_1 is assigned then ρ_t solves its adjoint, namely

$$-\frac{\partial \rho}{\partial t} = \frac{1}{2}\Delta\rho$$

with $\rho(\cdot, 1) = \rho_1$. On the contrary, with both boundary data the problem is more difficult but also much more interesting. In addition, a part from determining ρ_t, a further question arises, which already suggests a large deviation approach to the problem, as we will see later on.

> Imaginez que vous observez un système de particules en diffusion, qui soient en équilibre thermodynamique. Admettons qu'à un instant donné t_0 vous les ayez trouvées en répartition à peu près uniforme et qu'à $t_1 > t_0$ vous ayez trouvé un écart spontané et considérable par rapport à cette uniformité. On vous demande de quelle manière cet écart s'est produit. Quelle en est la manière la plus probable?[5]

As regards the first question, namely the determination of ρ_t, for sake of simplicity let us consider a 1-dimensional situation, so that $x_0, x_1 \in \mathbb{R}$, and let $p(x, t)$ be the fundamental solution of the heat equation, namely

$$p(x, t) := \frac{1}{\sqrt{2\pi t}} \exp\Big(-\frac{x^2}{2t}\Big).$$

Then consider a large number N of independent particles that at t_0 are in x_0 and whose movement is described by a standard Brownian motion. We are interested in those particles that at t_1 are near x_1, say belong to $[x_1 - \varepsilon, x_1 + \varepsilon]$: their number amounts to

$$n_1 = N \int_{x_1-\varepsilon}^{x_1+\varepsilon} p(x - x_0, t_1 - t_0)\mathrm{d}x.$$

On the other hand, among the N particles, those belonging to $[x_1 - \varepsilon, x_1 + \varepsilon]$ at t_1 and to $[x_t - \varepsilon, x_t + \varepsilon]$ at the intermediate time $t_0 < t < t_1$ are given by

$$\begin{aligned} n &= N \int_{x_t-\varepsilon}^{x_t+\varepsilon} \int_{x_1-\varepsilon}^{x_1+\varepsilon} p(y - x_0, t - t_0) g(x - y, t_1 - t)\mathrm{d}x\mathrm{d}y \\ &= N \int_{x_t-\varepsilon}^{x_t+\varepsilon} p(y - x_0, t - t_0)\mathrm{d}y \int_{x_1-\varepsilon}^{x_1+\varepsilon} p(x - x_t, t_1 - t)\mathrm{d}x \end{aligned}$$

where the second identity comes from the independence of Brownian increments. The answer to the question is the ratio n/n_1 and thus

$$\rho_t(x) = \frac{p(x - x_0, t - t_0)p(x_1 - x, t_1 - t)}{p(x_1 - x_0, t_1 - t_0)} \tag{3.1.3}$$

position at time t, but without telling us his result. The question reads then as follows: which probabilistic inference on the intermediate observation of our assistant can we deduce from the two observations?

[5]Imagine that you observe a system of diffusing particles, which are in thermodynamical equilibrium. Let us admit that at a given time t_0 you find them with almost uniform repartition and that at $t_1 > t_0$ you observe a spontaneous and significant deviation from this uniformity. You are asked to explain how this deviation occurred. What is its most likely behaviour?

This is the solution to the first question in the case the position of the particle at times t_0 and t_1 is exactly known. Let us now discuss the general situation with probability densities ρ_0 and ρ_1. In this case we consider a large number N of independent particles: we let them free to move at t_0 and approximately

$$N \int_{x_0-\varepsilon}^{x_0+\varepsilon} \rho_0(x)\mathrm{d}x \tag{3.1.4}$$

start from $[x_0 - \varepsilon, x_0 + \varepsilon]$. Then we observe that at time t_1 more or less

$$N \int_{x_1-\varepsilon}^{x_1+\varepsilon} \rho_1(y)\mathrm{d}y \tag{3.1.5}$$

particles belong to $[x_1 - \varepsilon, x_1 + \varepsilon]$ and this is surprising, because the expected value (based on the assumption that particles move according to a standard Brownian motion) is

$$N \int_{x_1-\varepsilon}^{x_1+\varepsilon} \int_{-\infty}^{+\infty} \rho_0(x)p(y-x, t_1-t_0)\mathrm{d}x\mathrm{d}y.$$

If we knew how many particles started from $[x_0-\varepsilon, x_0+\varepsilon]$ actually arrived in $[x_1-\varepsilon, x_1+\varepsilon]$, then it would be sufficient to multiply this number by (3.1.3) to conclude. Thus, in order to estimate such a number let us partition $\mathbb{R}$ in unit intervals $\{I_k\}_{k\in\mathbb{N}}$, define

$$a_k := N \int_{I_k} \rho_0(x)\mathrm{d}x \qquad \text{and} \qquad b_k := N \int_{I_k} \rho_1(y)\mathrm{d}y,$$

denote by p_{jk} the probability for a particle to start from I_j and arrive in I_k (by symmetry it holds $p_{jk} = p_{kj}$) and finally introduce c_{jk} as the number of particles starting from I_j and arriving in I_k. From these considerations we deduce that

$$\sum_{k\in\mathbb{N}} c_{jk} = a_j, \qquad \sum_{j\in\mathbb{N}} c_{jk} = b_k, \qquad \sum_{j\in\mathbb{N}} a_j = \sum_{k\in\mathbb{N}} b_k = N. \tag{3.1.6}$$

Of course the numbers c_{jk} are not known, any choice satisfying the conditions above is acceptable, but as $N \to \infty$ it is reasonable to choose those c_{jk} such that the associated evolution is the most probable one. To this aim, if we consider a single possible realization of the system, then the probability that such realization actually happens is given by

$$\prod_{j\in\mathbb{N}} \prod_{k\in\mathbb{N}} p_{jk}^{c_{jk}}.$$

However, notice that there are many other realizations with the same probability: indeed the a_j particles coming from I_j can be permuted $a_j!/\Pi_{k\in\mathbb{N}} c_{jk}!$ times and this is true for all $j \in \mathbb{N}$, whence

$$\prod_{j\in\mathbb{N}} \frac{a_j!}{\prod_{k\in\mathbb{N}} c_{jk}!}$$

different possible realizations. By multiplying the number of all the possible rearrangements associated to a certain admissible family $\{c_{jk}\}$ with its probability we get

$$\prod_{j\in\mathbb{N}} a_j! \prod_{j\in\mathbb{N}} \prod_{k\in\mathbb{N}} \frac{p_{jk}^{c_{jk}}}{c_{jk}!}$$

and now we have to look for those solutions $\{c_{jk}\}$ of (3.1.6) maximizing this quantity. We omit the computation (if interested, the reader can find it in [111], Section 7) and directly state Schrödinger's conclusion: it holds

$$c_{jk} = p_{jk} f_j g_k \tag{3.1.7}$$

where f_j and g_k are Lagrange multipliers determined by

$$\begin{cases} a_j = f_j \sum_{k\in\mathbb{N}} p_{jk} g_k & \forall j \in \mathbb{N} \\ b_k = g_k \sum_{j\in\mathbb{N}} p_{jk} f_j & \forall k \in \mathbb{N} \end{cases}.$$

In order to come back to the continuum case, observe that a_j and b_k are given by (3.1.4) and (3.1.5) respectively, f_j and g_k become functions f, g so that

$$f_j = \sqrt{N} \int_{I_j} f(x)\mathrm{d}x \qquad\qquad g_k = \sqrt{N} \int_{I_k} g(y)\mathrm{d}y$$

and finally p_{jk} is replaced by $p(x_1 - x_0, t_1 - t_0)$. Putting all together we finally get the **Schrödinger system**

$$\begin{cases} \rho_0(x_0) = f(x_0) \int_{-\infty}^{+\infty} p(x_1 - x_0, t_1 - t_0) g(x_1)\mathrm{d}x_1 \\ \rho_1(x_1) = g(x_1) \int_{-\infty}^{+\infty} p(x_1 - x_0, t_1 - t_0) f(x_0)\mathrm{d}x_0 \end{cases}$$

using the terminology adopted in [49] and [20] or, in a more concise way,

$$\begin{cases} \rho_0 = f \mathsf{h}_{t_1 - t_0} g \\ \rho_1 = g \mathsf{h}_{t_1 - t_0} f \end{cases} \tag{3.1.8}$$

where h_t is the heat semigroup associated to the kernel $p(x,t)$. On the other hand, since the c_{jk}'s are replced by a function too, called c, (3.1.7) becomes

$$c(x_0, x_1) = N p(x_1 - x_0, t_1 - t_0) f(x_0) g(x_1)$$

and if integrated on $[x_0 - \varepsilon, x_0 + \varepsilon] \times [x_1 - \varepsilon, x_1 + \varepsilon]$, this quantity represents the number of particles moving from $[x_0 - \varepsilon, x_0 + \varepsilon]$ to $[x_1 - \varepsilon, x_1 + \varepsilon]$, as desired; multiplying the expression above by (3.1.3) and integrating w.r.t. x_0 and x_1 over $\mathbb{R}^2$ (in order to take all possible starting points and destinations into account) we get

$$\rho_t(x) = \int_{-\infty}^{+\infty} p(x - x_0, t - t_0) f(x_0)\mathrm{d}x_0 \int_{-\infty}^{+\infty} p(x_1 - x, t_1 - t) g(x_1)\mathrm{d}x_1 = \mathsf{h}_{t-t_0} f \mathsf{h}_{t_1 - t} g. \tag{3.1.9}$$

Concerning the first question we have thus shown that

> ... die Antwort nicht durch eine Lösung der fokkerschen Gleichung geliefert wird, sondern [...] durch das Produkt der Lösungen zweier adjungierter Gleichungen, wobei nicht der einzelnen Lösung, sondern dem Produkt zeitliche Grenzbedingungen auferlegt sind.[6]

[6] ...the answer is not given by a solution of the Fokker equation, but by the product of two adjoint equation, where the time boundary conditions are not imposed on the single solutions but on their product.

Therefore, the problem addressed by Schrödinger is a classical physics example that is much closer to quantum mechanics than the diffusion picture described by (3.1.2a) and (3.1.2b). Indeed (3.1.9) is the Euclidean analogue of Born's statistical interpretation of the (L^2-norm of the) wave function; moreover since ρ_t is the product of two functions $\mathsf{h}_{t-t_0}f$ and $\mathsf{h}_{t_1-t}g$ solving the heat equation and its adjoint respectively, we gain a time symmetry which is commonly absent in diffusion models and on the contrary is typical in oscillatory phenomena (as quantum mechanics).

Nevertheless, this is not sufficient for concluding the equivalence between the system (3.1.1a)-(3.1.1b) and (3.1.2a)-(3.1.2b), because the evolution of ρ_t is not oscillatory at all. Secondly, the wave function ψ corresponds to two real-valued functions and therefore it is sufficient to provide the value of ψ at a single instant, while in the framework of Schrödinger's system we do not look at the values of f and g at the same time but at the values of their product at two different times. For this reason [111] finishes with the following question:

> Doit-on interpréter la remarque d'Eddington, citée plus haut, comme signalant la nécessité de modifier cette manière de voir en mécanique ondulatoire et prendre comme conditions aux limites les valeurs d'une seule probabilité réelle à deux instants différents?[7]

Besides, the assignment of f and g at the same time is completely meaningless from a physical point of view, because f and g are not probabilities, they are not observable, just a fictitious representation; but if we multiply them, then we recover the description of a real diffusion process. It is tempting to transfer this consideration to quantum mechanics and thus claim that ψ and $\overline{\psi}$ are fictitious representation for the real world description encoded in $|\psi|^2$. Although Schrödinger did not achieve to overcome the discrepancies between quantum world and the model described by (3.1.8), the fictitious character of the wave function puzzled many physicists and is nowadays shared by many different interpretations of quantum mechanics, e.g. the Copenhagen interpretation, Quantum Bayesianism and Stochastic Quantum Mechanics. Nagasawa's contribution is in the vein of the last quantum theory cited and in his monograph [100] the equivalence between Schrödinger's equation and diffusion processes is established in Theorem 4.1 (the reader is addressed to it for a rigorous discussion), by providing a direct correspondence between solutions of the former and the latter. Forgetting about regularity issues, such correspondence reads as:

(i) If the couple (f, g) of real valued functions satisfies the system (3.1.2a)-(3.1.2b) and we define
$$R := \frac{1}{2}\log(fg) \qquad\qquad S := \frac{1}{2}\log(g/f),$$
then the complex-valued function given by
$$\psi(t,x) := e^{R(t,x)+iS(t,x)} \tag{3.1.10}$$
is a solution to (3.1.1a) with potential
$$V = -c - 2\frac{\partial S}{\partial t} - |\nabla S|^2 - 2\mathbf{b}\cdot\nabla S. \tag{3.1.11}$$

[7] Shall one interpret previously quoted Eddington's remark as a hint for the necessity of modifying the usual way of looking at quantum mechanics and taking instead the values of a single real probability at two different times as boundary conditions?

(ii) If the complex-valued function ψ is a solution to (3.1.1a) in the form (3.1.10), then the functions

$$f(t,x) := e^{R(t,x)-S(t,x)} \qquad g(t,x) := e^{R(t,x)+S(t,x)}$$

are solutions to (3.1.2a)-(3.1.2b) with creation and killing

$$c = -V - 2\frac{\partial S}{\partial t} - |\nabla S|^2 - 2\mathbf{b}\cdot\nabla S.$$

We shall not investigate any further the parallelism between Schrödinger's equation and diffusion processes (a part from a final consideration in Section 5.5), as it is not the purpose of the manuscript, but we believe that, after the presentation of Schrödinger's physical motivations behind [110], it was our duty towards the reader to give an exhaustive and self-contained presentation of the topic, thus including Nagasawa's solution.

3.2 Existence problems and statistical physics

After having established the equivalence between the Schrödinger equations (3.1.1a) and (3.1.1b) and the system of diffusion equations (3.1.2a) and (3.1.2b), let us come back to (3.1.9): the intermediate probability densities ρ_t are fully determined once we know f and g, i.e. once we have solved (3.1.8). However existence and uniqueness of solutions for (3.1.8) are very delicate issues, because if in 1931's paper Schrödinger did not investigate the problem, suggesting that it was evident

> Die Existenz und Eindeutigkeit der Lösung (abgesehen vielleicht von besonders tückisch vorgegebenen ρ_0, ρ_1[8]) halte ich wegen der Vernünftigkeit der Fragestellung, die ganz eindeutig und scharf auf diese Gleichungen führt, für ausgemacht.[9]

in 1932's subsequent paper he backed off

> Je n'ai pas pu réussir à prouver ni qu'il existe toujours de telles solutions, ni qu'elles sont uniques. Mais j'en suis tout à fait convaincu.[10]

A first partial solution was given by Fortet in [49], later generalized by Beurling in [20] and finally solved by [69]. Jamison's argument relies on reciprocal processes, a particular class of stochastic processes whose definition was proposed by Bernstein in [19] and then deeply investigated by Jamison himself in [68]. However, for the existence and uniqueness results we will present in Chapter 4 a different approach is preferable and passes through Föllmer formulation of (3.1.8) as an entropy minimization problem. For this reason we shall now sketch how the large deviation feature hidden in the problem addressed by Schrödinger leads to a variational problem, following [81].

The heuristic described by the German physicist is already a statistical physics approach, hence the involvement of large deviations should not be surprising, but it is worth spending

[8] With the notation of the present manuscript.

[9] I give for free existence and uniqueness of the solution (maybe excluding particularly treacherous ρ_0, ρ_1) because of the reasonable formulation of the problem, which leads in a unique and clear way to these equations.

[10] I couldn't manage to prove either that there always exist solutions, nor that they are unique. But I am absolutely sure about that.

some words about it because it enables to make rigorous the passage from the discrete to the continuum case that we performed before in order to get (3.1.8). As a first step, we have to restate Schrödinger experiment within a probabilistic language: consider a large number N of independent particles moving in $\mathbb{R}^d$ and describe their evolution via a family $B^1, ..., B^N$ of independent Brownian motions starting at $x_0^1, ..., x_0^N$ respectively. Then introduce the empirical probability measure

$$Z^N := \frac{1}{N}\sum_{k=1}^N \delta_{B^k},$$

because it is an equivalent description of our system, define as well the empirical measure at time t as

$$Z_t^N := \frac{1}{N}\sum_{k=1}^N \delta_{B_t^k}$$

and assume that the initial positions are a (finer and finer) discretization of a certain probability distribution, namely

$$Z_0^N = \frac{1}{N}\sum_{k=1}^N \delta_{x_0^k} \rightharpoonup \mu_0$$

for some $\mu_0 \in \mathscr{P}(\mathbb{R})^d$. By the law of large numbers, Z^N converges in law to the Brownian motion B_{μ_0} with μ_0 as initial distribution as $N \to \infty$, so that in particular at time $t = 1$

$$Z_1^N \rightharpoonup (B_{\mu_0})_1 := \left(\int_{\mathbb{R}^d} p(\cdot - x, 1)\mathrm{d}\mu_0(x)\right)\mathcal{L}^d(\mathrm{d}y)$$

where $p(x,t)$ is the fundamental solution of the heat equation. Schrödinger's example now reads as follows: assume that the observation at $t = 1$ of the empirical measure is different from the expected profile $(B_{\mu_0})_1$ for all N, so that we can say that there exists $\mu_1 \in \mathscr{P}(\mathbb{R})^d$ such that Z_1^N belongs to a neighbourhood of it not including $(B_{\mu_0})_1$; although little probable, this is actually possible because N is always finite and thus we can condition with respect to this rare event and, knowing Z_0^N and Z_1^N, ask: what is the most plausible evolution from the initial configuration to the final one?. And what happens as $N \to \infty$? While the first question received an exhaustive answer in [110], for the second one the situation is more delicate; if one relies on Schrödinger's computations (see [111]) then a key role is played by Stirling's formula, but for a different approach which better fits to more general situations (non Gaussian evolution, abstract framework, etc.) Sanov's theorem is recommended, as pointed out by Föllmer in [47]. Such theorem ensures that for any family (Y^k) of i.i.d. random variables with common law $\mathbf{R} \in \mathscr{P}(C([0,1],\mathbb{R}^d))$, then the empirical measures (Z^k) associated to it satisfy the large deviation principle

$$\mathrm{Prob}(Z^k \in A) \overset{k\to\infty}{\asymp} \exp\left(-k \inf_{\mathbf{Q}\in A} H(\mathbf{Q}\,|\,\mathbf{R})\right),$$

where A belongs to some large class of measurable subsets of $\mathscr{P}(C([0,1],\mathbb{R}^d))$ and $H(\cdot\,|\,\cdot)$ is the Boltzmann-Shannon entropy functional, set equal to

$$H(\sigma\,|\,\tilde{\nu}) := \int \rho\log(\rho)\,\mathrm{d}\tilde{\nu}$$

if $\sigma = \rho\nu$ and $+\infty$ otherwise (see Chapter 4 for the precise definition). In our case we deal with i.i.d. Brownian motions such that $Z_0^k \rightharpoonup \mu_0$ as $k \to \infty$: as a consequence the statement of Sanov's theorem is modified in the following way

$$\text{(3.2.1)} \qquad \mathrm{Prob}(Z^n \in A) \overset{n\to\infty}{\asymp} \exp\Big(- n \inf_{\substack{\mathbf{Q}\in A \\ (\mathrm{e}_0)_*\mathbf{Q}=\mu_0}} \Big(H(\mathbf{Q}\,|\,\mathbf{R}) - H(\mu_0\,|\,(\mathrm{e}_0)_*\mathbf{R})\Big)\Big).$$

The proof of it can be found in [40] and [24]. With this result at our disposal, choose a distance d_P on $\mathscr{P}(\mathbb{R})^d$ consistent with the weak topology (for instance Prokhorov's one) and look at $\mathrm{Prob}(Z^k \in \cdot\,|\,Z^k \in U_\varepsilon)$, where

$$U_\varepsilon := \{\mathbf{Q} \in \mathscr{P}(C([0,1],\mathbb{R}^d)) \,:\, (\mathrm{e}_0)_*\mathbf{Q} = \mu_0,\, \mathsf{d}_P((\mathrm{e}_1)_*\mathbf{Q}, \mu_1) < \varepsilon\}$$

and by $Z^k \in U_\varepsilon$ we mean that its law belongs to U_ε. From (3.2.1) we see that for all $\varepsilon > 0$ and many $A \in \mathscr{P}(C([0,1],\mathbb{R}^d))$ it holds

$$\mathrm{Prob}(Z^k \in \cdot\,|\,Z^k \in U_\varepsilon) \overset{k\to\infty}{\asymp} \exp\Big(- k\Big(\inf_{\mathbf{Q}\in A\cap U_\varepsilon} H(\mathbf{Q}\,|\,\mathbf{R}) - \inf_{\mathbf{Q}\in U_\varepsilon} H(\mathbf{Q}\,|\,\mathbf{R})\Big)\Big)$$

whence, with some technical work,

$$\lim_{k\to\infty} \mathrm{Prob}(Z^k \in \cdot\,|\,Z^k \in U_\varepsilon) = \delta_{\mathbf{P}_\varepsilon},$$

$\mathbf{P}_\varepsilon$ being the (unique) solution of the minimization problem

$$\inf_{\mathbf{Q}\in U_\varepsilon} H(\mathbf{Q}\,|\,\mathbf{R}).$$

Uniqueness is trivial by the strict convexity of the entropy functional, while existence follows by standard arguments in calculus of variations (see the forthcoming Proposition 4.1.2). The passage to the limit as $\varepsilon \downarrow 0$ can be handled by means of Γ-convergence: it is not difficult to see that the previous minimization problem converges to

$$\text{(3.2.2)} \qquad \inf_{\substack{\mathbf{Q}\in\mathscr{P}(C([0,1],\mathbb{R}^d)) \\ (\mathrm{e}_0)_*\mathbf{Q}=\mu_0,\,(\mathrm{e}_1)_*\mathbf{Q}=\mu_1}} H(\mathbf{P}\,|\,\mathbf{R})$$

and $\mathbf{P}_\varepsilon \rightharpoonup \mathbf{P}$, where $\mathbf{P}$ is the unique solution to (3.2.2).

We have thus provided a physical motivation for (3.1.8) and a large deviation explanation about how to pass from (3.1.8) to (3.2.2).

3.3 An inspiring problem: applications and developments

We conclude the chapter with a non-exhaustive list of the different domains which have been influenced by Schrödinger's seminal works [110] and [111] throughout the decades until now. We have already cited the theory of reciprocal processes: the birth with Bernstein in [19] and the deep investigation performed by Jamison in [67], [68] and [69]; we add the survey [82] for the references therein and for a fresh revisitation of [68].

Passing to more recent developments, let us recall the following research fields:

- Euclidean quantum mechanics: see [123], [35], [34], [31] and also [99], [101]; exploiting the parallelism between quantum mechanics and (diffusion) stochastic processes, it aims to transfer results by analogy from one domain to the other.

- Stochastic optimal control: starting with the seminal papers [88] and [38], where the Schrödinger problem is translated in terms of stochastic control, and passing through the several contributions of Mikami [89], [90], [91], [92], [94], [95] and the crucial works [96] (later generalized by X. Tan and N. Touzi [119]) and [97], we conclude with a useful insight provided by the survey [93].

- Penalized Monge-Kantorovich problem: since in many cases the solutions of the optimal transport problem with cost c are not regular enough, a possible way to overcome this obstacle is by adding an entropic penalization, namely

$$\inf_{\boldsymbol{\gamma}\in\mathrm{Adm}(\mu_0,\mu_1)} \left\{ \int_{\mathrm{X}^2} c \,\mathrm{d}\boldsymbol{\gamma} + \frac{1}{k} H(\boldsymbol{\gamma} \,|\, \mathsf{R}) \right\},$$

  the advantage being that the solutions of this new problem are regular perturbations of the optimal transport one; see [107], [108] and the recent advancements [37], [17], [18], [103] and [23].

Further historical remarks as well as more bibliographical details can be found in the survey [81].

Chapter 4

On a Polish space

Let (X, τ) be a Polish space (as shall always be throughout the whole chapter), $\mu_0, \mu_1 \in \mathscr{P}(\mathrm{X})$ and R be a non-negative Radon measure on X^2, i.e. for every point $z \in \mathrm{X}^2$ there exists a neighbourhood with finite mass w.r.t. R. Recall that $\boldsymbol{\gamma} \in \mathscr{P}(\mathrm{X}^2)$ is called transport plan for μ_0, μ_1 provided $\pi^0_* \boldsymbol{\gamma} = \mu_0$ and $\pi^1_* \boldsymbol{\gamma} = \mu_1$, where $\pi^0, \pi^1 : \mathrm{X}^2 \to \mathrm{X}$ are the canonical projections. Motivated by the Schrödinger system (3.1.8), we are interested in finding a transport plan of the form

$$\boldsymbol{\gamma} = f \otimes g \, \mathsf{R}$$

for certain Borel functions $f, g : \mathrm{X} \to [0, \infty)$, where $f \otimes g(x, y) := f(x) g(y)$, because in the RCD setting (f, g) is a solution to (3.1.8) if and only if $f \otimes g \, \mathsf{R} \in \mathrm{Adm}(\mu_0, \mu_1)$, provided we choose R as $\mathsf{r}_1 \mathfrak{m} \otimes \mathfrak{m}$, where r_1 is the heat kernel at time $t = 1$. The advantage of this new perspective is the great generality, since it does not require the existence of heat kernels, and as we shall see in the forthcoming subsection this problem can be solved in a unique way in very abstract frameworks. Moreover the plan $\boldsymbol{\gamma}$ can be found as the minimum of

$$\boldsymbol{\gamma}' \quad \mapsto \quad H(\boldsymbol{\gamma}' \,|\, \mathsf{R})$$

among all transport plans from μ_0 to μ_1, where $H(\,\cdot\,|\,\cdot\,)$ is the Boltzmann-Shannon entropy. Hence, to resume the situation, at the end of Chapter 3 we have seen, at least formally, how the Schrödinger system leads to an entropy minimization problem, which is dynamical in spirit; now a static entropy minimization problem has just spotted. The link between all these different facets of the same topic is formally clear, or at least not difficult to guess, but our purpose is to make it rigorous. This is going to be performed in this chapter, which aims to be the first part of a user's guide for non-probabilists to the Schrödinger problem; in fact, all the probabilistic terminology commonly adopted in the literature is here translated into an analytic language.

As regards the structure of the chapter, in Section 4.1 Föllmer's version of the problem is established in a precise way on general Polish spaces as well as a dual and a dynamical formulation. For them we present the basic existence theorems and stress the relationship between static, dynamic and dual solutions; moreover, although less general than the results already known in the literature, we provide the reader with a partially new existence theorem, whose peculiarity is the characterization of the (unique) minimizer of the Schrödinger problem among all transport plans between the marginal constraints.

In Section 4.2, mimicking some classical results of optimal transport theory (convexity of the cost, restriction property and stability) we build a toolbox for the Schrödinger problem

that, to the best of our knowledge, is still missing in the literature and hopefully might be used by the working mathematician.

4.1 Statement of the problem and first results

On the road to show the equivalence of the three problems, the first obstacle is the correct definition of the Boltzmann-Shannon entropy on a Polish space (see [80] for more details).

Given a σ-finite measure ν on a Polish space (Y, τ'), there exists a measurable function $W : \mathrm{Y} \to [0, \infty)$ such that

$$z_W := \int e^{-W} \mathrm{d}\nu < +\infty.$$

Introducing the probability measure $\nu_W := z_W^{-1} e^{-W} \nu$, for any $\sigma \in \mathscr{P}(\mathrm{Y})$ such that $\int W \mathrm{d}\sigma < +\infty$ the Boltzmann-Shannon entropy is defined as

$$H(\sigma \,|\, \nu) := H(\sigma \,|\, \nu_W) - \int W \mathrm{d}\sigma - \log z_W \tag{4.1.1}$$

where $H(\sigma \,|\, \nu_W)$ is in turn defined as

$$H(\sigma \,|\, \tilde{\nu}) := \begin{cases} \displaystyle\int \rho \log(\rho) \, \mathrm{d}\tilde{\nu} & \text{if } \sigma = \rho\tilde{\nu} \\ +\infty & \text{if } \sigma \not\ll \tilde{\nu} \end{cases}$$

for all $\tilde{\nu} \in \mathscr{P}(\mathrm{Y})$. The definition is meaningful, because if $\int W' \mathrm{d}\sigma < +\infty$ for another function W' such that $z_{W'} < +\infty$, then

$$H(\sigma \,|\, \nu_W) - \int W \mathrm{d}\sigma - \log z_W = H(\sigma \,|\, \nu_{W'}) - \int W' \mathrm{d}\sigma - \log z_{W'}.$$

Hence $H(\,\cdot\,|\,\nu)$ is well-defined for all $\sigma \in \mathscr{P}(\mathrm{Y})$ such that $\int W \mathrm{d}\sigma < +\infty$ for some non-negative measurable function W with $z_W < +\infty$. We shall denote this set by $\mathscr{P}^H(\mathrm{Y}, \nu)$, while the family of probability measures with finite entropy w.r.t. ν will be denoted by $D(H(\cdot\,|\,\nu))$, namely the domain of the relative entropy w.r.t. ν. The next lemma shows that measures belonging to $\mathscr{P}^H(\mathrm{Y}, \nu)$ or $D(H(\cdot\,|\,\nu))$ enjoy some integrability properties.

Lemma 4.1.1. *Given a σ-finite measure ν on a Polish space (Y, τ'), the following hold true:*

(i) if $\sigma \in \mathscr{P}^H(\mathrm{Y}, \nu)$, then either $\sigma \not\ll \nu$ or $(\log(\rho))^- \in L^1(\sigma)$, where $\rho := \frac{\mathrm{d}\sigma}{\mathrm{d}\nu}$;

(ii) if $\sigma \in D(H(\cdot\,|\,\nu))$, then $\sigma = \rho\nu$ and $\log(\rho) \in L^1(\sigma)$.

proof
(i) Let $\sigma \in \mathscr{P}^H(\mathrm{Y}, \nu)$ and assume that $\sigma \ll \nu$. Then, using the same notation as in (4.1.1),

$$\begin{aligned} \int \left(\log \frac{\mathrm{d}\sigma}{\mathrm{d}\nu} \right)^- \mathrm{d}\sigma &= \int \left(\log \frac{\mathrm{d}\sigma}{\mathrm{d}\nu_W} - W - \log z_W \right)^- \mathrm{d}\sigma \\ &\leq \int \frac{\mathrm{d}\sigma}{\mathrm{d}\nu_W} \left(\log \frac{\mathrm{d}\sigma}{\mathrm{d}\nu_W} \right)^- \mathrm{d}\nu_W + \|W\|_{L^1(\sigma)} + |\log z_W|. \end{aligned}$$

Since $z(\log z)^- \leq e^{-1}$ and ν_W is a probability measure by construction, the thesis follows.

(ii) Let $\sigma \in D(H(\cdot\,|\,\nu))$, $\sigma = \rho\nu$ and argue as above to get

$$\int \big(\log(\rho)\big)^+ \mathrm{d}\sigma \leq \int \left(\log \frac{\mathrm{d}\sigma}{\mathrm{d}\nu_W}\right)^+ \mathrm{d}\sigma + \|W\|_{L^1(\sigma)} + |\log z_W|.$$

The fact that $H(\sigma\,|\,\nu)$ is finite implies in particular that so is $H(\sigma\,|\,\nu_W)$ and since ν_W is a probability measure, this entails $\log(\rho_W) \in L^1(\sigma)$, where ρ_W is the Radon-Nikodym derivative of σ w.r.t. ν_W. Therefore $(\log\rho)^+ \in L^1(\sigma)$ and by (i) this is sufficient to conclude. □

With this said, let (X, τ) be a Polish space, $\mathbf{R}$ a non-negative Radon measure on the path space $C([0,1],\mathrm{X})$ and $\mu_0, \mu_1 \in \mathscr{P}(\mathrm{X})$. Since $C([0,1],\mathrm{X})$ is Polish too, $\mathbf{R}$ turns out to be a σ-finite measure. Then the **dynamical Schrödinger problem** associated to $\mathbf{R}$ with initial and final marginals respectively given by μ_0 and μ_1 is the minimization problem

$$(S_{dyn}) \qquad \inf\{H(\mathbf{Q}\,|\,\mathbf{R}) \,:\, \mathbf{Q} \in \mathscr{P}^H(C([0,1],\mathrm{X}),\mathbf{R}),\ (\mathrm{e}_0)_*\mathbf{Q} = \mu_0,\ (\mathrm{e}_1)_*\mathbf{Q} = \mu_1\}$$

which is precisely the one obtained from the Schrödinger system via a large deviations argument. The minimal value is called **dynamical entropic cost** and denoted by $\mathscr{I}_d(\mu_0,\mu_1)$.

If we define $\mathsf{R} := (\mathrm{e}_0, \mathrm{e}_1)_*\mathbf{R}$, then (S_{dyn}) gives rise to a static version of the same problem, analogously to (1.1.9) and (1.1.8): we will refer to it as **(static) Schrödinger problem** and it is defined, for the marginal constraints μ_0 and μ_1, in the following way

$$(\mathrm{S}) \qquad \inf\{H(\boldsymbol{\gamma}'\,|\,\mathsf{R}) \,:\, \boldsymbol{\gamma}' \in \mathscr{P}^H(\mathrm{X}^2,\mathsf{R}),\ \pi^0_*\boldsymbol{\gamma}' = \mu_0,\ \pi^1_*\boldsymbol{\gamma}' = \mu_1\}.$$

Its minimal value is called **static entropic cost** and denoted by $\mathscr{I}_s(\mu_0,\mu_1)$. The link between (S_{dyn}) and (S) is very strong and can be understood in terms of disintegration, because on the one hand the static Schrödinger problem is a projected version of the dynamical one and on the other hand the solution of the dynamical problem can be regarded as a superposition of $\mathbf{R}$-bridges weighted with the solution of the static problem, as proved in [47]. In more precise terms:

- if $\mathbf{P}$ is a solution to (S_{dyn}), then $(\mathrm{e}_0,\mathrm{e}_1)_*\mathbf{P}$ is a solution to (S);
- if $\boldsymbol{\gamma}$ is a solution to (S), then

  $$\mathbf{P}(\cdot) := \int_{\mathrm{X}^2} \mathsf{R}^{xy}(\cdot)\mathrm{d}\boldsymbol{\gamma}(x,y) \tag{4.1.2}$$

  is a solution to (S_{dyn}), where R^{xy} denotes the disintegration of $\mathbf{R}$ w.r.t. $(\mathrm{e}_0,\mathrm{e}_1)$, which in particular implies that $\boldsymbol{\gamma} = (\mathrm{e}_0,\mathrm{e}_1)_*\mathbf{P}$ by construction.

This is a consequence of the following identity, which is a particular case of (A.2.2),

$$H(\mathbf{P}\,|\,\mathbf{R}) = H(\boldsymbol{\gamma}\,|\,\mathsf{R}) + \int_{\mathrm{X}^2} H(\mathsf{P}^{xy}\,|\,\mathsf{R}^{xy})\mathrm{d}\boldsymbol{\gamma}(x,y),$$

where P^{xy} is the disintegration of $\mathbf{P}$ w.r.t. $(\mathrm{e}_0,\mathrm{e}_1)$ and $\boldsymbol{\gamma} = (\mathrm{e}_0,\mathrm{e}_1)_*\mathbf{P}$. From this identity and the fact that disintegrations are probability measures it follows that $H(\mathbf{P}\,|\,\mathbf{R}) \geq H(\boldsymbol{\gamma}\,|\,\mathsf{R})$ with equality if and only if $\mathbf{P}$ and $\mathbf{R}$ share $\boldsymbol{\gamma}$-almost the same bridges. Namely, for a solution $\mathbf{P}$ to (S_{dyn}) $\mathsf{P}^{xy} = \mathsf{R}^{xy}$ for $\boldsymbol{\gamma}$-a.e. $(x,y) \in \mathrm{X}^2$, which means that $\mathbf{P}$ and $\mathbf{R}$ have $\boldsymbol{\gamma}$-almost the same disintegration. This also tells us that

$$\inf H(\mathbf{Q}\,|\,\mathbf{R}) = \inf H(\boldsymbol{\gamma}'\,|\,\mathsf{R}),$$

where the infima are taken as in (S_{dyn}) and (S) respectively. As a byproduct, dynamical and static entropic costs coincide, thus we can lighten the notation and define the **entropic cost** tout court as

$$\mathscr{I}(\mu_0, \mu_1) := \mathscr{I}_d(\mu_0, \mu_1) = \mathscr{I}_s(\mu_0, \mu_1).$$

Hence the relationship between static and dynamical Schrödinger problem is clear and the existence of a solution for the former provides a solution for the latter and viceversa. However, how can we decide if a solution actually exists? A preliminary and rough answer is given by the following result, which only relies on Weierstrass' theorem.

Proposition 4.1.2. *Let* (X, τ) *be a Polish space,* R *a non-negative Radon measure on* X^2 *and* $\mu_0, \mu_1 \in \mathscr{P}(\mathrm{X})$*. Assume that there exists a Borel function* $B : \mathrm{X} \to [0, \infty)$ *such that*

$$\int_{\mathrm{X}^2} e^{-B(x)-B(y)} \mathrm{d}\mathsf{R}(x, y) < \infty \qquad \int B \, \mathrm{d}\mu_0 < \infty \qquad \int B \, \mathrm{d}\mu_1 < \infty.$$

Then (S_{dyn}) *and* (S) *admit a unique solution if and only if there exists* $\boldsymbol{\gamma} \in \mathrm{Adm}(\mu_0, \mu_1)$ *with finite entropy w.r.t.* R*.*

proof Existence follows by the direct method of calculus of variations: the class of transport plans is not empty, narrowly compact (see e.g. [5]) and $H(\,\cdot\,|\,\mathsf{R})$ is well-defined therein; indeed by assumption $\int W \mathrm{d}\sigma < +\infty$ with $W(x, y) := B(x) + B(y)$ for all transport plan σ. Moreover $H(\,\cdot\,|\,\mathsf{R})$ is narrowly lower semicontinuous on the whole set where it is well defined (see Corollary 2.3 in [80]), hence in particular on the class of transport plans.

Since $H(\,\cdot\,|\,\mathsf{R})$ is strictly convex, uniqueness is equivalent to $D(H(\cdot\,|\,\mathsf{R})) \neq \emptyset$. □

A natural question now arises: under which necessary and sufficient conditions can we find a transport plan as above, thus granting existence and uniqueness of the solution to (S)? First of all, still using (A.2.2) we see that for any $\boldsymbol{\gamma} \in \mathrm{Adm}(\mu_0, \mu_1)$

$$H(\mu_0\,|\,\mathsf{R}_0) \leq H(\boldsymbol{\gamma}\,|\,\mathsf{R}) \qquad H(\mu_1\,|\,\mathsf{R}_1) \leq H(\boldsymbol{\gamma}\,|\,\mathsf{R})$$

where $\mathsf{R}_0 := \pi^0_*\mathsf{R}$ and $\mathsf{R}_1 := \pi^1_*\mathsf{R}$, so that $H(\mu_0\,|\,\mathsf{R}_0) < \infty$ and $H(\mu_1\,|\,\mathsf{R}_1) < \infty$ are preliminary necessary conditions for the problem to be non-trivial. As regards sufficiency, the answer is not so clear, at least in terms of *minimal* set of assumptions; following [81], a possible choice is the one below.

Proposition 4.1.3. *Let* (X, τ) *be a Polish space,* R *a non-negative Radon measure on* X^2 *and* $\mu_0, \mu_1 \in \mathscr{P}(\mathrm{X})$*. Assume that* $\mathsf{R}_0 = \mathsf{R}_1 =: \mathfrak{m}$ *and that there exist some measurable functions* $\alpha, \beta : \mathrm{X} \to [0, \infty)$ *such that* $\mathrm{d}\mathsf{R}(x, y) \geq e^{-\alpha(x)-\alpha(y)} \mathrm{d}(\mathfrak{m} \otimes \mathfrak{m})(x, y)$ *and*

$$\int_{\mathrm{X}^2} e^{-\beta(x)-\beta(y)} \mathrm{d}\mathsf{R}(x, y) < \infty.$$

Suppose in addition that the constraints μ_0, μ_1 *are such that* $H(\mu_0\,|\,\mathfrak{m}), H(\mu_1\,|\,\mathfrak{m}) < \infty$ *and*

$$\int (\alpha + \beta) \mathrm{d}\mu_0 < \infty \qquad \int (\alpha + \beta) \mathrm{d}\mu_1 < \infty.$$

Then (S_{dyn}) *and* (S) *admit a unique solution.*

Of course, the same conclusion can be obtained from a different set of hypotheses, because a sufficiently large degree of freedom on them is possible. Indeed, as it will be clear from the proof of Proposition 4.1.5, all sufficient conditions for existence are formulated in order to grant the existence of a transport plan $\boldsymbol{\gamma}' \in \mathrm{Adm}(\mu_0, \mu_1)$ with finite entropy: a natural candidate is $\boldsymbol{\gamma}' = \mu_0 \otimes \mu_1$. For this reason, a strengthening of the assumptions on R (resp. on μ_0, μ_1) allows to weaken those on μ_0, μ_1 (resp. R), for instance.

However, the result above is not completely satisfactory, because it does not fully describe the structure of the minimizers of (S). In fact, as shown in [78] (in the paper R is a probability measure but using (4.1.1) this assumption can be removed) when a solution $\boldsymbol{\gamma}$ to (S) exists, then there always exist two functions $f, g : \mathrm{X} \to [0, \infty)$ such that

$$\frac{\mathrm{d}\boldsymbol{\gamma}}{\mathrm{d}\mathsf{R}} = f \otimes g, \qquad \boldsymbol{\gamma}\text{-a.e.}$$

and $f \otimes g$ is Borel. Nevertheless, we do not know yet whether $\boldsymbol{\gamma} = f \otimes g\, \mathsf{R}$ and f, g are Borel themselves. Counterexamples actually may exist, as shown in [48] and [78]. A possible way to overcome this obstacle is proposed in [81] and it is the following.

Proposition 4.1.4. *Let* (X, τ) *be a Polish space,* R *a non-negative Radon measure on* X^2 *and* $\mu_0, \mu_1 \in \mathscr{P}(\mathrm{X})$*. Assume that the same hypotheses of Proposition 4.1.3 hold true and suppose in addition that:*

(i) $\mathfrak{m} \otimes \mathfrak{m} \ll \mathsf{R}$ *or* $\mathsf{R} \ll \mathfrak{m} \otimes \mathfrak{m}$*;*

(ii) (μ_0, μ_1) *belongs to the intrinsic core of the set of all admissible constraints, i.e. the ones for which* (S) *has finite value; this means that there exist two (different) admissible constraints* (ν_0, ν_1) *and* (ν_0', ν_1') *such that*

$$(\mu_0, \mu_1) = (1 - t)(\nu_0, \nu_1) + t(\nu_0', \nu_1')$$

for some $t \in (0, 1)$*.*

Then (S) *admits a unique solution* $\boldsymbol{\gamma}$ *and* $\boldsymbol{\gamma} = f \otimes g\, \mathsf{R}$ *for suitable positive Borel functions* f, g*, which are uniquely determined up to the transformation* $(f, g) \mapsto (cf, g/c)$ *for some* $c > 0$*.*

From (4.1.2) it follows that the unique solution to (S_{dyn}) is then given by

$$\mathrm{d}\mathbf{P}(\gamma) = f(\gamma_0) g(\gamma_1) \mathrm{d}\mathbf{R}(\gamma),$$

which is a time-symmetric version of Doob's h-transform (see [41] and the monograph [42]). As we will better see in the next chapter, this representation of the dynamical minimizer as doubly conditioned process is perfectly in line with the physical interpretation we pointed out in the historical introduction.

Yet, for our purposes we prefer to present a different point of view, because assumption (ii) implies that $\frac{\mathrm{d}\boldsymbol{\gamma}}{\mathrm{d}\mathsf{R}} > 0$ R-a.e. and this is too restrictive, because it rules out the case of marginals with bounded supports.

The following proposition collects the basic properties of the minimizer of the Schrödinger problem: points (i) and (ii) of the statement are already known in the literature on the subject (see in particular [78], [22] and [107]), but we provide the proof for reader's sake; points (iii) and (iv) are actually new, because even if an analogous result can be found in [36] (cf. Theorem 3.1 therein) the strategy is different and our (local) L^∞ control is new. A complete proof has already been presented in [63] for the compact case; here we adapt the same argument with slight modifications, which will be pointed out.

Proposition 4.1.5. *Let* $(\mathrm{X}, \tau, \mathfrak{m})$ *be a Polish space equipped with a non-negative Radon measure* $\mathfrak{m}$ *and let* R *be a non-negative Radon measure on* X^2 *such that* $\mathsf{R}_0 = \mathsf{R}_1 = \mathfrak{m}$ *and*

$$\mathfrak{m} \otimes \mathfrak{m} \ll \mathsf{R} \ll \mathfrak{m} \otimes \mathfrak{m}.$$

Let $\mu_0 = \rho_0 \mathfrak{m}$ *and* $\mu_1 = \rho_1 \mathfrak{m}$ *be Borel probability measures and assume that there exists a Borel function* $B : \mathrm{X} \to [0, \infty)$ *such that*

$$\int_{\mathrm{X}^2} e^{-B(x)-B(y)} \mathrm{d}\mathsf{R}(x, y) < \infty \qquad \int B \, \mathrm{d}\mu_0 < \infty \qquad \int B \, \mathrm{d}\mu_1 < \infty. \tag{4.1.3}$$

Suppose in addition that

$$H(\mu_0 \otimes \mu_1 \,|\, \mathsf{R}) < +\infty. \tag{4.1.4}$$

Then:

(i) There exists a unique minimizer $\boldsymbol{\gamma}$ *of* $H(\,\cdot\,|\,\mathsf{R})$ *among all transport plans from* μ_0 *to* μ_1.

(ii) $\boldsymbol{\gamma} = f \otimes g \mathsf{R}$ *for appropriate Borel functions* $f, g : \mathrm{X} \to [0, \infty)$ *which are unique up to the trivial transformation* $(f, g) \mapsto (cf, g/c)$ *for some* $c > 0$.

(iii) Assume that

$$\mathsf{R} \geq (\alpha \mathfrak{m}) \otimes (\alpha \mathfrak{m}) \tag{4.1.5}$$

for a suitable Borel function $\alpha : \mathrm{X} \to (0, \infty)$ *locally bounded away from 0 and* ∞*, i.e. for every point* $x \in \mathrm{X}$ *there exist a neighbourhood* U_x *and constants* $0 < c_x \leq C_x < \infty$ *such that*

$$c_x \leq \alpha \leq C_x \qquad \mathfrak{m}\text{-a.e. in } U_x.$$

Suppose also that $(\log \alpha)^- \in L^1(\mu_0)$, $(\log \alpha)^- \in L^1(\mu_1)$ *and* $\mu_0, \mu_1 \in D(H(\cdot\,|\,\mathfrak{m}))$ *with* $\rho_0, \rho_1 \in L^\infty_{loc}(\mathfrak{m})$. *Then:*

(a) (4.1.4) *holds;*

(b) $f, g \in L^\infty_{loc}(\mathfrak{m})$ *and* $\boldsymbol{\gamma}$ *is the only transport plan which can be written as* $f' \otimes g' \mathsf{R}$ *for Borel functions* $f', g' : \mathrm{X} \to [0, \infty)$.

(iv) Instead of (iii), assume that

$$\mathsf{R} \geq c \mathfrak{m} \otimes \mathfrak{m} \tag{4.1.6}$$

in $P_0 \times P_1$*, where* $P_0 := \{\rho_0 > 0\}$ *and* $P_1 := \{\rho_1 > 0\}$*, for a suitable constant* $c > 0$. *Suppose also that* $\mu_0, \mu_1 \in \mathscr{P}^H(\mathrm{X}, \mathfrak{m})$ *with* $\rho_0, \rho_1 \in L^\infty(\mathfrak{m})$. *Then:*

(a) (4.1.4) *holds;*

(b) $f, g \in L^\infty(\mathfrak{m})$ *and* $\boldsymbol{\gamma}$ *is the only transport plan which can be written as* $f' \otimes g' \mathsf{R}$ *for Borel functions* $f', g' : \mathrm{X} \to [0, \infty)$.

Let us mention that, with minor adaptations, the proof of claim (a) of point (iii) also provides a demonstration of Proposition 4.1.3. As regards the necessary conditions $H(\mu_0 \,|\, \mathsf{R}_0) < \infty$ and $H(\mu_1 \,|\, \mathsf{R}_1) < \infty$, they are already included in (4.1.4) and they are not asked in statements (iii) and (iv) Furthermore, the assumption $\mathsf{R}_0 = \mathsf{R}_1 = \mathfrak{m}$ is only needed for claim (b) of points (iii) and (iv).

proof As already pointed out, the fact that (X, τ) is Polish implies that $\mathfrak{m}$ and R are σ-finite, so that it is possible to give a meaning to $H(\cdot \,|\, \mathfrak{m})$ and $H(\cdot \,|\, \mathsf{R})$ as discussed above.

(i) See Proposition 4.1.2. In particular, from its proof we know that $H(\cdot \,|\, \mathsf{R})$ is well defined in $\mathrm{Adm}(\mu_0, \mu_1)$.

(ii) The uniqueness part of the claim is trivial, so we concentrate on existence. Finiteness of the entropy in particular grants that $\boldsymbol{\gamma} \ll \mathsf{R}$. Put $p := \frac{\mathrm{d}\boldsymbol{\gamma}}{\mathrm{d}\mathsf{R}}$ and let $P_0 := \{\rho_0 > 0\}$, $P_1 := \{\rho_1 > 0\}$. We start claiming that

$$p > 0 \quad \mathfrak{m} \otimes \mathfrak{m}\text{-a.e. on } P_0 \times P_1. \tag{4.1.7}$$

Notice that since $\mathfrak{m} \otimes \mathfrak{m}$ and R are mutually absolutely continuous, the claim makes sense and arguing by contradiction we shall assume that $\mathsf{R}(Z) > 0$, where $Z := (P_0 \times P_1) \cap \{p = 0\}$.

Let $s := \frac{\mathrm{d}(\mu_0 \otimes \mu_1)}{\mathrm{d}\mathsf{R}}$ and for $\lambda \in (0,1)$ let us define $\Phi(\lambda) : \mathrm{X}^2 \to \mathbb{R}$ by

$$\Phi(\lambda) := \frac{u(p + \lambda(s - p)) - u(p)}{\lambda}, \qquad \text{where} \quad u(z) := z \log(z).$$

The convexity of u grants that $\Phi(\lambda) \le u(s) - u(p) \in L^1(\mathrm{X}^2, \mathsf{R})$ (recall that $H(\mu_0 \otimes \mu_1 \,|\, \mathsf{R}) < \infty$ by assumption and Lemma 4.1.1 implies the desired integrability) and that $\Phi(\lambda)$ is monotone decreasing as $\lambda \downarrow 0$. Moreover, on Z we have $\Phi(\lambda) \downarrow -\infty$ R-a.e. as $\lambda \downarrow 0$, thus the monotone convergence theorem ensures that

$$\lim_{\lambda \downarrow 0} \frac{H(\boldsymbol{\gamma} + \lambda(\mu_0 \otimes \mu_1 - \boldsymbol{\gamma}) \,|\, \mathsf{R}) - H(\boldsymbol{\gamma} \,|\, \mathsf{R})}{\lambda} = -\infty.$$

Since $\boldsymbol{\gamma} + \lambda(\mu_0 \otimes \mu_1 - \boldsymbol{\gamma})$ is a transport plan from μ_0 to μ_1 for $\lambda \in (0,1)$, this is in contradiction with the minimality of $\boldsymbol{\gamma}$ which grants that the left hand side is non-negative, hence Z is R-negligible, as desired.

Let us now pick $h \in L^\infty(\mathrm{X}^2, \boldsymbol{\gamma})$ such that $\pi^0_*(h\boldsymbol{\gamma}) = \pi^1_*(h\boldsymbol{\gamma}) = 0$ and $\varepsilon \in (0, \|h\|^{-1}_{L^\infty(\boldsymbol{\gamma})})$. Then $(1 + \varepsilon h)\boldsymbol{\gamma}$ is a transport plan from μ_0 to μ_1 and noticing that hp is well defined R-a.e. we have

$$\begin{aligned}
\|u((1+\varepsilon h)p)\|_{L^1(\mathsf{R})} &= \int |(1+\varepsilon h)p \log((1+\varepsilon h)p)| \,\mathrm{d}\mathsf{R} \\
&\le \int (1+\varepsilon h)p \,|\log p| \,\mathrm{d}\mathsf{R} + \int (1+\varepsilon h) \,|\log(1+\varepsilon h)| \,\mathrm{d}\boldsymbol{\gamma} \\
&\le \|1+\varepsilon h\|_{L^\infty(\boldsymbol{\gamma})} \|p \log p\|_{L^1(\mathsf{R})} + \|(1+\varepsilon h)\log(1+\varepsilon h)\|_{L^\infty(\boldsymbol{\gamma})},
\end{aligned}$$

so that $u((1+\varepsilon h)p) \in L^1(\mathsf{R})$. Then again by the monotone convergence theorem we get

$$\lim_{\varepsilon \downarrow 0} \frac{H((1+\varepsilon h)\boldsymbol{\gamma} \,|\, \mathsf{R}) - H(\boldsymbol{\gamma} \,|\, \mathsf{R})}{\varepsilon} = \int \lim_{\varepsilon \downarrow 0} \frac{u((1+\varepsilon h)p) - u(p)}{\varepsilon} \,\mathrm{d}\mathsf{R} = \int hp(\log p + 1) \,\mathrm{d}\mathsf{R}.$$

By the minimality of $\boldsymbol{\gamma}$ we know that the left hand side in this last identity is non-negative, thus after running the same computation with $-h$ in place of h and noticing that the choice of h grants that $\int hp\,\mathsf{dR} = \int h\,\mathrm{d}\boldsymbol{\gamma} = 0$ we obtain

$$\int hp\log(p)\,\mathsf{dR} = 0 \qquad \forall h \in L^\infty(\boldsymbol{\gamma}) \text{ such that } \pi^0_*(h\boldsymbol{\gamma}) = \pi^1_*(h\boldsymbol{\gamma}) = 0. \tag{4.1.8}$$

The rest of the argument is better understood by introducing the spaces $V, {}^\perp W \subset L^1(\boldsymbol{\gamma})$ and $V^\perp, W \subset L^\infty(\boldsymbol{\gamma})$ as follows

$$\begin{aligned}
V &:= \big\{f \in L^1(\boldsymbol{\gamma}) \ : \ f = \varphi \oplus \psi \text{ for some } \varphi \in L^0(\mathfrak{m}_{|P_0}),\ \psi \in L^0(\mathfrak{m}_{|P_1})\big\},\\
W &:= \big\{h \in L^\infty(\boldsymbol{\gamma}) \ : \ \pi^0_*(h\boldsymbol{\gamma}) = \pi^1_*(h\boldsymbol{\gamma}) = 0\big\},\\
V^\perp &:= \big\{h \in L^\infty(\boldsymbol{\gamma}) \ : \ \int fh\,\mathrm{d}\boldsymbol{\gamma} = 0\ \forall f \in V\big\},\\
{}^\perp W &:= \big\{f \in L^1(\boldsymbol{\gamma}) : \int fh\,\mathrm{d}\boldsymbol{\gamma} = 0\ \forall h \in W\big\},
\end{aligned}$$

where here and in the following the function $\varphi \oplus \psi$ is defined as $\varphi \oplus \psi(x,y) := \varphi(x) + \psi(y)$. Notice that the Euler equation (4.1.8) reads as $\log(p) \in {}^\perp W$ and our thesis as $\log(p) \in V$; hence to conclude it is sufficient to show that ${}^\perp W \subset V$.

Claim 1: V is a closed subspace of $L^1(\boldsymbol{\gamma})$.

We start claiming that $f \in V$ if and only if $f \in L^1(\boldsymbol{\gamma})$ and

$$f(x,y) + f(x',y') = f(x,y') + f(x',y) \quad \mathfrak{m}\otimes\mathfrak{m}\otimes\mathfrak{m}\otimes\mathfrak{m}\text{-a.e. } (x,x',y,y') \in P_0^2 \times P_1^2. \tag{4.1.9}$$

Indeed the 'only if' follows trivially from $\boldsymbol{\gamma} \ll \mathfrak{m}\otimes\mathfrak{m}$ and the definition of V. For the 'if' we apply Fubini's theorem to get the existence of $x' \in P_0$ and $y' \in P_1$ such that

$$f(x,y) + f(x',y') = f(x,y') + f(x',y) \quad \mathfrak{m}\otimes\mathfrak{m}\text{-a.e. } x,y \in P_0 \times P_1.$$

Thus $f = f(\cdot,y') \oplus (f(x',\cdot) - f(x',y'))$, as desired.

Now notice that since (4.1.7) grants that $(\mathfrak{m}\otimes\mathfrak{m})_{|P_0\times P_1} \ll \boldsymbol{\gamma}$, we see that the condition (4.1.9) is closed w.r.t. $L^1(\boldsymbol{\gamma})$-convergence.

Claim 2: $V^\perp \subset W$.

Let $h \in L^\infty(\boldsymbol{\gamma}) \setminus W$, so that either the first or second marginal of $h\boldsymbol{\gamma}$ is non-zero. Say the first. Thus since $\pi^0_*\boldsymbol{\gamma} = \mu_0$ we have $\pi^0_*(h\boldsymbol{\gamma}) = f_0\mu_0$ for some $f_0 \in L^\infty(\mu_0) \setminus \{0\}$. Then the function $f := f_0 \oplus 0 = f_0 \circ \pi^0$ belongs to V and we have

$$\int hf\,\mathrm{d}\boldsymbol{\gamma} = \int f_0 \circ \pi^0\,\mathrm{d}(h\boldsymbol{\gamma}) = \int f_0\,\mathrm{d}\pi^0_*(h\boldsymbol{\gamma}) = \int f_0^2\,\mathrm{d}\mu_0 > 0,$$

so that $h \notin V^\perp$.

Claim 3: ${}^\perp W \subset V$.

Let $f \in L^1(\boldsymbol{\gamma}) \setminus V$, use the fact that V is closed and the Hahn-Banach theorem to find $h \in L^\infty(\boldsymbol{\gamma}) \sim L^1(\boldsymbol{\gamma})^*$ such that $\int fh\,\mathrm{d}\boldsymbol{\gamma} \neq 0$ and $\int \tilde{f}h\,\mathrm{d}\boldsymbol{\gamma} = 0$ for every $\tilde{f} \in V$. Thus $h \in V^\perp$ and hence by the previous step $h \in W$. The fact that $\int fh\,\mathrm{d}\boldsymbol{\gamma} \neq 0$ shows that $f \notin {}^\perp W$, as desired.

(iii) Starting with (a), observe that by direct computations we have

$$\begin{aligned} H(\mu_0 \otimes \mu_1 \,|\, \mathsf{R}) &= H(\mu_0 \,|\, \mathfrak{m}) + H(\mu_1 \,|\, \mathfrak{m}) + \int \log\left(\frac{\mathrm{d}(\mathfrak{m}\otimes\mathfrak{m})}{\mathrm{d}\mathsf{R}}\right)\rho_0\otimes\rho_1 \mathrm{d}(\mathfrak{m}\otimes\mathfrak{m}) \\ &\le H(\mu_0 \,|\, \mathfrak{m}) + H(\mu_1 \,|\, \mathfrak{m}) - \int \log(\alpha)\mathrm{d}\mu_0 - \int \log(\alpha)\mathrm{d}\mu_1 \end{aligned}$$

and recalling that $\mu_0, \mu_1 \in D(H(\cdot \,|\, \mathfrak{m}))$, our assumptions grant that the right hand side is finite.

Passing to (b), let σ be a transport plan from μ_0 to μ_1 such that $\sigma = f' \otimes g'\mathsf{R}$ for suitable non-negative Borel functions f', g'. We claim that in this case it holds $f', g' \in L^\infty_{loc}(\mathfrak{m})$, leading in particular to the claim in the statement about $\boldsymbol{\gamma}$.

By disintegrating R w.r.t. π^0, from $\pi^0_*(f' \otimes g'\mathsf{R}) = \rho_0\mathfrak{m}$ and $\mathsf{R}_0 = \mathfrak{m}$ we get that

$$f'(x)\int g'(y)\,\mathrm{d}\mathsf{R}^x(y) = \rho_0(x) < +\infty, \quad \text{for } \mathfrak{m}\text{-a.e. } x \tag{4.1.10}$$

whence $g' \in L^1(\mathsf{R}^x)$ for $\mathfrak{m}$-a.e. x. Since from (4.1.5) we have that $\mathsf{R}^x \ge \alpha(x)(\alpha\mathfrak{m})$ for $\mathfrak{m}$-a.e. $x \in \mathrm{X}$, we see that $g' \in L^1(\alpha\mathfrak{m})$ with

$$\alpha(x)\|g'\|_{L^1(\alpha\mathfrak{m})} \le \int g'(y)\,\mathrm{d}\mathsf{R}^x(y) \quad \text{for } \mathfrak{m}\text{-a.e. } x \in \mathrm{X}$$

and thus (4.1.10) yields

$$f' \le \frac{\rho_0}{\alpha\|g'\|_{L^1(\alpha\mathfrak{m})}}, \quad \mathfrak{m}\text{-a.e.},$$

which is the desired local L^∞ bound on f', since α is locally bounded away from 0. By interchanging the roles of f' and g', the same conclusion follows for g'.

For the uniqueness of $\boldsymbol{\gamma}$, put $\varphi := \log f'$, $\psi := \log g'$, notice that by what we have just proved

$$\int \varphi\,\mathrm{d}\sigma = \int \varphi\,\mathrm{d}\mu_0 \le H(\mu_0 \,|\, \mathfrak{m}) - \int \log(\alpha)\mathrm{d}\mu_0 - \log\|g'\|_{L^1(\alpha\mathfrak{m})}$$

and our assumptions grant that the right-hand side is finite. On the other hand

$$\int \varphi \oplus \psi\,\mathrm{d}\sigma = H(\sigma \,|\, \mathsf{R}) > -\infty \tag{4.1.11}$$

because, as already remarked in Proposition 4.1.2, (4.1.3) implies that $H(\cdot \,|\, \mathsf{R})$ is well-defined on $\mathrm{Adm}(\mu_0, \mu_1)$. From these two facts we infer that

$$\varphi \circ \pi^0, \psi \circ \pi^1 \in L^1(\sigma). \tag{4.1.12}$$

Putting for brevity $p' := f' \otimes g'$ and arguing as before to justify the passage to the limit inside the integral we get

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}\lambda} H\big((1-\lambda)\sigma + \lambda\boldsymbol{\gamma} \,|\, \mathsf{R}\big)_{|\lambda=0^+} &= \int (p - p')\log(p')\,\mathrm{d}\mathsf{R} \\ &= \int \varphi \oplus \psi\,\mathrm{d}(\boldsymbol{\gamma} - \sigma) \\ \text{(by (4.1.12))} \qquad &= \int \varphi\,\mathrm{d}\pi^0_*(\boldsymbol{\gamma} - \sigma) + \int \psi\,\mathrm{d}\pi^1_*(\boldsymbol{\gamma} - \sigma) \\ \text{(because } \sigma \text{ and } \boldsymbol{\gamma} \text{ have the same marginals)} \qquad &= 0. \end{aligned}$$

This equality and the convexity of $H(\,\cdot\,|\,\mathsf{R})$ yield $H(\sigma\,|\,\mathsf{R}) \le H(\boldsymbol{\gamma}\,|\,\mathsf{R})$ and being $\boldsymbol{\gamma}$ the unique minimum of $H(\,\cdot\,|\,\mathsf{R})$ among transport plans from μ_0 to μ_1, we conclude that $\sigma = \boldsymbol{\gamma}$.

(iv) Arguing as for point (*iii*), it is easy to see that

$$H(\mu_0 \otimes \mu_1\,|\,\mathsf{R}) \le H(\mu_0\,|\,\mathfrak{m}) + H(\mu_1\,|\,\mathfrak{m}) - \log(c) < +\infty.$$

For (*b*) let σ be a transport plan from μ_0 to μ_1 such that $\sigma = f' \otimes g'\mathsf{R}$ for suitable non-negative Borel functions f', g'. We claim that in this case it holds $f', g' \in L^\infty(\mathfrak{m})$, leading in particular to the claim in the statement about $\boldsymbol{\gamma}$.

By disintegrating R w.r.t. π^0, from $\pi^0_*(f' \otimes g'\mathsf{R}) = \rho_0\mathfrak{m}$ and $\mathsf{R}_0 = \mathfrak{m}$ we get (4.1.10) again, whence $g' \in L^1(\mathsf{R}^x)$ for $\mathfrak{m}$-a.e. x. Notice then that the sets where f' and g' are positive must coincide with P_0 and P_1 respectively, up to $\mathfrak{m}$-negligible sets, so that nothing changes in (4.1.10) if we restrict the integral to P_1. Moreover, since from (4.1.6) we have that $\mathsf{R}^x \ge c\mathfrak{m}$ in P_1 for $\mathfrak{m}$-a.e. $x \in P_0$, we see that $g' \in L^1(\mathfrak{m})$ with

$$c\|g'\|_{L^1(\mathfrak{m})} \le \int g'(y)\,\mathrm{d}\mathsf{R}^x(y) \quad \text{for } \mathfrak{m}\text{-a.e. } x \in P_0$$

and thus (4.1.10) yields

$$f' \le \frac{\|\rho_0\|_{L^\infty(\mathfrak{m})}}{c\|g'\|_{L^1(\mathfrak{m})}}, \quad \mathfrak{m}\text{-a.e. in } P_0,$$

which is the desired L^∞ bound on f', because in $\mathrm{X} \setminus P_0$ we already know that f' vanishes $\mathfrak{m}$-a.e. By interchanging the roles of f' and g', the same conclusion follows for g'.

For the uniqueness of $\boldsymbol{\gamma}$, put $\varphi := \log f'$, $\psi := \log g'$ and notice that, by what we have just proved, they are bounded from above. Together with (4.1.11), this implies that (4.1.12) holds also in this case and thus, arguing as in (*iii*), the conclusion follows. □

Although pretty artificial in a Polish space, controls (4.1.5) and (4.1.6) become natural when moving to the RCD framework, as it will be better understood in the next chapter (see Theorem 5.1.1).

From the proof of (*iii*), and more precisely from (4.1.10) and the analogous identity for ρ_1, we see that

$$f(x)\int g(y)\,\mathrm{d}\mathsf{R}^x(y) = \rho_0(x) \qquad \text{for } \mathfrak{m}\text{-a.e. } x$$
$$g(y)\int f(x)\,\mathrm{d}\mathsf{R}^y(x) = \rho_1(y) \qquad \text{for } \mathfrak{m}\text{-a.e. } y$$

and this is the way (3.1.8) has to be interpreted in a Polish space. This means that the couple (f, g) provided by the minimizer $\boldsymbol{\gamma}$ of (S) is a solution to the Schrödinger system. As the passage from (3.1.8) to (S) is trivial, the circle is now closed: under a suitable set of assumptions granting existence of a solution for (S) (e.g. those of Proposition 4.1.4 or Proposition 4.1.5) (3.1.8), (S) and (S_{dyn}) are all equivalent.

For sake of information, let us mention that (S) and (S_{dyn}) admit dual formulations. Under the same assumptions of Proposition 4.1.4 and using the same notations introduced therein, they read as follows

$$\text{(D)} \qquad \sup_{\varphi,\psi \in C_B(\mathrm{X})} \left\{ \int \varphi\,\mathrm{d}\mu_0 + \int \psi\,\mathrm{d}\mu_1 - \log \int_{\mathrm{X}^2} e^{\varphi\oplus\psi}\mathrm{d}\mathsf{R} \right\}$$

$$(D_{dyn}) \qquad \sup_{\varphi,\psi\in C_B(\mathrm{X})} \left\{ \int \varphi \,\mathrm{d}\mu_0 + \int \psi \,\mathrm{d}\mu_1 - \log \int e^{\varphi(\gamma_0)+\psi(\gamma_1)} \mathrm{d}\mathbf{R} \right\}$$

where $C_B(\mathrm{X})$ denotes the space of continuous functions $\phi : \mathrm{X} \to \mathbb{R}$ such that $\phi/\max\{B,1\} \in L^\infty(\mathfrak{m})$. The nature of both maximization problems is strongly connected with the variational representation of the relative entropy, namely

$$H(\mu \,|\, \nu) = \sup \left\{ \int u \,\mathrm{d}\mu - \log \int e^u \,\mathrm{d}\nu \,:\, u \in C_B(\mathrm{X}) \right\} \tag{4.1.14}$$

valid for any σ-finite measure ν and for any probability measure μ for which $H(\mu \,|\, \nu)$ is well defined.

It is proved in [77] that (S) and (D) have the same value (in the paper R is a probability measure but taking (4.1.1) into account we can handle the case of a reference measure with infinite mass), hence (S_{dyn}) and (D_{dyn}) too but in general there is no reason for the dual attainment to be realized within $C_B(\mathrm{X}) \times C_B(\mathrm{X})$. Nevertheless, the fact that primal and dual problem have the same value allows to completely describe the maximizing couples, because via a heuristic argument if the maximal value of (D) is attained at (φ, ψ) and $\boldsymbol{\gamma}$ is the minimizer of (S), then

$$H(\boldsymbol{\gamma} \,|\, \mathsf{R}) = \int_{\mathrm{X}^2} \varphi \oplus \psi \,\mathrm{d}\boldsymbol{\gamma} - \log \int_{\mathrm{X}^2} e^{\varphi\oplus\psi} \mathrm{d}\mathsf{R}$$

and this identity together with the case of equality in (4.1.14) entails that $\boldsymbol{\gamma} = \exp(\varphi \oplus \psi)\mathsf{R}$. On the other hand we know that $\boldsymbol{\gamma} = f \otimes g\mathsf{R}$ with f, g uniquely determined up to the trivial transformation $(f, g) \mapsto (cf, g/c)$ for some $c > 0$. Therefore also φ, ψ are uniquely determined up to $(\varphi, \psi) \mapsto (\varphi + c', \psi - c')$ and they must be of the form $\varphi = \log(f)$, $\psi = \log(g)$.

The only defect is that, as already pointed out, in general (D) has no solutions in $C_B(\mathrm{X})$. Yet, when a solution to (S) exists and can be written with a factorized density, as in Proposition 4.1.4, it is still meaningful to set

$$\varphi := \log(f) \qquad\qquad \psi := \log(g)$$

where f and g are positive and $-\infty$ otherwise. By analogy with the optimal transport case, φ and ψ are called **Schrödinger potentials**.

4.2 A toolbox for the Schrödinger problem

Aim of this part is to look at those basic properties which are known to be true for solutions of the optimal transport problem and wonder if, up to suitable adaptations, they hold for the minimizers of the Schrödinger problem too. Some examples are the convexity of the cost, the restriction property and, most of all, the stability, which is fundamental from an applicative/algorithmic point of view. The investigation, as in the previous section, is carried out in the abstract framework of Polish spaces.

Let us begin with the proof of the convexity of the entropic cost, seen as a function of the marginal constraints.

Proposition 4.2.1 (Convexity of the entropic cost)**.** *Let* R *be a non-negative Radon measure on* X^2 *such that* $\mathsf{R}_0 = \mathsf{R}_1 =: \mathfrak{m}$*, let* (Θ, λ) *be a probability space and* $\vartheta \mapsto \mu_\vartheta$*,* $\vartheta \mapsto \nu_\vartheta$ *be two*

measurable functions from Θ *to* $D(H(\cdot\,|\,\mathfrak{m}))$*, the space of probability measures on* X *with finite entropy w.r.t.* $\mathfrak{m}$*. Define*

$$\mu := \int_\Theta \mu_\vartheta \mathrm{d}\lambda(\vartheta), \quad \nu := \int_\Theta \nu_\vartheta \mathrm{d}\lambda(\vartheta).$$

If we assume that R *satisfies the same assumptions as in Proposition 4.1.3 and, using the notation introduced therein,* μ, ν *are such that*

$$\int (\alpha+\beta)\mathrm{d}\mu < \infty \qquad \int (\alpha+\beta)\mathrm{d}\nu < \infty \tag{4.2.1}$$

then

$$\mathscr{I}(\mu,\nu) \le \int_\Theta \mathscr{I}(\mu_\vartheta,\nu_\vartheta)\lambda(\mathrm{d}\vartheta)$$

where $\mathscr{I}$ *denotes the entropic cost w.r.t.* R*.*

proof As a first step, notice that (4.2.1) implies

$$\int (\alpha+\beta)\mathrm{d}\mu_\vartheta < +\infty, \qquad \int (\alpha+\beta)\mathrm{d}\nu_\vartheta < +\infty \tag{4.2.2}$$

for λ-a.e. $\vartheta \in \Theta$ and, on the other hand, the entropy of μ_ϑ and ν_ϑ w.r.t. $\mathfrak{m}$ is finite by construction. Hence, for λ-a.e. $\vartheta \in \Theta$ all the assumptions of Proposition 4.1.3 are satisfied by $\mathsf{R}, \mu_\vartheta, \nu_\vartheta$: denote by $\boldsymbol{\gamma}_\vartheta$ the unique solution of the Schrödinger problem associated to them. Let us then define

$$\boldsymbol{\gamma} := \int_\Theta \boldsymbol{\gamma}_\vartheta \mathrm{d}\lambda(\vartheta)$$

and observe that it is a competitor in the Schrödinger problem, because $\pi^0_*\boldsymbol{\gamma} = \mu$ and $\pi^1_*\boldsymbol{\gamma} = \nu$; in addition $\boldsymbol{\gamma} \ll \mathsf{R}$, because $\boldsymbol{\gamma}_\vartheta \ll \mathsf{R}$ for λ-a.e. $\vartheta \in \Theta$, so that we can consider the Radon-Nikodym derivative of $\boldsymbol{\gamma}$ w.r.t. R. As a consequence,

$$\mathscr{I}(\mu,\nu) \le H(\boldsymbol{\gamma}\,|\,\mathsf{R}) = \int_{\mathrm{X}^2} \left(\int_\Theta \frac{\mathrm{d}\boldsymbol{\gamma}_\vartheta}{\mathrm{d}\mathsf{R}} \mathrm{d}\lambda(\vartheta) \right) \log \left(\int_\Theta \frac{\mathrm{d}\boldsymbol{\gamma}_\vartheta}{\mathrm{d}\mathsf{R}} \mathrm{d}\lambda(\vartheta) \right) \mathrm{d}\mathsf{R}$$

and if we now use Jensen's inequality we get

$$\begin{aligned}\mathscr{I}(\mu,\nu) &\le \int_{\mathrm{X}^2} \int_\Theta \frac{\mathrm{d}\boldsymbol{\gamma}_\vartheta}{\mathrm{d}\mathsf{R}} \log\left(\frac{\mathrm{d}\boldsymbol{\gamma}_\vartheta}{\mathrm{d}\mathsf{R}}\right) \mathrm{d}\lambda(\vartheta)\mathrm{d}\mathsf{R} = \int_\Theta \int_{\mathrm{X}^2} \frac{\mathrm{d}\boldsymbol{\gamma}_\vartheta}{\mathrm{d}\mathsf{R}} \log\left(\frac{\mathrm{d}\boldsymbol{\gamma}_\vartheta}{\mathrm{d}\mathsf{R}}\right) \mathrm{d}\mathsf{R}\mathrm{d}\lambda(\vartheta) \\ &= \int_\Theta H(\boldsymbol{\gamma}_\vartheta\,|\,\mathsf{R})\mathrm{d}\lambda(\vartheta) = \int_\Theta \mathscr{I}(\mu_\vartheta,\nu_\vartheta)\mathrm{d}\lambda(\vartheta)\end{aligned}$$

where the last identity comes from the fact that $\boldsymbol{\gamma}_\vartheta$ is optimal for the constraint $(\mu_\vartheta, \nu_\vartheta)$ and the reference measure R for λ-a.e. $\vartheta \in \Theta$, thus concluding. □

It is worth mentioning a couple of further remarks about the proposition that we have just shown. First of all, while (4.2.1) implies (4.2.2), we can not say that the finiteness of $H(\mu\,|\,\mathfrak{m})$ and $H(\nu\,|\,\mathfrak{m})$ entails the finiteness of $H(\mu_\vartheta\,|\,\mathfrak{m})$ and $H(\nu_\vartheta\,|\,\mathfrak{m})$ for λ-a.e. $\vartheta \in \Theta$, whence the necessity of the hypothesis $\mu_\vartheta, \nu_\vartheta \in D(H(\cdot\,|\,\mathfrak{m}))$. Let us also point out that we did not assume $H(\mu\,|\,\mathfrak{m})$ and $H(\nu\,|\,\mathfrak{m})$ to be finite; however, one can observe that

$$\begin{aligned}H(\mu\,|\,\mathfrak{m}) &= \int \left(\int_\Theta \frac{\mathrm{d}\mu_\vartheta}{\mathrm{d}\mathfrak{m}} \mathrm{d}\lambda(\vartheta) \right) \log \left(\int_\Theta \frac{\mathrm{d}\mu_\vartheta}{\mathrm{d}\mathfrak{m}} \mathrm{d}\lambda(\vartheta) \right) \mathrm{d}\mathfrak{m} \\ &\le \int_\Theta \int \frac{\mathrm{d}\mu_\vartheta}{\mathrm{d}\mathfrak{m}} \log\left(\frac{\mathrm{d}\mu_\vartheta}{\mathrm{d}\mathfrak{m}}\right) \mathrm{d}\mathfrak{m}\mathrm{d}\lambda(\vartheta) = \int_\Theta H(\mu_\vartheta\,|\,\mathfrak{m})\mathrm{d}\lambda(\vartheta)\end{aligned} \tag{4.2.3}$$

where the inequality is due again to Jensen's inequality and in a completely analogous way

$$H(\nu \,|\, \mathfrak{m}) \leq \int_\Theta H(\nu_\vartheta \,|\, \mathfrak{m}) \mathrm{d}\lambda(\vartheta). \tag{4.2.4}$$

Hence, if we assume the right-hand sides of (4.2.3) and (4.2.4) to be finite, then $H(\mu \,|\, \mathfrak{m})$ and $H(\nu \,|\, \mathfrak{m})$ are finite too, which in particular yields finiteness of the entropic cost and existence of a (unique) solution to (S). The opposite implication is false in general, as we have already said, but it turns out to be true if Θ is a countable set and the proof goes as follows.

Let $\Theta = (\vartheta_i)_{i\in\mathbb{N}}$ and $\lambda = \sum_{i\in\mathbb{N}} \alpha_i \delta_{\vartheta_i}$ with $\alpha_i > 0$ for every $i \in \mathbb{N}$ and $\sum_{i\in\mathbb{N}} \alpha_i = 1$. Then observe that

$$H(\mu \,|\, \mathfrak{m}) = \iint_\Theta \frac{\mathrm{d}\mu_\vartheta}{\mathrm{d}\mathfrak{m}} \log\Big(\frac{\mathrm{d}\mu}{\mathrm{d}\mathfrak{m}}\Big) \lambda(\mathrm{d}\vartheta)\mathrm{d}\mathfrak{m} = \int_\Theta \int \Big(\frac{\mathrm{d}\mu_\vartheta}{\mathrm{d}\mathfrak{m}} \log\Big(\frac{\mathrm{d}\mu}{\mathrm{d}\mathfrak{m}}\Big)\mathrm{d}\mathfrak{m}\Big)\lambda(\mathrm{d}\vartheta)$$

and since $H(\mu \,|\, \mathfrak{m})$ is finite, we deduce that

$$\int \log\Big(\frac{\mathrm{d}\mu}{\mathrm{d}\mathfrak{m}}\Big)\mathrm{d}\mu_\vartheta < +\infty, \qquad \lambda\text{-a.e.}$$

whence

$$\int \log\Big(\frac{\mathrm{d}\mu}{\mathrm{d}\mathfrak{m}}\Big)\mathrm{d}\mu_\vartheta < +\infty, \qquad \forall \vartheta \in \Theta$$

By applying Jensen's inequality to $z \mapsto -\log z$ we deduce, for any $\vartheta \in \Theta$,

$$\begin{aligned}\int \log\Big(\frac{\mathrm{d}\mu}{\mathrm{d}\mathfrak{m}}\Big)\mathrm{d}\mu_\vartheta &= \int \log\Big(\int_\Theta \frac{\mathrm{d}\mu_\varphi}{\mathrm{d}\mathfrak{m}}\mathrm{d}\lambda(\varphi)\Big)\mathrm{d}\mu_\vartheta \geq \iint_\Theta \log\Big(\frac{\mathrm{d}\mu_\varphi}{\mathrm{d}\mathfrak{m}}\Big)\mathrm{d}\lambda(\varphi)\mathrm{d}\mu_\vartheta \\ &= \int_\Theta \int \frac{\mathrm{d}\mu_\vartheta}{\mathrm{d}\mathfrak{m}} \log\Big(\frac{\mathrm{d}\mu_\varphi}{\mathrm{d}\mathfrak{m}}\Big)\mathrm{d}\mathfrak{m}\mathrm{d}\lambda(\varphi)\end{aligned}$$

and combining this inequality with previous information we get

$$\int \frac{\mathrm{d}\mu_\vartheta}{\mathrm{d}\mathfrak{m}} \log\Big(\frac{\mathrm{d}\mu_\varphi}{\mathrm{d}\mathfrak{m}}\Big)\mathrm{d}\mathfrak{m} < +\infty, \qquad \forall \vartheta, \varphi \in \Theta \tag{4.2.5}$$

and in particular, for $\varphi = \vartheta$, $H(\mu_\vartheta \,|\, \mathfrak{m}) < +\infty$. A completely analogous argument works for ν.

If Θ were not countable, then (4.2.5) would be true for $\lambda \otimes \lambda$-a.e. $(\vartheta, \varphi) \in \Theta^2$ and this is not enough to conclude, because there could exist a set Θ' of λ-positive measure such that (4.2.5) fails everywhere on $\{(\vartheta, \vartheta) : \vartheta \in \Theta'\}$.

The second good news about minimizers of (S) and (S_{dyn}) is that they are locally optimal, in analogy with optimal transport. The rough idea is the following: if $\boldsymbol{\gamma}$ represents the most likely coupling from μ_0 to μ_1 according to R, namely the most likely way to transfer particles from the initial distribution μ_0 onto the final one μ_1 knowing that they should move according to R, then also partial transfers of mass are optimally described by $\boldsymbol{\gamma}$. The precise statement is the following.

Proposition 4.2.2 (Space restriction property)**.** *Let $\mathbf{R}$ be a non-negative Radon measure on $C([0,1], \mathrm{X})$ and $\mu_0, \mu_1 \in \mathscr{P}(\mathrm{X})$ be such that $\mathsf{R} := (\mathrm{e}_0, \mathrm{e}_1)_*\mathbf{R}, \mu_0, \mu_1$ satisfy the assumption of Proposition 4.1.3. Let $\boldsymbol{\gamma}$ and $\mathbf{P}$ be the unique solutions to* (S) *and* (S_{dyn}) *respectively. Then the following hold:*

(i) (dynamical version) if $\Omega \subset C([0,1],\mathrm{X})$ is a Borel set of curves such that $\mathbf{P}(\Omega) > 0$, then the probability measure $\mathbf{P}'$ defined by

$$\mathbf{P}'(A) := \frac{1}{\mathbf{P}(\Omega)}\mathbf{P}(A \cap \Omega) \qquad \forall A \in \mathscr{B}(C([0,1],\mathrm{X}))$$

is the unique solution to the Schrödinger problem with constraints $(\mathrm{e}_0)_*\mathbf{P}'$, $(e_1)_*\mathbf{P}'$ and reference measure $\mathbf{R}'(\cdot) := \mathbf{R}(\cdot \cap \Omega)$;

(ii) (static version) if $\Omega \subset \mathrm{X}^2$ is a Borel set such that $\boldsymbol{\gamma}(\Omega) > 0$, then the probability measure $\boldsymbol{\gamma}'$ defined by

$$\boldsymbol{\gamma}'(A) := \frac{1}{\boldsymbol{\gamma}(\Omega)}\boldsymbol{\gamma}(A \cap \Omega) \qquad \forall A \in \mathscr{B}(\mathrm{X}^2)$$

is the unique solution to the Schrödinger problem with constraints $\pi^0_*\boldsymbol{\gamma}'$, $\pi^1_*\boldsymbol{\gamma}'$ and reference measure $\mathsf{R}'(\cdot) := \mathsf{R}(\cdot \cap \Omega)$.

proof

(i) Let Ω be as in the statement. As a first remark, as a byproduct of Lemma 4.1.1 we have

$$\int \left(\log \frac{\mathrm{d}\mathbf{P}}{\mathrm{d}\mathbf{R}}\right)^- \mathrm{d}\mathbf{P} < +\infty$$

and this implies that $H(\mathbf{P}' \,|\, \mathbf{R}')$ is finite, because

$$\begin{aligned} H(\mathbf{P}' \,|\, \mathbf{R}') &= \int_\Omega \log\left(\frac{1}{\mathbf{P}(\Omega)}\frac{\mathrm{d}\mathbf{P}}{\mathrm{d}\mathbf{R}}\right)\frac{1}{\mathbf{P}(\Omega)}\mathrm{d}\mathbf{P} = \log\frac{1}{\mathbf{P}(\Omega)} + \frac{1}{\mathbf{P}(\Omega)}\int_\Omega \log\frac{\mathrm{d}\mathbf{P}}{\mathrm{d}\mathbf{R}}\mathrm{d}\mathbf{P} \\ &\le \log\frac{1}{\mathbf{P}(\Omega)} + \frac{1}{\mathbf{P}(\Omega)}\int_\Omega \left(\log\frac{\mathrm{d}\mathbf{P}}{\mathrm{d}\mathbf{R}}\right)^+ \mathrm{d}\mathbf{P} < +\infty. \end{aligned}$$

By Proposition 4.1.2 existence for the restricted Schrödinger problem is thus ensured. Now assume by contradiction that the thesis is false. Then there exists a probability measure $\mathbf{P}''$ (concentrated on Ω) such that

$$\int_\Omega \log\frac{\mathrm{d}\mathbf{P}''}{\mathrm{d}\mathbf{R}'}\mathrm{d}\mathbf{P}'' = H(\mathbf{P}'' \,|\, \mathbf{R}') < H(\mathbf{P}' \,|\, \mathbf{R}') = \int_\Omega \log\frac{\mathrm{d}\mathbf{P}'}{\mathrm{d}\mathbf{R}'}\mathrm{d}\mathbf{P}' \tag{4.2.6}$$

and $\mathbf{P}''_0 = \mathbf{P}'_0$, $\mathbf{P}''_1 = \mathbf{P}'_1$. Now, let us consider

$$\hat{\mathbf{P}} := (\mathbf{P} - \mathbf{P}_{|\Omega}) + \mathbf{P}(\Omega)\mathbf{P}'' = \mathbf{P}_{|\Omega^c} + \mathbf{P}(\Omega)\mathbf{P}''$$

and observe that $\hat{\mathbf{P}}$ is a probability measure on $C([0,1],\mathrm{X})$; such a measure can be equivalently rewritten as

$$\hat{\mathbf{P}} = \mathbf{P} + \mathbf{P}(\Omega)(\mathbf{P}'' - \mathbf{P}')$$

and this shows that $\hat{\mathbf{P}}$ has the same marginals as $\mathbf{P}$, so that it is a competitor in (S_{dyn}). In addition, by the very definition of $\hat{\mathbf{P}}$ we also deduce that $\hat{\mathbf{P}} = \mathbf{P}(\Omega)\mathbf{P}''$ on Ω and $\hat{\mathbf{P}} = \mathbf{P}$ on Ω^c, so that

$$\begin{aligned} H(\hat{\mathbf{P}} \,|\, \mathbf{R}) &= \int_\Omega \log\frac{\mathrm{d}\hat{\mathbf{P}}}{\mathrm{d}\mathbf{R}}\mathrm{d}\hat{\mathbf{P}} + \int_{\Omega^c} \log\frac{\mathrm{d}\hat{\mathbf{P}}}{\mathrm{d}\mathbf{R}}\mathrm{d}\hat{\mathbf{P}} \\ &= \int_\Omega \log\left(\mathbf{P}(\Omega)\frac{\mathrm{d}\mathbf{P}''}{\mathrm{d}\mathbf{R}}\right)\mathbf{P}(\Omega)\mathrm{d}\mathbf{P}'' + \int_{\Omega^c} \log\frac{\mathrm{d}\mathbf{P}}{\mathrm{d}\mathbf{R}}\mathrm{d}\mathbf{P} \\ &= \underbrace{\mathbf{P}(\Omega)\log\mathbf{P}(\Omega) + \mathbf{P}(\Omega)\int_\Omega \log\frac{\mathrm{d}\mathbf{P}''}{\mathrm{d}\mathbf{R}}\mathrm{d}\mathbf{P}''}_{\alpha} + \int_{\Omega^c} \log\frac{\mathrm{d}\mathbf{P}}{\mathrm{d}\mathbf{R}}\mathrm{d}\mathbf{P} \end{aligned}$$

Now observe that

$$\int_\Omega \log \frac{\mathrm{d}\mathbf{P}''}{\mathrm{d}\mathbf{R}} \mathrm{d}\mathbf{P}'' = \int_\Omega \log \frac{\mathrm{d}\mathbf{P}''}{\mathrm{d}\mathbf{R}'} \mathrm{d}\mathbf{P}''$$

and the same identity holds with $\mathbf{P}'$ in place of $\mathbf{P}''$. Hence, by this fact and by (4.2.6) we deduce

$$\begin{aligned}\alpha &= \mathbf{P}(\Omega)\log\mathbf{P}(\Omega) + \mathbf{P}(\Omega)\int_\Omega \log\frac{\mathrm{d}\mathbf{P}''}{\mathrm{d}\mathbf{R}'}\mathrm{d}\mathbf{P}'' < \mathbf{P}(\Omega)\log\mathbf{P}(\Omega) + \mathbf{P}(\Omega)\int_\Omega \log\frac{\mathrm{d}\mathbf{P}'}{\mathrm{d}\mathbf{R}'}\mathrm{d}\mathbf{P}' \\ &= \mathbf{P}(\Omega)\log\mathbf{P}(\Omega) + \mathbf{P}(\Omega)\int_\Omega \log\frac{\mathrm{d}\mathbf{P}'}{\mathrm{d}\mathbf{R}}\mathrm{d}\mathbf{P}' = \int_\Omega \log\frac{\mathrm{d}\mathbf{P}}{\mathrm{d}\mathbf{R}}\mathrm{d}\mathbf{P}.\end{aligned}$$

In conclusion

$$H(\hat{\mathbf{P}}\,|\,\mathbf{R}) < \int_\Omega \log\frac{\mathrm{d}\mathbf{P}}{\mathrm{d}\mathbf{R}}\mathrm{d}\mathbf{P} + \int_{\Omega^c} \log\frac{\mathrm{d}\mathbf{P}}{\mathrm{d}\mathbf{R}}\mathrm{d}\mathbf{P} = H(\mathbf{P}\,|\,\mathbf{R})$$

and this is impossible by the minimality of $\mathbf{P}$.
(ii) As the proof of (i) is purely set-theoretic and the dynamic aspect of R is not involved, we can obviously restate this result for the static Schrödinger problem. □

This result admits a natural interpretation in terms of conditioning. In fact, if we look at $\boldsymbol{\gamma}$ as the joint law of a random vector (Z_0, Z_1) and $\Omega \subset \mathrm{X}^2$ is such that $\boldsymbol{\gamma}(\Omega) > 0$, then the law of (Z_0, Z_1) conditioned to lie in Ω is optimal for the Schrödinger problem with reference measure R' and constraint (μ_0', μ_1'), where μ_0' (resp. μ_1') is the law of Z_0 (resp. Z_1) conditioned by the event $(Z_0, Z_1) \in \Omega$. For the dynamical problem an analogous explanation is possible and a standard conditioning is $Z_t \in A$ with $A \subset \mathrm{X}$, where $(Z_t)_{t\in[0,1]}$ is a stochastic process whose law is $\mathbf{P}$ (existence of such a process is possible under precise assumptions on $\mathbf{R}$, see [81]).

We conclude the section with a stability result. Although quite *ad hoc* assumptions are formulated in order to overcome the difficulty of working in a Polish space, the conclusion is still useful, as witnessed by the applications provided in Section 6.3.

Theorem 4.2.3 (Stability)**.** *Let* $(\mathrm{X}, \tau, \mathfrak{m})$ *be a Polish space equipped with a non-negative Radon measure,* R *a non-negative Radon measure on* X^2 *such that* $\mathsf{R}_0 = \mathsf{R}_1 = \mathfrak{m}$ *and* $\alpha, \beta : \mathrm{X} \to (0, \infty)$ *Borel functions locally bounded away from 0 and* ∞ *(in the sense of Proposition 4.1.5) such that*

$$(\alpha\mathfrak{m}) \otimes (\alpha\mathfrak{m}) \le \mathsf{R} \le (\beta\mathfrak{m}) \otimes (\beta\mathfrak{m}). \tag{4.2.7}$$

Let $\mu_0^k = \rho_0^k\mathfrak{m}$ *and* $\mu_1^k = \rho_1^k\mathfrak{m}$ *be Borel probability measures such that* ρ_0^k, ρ_1^k *are uniformly locally bounded. Assume that the Schrödinger problem associated to* μ_0^k, μ_1^k *and* R *admits a unique solution, denoted by* $\boldsymbol{\gamma}^k$*. Assume that* $\mu_0^k \rightharpoonup \mu_0$ *and* $\mu_1^k \rightharpoonup \mu_1$ *as* $k \to \infty$ *for suitable* $\mu_0, \mu_1 \in \mathscr{P}(\mathrm{X})$*.*

Then:

(i) there exists $\boldsymbol{\gamma} \in \mathrm{Adm}(\mu_0, \mu_1)$ *which can be written as* $\boldsymbol{\gamma} = f \otimes g\mathsf{R}$ *for suitable Borel functions* $f, g \in L^\infty_{loc}(\mathfrak{m})$ *such that* $\boldsymbol{\gamma}^k \rightharpoonup \boldsymbol{\gamma}$ *as* $k \to \infty$*, up to extract a subsequence;*

(ii) if we further suppose that μ_0, μ_1 *are such that* (4.1.3) *holds,* $(\log\alpha)^- \in L^1(\mu_0)$*,* $(\log\alpha)^- \in L^1(\mu_1)$ *and*

$$\liminf_{k\to\infty} H(\boldsymbol{\gamma}^k\,|\,\mathsf{R}) < +\infty, \tag{4.2.8}$$

then the entropic cost $\mathscr{I}(\mu_0,\mu_1)$ is finite, $\boldsymbol{\gamma}$ is the unique solution to the Schrödinger problem relative to μ_0,μ_1,R and the whole sequence $(\boldsymbol{\gamma}^k)$ converges to it;

(iii) given a non-negative Radon measure $\mathbf{R}$ on $C([0,1],\mathrm{X})$ such that $(\mathrm{e}_0,\mathrm{e}_1)_*\mathbf{R}=\mathsf{R}$, if we denote by $\mathbf{P}^k$ the unique solution to (S_{dyn}) associated to $\mu_0^k,\mu_1^k,\mathbf{R}$ and assume (together with all the hypotheses of point (ii)) that

$$\mathrm{X}^2\ni(x,y)\quad\mapsto\quad\mathsf{R}^{xy}\in\mathscr{P}(C([0,1],\mathrm{X}))$$

is continuous, where $\{\mathsf{R}^{xy}\}_{x,y\in\mathrm{X}}$ is the disintegration of $\mathbf{R}$ w.r.t. $(\mathrm{e}_0,\mathrm{e}_1)$ and the target is endowed with the narrow topology, then $\mathbf{P}^k\rightharpoonup\mathbf{P}$ as $k\to\infty$, where $\mathbf{P}$ is the unique solution to the Schrödinger problem relative to $\mu_0,\mu_1,\mathbf{R}$.

proof First of all $\{\boldsymbol{\gamma}^k\}$ is a tight family, thus there exists $\boldsymbol{\gamma}\in\mathrm{Adm}(\mu_0,\mu_1)$ such that, up to extract a subsequence, $\boldsymbol{\gamma}^k\rightharpoonup\boldsymbol{\gamma}$. On the other hand, by point (ii) of Proposition 4.1.5 we know that $\boldsymbol{\gamma}^k=f^k\otimes g^k\mathsf{R}$ for all $k\in\mathbb{N}$ for uniquely (up to normalization) determined Borel functions $f^k,g^k:\mathrm{X}\to[0,\infty)$; by the proof of point (iii) of the same proposition we also see that the present assumptions are sufficient to get

$$f^k\le\frac{\rho_0^k}{\alpha\|g^k\|_{L^1(\alpha\mathfrak{m})}},\quad g^k\le\frac{\rho_1^k}{\alpha\|f^k\|_{L^1(\alpha\mathfrak{m})}},\qquad\mathfrak{m}\text{-a.e.}$$

Thus let us normalize (f^k,g^k) in such a way that $\|g^k\|_{L^1(\alpha\mathfrak{m})}=1$. This immediately yields that the sequence (f^k) is uniformly locally bounded: for all $x\in\mathrm{X}$ there exist a neighbourhood U_x of x and $C_x>0$ such that $f^k\le C_x$ $\mathfrak{m}$-a.e. in U_x for all $k\in\mathbb{N}$. Since X is separable, a diagonalization argument ensures the existence of a locally bounded Borel function f and a subsequence (not relabeled) such that, for all $x\in\mathrm{X}$ and U_x as above,

$$f^k\overset{*}{\rightharpoonup}f\quad\text{in }L^\infty(U_x,\mathfrak{m})\quad\text{as }k\to\infty.$$

In order to get the same information on (g^k), we need to show that there exists $c>0$ such that $\|f^k\|_{L^1(\alpha\mathfrak{m})}\ge c$ for all $k\in\mathbb{N}$. If this were not the case, then there would exist a subsequence, not relabeled, such that $\|f^k\|_{L^1(\alpha\mathfrak{m})}\to0$ as $k\to\infty$; given a neighbourhood U where α is away from 0, say $\alpha\ge\delta>0$, this implies $\|f^k\|_{L^1(U,\mathfrak{m})}\to0$. Now take $\phi\in C_b(\mathrm{X})$ non-negative, not identically zero, supported in U and define $\xi(x,y):=\phi(x)\phi(y)$, $\xi\in C_b(\mathrm{X}^2)$. On the one hand

$$\int_{\mathrm{X}^2}\xi\,\mathrm{d}\boldsymbol{\gamma}^k\to\int_{\mathrm{X}^2}\xi\,\mathrm{d}\boldsymbol{\gamma}\qquad\text{as }k\to\infty$$

since $\boldsymbol{\gamma}^k\rightharpoonup\boldsymbol{\gamma}$. On the other hand, from $\|g^k\|_{L^1(\alpha\mathfrak{m})}=1$ we deduce that $\|g^k\|_{L^1(U,\mathfrak{m})}\le\delta^{-1}$; together with the second inequality in (4.2.7), the boundedness of β in U (up to restrict U, if necessary) and the uniform boundedness of (f^k), this fact entails

$$\int_{\mathrm{X}^2}\xi\,\mathrm{d}\boldsymbol{\gamma}^k\le\int\phi f^k\beta\,\mathrm{d}\mathfrak{m}\int\phi g^k\beta\,\mathrm{d}\mathfrak{m}\le M\int_U f^k\,\mathrm{d}\mathfrak{m}\to0\qquad\text{as }k\to\infty$$

whence a contradiction, because $\int_{\mathrm{X}^2}\xi\,\mathrm{d}\boldsymbol{\gamma}>0$. Thus (g^k) is uniformly locally bounded too and there exist a locally bounded Borel function g as well as a subsequence (not relabeled) such that g^k converges to g in the sense already discussed above for f^k.

Now we claim that $\boldsymbol{\gamma} = f \otimes g\mathsf{R}$. To this aim, let U, V be two neighbourhoods where $(f^k), (g^k)$ respectively are uniformly bounded; without loss of generality, we can also assume that α, β are bounded away from 0 and ∞ therein and $\mathfrak{m}(U), \mathfrak{m}(V)$ are finite; furthermore, let $\phi_1, \phi_2 \in C_b(\mathrm{X})$ be non-negative, not identically zero and supported in U, V respectively. Notice that

$$\int_{\mathrm{X}^2} \phi_1 \otimes \phi_2 \, \mathrm{d}\boldsymbol{\gamma}^k = \int \phi_1(x) f^k(x) \Big(\underbrace{\int \phi_2(y) g^k(y) \mathrm{d}\mathsf{R}^x(y)}_{=:\Phi_k(x)} \Big) \mathrm{d}\mathfrak{m}(x) \tag{4.2.9}$$

and $\Phi_k \to \Phi$ in $L^1(U, \mathfrak{m})$, where

$$\Phi(x) := \int \phi_2(y) g(y) \mathrm{d}\mathsf{R}^x(y).$$

In fact, α, β are bounded away from 0 and ∞ in V and from (4.2.7) we see that $\alpha(x)(\alpha\mathfrak{m}) \leq \mathsf{R}^x \leq \beta(x)(\beta\mathfrak{m})$ for $\mathfrak{m}$-a.e. $x \in \mathrm{X}$, so that weak-* convergence in $L^\infty(V, \mathfrak{m})$ and in $L^\infty(V, \mathsf{R}^x)$ are equivalent for $\mathfrak{m}$-a.e. $x \in \mathrm{X}$; since $\phi_2 \in L^1(\mathsf{R}^x)$ for $\mathfrak{m}$-a.e. $x \in X$, we deduce that $\Phi_k \to \Phi$ $\mathfrak{m}$-a.e. in U. Moreover, the functions Φ_k are uniformly bounded in U and $\mathfrak{m}(U) < \infty$, whence the desired L^1-convergence by the dominated convergence theorem. Now it is not difficult to prove the claim, because

$$\begin{aligned}
&\left| \int \phi_1 f^k \Phi_k \, \mathrm{d}\mathfrak{m} - \int \phi_1 f \Phi \, \mathrm{d}\mathfrak{m} \right| \\
&\qquad \leq \left| \int \phi_1 f^k \Phi_k \, \mathrm{d}\mathfrak{m} - \int \phi_1 f^k \Phi \, \mathrm{d}\mathfrak{m} \right| + \left| \int \phi_1 f^k \Phi \, \mathrm{d}\mathfrak{m} - \int \phi_1 f \Phi \, \mathrm{d}\mathfrak{m} \right| \\
&\qquad \leq \|\phi_1\|_{L^\infty(\mathfrak{m})} \sup_{k \in \mathbb{N}} \|f^k\|_{L^\infty(U,\mathfrak{m})} \|\Phi_k - \Phi\|_{L^1(U,\mathfrak{m})} + \left| \int \phi_1 \Phi \big(f^k - f\big) \mathrm{d}\mathfrak{m} \right|
\end{aligned}$$

and $\phi_1 \Phi \in L^1(U, \mathfrak{m})$, so that also the second term on the right-hand side vanishes as $k \to \infty$ because of $f^k \overset{*}{\rightharpoonup} f$ in $L^\infty(U, \mathfrak{m})$. Recalling (4.2.9), this means that

$$\lim_{k \to \infty} \int_{\mathrm{X}^2} \phi_1 \otimes \phi_2 \, \mathrm{d}\boldsymbol{\gamma}^k = \int \phi_1 f \Phi \, \mathrm{d}\mathfrak{m} = \int_{\mathrm{X}^2} \phi_1 \otimes \phi_2 \, \mathrm{d}(f \otimes g\mathsf{R}).$$

Coupling this information with $\boldsymbol{\gamma}^k \rightharpoonup \boldsymbol{\gamma}$ and the arbitrariness of ϕ_1, ϕ_2 the claim and thus the conclusion follow.

As regards (ii), by lower semicontinuity of the entropy and by definition of the entropic cost

$$\mathscr{I}(\mu_0, \mu_1) \leq H(\boldsymbol{\gamma} \,|\, \mathsf{R}) < +\infty.$$

By Proposition 4.1.2 this grants existence and uniqueness of the solution for the Schrödinger problem associated to μ_0, μ_1, R and by point (iii) of Proposition 4.1.5 such a solution must be $\boldsymbol{\gamma}$: indeed, by (A.2.2) we see that μ_0, μ_1 have finite entropy w.r.t. $\mathfrak{m}$, so that all the assumptions of point (iii) are fulfilled.

Finally, point (iii) is nothing but a straightforward consequence of the definition of weak convergence and of the disintegration $\mathbf{R} = \int_{\mathrm{X}^2} \mathsf{R}^{xy} \mathrm{d}\mathsf{R}(x, y)$. $\square$

Chapter 5

On a $\mathsf{RCD}^*(K, N)$ space

As a continuation of the previous one, this chapter is the second part of the user's guide for non-probabilists to the Schrödinger problem. We now move to the RCD framework, where the bounds on both the curvature and the dimension allow a better investigation of the dynamical Schrödinger problem and entail a deeper knowledge of the so called entropic interpolations and of the behaviour of the entropy along them.

In Section 5.1 we fix the setting we work within and introduce the main dynamical objects we shall consider from that moment on throughout the whole manuscript. The idea is the following: given a minimizer of (S_{dyn}), we associate to it a curve of measures which is an interpolation between the marginal constraints and a curve of functions linked to the dual problem, in complete analogy with optimal transport, where such dynamical objects are Wasserstein geodesics and interpolated Kantorovich potentials. Then, inspired by the formal considerations of [81], we endeavour the time evolution of these measures and functions by indicating the 'PDEs' they solve.

On the contrary the results of Section 5.2 are completely new, even in the smooth setting, and strongly rely on the functional inequalities investigated in Chapter 2, namely Hamilton's gradient estimate and Li-Yau inequality. Main achievements of this part are:

- a uniform Gaussian control of the densities along entropic interpolations;
- local equi-Lipschitz continuity of the Schrödinger potentials with explicit control of the local Lipschitz constant.

As regards Section 5.3, we show that Schrödinger potentials satisfy Bochner inequality in a suitable sense and this technical result is essential for the computation of the first and second derivatives of the relative entropy along entropic interpolations. Such computation had already been performed in [76] by formal arguments; here we verify that the same procedure is fully justifiable also in the case of $\mathsf{RCD}^*(K, N)$ spaces. As a byproduct we obtain crucial uniform estimates, in particular a uniform weighted L^2 control of the Hessian of Schrödinger potentials (see Lemma 5.3.4), and the vanishing of certain quantities (see Lemma 5.3.5): all these results will play a key role in Chapter 6 and Chapter 7.

Aim of Section 5.4 is to establish a further parallelism between optimal transport and Schrödinger problem, inspired by the Benamou-Brenier formula for the Wasserstein distance. Our contribution is an analogous dynamical representation for the entropic cost in compact $\mathsf{RCD}^*(K, N)$ spaces, which was still missing in the non-smooth framework; the proof relies on purely analytical arguments.

We conclude with a physical perspective: in Section 5.5 all the interpolating quantities we have introduced at the beginning of the chapter are reinterpreted as physical objects with a precise meaning, thus adopting Nelson's stochastic mechanics point of view (for a detailed dissertation on the subject, we suggest the monograph [102]).

5.1 The setting

The properties we stated in Proposition 4.1.5 are valid in the very general framework of Polish spaces. We shall now restate them in the form we shall need in the context of RCD spaces.

Recall that on a finite-dimensional $\mathsf{RCD}^*(K,N)$ space $(\mathrm{X},\mathsf{d},\mathfrak{m})$ the reference measure $\mathfrak{m}$ satisfies (1.2.7) and because of (1.2.5) M can be any positive constant, so that we can choose $W=\mathsf{d}^2(\cdot,\bar{x})$ for any $\bar{x}\in\mathrm{X}$ in (4.1.1). Setting $z:=\int e^{-\mathsf{d}^2(\cdot,\bar{x})}\mathrm{d}\mathfrak{m}$ and

$$\tilde{\mathfrak{m}}:=z^{-1}e^{-\mathsf{d}^2(\cdot,\bar{x})}\mathfrak{m}, \tag{5.1.1}$$

definition (4.1.1) then becomes

$$H(\mu\,|\,\mathfrak{m})=H(\mu\,|\,\tilde{\mathfrak{m}})-\int\mathsf{d}^2(\cdot,\bar{x})\mathrm{d}\mu-\log z \tag{5.1.2}$$

and this shows that $H(\,\cdot\,|\,\mathfrak{m})$ is well defined on $\mathscr{P}_2(\mathrm{X})$. Let us also remind that on RCD spaces there is a well defined heat kernel $\mathsf{r}_\varepsilon[x](y)$ (see (1.2.8), (1.2.9a), (1.2.9b) and (1.2.9c)). The choice of working with $\mathsf{r}_{\varepsilon/2}$ in the theorem below is convenient for the computations we will do later on.

Theorem 5.1.1. *Let $(\mathrm{X},\mathsf{d},\mathfrak{m})$ be a $\mathsf{RCD}^*(K,N)$ space with $K\in\mathbb{R}$ and $N\in[1,\infty)$ endowed with a non-negative Radon measure $\mathfrak{m}$. For $\varepsilon>0$ define the measure $\mathsf{R}^{\varepsilon/2}$ on X^2 as*

$$\mathrm{d}\mathsf{R}^{\varepsilon/2}(x,y):=\mathsf{r}_{\varepsilon/2}[x](y)\,\mathrm{d}\mathfrak{m}(x)\,\mathrm{d}\mathfrak{m}(y).$$

Also, let $\mu_0,\mu_1\in\mathscr{P}_2(\mathrm{X})$ be Borel probability measures with bounded densities.

Then there exist and are uniquely $\mathfrak{m}$-a.e. determined (up to multiplicative constants) two Borel non-negative functions $f^\varepsilon,g^\varepsilon:\mathrm{X}\to[0,\infty)$ such that $f^\varepsilon\otimes g^\varepsilon\mathsf{R}^{\varepsilon/2}$ is a transport plan from μ_0 to μ_1. In addition, $f^\varepsilon,g^\varepsilon$ belong to $L^\infty_{loc}(\mathrm{X})$ and their supports are included in $\mathrm{supp}(\mu_0)$ and $\mathrm{supp}(\mu_1)$ respectively.

If μ_0,μ_1 also have bounded supports, then $f^\varepsilon,g^\varepsilon$ belong to $L^\infty(\mathrm{X})$.

Finally, if the densities of μ_0,μ_1 also belong to $\mathrm{Test}(\mathrm{X})$, then $f^\varepsilon,g^\varepsilon\in\mathrm{Test}(\mathrm{X})$ as well.

proof It follows from point (*iii*) of Proposition 4.1.5 and the Gaussian estimates: indeed, start observing that $\mathsf{R}_0=\mathsf{R}_1=\mathfrak{m}$ and if we set $B:=\mathsf{d}^2(\cdot,\bar{x})$ with $\bar{x}\in\mathrm{X}$ arbitrarily fixed, then the second and third in (4.1.3) are authomatically satisfied; for the first one

$$\int_{\mathrm{X}^2}e^{-B\oplus B}\mathrm{d}\mathsf{R}=\int\Big(e^{-\mathsf{d}^2(y,\bar{x})}\mathrm{d}\mathsf{R}^x(y)\Big)e^{-\mathsf{d}^2(x,\bar{x})}\mathrm{d}\mathfrak{m}(x),$$

thus it is now sufficient to use the fact that $e^{-\mathsf{d}^2(y,\bar{x})}\le 1$, R^x is a probability measure and (1.2.7) to conclude. Furthermore from the left-hand side of (1.2.11) together with the symmetry of the heat kernel we know that

$$\mathsf{r}_{\varepsilon/2}[x](y)\ge\frac{C}{\min\{\mathfrak{m}_x,\mathfrak{m}_y\}}\exp\Big(-\frac{\mathsf{d}^2(x,y)}{\varepsilon}\Big)$$

where $\mathfrak{m}_x := \mathfrak{m}(B_{\sqrt{\varepsilon/2}}(x))$ and analogously for $\mathfrak{m}_y$; fixing $\bar{x} \in \mathrm{X}$ and using the triangle inequality we get

$$\mathsf{r}_{\varepsilon/2}[x](y) \geq \underbrace{\sqrt{\frac{C}{\mathfrak{m}_x}} \exp\Big(-\frac{2\mathsf{d}^2(x,\bar{x})}{\varepsilon}\Big)}_{:=\alpha(x)} \sqrt{\frac{C}{\mathfrak{m}_y}} \exp\Big(-\frac{2\mathsf{d}^2(y,\bar{x})}{\varepsilon}\Big),$$

thus obtaining α satisfying (4.1.5). It is also not difficult to see that $(\log\alpha)^- \in L^1(\mu_0)$ and $(\log\alpha)^- \in L^1(\mu_1)$, because

$$\big(\log\alpha(x)\big)^- \leq \frac{1}{2}(\log C)^- + \frac{1}{2}(-\log\mathfrak{m}_x)^- + \frac{2}{\varepsilon}\mathsf{d}^2(x,\bar{x})$$

and (1.2.5) allows to handle the second term: in fact, for any $\bar{x} \in \mathrm{X}$ a priori fixed we have $\mathfrak{m}_x \leq \mathfrak{m}(B_{\mathsf{d}(x,\bar{x})+\sqrt{\varepsilon/2}}(\bar{x}))$ and $\mathfrak{m}(B_{\mathsf{d}(x,\bar{x})+\sqrt{\varepsilon/2}}(\bar{x}))$ can be estimated by means of (1.2.5). Since all the assumptions required by Proposition 4.1.5 (*iii*) are verified, the conclusion follows.

For the case of compactly supported marginals, still the Gaussian estimates (1.2.11) on the heat kernel grant that there are constants $0 < c_\varepsilon \leq C_\varepsilon < +\infty$ such that

$$c_\varepsilon \mathfrak{m}\otimes\mathfrak{m} \leq \mathsf{R}^\varepsilon \leq C_\varepsilon \mathfrak{m}\otimes\mathfrak{m}$$

in $\mathrm{supp}(\mu_0)\times\mathrm{supp}(\mu_1)$ and thus point (*iv*) of Proposition 4.1.5 applies.

For the last part of the statement, notice that thanks to the representation formula (1.2.9c), the fact that $\pi^0_*(f^\varepsilon \otimes g^\varepsilon \mathsf{R}^{\varepsilon/2}) = \rho_0\mathfrak{m}$ reads as

$$f^\varepsilon \mathsf{h}_{\varepsilon/2}(g^\varepsilon) = \rho_0.$$

Now notice that by (1.2.15) we have $\mathsf{h}_{\varepsilon/2}(g^\varepsilon) \in \mathrm{Test}(\mathrm{X})$ and by (1.2.11) $\mathsf{h}_{\varepsilon/2}(g^\varepsilon)$ is locally bounded away from 0, so that $\frac{1}{\mathsf{h}_{\varepsilon/2}(g^\varepsilon)} \in \mathrm{Test}_{loc}(\mathrm{X})$. Since $\mathrm{Test}_{loc}(\mathrm{X})$ is an algebra we conclude that $f^\varepsilon = \frac{\rho_0}{\mathsf{h}_\varepsilon(g^\varepsilon)} \in \mathrm{Test}_{loc}(\mathrm{X})$ and since ρ_0 has compact support $f^\varepsilon \in \mathrm{Test}(\mathrm{X})$ *tout court*. The same applies to g^ε. □

We also present an analogous result in the compact setting, since it will be useful for some applications and remarks pointed out in Section 6.3.

Theorem 5.1.2. *Let* $(\mathrm{X},\mathsf{d},\mathfrak{m})$ *be a compact* $\mathsf{RCD}^*(K,N)$ *space with* $K\in\mathbb{R}$, $N\in[1,\infty)$ *and* $\mathfrak{m}\in\mathscr{P}(\mathrm{X})$. *For* $\varepsilon>0$ *define* $\mathsf{R}^{\varepsilon/2}\in\mathscr{P}(\mathrm{X}^2)$ *as in Theorem 5.1.1. Also, let* $\mu_0,\mu_1\in\mathscr{P}(\mathrm{X})$ *be Borel probability measures with bounded densities.*

Then there exist and are uniquely $\mathfrak{m}$*-a.e. determined (up to multiplicative constants) two Borel non-negative functions* $f^\varepsilon, g^\varepsilon : \mathrm{X}\to[0,\infty)$ *such that* $f^\varepsilon\otimes g^\varepsilon\mathsf{R}^{\varepsilon/2}$ *is a transport plan from* μ_0 *to* μ_1. *In addition,* $f^\varepsilon, g^\varepsilon$ *belong to* $L^\infty(\mathfrak{m})$.

Moreover, if the densities of μ_0,μ_1 *belong to* $\mathrm{Test}^\infty_{>0}(\mathrm{X})$, *then* $f^\varepsilon, g^\varepsilon\in\mathrm{Test}^\infty_{>0}(\mathrm{X})$ *as well.*

proof For the first part of the statement no comments are required. For the second one, we argue as for the last part of Theorem 5.1.1: indeed, $\pi^0_*(f^\varepsilon\otimes g^\varepsilon\mathsf{R}^{\varepsilon/2}) = \rho_0\mathfrak{m}$ reads as $f^\varepsilon\mathsf{h}_{\varepsilon/2}(g^\varepsilon) = \rho_0$; thus by (2.1.1) we have $\mathsf{h}_{\varepsilon/2}(g^\varepsilon)\in\mathrm{Test}^\infty_{>0}(\mathrm{X})$, and therefore from (2.1.7) applied with $\varphi(z) := z^{-1}$ we deduce that $\frac{1}{\mathsf{h}_{\varepsilon/2}(g^\varepsilon)}\in\mathrm{Test}^\infty_{>0}(\mathrm{X})$. Since $\mathrm{Test}^\infty(\mathrm{X})$ is an algebra we conclude that $f^\varepsilon = \frac{\rho_0}{\mathsf{h}_\varepsilon(g^\varepsilon)}\in\mathrm{Test}^\infty_{>0}(\mathrm{X})$. The same applies to g^ε. □

With this said, let us fix once for all the assumptions and notations which we shall use from now on.

Setting 5.1.3. $(\mathrm{X},\mathrm{d},\mathfrak{m})$ is a $\mathsf{RCD}^*(K,N)$ space with $K\in\mathbb{R}$ and $N\in[1,\infty)$, $\mathfrak{m}$ is a Borel non-negative measure and $\mu_0=\rho_0\mathfrak{m}$, $\mu_1=\rho_1\mathfrak{m}$ are two absolutely continuous Borel probability measures with bounded densities and supports.

For any $\varepsilon>0$ we consider the couple $(f^\varepsilon,g^\varepsilon)$ given by Theorem 5.1.1 normalized in such a way that

$$\int \log(\mathsf{h}_{\frac{\varepsilon}{2}}f^\varepsilon)\rho_1\,\mathrm{d}\mathfrak{m}=0,$$

then we set $\rho_0^\varepsilon:=\rho_0$, $\rho_1^\varepsilon:=\rho_1$, $\mu_0^\varepsilon:=\mu_0$, $\mu_1^\varepsilon:=\mu_1$ and

$$\begin{cases} f_t^\varepsilon:=\mathsf{h}_{\varepsilon t/2}f^\varepsilon \\ \varphi_t^\varepsilon:=\varepsilon\log f_t^\varepsilon \\ \text{for } t\in(0,1] \end{cases}\qquad \begin{cases} g_t^\varepsilon:=\mathsf{h}_{\varepsilon(1-t)/2}g^\varepsilon \\ \psi_t^\varepsilon:=\varepsilon\log g_t^\varepsilon \\ \text{for } t\in[0,1) \end{cases}\qquad \begin{cases} \rho_t^\varepsilon:=f_t^\varepsilon g_t^\varepsilon \\ \mu_t^\varepsilon:=\rho_t^\varepsilon\mathfrak{m} \\ \vartheta_t^\varepsilon:=\frac12(\psi_t^\varepsilon-\varphi_t^\varepsilon) \\ \text{for } t\in(0,1) \end{cases}$$

In order to investigate the time behaviour of the functions just defined, besides the notions of local calculus introduced in Section 2.2 we also need the corresponding weighted ones. The weight we will always consider is e^{-V} with $V:\mathrm{X}\to[0,\infty)$ a continuous map such that $e^{-V}\mathfrak{m}$ is a finite measure; thanks to (1.2.5) and (1.2.7) $V=M\mathsf{d}^2(\cdot,\bar x)$ is an admissible choice for any $M>0$ and $\bar x\in\mathrm{X}$. For $L^2(e^{-V}\mathfrak{m})$ no comments are required. The weighted Sobolev space and the weighted domains of the (local) divergence and Laplacian are denoted and defined as follows:

$$\begin{aligned} W^{1,2}(\mathrm{X},e^{-V}\mathfrak{m}) &:= \{f\in W^{1,2}_{loc}(\mathrm{X})\,:\,f,|Df|\in L^2(e^{-V}\mathfrak{m})\} \\ D(\mathrm{div},e^{-V}\mathfrak{m}) &:= \{W\in D_{loc}(\mathrm{div})\,:\,\mathrm{div}(W)\in L^2(e^{-V}\mathfrak{m})\} \\ D(\Delta,e^{-V}\mathfrak{m}) &:= \{f\in D_{loc}(\Delta)\,:\,\Delta f\in L^2(e^{-V}\mathfrak{m})\} \end{aligned}$$

where $|Df|,\mathrm{div},\Delta$ are the local objects already introduced in Section 2.2. In order to avoid possible confusion, it is worth stressing that the definitions above are not the ones built over the metric measure space $(\mathrm{X},\mathsf{d},e^{-V}\mathfrak{m})$: we always work within $(\mathrm{X},\mathsf{d},\mathfrak{m})$ and we ask $|Df|,\mathrm{div}(v),\Delta f$ to be 2-integrable w.r.t. $e^{-V}\mathfrak{m}$, the advantage being the fact that $L^2(e^{-V}\mathfrak{m})$ and $W^{1,2}(\mathrm{X},e^{-V}\mathfrak{m})$ are Hilbert spaces, unlike $L^2_{loc}(\mathfrak{m})$ and $W^{1,2}_{loc}(\mathrm{X})$.

The following proposition collects the basic properties of the functions defined in Setting 5.1.3 and the respective 'PDEs' solved. The proof is intentionally naïf and rough: we argue by direct computations and only rely on the Gaussian estimates (1.2.11) for the heat kernel. We choose this approach because on the one hand it is sufficient for our purposes, at least for the moment, and on the other hand it makes clearer the importance of the results carried out in Chapter 2. A comparison between the forthcoming estimates (5.1.3), (5.1.4) and those of Section 5.2 and Section 5.3 is a bright evidence.

Proposition 5.1.4. *With the same assumptions and notation as in Setting 5.1.3, the following holds.*

All the functions are well defined and for any $\varepsilon>0$:

a) $f_t^\varepsilon, g_t^\varepsilon, \rho_t^\varepsilon$ *belong to* $\mathrm{Test}(\mathrm{X})$ *for all* $t \in I$*, where* I *is the respective domain of definition (for* (ρ_t^ε) *we pick* $I = (0,1)$*);*

b) $\varphi_t^\varepsilon, \psi_t^\varepsilon, \vartheta_t^\varepsilon$ *belong to* $\mathrm{Test}_{loc}(\mathrm{X})$ *for all* $t \in I$*, where* I *is the respective domain of definition.*

For any $\varepsilon > 0$ *and* $C \subset I$ *compact all the curves* $(f_t^\varepsilon), (g_t^\varepsilon), (\rho_t^\varepsilon)$ *belong to* $AC(C, W^{1,2}(\mathrm{X}))$ *and* $(\varphi_t^\varepsilon), (\psi_t^\varepsilon), (\vartheta_t^\varepsilon)$ *to* $AC(C, W^{1,2}(\mathrm{X}, e^{-V}\mathfrak{m}))$*, where* I *is the respective domain of definition (for* (ρ_t^ε) *we pick* $I = (0,1)$*) and* $V = M\mathsf{d}^2(\cdot, \bar{x})$ *with* $\bar{x} \in \mathrm{X}$ *and* $M = M(K, N, \rho_0, \rho_1, C, \bar{x}) > 0$*; their time derivatives are given by the following expressions for a.e.* $t \in [0,1]$*:*

$$\frac{\mathrm{d}}{\mathrm{d}t} f_t^\varepsilon = \frac{\varepsilon}{2}\Delta f_t^\varepsilon \qquad \frac{\mathrm{d}}{\mathrm{d}t} g_t^\varepsilon = -\frac{\varepsilon}{2}\Delta g_t^\varepsilon$$
$$\frac{\mathrm{d}}{\mathrm{d}t}\varphi_t^\varepsilon = \frac{1}{2}|\nabla\varphi_t^\varepsilon|^2 + \frac{\varepsilon}{2}\Delta\varphi_t^\varepsilon \qquad -\frac{\mathrm{d}}{\mathrm{d}t}\psi_t^\varepsilon = \frac{1}{2}|\nabla\psi_t^\varepsilon|^2 + \frac{\varepsilon}{2}\Delta\psi_t^\varepsilon$$
$$\frac{\mathrm{d}}{\mathrm{d}t}\rho_t^\varepsilon + \mathrm{div}(\rho_t^\varepsilon \nabla\vartheta_t^\varepsilon) = 0 \qquad \frac{\mathrm{d}}{\mathrm{d}t}\vartheta_t^\varepsilon + \frac{|\nabla\vartheta_t^\varepsilon|^2}{2} = -\frac{\varepsilon^2}{8}\Big(2\Delta\log\rho_t^\varepsilon + |\nabla\log\rho_t^\varepsilon|^2\Big).$$

Moreover, for every $\varepsilon > 0$ *we have:*

i)

$$\sup_{t\in C} \|h_t^\varepsilon\|_{L^\infty(\mathrm{X})} + \mathrm{Lip}(h_t^\varepsilon) + \|\Delta h_t^\varepsilon\|_{W^{1,2}(\mathrm{X})} < \infty \tag{5.1.3}$$

if (h_t^ε) *is equal to any of* $(f_t^\varepsilon), (g_t^\varepsilon), (\rho_t^\varepsilon)$ *and*

$$\sup_{t\in C} \|e^{-V}h_t^\varepsilon\|_{L^\infty(\mathrm{X})} + \|e^{-V}\mathrm{lip}(h_t^\varepsilon)\|_{L^\infty(\mathrm{X})} + \|\Delta h_t^\varepsilon\|_{W^{1,2}(\mathrm{X}, e^{-V}\mathfrak{m})} < \infty \tag{5.1.4}$$

if (h_t^ε) *is equal to any of* $(\varphi_t^\varepsilon), (\psi_t^\varepsilon), (\vartheta_t^\varepsilon)$*; in both cases,* C *is a compact subset of the respective domain of definition (for* (ρ_t^ε) *we pick* $I = (0,1)$*),*

ii) $\mu_t^\varepsilon \in \mathscr{P}_2(\mathrm{X})$ *for every* $t \in [0,1]$ *and* $(\rho_t^\varepsilon) \in C([0,1], L^2(\mathrm{X}))$*,*

iii) we have $f_t^\varepsilon \to f^\varepsilon$ *and* $g_t^\varepsilon \to g^\varepsilon$ *in* $L^2(\mathrm{X})$ *as* $t \downarrow 0$ *and* $t \uparrow 1$ *respectively.*

proof Recalling (1.2.15) we see that $f_{t_0}^\varepsilon \in \mathrm{Test}(\mathrm{X})$ for any $t_0 > 0$. Then the maximum principle for the heat flow, the fact that it is a contraction in $W^{1,2}(\mathrm{X})$ and the Bakry-Émery gradient estimates (1.2.13) together with the Sobolev-to-Lipschitz property grant that (5.1.3) holds for (f_t^ε). The same arguments apply to (g_t^ε). Then the bound (5.1.3) for (ρ_t^ε) follows from the Leibniz rules for the gradient and Laplacian and thanks to this bound we see that the curves $(f_t^\varepsilon), (g_t^\varepsilon), (\rho_t^\varepsilon)$ belong to $AC(C, L^2(\mathrm{X}))$. As regards $\varphi_t^\varepsilon, \psi_t^\varepsilon, \vartheta_t^\varepsilon$, let us fix $\varepsilon > 0$: by Proposition 2.2.5 *b*) is true. Then use the representation formula (1.2.9c), the Gaussian estimates (1.2.11) and the fact that ρ_0 and f^ε have the same support to get

$$\begin{aligned} &\frac{C_1}{V_{\varepsilon t/2}}\|f^\varepsilon\|_{L^1(\mathfrak{m})}\exp\Big(-\frac{C_2\mathsf{d}^2(\cdot,\bar{x})}{\varepsilon t} - \frac{C_3}{\varepsilon t}\Big) \le \\ &\qquad \le f_t^\varepsilon \le \frac{C_4}{v_{\varepsilon t/2}}\|f^\varepsilon\|_{L^1(\mathfrak{m})}\exp\Big(-\frac{C_5\mathsf{d}^2(\cdot,\bar{x})}{\varepsilon t} + \frac{C_6}{\varepsilon t}\Big) \end{aligned} \tag{5.1.5}$$

for all $t \in (0,1]$, where $\bar{x} \in \mathrm{X}$ is arbitrary, all the constants $C_1, ..., C_6$ are positive and only depend on K, N, $\mathrm{supp}(\rho_0)$, $\mathrm{supp}(\rho_1)$, $\bar{x}$ and

$$v_s := \inf_{y\in\mathrm{supp}(\rho_0)} \mathfrak{m}(B_{\sqrt{s}}(y)) \qquad V_s := \sup_{y\in\mathrm{supp}(\rho_0)} \mathfrak{m}(B_{\sqrt{s}}(y)), \tag{5.1.6}$$

paying attention to the fact that for any $s>0$ it holds $0<v_s\le V_s<\infty$. This two-sided bound implies that $e^{-V}\varphi_t^\varepsilon\in L^\infty(\mathfrak{m})$ and $\varphi_t^\varepsilon\in L^2(e^{-V}\mathfrak{m})$ with $V=M\mathsf{d}^2(\cdot,\bar x)$ for any $\bar x\in\mathrm{X}$ and $M>0$. Moreover, taking into account (2.2.9) and (2.2.10) and the fact that $|\nabla f_t^\varepsilon|^2\in W^{1,2}(\mathrm{X})$ by (1.2.16), the same bound implies that $e^{-V}|\nabla\varphi_t^\varepsilon|\in L^\infty(\mathfrak{m})$, $|\nabla\varphi_t^\varepsilon|\in L^2(e^{-V}\mathfrak{m})$, $\varphi_t^\varepsilon\in D(\Delta,e^{-V}\mathfrak{m})$ and $\Delta\varphi_t^\varepsilon\in W^{1,2}(\mathrm{X},e^{-V}\mathfrak{m})$ with V as above and $M>\frac{3C_2}{\varepsilon t^-}$, C_2 being the constant appearing in (5.1.5). In fact

$$\begin{aligned}&\nabla\varphi_t^\varepsilon=\varepsilon\frac{\nabla f_t^\varepsilon}{f_t^\varepsilon},\qquad \Delta\varphi_t^\varepsilon=\varepsilon\Big(\frac{\Delta f_t^\varepsilon}{f_t^\varepsilon}-\frac{|\nabla f_t^\varepsilon|^2}{(f_t^\varepsilon)^2}\Big),\\ &|\nabla\Delta\varphi_t^\varepsilon|\le\varepsilon\Big(\frac{|\nabla\Delta f_t^\varepsilon|}{f_t^\varepsilon}+\frac{\Delta f_t^\varepsilon|\nabla f_t^\varepsilon|}{(f_t^\varepsilon)^2}+\frac{|\nabla|\nabla f_t^\varepsilon|^2|}{(f_t^\varepsilon)^2}+\frac{2|\nabla f_t^\varepsilon|^3}{(f_t^\varepsilon)^3}\Big),\end{aligned}\tag{5.1.7}$$

whence the claim. Thus all the norms appearing in (5.1.4) exist and the bound makes sense. In order to prove it, just look at (5.1.7), use (5.1.3) for (f_t^ε), notice that (1.2.16) gives us a locally (in t) uniform control on the $L^2(\mathfrak{m})$-norm of $|\nabla|\nabla f_t^\varepsilon|^2|$ while (5.1.5) entails a locally (in t) uniform two-sided bound for (f_t^ε), i.e.

$$\frac{C_1}{V_{\varepsilon t^+/2}}\|f^\varepsilon\|_{L^1(\mathfrak{m})}\exp\Big(-\frac{C_2\mathsf{d}^2(\cdot,\bar x)}{\varepsilon t^-}-\frac{C_3}{\varepsilon t^-}\Big)\le f_t^\varepsilon\le\frac{C_4}{v_{\varepsilon t^-/2}}\|f^\varepsilon\|_{L^1(\mathfrak{m})}\exp\Big(-\frac{C_5\mathsf{d}^2(\cdot,\bar x)}{\varepsilon t^+}+\frac{C_6}{\varepsilon t^-}\Big)$$

where $t^-:=\inf_C t$ and $t^+:=\sup_C t$. For the same reason we also deduce that (φ_t^ε) belongs to $AC(C,L^2(e^{-V}\mathfrak{m}))$ with V as above. The same arguments apply to (g_t^ε) and thus we get the same conclusions for (ψ_t^ε) and $(\vartheta_t^\varepsilon)$.

The equations for $\frac{\mathrm{d}}{\mathrm{d}t}\varphi_t^\varepsilon$ and $\frac{\mathrm{d}}{\mathrm{d}t}\psi_t^\varepsilon$ are easily derived, for $\frac{\mathrm{d}}{\mathrm{d}t}\rho_t^\varepsilon$ we notice that $\varepsilon\log\rho_t^\varepsilon=\varphi_t^\varepsilon+\psi_t^\varepsilon$ and thus

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\rho_t^\varepsilon&=\rho_t^\varepsilon\frac{\mathrm{d}}{\mathrm{d}t}\log\rho_t^\varepsilon=\rho_t^\varepsilon\frac{1}{\varepsilon}\Big(\frac{|\nabla\varphi_t^\varepsilon|^2}{2}-\frac{|\nabla\psi_t^\varepsilon|^2}{2}+\frac{\varepsilon}{2}\Delta\varphi_t^\varepsilon-\frac{\varepsilon}{2}\Delta\psi_t^\varepsilon\Big)\\&=\rho_t^\varepsilon\Big(-\langle\nabla\vartheta_t^\varepsilon,\nabla\log\rho_t^\varepsilon\rangle-\Delta\vartheta_t^\varepsilon\Big)=-\langle\nabla\vartheta_t^\varepsilon,\nabla\rho_t^\varepsilon\rangle-\rho_t^\varepsilon\Delta\vartheta_t^\varepsilon=-\mathrm{div}(\rho_t^\varepsilon\nabla\vartheta_t^\varepsilon)\end{aligned}$$

and for $\frac{\mathrm{d}}{\mathrm{d}t}\vartheta_t^\varepsilon$ we observe that

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\vartheta_t^\varepsilon+\frac{|\nabla\vartheta_t^\varepsilon|^2}{2}&=-\frac{|\nabla\psi_t^\varepsilon|^2}{4}-\frac{\varepsilon}{4}\Delta\psi_t^\varepsilon-\frac{|\nabla\varphi_t^\varepsilon|^2}{4}-\frac{\varepsilon}{4}\Delta\varphi_t^\varepsilon+\frac{|\nabla\psi_t^\varepsilon|^2}{8}+\frac{|\nabla\varphi_t^\varepsilon|^2}{8}-\frac{\langle\nabla\psi_t^\varepsilon,\nabla\varphi_t^\varepsilon\rangle}{4}\\&=-\frac{\varepsilon^2}{4}\Delta\log\rho_t^\varepsilon-\frac{1}{8}\Big(|\nabla\psi_t^\varepsilon|^2+|\nabla\varphi_t^\varepsilon|^2+2\langle\nabla\varphi_t^\varepsilon,\nabla\psi_t^\varepsilon\rangle\Big)\\&=-\frac{\varepsilon^2}{8}\Big(2\Delta\log\rho_t^\varepsilon+|\nabla\log\rho_t^\varepsilon|^2\Big).\end{aligned}$$

The fact that (ρ_t^ε) is absolutely continuous with values in $W^{1,2}(\mathrm{X})$ then follows by rewriting its derivative as

$$\frac{\mathrm{d}}{\mathrm{d}t}\rho_t^\varepsilon=\frac{\varepsilon}{2}g_t^\varepsilon\Delta f_t^\varepsilon-\frac{\varepsilon}{2}f_t^\varepsilon\Delta g_t^\varepsilon$$

and using the bound (5.1.3) for $f_t^\varepsilon,g_t^\varepsilon$. The absolute continuity of $(\varphi_t^\varepsilon),(\psi_t^\varepsilon),(\vartheta_t^\varepsilon)$ when seen with values in $W^{1,2}(\mathrm{X},e^{-V}\mathfrak{m})$ is a direct consequence of the expressions for their derivatives, rewritten as

$$\frac{\mathrm{d}}{\mathrm{d}t}\varphi_t^\varepsilon=\frac{\varepsilon}{2}\frac{\Delta f_t^\varepsilon}{f_t^\varepsilon}\qquad\frac{\mathrm{d}}{\mathrm{d}t}\psi_t^\varepsilon=-\frac{\varepsilon}{2}\frac{\Delta g_t^\varepsilon}{g_t^\varepsilon}\qquad\frac{\mathrm{d}}{\mathrm{d}t}\vartheta_t^\varepsilon=-\frac{\varepsilon}{4}\frac{\Delta f_t^\varepsilon}{f_t^\varepsilon}-\frac{\varepsilon}{4}\frac{\Delta g_t^\varepsilon}{g_t^\varepsilon},$$

and the bounds (5.1.3), (5.1.5) for $f_t^\varepsilon, g_t^\varepsilon$ in conjunction with (1.2.16).

It is clear that $\rho_t^\varepsilon \geq 0$ for every ε, t, hence the identity

$$\int \rho_t^\varepsilon \, \mathrm{d}\mathfrak{m} = \int \mathsf{h}_{\varepsilon t/2} f^\varepsilon \mathsf{h}_{\varepsilon(1-t)/2} g^\varepsilon \, \mathrm{d}\mathfrak{m} = \int f^\varepsilon \mathsf{h}_{\varepsilon/2} g^\varepsilon \, \mathrm{d}\mathfrak{m} = \int \rho_0^\varepsilon \, \mathrm{d}\mathfrak{m} = 1$$

shows that $\mu_t^\varepsilon \in \mathscr{P}(\mathrm{X})$. The fact that μ_t^ε has finite second moment then follows from (5.1.5) written for $f_t^\varepsilon, g_t^\varepsilon$ for $t \in (0,1)$, while it is trivial for $t = 0, 1$.

Due to the continuity of $[0,\infty) \ni t \mapsto \mathsf{h}_t h \in L^2(\mathrm{X})$ for every $h \in L^2(\mathrm{X})$, the claimed continuities in L^2 for the f's and g's follow. Then for what concerns the ρ's, we need to check that for every $\varepsilon > 0$ we have

$$\rho_0 = f^\varepsilon \mathsf{h}_{\varepsilon/2} g^\varepsilon \qquad\qquad \rho_1 = g^\varepsilon \mathsf{h}_{\varepsilon/2} f^\varepsilon. \tag{5.1.8}$$

As already noticed in the proof of Theorem 5.1.1, these are equivalent to the fact that $f^\varepsilon \otimes g^\varepsilon \, \mathsf{R}^{\varepsilon/2}$ is a transport plan from μ_0 to μ_1; hence, (5.1.8) holds by the very choice of $(f^\varepsilon, g^\varepsilon)$ made. □

Using the terminology adopted in the literature (see [81]) we shall refer to:

- φ_t^ε and ψ_t^ε as Schrödinger potentials, in connection with Kantorovich ones;
- $(\mu_t^\varepsilon)_{t\in[0,1]}$ as entropic interpolation, in analogy with displacement one.

However the motivation behind the definition of (μ_t^ε) is not clear yet. Is there any link with the dynamical Schrödinger problem? The answer is affirmative and relies on the classical mechanics version of Born's formula (3.1.9) anticipated in Chapter 3.

Within the RCD setting and relying on Theorem 5.1.1 and Proposition 5.1.4 we can prove it in a rigorous way, but first let us briefly review the notion of 'Brownian motion' (or more precisely diffusion process) in a $\mathsf{RCD}(K,\infty)$ space and interpret in probabilistic terms the analytical properties of the heat kernel presented in Section 1.2. By [7] we know that there exists a unique (in law) Markov family $\{\mathbf{R}^x\}_{x\in\mathrm{supp}(\mathfrak{m})}$ of probability measures on $C([0,\infty),\mathrm{X})$ such that

$$\mathsf{h}_t f(x) = \int f(\gamma_t) \mathrm{d}\mathbf{R}^x(\gamma) \qquad \forall t \geq 0$$

for all $f \in C_b(\mathrm{X})$ and $\mathfrak{m}$-a.e. $x \in \mathrm{supp}(\mathfrak{m})$. Looking at (1.2.9c) this tells us that $(\mathrm{e}_t)_*\mathbf{R}^x = \mathsf{r}_t[x]\mathfrak{m}$ and if we set $\mathbf{R} := \int \mathbf{R}^x \mathrm{d}\mathfrak{m}(x)$, then the canonical process $\mathbf{Z} = (Z_t)_{t\geq 0}$, $Z_t(\gamma) := \gamma_t$, is the unique (in law) Markov process concentrated on $C([0,\infty),\mathrm{X})$ with $Z_0 \sim \mathfrak{m}$ and transition probabilities given by the heat kernel, i.e.

$$\mathbf{R}(Z_{s+t} \in A \,|\, Z_s = x) = \int_A \mathsf{r}_t[x](y) \mathrm{d}\mathfrak{m}(y) \qquad \forall s,t \geq 0,\ A \in \mathscr{B}(\mathrm{X})$$

for $\mathfrak{m}$-a.e. $x \in \mathrm{supp}(\mathfrak{m})$: such a process is called **Brownian motion**. This allows to rewrite the representation formula (1.2.9c) in probabilistic terms, namely

$$\mathsf{h}_t f(x) = \mathsf{E}_{\mathbf{R}^x} f(Z_t) = \mathsf{E}_{\mathbf{R}}[f(Z_t) \,|\, Z_0 = x], \tag{5.1.9}$$

see also Appendix A.2 for an overview on the probabilistic notations here adopted. Now let us show the Euclidean analogue of Born's formula; since the Schrödinger problem is formulated on $C([0,1],\mathrm{X})$ instead of $C([0,\infty),\mathrm{X})$, from now on $\mathbf{R}$ (resp. $\mathbf{R}^{\varepsilon/2}$) shall denote the law of

$(Z_t)_{t\in[0,1]}$ (resp. $(Z_{\varepsilon t/2})_{t\in[0,1]}$) instead of $\mathbf{Z}$'s one: this does not change the fact that $(\mathrm{e}_t)_*\mathbf{R}=\mathfrak{m}$ and $(\mathrm{e}_t)_*\mathbf{R}^x=\mathsf{r}_t[x]\mathfrak{m}$ for all $t\in(0,1]$ (resp. $(\mathrm{e}_t)_*\mathbf{R}^{\varepsilon/2}=\mathfrak{m}$ and $(\mathrm{e}_t)_*(\mathbf{R}^{\varepsilon/2})^x=\mathsf{r}_{\varepsilon t/2}[x]\mathfrak{m}$). Recalling that the unique solution to (S_{dyn}) with $\mathbf{R}^{\varepsilon/2}$ as reference measure is given by

$$\mathrm{d}\mathbf{P}^{\varepsilon/2}(\gamma)=f^\varepsilon(\gamma_0)g^\varepsilon(\gamma_1)\mathrm{d}\mathbf{R}^{\varepsilon/2}(\gamma),$$

from these facts and the Markov property of $\mathbf{R}^{\varepsilon/2}$ (see Appendix A.4 and (A.4.1)) we deduce that

$$\begin{aligned}\frac{\mathrm{d}\mathbf{P}^\varepsilon_t}{\mathrm{d}\mathfrak{m}}(x)=\frac{\mathrm{d}\mathbf{P}^\varepsilon_t}{\mathrm{d}\mathbf{R}^\varepsilon_t}(x)&\overset{(A.2.1)}{=}\mathsf{E}_{\mathbf{R}}\left[\frac{\mathrm{d}\mathbf{P}^\varepsilon}{\mathrm{d}\mathbf{R}^\varepsilon}\,\middle|\,Z_t=x\right]=\mathsf{E}_{\mathbf{R}}[f^\varepsilon(Z_0)g^\varepsilon(Z_1)\,|\,Z_t=x]\\&=\mathsf{E}_{\mathbf{R}}[f^\varepsilon(Z_0)\,|\,Z_t=x]\mathsf{E}_{\mathbf{R}}[g^\varepsilon(Z_1)\,|\,Z_t=x]\overset{(5.1.9)}{=}f^\varepsilon_t(x)g^\varepsilon_t(x)\end{aligned}$$

whence (3.1.9). This tells us that the entropic interpolation (μ^ε_t) is the marginal flow of the solution to the dynamical Schrödinger problem associated to μ_0,μ_1 and $\mathbf{R}^{\varepsilon/2}$, namely $\mu^\varepsilon_t=(\mathrm{e}_t)_*\mathbf{P}^{\varepsilon/2}$ for all $t\in[0,1]$, thus renforcing the connection between the analytic and probabilistic approaches to the topic.

5.2 Uniform estimates for the densities and the potentials

We start investigating the continuity of several functions defined in terms of Schrödinger potentials and densities of entropic interpolations.

Lemma 5.2.1. *With the same assumptions and notation as in Setting 5.1.3, the following holds.*

For any $\varepsilon>0$ and for any $p<\infty$ the maps

$$\begin{array}{lll}(0,1)\ni t&\mapsto&\rho^\varepsilon_t|h^\varepsilon_t|^p\in L^1(\mathfrak{m})\\(0,1)\ni t&\mapsto&f^\varepsilon_t|\nabla h^\varepsilon_t|^p\in L^1(\mathfrak{m})\\(0,1)\ni t&\mapsto&|\nabla\rho^\varepsilon_t||\nabla h^\varepsilon_t|^p\in L^1(\mathfrak{m})\end{array}\qquad\begin{array}{lll}(0,1)\ni t&\mapsto&\rho^\varepsilon_t|\nabla h^\varepsilon_t|^p\in L^1(\mathfrak{m})\\(0,1)\ni t&\mapsto&g^\varepsilon_t|\nabla h^\varepsilon_t|^p\in L^1(\mathfrak{m})\\(0,1)\ni t&\mapsto&\Delta\rho^\varepsilon_t||\nabla h^\varepsilon_t|^p\in L^1(\mathfrak{m})\end{array}$$

are all continuous w.r.t. the strong topology, where h^ε_t is equal to any of $\varphi^\varepsilon_t,\psi^\varepsilon_t,\vartheta^\varepsilon_t,\log\rho^\varepsilon_t$.

proof Fix $\varepsilon>0$ and recall that by Proposition 5.1.4 we know that $(f^\varepsilon_t),(g^\varepsilon_t),(\rho^\varepsilon_t)$ belong to $C([0,1],L^2(\mathrm{X}))\cap AC_{loc}((0,1),W^{1,2}(\mathrm{X}))$ and $(h^\varepsilon_t)\in AC_{loc}(I,W^{1,2}(\mathrm{X},e^{-V}\mathfrak{m}))$ with $V=M\mathsf{d}^2(\cdot,\bar x)$ for some $\bar x\in\mathrm{X}$ and $M>0$ sufficiently large. This means that for any $t\in[0,1]$ and any sequence $(t_n)_{n\in\mathbb{N}}\subset[0,1]$ converging to t there exists a subsequence $(t_{n_k})_{k\in\mathbb{N}}$ such that

$$f^\varepsilon_{t_{n_k}}\to f^\varepsilon_t\quad g^\varepsilon_{t_{n_k}}\to g^\varepsilon_t\quad\rho^\varepsilon_{t_{n_k}}\to\rho^\varepsilon_t\qquad\mathfrak{m}\text{-a.e.}$$

as $k\to\infty$; if $t\in I$ then

$$h^\varepsilon_{t_{n_k}}\to h^\varepsilon_t\quad|\nabla h^\varepsilon_{t_{n_k}}|\to|\nabla h^\varepsilon_t|\quad\Delta h^\varepsilon_{t_{n_k}}\to\Delta h^\varepsilon_t\qquad\mathfrak{m}\text{-a.e.}$$

and if $t\in(0,1)$ then we also have

$$|\nabla\rho^\varepsilon_{t_{n_k}}|\to|\nabla\rho^\varepsilon_t|\quad\Delta\rho^\varepsilon_{t_{n_k}}\to\Delta\rho^\varepsilon_t\qquad\mathfrak{m}\text{-a.e.}$$

Therefore, it is sufficient to show that all the functions appearing in the statement belong to $L^\infty_{loc}((0,1),L^1(\mathrm{X}))$. To this aim, consider $t\mapsto\rho^\varepsilon_t|\nabla\varphi^\varepsilon_t|^p$ and start observing that by (2.3.3) we

know that for any $\delta \in (0,1]$ and $\bar{x} \in \mathrm{X}$ it holds $|\nabla \varphi_t^\varepsilon| \le C_\delta(1+\mathsf{d}(\cdot,\bar{x}))$ for all $t \in [\delta, 1-\delta]$, whence

$$|\nabla \varphi_t^\varepsilon|^p \le C_{\delta,p}(1+\mathsf{d}^p(\cdot,\bar{x})) \tag{5.2.1}$$

for any $p<\infty$. On the other hand, the estimate for ρ_t^ε is more difficult. Let

$$v_s := \inf_{y\in \mathrm{supp}(\rho_0)\cup \mathrm{supp}(\rho_1)} \mathfrak{m}(B_{\sqrt{s}}(y)) \qquad V_s := \sup_{y\in \mathrm{supp}(\rho_0)\cup \mathrm{supp}(\rho_1)} \mathfrak{m}(B_{\sqrt{s}}(y))$$

and observe that for any $s>0$ we have $0<v_s\le V_s<\infty$. By means of the representation formula (1.2.9c), the Gaussian estimates (1.2.11) and since $t\in[\delta,1-\delta]$, we get

$$\begin{aligned}
\rho_t^\varepsilon(x) &\le \int f^\varepsilon(y) \frac{C_1}{\mathfrak{m}(B_{\sqrt{\varepsilon t/2}}(y))} \exp\Big(-\frac{\mathsf{d}^2(x,y)}{3\varepsilon t}+\frac{C_2\varepsilon t}{2}\Big)\mathrm{d}\mathfrak{m}(y)\\
&\quad\times \int g^\varepsilon(y) \frac{C_1}{\mathfrak{m}(B_{\sqrt{\varepsilon(1-t)/2}}(y))} \exp\Big(-\frac{\mathsf{d}^2(x,y)}{3\varepsilon(1-t)}+\frac{C_2\varepsilon(1-t)}{2}\Big)\mathrm{d}\mathfrak{m}(y)\\
&\le \frac{C_3}{v^2_{\varepsilon\delta/2}}\Big(\int f^\varepsilon \mathrm{d}\mathfrak{m}\int g^\varepsilon \mathrm{d}\mathfrak{m}\Big)e^{C_4(1+\mathsf{d}^2(x,\bar{x}))}
\end{aligned}$$

with C_3, C_4 depending on $K, N, \varepsilon, \delta, \bar{x}$, $\mathrm{supp}(\rho_0)$ and $\mathrm{supp}(\rho_1)$. Using the first inequality in (1.2.11) and the fact that $f^\varepsilon \otimes g^\varepsilon \mathsf{R}^{\varepsilon/2}$ is a probability measure (by construction - recall our Setting 5.1.3) we then obtain

$$\int f^\varepsilon \mathrm{d}\mathfrak{m}\int g^\varepsilon \mathrm{d}\mathfrak{m} \le C_1 V_{\varepsilon/2} e^{\frac{D^2}{\varepsilon}} \int f^\varepsilon \otimes g^\varepsilon \mathsf{R}^{\varepsilon/2} = C_1 V_{\varepsilon/2} e^{\frac{D^2}{\varepsilon}} \tag{5.2.2}$$

where D is the diameter of $\mathrm{supp}(\rho_0)\cup\mathrm{supp}(\rho_1)$. Plugging these pieces of information together we deduce that

$$\rho_t^\varepsilon(x) \le \frac{C_1C_3V_{\varepsilon/2}}{v^2_{\varepsilon\delta/2}} \exp\Big(\frac{D^2}{\varepsilon}+C_4(1+\mathsf{d}^2(x,\bar{x}))\Big) \tag{5.2.3}$$

and coupling this inequality with (5.2.1) we see that

$$\int |\nabla\varphi_t^\varepsilon|^p \rho_t^\varepsilon \mathrm{d}\mathfrak{m} \le C_{\delta,p} M \int e^{-C\mathsf{d}^2(\cdot,\bar{x})}(1+\mathsf{d}^p(\cdot,\bar{x}))\mathrm{d}\mathfrak{m} \qquad \forall t\in[\delta,1-\delta].$$

The integral on the right-hand side can be written as

$$\int_0^\infty e^{-Cr^2}(1+r^p)\mathrm{d}T_*\mathfrak{m},$$

where $T:\mathrm{X}\to[0,\infty)$ is defined as $T=\mathsf{d}(\cdot,\bar{x})$, and by the Bishop-Gromov inequality in spherical form (1.2.3) such integral is finite. By the arbitrariness of $\delta\in(0,1]$ we have thus shown that $t\mapsto \rho_t^\varepsilon|\nabla\varphi_t^\varepsilon|^p$ is in $L^\infty_{loc}((0,1),L^1(\mathrm{X}))$ and the same argument applies to $\psi_t^\varepsilon, \vartheta_t^\varepsilon, \log\rho_t^\varepsilon$.

With slight modifications we can also handle the case of $t\mapsto f_t^\varepsilon|\nabla h_t^\varepsilon|^p$ and $t\mapsto g_t^\varepsilon|\nabla h_t^\varepsilon|^p$: in fact, fixing $[t_0,t_1]\subset(0,1)$, it is sufficient to replace (5.2.3) with the second inequality in (5.1.5) and couple it with (5.2.1) so that

$$\sup_{t\in[t_0,t_1]}\int |\nabla\varphi_t^\varepsilon|^p f_t^\varepsilon \mathrm{d}\mathfrak{m} \le \sup_{t\in[t_0,t_1]} C_1\int e^{\frac{C_2}{\varepsilon t}-\frac{C_3\mathsf{d}^2(\cdot,\bar{x})}{\varepsilon t}}(1+\mathsf{d}^p(\cdot,\bar{x}))\mathrm{d}\mathfrak{m}<\infty \tag{5.2.4}$$

for suitable positive constants C_1, C_2, C_3 independent of $t \in [t_0, t_1]$. The arbitrariness of $t_0, t_1 \in (0,1)$ gives the conclusion and the same holds for $\psi_t^\varepsilon, \vartheta_t^\varepsilon, \log \rho_t^\varepsilon$; for g_t^ε use the analogous of (5.1.5).

Passing to $t \mapsto \rho_t^\varepsilon |\varphi_t^\varepsilon|^p$, fix again $\delta \in (0,1)$ and use both the first and the second inequality in (5.1.5) to deduce that

$$|\varphi_t^\varepsilon|^p \le C(1 + \mathsf{d}^p(\cdot, \bar{x})) \qquad \forall t \in [\delta, 1]$$

for a suitable positive constant C independent of $t \in [\delta, 1]$; from this inequality, (5.2.3) and arguing as for $t \mapsto \rho_t^\varepsilon |\nabla \varphi_t^\varepsilon|^p$ we conclude. The same holds for $\psi_t^\varepsilon, \vartheta_t^\varepsilon, \log \rho_t^\varepsilon$.

Finally, for $t \mapsto |\nabla \rho_t^\varepsilon||\nabla h_t^\varepsilon|^p$ and $t \mapsto \Delta \rho_t^\varepsilon |\nabla h_t^\varepsilon|^p$ notice that

$$\begin{aligned}
|\nabla \rho_t^\varepsilon||\nabla h_t^\varepsilon|^p =& \rho_t^\varepsilon |\nabla \log \rho_t^\varepsilon||\nabla h_t^\varepsilon|^p \le \frac{1}{2}\rho_t^\varepsilon |\nabla \log \rho_t^\varepsilon|^2 + \frac{1}{2}\rho_t^\varepsilon |\nabla h_t^\varepsilon|^{2p} \\
\Delta \rho_t^\varepsilon |\nabla h_t^\varepsilon|^p =& (\Delta f_t^\varepsilon) g_t^\varepsilon |\nabla h_t^\varepsilon|^p + (\Delta g_t^\varepsilon) f_t^\varepsilon |\nabla h_t^\varepsilon|^p + 2\varepsilon^{-2} \rho_t^\varepsilon \langle \nabla \varphi_t^\varepsilon, \nabla \psi_t^\varepsilon \rangle |\nabla h_t^\varepsilon|^p \\
\le& (\Delta f_t^\varepsilon) g_t^\varepsilon |\nabla h_t^\varepsilon|^p + (\Delta g_t^\varepsilon) f_t^\varepsilon |\nabla h_t^\varepsilon|^p + \frac{1}{2\varepsilon^2}\rho_t^\varepsilon |\nabla \varphi_t^\varepsilon|^4 \\
&+ \frac{1}{2\varepsilon^2}\rho_t^\varepsilon |\nabla \psi_t^\varepsilon|^4 + \frac{1}{\varepsilon^2}\rho_t^\varepsilon |\nabla h_t^\varepsilon|^{2p},
\end{aligned}$$

use the a priori estimate (1.2.12b) to deduce that $\Delta f_t^\varepsilon, \Delta g_t^\varepsilon \in L^\infty_{loc}((0,1), L^2(\mathfrak{m}))$ and point out that $t \mapsto f_t^\varepsilon |\nabla h_t^\varepsilon|^p$ and $t \mapsto g_t^\varepsilon |\nabla h_t^\varepsilon|^p$ also belong to $L^\infty_{loc}((0,1), L^2(\mathfrak{m}))$. This last claim is a straightforward generalization of (5.2.4) from L^1 to any L^p-norm. □

Now let us collect information about quantities which remain bounded as $\varepsilon \downarrow 0$. The first result is a great improvement of (5.1.3) and could not be fully appreciated in the compact case considered in [63].

Proposition 5.2.2 (uniform L^∞ bound on the densities)**.** *With the same assumptions and notations as in Setting 5.1.3 the following holds.*

For every $\bar{x} \in \mathrm{X}$ there exist constants $C, M > 0$ which depends on $K, N, \bar{x}, \rho_0, \rho_1$ such that

$$\rho_t^\varepsilon \le M e^{-C\mathsf{d}^2(\cdot, \bar{x})} \qquad \mathfrak{m}\text{-a.e.} \tag{5.2.5}$$

for every $t \in [0,1]$ and for every $\varepsilon \in (0,1)$.

proof We claim that there exists a constant $M' > 0$ which depends on K, N and the diameters of the supports of ρ_0, ρ_1 such that

$$\|\rho_t^\varepsilon\|_{L^\infty(\mathrm{X})} \le M' \max\{\|\rho_0\|_{L^\infty(\mathrm{X})}, \|\rho_1\|_{L^\infty(\mathrm{X})}\} \tag{5.2.6}$$

for every $t \in [0,1]$ and for every $\varepsilon \in (0,1)$.

Fix $\varepsilon > 0$. We know that $(\rho_t^\varepsilon) \in C([0,1], L^2(\mathrm{X})) \cap AC_{loc}((0,1), L^2(\mathrm{X}))$ from Proposition 5.1.4 and $\rho_t^\varepsilon \le C_\varepsilon$ for all $t \in [0,1]$ by the maximum principle, thus for any $p \ge 2$ the function $E_p : [0,1] \to [0,\infty)$ defined by

$$E_p(t) := \int (\rho_t^\varepsilon)^p \, \mathrm{d}\mathfrak{m},$$

belongs to $C([0,1]) \cap AC_{loc}((0,1))$. An application of the dominated convergence theorem grants that its derivative can be computed passing the limit inside the integral, obtaining

$$\frac{\mathrm{d}}{\mathrm{d}t} E_p(t) = p \int (\rho_t^\varepsilon)^{p-1} \frac{\mathrm{d}}{\mathrm{d}t} \rho_t^\varepsilon \mathrm{d}\mathfrak{m} = -p \int (\rho_t^\varepsilon)^{p-1} \mathrm{div}(\rho_t^\varepsilon \nabla \vartheta_t^\varepsilon) \, \mathrm{d}\mathfrak{m}. \tag{5.2.7}$$

However, integration by parts formula can not be applied now, because $|\nabla\vartheta_t^\varepsilon|$ is only locally integrable. To overcome the problem, let $\bar{x}\in \mathrm{X}, R>0$ and $\chi_R\in \mathrm{Test}^\infty(\mathrm{X})$ be a cut-off function with support in $B_{R+1}(\bar{x})$ such that $\chi_R\equiv 1$ in $B_R(\bar{x})$; then

$$\begin{aligned}
p\int \chi_R(\rho_t^\varepsilon)^{p-1}\mathrm{div}(\rho_t^\varepsilon\nabla\vartheta_t^\varepsilon)\,\mathrm{d}\mathfrak{m} &= -p\int\langle\nabla(\rho_t^\varepsilon)^{p-1},\nabla\vartheta_t^\varepsilon\rangle\chi_R\rho_t^\varepsilon\,\mathrm{d}\mathfrak{m} - p\int\langle\nabla\chi_R,\nabla\vartheta_t^\varepsilon\rangle(\rho_t^\varepsilon)^p\,\mathrm{d}\mathfrak{m}\\
&= -(p-1)\int\langle\nabla(\rho_t^\varepsilon)^{p},\nabla\vartheta_t^\varepsilon\rangle\chi_R\,\mathrm{d}\mathfrak{m} - p\int\langle\nabla\chi_R,\nabla\vartheta_t^\varepsilon\rangle(\rho_t^\varepsilon)^p\,\mathrm{d}\mathfrak{m}\\
&= (p-1)\int\chi_R(\rho_t^\varepsilon)^p\Delta\vartheta_t^\varepsilon\,\mathrm{d}\mathfrak{m} - \int(\rho_t^\varepsilon)^p\langle\nabla\vartheta_t^\varepsilon,\nabla\chi_R\rangle\,\mathrm{d}\mathfrak{m}.
\end{aligned}$$

On the one hand, $(\rho_t^\varepsilon)^{p-1}\mathrm{div}(\rho_t^\varepsilon\nabla\vartheta_t^\varepsilon)\in L^1(\mathrm{X})$ because $\mathrm{div}(\rho_t^\varepsilon\nabla\vartheta_t^\varepsilon)\in L^2(\mathrm{X})$ by Proposition 5.1.4 and thus, by the dominated convergence theorem, the left-hand side converges to itself without χ_R as $R\to\infty$. On the other hand, the uniform boundedness of $\||\nabla\chi_R|\|_{L^\infty(\mathrm{X})}$ w.r.t. R and the fact that $(\rho_t^\varepsilon)^p|\nabla\vartheta_t^\varepsilon|\in L^1(\mathrm{X})$ by Lemma 5.2.1 imply that the second term on the right-hand side vanishes in the limit. As regards the first one, we claim that $(\rho_t^\varepsilon)^p\Delta\vartheta_t^\varepsilon\in L^1(\mathrm{X})$: indeed

$$\Delta\vartheta_t^\varepsilon = \frac{\varepsilon}{2}\Big(\frac{\Delta g_t^\varepsilon}{g_t^\varepsilon}-\frac{\Delta f_t^\varepsilon}{f_t^\varepsilon}\Big)-\frac{1}{2\varepsilon}\Big(|\nabla\psi_t^\varepsilon|^2-|\nabla\varphi_t^\varepsilon|^2\Big)$$

and by the fact that $\Delta f_t^\varepsilon,\Delta g_t^\varepsilon\in L^2(\mathrm{X})$ together with (2.3.3) and (5.2.3), the conclusion follows along the same lines pointed out in Lemma 5.2.1. Hence, the claim ensures that the first term on the right-hand side converges to itself without χ_R as $R\to\infty$, so that all in all

$$p\int(\rho_t^\varepsilon)^{p-1}\mathrm{div}(\rho_t^\varepsilon\nabla\vartheta_t^\varepsilon)\,\mathrm{d}\mathfrak{m} = (p-1)\int(\rho_t^\varepsilon)^p\Delta\vartheta_t^\varepsilon\,\mathrm{d}\mathfrak{m}$$

and plugging this identity into (5.2.7) we get

$$\frac{\mathrm{d}}{\mathrm{d}t}E_p(t) = -(p-1)\int(\rho_t^\varepsilon)^p\Delta\vartheta_t^\varepsilon\,\mathrm{d}\mathfrak{m}.$$

Now notice that $\vartheta_t^\varepsilon=\psi_t^\varepsilon-\frac{\varepsilon}{2}\log\rho_t^\varepsilon$ to get

$$\frac{\mathrm{d}}{\mathrm{d}t}E_p(t) = -(p-1)\int(\rho_t^\varepsilon)^p\Delta\psi_t^\varepsilon\mathrm{d}\mathfrak{m}+\frac{\varepsilon}{2}(p-1)\int(\rho_t^\varepsilon)^p\Delta\log\rho_t^\varepsilon\mathrm{d}\mathfrak{m}. \tag{5.2.8}$$

On the one hand, using again the same cut-off technique to motivate integration by parts,

$$\int(\rho_t^\varepsilon)^p\Delta\log\rho_t^\varepsilon\mathrm{d}\mathfrak{m} = -p\int(\rho_t^\varepsilon)^{p-1}\langle\nabla\rho_t^\varepsilon,\nabla\log\rho_t^\varepsilon\rangle\mathrm{d}\mathfrak{m} = -p\int(\rho_t^\varepsilon)^{p-2}|\nabla\rho_t^\varepsilon|^2\mathrm{d}\mathfrak{m}\le 0$$

and on the other one, choosing $\delta:=\frac{1}{2}$ and $T=1$ in (2.3.6) and any point $\bar{x}\in\mathrm{X}$ we get the existence of a constant $C>0$ depending on $K,N,\bar{x}$ and the diameters of the supports of ρ_0,ρ_1 such that $\Delta\psi_t^\varepsilon\ge -C(1+\mathsf{d}^2(\cdot,\bar{x}))$ for any $t\in[0,1/2]$, so that we have

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}E_p(t) &\le C(p-1)E_p(t)+C(p-1)\int(\rho_t^\varepsilon)^p\mathsf{d}^2(\cdot,\bar{x})\mathrm{d}\mathfrak{m}\\
&= C(p-1)E_p(t)+C(p-1)\Big(\int_{B_R(\bar{x})}(\rho_t^\varepsilon)^p\mathsf{d}^2(\cdot,\bar{x})\mathrm{d}\mathfrak{m}+\int_{\mathrm{X}\setminus B_R(\bar{x})}(\rho_t^\varepsilon)^p\mathsf{d}^2(\cdot,\bar{x})\mathrm{d}\mathfrak{m}\Big)\\
&\le C(p-1)(1+R^2)E_p(t)+C(p-1)\int_{\mathrm{X}\setminus B_R(\bar{x})}(\rho_t^\varepsilon)^p\mathsf{d}^2(\cdot,\bar{x})\mathrm{d}\mathfrak{m},\qquad \forall t\in[0,1/2]
\end{aligned}$$

where $R>0$ is to be fixed later. Now let us handle the rightmost term above. To this aim, we can follow the argument already adopted in the proof of Lemma 5.2.1 for the upper bound on ρ_t^ε, but the previous rough estimate has to be made more precise: the dependence on t,ε shall be explicit. Keeping the same notations introduced in the lemma, notice that by means of the representation formula (1.2.9c), the Gaussian estimates (1.2.11) and since $t\in[0,1/2]$, we get

$$\begin{aligned}\rho_t^\varepsilon(x) &\le \int f^\varepsilon(y)\frac{C_1}{\mathfrak{m}(B_{\sqrt{\varepsilon t/2}}(y))}\exp\Big(-\frac{\mathsf{d}^2(x,y)}{3\varepsilon t}+\frac{C_2\varepsilon t}{2}\Big)\mathrm{d}\mathfrak{m}(y)\\ &\quad\times\int g^\varepsilon(y)\frac{C_1}{\mathfrak{m}(B_{\sqrt{\varepsilon(1-t)/2}}(y))}\exp\Big(-\frac{\mathsf{d}^2(x,y)}{3\varepsilon(1-t)}+\frac{C_2\varepsilon(1-t)}{2}\Big)\mathrm{d}\mathfrak{m}(y)\\ &\le \frac{C_3}{v_{\varepsilon t/2}v_{\varepsilon/4}}\int f^\varepsilon\mathrm{d}\mathfrak{m}\int g^\varepsilon\mathrm{d}\mathfrak{m}\exp\Big(\frac{C_4}{\varepsilon t}-\frac{C_5\mathsf{d}^2(x,\bar x)}{\varepsilon t}\Big)\end{aligned}$$

with C_3,C_4,C_5 depending on $K,N,\bar x$, $\mathrm{supp}(\rho_0)$ and $\mathrm{supp}(\rho_1)$. Plugging (5.2.2) in the inequality above we deduce that

$$\rho_t^\varepsilon(x)\le\frac{C_1C_3V_{\varepsilon/2}}{v_{\varepsilon t/2}v_{\varepsilon/4}}\exp\Big(\frac{D^2}{\varepsilon}+\frac{C_4}{\varepsilon t}-\frac{C_5\mathsf{d}^2(x,\bar x)}{\varepsilon t}\Big).$$

Now observe that by the Bishop-Gromov inequality (1.2.2) with $R=1$ $v_s\ge C_6s^N$ for $s\in(0,1)$, with C_6 depending on K,N and D, while by the compactness of the supports of ρ_0 and ρ_1 we have $V_{\varepsilon/2}\le C_7$ with C_7 depending on the same parameters as C_6. Hence

$$\rho_t^\varepsilon(x)\le\frac{C_8}{\varepsilon^Nt^{N/2}}\exp\Big(\frac{D^2}{\varepsilon}+\frac{C_4}{\varepsilon t}-\frac{C_5\mathsf{d}^2(x,\bar x)}{\varepsilon t}\Big)\tag{5.2.9}$$

and if we raise the inequality to the power p, multiply both sides by $\mathsf{d}^2(\cdot,\bar x)$, integrate over $\mathrm{X}\setminus B_R(\bar x)$, rewrite the integral in terms of the push-forward measure $T_*\mathfrak{m}$ (where $T=\mathsf{d}(\cdot,\bar x)$) and use the Bishop-Gromov inequality in spherical form (1.2.3) to control $T_*\mathfrak{m}$, then we get

$$\int_{\mathrm{X}\setminus B_R(\bar x)}(\rho_t^\varepsilon)^p\mathsf{d}^2(\cdot,\bar x)\mathrm{d}\mathfrak{m}\le\frac{C_8^p}{\varepsilon^{Np}t^{Np/2}}e^{\frac{pD^2}{\varepsilon}+\frac{pC_4}{\varepsilon t}}\int_R^\infty C_9r^2e^{-\frac{pC_5r^2}{\varepsilon t}}e^{C_9r}\mathrm{d}r$$

with C_9 depending on $K,N,\mathrm{supp}(\rho_0)$ and $\mathrm{supp}(\rho_1)$. With explicit manipulations this yields

$$\int_{\mathrm{X}\setminus B_R(\bar x)}(\rho_t^\varepsilon)^p\mathsf{d}^2(\cdot,\bar x)\mathrm{d}\mathfrak{m}\le\left(\frac{C_{10}}{(\varepsilon^2t)^{N/2}}e^{\frac{D^2}{\varepsilon}+\frac{C_4}{\varepsilon t}-\frac{C_5R^2}{\varepsilon t}}\right)^p$$

where C_{10} continuously depends on R. If we choose $R>\sqrt{(D^2+C_4)/C_5}$, then there exists a constant c independent of $\varepsilon,t\in(0,1)$ such that the right-hand side is bounded from above by c^p. We have thus obtained $E_p'\le C'(p-1)E_p+c^p$ for all $t\in[0,1/2]$, whence by Gronwall's inequality

$$E_p(t)\le\left(E_p(0)+\frac{c^p}{C'(p-1)}\right)e^{C'(p-1)t}-\frac{c^p}{C'(p-1)}\le\left(E_p(0)+\frac{c^p}{C'(p-1)}\right)e^{C'(p-1)}$$

for all $t\in[0,1/2]$. Passing to the p-th roots, writing $E_p(t)=\|\rho_t^\varepsilon\|^{p-1}_{L^{p-1}(\mu_t^\varepsilon)}$ and observing that, being μ_t^ε a probability measure, we have $\|h\|_{L^p(\mu_t^\varepsilon)}\uparrow\|h\|_{L^\infty(\mu_t^\varepsilon)}$ as $p\to\infty$, we obtain

$$\|\rho_t^\varepsilon\|_{L^\infty}\le e^{C'}\|\rho_0\|_{L^\infty},\quad\forall t\in[0,1/2].$$

Switching the roles of ρ_0 and ρ_1 we get the analogous control for $t \in [\frac{1}{2}, 1]$, whence the claim (5.2.6) with $M' := e^{\tilde{C}}$. As regards (5.2.5), for $t = 0, 1$ there is nothing to prove; for $t \in (0, 1)$ let us look back at (5.2.9) and observe that

$$\frac{C_8}{\varepsilon^N t^{N/2}} e^{\frac{D^2}{\varepsilon} + \frac{C_4}{\varepsilon t}} e^{-\frac{C_5 \mathsf{d}^2(x,\bar{x})}{\varepsilon t}} \leq C'' e^{-\tilde{C}\mathsf{d}^2(x,\bar{x})} \qquad \mathfrak{m}\text{-a.e. in } \mathrm{X} \setminus B_R(\bar{x})$$

with $C'', \tilde{C}$ independent of $\varepsilon, t \in (0, 1)$. This bound together with (5.2.6) gives (5.2.5) with $C = \tilde{C}$ and

$$M := \max\{M' e^{\tilde{C}R^2} \max\{\|\rho_0\|_{L^\infty(\mathfrak{m})}, \|\rho_1\|_{L^\infty(\mathfrak{m})}\}, C''\}.$$

□

The second result we present is mostly an easy consequence or a rewriting of the estimates obtained in Chapter 2, but it is worth a second statement because it improves (5.1.4) in a quantitative sense.

Proposition 5.2.3 (Locally uniform Lipschitz and Laplacian controls for the potentials)**.** *With the same assumptions and notations as in Setting 5.1.3 the following holds.*

For all $\delta \in (0, 1)$ and $\bar{x} \in \mathrm{X}$ there exists $C > 0$ which only depends on $K, N, \delta, \bar{x}$ such that

$$\mathrm{lip}(\varphi_t^\varepsilon) \leq C\big(1 + \mathsf{d}(\cdot, \bar{x})\big), \qquad \mathfrak{m}\text{-a.e.} \tag{5.2.10a}$$

$$\Delta\varphi_t^\varepsilon \geq -C\big(1 + \mathsf{d}^2(\cdot, \bar{x})\big), \qquad \mathfrak{m}\text{-a.e.} \tag{5.2.10b}$$

and for all $\delta \in (0, 1)$ and bounded Borel set B there exists $C' > 0$ which only depends on K, N, δ, B such that

$$\|\Delta\varphi_t^\varepsilon\|_{L^1(B,\mathfrak{m})} \leq C' \tag{5.2.11}$$

for every $t \in [\delta, 1]$ and $\varepsilon \in (0, 1)$. Analogous bounds hold for the ψ_t^ε's in the time interval $[0, 1 - \delta]$.

proof Fix $\delta \in (0, 1)$ and $\bar{x} \in \mathrm{X}$ as in the statement and notice that the bound (2.3.3) yields

$$|\nabla\varphi_t^\varepsilon| = \varepsilon|\nabla \log \mathsf{h}_{\frac{\varepsilon t}{2}} f^\varepsilon| \leq C\big(1 + \mathsf{d}(\cdot, \bar{x})\big) \qquad \forall t \in [\delta, 1],\ \varepsilon \in (0, 1).$$

Thus recalling the Sobolev-to-Lipschitz property (1.2.14) we obtain the bound (5.2.10a). The bound (5.2.10b) is a restatement of (2.3.6). Finally, let B be as in the statement and χ_B a Lipschitz cut-off function identically equal to 1 on B and with bounded support, notice that $|h| = h + 2h^-$ whence

$$\begin{aligned}\int_B |\Delta\varphi_t^\varepsilon|\,\mathrm{d}\mathfrak{m} &= \int_B \chi_B|\Delta\varphi_t^\varepsilon|\,\mathrm{d}\mathfrak{m} \leq \int \chi_B|\Delta\varphi_t^\varepsilon|\,\mathrm{d}\mathfrak{m} = \int \chi_B\Delta\varphi_t^\varepsilon\,\mathrm{d}\mathfrak{m} + 2\int \chi_B(\Delta\varphi_t^\varepsilon)^-\,\mathrm{d}\mathfrak{m} \\ &= -\int \langle\nabla\chi_B, \nabla\varphi_t^\varepsilon\rangle\mathrm{d}\mathfrak{m} + 2\int \chi_B(\Delta\varphi_t^\varepsilon)^-\,\mathrm{d}\mathfrak{m}\end{aligned}$$

and take into account (5.2.10a) to estimate the first term in the final expression and (5.2.10b) for the second one: (5.2.11) then follows.

The bounds for ψ_t^ε are obtained in the same way. □

5.3 The entropy along entropic interpolations

In [76] Léonard computed the first and second derivatives of the relative entropy along entropic interpolations and in [63] we verified that his computations are fully justifiable in the case of compact $\mathsf{RCD}^*(K,N)$ spaces. Here we would like to show that the same holds true also in the non-compact setting. As we shall see later on, these formulas will be the crucial tool for showing that the acceleration of the entropic interpolation goes to 0 in a suitable sense.

As a first step, let us slightly improve the statement of Lemma 5.2.1.

Lemma 5.3.1. *With the same assumptions and notation as in Setting 5.1.3, the following holds.*

For any $\varepsilon > 0$ and for any $p < \infty$ the maps

$$\begin{array}{rcl} I \ni t & \mapsto & \rho_t^\varepsilon |h_t^\varepsilon|^p \in L^1(\mathfrak{m}) \\ (0,1) \ni t & \mapsto & f_t^\varepsilon |\nabla h_t^\varepsilon|^p \in L^1(\mathfrak{m}) \\ (0,1) \ni t & \mapsto & |\nabla \rho_t^\varepsilon||\nabla h_t^\varepsilon|^p \in L^1(\mathfrak{m}) \end{array} \qquad \begin{array}{rcl} I \ni t & \mapsto & \rho_t^\varepsilon |\nabla h_t^\varepsilon|^p \in L^1(\mathfrak{m}) \\ (0,1) \ni t & \mapsto & g_t^\varepsilon |\nabla h_t^\varepsilon|^p \in L^1(\mathfrak{m}) \\ (0,1) \ni t & \mapsto & \Delta \rho_t^\varepsilon||\nabla h_t^\varepsilon|^p \in L^1(\mathfrak{m}) \end{array}$$

are all continuous w.r.t. the strong topology, where h_t^ε is equal to any of $\varphi_t^\varepsilon, \psi_t^\varepsilon, \vartheta_t^\varepsilon, \log \rho_t^\varepsilon$ and I is the respective domain of definition (for $(\log \rho_t^\varepsilon)$ we pick $I = (0,1)$).

proof The only change w.r.t. Lemma 5.2.1 regards $\rho_t^\varepsilon |h_t^\varepsilon|^p$ and $\rho_t^\varepsilon |\nabla h_t^\varepsilon|^p$: for the proof it is sufficient to follow the same argument, just using (5.2.5) instead of (5.2.3). □

Secondly, although $\mathbf{\Delta}|\nabla \varphi_t^\varepsilon|^2$ is not a Radon measure in general because of lack of integrability (and the same is true for $\psi_t^\varepsilon, \vartheta_t^\varepsilon, \log \rho_t^\varepsilon$), it turns out that $\rho_t^\varepsilon \mathbf{\Delta}|\nabla \varphi_t^\varepsilon|^2$ is a finite measure and $\rho_t^\varepsilon \mathbf{\Gamma}_2(\varphi_t^\varepsilon)$ too, so that a weighted Bochner inequality for the Schrödinger potentials can be deduced.

Lemma 5.3.2. *With the same assumptions and notations as in Setting 5.1.3, the following holds.*

For any $\varepsilon > 0$ and $t \in (0,1)$ the set functions

$$\begin{array}{l} \rho_t^\varepsilon \mathbf{\Delta}|\nabla \varphi_t^\varepsilon|^2,\ \rho_t^\varepsilon \mathbf{\Delta}|\nabla \psi_t^\varepsilon|^2,\ \rho_t^\varepsilon \mathbf{\Delta}|\nabla \vartheta_t^\varepsilon|^2,\ \rho_t^\varepsilon \mathbf{\Delta}|\nabla \log \rho_t^\varepsilon|^2 \\ \rho_t^\varepsilon \mathbf{\Gamma}_2(\varphi_t^\varepsilon),\ \rho_t^\varepsilon \mathbf{\Gamma}_2(\psi_t^\varepsilon),\ \rho_t^\varepsilon \mathbf{\Gamma}_2(\vartheta_t^\varepsilon),\ \rho_t^\varepsilon \mathbf{\Gamma}_2(\log \rho_t^\varepsilon) \end{array}$$

are well defined finite Borel measures on X *and the following integration by parts identity holds:*

$$\int \rho_t^\varepsilon \mathrm{d}\mathbf{\Delta}|\nabla h_t^\varepsilon|^2 = \int \Delta \rho_t^\varepsilon |\nabla h_t^\varepsilon|^2 \mathrm{d}\mathfrak{m} \tag{5.3.1}$$

where h_t^ε is equal to any of $\varphi_t^\varepsilon, \psi_t^\varepsilon, \vartheta_t^\varepsilon, \log \rho_t^\varepsilon$.

Finally, the following weighted Bochner inequalities are satisfied:

$$\rho_t^\varepsilon \mathbf{\Gamma}_2(h_t^\varepsilon) \geq \rho_t^\varepsilon \big(|\mathrm{Hess}(h_t^\varepsilon)|_{\mathsf{HS}}^2 + K|\mathrm{d}h_t^\varepsilon|^2\big)\mathfrak{m} \tag{5.3.2a}$$

$$\rho_t^\varepsilon \mathbf{\Gamma}_2(h_t^\varepsilon) \geq \rho_t^\varepsilon \Big(\frac{(\Delta h_t^\varepsilon)^2}{N} + K|\mathrm{d}h_t^\varepsilon|^2\Big)\mathfrak{m} \tag{5.3.2b}$$

where h_t^ε is as above.

proof Fix $\varepsilon > 0$, $t \in (0,1)$, recall that $f_t^\varepsilon, \rho_t^\varepsilon \in \mathrm{Test}(\mathrm{X})$ and $\varphi_t^\varepsilon \in \mathrm{Test}_{loc}(\mathrm{X})$ by Proposition 5.1.4: as a consequence, $|\nabla\varphi_t^\varepsilon|^2 \in D_{loc}(\mathbf{\Delta})$ with

$$\mathbf{\Delta}|\nabla\varphi_t^\varepsilon|^2 = \varepsilon^2\rho_t^\varepsilon\frac{\mathbf{\Delta}|\nabla f_t^\varepsilon|^2}{(f_t^\varepsilon)^2} + \varepsilon^2\rho_t^\varepsilon|\nabla f_t^\varepsilon|^2\Delta\Big(\frac{1}{(f_t^\varepsilon)^2}\Big) + 2\varepsilon^2\rho_t^\varepsilon\langle\nabla|\nabla f_t^\varepsilon|^2, \nabla\Big(\frac{1}{(f_t^\varepsilon)^2}\Big)\rangle \tag{5.3.3}$$

so that $\mathbf{\Delta}|\nabla\varphi_t^\varepsilon|^2$ and $\mathbf{\Gamma}_2(\varphi_t^\varepsilon)$ have a meaning, in the sense provided in Section 2.2. Secondly, define $\alpha_1, \alpha_2, \alpha_3$ as the first, the second and the third term on the right-hand side of (5.3.3) respectively. By (1.2.23) applied to f_t^ε we get

$$\begin{aligned}\alpha_1 &\geq 2\frac{\varepsilon^2\rho_t^\varepsilon}{(f_t^\varepsilon)^2}\Big(K|\nabla f_t^\varepsilon|^2 + \langle\nabla f_t^\varepsilon, \nabla\Delta f_t^\varepsilon\rangle\Big) = 2K\rho_t^\varepsilon|\nabla\varphi_t^\varepsilon|^2 + 2\varepsilon g_t^\varepsilon\langle\nabla\varphi_t^\varepsilon, \nabla\Delta f_t^\varepsilon\rangle\\ &\geq 2K\rho_t^\varepsilon|\nabla\varphi_t^\varepsilon|^2 - \varepsilon g_t^\varepsilon|\nabla\varphi_t^\varepsilon|^2 - \varepsilon g_t^\varepsilon|\nabla\Delta f_t^\varepsilon|^2\end{aligned}$$

and by Lemma 5.3.1 together with $g_t^\varepsilon \in \mathrm{Test}(\mathrm{X})$ the right-hand side is integrable. As regards α_2, by the following computations

$$\begin{aligned}\alpha_2 &= 2\varepsilon^2\rho_t^\varepsilon\left(3\frac{|\nabla f_t^\varepsilon|^4}{(f_t^\varepsilon)^4} - \frac{|\nabla f_t^\varepsilon|^2\Delta f_t^\varepsilon}{(f_t^\varepsilon)^3}\right) = 6\varepsilon^{-2}\rho_t^\varepsilon|\nabla\varphi_t^\varepsilon|^4 - 2g_t^\varepsilon\Delta f_t^\varepsilon|\nabla\varphi_t^\varepsilon|^4\\ &\geq 6\varepsilon^{-2}\rho_t^\varepsilon|\nabla\varphi_t^\varepsilon|^4 - g_t^\varepsilon|\Delta f_t^\varepsilon|^2 - g_t^\varepsilon|\nabla\varphi_t^\varepsilon|^8\end{aligned}$$

and still by Lemma 5.3.1 we see that it is integrable and the same holds for α_3, because

$$\begin{aligned}\alpha_3 &= -4\varepsilon^2\frac{\rho_t^\varepsilon}{(f_t^\varepsilon)^3}\langle\nabla|\nabla f_t^\varepsilon|^2, \nabla f_t^\varepsilon\rangle = -8g_t^\varepsilon\mathrm{Hess}(f_t^\varepsilon)(\nabla\varphi_t^\varepsilon, \nabla\varphi_t^\varepsilon)\\ &\geq -4g_t^\varepsilon|\mathrm{Hess}(f_t^\varepsilon)|_{\mathsf{HS}}^2 - 4g_t^\varepsilon|\nabla\varphi_t^\varepsilon|^4.\end{aligned}$$

Hence the negative part of $\rho_t^\varepsilon\mathbf{\Delta}|\nabla\varphi_t^\varepsilon|^2$ is a finite measure on X. As regards the positive one, let $\bar{x} \in \mathrm{X}$, $R > 0$ and $\chi_R \in \mathrm{Test}^\infty(\mathrm{X})$ be a cut-off function with support in $B_{R+1}(\bar{x})$ such that $\chi_R \equiv 1$ in $B_R(\bar{x})$; then

$$\begin{aligned}\int \chi_R\rho_t^\varepsilon \mathrm{d}\mathbf{\Delta}|\nabla\varphi_t^\varepsilon|^2 &= \int \Delta(\chi_R\rho_t^\varepsilon)|\nabla\varphi_t^\varepsilon|^2\mathrm{d}\mathfrak{m}\\ &= \int \Delta\chi_R\rho_t^\varepsilon|\nabla\varphi_t^\varepsilon|^2\mathrm{d}\mathfrak{m} + \int \chi_R\Delta\rho_t^\varepsilon|\nabla\varphi_t^\varepsilon|^2\mathrm{d}\mathfrak{m} + 2\int\langle\nabla\chi_R, \nabla\rho_t^\varepsilon\rangle|\nabla\varphi_t^\varepsilon|^2\mathrm{d}\mathfrak{m}\end{aligned}$$

and by Fatou's lemma

$$\int \rho_t^\varepsilon \mathrm{d}(\mathbf{\Delta}|\nabla\varphi_t^\varepsilon|^2)^+ \leq \liminf_{R\to\infty}\int \chi_R\rho_t^\varepsilon\mathrm{d}(\mathbf{\Delta}|\nabla\varphi_t^\varepsilon|^2)^+$$

while by Lebesgue's dominated convergence theorem $\chi_R \to 1$ in $L^1(\rho_t^\varepsilon(\mathbf{\Delta}|\nabla\varphi_t^\varepsilon|^2)^-)$, whence

$$\int \rho_t^\varepsilon \mathrm{d}\mathbf{\Delta}|\nabla\varphi_t^\varepsilon|^2 \leq \liminf_{R\to\infty}\int \chi_R\rho_t^\varepsilon\mathrm{d}\mathbf{\Delta}|\nabla\varphi_t^\varepsilon|^2. \tag{5.3.4}$$

On the other hand, the fact that $\||\nabla\chi_R|\|_{L^\infty(\mathfrak{m})}, \|\Delta\chi_R\|_{L^\infty(\mathfrak{m})}$ are uniformly bounded as $R \to \infty$ by Lemma 2.2.1 and $\rho_t^\varepsilon|\nabla\varphi_t^\varepsilon|^2, \Delta\rho_t^\varepsilon|\nabla\varphi_t^\varepsilon|^2, |\nabla\rho_t^\varepsilon||\nabla\varphi_t^\varepsilon|^2 \in L^1(\mathfrak{m})$ by Lemma 5.3.1 allows us to apply the dominated convergence theorem and infer that

$$\begin{aligned}&\lim_{R\to\infty}\left(\int \Delta\chi_R\rho_t^\varepsilon|\nabla\varphi_t^\varepsilon|^2\mathrm{d}\mathfrak{m} + \int \chi_R\Delta\rho_t^\varepsilon|\nabla\varphi_t^\varepsilon|^2\mathrm{d}\mathfrak{m} + 2\int\langle\nabla\chi_R, \nabla\rho_t^\varepsilon\rangle|\nabla\varphi_t^\varepsilon|^2\mathrm{d}\mathfrak{m}\right)\\ &\quad= \int \Delta\rho_t^\varepsilon|\nabla\varphi_t^\varepsilon|^2\mathrm{d}\mathfrak{m} < +\infty.\end{aligned}$$

Coupling this information with (5.3.4) we finally get that $\rho_t^\varepsilon \mathbf{\Delta}|\nabla\varphi_t^\varepsilon|^2$ is a finite measure and this information implies that $\chi_R \to 1$ in $L^1(\rho_t^\varepsilon \mathbf{\Delta}|\nabla\varphi_t^\varepsilon|^2)$. Hence equality actually holds in (5.3.4) with the limit in place of the lim inf and (5.3.1) is satisfied. Taking (5.1.7) into account, we notice that

(5.3.5)
$$\begin{aligned}
\rho_t^\varepsilon|\langle\nabla\varphi_t^\varepsilon,\nabla\Delta\varphi_t^\varepsilon\rangle| &\le \varepsilon\rho_t^\varepsilon|\nabla\varphi_t^\varepsilon|\left(\frac{|\nabla\Delta f_t^\varepsilon|}{f_t^\varepsilon}+\frac{\Delta f_t^\varepsilon|\nabla f_t^\varepsilon|}{(f_t^\varepsilon)^2}+\frac{|\nabla|\nabla f_t^\varepsilon|^2|}{(f_t^\varepsilon)^2}+\frac{2|\nabla f_t^\varepsilon|^3}{(f_t^\varepsilon)^3}\right)\\
&= \varepsilon g_t^\varepsilon|\nabla\varphi_t^\varepsilon||\nabla\Delta f_t^\varepsilon| + g_t^\varepsilon\Delta f_t^\varepsilon|\nabla\varphi_t^\varepsilon|^2 + 2\varepsilon^{-1}\rho_t^\varepsilon|\mathrm{Hess}(f_t^\varepsilon)|_{\mathsf{HS}}|\nabla\varphi_t^\varepsilon|^3 + 2\varepsilon^{-2}\rho_t^\varepsilon|\nabla\varphi_t^\varepsilon|^4\\
&\le \frac{1}{2}\varepsilon g_t^\varepsilon|\nabla\varphi_t^\varepsilon|^2 + \frac{1}{2}\varepsilon g_t^\varepsilon|\nabla\Delta f_t^\varepsilon|^2 + \varepsilon^{-1}\rho_t^\varepsilon|\mathrm{Hess}(f_t^\varepsilon)|_{\mathsf{HS}}^2 + \varepsilon^{-1}\rho_t^\varepsilon|\nabla\varphi_t^\varepsilon|^6 + 2\varepsilon^{-2}\rho_t^\varepsilon|\nabla\varphi_t^\varepsilon|^4
\end{aligned}$$

and since all the terms appearing on the right-hand side are integrable, we have just proved that $\rho_t^\varepsilon\mathbf{\Gamma}_2(\varphi_t^\varepsilon)$ is a finite measure too.

Now let us look at (5.3.2a) and observe that it can not be trivially deduced from (1.2.23) applied to φ_t^ε, because $\varphi_t^\varepsilon \notin \mathrm{Test}(\mathrm{X})$. To overcome this problem let $E \subset \mathrm{X}$ be a bounded Borel set and let $\chi \in \mathrm{Test}^\infty(\mathrm{X})$ be a cut-off function with bounded support such that $\chi \equiv 1$ in E. Then, as already remarked in Proposition 2.2.5, $\chi\varphi_t^\varepsilon \in \mathrm{Test}(\mathrm{X})$ and (1.2.23) holds for it, whence

$$\mathbf{\Gamma}_2(\varphi_t^\varepsilon)(E) \ge \int_E \big(|\mathrm{Hess}(\varphi_t^\varepsilon)|_{\mathsf{HS}}^2 + K|\mathrm{d}\varphi_t^\varepsilon|^2\big)\mathrm{d}\mathfrak{m}$$

because $\chi \equiv 1$ in E. Hence φ_t^ε satisfies (1.2.23) on all bounded Borel sets. By multiplying both sides of the inequality above by ρ_t^ε, using the fact that $\rho_t^\varepsilon\mathbf{\Gamma}_2(\varphi_t^\varepsilon)$ is a measure and arguing by σ-additivity, (5.3.2a) then follows and an analogous argument provides us with (5.3.2b).

The claim for $\psi_t^\varepsilon, \vartheta_t^\varepsilon, \log\rho_t^\varepsilon$ are obtained following the same lines. □

Now we are in the position for motivating Léonard's computations, thus getting the formulas for the first and second dervative of the entropy along entropic interpolations.

Proposition 5.3.3. *With the same assumptions and notations as in Setting 5.1.3 the following holds.*

For any $\varepsilon > 0$ the map $t \mapsto H(\mu_t^\varepsilon\,|\,\mathfrak{m})$ belongs to $C([0,1]) \cap C^2(0,1)$ and for every $t \in (0,1)$ it holds

$$\frac{\mathrm{d}}{\mathrm{d}t}H(\mu_t^\varepsilon\,|\,\mathfrak{m}) = \int \langle\nabla\rho_t^\varepsilon,\nabla\vartheta_t^\varepsilon\rangle\,\mathrm{d}\mathfrak{m} = \frac{1}{2\varepsilon}\int\big(|\nabla\psi_t^\varepsilon|^2 - |\nabla\varphi_t^\varepsilon|^2\big)\rho_t^\varepsilon\,\mathrm{d}\mathfrak{m}, \tag{5.3.6a}$$

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}H(\mu_t^\varepsilon\,|\,\mathfrak{m}) = \int \rho_t^\varepsilon\,\mathrm{d}\big(\mathbf{\Gamma}_2(\vartheta_t^\varepsilon) + \tfrac{\varepsilon^2}{4}\mathbf{\Gamma}_2(\log(\rho_t^\varepsilon))\big) = \frac{1}{2}\int \rho_t^\varepsilon\,\mathrm{d}\big(\mathbf{\Gamma}_2(\varphi_t^\varepsilon) + \mathbf{\Gamma}_2(\psi_t^\varepsilon)\big). \tag{5.3.6b}$$

proof All the integrals in (5.3.6a) and (5.3.6b) make sense by Lemma 5.3.1 and Lemma 5.3.2 and the equality of the two expressions for both the first and second derivative follows from $\vartheta_t^\varepsilon = \frac{\psi_t^\varepsilon - \varphi_t^\varepsilon}{2}$, $\varepsilon\log\rho_t^\varepsilon = \psi_t^\varepsilon + \varphi_t^\varepsilon$ and the fact that $\mathbf{\Gamma}_2(\cdot)$ is a quadratic form. In addition, recall (5.1.1) and the fact that $(\rho_t^\varepsilon) \in C([0,1], L^2(\mathfrak{m}))$ by Proposition 5.1.4: from these facts we see that $[0,1] \ni t \mapsto H(\mu_t^\varepsilon\,|\,\tilde{\mathfrak{m}})$ is continuous; moreover, (5.2.5) and $(\rho_t^\varepsilon) \in C([0,1], L^2(\mathfrak{m}))$ imply that

$$[0,1] \ni t \mapsto \int \mathsf{d}^2(\cdot,\bar{x})\mathrm{d}\mu_t^\varepsilon$$

is continuous too for any $\bar{x} \in \mathrm{X}$, whence by (5.1.2) the continuity of $[0,1] \ni t \mapsto H(\mu_t^\varepsilon\,|\,\mathfrak{m})$.

Now fix $\varepsilon > 0$ and recall from Proposition 5.1.4 that $(\rho_t^\varepsilon) \in AC_{loc}((0,1), L^2(\mathrm{X}))$ and that it is, locally in $t \in (0,1)$ and in space, uniformly bounded away from 0 and ∞. Therefore, for given

$\bar{x} \in \mathrm{X}$ and $R > 0$, if we pick a cut-off function $\chi_R \in \mathrm{Test}^\infty(\mathrm{X})$ with support in $B_{R+1}(\bar{x})$ and such that $\chi_R \equiv 1$ in $B_R(\bar{x})$, then for $u(z) := z \log z$ we have that $(0,1) \ni t \mapsto \chi_R u(\rho_t^\varepsilon) \in L^2(\mathrm{X})$ is absolutely continuous. In particular, so is $\int \chi_R u(\rho_t^\varepsilon)\,\mathrm{d}\mathfrak{m}$ and it is then clear that

$$\frac{\mathrm{d}}{\mathrm{d}t}\int \chi_R u(\rho_t^\varepsilon)\mathrm{d}\mathfrak{m} = \int \chi_R(\log(\rho_t^\varepsilon)+1)\frac{\mathrm{d}}{\mathrm{d}t}\rho_t^\varepsilon\,\mathrm{d}\mathfrak{m}, \qquad \text{a.e. } t.$$

Using the formula for $\frac{\mathrm{d}}{\mathrm{d}t}\rho_t^\varepsilon$ provided by Proposition 5.1.4 we then get

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\int \chi_R u(\rho_t^\varepsilon)\,\mathrm{d}\mathfrak{m} &= -\int \chi_R(\log(\rho_t^\varepsilon)+1)\mathrm{div}(\rho_t^\varepsilon\nabla\vartheta_t^\varepsilon)\mathrm{d}\mathfrak{m} = \int \langle\nabla(\chi_R(\log(\rho_t^\varepsilon)+1),\nabla\vartheta_t^\varepsilon\rangle\rho_t^\varepsilon\mathrm{d}\mathfrak{m}\\ &= \int \chi_R\,\langle\nabla\rho_t^\varepsilon,\nabla\vartheta_t^\varepsilon\rangle\,\mathrm{d}\mathfrak{m} + \int \langle\nabla\chi_R,\nabla\vartheta_t^\varepsilon\rangle\,(\log\rho_t^\varepsilon+1)\rho_t^\varepsilon\mathrm{d}\mathfrak{m}\end{aligned}$$

and integrating this identity over $[t_0,t_1]\subset(0,1)$ yields

$$\begin{aligned}&\int \chi_R u(\rho_{t_1}^\varepsilon)\mathrm{d}\mathfrak{m} - \int \chi_R u(\rho_{t_0}^\varepsilon)\mathrm{d}\mathfrak{m}\\ &\qquad = \iint_{t_0}^{t_1} \chi_R\,\langle\nabla\rho_t^\varepsilon,\nabla\vartheta_t^\varepsilon\rangle\,\mathrm{d}t\mathrm{d}\mathfrak{m} + \iint_{t_0}^{t_1}\langle\nabla\chi_R,\nabla\vartheta_t^\varepsilon\rangle\,(\log\rho_t^\varepsilon+1)\rho_t^\varepsilon\mathrm{d}t\mathrm{d}\mathfrak{m}.\end{aligned} \tag{5.3.7}$$

On the one hand $\rho_t^\varepsilon\log\rho_t^\varepsilon \in L^1(\mathfrak{m})$ by Lemma 5.3.1 if $t\in(0,1)$ and trivially if $t=0,1$, thus by the dominated convergence theorem the left-hand side converges to itself without χ_R as $R\to\infty$. On the other hand, $|\nabla\rho_t^\varepsilon||\nabla\vartheta_t^\varepsilon| \in L^\infty_{loc}((0,1),L^1(\mathrm{X}))$ by Lemma 5.3.1, so that again by the dominated convergence theorem the first integral on the right-hand side of (5.3.7) converges to itself without χ_R as $R\to\infty$. Moreover, the fact that $\rho_t^\varepsilon(\log\rho_t^\varepsilon+1)|\nabla\vartheta_t^\varepsilon| \in L^\infty_{loc}((0,1),L^1(\mathrm{X}))$ (as a byproduct of Lemma 5.3.1), $\||\nabla\chi_R|\|_{L^\infty(\mathfrak{m})}$ is uniformly bounded as $R\to\infty$ (Lemma 2.2.1) and $|\nabla\chi_R|$ is supported in $B_{R+1}(\bar{x})\setminus B_R(\bar{x})$ entail that the second integral on the right-hand side of (5.3.7) vanishes in the limit. Therefore, plugging all these pieces of information together we get

$$H(\mu_{t_1}^\varepsilon\,|\,\mathfrak{m}) - H(\mu_{t_0}^\varepsilon\,|\,\mathfrak{m}) = \iint_{t_0}^{t_1}\langle\nabla\rho_t^\varepsilon,\nabla\vartheta_t^\varepsilon\rangle\,\mathrm{d}t\mathrm{d}\mathfrak{m}$$

and the fact that $|\nabla\rho_t^\varepsilon||\nabla\vartheta_t^\varepsilon| \in C((0,1),L^1(\mathrm{X}))$ by Lemma 5.3.1 implies that $t\mapsto H(\mu_t^\varepsilon\,|\,\mathfrak{m})$ is C^1 and the formula for the first derivative holds for all $t\in(0,1)$.

For (5.3.6b), recall that from Proposition 5.1.4 we know that $(\rho_t^\varepsilon)\in AC_{loc}((0,1),W^{1,2}(\mathrm{X}))$ and $(\vartheta_t^\varepsilon)\in AC_{loc}((0,1),W^{1,2}(\mathrm{X},e^{-V}\mathfrak{m}))$ with $V = M\mathsf{d}^2(\cdot,\bar{x})$, for some $\bar{x}\in\mathrm{X}$ and $M>0$ sufficiently large. Hence, for such $\bar{x}$ and for $R>0$ pick a cut-off function $\chi_R\in\mathrm{Test}^\infty(\mathrm{X})$ with the same properties as above and observe that $(0,1)\ni t\mapsto \chi_R\,\langle\nabla\rho_t^\varepsilon,\nabla\vartheta_t^\varepsilon\rangle \in L^2(\mathrm{X})$ is absolutely continuous. In particular, so is $\int\chi_R\,\langle\nabla\rho_t^\varepsilon,\nabla\vartheta_t^\varepsilon\rangle\,\mathrm{d}\mathfrak{m}$ and

$$\frac{\mathrm{d}}{\mathrm{d}t}\int\chi_R\langle\nabla\rho_t^\varepsilon,\nabla\vartheta_t^\varepsilon\rangle\mathrm{d}\mathfrak{m} = \int\chi_R\Big(\langle\nabla\frac{\mathrm{d}}{\mathrm{d}t}\rho_t^\varepsilon,\nabla\vartheta_t^\varepsilon\rangle + \langle\nabla\rho_t^\varepsilon,\nabla\frac{\mathrm{d}}{\mathrm{d}t}\vartheta_t^\varepsilon\rangle\Big)\mathrm{d}\mathfrak{m}, \qquad \text{a.e. } t.$$

Thus from the formulas for $\frac{\mathrm{d}}{\mathrm{d}t}\rho_t^\varepsilon, \frac{\mathrm{d}}{\mathrm{d}t}\vartheta_t^\varepsilon$ provided in Proposition 5.1.4 and integrating over $[t_0,t_1]\subset(0,1)$ we obtain

$$\begin{aligned}\int\chi_R\langle\nabla\rho_{t_1}^\varepsilon,\nabla\vartheta_{t_1}^\varepsilon\rangle\mathrm{d}\mathfrak{m} - \int\chi_R\langle\nabla\rho_{t_0}^\varepsilon,\nabla\vartheta_{t_0}^\varepsilon\rangle\mathrm{d}\mathfrak{m} &= \underbrace{\iint_{t_0}^{t_1} -\chi_R\,\langle\nabla(\mathrm{div}(\rho_t^\varepsilon\nabla\vartheta_t^\varepsilon)),\nabla\vartheta_t^\varepsilon\rangle\,\mathrm{d}t\mathrm{d}\mathfrak{m}}_{A}\\ &\quad + \underbrace{\iint_{t_0}^{t_1}\chi_R\langle\nabla\rho_t^\varepsilon,\nabla\big(-\tfrac{1}{2}|\nabla\vartheta_t^\varepsilon|^2 - \tfrac{\varepsilon^2}{4}\Delta\log(\rho_t^\varepsilon) - \tfrac{\varepsilon^2}{8}|\nabla\log(\rho_t^\varepsilon)|^2\big)\rangle\mathrm{d}t\mathrm{d}\mathfrak{m}}_{B}.\end{aligned}$$

From $\langle\nabla\rho_t^\varepsilon,\nabla\vartheta_t^\varepsilon\rangle\in L^1(\mathfrak{m})$ (Lemma 5.3.1) it is easy to see that the left-hand side converges to itself without χ_R as $R\to\infty$. As regards A, integration by parts and Leibniz rule imply that it can be rewritten as

$$\begin{aligned}A&=\iint_{t_0}^{t_1}\mathrm{div}(\rho_t^\varepsilon\nabla\vartheta_t^\varepsilon)\langle\nabla\chi_R,\nabla\vartheta_t^\varepsilon\rangle\mathrm{d}t\mathrm{d}\mathfrak{m}+\iint_{t_0}^{t_1}\chi_R\mathrm{div}(\rho_t^\varepsilon\nabla\vartheta_t^\varepsilon)\Delta\vartheta_t^\varepsilon\mathrm{d}t\mathrm{d}\mathfrak{m}\\&=\iint_{t_0}^{t_1}\langle\nabla\rho_t^\varepsilon,\nabla\vartheta_t^\varepsilon\rangle\langle\nabla\chi_R,\nabla\vartheta_t^\varepsilon\rangle\mathrm{d}t\mathrm{d}\mathfrak{m}+\iint_{t_0}^{t_1}\rho_t^\varepsilon\Delta\vartheta_t^\varepsilon\langle\nabla\chi_R,\nabla\vartheta_t^\varepsilon\rangle\mathrm{d}t\mathrm{d}\mathfrak{m}\\&\quad-\iint_{t_0}^{t_1}\chi_R\rho_t^\varepsilon\langle\nabla\vartheta_t^\varepsilon,\nabla\Delta\vartheta_t^\varepsilon\rangle\mathrm{d}t\mathrm{d}\mathfrak{m}-\iint_{t_0}^{t_1}\rho_t^\varepsilon\Delta\vartheta_t^\varepsilon\langle\nabla\vartheta_t^\varepsilon,\nabla\chi_R\rangle\mathrm{d}t\mathrm{d}\mathfrak{m}\\&=\iint_{t_0}^{t_1}\langle\nabla\rho_t^\varepsilon,\nabla\vartheta_t^\varepsilon\rangle\langle\nabla\chi_R,\nabla\vartheta_t^\varepsilon\rangle\mathrm{d}t\mathrm{d}\mathfrak{m}-\iint_{t_0}^{t_1}\chi_R\rho_t^\varepsilon\langle\nabla\vartheta_t^\varepsilon,\nabla\Delta\vartheta_t^\varepsilon\rangle\mathrm{d}t\mathrm{d}\mathfrak{m}.\end{aligned}$$

Thanks to (5.3.5), which also tells us that $t\mapsto\rho_t^\varepsilon\langle\nabla\vartheta_t^\varepsilon,\nabla\Delta\vartheta_t^\varepsilon\rangle$ belongs to $L^\infty_{loc}((0,1),L^1(\mathfrak{m}))$, the second integral on the right-hand side converges to itself without χ_R as $R\to\infty$, while for the first one notice that

$$\left|\iint_{t_0}^{t_1}\langle\nabla\rho_t^\varepsilon,\nabla\vartheta_t^\varepsilon\rangle\langle\nabla\chi_R,\nabla\vartheta_t^\varepsilon\rangle\mathrm{d}t\mathrm{d}\mathfrak{m}\right|\le\iint_{t_0}^{t_1}|\nabla\chi_R||\nabla\rho_t^\varepsilon||\nabla\vartheta_t^\varepsilon|^2\mathrm{d}t\mathrm{d}\mathfrak{m}.$$

Since $|\nabla\rho_t^\varepsilon||\nabla\vartheta_t^\varepsilon|^2\in L^\infty_{loc}((0,1),L^1(\mathfrak{m}))$ (Lemma 5.3.1), $\||\nabla\chi_R|\|_{L^\infty(\mathfrak{m})}$ is uniformly bounded as $R\to\infty$ (Lemma 2.2.1) and $|\nabla\chi_R|\to0$ $\mathfrak{m}$-a.e. as $R\to\infty$, by the dominated convergence theorem we deduce that this integral vanishes in the limit. Passing to B, let us rewrite it as

$$\begin{aligned}B&=\iint_{t_0}^{t_1}\frac12|\nabla\vartheta_t^\varepsilon|^2\mathrm{div}(\chi_R\nabla\rho_t^\varepsilon)-\frac{\varepsilon^2}{4}\chi_R\langle\nabla\rho_t^\varepsilon,\nabla\Delta\log\rho_t^\varepsilon\rangle+\frac{\varepsilon^2}{8}|\nabla\log\rho_t^\varepsilon|^2\mathrm{div}(\chi_R\nabla\rho_t^\varepsilon)\mathrm{d}\mathfrak{m}\\&=\iint_{t_0}^{t_1}\frac12\chi_R\Delta\rho_t^\varepsilon|\nabla\vartheta_t^\varepsilon|^2-\frac{\varepsilon^2}{4}\chi_R\rho_t^\varepsilon\langle\nabla\log\rho_t^\varepsilon,\nabla\Delta\log\rho_t^\varepsilon\rangle+\frac{\varepsilon^2}{8}\chi_R\Delta\rho_t^\varepsilon|\nabla\log\rho_t^\varepsilon|^2\mathrm{d}\mathfrak{m}\qquad(5.3.8a)\\&\quad+\iint_{t_0}^{t_1}\frac12|\nabla\vartheta_t^\varepsilon|^2\langle\nabla\chi_R,\nabla\rho_t^\varepsilon\rangle+\frac{\varepsilon^2}{8}|\nabla\log\rho_t^\varepsilon|^2\langle\nabla\chi_R,\nabla\rho_t^\varepsilon\rangle\mathrm{d}\mathfrak{m}.\qquad(5.3.8b)\end{aligned}$$

On the one hand, observe that $\Delta\rho_t^\varepsilon|\nabla\vartheta_t^\varepsilon|^2,\Delta\rho_t^\varepsilon|\nabla\log\rho_t^\varepsilon|^2\in L^\infty_{loc}((0,1),L^1(\mathfrak{m}))$ by Lemma 5.3.1 and $\rho_t^\varepsilon\langle\nabla\log\rho_t^\varepsilon,\nabla\Delta\log\rho_t^\varepsilon\rangle\in L^\infty_{loc}((0,1),L^1(\mathfrak{m}))$ by (5.3.5): these facts and the dominated convergence theorem entail that (5.3.8a) converges to itself without χ_R as $R\to\infty$. On the other hand, the fact that $|\nabla\rho_t^\varepsilon||\nabla\vartheta_t^\varepsilon|^2,|\nabla\rho_t^\varepsilon||\nabla\log\rho_t^\varepsilon|^2\in L^\infty_{loc}((0,1),L^1(\mathfrak{m}))$ by Lemma 5.3.1 and the aforementioned properties of χ_R imply that (5.3.8b) vanishes. Therefore

$$\begin{aligned}\int\langle\nabla\rho_{t_1}^\varepsilon,\nabla\vartheta_{t_1}^\varepsilon\rangle\mathrm{d}\mathfrak{m}\quad&-\quad\int\langle\nabla\rho_{t_0}^\varepsilon,\nabla\vartheta_{t_0}^\varepsilon\rangle\mathrm{d}\mathfrak{m}\\&=\quad\iint_{t_0}^{t_1}\frac12\Delta\rho_t^\varepsilon|\nabla\vartheta_t^\varepsilon|^2-\rho_t^\varepsilon\langle\nabla\vartheta_t^\varepsilon,\nabla\Delta\vartheta_t^\varepsilon\rangle\mathrm{d}t\mathrm{d}\mathfrak{m}\\&\qquad+\iint_{t_0}^{t_1}\frac{\varepsilon^2}{8}\Delta\rho_t^\varepsilon|\nabla\log\rho_t^\varepsilon|^2-\frac{\varepsilon^2}{4}\rho_t^\varepsilon\langle\nabla\log\rho_t^\varepsilon,\nabla\Delta\log\rho_t^\varepsilon\rangle\mathrm{d}t\mathrm{d}\mathfrak{m}\\&\overset{(5.3.1)}{=}\iint_{t_0}^{t_1}\rho_t^\varepsilon\,\mathrm{d}t\mathrm{d}\big(\mathbf{\Gamma}_2(\vartheta_t^\varepsilon)+\tfrac{\varepsilon^2}{4}\mathbf{\Gamma}_2(\log(\rho_t^\varepsilon))\big).\end{aligned}$$

To obtain C^2 regularity and that the formula for the second derivative is valid for any $t\in(0,1)$ it is sufficient to check the continuity of $t\mapsto\int\rho_t^\varepsilon\,\mathrm{d}\mathbf{\Gamma}_2(\vartheta_t^\varepsilon)$, as the continuity of the other term

follows along the same lines. To this aim, by Lemma 5.3.1 and (5.3.1) we deduce the continuity of $t \mapsto \int \rho_t^\varepsilon \, \mathrm{d}\mathbf{\Delta}(\vartheta_t^\varepsilon)$; for $t \mapsto \rho_t^\varepsilon \langle \nabla \vartheta_t^\varepsilon, \nabla \Delta \vartheta_t^\varepsilon \rangle$ use (5.3.5) and argue in the same way as in the proof of Lemma 5.3.1. □

As a first consequence of the formulas just obtained, we show that some quantities remain bounded as $\varepsilon \downarrow 0$:

Lemma 5.3.4 (Bounded quantities)**.** *With the same assumptions and notations of Setting 5.1.3 we have*

$$\sup_{\varepsilon \in (0,1),\, t \in [0,1]} \int \mathsf{d}^2(\cdot, \bar{x}) \rho_t^\varepsilon \mathrm{d}\mathfrak{m} < \infty, \tag{5.3.9a}$$

$$\sup_{\varepsilon \in (0,1),\, t \in [0,1]} |H(\mu_t^\varepsilon \,|\, \mathfrak{m})| < \infty, \tag{5.3.9b}$$

$$\sup_{\varepsilon \in (0,1)} \iint_0^1 \Big(|\nabla \vartheta_t^\varepsilon|^2 + \varepsilon^2 |\nabla \log \rho_t^\varepsilon|^2 \Big) \rho_t^\varepsilon \, \mathrm{d}t \, \mathrm{d}\mathfrak{m} < \infty, \tag{5.3.9c}$$

$\bar{x} \in \mathrm{X}$ *being completely arbitrary in* (5.3.9a)*, and for any* $\delta \in (0, \frac{1}{2})$

$$\sup_{\varepsilon \in (0,1)} \iint_\delta^{1-\delta} \Big(|\mathrm{Hess}(\vartheta_t^\varepsilon)|_{\mathsf{HS}}^2 + \varepsilon^2 |\mathrm{Hess}(\log \rho_t^\varepsilon)|_{\mathsf{HS}}^2 \Big) \rho_t^\varepsilon \, \mathrm{d}t \, \mathrm{d}\mathfrak{m} < \infty, \tag{5.3.10a}$$

$$\sup_{\varepsilon \in (0,1)} \iint_\delta^{1-\delta} \Big(|\Delta \vartheta_t^\varepsilon|^2 + \varepsilon^2 |\Delta \log \rho_t^\varepsilon|^2 \Big) \rho_t^\varepsilon \, \mathrm{d}t \, \mathrm{d}\mathfrak{m} < \infty. \tag{5.3.10b}$$

proof We start with (5.3.9a) and observe that it immediately follows from (5.2.5). As regards (5.3.9b), argue as in the proof of Proposition 5.3.3: taking into account (5.1.2) and the notations used therein with $\mu = \mu_t^\varepsilon$ and $V = \mathsf{d}^2(\cdot, \bar{x})$ (where $\bar{x}$ is any point of X), $H(\mu_t^\varepsilon \,|\, \tilde{\mathfrak{m}})$ is bounded from below uniformly in $t \in [0,1]$ and $\varepsilon \in (0,1)$, whereas by Proposition 5.2.2 $H(\mu_t^\varepsilon \,|\, \tilde{\mathfrak{m}})$ is also bounded from above uniformly in $t \in [0,1]$ and $\varepsilon \in (0,1)$. Moreover, (5.3.9a) ensures that $\int V \mathrm{d}\mu_t^\varepsilon$ is uniformly bounded too, whence the desired (5.3.9b).

Let us now pass to (5.3.9c) and observe that Proposition 5.2.3 together with (5.3.9a) grants

$$\sup_{\varepsilon \in (0,1)} \iint_{\frac{1}{2}}^1 |\nabla \varphi_t^\varepsilon|^2 \rho_t^\varepsilon \, \mathrm{d}t \, \mathrm{d}\mathfrak{m} + \iint_0^{\frac{1}{2}} |\nabla \psi_t^\varepsilon|^2 \rho_t^\varepsilon \, \mathrm{d}t \, \mathrm{d}\mathfrak{m} < \infty. \tag{5.3.11}$$

As a second step, notice that (5.3.6a) gives

$$\begin{aligned} \iint_0^{\frac{1}{2}} |\nabla \varphi_t^\varepsilon|^2 \rho_t^\varepsilon \, \mathrm{d}t \, \mathrm{d}\mathfrak{m} &= \iint_0^{\frac{1}{2}} |\nabla \psi_t^\varepsilon|^2 \rho_t^\varepsilon \, \mathrm{d}t \, \mathrm{d}\mathfrak{m} - 2\varepsilon \int_0^{\frac{1}{2}} \frac{\mathrm{d}}{\mathrm{d}t} H(\mu_t^\varepsilon \,|\, \mathfrak{m}) \, \mathrm{d}t \\ &= \iint_0^{\frac{1}{2}} |\nabla \psi_t^\varepsilon|^2 \rho_t^\varepsilon \, \mathrm{d}t \, \mathrm{d}\mathfrak{m} + 2\varepsilon \big(H(\mu_0 \,|\, \mathfrak{m}) - H(\mu_t^\varepsilon \,|\, \mathfrak{m}) \big) \end{aligned}$$

so that taking into account (5.3.9b) and (5.3.11) we see that the right hand side is uniformly bounded for $\varepsilon \in (0,1)$. Using again (5.3.11) we deduce that

$$\sup_{\varepsilon \in (0,1)} \iint_0^1 |\nabla \varphi_t^\varepsilon|^2 \rho_t^\varepsilon \, \mathrm{d}t \, \mathrm{d}\mathfrak{m} < \infty.$$

A symmetric argument provides the analogous bound for (ψ_t^ε) and thus recalling that $\vartheta_t^\varepsilon = \frac{1}{2}(\psi_t^\varepsilon - \varphi_t^\varepsilon)$ and $\varepsilon \log \rho_t^\varepsilon = \psi_t^\varepsilon + \varphi_t^\varepsilon$ we obtain (5.3.9c).

Now use the fact that $\vartheta_t^\varepsilon = -\varphi_t^\varepsilon + \frac{\varepsilon}{2}\log \rho_t^\varepsilon$ in conjunction with (5.3.6a) to get

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t} H(\mu_t^\varepsilon \,|\, \mathfrak{m})_{|t=\delta} &= -\int \langle \nabla \rho_\delta^\varepsilon, \nabla \varphi_\delta^\varepsilon \rangle \mathrm{d}\mathfrak{m} + \frac{\varepsilon}{2}\int \langle \nabla \rho_\delta^\varepsilon, \nabla \log \rho_\delta^\varepsilon \rangle \mathrm{d}\mathfrak{m} \\ &= \int \rho_\delta^\varepsilon \Delta \varphi_\delta^\varepsilon \mathrm{d}\mathfrak{m} + \frac{\varepsilon}{2}\int \frac{|\nabla \rho_\delta^\varepsilon|^2}{\rho_\delta^\varepsilon}\mathrm{d}\mathfrak{m} \geq \int \rho_\delta^\varepsilon \Delta \varphi_\delta^\varepsilon \mathrm{d}\mathfrak{m}.\end{aligned}$$

Recalling the lower bound (5.2.10b) and (5.3.9a), we get that for some constant C_δ independent on ε it holds

$$\frac{\mathrm{d}}{\mathrm{d}t} H(\mu_t^\varepsilon \,|\, \mathfrak{m})_{|t=\delta} \geq -C_\delta \qquad \forall \varepsilon \in (0,1)$$

and an analogous argument starting from $\vartheta_t^\varepsilon = \psi_t^\varepsilon - \frac{\varepsilon}{2}\log \rho_t^\varepsilon$ yields $\frac{\mathrm{d}}{\mathrm{d}t} H(\mu_t^\varepsilon \,|\, \mathfrak{m})_{|t=1-\delta} \leq C_\delta$ for every $\varepsilon \in (0,1)$. Therefore

$$\sup_{\varepsilon \in (0,1)} \int_\delta^{1-\delta} \frac{\mathrm{d}^2}{\mathrm{d}t^2} H(\mu_t^\varepsilon \,|\, \mathfrak{m}) = \sup_{\varepsilon \in (0,1)} \left(\frac{\mathrm{d}}{\mathrm{d}t} H(\mu_t^\varepsilon \,|\, \mathfrak{m})_{|t=1-\delta} - \frac{\mathrm{d}}{\mathrm{d}t} H(\mu_t^\varepsilon \,|\, \mathfrak{m})_{|t=\delta} \right) < \infty.$$

The bounds (5.3.10a) and (5.3.10b) then come from this last inequality used in conjunction with (5.3.9c) and the weighted Bochner inequalities (5.3.2a) and (5.3.2b) respectively. □

With the help of the previous lemma we can now prove that some crucial quantities vanish in the limit $\varepsilon \downarrow 0$; as we shall see in the proof of our main theorem 7.1.2, this is what we will need to prove that the acceleration of the entropic interpolations goes to 0 (in a suitable sense) as ε goes to zero.

Lemma 5.3.5 (Vanishing quantities)**.** *With the same assumptions and notations of Setting 5.1.3, for any $\delta \in (0, \frac{1}{2})$ we have*

$$\lim_{\varepsilon \downarrow 0} \varepsilon^2 \iint_\delta^{1-\delta} \rho_t^\varepsilon |\Delta \log \rho_t^\varepsilon| \,\mathrm{d}t\,\mathrm{d}\mathfrak{m} = 0, \tag{5.3.12a}$$

$$\lim_{\varepsilon \downarrow 0} \varepsilon^2 \iint_\delta^{1-\delta} \rho_t^\varepsilon |\nabla \log \rho_t^\varepsilon|^2 \,\mathrm{d}t\,\mathrm{d}\mathfrak{m} = 0, \tag{5.3.12b}$$

$$\lim_{\varepsilon \downarrow 0} \varepsilon^2 \iint_\delta^{1-\delta} \rho_t^\varepsilon |\Delta \log \rho_t^\varepsilon||\nabla \log \rho_t^\varepsilon| \,\mathrm{d}t\,\mathrm{d}\mathfrak{m} = 0, \tag{5.3.12c}$$

$$\lim_{\varepsilon \downarrow 0} \varepsilon^2 \iint_\delta^{1-\delta} \rho_t^\varepsilon |\nabla \log \rho_t^\varepsilon|^3 \,\mathrm{d}t\,\mathrm{d}\mathfrak{m} = 0. \tag{5.3.12d}$$

proof For (5.3.12a) we notice that

$$\varepsilon^2 \iint_\delta^{1-\delta} \rho_t^\varepsilon |\Delta \log \rho_t^\varepsilon| \,\mathrm{d}t\,\mathrm{d}\mathfrak{m} \leq \varepsilon\sqrt{1-2\delta}\sqrt{\varepsilon^2 \iint_\delta^{1-\delta} \rho_t^\varepsilon |\Delta \log \rho_t^\varepsilon|^2 \,\mathrm{d}t\,\mathrm{d}\mathfrak{m}}$$

and the fact that, by (5.3.10b), the last square root is uniformly bounded in $\varepsilon \in (0,1)$.

For (5.3.12b) we start observing that Lemma 5.3.6 below applies to ρ_t^ε, because by Proposition 5.1.4 $\rho_t^\varepsilon \in \mathrm{Test(X)} \cap L^1(\mathfrak{m})$ and

$$\Delta \rho_t^\varepsilon = f_t^\varepsilon \Delta g_t^\varepsilon + g_t^\varepsilon \Delta f_t^\varepsilon + 2\langle \nabla f_t^\varepsilon, \nabla g_t^\varepsilon \rangle \in L^1(\mathfrak{m}).$$

Hence, from the identity $\rho_t^\varepsilon|\nabla\log\rho_t^\varepsilon|^2 = -\rho_t^\varepsilon\Delta\log\rho_t^\varepsilon + \Delta\rho_t^\varepsilon$ and the fact that $\int\Delta\rho_t^\varepsilon\,\mathrm{d}\mathfrak{m} = 0$ we get

$$\varepsilon^2\iint_\delta^{1-\delta}\rho_t^\varepsilon|\nabla\log\rho_t^\varepsilon|^2\,\mathrm{d}t\,\mathrm{d}\mathfrak{m} = -\varepsilon^2\iint_\delta^{1-\delta}\rho_t^\varepsilon\Delta\log\rho_t^\varepsilon\,\mathrm{d}t\,\mathrm{d}\mathfrak{m} \le \varepsilon^2\iint_\delta^{1-\delta}\rho_t^\varepsilon|\Delta\log\rho_t^\varepsilon|\,\mathrm{d}t\,\mathrm{d}\mathfrak{m}$$

and then conclude by (5.3.12a).

For (5.3.12c) we observe that

$$\begin{aligned}\varepsilon^2\iint_\delta^{1-\delta}&\rho_t^\varepsilon|\Delta\log\rho_t^\varepsilon||\nabla\log\rho_t^\varepsilon|\,\mathrm{d}t\,\mathrm{d}\mathfrak{m}\\ &\le\sqrt{\varepsilon^2\iint_\delta^{1-\delta}\rho_t^\varepsilon|\Delta\log\rho_t^\varepsilon|^2\,\mathrm{d}t\,\mathrm{d}\mathfrak{m}}\sqrt{\varepsilon^2\iint_\delta^{1-\delta}\rho_t^\varepsilon|\nabla\log\rho_t^\varepsilon|^2\,\mathrm{d}t\,\mathrm{d}\mathfrak{m}},\end{aligned}$$

and use the fact that the first square root in the right hand side is bounded (by (5.3.10b)) and the second one goes to 0 (by (5.3.12b)).

To prove (5.3.12d) we start again from the identity $\rho_t^\varepsilon|\nabla\log\rho_t^\varepsilon|^2 = -\rho_t^\varepsilon\Delta\log\rho_t^\varepsilon + \Delta\rho_t^\varepsilon$ to get

$$\iint_\delta^{1-\delta}\rho_t^\varepsilon|\nabla\log\rho_t^\varepsilon|^3\,\mathrm{d}t\,\mathrm{d}\mathfrak{m} = -\iint_\delta^{1-\delta}\rho_t^\varepsilon\Delta\log\rho_t^\varepsilon|\nabla\log\rho_t^\varepsilon|\,\mathrm{d}t\,\mathrm{d}\mathfrak{m} + \iint_\delta^{1-\delta}\Delta\rho_t^\varepsilon|\nabla\log\rho_t^\varepsilon|\,\mathrm{d}t\,\mathrm{d}\mathfrak{m}.$$

After a multiplication by ε^2 we see that the first integral on the right-hand side vanishes as $\varepsilon\downarrow 0$ thanks to (5.3.12c). For the second we start noticing that an application of the dominated convergence theorem ensures that

$$\iint_\delta^{1-\delta}\Delta\rho_t^\varepsilon|\nabla\log\rho_t^\varepsilon|\,\mathrm{d}t\,\mathrm{d}\mathfrak{m} = \lim_{\eta\downarrow 0}\iint_\delta^{1-\delta}\Delta\rho_t^\varepsilon\sqrt{\eta+|\nabla\log\rho_t^\varepsilon|^2}\,\mathrm{d}t\,\mathrm{d}\mathfrak{m}, \tag{5.3.13}$$

then we observe that for every $\eta > 0$ the map $z\mapsto\sqrt{\eta+z}$ is in $C^1([0,\infty))$ and since $|\nabla\log\rho_t^\varepsilon|^2\in W^{1,2}(\mathrm{X})$ we deduce that $\sqrt{\eta+|\nabla\log\rho_t^\varepsilon|^2}\in W^{1,2}(\mathrm{X})$ as well. Thus by the chain rule for gradients and the Leibniz rule (1.2.21) it holds

$$\begin{aligned}\Big|\iint_\delta^{1-\delta}&\Delta\rho_t^\varepsilon\sqrt{\eta+|\nabla\log\rho_t^\varepsilon|^2}\,\mathrm{d}t\,\mathrm{d}\mathfrak{m}\Big|\\ &=\Big|\iint_\delta^{1-\delta}\frac{\rho_t^\varepsilon}{2\sqrt{\eta+|\nabla\log\rho_t^\varepsilon|^2}}\langle\nabla\log\rho_t^\varepsilon,\nabla|\nabla\log\rho_t^\varepsilon|^2\rangle\,\mathrm{d}t\,\mathrm{d}\mathfrak{m}\Big|\\ &=\Big|\iint_\delta^{1-\delta}\frac{\rho_t^\varepsilon}{\sqrt{\eta+|\nabla\log\rho_t^\varepsilon|^2}}\mathrm{Hess}(\log\rho_t^\varepsilon)(\nabla\log\rho_t^\varepsilon,\nabla\log\rho_t^\varepsilon)\,\mathrm{d}t\,\mathrm{d}\mathfrak{m}\Big|\\ &\le\iint_\delta^{1-\delta}\rho_t^\varepsilon|\mathrm{Hess}(\log\rho_t^\varepsilon)|_{\mathsf{HS}}\,|\nabla\log\rho_t^\varepsilon|\,\mathrm{d}t\,\mathrm{d}\mathfrak{m}\end{aligned}$$

and being this true for any $\eta > 0$, from (5.3.13) we obtain

$$\begin{aligned}\varepsilon^2\Big|\iint_\delta^{1-\delta}\Delta\rho_t^\varepsilon|\nabla\log\rho_t^\varepsilon|\,\mathrm{d}t\,\mathrm{d}\mathfrak{m}\Big| &\le\varepsilon^2\iint_\delta^{1-\delta}\rho_t^\varepsilon|\mathrm{Hess}(\log\rho_t^\varepsilon)|_{\mathsf{HS}}|\nabla\log\rho_t^\varepsilon|\,\mathrm{d}t\,\mathrm{d}\mathfrak{m}\\ &\le\sqrt{\varepsilon^2\iint_\delta^{1-\delta}\rho_t^\varepsilon|\mathrm{Hess}(\log\rho_t^\varepsilon)|_{\mathsf{HS}}^2\,\mathrm{d}t\,\mathrm{d}\mathfrak{m}}\\ &\quad\times\sqrt{\varepsilon^2\iint_\delta^{1-\delta}\rho_t^\varepsilon|\nabla\log\rho_t^\varepsilon|^2\,\mathrm{d}t\,\mathrm{d}\mathfrak{m}}.\end{aligned}$$

In this last expression the first square root is uniformly bounded in $\varepsilon \in (0,1)$ by (5.3.10a), while the second one vanishes as $\varepsilon \downarrow 0$ thanks to (5.3.12b). □

Lemma 5.3.6. *Let* $(\mathrm{X}, \mathsf{d}, \mathfrak{m})$ *be a* $\mathsf{RCD}(K,\infty)$ *space with* $K \in \mathbb{R}$ *endowed with a Borel non-negative measure* $\mathfrak{m}$ *which is finite on bounded sets and* $h \in D(\Delta) \cap L^1(\mathfrak{m})$ *with* $\Delta h \in L^1(\mathfrak{m})$. *Then*

$$\int \Delta h \mathrm{d}\mathfrak{m} = 0$$

proof Let $\bar{x} \in \mathrm{X}$, $R > 0$ and $\chi_R \in \mathrm{Test}^\infty(\mathrm{X})$ be a cut-off function built as in Lemma 2.2.1 such that $\chi_R \equiv 1$ in $B_R(\bar{x})$ and $\chi_R \equiv 0$ in $\mathrm{X} \setminus B_{R+1}(\bar{x})$. Then

$$\begin{aligned} \left| \int \chi_R \Delta h \mathrm{d}\mathfrak{m} \right| &= \left| \int \Delta \chi_R h \mathrm{d}\mathfrak{m} \right| = \left| \int_{\mathrm{X} \setminus B_{R+1}(\bar{x})} \Delta \chi_R h \mathrm{d}\mathfrak{m} \right| \\ &\leq \|\Delta \chi_R\|_{L^\infty(\mathfrak{m})} \int_{\mathrm{X} \setminus B_{R+1}(\bar{x})} h \mathrm{d}\mathfrak{m} \end{aligned} \tag{5.3.14}$$

and by Lemma 2.2.1 $\|\Delta \chi_R\|_{L^\infty(\mathfrak{m})}$ is uniformly bounded in R, while

$$\int_{\mathrm{X} \setminus B_{R+1}(\bar{x})} h \mathrm{d}\mathfrak{m} \to 0 \qquad \text{as } R \to \infty$$

since $h \in L^1(\mathfrak{m})$. As a consequence, the last term in (5.3.14) vanishes as $R \to \infty$, while the left hand side converges to $|\int \Delta h \mathrm{d}\mathfrak{m}|$ by the dominated convergence theorem and this gives us the conclusion. □

5.4 A Benamou-Brenier formula for the entropic cost

We conclude the section with an interesting application of the 'PDE' viewpoint provided in Proposition 5.1.4: a Benamou-Brenier-type formula for the Schrödinger problem. To this aim let us first define

$$\begin{aligned} \varphi_0^\varepsilon &:= \varepsilon \log(f^\varepsilon) \qquad \text{in } \operatorname{supp}(\mu_0) \\ \psi_1^\varepsilon &:= \varepsilon \log(g^\varepsilon) \qquad \text{in } \operatorname{supp}(\mu_1) \end{aligned} \tag{5.4.1}$$

and, recalling the definition of the entropic cost given in Chapter 4, introduce the slowed down entropic cost $\mathscr{I}_\varepsilon$ relative to $\mathsf{R}^{\varepsilon/2}$ as

$$\mathscr{I}_\varepsilon(\mu_0, \mu_1) := \inf_{\boldsymbol{\gamma}' \in \mathrm{Adm}(\mu_0,\mu_1)} H(\boldsymbol{\gamma}' \,|\, \mathsf{R}^{\varepsilon/2}).$$

With this said we are now in the position to give a threefold dynamic representation of the entropic cost.

Proposition 5.4.1. *With the same assumptions and notations as in Setting 5.1.3 and here above, for any* $\varepsilon > 0$ *the following holds:*

$$\begin{aligned} \varepsilon \mathscr{I}_\varepsilon(\mu_0, \mu_1) &= \frac{\varepsilon}{2}\Big(H(\mu_0 \,|\, \mathfrak{m}) + H(\mu_1 \,|\, \mathfrak{m})\Big) + \iint_0^1 \Big(\frac{|\nabla \vartheta_t^\varepsilon|^2}{2} + \frac{\varepsilon^2}{8}|\nabla \log \rho_t^\varepsilon|^2\Big)\rho_t^\varepsilon \mathrm{d}t\mathrm{d}\mathfrak{m} \\ &= \varepsilon H(\mu_0 \,|\, \mathfrak{m}) + \iint_0^1 \frac{|\nabla \psi_t^\varepsilon|^2}{2} \rho_t^\varepsilon \mathrm{d}t\mathrm{d}\mathfrak{m} \\ &= \varepsilon H(\mu_1 \,|\, \mathfrak{m}) + \iint_0^1 \frac{|\nabla \varphi_t^\varepsilon|^2}{2} \rho_t^\varepsilon \mathrm{d}t\mathrm{d}\mathfrak{m}. \end{aligned} \tag{5.4.2}$$

proof Fix $\varepsilon > 0$ and let us prove the first identity in (5.4.2). To this aim fix $\bar{x} \in \mathrm{X}$, $R > 0$ and let $\chi_R \in \mathrm{Test}^\infty(\mathrm{X})$ be a cut-off function as in Lemma 2.2.1 with support in $B_{R+1}(\bar{x})$ such that $\chi_R \equiv 1$ on $B_R(\bar{x})$; using the notations of Setting 5.1.3, by Proposition 5.1.4 $t \mapsto \int \chi_R \vartheta_t^\varepsilon \rho_t^\varepsilon \,\mathrm{d}\mathfrak{m}$ belongs to $AC_{loc}((0,1))$ with

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\int \chi_R \vartheta_t^\varepsilon \rho_t^\varepsilon \,\mathrm{d}\mathfrak{m} &= \int \chi_R\Big(-\frac{|\nabla \vartheta_t^\varepsilon|^2}{2} - \frac{\varepsilon^2}{4}\Delta \log \rho_t^\varepsilon - \frac{\varepsilon^2}{8}|\nabla \log \rho_t^\varepsilon|^2\Big)\rho_t^\varepsilon \,\mathrm{d}\mathfrak{m} \\ &\quad - \int \chi_R \vartheta_t^\varepsilon \mathrm{div}(\rho_t^\varepsilon \nabla \vartheta_t^\varepsilon)\mathrm{d}\mathfrak{m} \qquad \text{a.e. } t \in (0,1).\end{aligned}$$

Integration by parts formula and integration in time on $[\delta, 1-\delta]$ then yield

$$\begin{aligned}&\int \chi_R \vartheta_{1-\delta}^\varepsilon \rho_{1-\delta}^\varepsilon \,\mathrm{d}\mathfrak{m} - \int \chi_R \vartheta_\delta^\varepsilon \rho_\delta^\varepsilon \,\mathrm{d}\mathfrak{m} \\ &\qquad = \iint_\delta^{1-\delta} \chi_R\Big(\frac{|\nabla \vartheta_t^\varepsilon|^2}{2} + \frac{\varepsilon^2}{8}|\nabla \log \rho_t^\varepsilon|^2\Big)\rho_t^\varepsilon \,\mathrm{d}\mathfrak{m} + \iint_\delta^{1-\delta} \langle \nabla \chi_R, \nabla \rho_t^\varepsilon\rangle \mathrm{d}\mathfrak{m} \\ &\qquad\quad + \iint_\delta^{1-\delta} \vartheta_t^\varepsilon \langle \nabla \chi_R, \nabla \vartheta_t^\varepsilon\rangle \rho_t^\varepsilon \mathrm{d}\mathfrak{m} \qquad \text{a.e. } t \in (0,1)\end{aligned}$$

and the arguments by dominated convergence theorem already explained in the proof of Proposition 5.3.3 enable the passage to the limit as $R \to \infty$, thus getting

$$\int \vartheta_{1-\delta}^\varepsilon \rho_{1-\delta}^\varepsilon \,\mathrm{d}\mathfrak{m} - \int \vartheta_\delta^\varepsilon \rho_\delta^\varepsilon \,\mathrm{d}\mathfrak{m} = \iint_\delta^{1-\delta} \Big(\frac{|\nabla \vartheta_t^\varepsilon|^2}{2} + \frac{\varepsilon^2}{8}|\nabla \log \rho_t^\varepsilon|^2\Big)\rho_t^\varepsilon \,\mathrm{d}\mathfrak{m}.$$

Now let $\delta \downarrow 0$: convergence of the right-hand side is trivial by monotonicity. For the left-hand side consider $t \mapsto \int \vartheta_t^\varepsilon \rho_t^\varepsilon \,\mathrm{d}\mathfrak{m}$, use the identity $\vartheta_t^\varepsilon = \psi_t^\varepsilon - \frac{\varepsilon}{2}\log \rho_t^\varepsilon$ and observe that

$$t \mapsto \int \psi_t^\varepsilon \rho_t^\varepsilon \,\mathrm{d}\mathfrak{m} \quad \text{and} \quad t \mapsto -\frac{\varepsilon}{2}H(\mu_t^\varepsilon \,|\, \mathfrak{m})$$

are both continuous at $t = 0$, the former by Lemma 5.3.1 and the latter by Proposition 5.3.3. This implies that

$$\lim_{\delta \downarrow 0} \int \vartheta_\delta^\varepsilon \rho_\delta^\varepsilon \,\mathrm{d}\mathfrak{m} = \int \psi_0^\varepsilon \rho_0 \,\mathrm{d}\mathfrak{m} - \frac{\varepsilon}{2}H(\mu_0 \,|\, \mathfrak{m}).$$

The same argument with the identity $\vartheta_t^\varepsilon = -\varphi_t^\varepsilon + \frac{\varepsilon}{2}\log \rho_t^\varepsilon$ allows us to handle $\int \vartheta_{1-\delta}^\varepsilon \rho_{1-\delta}^\varepsilon \,\mathrm{d}\mathfrak{m}$ too, so that

$$\begin{aligned}-\int \psi_0^\varepsilon \rho_0 \,\mathrm{d}\mathfrak{m} - \int \varphi_1^\varepsilon \rho_1 \,\mathrm{d}\mathfrak{m} + \frac{\varepsilon}{2}\Big(H(\mu_0 \,|\, \mathfrak{m}) + H(\mu_1 \,|\, \mathfrak{m})\Big) \\ = \iint_0^1 \Big(\frac{|\nabla \vartheta_t^\varepsilon|^2}{2} + \frac{\varepsilon^2}{8}|\nabla \log \rho_t^\varepsilon|^2\Big)\rho_t^\varepsilon \,\mathrm{d}\mathfrak{m}.\end{aligned}$$

Thanks to the identity $\varphi_0^\varepsilon + \psi_0^\varepsilon = \varepsilon \log \rho_0$ in $\mathrm{supp}(\mu_0)$ and the analogous one in $t = 1$, this is in turn equivalent to

$$\begin{aligned}\int \varphi_0^\varepsilon \rho_0 \,\mathrm{d}\mathfrak{m} + \int \psi_1^\varepsilon \rho_1 \,\mathrm{d}\mathfrak{m} - \frac{\varepsilon}{2}\Big(H(\mu_0 \,|\, \mathfrak{m}) + H(\mu_1 \,|\, \mathfrak{m})\Big) \\ = \iint_0^1 \Big(\frac{|\nabla \vartheta_t^\varepsilon|^2}{2} + \frac{\varepsilon^2}{8}|\nabla \log \rho_t^\varepsilon|^2\Big)\rho_t^\varepsilon \,\mathrm{d}\mathfrak{m}\end{aligned}$$

and now it is sufficient to observe that by (5.4.1)

$$\varepsilon\mathscr{I}_\varepsilon(\mu_0,\mu_1)=\varepsilon H\big(f^\varepsilon\otimes g^\varepsilon\mathsf{R}^{\varepsilon/2}\,|\,\mathsf{R}^{\varepsilon/2}\big)=\int\varphi_0^\varepsilon\mathrm{d}\mu_0+\int\psi_1^\varepsilon\mathrm{d}\mu_1.$$

For the second and third identities in (5.4.2), the argument closely follows the one we have just presented. Indeed, it is just a matter of computation to rewrite the continuity equation solved by $(\rho_t^\varepsilon,\vartheta_t^\varepsilon)$ as forward and backward Fokker-Planck equations with velocity fields given by $\nabla\psi_t^\varepsilon$ and $\nabla\varphi_t^\varepsilon$ respectively, i.e.

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\rho_t^\varepsilon+\mathrm{div}(\rho_t^\varepsilon\nabla\psi_t^\varepsilon)&=\frac{\varepsilon}{2}\Delta\rho_t^\varepsilon\\ -\frac{\mathrm{d}}{\mathrm{d}t}\rho_t^\varepsilon+\mathrm{div}(\rho_t^\varepsilon\nabla\varphi_t^\varepsilon)&=\frac{\varepsilon}{2}\Delta\rho_t^\varepsilon\end{aligned}$$

where the time derivatives are meant as in Proposition 5.1.4. Therefore, arguing as above it is not difficult to see that

$$\int\varphi_1^\varepsilon\mathrm{d}\mu_1-\int\varphi_0^\varepsilon\mathrm{d}\mu_0=-\iint_0^1\frac{|\nabla\varphi_t^\varepsilon|^2}{2}\rho_t^\varepsilon\mathrm{d}t\mathrm{d}\mathfrak{m}$$

and an analogous identity holds true for ψ_t^ε. Finally using the identities $\varphi_0^\varepsilon+\psi_0^\varepsilon=\varepsilon\log\rho_0$ in $\mathrm{supp}(\mu_0)$ and $\varphi_1^\varepsilon+\psi_1^\varepsilon=\varepsilon\log\rho_1$ in $\mathrm{supp}(\mu_1)$, the entropic cost can be rewritten as

$$\varepsilon\mathscr{I}_\varepsilon(\mu_0,\mu_1)=\varepsilon H(\mu_0\,|\,\mathfrak{m})+\int\psi_1^\varepsilon\mathrm{d}\mu_1-\int\psi_0^\varepsilon\mathrm{d}\mu_0=\varepsilon H(\mu_1\,|\,\mathfrak{m})+\int\varphi_0^\varepsilon\mathrm{d}\mu_0-\int\varphi_1^\varepsilon\mathrm{d}\mu_1,$$

whence the conclusion. □

Thus the entropic cost can be expressed as an action functional in three different ways: in terms of $(\rho_t^\varepsilon,\vartheta_t^\varepsilon)$, $(\rho_t^\varepsilon,\psi_t^\varepsilon)$ and $(\rho_t^\varepsilon,\varphi_t^\varepsilon)$. As we have just seen, three different 'PDEs' are associated to these couples, so that three different minimization problems can be introduced, namely

$$\begin{aligned}&\frac{\varepsilon}{2}\Big(H(\mu_0\,|\,\mathfrak{m})+H(\mu_1\,|\,\mathfrak{m})\Big)+\inf_{(\nu,v)}\Big\{\iint_0^1\Big(\frac{|v_t|^2}{2}+\frac{\varepsilon^2}{8}|\nabla\log\nu_t|^2\Big)\nu_t\mathrm{d}t\mathrm{d}\mathfrak{m}\Big\}\\ &\varepsilon H(\mu_0\,|\,\mathfrak{m})+\inf_{(\nu,v)}\iint_0^1\frac{|v_t|^2}{2}\nu_t\mathrm{d}t\mathrm{d}\mathfrak{m}\\ &\varepsilon H(\mu_1\,|\,\mathfrak{m})+\inf_{(\nu,v)}\iint_0^1\frac{|v_t|^2}{2}\nu_t\mathrm{d}t\mathrm{d}\mathfrak{m}\end{aligned}\tag{5.4.3}$$

the infima being taken among all suitable solutions of the corresponding 'PDEs'. In line with the smooth theory, a natural guess is that all the infima coincide and they are attained if and only if $(\nu_t,v_t)=(\rho_t^\varepsilon,\nabla\vartheta_t^\varepsilon)$, $(\nu_t,v_t)=(\rho_t^\varepsilon,\nabla\psi_t^\varepsilon)$ and $(\nu_t,v_t)=(\rho_t^\varepsilon,\nabla\varphi_t^\varepsilon)$ respectively.

We shall now investigate the first minimization problem and for sake of simplicity we will assume the space $(\mathrm{X},\mathsf{d},\mathfrak{m})$ to be compact and $\rho_0,\rho_1\in\mathrm{Test}^\infty_{>0}(\mathrm{X})$. In this framework, from Theorem 5.1.2 we know that $f^\varepsilon,g^\varepsilon\in\mathrm{Test}^\infty_{>0}(\mathrm{X})$ for all $\varepsilon>0$ as well; arguing as in Proposition 5.1.4 together with the compactness of X, which ensures that all the objects defined in Setting 5.1.3 belong to $\mathrm{Test}^\infty_{>0}(\mathrm{X})$, it is not difficult to see that all the curves

$$(f_t^\varepsilon),(g_t^\varepsilon),(\rho_t^\varepsilon),(\varphi_t^\varepsilon),(\psi_t^\varepsilon),(\vartheta_t^\varepsilon)\in AC([0,1],W^{1,2}(\mathrm{X}))\tag{5.4.4}$$

after having defined $\varphi_0^\varepsilon := \varepsilon \log(f^\varepsilon)$, $\psi_1^\varepsilon := \varepsilon \log(g^\varepsilon)$ (unlike (5.4.1), now these definitions make sense on the whole X) and set $\vartheta_t^\varepsilon := (\psi_t^\varepsilon - \varphi_t^\varepsilon)/2$ on the whole $[0,1]$. In particular this means that the time derivatives of $(f_t^\varepsilon), (g_t^\varepsilon), (\rho_t^\varepsilon), (\varphi_t^\varepsilon), (\psi_t^\varepsilon), (\vartheta_t^\varepsilon)$, explicitly written in Proposition 5.1.4, have to be understood as the strong $W^{1,2}$-limit of the incremental ratios. Furthermore

$$\sup_{t\in[0,1]} \|h_t^\varepsilon\|_{L^\infty} + \mathrm{Lip}(h_t^\varepsilon) + \|\Delta h_t^\varepsilon\|_{W^{1,2}} + \|\Delta h_t^\varepsilon\|_{L^\infty} < \infty,$$

where (h_t^ε) is equal to any of $(f_t^\varepsilon), (g_t^\varepsilon), (\rho_t^\varepsilon), (\varphi_t^\varepsilon), (\psi_t^\varepsilon), (\vartheta_t^\varepsilon)$. The details of the proof can be found in [63].

We shall also make use of the following simple lemma valid on general metric measure spaces.

Lemma 5.4.2. *Let $(\mathrm{Y}, \mathsf{d}_\mathrm{Y}, \mathfrak{m}_\mathrm{Y})$ be a complete separable metric measure space endowed with a non-negative measure $\mathfrak{m}_\mathrm{Y}$ which is finite on bounded sets and assume that $W^{1,2}(\mathrm{Y})$ is separable. Let $\boldsymbol{\pi}$ be a test plan and $f \in W^{1,2}(\mathrm{Y})$. Then $t \mapsto \int f \circ \mathrm{e}_t \,\mathrm{d}\boldsymbol{\pi}$ is absolutely continuous and*

$$\Big|\frac{\mathrm{d}}{\mathrm{d}t}\int f\circ \mathrm{e}_t\,\mathrm{d}\boldsymbol{\pi}\Big| \le \int |\mathrm{d}f|(\gamma_t)|\dot\gamma_t|\,\mathrm{d}\boldsymbol{\pi}(\gamma) \qquad \text{a.e. } t\in[0,1], \tag{5.4.5}$$

where the exceptional set can be chosen to be independent on f.

Moreover, if $(f_t) \in AC([0,1], L^2(\mathrm{Y})) \cap L^\infty([0,1], W^{1,2}(\mathrm{Y}))$, then the map $t \mapsto \int f_t \circ \mathrm{e}_t \,\mathrm{d}\boldsymbol{\pi}$ is also absolutely continuous and

$$\frac{\mathrm{d}}{\mathrm{d}s}\Big(\int f_s\circ \mathrm{e}_s\,\mathrm{d}\boldsymbol{\pi}\Big)_{|s=t} = \int \big(\frac{\mathrm{d}}{\mathrm{d}s} f_s|_{s=t}\big)\circ \mathrm{e}_t\,\mathrm{d}\boldsymbol{\pi} + \frac{\mathrm{d}}{\mathrm{d}s}\Big(\int f_t\circ \mathrm{e}_s\,\mathrm{d}\boldsymbol{\pi}\Big)_{|s=t} \qquad \text{a.e. } t\in[0,1].$$

proof The absolute continuity of $t \mapsto \int f \circ \mathrm{e}_t \,\mathrm{d}\boldsymbol{\pi}$ and the bound (5.4.5) are trivial consequences of the definitions of test plans and Sobolev functions, see in particular (1.1.4). The fact that the exceptional set can be chosen independently on f follows from the separability of $W^{1,2}(\mathrm{Y})$ and standard approximation procedures, carried out, for instance, in [56].

For the second part, we start noticing that the second derivative in the right hand side exists for a.e. t thanks to what we have just proved, so that the claim makes sense. The absolute continuity follows from the fact that for any $t_0, t_1 \in [0,1]$, $t_0 < t_1$ it holds

$$\begin{aligned}\Big|\int f_{t_1}\circ \mathrm{e}_{t_1} - f_{t_0}\circ \mathrm{e}_{t_0}\,\mathrm{d}\boldsymbol{\pi}\Big| &\le \Big|\int f_{t_1}\circ \mathrm{e}_{t_1} - f_{t_1}\circ \mathrm{e}_{t_0}\,\mathrm{d}\boldsymbol{\pi}\Big| + \Big|\int f_{t_1} - f_{t_0}\,\mathrm{d}(\mathrm{e}_{t_0})_*\boldsymbol{\pi}\Big| \\ &\le \iint_{t_0}^{t_1} |\mathrm{d}f_{t_1}|(\gamma_t)|\dot\gamma_t|\,\mathrm{d}t\,\mathrm{d}\boldsymbol{\pi}(\gamma) + \iint_{t_0}^{t_1}\Big|\frac{\mathrm{d}}{\mathrm{d}t}f_t\Big|\,\mathrm{d}t\,\mathrm{d}(\mathrm{e}_{t_0})_*\boldsymbol{\pi}\end{aligned}$$

and our assumptions on (f_t) and $\boldsymbol{\pi}$. Now fix a point t of differentiability for (f_t) and observe that the fact that $\frac{f_{t+h}-f_t}{h}$ strongly converges in $L^2(\mathrm{Y})$ to $\frac{\mathrm{d}}{\mathrm{d}t}f_t$ and $(\mathrm{e}_{t+h})_*\boldsymbol{\pi}$ weakly converges to $(\mathrm{e}_t)_*\boldsymbol{\pi}$ as $h \to 0$ and the densities are equibounded is sufficient to get

$$\lim_{h\to 0}\int \frac{f_{t+h}-f_t}{h}\circ \mathrm{e}_{t+h}\,\mathrm{d}\boldsymbol{\pi} = \int \frac{\mathrm{d}}{\mathrm{d}t}f_t\circ \mathrm{e}_t\,\mathrm{d}\boldsymbol{\pi} = \lim_{h\to 0}\int \frac{f_{t+h}-f_t}{h}\circ \mathrm{e}_t\,\mathrm{d}\boldsymbol{\pi}.$$

Hence the conclusion comes dividing by h the trivial identity

$$\begin{aligned}\int f_{t+h}\circ \mathrm{e}_{t+h} - f_t\circ \mathrm{e}_t\,\mathrm{d}\boldsymbol{\pi} = \int f_t\circ \mathrm{e}_{t+h} - f_t\circ \mathrm{e}_t\,\mathrm{d}\boldsymbol{\pi} + \int f_{t+h}\circ \mathrm{e}_t - f_t\circ \mathrm{e}_t\,\mathrm{d}\boldsymbol{\pi} + \\ + \int (f_{t+h}-f_t)\circ \mathrm{e}_{t+h} - (f_{t+h}-f_t)\circ \mathrm{e}_t\,\mathrm{d}\boldsymbol{\pi}\end{aligned}$$

and letting $h \to 0$. □

We are now able to prove that the first minimization problem in (5.4.3) provides a further variational representation of the entropic cost and $(\rho_t^\varepsilon, \nabla\vartheta_t^\varepsilon)$ is the unique minimizer.

Theorem 5.4.3 (Benamou-Brenier formula for the entropic cost)**.** *Let* $(\mathrm{X}, \mathsf{d}, \mathfrak{m})$ *be a compact* $\mathsf{RCD}^*(K,N)$ *space with* $K \in \mathbb{R}$, $N \in [1,\infty)$ *and* $\mathfrak{m} \in \mathscr{P}(\mathrm{X})$ *and let* $\mu_0, \mu_1 \in \mathscr{P}(\mathrm{X})$ *be Borel probability measures whose densities belong to* $\mathrm{Test}^\infty_{>0}(\mathrm{X})$.

Then

$$\varepsilon\mathscr{I}_\varepsilon(\mu_0,\mu_1) = \frac{\varepsilon}{2}\Big(H(\mu_0\,|\,\mathfrak{m}) + H(\mu_1\,|\,\mathfrak{m})\Big) + \inf_{(\nu,v)}\Big\{\iint_0^1 \Big(\frac{|v_t|^2}{2} + \frac{\varepsilon^2}{8}|\nabla\log\nu_t|^2\Big)\nu_t\mathrm{d}t\mathrm{d}\mathfrak{m}\Big\} \tag{5.4.6}$$

where the entropic cost is relative to $\mathsf{R}^{\varepsilon/2}$ *defined as in Theorem 5.1.1 and the infimum is taken among all couples* $(\nu\mathfrak{m}, v)$ *solving the continuity equation in the sense of Theorem 1.1.4 under the constraint* $\nu_0\mathfrak{m} = \mu_0$ *and* $\nu_1\mathfrak{m} = \mu_1$.

proof As a preliminary remark, By Theorem 5.1.2 $\mathscr{I}_\varepsilon(\mu_0,\mu_1)$ is finite. Thus, given a solution (ν, v) of the continuity equation in the sense of the statement, without loss of generality we can assume that

$$\iint_0^1 \Big(\frac{|v_t|^2}{2} + \frac{\varepsilon^2}{8}|\nabla\log\nu_t|^2\Big)\nu_t\mathrm{d}t\mathrm{d}\mathfrak{m} < +\infty.$$

By Theorem 1.1.4 (ν_t) is W_2-absolutely continuous and the lifting $\boldsymbol{\pi}$ of (ν_t) is a test plan; moreover, $(\vartheta_t^\varepsilon) \in AC([0,1], L^2(\mathrm{X})) \cap L^\infty([0,1], W^{1,2}(\mathrm{X}))$ by what explained before. Thus Lemma 5.4.2 applies to $\boldsymbol{\pi}$ and $t \mapsto \vartheta_t^\varepsilon$, whence

$$\frac{\mathrm{d}}{\mathrm{d}s}\Big(\int \vartheta_s^\varepsilon\nu_s\mathrm{d}\mathfrak{m}\Big)_{|s=t} = \int \big(\frac{\mathrm{d}}{\mathrm{d}s}\vartheta_s^\varepsilon{}_{|s=t}\big)\nu_t\mathrm{d}\mathfrak{m} + \frac{\mathrm{d}}{\mathrm{d}s}\Big(\int \vartheta_t^\varepsilon\nu_s\mathrm{d}\mathfrak{m}\Big)_{|s=t}$$

for a.e. $t \in [0,1]$. For the first term on the right-hand side, the fact that $\vartheta_t^\varepsilon = (\psi_t^\varepsilon - \varphi_t^\varepsilon)/2$ and the 'PDEs' solved by $\psi_t^\varepsilon, \varphi_t^\varepsilon$ yield

$$\begin{aligned}\int \big(\frac{\mathrm{d}}{\mathrm{d}s}\vartheta_s^\varepsilon{}_{|s=t}\big)\nu_t\mathrm{d}\mathfrak{m} &= -\int \Big(\frac{|\nabla\psi_t^\varepsilon|^2}{4} + \frac{\varepsilon}{4}\Delta\psi_t^\varepsilon + \frac{|\nabla\varphi_t^\varepsilon|^2}{4} + \frac{\varepsilon}{4}\Delta\varphi_t^\varepsilon\Big)\nu_t\mathrm{d}\mathfrak{m}\\ &= \int \Big(-\frac{|\nabla\psi_t^\varepsilon|^2}{4} - \frac{|\nabla\varphi_t^\varepsilon|^2}{4} + \frac{\varepsilon}{4}\langle\nabla(\psi_t^\varepsilon + \varphi_t^\varepsilon), \nabla\log\nu_t\rangle\Big)\nu_t\mathrm{d}\mathfrak{m}\end{aligned}$$

and by Young's inequality

$$\varepsilon\langle\nabla(\psi_t^\varepsilon + \varphi_t^\varepsilon), \nabla\log\nu_t\rangle \le \frac{1}{2}|\nabla(\psi_t^\varepsilon + \varphi_t^\varepsilon)|^2 + \frac{\varepsilon^2}{2}|\nabla\log\nu_t|^2$$

with equality if and only if $\nabla(\psi_t^\varepsilon + \varphi_t^\varepsilon) = \varepsilon\nabla\log\nu_t$ $\mathfrak{m}$-a.e. On the other hand, the fact that (ν, v) is a solution of the continuity equation and $\vartheta_t^\varepsilon \in W^{1,2}(\mathrm{X})$ imply that

$$\frac{\mathrm{d}}{\mathrm{d}s}\Big(\int \vartheta_t^\varepsilon\nu_s\mathrm{d}\mathfrak{m}\Big)_{|s=t} = \frac{1}{2}\int \langle\nabla(\psi_t^\varepsilon - \varphi_t^\varepsilon), v_t\rangle\nu_t\mathrm{d}\mathfrak{m}$$

and by Young's inequality

$$\langle\nabla(\psi_t^\varepsilon - \varphi_t^\varepsilon), v_t\rangle \le \frac{1}{4}|\nabla(\psi_t^\varepsilon - \varphi_t^\varepsilon)|^2 + |v_t|^2$$

with equality if and only if $v_t = \nabla(\psi_t^\varepsilon - \varphi_t^\varepsilon)/2$ $\mathfrak{m}$-a.e. Plugging these observations together and integrating over $[0,1]$ (here the fact that $(\vartheta_t^\varepsilon)$ is absolutely continuous on $[0,1]$, and not only locally on $(0,1)$, is crucial) we deduce that

$$\int \vartheta_1^\varepsilon \,\mathrm{d}\mu_1 - \int \vartheta_0^\varepsilon \,\mathrm{d}\mu_0 \leq \iint_0^1 \Big(\frac{|v_t|^2}{2} + \frac{\varepsilon^2}{8}|\nabla \log \nu_t|^2\Big)\nu_t \mathrm{d}t\mathrm{d}\mathfrak{m}$$

Thanks to the identities $\vartheta_1^\varepsilon = \psi_1^\varepsilon - \frac{\varepsilon}{2}\log\rho_1$ and $\vartheta_0^\varepsilon = -\varphi_0^\varepsilon + \frac{\varepsilon}{2}\log\rho_0$ the left-hand side can be rewritten as in Proposition 5.4.1, thus getting

$$\varepsilon\mathscr{I}_\varepsilon(\mu_0,\mu_1) = \frac{\varepsilon}{2}\Big(H(\mu_0\,|\,\mathfrak{m}) + H(\mu_1\,|\,\mathfrak{m})\Big) + \inf_{(\nu,v)}\Big\{\iint_0^1 \Big(\frac{|v_t|^2}{2} + \frac{\varepsilon^2}{8}|\nabla\log\nu_t|^2\Big)\nu_t\mathrm{d}t\mathrm{d}\mathfrak{m}\Big\}$$

and the infimum is attained if and only if

$$\nabla(\psi_t^\varepsilon + \varphi_t^\varepsilon) = \varepsilon\nabla\log\nu_t \qquad \text{and} \qquad v_t = \frac{1}{2}\nabla(\psi_t^\varepsilon - \varphi_t^\varepsilon)$$

and this concludes the proof. □

We have thus obtained a dynamical representation for the entropic cost, which is in line with the celebrated Benamou-Brenier formula because the infimum is taken among all solutions to the continuity equation. However, because of the entropic nature of the problem, the standard kinetic energy functional is penalized by the Fisher information and this implicitly forces to consider in (5.4.6) solutions of the continuity equation that are not just integrable (as a genuinely weak solution should be) but also somehow regular, namely $\log\nu_t \in W^{1,2}_{loc}(\mathrm{X})$.

In order to remove such penalization from the functional and in analogy with the smooth case (see for instance Proposition 4.1 in [81] and Theorem 5.1 in [52]), one could wonder whether the entropic cost can be also represented in terms of the second or third minimization problem in (5.4.3), namely

$$\begin{aligned}\varepsilon\mathscr{I}_\varepsilon(\mu_0,\mu_1) &= \varepsilon H(\mu_0\,|\,\mathfrak{m}) + \inf_{(\nu,v)}\iint_0^1 \frac{|v_t|^2}{2}\nu_t\mathrm{d}t\mathrm{d}\mathfrak{m}\\ &= \varepsilon H(\mu_1\,|\,\mathfrak{m}) + \inf_{(\nu,v)}\iint_0^1 \frac{|v_t|^2}{2}\nu_t\mathrm{d}t\mathrm{d}\mathfrak{m}\end{aligned} \tag{5.4.7}$$

where the infima are taken among all couples (ν, v) such that the map $t \mapsto |v_t|^2\nu_t\mathrm{d}\mathfrak{m}$ is Borel, belongs to $L^1(0,1)$, $\nu_t \leq C$ for all $t \in [0,1]$ for some $C > 0$ and, respectively,

(i) the forward Fokker-Planck equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\nu_t + \mathrm{div}(v_t\nu_t) = \frac{\varepsilon}{2}\Delta\nu_t$$

is satisfied in the following sense: for any $f \in D(\Delta)$ the map $[0,1] \ni t \mapsto \int f\nu_t\,\mathrm{d}\mathfrak{m}$ is absolutely continuous and it holds

$$\frac{\mathrm{d}}{\mathrm{d}t}\int f\nu_t\,\mathrm{d}\mathfrak{m} = \int\Big(\mathrm{d}f(v_t) + \frac{\varepsilon}{2}\Delta f\Big)\nu_t\mathrm{d}\mathfrak{m} \qquad \text{-a.e. } t;$$

(ii) the backward Fokker-Planck equation

$$-\frac{\mathrm{d}}{\mathrm{d}t}\nu_t + \mathrm{div}(v_t\nu_t) = \frac{\varepsilon}{2}\Delta\nu_t$$

is satisfied as in the forward sense, up to a change of sign.

In this direction, the '$\geq$' inequalities in (5.4.7) are easily satisfied because of (5.4.2) and the fact that $(\rho_t^\varepsilon, \psi_t^\varepsilon)$, $(\rho_t^\varepsilon, \varphi_t^\varepsilon)$ are solutions to the forward and backward Fokker-Planck equation respectively in the sense described above. For the converse inequalities, if we further assume that $\log \nu_t \in W^{1,2}_{loc}(\mathrm{X})$ with

$$\iint_0^1 |\nabla \log \nu_t|^2 \nu_t \mathrm{d}t\mathrm{d}\mathfrak{m} < \infty,$$

then the problem trivializes, because a solution (ν, v) to the forward Fokker-Planck equation satisfying this condition can be seen as a solution (ν, w) to the continuity equation with $w = v - \frac{\varepsilon}{2}\nabla \log \nu$ and $t \mapsto |w_t|^2 \nu_t \mathrm{d}\mathfrak{m}$ belonging to $L^1(0,1)$, so that the same strategy adopted in Theorem 5.4.3 can be set up. Without this further assumption, the problem seems to be still open in the RCD setting.

Remark 5.4.4. In the smooth case, both in the already cited Proposition 4.1 in [81] and Theorem 5.1 in [52], in order to establish the analogue (5.4.7) of the Benamou-Brenier formula a key role is played by a lifting result for solutions of the Fokker-Planck equation (see Mik90 and CL95) and by Girsanov's theorem. By the former to any distributional solution (ν, v) of the Fokker-Planck equation with

$$\iint_0^1 \frac{|v_t|^2}{2} \nu_t \mathrm{d}t\mathrm{d}\mathfrak{m} < \infty$$

we can associate a path measure $\mathbf{Q}$ such that $(\mathrm{e}_t)_*\mathbf{Q} = \nu_t \mathfrak{m}$ for all $t \in [0,1]$ and $\mathbf{Q}$ solves the martingale problem with forward stochastic derivative $\partial_t + \frac{\varepsilon}{2}\Delta + v_t \cdot \nabla$. The latter together with uniqueness of solutions to such martingale problem implies that

$$H(\mathbf{Q} \,|\, \mathbf{R}^{\varepsilon/2}) = H(\mu_0 \,|\, \mathfrak{m}) + \iint_0^1 \frac{|v_t|^2}{2} \nu_t \mathrm{d}t\mathrm{d}\mathfrak{m},$$

where $\mathbf{R}^{\varepsilon/2}$ is the law of the slowed down Brownian motion (see also the end of Section 5.1) and $\mathfrak{m}$ is the Lebesgue measure in [81] or $\mathbf{R}^{\varepsilon/2}$ is the law of a diffusion Markov process on $\mathbb{R}^d$ and $\mathfrak{m}$ is its invariant measure in [52]. Taking into account the dynamical formulation (S_{dyn}), this is sufficient to deduce the '$\leq$' inequality in (5.4.7).

In the metric framework, to the best of our knowledge these two ingredients are not available yet. Partial results for the lifting of solutions to the Fokker-Planck equation have been obtained by Trevisan in [120], relying on a superposition principle, but this requires further regularity assumptions on (v_t), more precisely some integrability condition on Δv_t. ■

5.5 A physical digression

In the already cited Nagasawa's monograph [100], a crucial step in the proof of the equivalence between Schrödinger's equation and diffusion processes is the fact that the stochastic process associated to (3.1.2a)-(3.1.2b) can be equivalently represented by two diffusion equations of a different kind:

$$\frac{\partial u}{\partial t} + \frac{1}{2}\Delta u + \big(\mathbf{b}(t,x) + \mathbf{a}(t,x)\big) \cdot \nabla u = 0 \tag{5.5.1a}$$

$$-\frac{\partial v}{\partial t} + \frac{1}{2}\Delta v + \big(-\mathbf{b}(t,x) + \hat{\mathbf{a}}(t,x)\big) \cdot \nabla v = 0 \tag{5.5.1b}$$

where the creation and killing term c has disappeared and on the contrary additional drifts $\mathbf{a}$ and $\hat{\mathbf{a}}$ have spotted, linked together by the relationship

$$\frac{\mathbf{a}+\hat{\mathbf{a}}}{2} = \nabla R \qquad\qquad \frac{\mathbf{a}-\hat{\mathbf{a}}}{2} = \nabla S,$$

R, S being suitable functions (cf. with the end of Section 3.1). The advantage of this change of perspective is the fact that the solutions of (5.5.1a)-(5.5.1b) are the transition probabilities of the stochastic process lying behind, while this is not true for the solutions of (3.1.2a)-(3.1.2b); this also means that the physical meaning is encoded in this new system of diffusion equations and not in the old one, as already suggested. Furthermore the distribution density $\mu = fg$ solve a pair of Fokker-Planck equations

$$\begin{aligned} -\frac{\partial\mu}{\partial t} + \frac{1}{2}\Delta\mu - \operatorname{div}(\mathbf{b}(t,x)+\mathbf{a}(t,x))\mu\big) &= 0 \\ \frac{\partial\mu}{\partial t} + \frac{1}{2}\Delta\mu - \operatorname{div}(-\mathbf{b}(t,x)+\hat{\mathbf{a}}(t,x))\mu\big) &= 0 \end{aligned}$$

and a continuity equation

$$\frac{\partial\mu}{\partial t} + \operatorname{div}\Big(\Big(\mathbf{b} + \frac{\mathbf{a}-\hat{\mathbf{a}}}{2}\Big)\mu\Big) = 0.$$

This is perfectly in line with what we have seen in this chapter, in particular in Proposition 5.1.4 and Proposition 5.4.1, up to set $\mathbf{b} \equiv 0$, rescale properly and replace $\mathbf{a}, \hat{\mathbf{a}}$ by $\nabla\psi^\varepsilon, \nabla\varphi^\varepsilon$ respectively and R, S by $\log\rho^\varepsilon, \vartheta^\varepsilon$ respectively.

Still concerning Schrödinger's equation, the PDE framework provided in Proposition 5.1.4 and describing the evolution of entropic interpolations allows us to check, at least formally, that

$$\Psi_t := e^{R+iS} = \sqrt{\rho_t^\varepsilon}e^{i\vartheta_t^\varepsilon}$$

is a solution to the linear Schrödinger equation (3.1.1a) without drift, as claimed at the end of Section 3.1. Indeed, assuming for sake of simplicity $\varepsilon = 1$ (whence the drop of the apex ε in the notation), deriving Ψ_t in time and using

$$\begin{aligned} &\frac{\mathrm{d}}{\mathrm{d}t}\rho_t + \operatorname{div}(\rho_t\nabla\vartheta_t) = 0 \\ &\frac{\mathrm{d}}{\mathrm{d}t}\vartheta_t + \frac{|\nabla\vartheta_t|^2}{2} = -\frac{1}{8}\big(2\Delta\log\rho_t + |\nabla\log\rho_t|^2\big) \end{aligned}$$

we see that, on the one hand,

$$\Delta\Psi_t = e^{i\vartheta_t}\Big(\Delta\sqrt{\rho_t} + i\frac{\langle\nabla\rho_t,\nabla\vartheta_t\rangle}{\sqrt{\rho_t}} + i\sqrt{\rho_t}\Delta\vartheta_t - \sqrt{\rho_t}|\nabla\vartheta_t|^2\Big)$$

and on the other hand

$$\begin{aligned} \frac{\partial\Psi_t}{\partial t} &= -\frac{e^{i\vartheta_t}}{2}\Big(\frac{1}{\sqrt{\rho_t}}\operatorname{div}(\rho_t\nabla\vartheta_t) + i\sqrt{\rho_t}|\nabla\vartheta_t|^2 + \frac{i}{2}\sqrt{\rho_t}\Delta\log\rho_t + \frac{i}{4}\sqrt{\rho_t}|\nabla\log\rho_t|^2\Big) \\ &= -\frac{e^{i\vartheta_t}}{2}\Big(\frac{\langle\nabla\rho_t,\nabla\vartheta_t\rangle}{\sqrt{\rho_t}} + \sqrt{\rho_t}\Delta\vartheta_t + i\sqrt{\rho_t}|\nabla\vartheta_t|^2 + i\Delta\sqrt{\rho_t}\Big). \end{aligned}$$

Putting the two identities together we deduce that

$$i\frac{\partial \Psi_t}{\partial t} + \frac{1}{2}\Delta \Psi_t - \frac{\Delta\sqrt{\rho_t}}{\sqrt{\rho_t}}\Psi_t = 0. \tag{5.5.2}$$

According to Nagasawa's monograph [100], the potential V that has to be considered in (3.1.1a) in such a way that Ψ is a solution is given by (3.1.11), which in this case reads as

$$V = -2\frac{\partial S}{\partial t} - |\nabla S|^2$$

and since $S = \vartheta$, this yields

$$V = \frac{1}{2}\Delta \log \rho_t + \frac{1}{4}|\nabla \log \rho_t|^2 = \frac{\Delta\sqrt{\rho_t}}{\sqrt{\rho_t}}.$$

For sake of information, physicists refer to this functional as Bohm potential. Hence (5.5.2) is precisely the linear Schrödinger equation (3.1.1a) without drift.

As a concluding remark, it is worth recalling that in Nelson's monograph [102] the gradients of the functions R, S (or $\log \rho^\varepsilon$ and ϑ^ε within our language) have a precise nomenclature, because of their physical meaning:

$$\mathsf{v}_t^{\varepsilon,cu} := \frac{1}{2}\nabla\big(\psi_t^\varepsilon - \varphi_t^\varepsilon\big) = \nabla \vartheta_t^\varepsilon$$

is called *current velocity* because it is the velocity field driving the evolution of ρ_t^ε, while

$$\mathsf{v}_t^{\varepsilon,os} := \frac{1}{2}\nabla\big(\psi_t^\varepsilon + \varphi_t^\varepsilon\big) = \frac{\varepsilon}{2}\nabla \log \rho_t^\varepsilon$$

is the *osmotic velocity*, since it describes the diffusion effects of Brownian motion; this is clear when one integrates the squared norm of the osmotic velocity w.r.t. μ_t^ε, the result being the Fisher information. Looking at the results of this chapter from this physical point of view, we see that (5.3.9c) becomes

$$\sup_{\varepsilon\in(0,1)} \iint_0^1 |\mathsf{v}_t^{\varepsilon,cu}|^2 \rho_t^\varepsilon \mathrm{d}t\mathrm{d}\mathfrak{m} < \infty$$

and provides a uniform control on the kinetic energy of the system during the slowing down procedure (i.e. as $\varepsilon \downarrow 0$ and the Brownian motion described by R^ε is progressively freezed), while (5.3.12b), which reads as

$$\lim_{\varepsilon\downarrow 0} \iint_\delta^{1-\delta} |\mathsf{v}_t^{\varepsilon,os}|^2 \rho_t^\varepsilon \mathrm{d}t\mathrm{d}\mathfrak{m} = 0,$$

suggests that the diffusion effect disappears in the limit, in line with the physical interpretation; the mathematical proof is postponed to the next chapter (cf. Lemma 6.3.5).

Part III

Geometry of RCD spaces

Chapter 6

From entropic to displacement interpolations

In this chapter the uniform estimates on interpolations and potentials as well as the information on the behaviour of the entropy along entropic interpolations obtained in Chapter 5 are widely exploited. The purpose is to understand what happens as $\varepsilon \downarrow 0$ and investigate the limit case. Inspired by the abstract results pointed out by Léonard in [79], we establish strong links between Schrödinger problem and optimal transport, a deep relationship that in many aspects had not been conjectured yet. Indeed, not only do we link entropic interpolations to displacement ones (in line with [79]), but we also pass from Schrödinger potentials to Kantorovich ones and explore the geometric information hidden in (5.3.6a) and (5.3.6b), underlying some remarkable consequences.

In Section 6.1 a compactness result is proved: we show that Schrödinger potentials and entropic interpolations converge in a suitable sense, up to subsequences, to limit real-valued functions and curves of measures respectively.

Aim of Section 6.2 is then to characterize these limits. On the one hand, we show that the limit potentials are not only Kantorovich potentials but also that their evolution is given by the Hopf-Lax semigroup, thus proving that the viscous solution of the Hamilton-Jacobi equation can be obtained via a vanishing viscosity method, in accordance with the smooth case. On the other hand, the limit interpolation is the (unique) Wasserstein geodesic between the marginal constraints. We also discuss Γ-convergence of the rescaled Schrödinger problems to Monge-Kantorovich one.

The formulas for first and second derivatives of the entropy along entropic interpolations are the core of Section 6.3. From (5.3.6b) one can guess that the Boltzmann entropy along entropic interpolations is more convex than along displacement ones; more precisely, it is (K, N)-convex in the sense of Erbar-Kuwada-Sturm ([45]). By means of the compactness and characterization results of the first two sections, this allows us to recover in a simple way some well known results in the theory of RCD^* spaces: for instance, the HWI inequality and the $\mathsf{CD}^e(K, N)$ condition.

6.1 Compactness

Starting from the uniform estimates discussed in Section 5.2, let us first prove that when we pass to the limit as $\varepsilon \downarrow 0$, up to subsequences Schrödinger potentials and entropic interpolations

converge in a suitable sense to limit potentials and interpolations.

Proposition 6.1.1 (Compactness)**.** *With the same assumptions and notations as in Setting 5.1.3 the following holds.*

For any sequence $\varepsilon_n \downarrow 0$ there exists a subsequence, not relabeled, so that:

(i) the curves $(\mu_t^{\varepsilon_n})$ uniformly converge in $(\mathscr{P}_2(\mathrm{X}), W_2)$ to a limit curve (μ_t) which belongs to $AC([0,1],(\mathscr{P}_2(\mathrm{X}),W_2))$. Moreover, there is $M>0$ so that

$$\mu_t \le M\mathfrak{m} \qquad \forall t \in [0,1] \tag{6.1.1}$$

and setting $\rho_t := \frac{\mathrm{d}\mu_t}{\mathrm{d}\mathfrak{m}}$ it holds

$$\rho_t^{\varepsilon_n} \overset{*}{\rightharpoonup} \rho_t \quad \text{in } L^\infty(\mathfrak{m}) \qquad \forall t \in [0,1]. \tag{6.1.2}$$

(ii) for all $t \in I$ the functions $\varphi_t^{\varepsilon_n}, \psi_t^{\varepsilon_n}$ converge locally uniformly on X to locally Lipschitz limit functions φ_t, ψ_t respectively and, for any Lipschitz cut-off function χ with bounded support, the curves $(\chi\varphi_t^{\varepsilon_n}), (\chi\psi_t^{\varepsilon_n})$ converge locally uniformly on I with values in $L^1(\mathrm{X})$ to limit curves $(\chi\varphi_t), (\chi\psi_t) \in AC_{loc}(I, L^1(\mathrm{X}))$ with $\mathrm{Lip}(\chi\varphi_t), \mathrm{Lip}(\chi\psi_t)$ locally bounded for $t \in I$, where $I := (0,1]$ for the φ's and $I := [0,1)$ for the ψ's. Moreover for every $t \in (0,1)$ it holds

$$\begin{aligned} &\varphi_t + \psi_t \le 0 && \text{on } \mathrm{X}, \\ &\varphi_t + \psi_t = 0 && \text{on } \mathrm{supp}(\mu_t). \end{aligned} \tag{6.1.3}$$

Similarly, the curves $(\vartheta_t^{\varepsilon_n})$ converge in $(0,1)$ to the limit curve $t \mapsto \vartheta_t := \frac{1}{2}(\psi_t - \varphi_t)$ in the same sense as above.

proof

(i) Fix $\varepsilon \in (0,1)$; we want to apply Theorem 1.1.4 to (μ_t^ε) and $(\nabla\vartheta_t^\varepsilon)$. The continuity of $t \mapsto \rho_t^\varepsilon \in L^2(\mathrm{X})$ granted by Proposition 5.1.4 yields weak continuity of (μ_t) and (1.1.10a) is a consequence of (5.2.5). From the bound (5.3.9c) it follows (1.1.10b) and from the formula for $\frac{\mathrm{d}}{\mathrm{d}t}\rho_t^\varepsilon$ given in Proposition 5.1.4 and again the L^2-continuity of (ρ_t^ε) on $[0,1]$ it easily follows that (μ_t^ε) and $(\vartheta_t^\varepsilon)$ solve the continuity equation in the sense of Theorem 1.1.4. The conclusion of such theorem ensures that (μ_t^ε) is W_2-absolutely continuous with

$$\int_0^1 |\dot\mu_t^\varepsilon|^2 \,\mathrm{d}t = \iint_0^1 |\nabla\vartheta_t^\varepsilon|^2 \rho_t^\varepsilon \,\mathrm{d}t\,\mathrm{d}\mathfrak{m}.$$

The bound (5.3.9c) grants that the right hand side is uniformly bounded in $\varepsilon \in (0,1)$ and since $\{(\mu_t^\varepsilon)\}_\varepsilon$ is tight and 2-uniformly integrable by (5.2.5) (hence W_2-compact), this is sufficient to ensure the compactness of the family $\{(\mu_t^\varepsilon)\}_\varepsilon$ in $C([0,1],(\mathscr{P}_2(\mathrm{X}),W_2))$ and, by the lower semicontinuity of the kinetic energy, the fact that any limit curve (μ_t) is absolutely continuous. The bound (6.1.1) is then a direct consequence of the uniform bound (5.2.6) and the convergence property (6.1.2) comes from the weak convergence of the measures and the uniform bound on the densities.

(ii) Let $\mathcal{B}$ be a countable family of increasing nested closed balls covering X. From the formula for $\frac{\mathrm{d}}{\mathrm{d}t}\varphi_t^\varepsilon$ provided in Proposition 5.1.4, for any $B \in \mathcal{B}$ and Lipschitz cut-off function χ with bounded support such that $\chi \equiv 1$ in B we obtain

$$\|\chi(\varphi_t^\varepsilon - \varphi_s^\varepsilon)\|_{L^1(\mathfrak{m})} \le \iint_t^s \chi\Big(\frac{|\nabla\varphi_r^\varepsilon|^2}{2} + \frac{\varepsilon}{2}|\Delta\varphi_r^\varepsilon|\Big)\mathrm{d}r\,\mathrm{d}\mathfrak{m} \qquad \forall\varepsilon>0,\ \forall t,s \in (0,1],\ t<s.$$

Thus for $\delta \in (0,1)$ the estimates (5.2.10a) and (5.2.11) give

$$\|\chi(\varphi_t^\varepsilon - \varphi_s^\varepsilon)\|_{L^1(\mathfrak{m})} \le C'_\delta |s-t| \qquad \forall \varepsilon \in (0,1),\ \forall t,s \in [\delta,1],\ t<s. \tag{6.1.4}$$

Now notice that for $h \in \mathrm{LIP}_c(\mathrm{X})$ and $\mu \in \mathscr{P}(\mathrm{X})$, integrating in y w.r.t. μ the trivial inequality $h(x) \le h(y) + \mathsf{d}(x,y)\,\mathrm{Lip}(h)$ yields

$$h(x)^+ \le \Big(\int h\,\mathrm{d}\mu + D\,\mathrm{Lip}(h)\Big)^+ \le \Big|\int h\,\mathrm{d}\mu\Big| + D\,\mathrm{Lip}(h)$$

where D is the diameter of $\mathrm{supp}(h)$ and since a similar bound can be obtained for $h(x)^-$ we get

$$\|h\|_{L^\infty(\mathfrak{m})} \le \Big|\int h\,\mathrm{d}\mu\Big| + D\,\mathrm{Lip}(h) \le \|h\|_{L^1(\mu)} + D\,\mathrm{Lip}(h). \tag{6.1.5}$$

Choosing $\mu := \mu_1$ and $h := \chi\varphi_1^\varepsilon$ and recalling that the normalization chosen for $(f^\varepsilon, g^\varepsilon)$ in Setting 5.1.3 reads as $\int \varphi_1^\varepsilon\,\mathrm{d}\mu_1 = 0$, we see that if $\mathrm{supp}(\mu_1) \subset B$ (which is always the case, up to choose B sufficiently large), then $\int \chi\varphi_1^\varepsilon\,\mathrm{d}\mu_1 = 0$ too and thus from the first inequality above we deduce that $\{\chi\varphi_1^\varepsilon\}_{\varepsilon\in(0,1)}$ is uniformly bounded in $L^\infty(\mathfrak{m})$. Using this information together with (5.2.10a), (6.1.4) and the second inequality in (6.1.5) with $\mu := \alpha\mathfrak{m}|_{\mathrm{supp}(\chi)}$ (α being the normalization constant) we conclude that

$$\sup_{\varepsilon\in(0,1)} \sup_{t\in[\delta,1]} \|\chi\varphi_t^\varepsilon\|_{L^\infty(\mathfrak{m})} < \infty. \tag{6.1.6}$$

By Ascoli-Arzelà's theorem, this uniform bound and the equi-Lipschitz continuity in space (given by (5.2.10a) and the fact that χ has bounded support) together with the equi-Lipschitz continuity in time given by (6.1.4) give compactness in $C([\delta,1], L^1(\mathrm{X}))$; it is clear then that any limit curve (ζ_t) belongs to $\mathrm{LIP}([\delta,1], L^1(\mathrm{X}))$ and that $\sup_{t\in[\delta,1]} \mathrm{Lip}(\zeta_t) < \infty$. By repeating the same argument for all $B \in \mathcal{B}$, a diagonalization argument provides us with a family of curves $\{(\zeta_t^i)\}_{i\in\mathbb{N}} \subset \mathrm{LIP}([\delta,1], L^1(\mathrm{X}))$ such that $\zeta_t^i = \zeta_t^j$ $\mathfrak{m}$-a.e. in $B_i \cap B_j$ for all $t \in [\delta,1]$. Hence, putting $\varphi_t := \zeta_t^i$ in B_i for all $i \in \mathbb{N}$, it is easy to see that $(\varphi_t) \in \mathrm{LIP}([\delta,1], L^1(\mathrm{Y}))$ for any bounded Borel set $\mathrm{Y} \subset \mathrm{X}$. A further diagonalization argument (in time) and the arbitrariness of $\delta \in (0,1)$ provide us the required results on $(0,1]$.

The argument for the ψ_t^ε's follows the same lines provided we are able to show that for some $t \in [0,1)$ the functions $\chi\psi_t^\varepsilon$ are uniformly bounded, which is the same as to prove that ψ_t^ε are locally uniformly bounded. To see this, observe that from the estimate (5.2.6) it follows that on a given bounded Borel set B

$$-e^{-1}\mathfrak{m}(B) \le \int_B \rho_t^\varepsilon \log \rho_t^\varepsilon \mathrm{d}\mathfrak{m} \le M \log M \mathfrak{m}(B) \qquad \forall \varepsilon \in (0,1),\ t \in [0,1],$$

thus multiplying the identity

$$\varphi_t^\varepsilon + \psi_t^\varepsilon = \varepsilon \log \rho_t^\varepsilon \qquad \forall t \in (0,1) \tag{6.1.7}$$

by ρ_t^ε and integrating on B we get

$$-\varepsilon e^{-1}\mathfrak{m}(B) \le \int_B \varphi_t^\varepsilon + \psi_t^\varepsilon\,\mathrm{d}\mu_t^\varepsilon \le \varepsilon M \log M \mathfrak{m}(B) \qquad \forall t \in (0,1). \tag{6.1.8}$$

Since we know that $\varphi^\varepsilon_{1/2}$ is locally uniformly bounded, this yields a uniform control on $\int_B \psi^\varepsilon_{1/2}\,\mathrm{d}\mu^\varepsilon_{1/2}$ and then we can proceed as before starting from the first inequality in (6.1.5) with $h := \psi^\varepsilon_{1/2}$ and $\mu := \alpha_\varepsilon \mu^\varepsilon_{1/2}|_B$, α_ε being the normalization constant. The fact that μ has compact support (although not h) and the tightness of $\{\mu^\varepsilon_{1/2}\}_{\varepsilon\in(0,1)}$, which follows from (5.2.5) and grants that α_ε is uniformly bounded (provided B has been chosen sufficiently large), entail the desired locally uniform bound for $\psi^\varepsilon_{1/2}$.

The claim for the $(\vartheta^\varepsilon_t)$ is now obvious.

Finally, to prove the first in (6.1.3) we pass to the limit in (6.1.7) recalling the uniform bound (5.2.6), then passing to the limit in (6.1.8) (by uniform convergence of functions and weak convergence of measures) we deduce that

$$\int_B \varphi_t + \psi_t \,\mathrm{d}\mu_t = 0,$$

which forces the second in (6.1.3) by the arbitrariness of B. □

6.2 Identification of the limit curve and potentials

We now show that the limit interpolation is the geodesic from μ_0 to μ_1 and the limit potentials are Kantorovich potentials that evolve according to the Hopf-Lax semigroup (recall formula (1.1.1)).

Proposition 6.2.1 (Limit curve and potentials)**.** *With the same assumptions and notations as in Setting 5.1.3 the following holds.*

The limit curve (μ_t) given by Proposition 6.1.1 is unique (i.e. independent on the sequence $\varepsilon_n \downarrow 0$) and is the only W_2-geodesic connecting μ_0 to μ_1.

For any Lipschitz cut-off function χ with bounded support and any limit curve (φ_t) given by Proposition 6.1.1, $(\chi\varphi_t)$ is in $AC_{loc}((0,1], C(\mathrm{X})) \cap L^\infty_{loc}((0,1], W^{1,2}(\mathrm{X}))$ and for any $t_0, t_1 \in (0,1]$, $t_0 < t_1$ we have

$$-\varphi_{t_1} = Q_{t_1-t_0}(-\varphi_{t_0}) \tag{6.2.1a}$$

$$\int \varphi_{t_0}\,\mathrm{d}\mu_{t_0} - \int \varphi_{t_1}\,\mathrm{d}\mu_{t_1} = \frac{1}{2(t_1-t_0)} W_2^2(\mu_{t_0}, \mu_{t_1}) \tag{6.2.1b}$$

and $-(t_1-t_0)\varphi_{t_1}$ is a Kantorovich potential from μ_{t_1} to μ_{t_0}.

Similarly, for any cut-off function χ as above and any limit curve (ψ_t) given by Proposition 6.1.1, $(\chi\psi_t)$ belongs to $AC_{loc}([0,1), C(\mathrm{X})) \cap L^\infty_{loc}([0,1), W^{1,2}(\mathrm{X}))$ and for every $t_0, t_1 \in [0,1)$, $t_0 < t_1$ we have

$$-\psi_{t_0} = Q_{t_1-t_0}(-\psi_{t_1}) \tag{6.2.2a}$$

$$\int \psi_{t_1}\,\mathrm{d}\mu_{t_1} - \int \psi_{t_0}\,\mathrm{d}\mu_{t_0} = \frac{1}{2(t_1-t_0)} W_2^2(\mu_{t_0}, \mu_{t_1}) \tag{6.2.2b}$$

and $-(t_1-t_0)\psi_{t_0}$ is a Kantorovich potential from μ_{t_0} to μ_{t_1}.

proof

Inequality $\leq$ in (6.2.1a). Pick $x, y \in \mathrm{X}$, $r > 0$, define

$$\nu^r_x := \frac{1}{\mathfrak{m}(B_r(x))}\mathfrak{m}|_{B_r(x)} \qquad\qquad \nu^r_y := \frac{1}{\mathfrak{m}(B_r(y))}\mathfrak{m}|_{B_r(y)}$$

and $\boldsymbol{\pi}^r$ as the only lifting of the only W_2-geodesic from ν_x^r to ν_y^r (recall point (i) of Theorem 1.2.6). Since ν_x^r, ν_y^r have compact support and $\boldsymbol{\pi}^r \in \mathrm{OptGeo}(\nu_x^r, \nu_y^r)$, there exist $\bar{x} \in \mathrm{X}$ and $R > 0$ sufficiently large such that

$$\mathrm{supp}((\mathrm{e}_t)_*\boldsymbol{\pi}^r) \subset B_R(\bar{x}), \qquad \forall t \in [0,1]. \tag{6.2.3}$$

Hence let χ be a Lipschitz cut-off function with bounded support such that $\chi \equiv 1$ in $B_R(\bar{x})$. Then, let $\varepsilon \in (0,1)$ and $0 < t_0 < t_1 \leq 1$, define $\tilde{\varphi}_t^\varepsilon := \chi\varphi_t^\varepsilon$ and observe that $(\tilde{\varphi}_t^\varepsilon) \in AC_{loc}((0,1], L^2(\mathfrak{m})) \cap L^\infty_{loc}((0,1], W^{1,2}(\mathrm{X}))$ by Proposition 5.1.4 and the compactness of the support of χ; thus, by Lemma 5.4.2 applied to $\boldsymbol{\pi}^r$ and $t \mapsto \tilde{\varphi}^\varepsilon_{(1-t)t_0+tt_1}$, we get

$$\frac{\mathrm{d}}{\mathrm{d}t}\int \tilde{\varphi}^\varepsilon_{(1-t)t_0+tt_1} \circ \mathrm{e}_t \,\mathrm{d}\boldsymbol{\pi}^r \geq \int (t_1 - t_0)\frac{\mathrm{d}}{\mathrm{d}s}\tilde{\varphi}^\varepsilon_s|_{s=(1-t)t_0+tt_1}(\gamma_t) - |\mathrm{d}\tilde{\varphi}^\varepsilon_{(1-t)t_0+tt_1}|(\gamma_t)|\dot{\gamma}_t| \,\mathrm{d}\boldsymbol{\pi}^r(\gamma).$$

As (6.2.3) implies that $\chi(\gamma_t) = 1$ for all $t \in [0,1]$ for $\boldsymbol{\pi}^r$-a.e. γ, $\tilde{\varphi}^\varepsilon$ can be replaced by φ^ε in the inequality above and, recalling the expression for $\frac{\mathrm{d}}{\mathrm{d}t}\varphi_t^\varepsilon$ and using Young's inequality, we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}\int \varphi^\varepsilon_{(1-t)t_0+tt_1} \circ \mathrm{e}_t \,\mathrm{d}\boldsymbol{\pi}^r \geq \int \varepsilon\frac{t_1 - t_0}{2}\Delta\varphi^\varepsilon_{(1-t)t_0+tt_1}(\gamma_t) - \frac{1}{2(t_1 - t_0)}|\dot{\gamma}_t|^2 \,\mathrm{d}\boldsymbol{\pi}^r(\gamma).$$

Integrating in time and using property i) of Theorem 1.2.6 (indeed, recall that $\boldsymbol{\pi}^r$ is optimal), we get

$$\int \varphi^\varepsilon_{t_1} \,\mathrm{d}\nu_y^r - \int \varphi^\varepsilon_{t_0} \,\mathrm{d}\nu_x^r \geq -\frac{1}{2(t_1 - t_0)}W_2^2(\nu_y^r, \nu_x^r) + \iint_0^1 \varepsilon\frac{t_1 - t_0}{2}\Delta\varphi^\varepsilon_{(1-t)t_0+tt_1} \circ \mathrm{e}_t \,\mathrm{d}t \,\mathrm{d}\boldsymbol{\pi}^r.$$

Let $\varepsilon \downarrow 0$ along the sequence (ε_n) for which $(\varphi_t^{\varepsilon_n})$ converges to our given (φ_t) in the sense of Proposition 6.1.1 and use the uniform bound (5.2.11) with $B = B_R(\bar{x})$ and the fact that $\boldsymbol{\pi}^r$ has bounded compression to deduce that

$$\int \varphi_{t_1} \,\mathrm{d}\nu_y^r - \int \varphi_{t_0} \,\mathrm{d}\nu_x^r \geq -\frac{1}{2(t_1 - t_0)}W_2^2(\nu_y^r, \nu_x^r)$$

and finally letting $r \downarrow 0$ we conclude from the arbitrariness of $x \in \mathrm{X}$ that

$$-\varphi_{t_1}(y) \leq Q_{t_1 - t_0}(-\varphi_{t_0})(y) \qquad \forall y \in \mathrm{X}. \tag{6.2.4}$$

Inequality $\geq$ in (6.2.1a). To prove the opposite inequality we fix $\bar{x} \in \mathrm{X}$, $R > 1$, again $0 < t_0 < t_1 \leq 1$ and we take a cut-off function $\chi \in \mathrm{Test}^\infty(\mathrm{X})$ with support in $B_{R+1}(\bar{x})$ such that $\chi \equiv 1$ in $B_R(\bar{x})$. Define the vector field $X_t^\varepsilon := \chi\nabla\varphi_t^\varepsilon$ and apply Theorem 1.2.5 to $((t_1 - t_0)X^\varepsilon_{(1-t)t_1+tt_0})$: the inequality

$$\mathrm{div}X_t^\varepsilon \geq \chi\Delta\varphi_t^\varepsilon - |\nabla\chi||\nabla\varphi_t^\varepsilon|$$

and the bounds (5.2.10a), (5.2.10b) on $\nabla\varphi_t^\varepsilon, \Delta\varphi_t^\varepsilon$ ensure that the theorem is applicable and we obtain existence of the regular Lagrangian flow F^ε. Put $\boldsymbol{\pi}^\varepsilon := (F_\cdot^\varepsilon)_*\mathfrak{m}|_{B_{R/2}(\bar{x})}$, where $F_\cdot^\varepsilon : \mathrm{X} \to C([0,1], \mathrm{X})$ is the $\mathfrak{m}$-a.e. defined map which sends x to $F_t^\varepsilon(x)$, and observe that the bound (1.2.25) and the identity (1.2.26) provided by Theorem 1.2.5 coupled with the estimates

(5.2.10a), (5.2.10b) on $\nabla\varphi_t^\varepsilon, \Delta\varphi_t^\varepsilon$ and the fact that $\chi \in \mathrm{Test}^\infty(\mathrm{X})$ ensure that $\boldsymbol{\pi}^\varepsilon$ is a test plan with

$$\sup_{\varepsilon\in(0,1)} \iint_0^1 |\dot\gamma_t|^2 \,\mathrm{d}t\,\mathrm{d}\boldsymbol{\pi}^\varepsilon(\gamma) < \infty \qquad \text{and} \qquad (\mathrm{e}_t)_*\boldsymbol{\pi}^\varepsilon \le C\mathfrak{m} \qquad \forall t\in[0,1],\ \varepsilon\in(0,1), \tag{6.2.5}$$

for some $C<\infty$. Moreover, from (5.2.10a) we know that $|X_t^\varepsilon| \le C'(R+2)$ for all $t\in[t_0,1]$ for some $C'<\infty$ which only depends on $t_0,\bar x$ and together with (1.2.26) this implies that for $\boldsymbol{\pi}^\varepsilon$-a.e. γ if $\gamma_0\in B_r(\bar x)$, then $\gamma_t \in B_{r+(t_1-t_0)C'(R+2)}(\bar x)$ for all $t\in[0,1]$. Take $r=R/2$ and suppose for the moment that

$$t_1-t_0 \le \frac{1}{6C'} \le \frac{R-r}{C'(R+2)} = \frac{R}{2C'(R+2)}, \tag{6.2.6}$$

the second inequality being true because $R>1$. This means that the measures $(\mathrm{e}_t)_*\boldsymbol{\pi}^\varepsilon$ are supported in $B_R(\bar x)$ for all $t\in[0,1]$ and $\chi(\gamma_t)=1$ for all $t\in[0,1]$ for $\boldsymbol{\pi}^\varepsilon$-a.e. γ. Thus, choosing a further cut-off function $\tilde\chi\in\mathrm{Test}^\infty(\mathrm{X})$ with bounded support such that $\tilde\chi\equiv 1$ in $B_R(\bar x)$ and defining $\tilde\varphi_t^\varepsilon := \tilde\chi\varphi_t^\varepsilon$, in such a way that $(\tilde\varphi_t^\varepsilon)\in AC_{loc}((0,1],L^2(\mathfrak{m}))\cap L^\infty_{loc}((0,1],W^{1,2}(\mathrm{X}))$, by Lemma 5.4.2 applied to $\boldsymbol{\pi}^\varepsilon$ and $t\mapsto\tilde\varphi^\varepsilon_{(1-t)t_1+tt_0}$ and by the fact that $\tilde\chi(\gamma_t)=1$ for all $t\in[0,1]$ for $\boldsymbol{\pi}^\varepsilon$-a.e. γ we obtain

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}\int \varphi^\varepsilon_{(1-t)t_1+tt_0}\circ\mathrm{e}_t\,\mathrm{d}\boldsymbol{\pi}^\varepsilon = \frac{\mathrm{d}}{\mathrm{d}t}\int \tilde\varphi^\varepsilon_{(1-t)t_1+tt_0}\circ\mathrm{e}_t\,\mathrm{d}\boldsymbol{\pi}^\varepsilon\\
&\quad= \int (t_0-t_1)\frac{\mathrm{d}}{\mathrm{d}s}\tilde\varphi^\varepsilon_s|_{s=(1-t)t_1+tt_0}\circ\mathrm{e}_t\,\mathrm{d}\boldsymbol{\pi}^\varepsilon + \frac{\mathrm{d}}{\mathrm{d}s}\int\tilde\varphi^\varepsilon_{(1-t)t_1+tt_0}\circ\mathrm{e}_s\,\mathrm{d}\boldsymbol{\pi}^\varepsilon|_{s=t}\\
&\quad= \int (t_0-t_1)\frac{\mathrm{d}}{\mathrm{d}s}\varphi^\varepsilon_s|_{s=(1-t)t_1+tt_0}\circ\mathrm{e}_t\,\mathrm{d}\boldsymbol{\pi}^\varepsilon + (t_1-t_0)\int \mathrm{d}\tilde\varphi^\varepsilon_{(1-t)t_1+tt_0}(X_t^\varepsilon)\circ\mathrm{e}_t\,\mathrm{d}\boldsymbol{\pi}^\varepsilon\\
&\quad= \int\Big(\frac{t_0-t_1}{2}|\mathrm{d}\varphi^\varepsilon_{(1-t)t_1+tt_0}|^2 + \varepsilon\frac{t_0-t_1}{2}\Delta\varphi^\varepsilon_{(1-t)t_1+tt_0} + (t_1-t_0)|\mathrm{d}\varphi^\varepsilon_{(1-t)t_1+tt_0}|^2\Big)\circ\mathrm{e}_t\,\mathrm{d}\boldsymbol{\pi}^\varepsilon\\
&\quad= \int\Big(\frac{t_1-t_0}{2}|\mathrm{d}\varphi^\varepsilon_{(1-t)t_1+tt_0}|^2 + \varepsilon\frac{t_0-t_1}{2}\Delta\varphi^\varepsilon_{(1-t)t_1+tt_0}\Big)\circ\mathrm{e}_t\,\mathrm{d}\boldsymbol{\pi}^\varepsilon.
\end{aligned}$$

Integrating in time and recalling (1.2.26) we deduce

$$\int \varphi^\varepsilon_{t_0}\circ\mathrm{e}_1 - \varphi^\varepsilon_{t_1}\circ\mathrm{e}_0\,\mathrm{d}\boldsymbol{\pi}^\varepsilon = \iint_0^1 \frac{1}{2(t_1-t_0)}|\dot\gamma_t|^2 + \varepsilon\frac{t_0-t_1}{2}\Delta\varphi^\varepsilon_{(1-t)t_1+tt_0}(\gamma_t)\,\mathrm{d}t\,\mathrm{d}\boldsymbol{\pi}^\varepsilon(\gamma). \tag{6.2.7}$$

Now, as before, we let $\varepsilon\downarrow 0$ along the sequence (ε_n) for which $(\varphi_t^{\varepsilon_n})$ converges to our given (φ_t) in the sense of Proposition 6.1.1: the first in (6.2.5) grants that $(\boldsymbol{\pi}^\varepsilon)$ is tight in $\mathscr{P}(C([0,1],\mathrm{X}))$ (because $\gamma\mapsto\int_0^1|\dot\gamma_t|^2\,\mathrm{d}t$ has compact sublevels) and thus up to pass to a subsequence, not relabeled, we can assume that $(\boldsymbol{\pi}^{\varepsilon_n})$ weakly converges to some $\boldsymbol{\pi}\in\mathscr{P}(C([0,1],\mathrm{X}))$. The second in (6.2.5) and the bound (5.2.11) with $B=B_R(\bar x)$ grant that the term with the Laplacian in (6.2.7) vanishes in the limit and thus taking into account the lower semicontinuity of the 2-energy we deduce that

$$\int \varphi_{t_0}\circ\mathrm{e}_1 - \varphi_{t_1}\circ\mathrm{e}_0\,\mathrm{d}\boldsymbol{\pi} \ge \frac{1}{2(t_1-t_0)}\iint_0^1|\dot\gamma_t|^2\,\mathrm{d}t\,\mathrm{d}\boldsymbol{\pi} \ge \frac{1}{2(t_1-t_0)}\int \mathsf{d}^2(\gamma_0,\gamma_1)\,\mathrm{d}\boldsymbol{\pi}(\gamma).$$

Now notice that (6.2.4) implies that

$$\frac{\mathsf{d}^2(\gamma_0,\gamma_1)}{2(t_1-t_0)} \ge \varphi_{t_0}(\gamma_1) - \varphi_{t_1}(\gamma_0) \tag{6.2.8}$$

for any curve γ, hence the above gives

$$\int \varphi_{t_0} \circ \mathrm{e}_1 - \varphi_{t_1} \circ \mathrm{e}_0 \,\mathrm{d}\boldsymbol{\pi} \geq \frac{1}{2(t_1 - t_0)} \int \mathsf{d}^2(\gamma_0, \gamma_1) \,\mathrm{d}\boldsymbol{\pi}(\gamma) \geq \int \varphi_{t_0} \circ \mathrm{e}_1 - \varphi_{t_1} \circ \mathrm{e}_0 \,\mathrm{d}\boldsymbol{\pi}$$

thus forcing the inequalities to be equalities. In particular, equality in (6.2.8) holds for $\boldsymbol{\pi}$-a.e. γ and since $(\mathrm{e}_0)_*\boldsymbol{\pi} = \mathfrak{m}|_{B_{R/2}(\bar{x})}$, this is the same as to say that for $\mathfrak{m}$-a.e. $y \in B_{R/2}(\bar{x})$ equality holds in (6.2.4). Since both sides of (6.2.4) are continuous in y, we deduce that equality holds for any $y \in B_{R/2}(\bar{x})$ and the arbitrariness of R in (6.2.6) allows us to say that equality actually holds for any $y \in \mathrm{X}$.

In order to get rid of (6.2.6), it is sufficient to take a partition $t_0 = s_0 < s_1 < ... < s_n = t_1$ such that $s_{i+1} - s_i \leq (6C')^{-1}$ for all $i = 0, ..., n-1$ and use the fact that Q_t is a semigroup. More precisely

$$\begin{aligned} -\varphi_{t_1} &= Q_{t_1 - s_{n-1}}(-\varphi_{s_{n-1}}) = (Q_{t_1 - s_{n-1}} \circ Q_{s_{n-1} - s_{n-2}})(-\varphi_{s_{n-2}}) \\ &= (Q_{t_1 - s_{n-1}} \circ Q_{s_{n-1} - s_{n-2}} \circ ... \circ Q_{s_1 - t_0})(-\varphi_{t_0}) = Q_{t_1 - t_0}(-\varphi_{t_0}) \end{aligned}$$

whence (6.2.1a) for all $t_0, t_1 \in (0, 1]$, $t_0 < t_1$.

Other properties of φ_t. From Proposition 6.1.1 we already know that, for any Lipschitz cut-off function χ with bounded support, $(\chi\varphi_t) \in AC_{loc}((0,1], L^1(\mathrm{X})) \cap L^\infty_{loc}((0,1], W^{1,2}(\mathrm{X}))$. Since φ_t is a real-valued function for all $t \in (0, 1]$, (6.2.1a) tells us that for all $x \in \mathrm{X}$ $t \mapsto \varphi_t(x)$ satisfies (1.1.2) for a.e. $t \in (0, 1]$ and since by Proposition 6.1.1 we know that $\mathrm{lip}(\varphi_t)$ is locally bounded in space and in $t \in (0, 1]$, this implies that for all $t_0, t_1 \in (0, 1]$ with $t_0 < t_1$

$$\|\chi(\varphi_{t_1} - \varphi_{t_0})\|_\infty \leq \sup_{x \in \mathrm{supp}(\chi)} |\varphi_{t_1}(x) - \varphi_{t_0}(x)| \leq \Big(\sup_{t \in [t_0, t_1]} \mathrm{Lip}(\varphi^\varepsilon_t|_{\mathrm{supp}(\chi)}) \Big) |t_1 - t_0|$$

whence $(\chi\varphi_t) \in AC_{loc}((0,1], C(\mathrm{X})) \cap L^\infty_{loc}((0,1], W^{1,2}(\mathrm{X}))$.

Up to extract a further subsequence - not relabeled - we can assume that the curves $(\mu^{\varepsilon_n}_t)$ converge to a limit curve (μ_t) as in Proposition 6.1.1. We claim that for any $t_0, t_1 \in (0, 1]$, $t_0 < t_1$ it holds

$$-\int \varphi_{t_1} \,\mathrm{d}\mu_{t_1} + \int \varphi_{t_0} \,\mathrm{d}\mu_{t_0} \geq \frac{1}{2(t_1 - t_0)} W_2^2(\mu_{t_0}, \mu_{t_1}). \tag{6.2.9}$$

To see this, fix $\bar{x} \in \mathrm{X}$ and $R > 0$, let $\chi_R \in \mathrm{Test}^\infty(\mathrm{X})$ be a cut-off function with support in $B_{R+1}(\bar{x})$ such that $\chi_R \equiv 1$ on $B_R(\bar{x})$ and observe that by Proposition 5.1.4 $t \mapsto \int \chi_R \varphi^\varepsilon_t \rho^\varepsilon_t \,\mathrm{d}\mathfrak{m}$ belongs to $C((0,1]) \cap AC_{loc}((0,1))$ with

$$\begin{aligned} -\frac{\mathrm{d}}{\mathrm{d}t} \int \chi_R \varphi^\varepsilon_t \rho^\varepsilon_t \,\mathrm{d}\mathfrak{m} = &\int \chi_R \Big(-\frac{|\nabla \varphi^\varepsilon_t|^2}{2} - \frac{\varepsilon}{2} \Delta \varphi^\varepsilon_t - \langle \nabla \varphi^\varepsilon_t, \nabla \vartheta^\varepsilon_t \rangle \Big) \rho^\varepsilon_t \,\mathrm{d}\mathfrak{m} \\ &+ \int \varphi^\varepsilon_t \langle \nabla \chi_R, \nabla \vartheta^\varepsilon_t \rangle \rho^\varepsilon_t \,\mathrm{d}\mathfrak{m} \qquad \text{a.e. } t \in (0, 1). \end{aligned}$$

Integrating and recalling that $\varphi^\varepsilon_t = \frac{\varepsilon}{2} \log \rho^\varepsilon_t - \vartheta^\varepsilon_t$ we deduce

$$\begin{aligned} -\int \chi_R \varphi^\varepsilon_{t_1} \,\mathrm{d}\mu^\varepsilon_{t_1} + \int \chi_R \varphi^\varepsilon_{t_0} \,\mathrm{d}\mu^\varepsilon_{t_0} = &\iint_{t_0}^{t_1} \chi_R \Big(\frac{|\nabla \vartheta^\varepsilon_t|^2}{2} - \frac{\varepsilon^2}{8} |\nabla \log \rho^\varepsilon_t|^2 - \frac{\varepsilon}{2} \Delta \varphi^\varepsilon_t \Big) \rho^\varepsilon_t \,\mathrm{d}t \,\mathrm{d}\mathfrak{m} \\ &+ \iint_{t_0}^{t_1} \varphi^\varepsilon_t \langle \nabla \chi_R, \nabla \vartheta^\varepsilon_t \rangle \rho^\varepsilon_t \,\mathrm{d}t \,\mathrm{d}\mathfrak{m} \end{aligned} \tag{6.2.10}$$

and an application of the dominated convergence theorem allows us to pass to the limit as $R \to \infty$. Indeed, by Proposition 5.1.4, (5.2.5) and the fact that $\mu_t^\varepsilon \in \mathscr{P}(\mathrm{X})$ it follows that $\varphi_t^\varepsilon \in L^1(\mu_t^\varepsilon)$, so that the left-hand side converges to itself without χ_R as $R \to \infty$. By (5.3.9c) we can handle the terms with $|\nabla \vartheta_t^\varepsilon|$ and $|\nabla \log \rho_t^\varepsilon|$ on the right-hand side, while for the one with $\Delta \varphi_t^\varepsilon$ notice that

$$\rho_t^\varepsilon \Delta \varphi_t^\varepsilon = \varepsilon g_t^\varepsilon \Delta f_t^\varepsilon - \varepsilon^{-1} \rho_t^\varepsilon |\nabla \varphi_t^\varepsilon|^2$$

and take into account that $(g_t^\varepsilon) \in C([0,1], L^2(\mathfrak{m}))$, $(\Delta f_t^\varepsilon) \in C((0,1], L^2(\mathfrak{m}))$ by Proposition 5.1.4 and $t \mapsto \rho_t^\varepsilon |\nabla \varphi_t^\varepsilon|^2$ belongs to $C((0,1], L^1(\mathfrak{m}))$ by Lemma 5.3.1. Finally

$$|\varphi_t^\varepsilon \langle \nabla \chi_R, \nabla \vartheta_t^\varepsilon \rangle \rho_t^\varepsilon| \le \frac{1}{2} \rho_t^\varepsilon |\varphi_t^\varepsilon|^2 |\nabla \chi_R| + \frac{1}{2} \rho_t^\varepsilon |\nabla \vartheta_t^\varepsilon|^2 |\nabla \chi_R|$$

and again by Lemma 5.3.1, (5.3.9c) and the fact that $\||\nabla \chi_R|\|_{L^\infty(\mathfrak{m})}$ is bounded as $R \to \infty$ we can pass to the limit also in the last term on the right-hand side of (6.2.10). Hence we get

$$-\int \varphi_{t_1}^\varepsilon \, \mathrm{d}\mu_{t_1}^\varepsilon + \int \varphi_{t_0}^\varepsilon \, \mathrm{d}\mu_{t_0}^\varepsilon = \iint_{t_0}^{t_1} \Big(\frac{|\nabla \vartheta_t^\varepsilon|^2}{2} - \frac{\varepsilon^2}{8} |\nabla \log \rho_t^\varepsilon|^2 - \frac{\varepsilon}{2} \Delta \varphi_t^\varepsilon \Big) \rho_t^\varepsilon \, \mathrm{d}t \, \mathrm{d}\mathfrak{m}$$

As already noticed in the proof of point (i) of Proposition 6.1.1, (μ_t^ε) and $(\nabla \vartheta_t^\varepsilon)$ satisfy the assumptions of Theorem 1.1.4, thus from such theorem we deduce that

$$\iint_{t_0}^{t_1} \frac{|\nabla \vartheta_t^\varepsilon|^2}{2} \rho_t^\varepsilon \, \mathrm{d}t \, \mathrm{d}\mathfrak{m} = \frac{1}{2} \int_{t_0}^{t_1} |\dot{\mu}_t^\varepsilon|^2 \, \mathrm{d}t \ge \frac{1}{2(t_1 - t_0)} W_2^2(\mu_{t_0}^\varepsilon, \mu_{t_1}^\varepsilon).$$

Therefore

$$-\int \varphi_{t_1}^\varepsilon \, \mathrm{d}\mu_{t_1}^\varepsilon + \int \varphi_{t_0}^\varepsilon \, \mathrm{d}\mu_{t_0}^\varepsilon \ge \frac{1}{2(t_1 - t_0)} W_2^2(\mu_{t_0}^\varepsilon, \mu_{t_1}^\varepsilon) + \iint_{t_0}^{t_1} \Big(-\frac{\varepsilon^2}{8} |\nabla \log \rho_t^\varepsilon|^2 - \frac{\varepsilon}{2} \Delta \varphi_t^\varepsilon \Big) \rho_t^\varepsilon \, \mathrm{d}t \, \mathrm{d}\mathfrak{m}.$$

We now pass to the limit in $\varepsilon = \varepsilon_n \downarrow 0$: the left hand side trivially converges to the left hand side of (6.2.9) while $W_2^2(\mu_{t_0}^{\varepsilon_n}, \mu_{t_1}^{\varepsilon_n}) \to W_2^2(\mu_{t_0}, \mu_{t_1})$, the contribution of the term with $|\nabla \log \rho_t^\varepsilon|$ vanishes by (5.3.12b) and so does the one with $\Delta \varphi_t^\varepsilon$, the reason being the following: arguing as for (5.3.12a), note that for any $\delta \in (0,1]$

$$\varepsilon \iint_\delta^1 \rho_t^\varepsilon |\Delta \varphi_t^\varepsilon| \, \mathrm{d}t \, \mathrm{d}\mathfrak{m} \le \varepsilon \sqrt{1-\delta} \sqrt{\iint_\delta^1 \rho_t^\varepsilon |\Delta \varphi_t^\varepsilon|^2 \, \mathrm{d}t \, \mathrm{d}\mathfrak{m}}$$

and the last square root is uniformly bounded in $\varepsilon \in (0,1)$ thanks to (5.3.10b) and (5.3.9c). Hence (6.2.9) is proved.

Now notice that (6.2.1a) can be rewritten as

$$-(t_1 - t_0)\varphi_{t_1} = \big((t_1 - t_0)\varphi_{t_0}\big)^c,$$

so that in particular $-(t_1 - t_0)\varphi_{t_1}$ is c-concave and $(-(t_1 - t_0)\varphi_{t_1})^c \ge (t_1 - t_0)\varphi_{t_0}$. Hence both (6.2.1b) and the fact that $-(t_1 - t_0)\varphi_{t_1}$ is a Kantorovich potential follow from

$$\begin{aligned} \frac{1}{2} W_2^2(\mu_{t_0}, \mu_{t_1}) &\ge \int -(t_1 - t_0)\varphi_{t_1} \, \mathrm{d}\mu_{t_1} + \int (-(t_1 - t_0)\varphi_{t_1})^c \, \mathrm{d}\mu_{t_0} \\ &\ge \int -(t_1 - t_0)\varphi_{t_1} \, \mathrm{d}\mu_{t_1} + \int (t_1 - t_0)\varphi_{t_0} \, \mathrm{d}\mu_{t_0} \overset{(6.2.9)}{\ge} \frac{1}{2} W_2^2(\mu_{t_0}, \mu_{t_1}) \end{aligned}$$

The claims about (ψ_t) are proved in the same way.
(μ_t) **is a geodesic**. Let $[t_0, t_1] \subset (0,1)$, pick $t \in [0,1]$ and put $t_0' := (1-t)t_1 + tt_0$. We know that $-(t_1 - t_0)\varphi_{t_1}$ and $-t(t_1 - t_0)\varphi_{t_1}$ are Kantorovich potentials from μ_{t_1} to μ_{t_0} and from μ_{t_1} to $\mu_{t_0'}$ respectively and thus by point (ii) of Theorem 1.2.6 we deduce

$$\begin{aligned} W_2^2(\mu_{t_0}, \mu_{t_1}) &= \int |\mathrm{d}((t_1 - t_0)\varphi_{t_1})|^2 \, \mathrm{d}\mu_{t_1} = \frac{1}{t^2} \int |\mathrm{d}((t_1 - t_0')\varphi_{t_1})|^2 \, \mathrm{d}\mu_{t_1} \\ &= \frac{(t_1 - t_0)^2}{(t_1 - t_0')^2} W_2^2(\mu_{t_1}, \mu_{t_0'}). \end{aligned}$$

Swapping the roles of t_0, t_1 and using the ψ's in place of the φ's we then get

$$W_2(\mu_{t_1'}, \mu_{t_0'}) = \frac{t_1' - t_0'}{t_1 - t_0} W_2(\mu_{t_1}, \mu_{t_0}) \qquad \forall [t_0', t_1'] \subset [t_0, t_1] \subset (0,1).$$

This grants that the restriction of (μ_t) to any interval $[t_0, t_1] \subset (0,1)$ is a constant speed geodesic. Since (μ_t) is continuous on the whole $[0,1]$, this gives the conclusion. Since in this situation the W_2-geodesic connecting μ_0 to μ_1 is unique (recall point (i) of Theorem 1.2.6), by the arbitrariness of the subsequences chosen we also proved the uniqueness of the limit curve (μ_t). □

Remark 6.2.2 (The vanishing viscosity limit)**.** The part of this last proposition concerning the properties of the φ_t^ε's is valid in a context wider than the one provided by Schrödinger problem: we could restate the result by saying that if (φ_t^ε) solves

$$\frac{\mathrm{d}}{\mathrm{d}t}\varphi_t^\varepsilon = \frac{1}{2}|\nabla \varphi_t^\varepsilon|^2 + \frac{\varepsilon}{2}\Delta \varphi_t^\varepsilon \tag{6.2.11}$$

and φ_0^ε uniformly converges to some φ_0, then φ_t^ε uniformly converges to $\varphi_t := -Q_t(-\varphi_0)$.

In this direction, it is worth recalling that in [2] and [50] it has been developed a theory of viscosity solutions for some first-order Hamilton-Jacobi equations on metric spaces. This theory applies in particular to the equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\varphi_t = \frac{1}{2}\mathrm{lip}(\varphi_t)^2 \tag{6.2.12}$$

whose only viscosity solution is given by the formula $\varphi_t := -Q_t(-\varphi_0)$.

Therefore, we have just proved that if one works not only on a metric space, but on a metric measure space which is a $\mathsf{RCD}^*(K,N)$ space, then the solutions of the viscous approximation (6.2.11) converge to the unique viscosity solution of (6.2.12), in accordance with the classical case. ■

Remark 6.2.3. It is not clear whether the 'full' families $\varphi_t^\varepsilon, \psi_t^\varepsilon$ converge as $\varepsilon \downarrow 0$ to a unique limit. This is related to the non-uniqueness of the Kantorovich potentials in the classical optimal transport problem. ■

We shall now make use of the following lemma. It could be directly deduced from the results obtained by Cheeger in [28]; however, the additional regularity assumptions on both the space and the function allow for a 'softer' argument based on the metric Brenier's theorem, which we propose and is new to our knowledge.

Lemma 6.2.4. *Let* $(\mathrm{Y}, \mathsf{d}_{\mathrm{Y}}, \mathfrak{m}_{\mathrm{Y}})$ *be a* $\mathsf{RCD}^*(K,N)$ *space with* $K \in \mathbb{R}$ *and* $N \in [1,\infty)$ *and let* $\phi : \mathrm{X} \to \mathbb{R} \cup \{-\infty\}$ *be a c-concave function not identically* $-\infty$*. Let* Ω *be the interior of the set* $\{\phi > -\infty\}$*. Then* ϕ *is locally Lipschitz on* Ω *and*

$$\operatorname{lip}\phi = |\mathrm{d}\phi|, \quad \mathfrak{m}\text{-}a.e. \text{ on } \Omega.$$

proof Lemma 3.3 in [62] grants that ϕ is locally Lipschitz on Ω and that $\partial^c\phi(x) \neq \emptyset$ for every $x \in \Omega$. The same lemma also grants that for $K \subset \Omega$ compact, the set $\cup_{x\in K}\partial^c\phi(x)$ is bounded. Recalling that $\partial^c\phi$ is the set of $(x,y) \in \mathrm{Y}^2$ such that

$$\phi(x) + \phi^c(y) = \frac{1}{2}\mathsf{d}^2(x,y)$$

and that ϕ, ϕ^c are upper semicontinuous, we see that $\partial^c\phi$ is closed. Hence for $K \subset \Omega$ compact the set $\cup_{x\in K}\partial^c\phi(x)$ is compact and not empty and thus by the Kuratowski-Ryll-Nardzewski Borel selection theorem (see Theorem A.2.1) we deduce the existence of a Borel map $T : \Omega \to \mathrm{Y}$ such that $T(x) \in \partial^c\phi(x)$ for every $x \in \Omega$.

Pick $\mu \in \mathscr{P}_2(\mathrm{Y})$ with $\operatorname{supp}(\mu) \subset\subset \Omega$ and $\mu \leq C\mathfrak{m}$ for some $C > 0$ and set $\nu := T_*\mu$. By construction, μ, ν have both bounded support, T is an optimal map and ϕ is a Kantorovich potential from μ to ν.

Hence point (*iii*) of Theorem 1.2.6 applies and since $\operatorname{lip}\phi = \max\{|D^+\phi|, |D^-\phi|\}$, by the arbitrariness of μ to conclude it is sufficient to show that $|D^+\phi| = |D^-\phi|$ $\mathfrak{m}$-a.e. This easily follows from the fact that $\mathfrak{m}$ is doubling and ϕ Lipschitz, see Proposition 2.7 in [6]. □

With this said, we can now show that the energies of the Schrödinger potentials converge in a localized sense to the energy of the limit ones:

Proposition 6.2.5. *With the same assumptions and notations as in Setting 5.1.3 the following holds.*

Let $\varepsilon_n \downarrow 0$ *be a sequence such that* $(\varphi_t^{\varepsilon_n}), (\psi_t^{\varepsilon_n})$ *converge to limit curves* $(\varphi_t), (\psi_t)$ *as in Proposition 6.1.1 and let* χ *be a cut-off function with bounded support. Then for every* $\delta \in (0,1)$ *we have*

$$\begin{aligned} \lim_{n\to\infty} \iint_\delta^1 \chi |\mathrm{d}\varphi_t^{\varepsilon_n}|^2 \,\mathrm{d}t\,\mathrm{d}\mathfrak{m} &= \iint_\delta^1 \chi |\mathrm{d}\varphi_t|^2 \,\mathrm{d}t\,\mathrm{d}\mathfrak{m}, \\ \lim_{n\to\infty} \iint_0^{1-\delta} \chi |\mathrm{d}\psi_t^{\varepsilon_n}|^2 \,\mathrm{d}t\,\mathrm{d}\mathfrak{m} &= \iint_0^{1-\delta} \chi |\mathrm{d}\psi_t|^2 \,\mathrm{d}t\,\mathrm{d}\mathfrak{m}. \end{aligned} \tag{6.2.13}$$

proof Fix $\delta \in (0,1)$, let χ be a cut-off function with bounded support, so that $t \mapsto \int \chi\varphi_t^\varepsilon \,\mathrm{d}\mathfrak{m}$ is in $AC([\delta,1])$, and notice that from the formula for $\frac{\mathrm{d}}{\mathrm{d}t}\varphi_t^\varepsilon$ we get

$$\int \chi\big(\varphi_1^\varepsilon - \varphi_\delta^\varepsilon\big)\mathrm{d}\mathfrak{m} = \frac{1}{2}\iint_\delta^1 \chi\big(|\mathrm{d}\varphi_t^\varepsilon|^2 + \varepsilon\Delta\varphi_t^\varepsilon\big)\mathrm{d}t\,\mathrm{d}\mathfrak{m}.$$

Choosing $\varepsilon := \varepsilon_n$, letting $n \to \infty$ and using the uniform bound (5.2.11) with $B = \operatorname{supp}(\chi)$ we obtain that

$$\lim_{n\to\infty} \frac{1}{2}\iint_\delta^1 \chi |\mathrm{d}\varphi_t^{\varepsilon_n}|^2 \,\mathrm{d}t\,\mathrm{d}\mathfrak{m} = \lim_{n\to\infty} \int \chi\big(\varphi_1^{\varepsilon_n} - \varphi_\delta^{\varepsilon_n}\big)\mathrm{d}\mathfrak{m} = \int \chi\big(\varphi_1 - \varphi_\delta\big)\mathrm{d}\mathfrak{m}. \tag{6.2.14}$$

Combining (1.1.2) and (6.2.1a) we see that for any $x \in \mathrm{X}$ it holds

$$\frac{\mathrm{d}}{\mathrm{d}t}\varphi_t(x) = \frac{1}{2}\big((\operatorname{lip}\varphi_t)(x)\big)^2 \qquad a.e.\ t \in [0,1].$$

By Fubini's theorem we see that the same identity holds for $\mathscr{L}^1 \times \mathfrak{m}$-a.e. $(t,x) \in [\delta,1] \times \mathrm{X}$. The identity (6.2.1a) also grants that φ_t is a multiple of a c-concave function, thus the thesis of Lemma 6.2.4 is valid for φ_t and recalling that $(\chi\varphi_t) \in AC_{loc}((0,1],L^1(\mathrm{X}))$ by Proposition 6.1.1 we deduce that

$$\int \chi(\varphi_1 - \varphi_\delta)\mathrm{d}\mathfrak{m} = \int_\delta^1 \frac{\mathrm{d}}{\mathrm{d}t}\int \chi\varphi_t \mathrm{d}\mathfrak{m}\,\mathrm{d}t = \iint_\delta^1 \chi\frac{|\mathrm{d}\varphi_t|^2}{2}\,\mathrm{d}t\,\mathrm{d}\mathfrak{m},$$

which together with (6.2.14) gives the first in (6.2.13). The second is proved in the same way. □

Corollary 6.2.6. *With the same assumptions and notations as in Setting 5.1.3 the following holds.*

Let $\varepsilon_n \downarrow 0$ be a sequence such that $(\varphi_t^{\varepsilon_n}),(\psi_t^{\varepsilon_n})$ converge to limit curves $(\varphi_t),(\psi_t)$ as in Proposition 6.1.1. Then for every $\delta \in (0,1)$ and for every continuous cut-off function χ with bounded support we have

$$\begin{array}{rcll} (\chi\mathrm{d}\varphi_t^{\varepsilon_n}) & \to & (\chi\mathrm{d}\varphi_t) & \textit{in} \quad L^2([\delta,1],L^2(T^*\mathrm{X})) \\ (\chi\mathrm{d}\psi_t^{\varepsilon_n}) & \to & (\chi\mathrm{d}\psi_t) & \textit{in} \quad L^2([0,1-\delta],L^2(T^*\mathrm{X})) \\ (\chi\mathrm{d}\varphi_t^{\varepsilon_n}\otimes\mathrm{d}\varphi_t^{\varepsilon_n}) & \to & (\chi\mathrm{d}\varphi_t\otimes\mathrm{d}\varphi_t) & \textit{in} \quad L^2([\delta,1],L^2((T^*)^{\otimes 2}\mathrm{X})) \\ (\chi\mathrm{d}\psi_t^{\varepsilon_n}\otimes\mathrm{d}\psi_t^{\varepsilon_n}) & \to & (\chi\mathrm{d}\psi_t\otimes\mathrm{d}\psi_t) & \textit{in} \quad L^2([0,1-\delta],L^2((T^*)^{\otimes 2}\mathrm{X})) \\ (\chi\mathrm{d}\varphi_t^{\varepsilon_n}\otimes\mathrm{d}\psi_t^{\varepsilon_n}) & \to & (\chi\mathrm{d}\varphi_t\otimes\mathrm{d}\psi_t) & \textit{in} \quad L^2([\delta,1-\delta],L^2((T^*)^{\otimes 2}\mathrm{X})) \end{array} \tag{6.2.15}$$

proof Let χ be as in the statement and start noticing that the closure of the differential grants that $\chi\mathrm{d}\varphi_t^{\varepsilon_n} \rightharpoonup \chi\mathrm{d}\varphi_t$ in $L^2(T^*\mathrm{X})$ for all $t \in (0,1]$. This and the fact that $(\chi\mathrm{d}\varphi_t^{\varepsilon_n})$ is equibounded in $L^2([\delta,1],L^2(T^*\mathrm{X}))$, as a direct consequence of (5.2.10a), are sufficient to ensure that $(\chi\mathrm{d}\varphi_t^{\varepsilon_n}) \rightharpoonup (\chi\mathrm{d}\varphi_t)$ in $L^2([\delta,1],L^2(T^*\mathrm{X}))$. Given that the first in (6.2.13) grants (local) convergence of the $L^2([\delta,1],L^2(T^*\mathrm{X}))$-norms, we deduce (local) strong convergence. This establishes the first limit.

Now observe that for every $\omega \in L^2([\delta,1],L^2(T^*\mathrm{X}))$ the fact that $\sqrt{\chi}|\mathrm{d}\varphi_t^{\varepsilon_n}|$ is uniformly bounded in $L^\infty([\delta,1]\times\mathrm{X})$ and the strong L^2-convergence just proved ensure that

$$\langle\sqrt{\chi}\mathrm{d}\varphi_t^{\varepsilon_n},\omega_t\rangle \quad\to\quad \langle\sqrt{\chi}\mathrm{d}\varphi_t,\omega_t\rangle \qquad \text{in} \quad L^2([\delta,1]\times\mathrm{X}).$$

It follows that for any $\omega_1,\omega_2 \in L^2([\delta,1],L^2(T^*\mathrm{X}))$ we have

$$\iint_\delta^1 \langle\sqrt{\chi}\mathrm{d}\varphi_t^{\varepsilon_n},\omega_{1,t}\rangle\,\langle\sqrt{\chi}\mathrm{d}\varphi_t^{\varepsilon_n},\omega_{2,t}\rangle\,\mathrm{d}t\,\mathrm{d}\mathfrak{m} \quad\to\quad \iint_\delta^1 \langle\sqrt{\chi}\mathrm{d}\varphi_t,\omega_{1,t}\rangle\,\langle\sqrt{\chi}\mathrm{d}\varphi_t,\omega_{2,t}\rangle\,\mathrm{d}t\,\mathrm{d}\mathfrak{m}$$

and thus to conclude it remains to prove that

$$\iint_\delta^1 |\chi\mathrm{d}\varphi_t^{\varepsilon_n}\otimes\mathrm{d}\varphi_t^{\varepsilon_n}|_{\mathsf{HS}}^2\,\mathrm{d}t\,\mathrm{d}\mathfrak{m} \quad\to\quad \iint_\delta^1 |\chi\mathrm{d}\varphi_t\otimes\mathrm{d}\varphi_t|_{\mathsf{HS}}^2\,\mathrm{d}t\,\mathrm{d}\mathfrak{m}.$$

Since $|v\otimes v|_{\mathsf{HS}}^2 = |v|^4$ this is a direct consequence of the fact that $\sqrt{\chi}|\mathrm{d}\varphi_t^{\varepsilon_n}|$ are uniformly bounded and converge to $\sqrt{\chi}|\mathrm{d}\varphi_t|$ in $L^2([\delta,1]\times\mathrm{X})$. Hence also the third limit is established.

The other claims follow by analogous arguments. □

The estimates that we have for the functions φ's tell nothing about their regularity as $t \downarrow 0$ and similarly little we know so far about the ψ's for $t \uparrow 1$. We now see in which sense limit functions φ_0, ψ_1 exist. This is not needed for the proof of our main result, but we believe it is relevant in its own.

Thus let us fix $\varepsilon_n \downarrow 0$ so that $\varphi_t^{\varepsilon_n} \to \varphi_t$ for $t \in (0,1]$ and $\psi_t^{\varepsilon_n} \to \psi_t$ for $t \in [0,1)$ as in Proposition 6.1.1. Then define the functions $\varphi_0, \psi_1 : \mathrm{X} \to \mathbb{R} \cup \{-\infty\}$ as

$$\begin{aligned} \varphi_0(x) &:= \inf_{t\in(0,1]} \varphi_t(x) = \lim_{t\downarrow 0} \varphi_t(x), \\ \psi_1(x) &:= \inf_{t\in[0,1)} \psi_t(x) = \lim_{t\uparrow 1} \psi_t(x). \end{aligned} \tag{6.2.16}$$

Notice that the fact that the inf are equal to the stated limits is a consequence of formulas (6.2.1a), (6.2.2a), which directly imply that for every $x \in \mathrm{X}$ the maps $t \mapsto \varphi_t(x)$ and $t \mapsto \psi_{1-t}(x)$ are non-decreasing.

The main properties of φ_0, ψ_1 are collected in the following proposition:

Proposition 6.2.7. *With the same assumptions and notations as in Setting 5.1.3 and for φ_0, ψ_1 defined by* (6.2.16) *the following holds.*

i) The functions $-\varphi_t$ (resp. $-\psi_t$) Γ-converge to $-\varphi_0$ (resp. $-\psi_1$) as $t \downarrow 0$ (resp. $t \uparrow 1$).

ii) For every $t \in (0,1]$ we have

$$Q_t(-\varphi_0) = -\varphi_t \qquad\qquad Q_t(-\psi_1) = -\psi_{1-t}.$$

iii) It holds

$$\varphi_0(x) = \begin{cases} -\psi_0(x) & \text{if } x \in \mathrm{supp}(\rho_0) \\ -\infty & \text{otherwise} \end{cases} \qquad\qquad \psi_1(x) = \begin{cases} -\varphi_1(x) & \text{if } x \in \mathrm{supp}(\rho_1) \\ -\infty & \text{otherwise} \end{cases}$$

iv) We have

$$\int \varphi_0 \rho_0 \, \mathrm{d}\mathfrak{m} + \int \psi_1 \rho_1 \, \mathrm{d}\mathfrak{m} = \frac{1}{2} W_2^2(\mu_0, \mu_1).$$

v) Define φ_0^ε on $\mathrm{supp}(\rho_0)$ as $\varphi_0^\varepsilon := \varepsilon \log(f^\varepsilon)$ and let $\varepsilon_n \downarrow 0$ be such that $\varphi_t^{\varepsilon_n}, \psi_t^{\varepsilon_n}$ converge to φ_t, ψ_t as $n \to \infty$ as in Proposition 6.1.1.

Then the functions $\rho_0 \varphi_0^{\varepsilon_n}$, set to be 0 on $\mathrm{X} \setminus \mathrm{supp}(\rho_0)$, converge to $\rho_0 \varphi_0$ in $L^\infty(\mathfrak{m})$ as $n \to \infty$.

With the analogous definition of $\rho_1 \psi_1^{\varepsilon_n}$ we have that these converge to $\rho_1 \psi_1$ in $L^\infty(\mathfrak{m})$ as $n \to \infty$.

proof We shall prove the claims for φ_0 only, as those for ψ_1 follow along similar lines.
(i) For the $\Gamma - \overline{\lim}$ inequality we simply observe that by definition $-\varphi_0(x) = \lim_{t\downarrow 0} -\varphi_t(x)$. To prove the $\Gamma - \underline{\lim}$ inequality, use the fact that $-\varphi_t \geq -\varphi_s$ for $0 < t \leq s$ and the continuity of φ_s: for given (x_t) converging to x we have

$$\underline{\lim}_{t\downarrow 0} -\varphi_t(x_t) \geq \underline{\lim}_{t\downarrow 0} -\varphi_s(x_t) = -\varphi_s(x) \qquad \forall s > 0.$$

The conclusion follows letting $s \downarrow 0$.

(ii) This claim follows from the general properties of Γ-convergence and fine manipulations of the Gaussian estimates (1.2.11). From $-\varphi_0 \geq -\varphi_s$ we deduce that

$$Q_t(-\varphi_0) \geq Q_t(-\varphi_s) \overset{(6.2.1a)}{=} -\varphi_{t+s} \qquad \forall s \in (0,1]$$

and thus letting $s \downarrow 0$ and using the continuity of $(0,1] \ni t \mapsto \varphi_t(x)$ for all $x \in \mathrm{X}$ we obtain $Q_t(-\varphi_0)(x) \geq -\varphi_t(x)$ for all $x \in \mathrm{X}$. For the opposite inequality, use the representation formula (1.2.9c), the Gaussian estimates (1.2.11) and the fact that ρ_0 and f^ε have the same support to get

$$\frac{C_1}{V_{\varepsilon t/2}} \|f^\varepsilon\|_{L^1(\mathfrak{m})} \exp\Big(-\frac{C_2 D_0^2(\cdot)}{\varepsilon t}\Big) \leq f_t^\varepsilon \leq \frac{C_3}{v_{\varepsilon t/2}} \|f^\varepsilon\|_{L^1(\mathfrak{m})} \exp\Big(-\frac{C_4 d_0^2(\cdot)}{\varepsilon t}\Big)$$

for all $t \in (0,1]$, where $v_{\varepsilon t/2}, V_{\varepsilon t/2}$ are defined as in (5.1.6), the constants $C_1, ..., C_4$ are positive and only depend on $K, N, \mathrm{supp}(\rho_0)$ and

$$d_0(x) := \inf_{y \in \mathrm{supp}(\rho_0)} \mathsf{d}(x,y) \qquad D_0(x) := \sup_{y \in \mathrm{supp}(\rho_0)} \mathsf{d}(x,y).$$

From this two-sided bound we deduce the following one

$$\begin{aligned} \varepsilon \log C_1 - \varepsilon \log V_{\varepsilon t/2} + \varepsilon \log \|f^\varepsilon\|_{L^1(\mathfrak{m})} - \frac{C_2 D_0^2(\cdot)}{t} &\leq \varphi_t^\varepsilon \\ &\leq \varepsilon \log C_3 - \varepsilon \log v_{\varepsilon t/2} + \varepsilon \log \|f^\varepsilon\|_{L^1(\mathfrak{m})} - \frac{C_4 d_0^2(\cdot)}{t} \end{aligned}$$

for all $t \in (0,1]$. We start claiming that

$$\lim_{\varepsilon \downarrow 0} \varepsilon \log v_\varepsilon = \lim_{\varepsilon \downarrow 0} \varepsilon \log V_\varepsilon = 0. \tag{6.2.17}$$

Indeed on one side since $\mathrm{supp}(\rho_0)$ is compact, we have $\varepsilon \log v_\varepsilon \leq \varepsilon \log V_\varepsilon \leq \varepsilon \log \mathfrak{m}(B)$, where B is a sufficiently large set containing all the ε-enlargements of $\mathrm{supp}(\rho_0)$ for $\varepsilon \in (0,1)$, so that $\limsup_\varepsilon \varepsilon \log V_\varepsilon \leq 0$. On the other one, letting C be the local doubling constant of B and D its diameter we have

$$\mathfrak{m}(B_{\sqrt{\varepsilon}}(y)) \geq C^{\log_2(D/\sqrt{\varepsilon})+1} \mathfrak{m}(B) \qquad \forall y \in \mathrm{supp}(\rho_0).$$

Thus $v_\varepsilon \geq C^{\log_2(D/\sqrt{\varepsilon})+1} \mathfrak{m}(B)$ from which it follows that $\varliminf_\varepsilon \varepsilon \log v_\varepsilon \geq 0$ and thus (6.2.17) is proved. Secondly, we claim that there exists a constant $M > 0$ such that

$$\varepsilon |\log \|f^\varepsilon\|_{L^1(\mathfrak{m})}| \leq M \qquad \forall \varepsilon \in (0,1). \tag{6.2.18}$$

To this aim start observing that by (5.2.2)

$$\varepsilon \log \Big(\|f^\varepsilon\|_{L^1(\mathfrak{m})} \|g^\varepsilon\|_{L^1(\mathfrak{m})} \Big) \leq \varepsilon \log C + \varepsilon \log V_{\varepsilon/2} + D^2$$

where C is a positive constant independent of ε and D is the diameter of $\mathrm{supp}(\rho_0) \cup \mathrm{supp}(\rho_1)$. On the other hand, taking into account the normalization choice for $(f^\varepsilon, g^\varepsilon)$ (see Setting 5.1.3), Jensen's inequality for $-\log$ and the L^1-contractivity of the heat flow it holds

$$0 = \int \log(\mathsf{h}_{\frac{\varepsilon}{2}} f^\varepsilon) \rho_1 \, \mathrm{d}\mathfrak{m} \leq \log \int \mathsf{h}_{\frac{\varepsilon}{2}} f^\varepsilon \rho_1 \, \mathrm{d}\mathfrak{m} \leq \log \Big(\|f^\varepsilon\|_{L^1(\mathfrak{m})} \|\rho_1\|_{L^\infty(\mathfrak{m})} \Big),$$

whence $\log \|f^\varepsilon\|_{L^1(\mathfrak{m})} \geq -\log \|\rho_1\|_{L^\infty(\mathfrak{m})}$ for all $\varepsilon \in (0,1)$. Arguing in an analogous way for g^ε and recalling that $\rho_1 = g^\varepsilon \mathsf{h}_{\varepsilon/2} f^\varepsilon$ we have

$$H(\mu_1 \,|\, \mathfrak{m}) = \int \rho_1 \log \rho_1 \, \mathrm{d}\mathfrak{m} = \int \log(g^\varepsilon)\rho_1 \, \mathrm{d}\mathfrak{m} \leq \log \int g^\varepsilon \rho_1 \, \mathrm{d}\mathfrak{m} \leq \log \Big(\|g^\varepsilon\|_{L^1(\mathfrak{m})} \|\rho_1\|_{L^\infty(\mathfrak{m})} \Big),$$

whence $\log \|g^\varepsilon\|_{L^1(\mathfrak{m})} \geq H(\mu_1 \,|\, \mathfrak{m}) - \log \|\rho_1\|_{L^\infty(\mathfrak{m})}$ for all $\varepsilon \in (0,1)$. Putting all these pieces of information together with (6.2.17), we see that (6.2.18) follows. Therefore, passing to the limit as $\varepsilon \downarrow 0$ in the previous two-sided bound and recalling the locally uniform convergence of φ_t^ε to φ_t for all $t \in (0,1]$, we get

$$\frac{C_4 d_0^2(\cdot)}{t} - M \leq -\varphi_t \leq \frac{C_2 D_0^2(\cdot)}{t} + M.$$

Hence if we fix $x \in \mathrm{X}$ and a sequence $t_n \downarrow 0$, by coercivity we can find $y_n \in \mathrm{X}$ such that

$$Q_t(-\varphi_{t_n})(x) = \frac{\mathsf{d}^2(x, y_n)}{2t} - \varphi_{t_n}(y_n).$$

Actually, the functions $(-\varphi_{t_n})_{n\in\mathbb{N}}$ are equi-coercive and thus, by compactness, up to pass to a subsequence we can assume that $y_n \to y$ for some $y \in \mathrm{X}$, so that taking into account the $\Gamma - \underline{\lim}$ inequality previously proved we get

$$\begin{aligned} \frac{\mathsf{d}^2(x,y)}{2t} - \varphi_0(y) &\leq \varliminf_{n\to\infty} \frac{\mathsf{d}^2(x,y_n)}{2t} - \varphi_{t_n}(y_n) = \varliminf_{n\to\infty} Q_t(-\varphi_{t_n})(x) \\ &\overset{(6.2.1a)}{=} \varliminf_{n\to\infty} -\varphi_{t_n+t}(x) = -\varphi_t(x) \end{aligned}$$

which shows that $Q_t(-\varphi_0)(x) \leq -\varphi_t(x)$, as desired.
(iii) For any $t \in (0,1]$ we have

$$\varphi_0 \leq \varphi_t \overset{(6.1.3)}{\leq} -\psi_t$$

so that letting $t \downarrow 0$ and using the continuity of $[0,1) \ni t \mapsto \psi_t(x)$ for all $x \in \mathrm{X}$ we deduce that

$$\varphi_0 \leq -\psi_0 \qquad \text{on } \mathrm{X}.$$

Now notice that the fact that $-\varphi_0 \leq \Gamma - \underline{\lim}(-\varphi_t)$ implies that

$$\varphi_0(\gamma_0) \geq \varlimsup_{t\downarrow 0} \varphi_t(\gamma_t) \qquad \forall \gamma \in C([0,1], \mathrm{X}). \tag{6.2.19}$$

Let $\boldsymbol{\pi}$ be the lifting of the W_2-geodesic (μ_t) (recall point (i) of Theorem 1.2.6); taking into account that the evaluation maps $\mathrm{e}_t : C([0,1], \mathrm{X}) \to \mathrm{X}$ are continuous and that $\mathrm{supp}(\boldsymbol{\pi})$ is a compact subset of $C([0,1], \mathrm{X})$, because given by constant speed geodesics running from the compact set $\mathrm{supp}(\rho_0)$ to the compact $\mathrm{supp}(\rho_1)$, it is easy to see that for every $\gamma \in \mathrm{supp}(\boldsymbol{\pi})$ and $t \in [0,1]$ we have $\gamma_t \in \mathrm{supp}(\mu_t)$ and viceversa for every $x \in \mathrm{supp}(\mu_t)$ there is $\gamma \in \mathrm{supp}(\boldsymbol{\pi})$ with $\gamma_t = x$.

Thus let $x \in \mathrm{supp}(\rho_0) = \mathrm{supp}(\mu_0)$ and find $\gamma \in \mathrm{supp}(\boldsymbol{\pi})$ with $\gamma_0 = x$: from the fact that $\gamma_t \in \mathrm{supp}(\mu_t)$ and (6.1.3) we get

$$\varphi_0(x) \overset{(6.2.19)}{\geq} \varlimsup_{t\downarrow 0} \varphi_t(\gamma_t) = \varlimsup_{t\downarrow 0} -\psi_t(\gamma_t) = -\psi_0(x).$$

Thus to conclude it remains to prove that $\varphi_0 = -\infty$ outside $\mathrm{supp}(\rho_0)$ and to this aim we shall use the Gaussian estimates (1.2.11). First notice that (6.2.17) still holds if we slightly change the definition of v_s and V_s as follows

$$v_s := \inf_{y\in \mathrm{supp}(\rho_0)\cup\mathrm{supp}(\rho_1)} \mathfrak{m}(B_{\sqrt{s}}(y)) \qquad V_s := \sup_{y\in \mathrm{supp}(\rho_0)\cup\mathrm{supp}(\rho_1)} \mathfrak{m}(B_{\sqrt{s}}(y)).$$

Then, observing that by construction we have $\mathrm{supp}(f^\varepsilon) = \mathrm{supp}(\rho_0)$ for every $\varepsilon > 0$, the second in (1.2.11) yields

$$f_t^\varepsilon(x) = \mathsf{h}_{\varepsilon t/2} f^\varepsilon(x) = \int f^\varepsilon(y)\mathsf{r}_{\varepsilon t/2}(x,y)\,\mathrm{d}\mathfrak{m}(y) \le \frac{C_2}{v_{\varepsilon t/2}} e^{-\frac{\mathsf{d}^2(x,\mathrm{supp}(\rho_0))}{3\varepsilon t}} \int f^\varepsilon\,\mathrm{d}\mathfrak{m},$$
$$g_t^\varepsilon(x) = \mathsf{h}_{\varepsilon(1-t)/2} g^\varepsilon(x) = \int g^\varepsilon(y)\mathsf{r}_{\varepsilon(1-t)/2}(x,y)\,\mathrm{d}\mathfrak{m}(y) \le \frac{C_2}{v_{\varepsilon(1-t)/2}} \int g^\varepsilon\,\mathrm{d}\mathfrak{m},$$

for every $t \in (0,1)$ and thus coupling these bounds with (5.2.2) we obtain

$$\rho_t^\varepsilon(x) = f_t^\varepsilon(x) g_t^\varepsilon(x) \le \frac{C_1 C_2^2 V_\varepsilon}{v_{\varepsilon(1-t)/2} v_{\varepsilon t/2}} e^{\frac{C_1 D^2}{\varepsilon}} e^{-\frac{\mathsf{d}^2(x,\mathrm{supp}(\rho_0))}{3\varepsilon t}} \qquad \forall x \in \mathrm{X},\ t \in (0,1).$$

Therefore recalling (6.2.17) we obtain

$$\overline{\lim_{\varepsilon\downarrow 0}}\, \varepsilon \log(\rho_t^\varepsilon(x)) \le C_1 D^2 - \frac{\mathsf{d}^2(x,\mathrm{supp}(\rho_0))}{3t} \qquad \forall x \in \mathrm{X},\ t \in (0,1). \tag{6.2.20}$$

Now let $\varepsilon_n \downarrow 0$ be the sequence such that $\varphi_t^{\varepsilon_n}, \psi_t^{\varepsilon_n}$ converge to φ_t, ψ_t as in Proposition 6.1.1 and put $S(x) := \sup_{\varepsilon\in(0,1), t\in[0,1/2]} |\psi_t^\varepsilon(x)| < \infty$. The inequality

$$\varphi_t(x) = \lim_{n\to\infty} \varphi_t^{\varepsilon_n}(x) \le \overline{\lim_{n\to\infty}}\, \varepsilon_n \log(\rho_t^{\varepsilon_n}(x)) - \lim_{n\to\infty} \psi_t^{\varepsilon_n}(x) \overset{(6.2.20)}{\le} S(x) + C_1 D^2 - \frac{\mathsf{d}^2(x,\mathrm{supp}(\rho_0))}{3t}$$

shows that if $x \notin \mathrm{supp}(\rho_0)$ we have $\varphi_0(x) = \lim_{t\downarrow 0} \varphi_t(x) = -\infty$, as desired.
(iv) By the point (iii) just proven we have

$$\int \varphi_0\rho_0\,\mathrm{d}\mathfrak{m} + \int \psi_1\rho_1\,\mathrm{d}\mathfrak{m} = -\int \psi_0\rho_0\,\mathrm{d}\mathfrak{m} - \int \varphi_1\rho_1\,\mathrm{d}\mathfrak{m}$$

so that taking into account the weak continuity of $t \mapsto \mu_t$, the fact that the measures μ_t have equibounded supports and the uniform continuity of $t \mapsto \chi\varphi_t$ (resp. $t \mapsto \chi\psi_t$) for t close to 1 (resp. close to 0), where χ is a cut-off function with bounded support and identically equal to 1 on a set containing $\mathrm{supp}(\mu_t)$ for all $t \in [0,1]$, we get

$$\begin{aligned}\int \varphi_0\rho_0\,\mathrm{d}\mathfrak{m} + \int \psi_1\rho_1\,\mathrm{d}\mathfrak{m} &= \lim_{t\downarrow 0} -\int \psi_t\rho_t\,\mathrm{d}\mathfrak{m} - \int \varphi_{1-t}\rho_{1-t}\,\mathrm{d}\mathfrak{m} \\ &\overset{(6.1.3)}{=} \lim_{t\downarrow 0} \int \varphi_t\rho_t\,\mathrm{d}\mathfrak{m} - \int \varphi_{1-t}\rho_{1-t}\,\mathrm{d}\mathfrak{m} \overset{(6.2.1\mathrm{b})}{=} \frac{1}{2} W_2^2(\mu_0,\mu_1).\end{aligned}$$

(v) Since $\rho_0 \in L^\infty(\mathfrak{m})$, we also have $\rho_0 \log(\rho_0) \in L^\infty(\mathfrak{m})$. The claim then follows from the identity $\rho_0\varphi_0^\varepsilon = \varepsilon\rho_0 \log\rho_0 - \rho_0\psi_0^\varepsilon$, the compactness of $\mathrm{supp}(\rho_0)$, the local uniform convergence of $\psi_0^{\varepsilon_n}$ to ψ_0 as $n \to \infty$ and the fact that $\psi_0 = -\varphi_0$ on $\mathrm{supp}(\rho_0)$. □

Remark 6.2.8 (Entropic and transportation cost)**.** For $\varepsilon > 0$ the *entropic cost* from ρ_0 to ρ_1 is defined as

$$\mathscr{I}_\varepsilon(\rho_0, \rho_1) := \inf H(\boldsymbol{\gamma} \,|\, \mathsf{R}^\varepsilon),$$

the infimum being taken among all transport plans $\boldsymbol{\gamma}$ from $\mu_0 := \rho_0 \mathfrak{m}$ to $\mu_1 := \rho_1 \mathfrak{m}$. Hence with our notation

$$\mathscr{I}_\varepsilon(\rho_0, \rho_1) = H\big(f^\varepsilon \otimes g^\varepsilon \mathsf{R}^\varepsilon \,|\, \mathsf{R}^\varepsilon\big) = \frac{1}{\varepsilon} \int \varphi_0^\varepsilon \oplus \psi_1^\varepsilon f^\varepsilon \otimes g^\varepsilon \, \mathrm{d}\mathsf{R}^\varepsilon = \frac{1}{\varepsilon}\Big(\int \varphi_0^\varepsilon \rho_0 \, \mathrm{d}\mathfrak{m} + \int \psi_1^\varepsilon \rho_1 \, \mathrm{d}\mathfrak{m}\Big)$$

and by $(iv), (v)$ of the previous proposition we get

$$\lim_{\varepsilon \downarrow 0} \varepsilon \mathscr{I}_\varepsilon(\rho_0, \rho_1) = \frac{1}{2} W_2^2(\mu_0, \mu_1). \tag{6.2.21}$$

In other words, after the natural rescaling the entropic cost converges to the quadratic transportation cost, thus establishing another link between the Schrödinger problem and the transport one.

We would like to emphasize that although this argument is new, the result is not, not even on $\mathsf{RCD}^*(K, N)$ spaces: Léonard proved in [79] that the same limit holds in a very abstract setting provided the heat kernel satisfies the appropriate large deviation principle

$$\varepsilon \log \mathsf{r}_t^\varepsilon(x, y) \sim -\frac{\mathsf{d}^2(x, y)}{2}.$$

Since recently such asymptotic behaviour for the heat kernel on $\mathsf{RCD}^*(K, N)$ spaces has been proved by Jiang-Li-Zhang in [71], Léonard's result applies. Thus in this remark we simply wanted to show an alternative proof of such limiting property.

The next step could be the investigation of the first order expansion of (6.2.21), namely the asymptotic behaviour as $\varepsilon \downarrow 0$ of

$$\mathscr{I}_\varepsilon(\rho_0, \rho_1) - \frac{1}{2\varepsilon} W_2^2(\mu_0, \mu_1)$$

or its Γ-limit when μ_0, μ_1 are not fixed. Concerning this second perspective, in the smooth case of $\mathbb{R}^d$ equipped with the Euclidean distance and the weighted measure $\nu := e^{-V} \mathscr{L}^d$, where V is a C^2-regular K-convex function with $K \in \mathbb{R}$, the answer is known from [43] and [46]. There it reads as

$$\mathscr{I}_\varepsilon(\cdot, \rho_1) - \frac{1}{2\varepsilon} W_2^2(\cdot, \mu_1) \xrightarrow{\Gamma} \frac{1}{2} H(\cdot \,|\, \nu) - \frac{1}{2} H(\mu_0 \,|\, \nu)$$

for $\varepsilon \downarrow 0$, where the Γ-limit is understood w.r.t. W_2-convergence. Pay attention to the fact that R^ε is associated to the metric measure space $(\mathbb{R}^d, \mathsf{d}_{\mathrm{Eucl}}, \nu)$, hence it is the joint law $(\mathrm{e}_0, \mathrm{e}_\varepsilon)_* \mathbf{R}$, where $\mathbf{R}$ is the law of a solution $(Z_t)_{t \geq 0}$ to the Fokker-Planck equation

$$\mathrm{d}Z_t = -\nabla V(Z_t) \mathrm{d}t + \sqrt{2} \mathrm{d}W_t$$

with $Z_0 \sim \nu$ and $(W_t)_{t \geq 0}$ a standard $\mathbb{R}^d$-valued Brownian motion. It would be interesting to generalize such result to Riemannian manifolds and possibly to the RCD framework. ■

6.3 Some consequences of the entropic approximation

As shown in [45], the $\mathsf{RCD}^*(K,N)$ condition is equivalent to the (K,N)-convexity of the Boltzmann entropy $H(\cdot\,|\,\mathfrak{m})$ on $\mathscr{P}_2(\mathrm{X}) \cap D(H(\cdot\,|\,\mathfrak{m}))$, i.e. given any two measures $\mu_0, \mu_1 \in \mathscr{P}_2(\mathrm{X})$ with finite entropy w.r.t. $\mathfrak{m}$ there exists a constant speed geodesic $(\mu_t)_{t\in[0,1]}$ connecting them such that the function $H(t) := H(\mu_t\,|\,\mathfrak{m})$ satisfies

$$H'' \geq KW_2^2(\mu_0,\mu_1) + \frac{1}{N}(H')^2 \tag{6.3.1}$$

in distribution sense on $[0,1]$. This is also called **entropic curvature-dimension condition** and denoted by $\mathsf{CD}^e(K,N)$. It is a matter of computation (see Lemma 2.2 in [45]) to show that this in turn equivalent to ask

$$U_N(\mu_t) \geq \sigma_{K,N}^{(1-t)}(W_2(\mu_0,\mu_1))U_N(\mu_0) + \sigma_{K,N}^{(t)}(W_2(\mu_0,\mu_1))U_N(\mu_1) \tag{6.3.2}$$

for all $t \in [0,1]$, where the distortion coefficients $\sigma_{K,N}^{(t)}$ are the ones already introduced in (1.2.2) and

$$U_N(\mu) := \exp\Big(-\frac{1}{N}H(\mu\,|\,\mathfrak{m})\Big).$$

This new curvature-dimension condition will play a crucial role throughout the whole section and, relying on the results of Chapters 5 and 6, we are going to study its connection with entropic interpolations.

Let us begin with a warm-up example, that enables the reader to understand the strategy that will be adopted with slight modifications in the subsequent applications; it also explains why (6.3.2) is more natural for Wasserstein geodesics, while (6.3.1) can be directly checked for entropic interpolations and better fits them.

For $K = 0$ (6.3.1) becomes the following ODE

$$y' = \frac{1}{N}y^2$$

so that by a comparison argument and the fact that $y(t) = -N/t$ is an exact solution to it we deduce that $H' \geq -N/t$ for all $t \in (0,1]$, whence by integration

$$H(\mu_1\,|\,\mathfrak{m}) - H(\mu_t\,|\,\mathfrak{m}) \geq N\log t \tag{6.3.3}$$

for all $t \in (0,1]$. The interest in this inequality could be the following: if $\mathfrak{m}$ is a probability measure and we choose $\mu_1 = \mathfrak{m}$, then $0 \leq H(\mu_t\,|\,\mathfrak{m}) \leq -N\log t$ for all $t \in (0,1]$, regardless of the initial condition μ_0; in this sense (6.3.3) shows a behaviour analogous to the Li-Yau inequality. A simpler way to deduce (6.3.3) is the following: for $K = 0$ the term $\sigma_{K,N}^{(1-t)}(W_2(\mu_0,\mu_1))U_N(\mu_0)$ is positive and thus can be neglected in (6.3.2), which then becomes

$$U_N(\mu_t) \geq tU_N(\mu_1)$$

and this is completely equivalent to (6.3.3) as long as $t \in (0,1]$.

As we are going to see below, the same inequality holds true along entropic interpolations too, for any slowing-down parameter ε, and it is for this reason that we have also presented the ODE approach, although less immediate. It is worth stressing that, unlike all the previous

results, here μ_0, μ_1 need not have bounded supports: we manage to remove this assumption by means of the stability of the Schrödinger problem (see Theorem 4.2.3). As a consequence, the entropic interpolation can not be defined as in Setting 5.1.3, but the problem can be easily overcome since by (4.1.2) and Theorem 5.1.1 the existence of a dynamical solution associated to μ_0, μ_1 and $\mathbf{R}^{\varepsilon/2}$ is granted and thus (μ_t^ε) can and shall be defined as the marginal flow of such solution.

Proposition 6.3.1. *Let* $(\mathrm{X}, \mathsf{d}, \mathfrak{m})$ *be a* $\mathsf{RCD}^*(0, N)$ *space where* $N \in [1, \infty)$ *and* $\mathfrak{m}$ *is a non-negative Radon measure; let* $\mu_0 = \rho_0 \mathfrak{m}$ *and* $\mu_1 = \rho_1 \mathfrak{m}$ *be two absolutely continuous Borel probability measures belonging to* $\mathscr{P}_2(\mathrm{X})$ *with bounded densities.*

Then

$$H(\mu_1 \,|\, \mathfrak{m}) - H(\mu_t^\varepsilon \,|\, \mathfrak{m}) \geq N \log t$$

for all $t \in (0, 1]$ *and for all* $\varepsilon > 0$*, where* $\mu_t^\varepsilon := (\mathrm{e}_t)_* \mathbf{P}^{\varepsilon/2}$*,* $\mathbf{P}^{\varepsilon/2}$ *being the (unique) dynamical solution of the Schrödinger problem associated to* $\mu_0, \mu_1, \mathbf{R}^{\varepsilon/2}$ *(see the end of Section 5.1 for the definition of* $\mathbf{R}^{\varepsilon/2}$*).*

proof Fix $\varepsilon > 0$, assume that ρ_0, ρ_1 also have bounded supports (so that the definition of μ_t^ε coincides with the one provided in Setting 5.1.3), set $H(t) := H(\mu_t^\varepsilon \,|\, \mathfrak{m})$ and observe that by Proposition 5.3.3 $H \in C([0, 1]) \cap AC_{loc}((0, 1))$ with

$$H(t) - H(1) = -\int_t^1 H'(s) \mathrm{d}s = -\iint_t^1 \langle \nabla \rho_s^\varepsilon, \nabla \vartheta_s^\varepsilon \rangle \mathrm{d}s \mathrm{d}\mathfrak{m}.$$

Using a cut-off argument analogous to the one adopted in Proposition 5.2.2 to handle (5.2.7) it is not difficult to see that

$$\int \langle \nabla \rho_t^\varepsilon, \nabla \vartheta_t^\varepsilon \rangle \mathrm{d}\mathfrak{m} = -\int \rho_t^\varepsilon \Delta \vartheta_t^\varepsilon \mathrm{d}\mathfrak{m},$$

thus plugging this identity into the previous line and recalling that $\vartheta_t^\varepsilon = \frac{\varepsilon}{2} \log \rho_t^\varepsilon - \varphi_t^\varepsilon$, we get

$$H(t) - H(1) = \frac{\varepsilon}{2} \iint_t^1 \rho_s^\varepsilon \Delta \log \rho_s^\varepsilon \mathrm{d}s \mathrm{d}\mathfrak{m} - \iint_t^1 \rho_s^\varepsilon \Delta \varphi_s^\varepsilon \mathrm{d}s \mathrm{d}\mathfrak{m}.$$

Now observe that Li-Yau inequality for φ_t^ε reads as $\Delta \varphi_t^\varepsilon \geq -N/t$ (because we have assumed $K = 0$), while the first term on the right-hand side can be integrated by parts arguing as above. This yields

$$H(t) - H(1) \leq -\iint_t^1 \rho_s^\varepsilon |\nabla \log \rho_s^\varepsilon|^2 \mathrm{d}s \mathrm{d}\mathfrak{m} + N \iint_t^1 s^{-1} \rho_s^\varepsilon \mathrm{d}s \mathrm{d}\mathfrak{m} \leq -N \log t$$

and thus the conclusion. In order to get rid of the boundedness assumption on $\mathrm{supp}(\mu_0)$ and $\mathrm{supp}(\mu_1)$, take $\mu_0, \mu_1 \in \mathscr{P}_2(\mathrm{X})$ with bounded densities; then fix $\bar{x} \in \mathrm{X}$, define

$$\mu_i^k := \Big(\int \rho_i^k \, \mathrm{d}\mathfrak{m} \Big)^{-1} \rho_i^k \mathfrak{m}, \quad \rho_i^k := \chi_k \rho_i \qquad \text{for } i = 0, 1$$

where χ_k is a Lipschitz cut-off function supported in $B_{k+1}(\bar{x})$ and identically equal to 1 on $B_k(\bar{x})$, and invoke Theorem 5.1.1 together with part *(iii)* of Theorem 4.2.3. The latter can actually be applied because if we disintegrate $\mathbf{R}^{\varepsilon/2}$ w.r.t. $(\mathrm{e}_0, \mathrm{e}_1)$ as

$$\mathbf{R}^{\varepsilon/2}(\cdot) = \int_{\mathrm{X}^2} (\mathbf{R}^{\varepsilon/2})^{xy}(\cdot) \mathrm{d}\mathsf{R}^{\varepsilon/2}(x, y),$$

then by [27] (see Theorem 1 therein) $(x,y)\mapsto(\mathbf{R}^{\varepsilon/2})^{xy}$ is continuous.

This grants that for all $k\in\mathbb{N}$ the dynamical Schrödinger problem associated to $\mu_0^k,\mu_1^k,\mathbf{R}^{\varepsilon/2}$ has a unique solution $\mathbf{P}^{k,\varepsilon/2}$ and $(\mathbf{P}^{k,\varepsilon/2})$ weakly converges to the unique solution of the dynamical Schrödinger problem associated to $\mu_0,\mu_1,\mathbf{R}^{\varepsilon/2}$, say $\mathbf{P}^{\varepsilon/2}$. This implies that

$$(\mathrm{e}_t)_*\mathbf{P}^{k,\varepsilon/2}\rightharpoonup(\mathrm{e}_t)_*\mathbf{P}^{\varepsilon/2}\quad\text{as } k\to\infty,\qquad\forall t\in[0,1]$$

and by lower semicontinuity of the entropy together with $H(\mu_1^k\,|\,\mathfrak{m})\to H(\mu_1\,|\,\mathfrak{m})$ this is sufficient to conclude. □

From this result it is now easy to recover (6.3.3).

Corollary 6.3.2. *Let* $(\mathrm{X},\mathsf{d},\mathfrak{m})$ *be a* $\mathsf{RCD}^*(0,N)$ *space where* $N\in[1,\infty)$ *and* $\mathfrak{m}$ *is a non-negative Radon measure; let* $\mu_0,\mu_1\in\mathscr{P}_2(\mathrm{X})$ *have finite entropy w.r.t.* $\mathfrak{m}$.

Then

$$H(\mu_1\,|\,\mathfrak{m})-H(\mu_t\,|\,\mathfrak{m})\geq N\log t$$

for all $t\in(0,1]$.

proof Denote by ρ_0,ρ_1 the densities of μ_0,μ_1 respectively and assume for the moment that they are bounded and with bounded support. Then, using the same notation of Setting 5.1.3, Proposition 6.3.1 applies to (μ_t^ε) and by Proposition 6.2.1 we know that $\mu_t^\varepsilon\rightharpoonup\mu_t$ as $\varepsilon\downarrow 0$, where (μ_t) is the unique Wasserstein geodesic between μ_0 and μ_1. By lower semicontinuity of the entropy, this yields (6.3.3).

To get rid of the compactness and boundedness assumptions on the densities, it is sufficient to argue by standard cut-off arguments: given $\mu_0,\mu_1\in\mathscr{P}_2(\mathrm{X})\cap D(H(\cdot\,|\,\mathfrak{m}))$, whose densities are denoted by ρ_0,ρ_1 respectively, define

$$\mu_i^k:=\Big(\int\rho_i^k\,\mathrm{d}\mathfrak{m}\Big)^{-1}\rho_i^k\mathfrak{m},\quad\rho_i^k:=\max\{k,\chi_k\rho_i\}\qquad\text{for } i=0,1$$

where χ_k is a Lipschitz cut-off function supported in $B_{k+1}(\bar{x})$ and identically equal to 1 on $B_k(\bar{x})$, $\bar{x}$ being arbitrarily fixed. If we denote by (μ_t^k) (resp. (μ_t)) the unique W_2-geodesic between μ_0^k and μ_1^k (resp. μ_0 and μ_1), then it is well known that $\mu_t^k\rightharpoonup\mu_t$ as $k\to\infty$ and since $H(\mu_1^k\,|\,\mathfrak{m})\to H(\mu_1\,|\,\mathfrak{m})$, by lower semicontinuity of the entropy this is enough to get (6.3.3) and thus conclude. □

After this warm-up example, let us show that the entropy is (K,N)-convex along entropic interpolations on $\mathsf{RCD}^*(K,N)$ spaces, provided $K\geq 0$, thus generalizing to the non-smooth framework the result of a forthcoming paper by I. Gentil, C. Léonard and L. Ripani. As in Proposition 6.3.1 and still relying on Theorem 4.2.3, in the following statement we allow μ_0,μ_1 to have possibly unbounded supports.

Proposition 6.3.3. *Let* $(\mathrm{X},\mathsf{d},\mathfrak{m})$ *be a* $\mathsf{RCD}^*(K,N)$ *space with* $K\geq 0$ *and* $N\in[1,\infty)$ *and let* $\mu_0,\mu_1\in\mathscr{P}_2(\mathrm{X})\cap D(H(\cdot\,|\,\mathfrak{m}))$ *with bounded densities. Then the entropy is* (K,N)*-convex along* (μ_t^ε) *for all* $\varepsilon>0$*, where* μ_t^ε *is defined as in Proposition 6.3.1.*

proof Fix $\varepsilon>0$, assume that μ_0,μ_1 also have bounded supports and observe that from (5.3.6b) and (5.3.2b) we know that for all $t\in(0,1)$ it holds

$$\begin{aligned}\frac{\mathrm{d}^2}{\mathrm{d}t^2}H(\mu_t^\varepsilon\,|\,\mathfrak{m})\geq{}&K\int|\nabla\vartheta_t^\varepsilon|^2\rho_t^\varepsilon\mathrm{d}\mathfrak{m}+\frac{1}{N}\int\big(\Delta\vartheta_t^\varepsilon\big)^2\rho_t^\varepsilon\mathrm{d}\mathfrak{m}\\&+\frac{\varepsilon^2}{4}\bigg(K\int|\nabla\log\rho_t^\varepsilon|^2\rho_t^\varepsilon\mathrm{d}\mathfrak{m}+\frac{1}{N}\int\big(\Delta\log\rho_t^\varepsilon\big)^2\rho_t^\varepsilon\mathrm{d}\mathfrak{m}\bigg).\end{aligned}\tag{6.3.4}$$

The term with $\Delta \log \rho_t^\varepsilon$ on the right-hand side as well as those with K can be neglected, because positive, while the remaining one can be estimated by Cauchy-Schwarz inequality as follows

$$\int \left(\Delta \vartheta_t^\varepsilon\right)^2 \rho_t^\varepsilon \mathrm{d}\mathfrak{m} \geq \left(\int \rho_t^\varepsilon \Delta \vartheta_t^\varepsilon \mathrm{d}\mathfrak{m} \right)^2$$

and now a cut-off argument completely analogous to the one adopted in the proof of Proposition 5.2.2 motivates the following integration by parts

$$\int \rho_t^\varepsilon \Delta \vartheta_t^\varepsilon \mathrm{d}\mathfrak{m} = - \int \langle \nabla \rho_t^\varepsilon, \nabla \vartheta_t^\varepsilon \rangle \, \mathrm{d}\mathfrak{m}.$$

Comparing this expression with (5.3.6a), we see that (6.3.1) along the entropic interpolation from μ_0 to μ_1 is established pointwise for all $t \in (0,1)$ and since $t \mapsto H(\mu_t^\varepsilon \,|\, \mathfrak{m})$ is continuous on $[0,1]$ by Proposition 5.3.3 it follows that (6.3.1) holds in the distribution sense on the whole $[0,1]$. This is in turn equivalent to the integrated version (6.3.2) and the advantage of the latter formulation is the stability w.r.t. W_2-convergence, which allows to get rid of the compactness assumption on the support of the densities. In fact, rewriting (6.3.2) as

$$H(\mu_t^\varepsilon \,|\, \mathfrak{m}) \leq -N \log \left(\sigma_{K,N}^{(1-t)}(W_2(\mu_0,\mu_1)) U_N(\mu_0) + \sigma_{K,N}^{(t)}(W_2(\mu_0,\mu_1)) U_N(\mu_1) \right)$$

and arguing as in the last part of Proposition 6.3.1, the conclusion follows because the same approximation procedure also entails that $W_2(\mu_0^k, \mu_1^k) \to W_2(\mu_0,\mu_1)$ as $k \to \infty$, besides $H(\mu_0^k \,|\, \mathfrak{m}) \to H(\mu_0 \,|\, \mathfrak{m})$ and $H(\mu_1^k \,|\, \mathfrak{m}) \to H(\mu_1 \,|\, \mathfrak{m})$. □

As a byproduct, this immediately yields an 'entropic' and alternative proof of the fact that the $\mathsf{CD}^*(K,N)$ condition implies $\mathsf{CD}^e(K,N)$ on infinitesimally Hilbertian spaces, at least for $K \geq 0$.

Corollary 6.3.4. *Let* $(\mathrm{X}, \mathsf{d}, \mathfrak{m})$ *be a* $\mathsf{RCD}^*(K,N)$ *space with* $K \geq 0$ *and* $N \in [1,\infty)$*. Then* $\mathsf{CD}^e(K,N)$ *holds.*

proof Let $\mu_0 = \rho_0 \mathfrak{m}$ and $\mu_1 = \rho_1 \mathfrak{m}$ be two absolutely continuous Borel probability measures with bounded densities and supports and for any $\varepsilon > 0$ introduce the notations of Setting 5.1.3. Proposition 6.3.3 applies to (μ_t^ε) and by Proposition 6.2.1 we know that $\mu_t^\varepsilon \rightharpoonup \mu_t$ as $\varepsilon \downarrow 0$, where (μ_t) is the unique Wasserstein geodesic between μ_0 and μ_1. Thus, rewriting (6.3.2) as in the proof of Proposition 6.3.3 and by lower semicontinuity of the entropy, this yields (6.3.2) along (μ_t).

To get rid of the compactness and boundedness assumptions on the densities, adopt the same approximation procedure described in the proof of Corollary 6.3.2 and notice that it also ensures $W_2(\mu_0^k, \mu_1^k) \to W_2(\mu_0,\mu_1)$ as $k \to \infty$. □

A further interesting consequence of (5.3.6a) and (5.3.6b) is the HWI inequality and it is worth recalling how it is deduced. In a $\mathsf{RCD}(K,\infty)$ space $(\mathrm{X}, \mathsf{d}, \mathfrak{m})$ it is well known that the Boltzmann entropy $H(\cdot \,|\, \mathfrak{m})$ is geodesically K-convex, which means that for all $\mu_0, \mu_1 \in \mathscr{P}_2(\mathrm{X})$ with finite entropy w.r.t. $\mathfrak{m}$ there exists a W_2-geodesic (μ_t) connecting them such that

$$H(\mu_t \,|\, \mathfrak{m}) \leq (1-t) H(\mu_0 \,|\, \mathfrak{m}) + t H(\mu_1 \,|\, \mathfrak{m}) - \frac{K}{2} t(1-t) W_2^2(\mu_0,\mu_1)$$

for all $t \in [0,1]$. This inequality can be rewritten as

$$\frac{H(\mu_t \,|\, \mathfrak{m}) - H(\mu_0 \,|\, \mathfrak{m})}{t} \leq H(\mu_1 \,|\, \mathfrak{m}) - H(\mu_0 \,|\, \mathfrak{m}) - \frac{K}{2}(1-t) W_2^2(\mu_0,\mu_1)$$

and denoting by ρ_0 the Radon-Nikodym derivative of μ_0 w.r.t. $\mathfrak{m}$ (without loss of generality we can assume that $\mu_0 \ll \mathfrak{m}$), choosing $\boldsymbol{\gamma} \in \mathrm{Opt}(\mu_0, \mu_1)$ and assuming that we can pass to the limit as $t \downarrow 0$, this implies

$$-\int \frac{|\nabla \rho_0|(x)}{\rho_0(x)} \mathsf{d}(x,y) \mathrm{d}\boldsymbol{\gamma}(x,y) \leq H(\mu_1 \,|\, \mathfrak{m}) - H(\mu_0 \,|\, \mathfrak{m}) - \frac{K}{2} W_2^2(\mu_0, \mu_1)$$

whence, by the Cauchy-Schwarz inequality,

$$H(\mu_0 \,|\, \mathfrak{m}) - H(\mu_1 \,|\, \mathfrak{m}) \leq W_2(\mu_0, \mu_1) \sqrt{\int \frac{|\nabla \rho_0|^2}{\rho_0} \mathrm{d}\mathfrak{m}} - \frac{K}{2} W_2^2(\mu_0, \mu_1).$$

This is the celebrated **HWI inequality**, so called because expressed in terms of the entropy functional $H(\cdot \,|\, \mathfrak{m})$, the Wasserstein distance W_2 and the Fisher information $I(\cdot \,|\, \mathfrak{m})$, defined by

$$I(\mu \,|\, \mathfrak{m}) := \begin{cases} 4 \int |\nabla \sqrt{\rho}|^2 \, \mathrm{d}\mathfrak{m} = \int_{\{\rho > 0\}} \frac{|\nabla \rho|^2}{\rho} \, \mathrm{d}\mathfrak{m} & \text{if } \mu = \rho \mathfrak{m},\ \sqrt{\rho} \in W^{1,2}(\mathrm{X}) \\ +\infty & \text{otherwise} \end{cases}$$

(see Lemma 4.10 in [6] for the equality of the two expressions in the case $\sqrt{\rho} \in W^{1,2}(\mathrm{X})$). It is a very powerful tool that allows to deduce with very little effort the logarithmic Sobolev inequality (if $K > 0$) and Talagrand inequality, provided $\mathfrak{m}$ is a probability measure.

However, the passage to the limit as $t \downarrow 0$ is a delicate step (see for instance Theorem 20.10 and Corollary 20.13 in [121] or Proposition 7.18 in [3] together with Section 7 of [6] or also Proposition 5.10 in [58]). Here we propose an 'entropic' approach that strongly relies on (5.3.6a) and (5.3.6b) in order to get an analogue of the HWI inequality for entropic interpolations, before passing to the limit as $\varepsilon \downarrow 0$. As a first tool we need the following lemma, whose main features are the discovery of a quantity which is preserved along the entropic evolution and the convergence of (the integrals of) current and osmotic velocities (see Section 5.5) to the squared Wasserstein distance and 0 respectively.

Lemma 6.3.5. *With the same assumptions and notations as in Setting 5.1.3 we have that, for any $\varepsilon > 0$, the map*

$$(0,1) \ni t \quad \mapsto \quad \int \Big(|\nabla \vartheta_t^\varepsilon|^2 - \frac{\varepsilon^2}{4} |\nabla \log \rho_t^\varepsilon|^2 \Big) \rho_t^\varepsilon \mathrm{d}\mathfrak{m} \tag{6.3.5}$$

is real-valued and constant. Moreover the following hold:

$$\lim_{\varepsilon \downarrow 0} \iint_0^1 |\nabla \vartheta_t^\varepsilon|^2 \rho_t^\varepsilon \mathrm{d}t \mathrm{d}\mathfrak{m} = 2 \lim_{\varepsilon \downarrow 0} \iint_0^1 t |\nabla \vartheta_t^\varepsilon|^2 \rho_t^\varepsilon \mathrm{d}t \mathrm{d}\mathfrak{m} = W_2^2(\mu_0, \mu_1), \tag{6.3.6a}$$

$$\lim_{\varepsilon \downarrow 0} \varepsilon^2 \iint_0^1 |\nabla \log \rho_t^\varepsilon|^2 \rho_t^\varepsilon \mathrm{d}t \mathrm{d}\mathfrak{m} = 0. \tag{6.3.6b}$$

proof The map is real-valued because by Lemma 5.2.1 $t \mapsto \rho_t^\varepsilon |\nabla \vartheta_t^\varepsilon|^2$ and $t \mapsto \rho_t^\varepsilon |\nabla \log \rho_t^\varepsilon|^2$ belong to $L^\infty_{loc}((0,1), L^1(\mathrm{X}))$. To prove that it is constant, by the regularizing properties of the heat flow we have that $t \mapsto \int \langle \nabla f_t^\varepsilon, \nabla g_t^\varepsilon \rangle \, \mathrm{d}\mathfrak{m}$ belongs to $AC_{loc}((0,1))$ with

$$\frac{\mathrm{d}}{\mathrm{d}t} \int \langle \nabla f_t^\varepsilon, \nabla g_t^\varepsilon \rangle \, \mathrm{d}\mathfrak{m} = \frac{\varepsilon}{2} \int \Big(\langle \nabla \Delta f_t^\varepsilon, \nabla g_t^\varepsilon \rangle - \langle \nabla f_t^\varepsilon, \nabla \Delta g_t^\varepsilon \rangle \Big) \mathrm{d}\mathfrak{m}, \qquad \text{a.e. } t$$

and integration by parts shows that the derivative vanishes a.e. It is now sufficient to recall the definition of ϑ_t^ε and the identity $\varphi_t^\varepsilon + \psi_t^\varepsilon = \varepsilon \log \rho_t^\varepsilon$ to infer that

$$\varepsilon^2 \int \langle \nabla f_t^\varepsilon, \nabla g_t^\varepsilon \rangle \, \mathrm{d}\mathfrak{m} = \int \langle \nabla \varphi_t^\varepsilon, \nabla \psi_t^\varepsilon \rangle \rho_t^\varepsilon \mathrm{d}\mathfrak{m} = - \int \Big(|\nabla \vartheta_t^\varepsilon|^2 - \frac{\varepsilon^2}{4} |\nabla \log \rho_t^\varepsilon|^2 \Big) \rho_t^\varepsilon \mathrm{d}\mathfrak{m}.$$

As regards (6.3.6a) and (6.3.6b), start observing that $(\mu_t^\varepsilon, \vartheta_t^\varepsilon)$ solves the continuity equation in the sense of Theorem 1.1.4, as already remarked in Proposition 6.1.1, so that by the Benamou-Brenier formula (Theorem 1.1.5)

$$\iint_0^1 |\nabla \vartheta_t^\varepsilon|^2 \rho_t^\varepsilon \, \mathrm{d}t \, \mathrm{d}\mathfrak{m} \geq W_2^2(\mu_0, \mu_1).$$

On the other hand, from the first identity in (5.4.2) and (6.2.21) we know that

$$\lim_{\varepsilon \downarrow 0} \iint_0^1 \Big(|\nabla \vartheta_t^\varepsilon|^2 + \frac{\varepsilon^2}{4} |\nabla \log \rho_t^\varepsilon|^2 \Big) \rho_t^\varepsilon \mathrm{d}t \mathrm{d}\mathfrak{m} = W_2^2(\mu_0, \mu_1).$$

Hence

$$\lim_{\varepsilon \downarrow 0} \iint_0^1 |\nabla \vartheta_t^\varepsilon|^2 \rho_t^\varepsilon \mathrm{d}t \mathrm{d}\mathfrak{m} = W_2^2(\mu_0, \mu_1) \qquad \text{and} \qquad \lim_{\varepsilon \downarrow 0} \varepsilon^2 \iint_0^1 |\nabla \log \rho_t^\varepsilon|^2 \rho_t^\varepsilon \mathrm{d}t \mathrm{d}\mathfrak{m} = 0.$$

To conclude the proof, denote by Q^ε the value of the constant function defined in (6.3.5) and integrate it on $[0,1]$: on the one hand, by what we have just written above

$$\lim_{\varepsilon \downarrow 0} Q^\varepsilon = W_2^2(\mu_0, \mu_1); \tag{6.3.7}$$

on the other hand

$$0 \leq \lim_{\varepsilon \downarrow 0} \varepsilon^2 \iint_0^1 t |\nabla \log \rho_t^\varepsilon|^2 \rho_t^\varepsilon \mathrm{d}t \mathrm{d}\mathfrak{m} \leq \lim_{\varepsilon \downarrow 0} \varepsilon^2 \iint_0^1 |\nabla \log \rho_t^\varepsilon|^2 \rho_t^\varepsilon \mathrm{d}t \mathrm{d}\mathfrak{m} = 0.$$

Thus multiplying Q^ε by t and integrating in time we get

$$\begin{aligned} \lim_{\varepsilon \downarrow 0} \iint_0^1 t |\nabla \vartheta_t^\varepsilon|^2 \rho_t^\varepsilon \mathrm{d}t \mathrm{d}\mathfrak{m} &= \lim_{\varepsilon \downarrow 0} \iint_0^1 t \Big(|\nabla \vartheta_t^\varepsilon|^2 - \frac{\varepsilon^2}{4} |\nabla \log \rho_t^\varepsilon|^2 \Big) \rho_t^\varepsilon \mathrm{d}t \mathrm{d}\mathfrak{m} \\ &= \lim_{\varepsilon \downarrow 0} \int_0^1 t Q^\varepsilon \mathrm{d}t = \frac{1}{2} W_2^2(\mu_0, \mu_1). \end{aligned}$$

□

We are now ready to prove the HWI inequality. For sake of simplicity we shall assume the space $(\mathrm{X}, \mathsf{d}, \mathfrak{m})$ to be compact and $\rho_0, \rho_1 \in \mathrm{Test}^\infty_{>0}(\mathrm{X})$, but we do not believe these assumptions to be crucial. As a consequence, the conclusion of Proposition 5.3.3 is reinforced: $t \mapsto H(\mu_t^\varepsilon \,|\, \mathfrak{m})$ is $C^2([0,1])$, which means that (5.3.6a) and (5.3.6b) hold for all $t \in [0,1]$. This is a consequence of the enhanced regularity of Schrödinger potentials (5.4.4) in the present framework; see also Proposition 5.5 in [63] for more details.

Proposition 6.3.6. *Let $(\mathrm{X}, \mathsf{d}, \mathfrak{m})$ be a compact $\mathsf{RCD}^*(K,N)$ space with $K \in \mathbb{R}$, $N \in [1, \infty)$ and $\mathfrak{m} \in \mathscr{P}(\mathrm{X})$; let $\mu_0, \mu_1 \in \mathscr{P}(\mathrm{X})$ and assume that $H(\mu_0 \,|\, \mathfrak{m}) < \infty$.*

Then

$$H(\mu_1 \,|\, \mathfrak{m}) - H(\mu_0 \,|\, \mathfrak{m}) \leq W_2(\mu_0, \mu_1) \sqrt{I(\mu_1 \,|\, \mathfrak{m})} - \frac{K}{2} W_2^2(\mu_0, \mu_1). \tag{6.3.8}$$

proof Fix $\varepsilon > 0$ and assume for the moment that $\mu_0 = \rho_0\mathfrak{m}$, $\mu_1 = \rho_1\mathfrak{m}$ with $\rho_0, \rho_1 \in \mathrm{Test}^\infty_{>0}(\mathrm{X})$. By Proposition 5.3.3 in the compact case we know that $t \mapsto H(\mu_t^\varepsilon \,|\, \mathfrak{m})$ belongs to $C^2([0,1])$, so that if we apply the standard calculus identity (valid for any C^2-regular function h)

$$h(1) - h(0) = h'(1) - \int_0^1 t h''(t)\mathrm{d}t$$

to $h(t) = H(\mu_t^\varepsilon \,|\, \mathfrak{m})$ and use (5.3.6a), (5.3.6b), (5.3.2a) neglecting the Hessian term in the last one, we get

$$H(\mu_1^\varepsilon \,|\, \mathfrak{m}) - H(\mu_0^\varepsilon \,|\, \mathfrak{m}) \le \int \langle \nabla\rho_1^\varepsilon, \nabla\vartheta_1^\varepsilon\rangle \,\mathrm{d}\mathfrak{m} - K \iint_0^1 t\Big(|\nabla\vartheta_t^\varepsilon|^2 + \frac{\varepsilon^2}{4}|\nabla \log\rho_t^\varepsilon|^2\Big)\rho_t^\varepsilon \mathrm{d}t\mathrm{d}\mathfrak{m}. \tag{6.3.9}$$

By the Cauchy-Schwarz inequality and the definition of Fisher information, the first term on the right-hand side can be estimated as

$$\int \langle \nabla\rho_1^\varepsilon, \nabla\vartheta_1^\varepsilon\rangle \,\mathrm{d}\mathfrak{m} \le \sqrt{\int |\nabla\vartheta_1^\varepsilon|^2 \mathrm{d}\mu_1}\sqrt{I(\mu_1 \,|\, \mathfrak{m})}$$

and, under the additional regularity assumptions on X and ρ_0, ρ_1, Lemma 6.3.5 can be strengthened as follows:

$$t \quad \mapsto \quad \int \Big(|\nabla\vartheta_t^\varepsilon|^2 - \frac{\varepsilon^2}{4}|\nabla \log\rho_t^\varepsilon|^2\Big)\rho_t^\varepsilon \mathrm{d}\mathfrak{m}$$

is well defined on $[0,1]$ and constant therein; this is immediately seen by observing that $t \mapsto \int \langle \nabla f_t^\varepsilon, \nabla g_t^\varepsilon\rangle \,\mathrm{d}\mathfrak{m}$ now belongs to $AC([0,1])$. Therefore

$$\int |\nabla\vartheta_1^\varepsilon|^2 \mathrm{d}\mu_1 = Q^\varepsilon + \frac{\varepsilon^2}{4} I(\mu_1 \,|\, \mathfrak{m}),$$

Q^ε being defined as the constant value of (6.3.5), and by (6.3.7) together with the finiteness of $I(\mu_1 \,|\, \mathfrak{m})$ we deduce that

$$\lim_{\varepsilon\downarrow 0} \int |\nabla\vartheta_1^\varepsilon|^2 \mathrm{d}\mu_1 = W_2^2(\mu_0, \mu_1). \tag{6.3.10}$$

For the second integral on the right-hand side in (6.3.9), by Lemma 6.3.5 we are allowed to pass to the limit as $\varepsilon \downarrow 0$ and we get $-\frac{K}{2}W_2^2(\mu_0, \mu_1)$. This fact and (6.3.10) entail (6.3.8).

It remains to remove the regularity assumptions on ρ_0 and ρ_1. To this aim, let us first assume that ρ_0, ρ_1 are bounded away from 0 and define

$$\mu_i^k := \big(\mathsf{h}_{1/k}\rho_i\big)\mathfrak{m} \qquad \text{for } i = 0, 1,$$

noticing that (6.3.8) applies to μ_0^k, μ_1^k. It is easy to see that $W_2(\mu_0^k, \mu_1^k) \to W_2(\mu_0, \mu_1)$, $H(\mu_0^k \,|\, \mathfrak{m}) \to H(\mu_0 \,|\, \mathfrak{m})$ and $H(\mu_1^k \,|\, \mathfrak{m}) \to H(\mu_1 \,|\, \mathfrak{m})$ as $k \to \infty$; as regards the Fisher information, if $\mu_1 \notin D(I(\cdot \,|\, \mathfrak{m}))$ there is nothing to prove, so that without loss of generality we can assume $I(\mu_1 \,|\, \mathfrak{m}) < \infty$ and by the regularizing properties of the heat semigroup together with the fact that $\rho_1 \ge c$ for some $c > 0$ we get $I(\mu_1^k \,|\, \mathfrak{m}) \to I(\mu_1 \,|\, \mathfrak{m})$ as $k \to \infty$.

In order to get rid of the fact that ρ_0, ρ_1 are bounded away from 0 too, define

$$\mu_i^k := \Big(\int \rho_i^k \,\mathrm{d}\mathfrak{m}\Big)^{-1} \rho_i^k \mathfrak{m}, \quad \rho_i^k := \rho_i + \frac{1}{k} \qquad \text{for } i = 0, 1$$

and observe that the converge of the Wasserstein distance and of the entropies are clear; for the Fisher information, still assuming that $I(\mu_1 \,|\, \mathfrak{m})$ is finite, it is sufficient to observe that

$$\limsup_{k\to\infty} I(\mu_1^k \,|\, \mathfrak{m}) = \limsup_{k\to\infty} \int \frac{|\nabla \rho_1|^2}{\rho_1 + 1/k} \mathrm{d}\mathfrak{m} \le \int \frac{|\nabla \rho_1|^2}{\rho_1} \mathrm{d}\mathfrak{m} = I(\mu_1 \,|\, \mathfrak{m}) \le \liminf_{k\to\infty} I(\mu_1^k \,|\, \mathfrak{m})$$

where the second inequality is motivated by the lower semicontinuity of Fisher information w.r.t. L^1 convergence. □

Remark 6.3.7 (Lagrangian and Hamiltonian interpretation)**.** From a heuristic point of view, the expression of the constant quantity Q^ε (Lemma 6.3.5) can be deduced by standard arguments in Lagrangian and Hamiltonian formalism. Indeed, motivated by the Benamou-Brenier formula for the entropic cost provided by Theorem 5.4.3 let us consider the action functional

$$\mathcal{A}(\nu, v) = \iint_0^1 \Big(\frac{|v_t|^2}{2} + \frac{\varepsilon^2}{8} |\nabla \log \nu_t|^2\Big) \nu_t \mathrm{d}t \mathrm{d}\mathfrak{m} \tag{6.3.11}$$

associated to the Lagrangian

$$\mathcal{L}(\nu, v) = \int \Big(\frac{|v|^2}{2}\nu + \frac{\varepsilon^2}{8} \frac{|\nabla \nu|^2}{\nu}\Big) \mathrm{d}\mathfrak{m}.$$

By means of Legendre's transform, the corresponding Hamiltonian is given by

$$\mathcal{H}(\nu, p) = \int \Big(\frac{|p|^2}{2\nu} - \frac{\varepsilon^2}{8} \frac{|\nabla \nu|^2}{\nu}\Big) \mathrm{d}\mathfrak{m}$$

and, at least formally, $\mathcal{H}$ is constant along the critical points of $\mathcal{A}$. Since ρ_0 and ρ_1 are prescribed, the Euler-Lagrange equation for (6.3.11) reads as

$$\begin{aligned} &\frac{\mathrm{d}}{\mathrm{d}t} \nu_t + \mathrm{div}(\nu_t v_t) = 0 \\ &\frac{\mathrm{d}}{\mathrm{d}t} v_t + \frac{|v_t|^2}{2} = -\frac{\varepsilon^2}{8} \big(2\Delta \log \nu_t + |\nabla \log \nu_t|^2\big) \end{aligned}$$

and by Proposition 5.1.4 we know that these 'PDEs' are satisfied along $(\rho_t^\varepsilon, \vartheta_t^\varepsilon)$. Finally, as in the Hamiltonian p represents a momentum density, it is natural to set $p_t := \nu_t v_t$. From these considerations, the guess on the existence of a constant quantity and on its expression follows. ■

Remark 6.3.8. As regards Lemma 6.3.5 some further comments are worthwhile. First of all, let us stress that by means of the newly discovered constant quantity Q^ε we are able to improve (5.3.12b) to (6.3.6b). Secondly, by (6.2.21) and the first dynamic representation of the entropic cost provided in Proposition 5.4.1 we already knew that

$$\lim_{\varepsilon\downarrow 0} \iint_0^1 \Big(|\nabla \vartheta_t^\varepsilon|^2 + \frac{\varepsilon^2}{4} |\nabla \log \rho_t^\varepsilon|^2\Big) \rho_t^\varepsilon \mathrm{d}t \mathrm{d}\mathfrak{m} = W_2^2(\mu_0, \mu_1).$$

Lemma 6.3.5 is a refinement, because it separately determines the behaviour of current and osmotic velocities as $\varepsilon \downarrow 0$. From a different viewpoint, starting with the first dynamic representation of the entropic cost provided in Proposition 5.4.1 and (6.3.6a), (6.3.6b) we get

$$\lim_{\varepsilon \downarrow 0} \varepsilon \mathscr{I}_\varepsilon(\rho_0, \rho_1) = \frac{1}{2} W_2^2(\mu_0, \mu_1)$$

again, but with a different argument w.r.t. the one adopted in Remark 6.2.8. ■

Chapter 7

Second order differentiation formula and applications

In this chapter the uniform estimates obtained in Chapter 5 and Lemma 5.3.5 on the vanishing quantities are diffusely invoked once more together with the convergence results of Section 6.1 and 6.2, in order to establish the second order differentiation formula anticipated in the Introduction (see Theorem 0.0.5 and the forthcoming Theorem 7.1.2 for an equivalent statement). Some consequences and applications of it are then provided.

The second order differentiation formula is obtained in Section 7.1. This reads as a first order differentiation formula for $t \mapsto \int \langle \nabla f, \nabla \phi_t \rangle \, \mathrm{d}\mu_t$, so that it is completely natural to ask whether ∇f can be replaced by a general vector field: the answer is affirmative and this is done in Corollary 7.1.3 along the same lines of Theorem 7.1.2.

As regards the applications of Theorem 7.1.2, we present where and how our main result can be used in the proof of the splitting theorem to simplify the strategy, after having recalled the various achievements on the subject and the structure of the paper [55].

7.1 Proof of the main theorem

We start with the following simple continuity statement:

Lemma 7.1.1. *With the same assumptions and notation as in Setting 5.1.3, let $t \mapsto \mu_t = \rho_t \mathfrak{m}$ be the W_2-geodesic from μ_0 to μ_1 and $(\varphi_t)_{t\in(0,1]}$ and $(\psi_t)_{t\in[0,1)}$ any couple of limit functions given by Proposition 6.1.1.*

Then the maps

$$\begin{array}{llll} (0,1] \ni t & \mapsto & \rho_t \, \mathrm{d}\varphi_t & \in L^2(T^*\mathrm{X}) \\ [0,1) \ni t & \mapsto & \rho_t \, \mathrm{d}\psi_t & \in L^2(T^*\mathrm{X}) \\ (0,1] \ni t & \mapsto & \rho_t \, \mathrm{d}\varphi_t \otimes \mathrm{d}\varphi_t & \in L^2((T^*)^{\otimes 2}\mathrm{X}) \\ [0,1) \ni t & \mapsto & \rho_t \, \mathrm{d}\psi_t \otimes \mathrm{d}\psi_t & \in L^2((T^*)^{\otimes 2}\mathrm{X}) \end{array}$$

are all continuous w.r.t. the strong topologies.

proof By Lemma 1.2.8 we know that for any $p < \infty$ we have $\rho_s \to \rho_t$ in $L^p(\mathfrak{m})$ as $s \to t$ and thus in particular $\sqrt{\rho_s} \to \sqrt{\rho_t}$ as $s \to t$. Moreover, the compactness of the supports of ρ_0 and ρ_1 implies that there exist $\bar{x} \in \mathrm{X}$ and $R > 0$ such that $\mathrm{supp}(\rho_t) \subset B_R(\bar{x})$ for all $t \in [0,1]$. Recalling Lemma 2.2.1, consider a cut-off function $\chi \in \mathrm{Test}^\infty(\mathrm{X})$ with support in $B_{R+1}(\bar{x})$

such that $\chi \equiv 1$ in $B_R(\bar{x})$. The closure of the differential and the fact that $\chi\varphi_s \to \chi\varphi_t$ weakly in $W^{1,2}(\mathrm{X})$ as $s \to t > 0$ (as a consequence of $(\chi\varphi_t) \in C((0,1], C(\mathrm{X})) \cap L^\infty_{loc}((0,1), W^{1,2}(\mathrm{X}))$, see Proposition 6.2.1) grant that $\mathrm{d}(\chi\varphi_s) \to \mathrm{d}(\chi\varphi_t)$ weakly in $L^2(T^*\mathrm{X})$ and thus $\chi\mathrm{d}\varphi_s \to \chi\mathrm{d}\varphi_t$ too. Together with the previous claim about the densities, the fact that the latter are uniformly bounded in $L^\infty(\mathfrak{m})$ and how χ is constructed, this is sufficient to conclude that $t \mapsto \sqrt{\rho_t}\mathrm{d}\varphi_t \in L^2(T^*\mathrm{X})$ is weakly continuous.

We now claim that $t \mapsto \sqrt{\rho_t}\mathrm{d}\varphi_t \in L^2(T^*\mathrm{X})$ is strongly continuous and to this aim we show that their $L^2(T^*\mathrm{X})$-norms are constant. To see this, recall that by Proposition 6.2.1 we know that for $t \in (0,1]$ the function $-(1-t)\psi_t$ is a Kantorovich potential from μ_t to μ_1 while from (6.1.3) and the locality of the differential we get that $|\mathrm{d}\varphi_t| = |\mathrm{d}\psi_t|$ μ_t-a.e., thus by point (iii) in Theorem 1.2.6 we have that

$$\int |\mathrm{d}\varphi_t|^2 \rho_t \,\mathrm{d}\mathfrak{m} = \frac{1}{(1-t)^2}\int |\mathrm{d}(1-t)\psi_t|^2 \rho_t \,\mathrm{d}\mathfrak{m} = \frac{1}{(1-t)^2} W_2^2(\mu_t,\mu_1) = W_2^2(\mu_0,\mu_1).$$

Multiplying the $\sqrt{\rho_t}\mathrm{d}\varphi_t$'s by $\sqrt{\rho_t}$ and using again the $L^2(\mathfrak{m})$-strong continuity of $\sqrt{\rho_t}$ and the uniform $L^\infty(\mathfrak{m})$-bound we conclude that $t \mapsto \rho_t\mathrm{d}\varphi_t \in L^2(T^*\mathrm{X})$ is strongly continuous, as desired.

To prove the strong continuity of $t \mapsto \rho_t\,\mathrm{d}\varphi_t \otimes \mathrm{d}\varphi_t \in L^2((T^*)^{\otimes 2}\mathrm{X})$ we argue as in Corollary 6.2.6: the strong continuity of $t \mapsto \sqrt{\rho_t}\mathrm{d}\varphi_t \in L^2(T^*\mathrm{X})$ and the fact that these are, locally in $t \in (0,1]$, uniformly bounded (thanks again to $\mathrm{supp}(\rho_t) \subset B_R(\bar{x})$ for all $t \in [0,1]$), grant both that $t \mapsto \|\rho_t\mathrm{d}\varphi_t \otimes \mathrm{d}\varphi_t\|_{L^2((T^*)^{\otimes 2}\mathrm{X})}$ is continuous and that $t \mapsto \rho_t\mathrm{d}\varphi_t \otimes \mathrm{d}\varphi_t \in L^2((T^*)^{\otimes 2}\mathrm{X})$ is weakly continuous.

The claims about the ψ_t's follow in the same way. □

We now have all the tools needed to prove our main result. Notice that we shall not make explicit use of Theorem 0.0.6 but rather reprove it for (the restriction to $[\delta, 1-\delta]$ of) entropic interpolations.

Theorem 7.1.2. *Let $(\mathrm{X}, \mathsf{d}, \mathfrak{m})$ be a $\mathsf{RCD}^*(K,N)$ space with $K \in \mathbb{R}$ and $N \in [1,\infty)$. Let $\mu_0, \mu_1 \in \mathscr{P}_2(\mathrm{X})$ be such that $\mu_0, \mu_1 \le C\mathfrak{m}$ for some $C > 0$, with compact supports and let (μ_t) be the unique W_2-geodesic connecting μ_0 to μ_1. Also, let $h \in H^{2,2}(\mathrm{X})$.*

Then the map

$$[0,1] \ni t \quad \mapsto \quad \int h \,\mathrm{d}\mu_t \ \in \mathbb{R}$$

belongs to $C^2([0,1])$ and the following formulas hold for every $t \in [0,1]$:

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\int h\,\mathrm{d}\mu_t &= \int \langle \nabla h, \nabla\phi_t\rangle \,\mathrm{d}\mu_t, \\ \frac{\mathrm{d}^2}{\mathrm{d}t^2}\int h\,\mathrm{d}\mu_t &= \int \mathrm{Hess}(h)(\nabla\phi_t, \nabla\phi_t)\,\mathrm{d}\mu_t, \end{aligned} \tag{7.1.1}$$

where ϕ_t is any function such that for some $s \neq t$, $s \in [0,1]$, the function $-(s-t)\phi_t$ is a Kantorovich potential from μ_t to μ_s.

proof For the given μ_0, μ_1 introduce the notation of Setting 5.1.3 and then find $\varepsilon_n \downarrow 0$ such that $(\varphi_t^{\varepsilon_n}), (\psi_t^{\varepsilon_n})$ converge to limit curves $(\varphi_t), (\psi_t)$ as in Proposition 6.1.1.

By Lemma 1.2.7 we know that the particular choice of the ϕ_t's as in the statement does not affect the right hand sides in (7.1.1), we shall therefore prove that such formulas hold for

the choice $\phi_t := \psi_t$, which is admissible thanks to Proposition 6.2.1 whenever $t < 1$. The case $t = 1$ can be achieved swapping the roles of μ_0, μ_1 or, equivalently, with the choice $\phi_t = -\varphi_t$ which is admissible for $t > 0$.

Fix $h \in H^{2,2}(\mathrm{X})$ with compact support and for $t \in [0,1]$ set

$$I_n(t) := \int h \, \mathrm{d}\mu_t^{\varepsilon_n} \qquad\qquad I(t) := \int h \, \mathrm{d}\mu_t.$$

The bound (5.2.6) grants that the I_n's are uniformly bounded and the convergence in (6.1.2) that $I_n(t) \to I(t)$ for any $t \in [0,1]$.

Since $(\rho_t^{\varepsilon_n}) \in AC_{loc}((0,1), W^{1,2}(\mathrm{X}))$ we have that $I_n \in AC_{loc}((0,1))$ and, recalling the formula for $\frac{\mathrm{d}}{\mathrm{d}t}\rho_t^\varepsilon$ given by Proposition 5.1.4, that

$$\frac{\mathrm{d}}{\mathrm{d}t} I_n(t) = \int h \frac{\mathrm{d}}{\mathrm{d}t}\rho_t^{\varepsilon_n} \, \mathrm{d}\mathfrak{m} = -\int h \, \mathrm{div}(\rho_t^{\varepsilon_n} \nabla \vartheta_t^{\varepsilon_n}) \mathrm{d}\mathfrak{m} = \int \langle \nabla h, \nabla \vartheta_t^{\varepsilon_n} \rangle \, \rho_t^{\varepsilon_n} \, \mathrm{d}\mathfrak{m}. \tag{7.1.2}$$

The fact that $\vartheta_t = \frac{\psi_t - \varphi_t}{2}$, the compactness of $\mathrm{supp}(h)$ and the bounds (5.2.6) and (5.2.10a) ensure that $\left|\frac{\mathrm{d}}{\mathrm{d}t} I_n(t)\right|$ is uniformly bounded in n and $t \in [t_0, t_1] \subset (0,1)$ and the compactness of $\mathrm{supp}(h)$ also allows us to use the convergence properties (6.2.15) and (6.1.2), which grant that

$$\iint_{t_0}^{t_1} \langle \nabla h, \nabla \vartheta_t^{\varepsilon_n} \rangle \, \rho_t^{\varepsilon_n} \, \mathrm{d}t \, \mathrm{d}\mathfrak{m} \quad \to \quad \iint_{t_0}^{t_1} \langle \nabla h, \nabla \vartheta_t \rangle \, \rho_t \, \mathrm{d}t \, \mathrm{d}\mathfrak{m}.$$

This is sufficient to pass to the limit in the distributional formulation of $\frac{\mathrm{d}}{\mathrm{d}t} I_n(t)$ and taking into account that $I \in C([0,1])$ we have just proved that $I \in AC_{loc}((0,1))$ with

$$\frac{\mathrm{d}}{\mathrm{d}t} I(t) = \int \langle \nabla h, \nabla \vartheta_t \rangle \, \rho_t \, \mathrm{d}\mathfrak{m} \tag{7.1.3}$$

for a.e. $t \in [0,1]$. Recalling that $\vartheta_t = \frac{\psi_t - \varphi_t}{2}$, (6.1.3) and the locality of the differential we see that

$$\nabla \vartheta_t = \nabla \psi_t \quad \rho_t \mathfrak{m}\text{-a.e.} \qquad \forall t \in [0,1), \tag{7.1.4}$$

and thus by Lemma 7.1.1 we see that the right hand side of (7.1.3) has a continuous representative in $t \in [0,1)$, which then implies that $I \in C^1([0,1))$ and that the first in (7.1.1) holds for any $t \in [0,1)$.

For the second derivative we assume in addition that $h \in \mathrm{Test}^\infty(\mathrm{X})$. Then we recall that $(\rho_t^{\varepsilon_n}) \in AC_{loc}((0,1), W^{1,2}(\mathrm{X}))$ and $(\vartheta_t^{\varepsilon_n}) \in AC_{loc}((0,1), W^{1,2}(\mathrm{X}, e^{-V}\mathfrak{m}))$ with $V = M\mathsf{d}^2(\cdot, \bar{x})$ for some $\bar{x} \in \mathrm{X}$ and $M > 0$ sufficiently large. Consider the rightmost side of (7.1.2) to get that $\frac{\mathrm{d}}{\mathrm{d}t} I_n(t) \in AC_{loc}((0,1))$ and

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2} I_n(t) = \int \langle \nabla h, \nabla \frac{\mathrm{d}}{\mathrm{d}t} \vartheta_t^{\varepsilon_n} \rangle \rho_t^{\varepsilon_n} + \langle \nabla h, \nabla \vartheta_t^{\varepsilon_n} \rangle \, \frac{\mathrm{d}}{\mathrm{d}t} \rho_t^{\varepsilon_n} \, \mathrm{d}\mathfrak{m}$$

for a.e. t, so that defining the 'acceleration' a_t^ε as

$$a_t^\varepsilon := -\Big(\frac{\varepsilon^2}{4} \Delta \log \rho_t^\varepsilon + \frac{\varepsilon^2}{8} |\nabla \log \rho_t^\varepsilon|^2 \Big) \tag{7.1.5}$$

and recalling the formula for $\frac{\mathrm{d}}{\mathrm{d}t}\vartheta_t^\varepsilon$ given by Proposition 5.1.4 we have

$$\begin{aligned}\frac{\mathrm{d}^2}{\mathrm{d}t^2}I_n(t) &= \int \langle \nabla h, \nabla\Big(-\frac{1}{2}|\nabla\vartheta_t^{\varepsilon_n}|^2 + a_t^{\varepsilon_n}\Big)\rangle \rho_t^{\varepsilon_n} - \langle \nabla h, \nabla\vartheta_t^{\varepsilon_n}\rangle \operatorname{div}(\rho_t^{\varepsilon_n}\nabla\vartheta_t^{\varepsilon_n})\,\mathrm{d}\mathfrak{m}\\ &= \int \Big(-\frac{1}{2}\langle \nabla h, \nabla|\nabla\vartheta_t^{\varepsilon_n}|^2\rangle + \langle \nabla(\langle \nabla h, \nabla\vartheta_t^{\varepsilon_n}\rangle), \nabla\vartheta_t^{\varepsilon_n}\rangle + \langle \nabla h, \nabla a_t^{\varepsilon_n}\rangle\Big)\rho_t^{\varepsilon_n}\,\mathrm{d}\mathfrak{m}\\ \text{(by (1.2.21))}\quad &= \int \mathrm{Hess}(h)(\nabla\vartheta_t^{\varepsilon_n}, \nabla\vartheta_t^{\varepsilon_n})\rho_t^{\varepsilon_n}\,\mathrm{d}\mathfrak{m} - \int \big(\Delta h + \langle \nabla h, \nabla\log\rho_t^{\varepsilon_n}\rangle\big)a_t^{\varepsilon_n}\rho_t^{\varepsilon_n}\,\mathrm{d}\mathfrak{m}.\end{aligned}$$

Since $\vartheta_t^\varepsilon = \frac{\psi_t^\varepsilon - \varphi_t^\varepsilon}{2}$ and $\mathrm{Hess}(h) \in L^2(T^{*\otimes 2}\mathrm{X})$ with compact support, up to multiply $\nabla\vartheta_t^\varepsilon$ by a cut-off function χ identically equal to 1 on $\mathrm{supp}(h)$ we can apply the limiting property (6.2.15) and thus, taking also (6.1.2) into account, we see that

$$\int \mathrm{Hess}(h)(\nabla\vartheta_t^{\varepsilon_n}, \nabla\vartheta_t^{\varepsilon_n})\rho_t^{\varepsilon_n}\,\mathrm{d}\mathfrak{m} \quad \overset{n\to\infty}{\to} \quad \int \mathrm{Hess}(h)(\nabla\vartheta_t, \nabla\vartheta_t)\rho_t\,\mathrm{d}\mathfrak{m} \qquad \text{in } L^1_{loc}(0,1)$$

and since $|\nabla h|, \Delta h \in L^\infty(\mathrm{X})$, by Lemma 5.3.5 we deduce that

$$\int \big(\Delta h + \langle \nabla h, \nabla\log\rho_t^{\varepsilon_n}\rangle\big)a_t^{\varepsilon_n}\rho_t^{\varepsilon_n}\,\mathrm{d}\mathfrak{m} \quad \to \quad 0 \qquad \text{in } L^1_{loc}(0,1).$$

Hence we can pass to the limit in the distributional formulation of $\frac{\mathrm{d}^2}{\mathrm{d}t^2}I_n$ to obtain that $\frac{\mathrm{d}}{\mathrm{d}t}I \in AC_{loc}((0,1))$ and

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}I(t) = \int \mathrm{Hess}(h)(\nabla\vartheta_t, \nabla\vartheta_t)\rho_t\,\mathrm{d}\mathfrak{m} \tag{7.1.6}$$

for a.e. t. Using again (7.1.4) and Lemma 7.1.1 we conclude that the right hand side of (7.1.6) is continuous on $[0,1)$, so that $I \in C^2([0,1))$ and the second in (7.1.1) holds for every $t \in [0,1)$.

It remains to remove the assumptions that h has compact support and $h \in \mathrm{Test}^\infty(\mathrm{X})$. Starting with the former, pick $h \in \mathrm{Test}^\infty(\mathrm{X})$ and a cut-off function $\chi_R \in \mathrm{Test}^\infty(\mathrm{X})$ with support in $B_{R+1}(\bar{x})$ and such that $\chi_R \equiv 1$ in $B_R(\bar{x})$ for some $\bar{x} \in \mathrm{X}$. Set $h_R := \chi_R h$ and observe that the boundedness of $\||\nabla\chi_R|\|_{L^\infty(\mathfrak{m})}$ and $\|\Delta\chi_R\|_{L^\infty(\mathfrak{m})}$ w.r.t. R implies that $h_R \to h$ in $W^{1,2}(\mathrm{X})$ and $\Delta h_R \to \Delta h$ in $L^2(\mathfrak{m})$, so that the bound (1.2.18) grants that $h_R \to h$ in $W^{2,2}(\mathrm{X})$. Since $\mathrm{Test}^\infty(\mathrm{X})$ is an algebra, $h_R \in \mathrm{Test}^\infty(\mathrm{X})$ for every $R > 0$, thus the conclusion of the theorem holds for the h_R's.

Now notice that we can choose the ϕ_t's to be uniformly Lipschitz (e.g. by taking $\phi_t := \psi_t$ for $t \geq 1/2$ and $\phi_t := -\varphi_t$ for $t < 1/2$). The uniform L^∞ estimates (1.2.27), the fact that all the densities ρ_t are supported in a compact set B independent of $t \in [0,1]$, the equi-Lipschitz continuity of ϕ_t therein (Proposition 5.2.3) and the L^2-convergence of $h_R, \nabla h_R, \mathrm{Hess}(h_R)$ to $h, \nabla h, \mathrm{Hess}(h)$ respectively grant that as $R \to \infty$ we have that

$$\begin{aligned}\int h_R\,\mathrm{d}\mu_t \quad &\to \quad \int h\,\mathrm{d}\mu_t\\ \int \langle \nabla h_R, \nabla\phi_t\rangle\,\mathrm{d}\mu_t \quad &\to \quad \int \langle \nabla h, \nabla\phi_t\rangle\,\mathrm{d}\mu_t\\ \int \mathrm{Hess}(h_R)(\nabla\phi_t, \nabla\phi_t)\,\mathrm{d}\mu_t \quad &\to \quad \int \mathrm{Hess}(h)(\nabla\phi_t, \nabla\phi_t)\,\mathrm{d}\mu_t\end{aligned}$$

uniformly in $t \in [0,1]$. This is sufficient to overcome the first assumption.

In order to remove also the second one, i.e. the fact that $h \in \mathrm{Test}^\infty(\mathrm{X})$, pick $h \in H^{2,2}(\mathrm{X})$ and put $h_s := \mathfrak{h}_s h$, where $\mathfrak{h}_s$ is the mollified heat flow introduced in (2.2.2). As already remarked in the proof of Lemma 2.2.1, as $s \downarrow 0$ we have $h_s \to h$ in $W^{1,2}(\mathrm{X})$ and $\Delta h_s \to \Delta h$ in $L^2(\mathfrak{m})$. Thus the bound (1.2.18) grants that $h_s \to h$ in $W^{2,2}(\mathrm{X})$. Furthermore, still by what pointed out in Lemma 2.2.1 we know that $h_s \in \mathrm{Test}^\infty(\mathrm{X})$ for every $s > 0$, so that the conclusion of the theorem hold for the h_s's. Now it is sufficient to argue as just done above and the conclusion follows. □

A natural generalization of (7.1.1) is then given by the following result.

Corollary 7.1.3. *Let $(\mathrm{X}, \mathsf{d}, \mathfrak{m})$ be a $\mathsf{RCD}^*(K, N)$ space with $K \in \mathbb{R}$ and $N \in [1, \infty)$. Let $\mu_0, \mu_1 \in \mathscr{P}_2(\mathrm{X})$ be such that $\mu_0, \mu_1 \le C\mathfrak{m}$ for some $C > 0$, with compact supports and let (μ_t) be the unique W_2-geodesic connecting μ_0 to μ_1. Let (ϕ_t) be any curve of functions such that for some $s \neq t$, $s \in [0, 1]$, the function $-(s - t)\phi_t$ is a Kantorovich potential from μ_t to μ_s. Also, let $W \in H^{1,2}_C(T\mathrm{X})$.*

Then the map

$$[0, 1] \ni \ t \quad \mapsto \quad \int \langle W, \nabla\phi_t\rangle \mathrm{d}\mu_t \ \in \mathbb{R}$$

belongs to $C^1([0, 1])$ and the following formula holds for every $t \in [0, 1]$:

$$\frac{\mathrm{d}}{\mathrm{d}t} \int \langle W, \nabla\phi_t\rangle \mathrm{d}\mu_t = \int \langle \nabla_{\nabla\phi_t} W, \nabla\phi_t\rangle \mathrm{d}\mu_t. \tag{7.1.7}$$

proof As in the proof of Theorem 7.1.2, for the given μ_0, μ_1 introduce the notation of Setting 5.1.3 and then find $\varepsilon_n \downarrow 0$ such that $(\varphi^{\varepsilon_n}_t), (\psi^{\varepsilon_n}_t)$ converge to limit curves $(\varphi_t), (\psi_t)$ as in Proposition 6.1.1. Furthermore, we know that it is enough to show that (7.1.7) holds for the choice $\phi_t := \psi_t$, which is admissible thanks to Proposition 6.2.1 whenever $t < 1$. The case $t = 1$ can be achieved swapping the roles of μ_0, μ_1 or, equivalently, with the choice $\phi_t = -\varphi_t$ which is admissible for $t > 0$.

Suppose for the moment that $W = g\nabla h$ for some $g, h \in \mathrm{Test}^\infty(\mathrm{X})$ with h compactly supported and, in analogy with the proof of Theorem 7.1.2, define

$$J_n(t) := \int g\langle \nabla h, \nabla\vartheta^{\varepsilon_n}_t\rangle \mathrm{d}\mu^{\varepsilon_n}_t \qquad J(t) := \int g\langle \nabla h, \nabla\phi_t\rangle \mathrm{d}\mu_t.$$

For the same reasons explained above for $\frac{\mathrm{d}}{\mathrm{d}t} I_n$, $J_n \in AC_{loc}((0, 1))$ with

$$\frac{\mathrm{d}}{\mathrm{d}t} J_n(t) = \int g\langle \nabla h, \nabla \frac{\mathrm{d}}{\mathrm{d}t}\vartheta^{\varepsilon_n}_t\rangle \rho^{\varepsilon_n}_t + g\,\langle \nabla h, \nabla\vartheta^{\varepsilon_n}_t\rangle\, \frac{\mathrm{d}}{\mathrm{d}t}\rho^{\varepsilon_n}_t \,\mathrm{d}\mathfrak{m}$$

for a.e. t, so that defining the 'acceleration' a^ε_t as in (7.1.5) and recalling the formula for $\frac{\mathrm{d}}{\mathrm{d}t}\vartheta^\varepsilon_t$ given by Proposition 5.1.4 we have

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t} J_n(t) &= \int g\langle \nabla h, \nabla\Big(-\frac{1}{2}|\nabla\vartheta^{\varepsilon_n}_t|^2 + a^{\varepsilon_n}_t\Big)\rangle \rho^{\varepsilon_n}_t - g\,\langle \nabla h, \nabla\vartheta^{\varepsilon_n}_t\rangle \,\mathrm{div}(\rho^{\varepsilon_n}_t \nabla\vartheta^{\varepsilon_n}_t)\,\mathrm{d}\mathfrak{m} \\
&= \int \Big(-\frac{1}{2} g\langle \nabla h, \nabla|\nabla\vartheta^{\varepsilon_n}_t|^2\rangle + \langle \nabla(g\,\langle \nabla h, \nabla\vartheta^{\varepsilon_n}_t\rangle), \nabla\vartheta^{\varepsilon_n}_t\rangle\Big)\rho^{\varepsilon_n}_t \,\mathrm{d}\mathfrak{m} \\
&\quad + \int g\,\langle \nabla h, \nabla a^{\varepsilon_n}_t\rangle\, \rho^{\varepsilon_n}_t \mathrm{d}\mathfrak{m} \\
\text{(by (1.2.21))} \quad &= \int \Big(\mathrm{Hess}(h)(\nabla\vartheta^{\varepsilon_n}_t, \nabla\vartheta^{\varepsilon_n}_t) + \langle \nabla g, \nabla\vartheta^{\varepsilon_n}_t\rangle\,\langle \nabla h, \nabla\vartheta^{\varepsilon_n}_t\rangle\Big)\rho^{\varepsilon_n}_t \,\mathrm{d}\mathfrak{m} \\
&\quad - \int \big(g\Delta h + g\,\langle \nabla h, \nabla\log\rho^{\varepsilon_n}_t\rangle + \langle \nabla g, \nabla h\rangle\big) a^{\varepsilon_n}_t \rho^{\varepsilon_n}_t \,\mathrm{d}\mathfrak{m}.
\end{aligned}$$

Arguing as in Theorem 7.1.2 (since $|\nabla h|, \Delta h, g, |\nabla g| \in L^\infty(\mathfrak{m})$) and taking into account that by (1.2.22) the covariant derivative of $W = g\nabla h$ is given by $\nabla W = \nabla g \otimes \nabla h + g(\mathrm{Hess}(h))^\sharp$, we can pass to the limit in the distributional formulation of $\frac{\mathrm{d}}{\mathrm{d}t} J_n$ to obtain that $J \in AC_{loc}((0,1))$ and

$$\frac{\mathrm{d}}{\mathrm{d}t} J(t) = \int \langle \nabla_{\nabla \vartheta_t} W, \nabla \vartheta_t \rangle \mathrm{d}\mu_t$$

for a.e. t. Using (7.1.4) and Lemma 7.1.1 we conclude that the right hand side above is continuous on $[0,1)$, so that $J \in C^1([0,1))$ and (7.1.7) holds for every $t \in [0,1)$.

The compactness assumption on $\mathrm{supp}(h)$ can be removed as in Theorem 7.1.2 as well as $\Delta g, \Delta h \in L^\infty(\mathfrak{m})$, thus only assuming $g, h \in \mathrm{Test(X)}$. By linearity, (7.1.7) now holds for any $W \in \mathrm{TestV(X)}$.

For the general case, pick $W \in H^{1,2}_C(T\mathrm{X})$ and notice that by definition of $H^{1,2}_C(T\mathrm{X})$ there exists a sequence $(W_n)_n \subset \mathrm{TestV(X)}$ converging to W in $W^{1,2}_C(T\mathrm{X})$ such that (7.1.7) holds for the W_n's. Therefore, arguing as in the end of Theorem 7.1.2 we have that

$$\begin{aligned} \int \langle W_n, \nabla \phi_t \rangle \, \mathrm{d}\mu_t \qquad &\to \qquad \int \langle W, \nabla \phi_t \rangle \, \mathrm{d}\mu_t \\ \int \langle \nabla_{\nabla \phi_t} W_n, \nabla \phi_t \rangle \, \mathrm{d}\mu_t \qquad &\to \qquad \int \langle \nabla_{\nabla \phi_t} W, \nabla \phi_t \rangle \, \mathrm{d}\mu_t \end{aligned}$$

uniformly in $t \in [0,1]$ and this is sufficient to conclude. □

As a concluding remark, in the proof of Theorem 7.1.2 the final part shows that $H^{2,2}(\mathrm{X})$ can also be defined as the $W^{2,2}(\mathrm{X})$-closure of the functions in $\mathrm{Test}^\infty(\mathrm{X})$ with compact support. Similarly, by Corollary 7.1.3 $H^{1,2}_C(\mathrm{X})$ can be equivalently introduced as the $W^{1,2}_C(\mathrm{X})$ closure of all finite sums of objects of the following kind: $g_i \nabla f_i$ with $f_i, g_i \in \mathrm{Test}^\infty(\mathrm{X})$ and f_i compactly supported.

7.2 The splitting theorem

For the understanding of this section, only two new notions need to be introduced: line and Busemann function. Although they are meaningful in a purely metric setting, we shall present them in the case of a $\mathsf{RCD}^*(0,N)$ space $(\mathrm{X}, \mathsf{d}, \mathfrak{m})$.

A curve $\gamma : \mathbb{R} \to \mathrm{X}$ is a **line** provided

$$\mathsf{d}(\gamma_t, \gamma_s) = |t - s|, \qquad \forall t, s \in \mathbb{R}$$

and to a line we can associate the **Busemann function** $\mathrm{b} : \mathrm{X} \to \mathbb{R}$ as follows

$$\mathrm{b}(x) := \lim_{t \to +\infty} t - \mathsf{d}(x, \gamma_t).$$

At a first sight only the forward time direction is involved in this definition and b seems to be associated only to the positive half-line $(\gamma_t)_{t \geq 0}$, but thanks to the $\mathsf{RCD}^*(0,N)$ assumption $\mathrm{b} \equiv -\mathrm{b}^-$ on $\mathrm{supp}(\mathfrak{m})$ (see Theorem 4.11 of [55]), where

$$\mathrm{b}^-(x) := \lim_{t \to +\infty} t - \mathsf{d}(x, \gamma_{-t}),$$

so that the geometric information of b^- is encoded in b.

7.2.1 Overview of the paper

In [30] Cheeger and Gromoll proved the celebrated splitting theorem in its original Riemannian version, which was later generalized by Cheeger and Colding (see [29]) to the class of spaces which are measured Gromov-Hausdorff limits of smooth Riemannian manifolds. After the introduction of Lott-Sturm-Villani's synthetic curvature-dimension condition $\mathsf{CD}(K,N)$ ([86], [115], [116]) and Ambrosio-Gigli-Savaré's $\mathsf{RCD}(K,\infty)$ condition ([7]), the question became whether the splitting theorem would hold in the new framework of $\mathsf{CD}(0,N)$ and $\mathsf{RCD}(0,N)$ spaces too. In the former a counter-example was built by Cordero-Erausquin, Sturm and Villani (see the last theorem in [121]) while the answer is affirmative for the RCD case, as stated in the theorem below, and was given by Gigli in [55].

Theorem 7.2.1. *Let* $(\mathrm{X},\mathsf{d},\mathfrak{m})$ *be a* $\mathsf{RCD}^*(0,N)$ *space with* $N<\infty$ *and assume that* $\mathrm{supp}(\mathfrak{m})$ *contains a line. Then* $(\mathrm{X},\mathsf{d},\mathfrak{m})$ *is isomorphic (that is, there exists a measure preserving isometry) to the product of the Euclidean line* $(\mathbb{R},\mathsf{d}_{\mathrm{Eucl}},\mathscr{L}^1)$ *and another space* $(\mathrm{X}',\mathsf{d}',\mathfrak{m}')$*, where the product distance* $\mathsf{d}'\times\mathsf{d}_{\mathrm{Eucl}}$ *is defined as*

$$\mathsf{d}'\times\mathsf{d}_{\mathrm{Eucl}}((x',t),(y',s)):=\sqrt{\mathsf{d}'(x',y')^2+|t-s|^2},\qquad \forall x',y'\in\mathrm{X}',\ t',s'\in\mathbb{R}. \tag{7.2.1}$$

Moreover:

- *if* $N\geq 2$*, then* $(\mathrm{X}',\mathsf{d}',\mathfrak{m}')$ *is a* $\mathsf{RCD}^*(0,N-1)$ *space,*
- *if* $N\in[1,2)$*, then* X' *is just a point.*

We address the reader to [55] for a detailed overview on the history of the splitting theorem, its several variants and generalizations and the literature on the subject. Here we focus our attention on the proof of Theorem 7.2.1 and how Theorem 7.1.2 can simplify it. To this aim, let us quickly sketch the strategy adopted in [55] and stress the fact that at that time second order differential calculus on RCD spaces had not been developed yet, so that a notion of Hessian was not available nor the Bochner identity. Main steps in the non-smooth approach to the splitting theorem are the following, which actually coincide with the chapters of [55]:

(i) Multiples of b are Kantorovich potentials: thanks to this first step a bridge between the gradient flow of b and optimal transport is created, because Wasserstein geodesics can be built via the push-forward of any measure $\mu\in\mathscr{P}_2(\mathrm{X})$ through the gradient flow of b and this makes possible to use the $\mathsf{RCD}^*(0,N)$ condition.

(ii) The gradient flow of b is measure-preserving: if we denote by $t\mapsto\mathsf{F}_t$ a gradient flow of b, then by the previous step $t\mapsto(\mathsf{F}_t)_*\mu$ is a W_2-geodesic. Using the fact that the Rényi entropy (1.2.1) is geodescally convex on $(\mathscr{P}_2(\mathrm{X}),W_2)$ together with the first order differentiation formula contained in Theorem 0.0.4 and showing that $\mathbf{\Delta}\mathrm{b}=0$, it follows that $H_N(\mu\,|\,\mathfrak{m})\leq H_N((\mathsf{F}_t)_*\mu\,|\,\mathfrak{m})$. Noticing that $t\mapsto\mathsf{F}_{-t}$ is a gradient flow of $-\mathrm{b}$ the opposite inequality then follows and by the arbitrariness of $\mu\in\mathscr{P}_2(\mathrm{X})$ we get that $(\mathsf{F}_t)_*\mathfrak{m}=\mathfrak{m}$ for every $t\in\mathbb{R}$.

(iii) The gradient flow of b preserves the distance: arguing as in the previous step and keeping the same notation, we deduce that for all $f\in W^{1,2}(\mathrm{X})$ and $t\in\mathbb{R}$ it holds $f\circ\mathsf{F}_t\in W^{1,2}(\mathrm{X})$ with $|\nabla(f\circ\mathsf{F}_t)|\leq|\nabla f|\circ\mathsf{F}_t$ $\mathfrak{m}$-a.e. Then the Sobolev-to-Lipschitz property allows to translate this Sobolev information into a metric one.

(iv) The quotient space isometrically embeds into the original one: having a well defined gradient flow of b on the whole $\mathrm{supp}(\mathfrak{m})$ allows to consider the quotient X' of X w.r.t. the orbits $t \mapsto \mathsf{F}_t$; the quotient distance d' and measure $\mathfrak{m}'$ follow naturally. In this part the crucial result is the fact that the natural projection $\pi : \mathrm{supp}(\mathfrak{m}) \to \mathrm{X}'$ has an isometric embedding ι as right inverse.

(v) "Pythagoras' theorem" holds: by means of ι and arguing in duality with Sobolev functions it is possible to show that d splits according to (7.2.1).

(vi) The quotient space has dimension $N-1$: by all the previous steps we know that $(\mathrm{X}, \mathsf{d}, \mathfrak{m})$ splits as the product of $(\mathbb{R}, \mathsf{d}_{\mathrm{Eucl}}, \mathscr{L}^1)$ and a $\mathsf{RCD}^*(0, N)$ space $(\mathrm{X}', \mathsf{d}', \mathfrak{m}')$, so that it remains to prove that $(\mathrm{X}', \mathsf{d}', \mathfrak{m}')$ is actually a $\mathsf{RCD}^*(0, N-1)$ space if $N \geq 2$ or is just a point if $N \in [1, 2)$.

7.2.2 Application of the main theorem

The second order differentiation formula comes into play in the third chapter of [55], enabling a direct proof of the fact that the gradient flow of the Busemann function preserves the distance. In fact, using an argument which closely follows the first step in the proof of Proposition 6.2.1, we can describe the behaviour of the Busemann function along geodesics.

Theorem 7.2.2. *Let* $(\mathrm{X}, \mathsf{d}, \mathfrak{m})$ *be a* $\mathsf{RCD}^*(0, N)$ *space with* $N \in [1, \infty)$*. Then* b *is geodesically affine, i.e. for all* $x, y \in \mathrm{X}$ *there exists a geodesic* γ *connecting them such that* $t \mapsto \mathrm{b}(\gamma_t)$ *is affine.*

proof Let $x, y \in \mathrm{X}$, $r > 0$, define

$$\mu_x^r := \frac{1}{\mathfrak{m}(B_r(x))} \mathfrak{m}|_{B_r(x)} \qquad \mu_y^r := \frac{1}{\mathfrak{m}(B_r(y))} \mathfrak{m}|_{B_r(y)}$$

and let (μ_t^r) be the only W_2-geodesic from μ_x^r to μ_y^r. Since μ_x^r, μ_y^r have compact support, there exist $\bar{x} \in \mathrm{X}$ and $R > 0$ sufficiently large such that

$$\mathrm{supp}(\mu_t^r) \subset B_R(\bar{x}), \quad \forall t \in [0, 1].$$

Thus let $\chi \in \mathrm{Test}^\infty(\mathrm{X})$ be a cut-off function (Lemma 2.2.1) with bounded support such that $\chi \equiv 1$ in $B_R(\bar{x})$. Now recall that $\mathrm{b} \in W^{1,2}_{loc}(\mathrm{X})$ and $\boldsymbol{\Delta}\mathrm{b} = 0$, as proved in [55], so that $\chi\mathrm{b} \in D(\Delta)$; by (1.2.18) this implies that $\chi\mathrm{b} \in H^{2,2}(\mathrm{X})$ with $\mathrm{Hess}(\chi\mathrm{b}) = 0$. Therefore, applying Theorem 7.1.2 to $h = \chi\mathrm{b}$ and (μ_t^r) and noticing that $\chi\mathrm{b} = \mathrm{b}$ $\mathfrak{m}$-a.e. in $\mathrm{supp}(\mu_t^r)$ by construction, we get that

$$[0, 1] \ni t \quad \mapsto \quad \int \mathrm{b} \, \mathrm{d}\mu_t^r \in \mathbb{R}$$

is affine. In order to localize this information by passing to the limit as $r \downarrow 0$, let $\boldsymbol{\pi}^r$ be the only lifting of (μ_t^r) (recall point (i) of Theorem 1.2.6) and observe that

$$\sup_{r \in (0,1)} \iint_0^1 |\dot{\gamma}_t|^2 \mathrm{d}t \mathrm{d}\boldsymbol{\pi}^r(\gamma) < \infty$$

because $W_2(\mu_x^r, \delta_x) \to 0$ and $W_2(\mu_y^r, \delta_y)$ as $r \downarrow 0$: this implies that the lifting $(\boldsymbol{\pi}^r)$ is tight in $\mathscr{P}(C([0, 1], \mathrm{X}))$. Hence, up to pass to a suitable subsequence, not relabeled, we can assume that

$(\boldsymbol{\pi}^r)$ weakly converges to some $\boldsymbol{\pi} \in \mathscr{P}(C([0,1],\mathrm{X}))$ concentrated on $\mathrm{Geo}(\mathrm{X})$ with $(\mathrm{e}_0)_*\boldsymbol{\pi} = \delta_x$ and $(\mathrm{e}_1)_*\boldsymbol{\pi} = \delta_y$.

Since geodesics are essentially unique in finite-dimensional $\mathsf{RCD}^*(K,N)$ spaces, in the sense of point (iv) of Theorem 1.2.6, we deduce that for all $x \in \mathrm{X}$ the following holds: for $\mathfrak{m}$-a.e. $y \in \mathrm{X}$ there exists a (unique) geodesic γ from x to y such that $t \mapsto \mathrm{b}(\gamma_t)$ is affine. To get rid of the 'a.e.' quantification fix $x, y \in \mathrm{X}$ and let (y_n) be such that $y_n \to y$ as $n \to \infty$ and $t \mapsto \mathrm{b}(\gamma_t^n)$ is affine for a suitable geodesic joining x and y_n; observing that (γ^n) is relatively compact in $C([0,1],\mathrm{X})$, by the Ascoli-Arzelà theorem there exist a convergent subsequence and a limit curve γ connecting x and y, which is actually a geodesic. It is now easy to see that $t \mapsto \mathrm{b}(\gamma_t)$ is affine. □

The fact that the gradient flow of b is an isometry is now a consequence of the following result, proved by Sturm in [117] and here adapted to our purposes: if $(\mathrm{X},\mathsf{d},\mathfrak{m})$ is a $\mathsf{RCD}^*(0,N)$ space with $N \in [1,\infty)$ and $f : \mathrm{X} \to \mathbb{R}$ is a continuous and geodesically convex function with at most quadratic decrease, then for all $x \in \mathrm{X}$ there exists a unique EVI-gradient flow (x_t) of f starting at x. Since it is well known that EVI-gradient flows are contractive, i.e.

$$\mathsf{d}(x_t, y_t) \le \mathsf{d}(x_0, y_0) \qquad \forall x_0, y_0 \in \mathrm{X}, t > 0$$

where (x_t) and (y_t) are the EVI-gradient flows of f starting at x_0 and y_0 respectively, and b is both geodesically convex and concave by the theorem above, it is sufficient to apply Sturm's result to b and $-\mathrm{b}$ to get

$$\mathsf{d}(x_t, y_t) = \mathsf{d}(x_0, y_0) \qquad \forall x_0, y_0 \in \mathrm{X}, t > 0. \tag{7.2.2}$$

Indeed, by uniqueness of the gradient flows of b and $-\mathrm{b}$, for all $T > 0$ the gradient flow (x_t) of b starting at x_0 and the gradient flow (x'_t) of $-\mathrm{b}$ starting at x_T are linked via the identity $x_t = x'_{T-t}$ for all $t \in [0,T]$; plugging this information into the contraction estimate for (x'_t) we get

$$\mathsf{d}(x_0, y_0) \le \mathsf{d}(x_t, y_t) \qquad \forall x_0, y_0 \in \mathrm{X}, t > 0$$

whence (7.2.2).

Appendix A

Auxiliary results

In this part we recall some definitions and general facts about Γ-convergence, selection and disintegration theorems (both from an analytic and probabilistic point of view) and Markov processes.

A.1 Γ-convergence

A sequence of functions $f^k : \mathrm{Y} \to \overline{\mathbb{R}}$ on a metric space $(\mathrm{Y}, \mathsf{d}_{\mathrm{Y}})$ Γ-converges to f, and we shall write $\Gamma - \lim_{k\to\infty} f^k = f$, if:

(a) for all $y \in \mathrm{Y}$ and for all sequence $y_n \to y$, $\liminf_{k\to\infty} f^k(y_k) \geq f(y)$;

(b) for all $y \in \mathrm{Y}$ there exists a sequence $y'_k \to y$ such that $\limsup_{k\to\infty} f^k(y'_k) \leq f(y)$.

In the case the sequence (f^k) is equi-coercive, i.e. for any $c \in \mathbb{R}$ there exists a compact set $K_c \subset \mathrm{Y}$ such that $\{f^k \leq c\} \subset K_c$ for all $k \in \mathbb{N}$, and $\Gamma - \lim_{k\to\infty} f^k = f$, then:

(i) $\inf_{\mathrm{Y}} f$ is attained and
$$\lim_{k\to\infty} \inf_{y\in\mathrm{Y}} f^k(y) = \min_{y\in\mathrm{Y}} f(y);$$

(ii) if $\inf_{\mathrm{Y}} f < \infty$, then given a precompact sequence $(y_k) \subset \mathrm{Y}$ such that $\lim_{k\to\infty} f^k(y_k) = \min_{\mathrm{Y}} f$ every cluster point of (y_k) is a minimizer of f.

A proof of these facts as well as a detailed presentation of the topic can be found, for instance, in the monograph [39].

A.2 Selection and disintegration theorems

Given two arbitrary non-empty sets Y, Y', by axiom of choice a multifunction $F : \mathrm{Y} \to \mathrm{Y}'$ always admits a selector, i.e. a function $s : \mathrm{Y} \to \mathrm{Y}'$ such that $s(y) \in F(y)$ for every $y \in \mathrm{Y}$. But when does a measurable multifunction admit a measurable selector?

An answer is given by the Kuratowski-Ryll-Nardzewski selection theorem. In order to state it, let Y be a non-empty set and $\mathscr{A}$ an algebra of sets on Y, i.e. a non-empty family of subsets of Y which is closed under complementations and finite unions. We will denote by $\mathscr{A}_\sigma$ the smallest family of subsets of Y containing $\mathscr{A}$ and closed under countable unions. If we further consider a Polish space Y', the graph of a multifunction $F : \mathrm{Y} \to \mathrm{Y}'$ is defined by $\{(y, y') \in \mathrm{Y} \times \mathrm{Y}' \,|\, y' \in F(y)\}$ and it is closed if $F(y)$ is closed in Y' for every $y \in \mathrm{Y}$.

Theorem A.2.1 (Kuratowski, Ryll-Nardzewski)**.** *Let* Y *be a non-empty set,* $\mathscr{A}$ *an algebra of sets on* Y *and* $\mathscr{A}_\sigma$ *as above. Furthermore, let* Y' *be a Polish space. Then every* $\mathscr{A}_\sigma$*-measurable multifunction* $F : \mathrm{Y} \to \mathrm{Y}'$ *with closed graph admits a* $\mathscr{A}_\sigma$*-measurable selector.*

The second measure-theoretic topic of this section is the disintegration theorem. The reader is addressed to the rich literature on the subject for a detailed discussion (see for instance fremlin, dellacherie-meyer or the state of the art paper pachl), here we are only interested in providing the notions needed for the understanding of the theorem statement and in presenting conditional probabilities as disintegrations, since throughout the manuscript we occasionally adopt a probabilistic point of view.

Definition A.2.2. *Let* $(\mathrm{Y}, \mathscr{F}, \mu)$ *and* $(\mathrm{Y}', \mathscr{G}, \nu)$ *be* σ*-finite measure spaces and let* $T : \mathrm{Y} \to \mathrm{Y}'$ *be a measurable map. A* (T, ν)*-disintegration is a family* $\{\mu^y\}_{y \in Y'}$ *of measures on* $\mathscr{F}$ *such that:*

(i) for every $y \in Y'$*,* μ^y *is a* σ*-finite measure and* $\mu^y(\{x \in \mathrm{Y} \,|\, T(x) \neq y\}) = 0$ *for* ν*-a.e.* $y \in \mathrm{Y}'$*, that is* μ^y *is concentrated on the fiber* $T^{-1}(\{y\})$*;*

(ii) for every non-negative μ*-measurable function* f *on* Y*, the function* $y \mapsto \int_{\mathrm{Y}} f \, \mathrm{d}\mu^y$ *is* ν*-measurable;*

(iii) for every non-negative μ*-measurable function* f *on* Y*, it holds*

$$\int_{\mathrm{Y}} f \, \mathrm{d}\mu = \int_{\mathrm{Y}'} \left(\int_{\mathrm{Y}} f(x) \mathrm{d}\mu^y(x) \right) \mathrm{d}\nu(y).$$

When $\nu = T_* \mu$*, a* (T, ν)*-disintegration is also called disintegration of* μ *w.r.t.* T*.*

In the case μ, ν and every μ^y are probability measures, the disintegration $\{\mu^y\}_{y \in Y'}$ is also called (especially in probability) a regular conditional probability.

For the next statement, recall that a σ-algebra $\mathscr{G}$ is countably generated if there exists a countable family of sets such that $\mathscr{G}$ coincides with the smallest σ-algebra containing such a family.

Theorem A.2.3 (Disintegration)**.** *Let* μ *be a non-negative* σ*-finite regular Borel measure on a metric space* (Y, d) *and let* T *be a measurable map from* Y *into* $(\mathrm{Y}', \mathscr{G}, \nu)$*, where* ν *is non-negative,* σ*-finite and such that* $T_* \mu \ll \nu$*. Assume that* $\mathscr{G}$ *is countably generated and contains all singletons. Then:*

(i) μ *admits a* (T, ν)*-disintegration;*

(ii) if $\{\mu^y\}_{y \in \mathrm{Y}'}, \{\mu'^y\}_{y \in \mathrm{Y}'}$ *are two such disintegrations, then* $\mu^y = \mu'^y$ *for* ν*-a.e.* $y \in \mathrm{Y}'$*;*

(iii) if in addition $T_* \mu = \nu$*, then* μ^y *is a probability measure for* ν*-a.e.* $y \in \mathrm{Y}'$*.*

We conclude the section with some applications of the disintegration theorem to push-forward and relative entropy.

Let $(\mathrm{Y}, \tau), (\mathrm{Y}', \tau')$ be two Polish spaces endowed with their Borel σ-algebras, $\phi : \mathrm{Y} \to \mathrm{Y}'$ a Borel function, μ a probability measure on Y and ν a non-negative Radon measure on Y. Assume that $\mu \ll \nu$. Then $\phi_* \mu \ll \phi_* \nu$ with Radon-Nikodym derivative explicitly given by

$$\frac{\mathrm{d}\phi_* \mu}{\mathrm{d}\phi_* \nu}(y') = \int_{\mathrm{Y}} \frac{\mathrm{d}\mu}{\mathrm{d}\nu}(y) \mathrm{d}\nu^{y'}(y) \qquad \text{for } \phi_* \nu\text{-a.e. } y' \in \mathrm{Y}',$$

where $\{\nu^{y'}\}_{y'\in \mathrm{Y}'}$ denotes the disintegration of $\phi_*\nu$ w.r.t. ϕ, or in probabilistic language (see the forthcoming Appendix A.3 for a concise explanation of the parallelism between conditioning and disintegration)

$$\frac{\mathrm{d}\phi_*\mu}{\mathrm{d}\phi_*\nu} = \mathsf{E}_\nu\left[\frac{\mathrm{d}\mu}{\mathrm{d}\nu}\,\middle|\,\phi = \cdot\right] \qquad \phi_*\nu\text{-a.e.} \tag{A.2.1}$$

Moreover the following additivity property holds true

$$H(\mu\,|\,\nu) = H(\phi_*\mu\,|\,\phi_*\nu) + \int_{\mathrm{Y}'} H(\mu^{y'}\,|\,\nu^{y'})\mathrm{d}\phi_*\mu(y') \tag{A.2.2}$$

and in particular this shows that entropy is decreasing under push-forward. In fact, as $\nu^{y'}$ is a probability measure for $\phi_*\nu$-a.e. $y' \in \mathrm{Y}'$, we see that $H(\phi_*\mu\,|\,\phi_*\nu) \le H(\mu\,|\,\nu)$ for all $\mu \in \mathscr{P}(\mathrm{Y})$ with equality if and only if

$$\mu^{y'} = \nu^{y'} \qquad \text{for } \phi_*\mu - a.e.\, y' \in \mathrm{Y}'.$$

The proof of both (A.2.1) and (A.2.2) can be found in [80].

A.3 Conditioning as disintegration

In probability conditioning is usually introduced via the abstract Kolmogorov approach, i.e. given a probability space $(\mathrm{Y}, \mathscr{F}, \mathsf{P})$, an integrable random variable $Z : \mathrm{Y} \to \mathbb{R}$ and a σ-algebra $\mathscr{G} \subset \mathscr{F}$, the conditional expectation of Z given $\mathscr{G}$ is denoted by $\mathsf{E}_\mathsf{P}[Z\,|\,\mathscr{G}]$ and defined as the unique $\mathscr{G}$-measurable integrable random variable such that

$$\int_A Z\,\mathrm{d}\mathsf{P} = \int_A \mathsf{E}_\mathsf{P}[Z\,|\,\mathscr{G}]\,\mathrm{d}\mathsf{P}, \qquad \forall A \in \mathscr{G}$$

and, for any event $A \in \mathscr{F}$, the conditional probability of A given $\mathscr{G}$ is defined as

$$\mathsf{P}(A\,|\,\mathscr{G}) := \mathsf{E}_\mathsf{P}[\mathbb{1}_A\,|\,\mathscr{G}].$$

Awareness is required, because inspite of the notations and of the names $\mathsf{E}_\mathsf{P}[Z\,|\,\mathscr{G}]$ and $\mathsf{P}(A\,|\,\mathscr{G})$ are not scalars but random variables. If we further introduce a space with σ-algebra (Y', Σ) and a function $f : \mathrm{Y} \to \mathrm{Y}'$, then

$$\mathsf{E}_\mathsf{P}[Z\,|\,f] := \mathsf{E}_\mathsf{P}[Z\,|\,\sigma(f)] \qquad \text{and} \qquad \mathsf{P}(A\,|\,f) := \mathsf{P}(A\,|\,\sigma(f))$$

where $\sigma(f)$ denotes the σ-algebra generated by f, i.e. the smallest σ-algebra that makes f Σ-measurable. It is now tempting to define $\mathsf{E}_\mathsf{P}[Z\,|\,f = y'] := \mathsf{E}_\mathsf{P}[Z\,|\,f](y)$ and $\mathsf{P}(A\,|\,f = y') := \mathsf{P}(A\,|\,f)(y)$ with $y \in f^{-1}(\{y'\})$ arbitrarily chosen; the definition is actually meaningful, as $\mathsf{E}_\mathsf{P}[Z\,|\,f]$ and $\mathsf{P}(A\,|\,f)$ are P-a.e. constant on $f^{-1}(\{y'\})$, but $\mathsf{P}(\cdot\,|\,f = y')$ may fail to be a measure and thus in general there is no reason for

$$\mathsf{P}(A \cap f^{-1}(B)) = \int_B \mathsf{P}(A\,|\,f = y')\mathrm{d}f_*\mathsf{P}(y'), \qquad \forall A \in \mathscr{F},\, B \in \Sigma \tag{A.3.1}$$

to hold. When this identity is satisfied together with the following two conditions:

(i) for every $A \in \mathscr{F}$, $y' \mapsto \mathsf{P}(A \,|\, f = y')$ is Σ-measurable;

(ii) for $f_*\mathsf{P}$-a.e. $y' \in \mathrm{Y}'$, $\mathsf{P}(\cdot \,|\, f = y')$ is a probability measure;

then $\{\mathsf{P}(\cdot \,|\, f = y')\}_{y' \in \mathrm{Y}'}$ is called a regular conditional probability. Notice that this definition is perfectly in line with the notion of disintegration given in Appendix A.2, because point (*iii*) of Definition A.2.2 is equivalent to (A.3.1).

Thus, if we are willing to sacrifice a little of the generality of Kolmogorov's abstract definition, we can actually see conditioning as a disintegration procedure and in this way we surely gain in rigour, always preserving that intuition so common in conditioning arguments.

A.4 Markov property

Let $(\mathrm{Y}, \mathscr{F})$ be a space with σ-algebra. A family $(\mathscr{F}_t)_{t \geq 0} \subset \mathscr{F}$ of σ-algebras is called a filtration if it is increasing, that is

$$t \leq s \qquad \Rightarrow \qquad \mathscr{F}_t \subset \mathscr{F}_s.$$

A stochastic process $\mathbf{Z} = (Z_t)_{t \geq 0}$ on the measure space $(\mathrm{Y}, \mathscr{F}, \mathsf{P})$ is said to be $(\mathscr{F}_t)$-adapted if Z_t is $\mathscr{F}_t$-measurable for all $t \geq 0$.

With this said a stochastic process $\mathbf{Z}$ is Markov w.r.t. the filtration $(\mathscr{F}_t)_{t \geq 0}$ if it is adapted to the filtration and there exists a family of functions $\{p(s, x; t, I)\}$, called transition probabilities, such that:

(i) for every I, $(s, x, t) \mapsto p(s, x; t, I)$ is Borel;

(ii) for every (s, x, t), $I \mapsto p(s, x; t, I)$ is a probability measure;

(iii) $p(s, x; s, I) = \delta_x(I)$;

(iv) the Chapman-Kolmogorov relation holds, i.e.

$$p(s, x; t, I) = \int_{\mathbb{R}} p(s, x; r, dz) p(r, z; t, I), \qquad s < r < t;$$

(v) $\mathsf{P}R(Z_t \in I \,|\, \mathscr{F}_s) = p(s, Z_s; t, I)$.

A standard choice for the filtration $(\mathscr{F}_t)_{t \geq 0}$ is the following: $\mathscr{F}_t$ is the smallest σ-algebra that makes measurable all random variables Z_s for $0 \leq s \leq t$. With this choice $\mathsf{P}(\cdot \,|\, \mathscr{F}_s)$ is commonly denoted by $\mathsf{P}(\cdot \,|\, Z_{[0,s]})$, as already explained in Appendix A.3, and it is the disintegration of P w.r.t. $Z_{[0,s]}$, $Z_{[0,s]}$ being defined as $Z_{[0,s]}(y) := (Z_r(y))_{0 \leq r \leq s}$. On the other hand, what if we disintegrate P w.r.t. Z_s? Since the transition probability $\{p(s, x; t, \cdot)\}_{x \in \mathrm{Y}}$ is nothing but the disintegration of $(Z_s, Z_t)_*\mathsf{P}$ w.r.t. Z_s, the Markov property encoded in property (v) above can be equivalently restated in the following, more common way

$$\mathsf{P}(\cdot \,|\, Z_{[0,t]}) = \mathsf{P}(\cdot \,|\, Z_t).$$

A third characterization is actually possible: for every $t \in [0, 1]$ it holds

$$\mathsf{P}(Z_{[0,t]} \in \cdot,\, Z_{[t,1]} \in \cdot \,|\, Z_t) = \mathsf{P}(Z_{[0,t]} \in \cdot \,|\, Z_t)\mathsf{P}(\, Z_{[t,1]} \in \cdot \,|\, Z_t).$$

In a nutshell this means that, conditionally on the present state of the system, future and past are independent. By integration this identity yields

$$\mathsf{E}_{\mathsf{P}}[fg \mid Z_t] = \mathsf{E}_h P[f \mid Z_t]\mathsf{E}_h P[g \mid Z_t] \tag{A.4.1}$$

for all $t \in [0,1]$ and $f, g : \mathrm{Y} \to \mathbb{R}$ which are $Z_{[0,t]}$ and $Z_{[t,1]}$-measurable respectively.

Observe that the heat kernel on a $\mathsf{RCD}^*(K,\infty)$ space satisfies requirements $(i)-(v)$, thus the Brownian motion defined in Section 5.1 is actually a Markov process.

Bibliography

[1] L. Ambrosio, *Transport equation and Cauchy problem for BV vector fields*, Invent. Math., 158 (2004), pp. 227–260.

[2] L. Ambrosio and J. Feng, *On a class of first order Hamilton-Jacobi equations in metric spaces*, J. Differential Equations, 256 (2014), pp. 2194–2245.

[3] L. Ambrosio and N. Gigli, *A user's guide to optimal transport*, in Modelling and Optimisation of Flows on Networks, Lecture Notes in Mathematics, Springer Berlin Heidelberg, 2013, pp. 1–155.

[4] L. Ambrosio, N. Gigli, A. Mondino, and T. Rajala, *Riemannian Ricci curvature lower bounds in metric measure spaces with σ-finite measure*, Trans. Amer. Math. Soc., 367 (2015), pp. 4661–4701.

[5] L. Ambrosio, N. Gigli, and G. Savaré, *Gradient flows in metric spaces and in the space of probability measures*, Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, Basel, second ed., 2008.

[6] ——, *Calculus and heat flow in metric measure spaces and applications to spaces with Ricci bounds from below*, Invent. Math., 195 (2014), pp. 289–391.

[7] ——, *Metric measure spaces with Riemannian Ricci curvature bounded from below*, Duke Math. J., 163 (2014), pp. 1405–1490.

[8] ——, *Bakry-Émery curvature-dimension condition and Riemannian Ricci curvature bounds*, The Annals of Probability, 43 (2015), pp. 339–404. arXiv:1209.5786.

[9] L. Ambrosio, A. Mondino, and G. Savaré, *On the Bakry-Émery condition, the gradient estimates and the Local-to-Global property of $RCD^*(K,N)$ metric measure spaces*, The Journal of Geometric Analysis, 26 (2014), pp. 1–33.

[10] ——, *Nonlinear diffusion equations and curvature conditions in metric measure spaces*. Preprint, arXiv:1509.07273, 2015.

[11] L. Ambrosio and D. Trevisan, *Well posedness of Lagrangian flows and continuity equations in metric measure spaces*, Anal. PDE, 7 (2014), pp. 1179–1234.

[12] K. Bacher and K.-T. Sturm, *Localization and tensorization properties of the curvature-dimension condition for metric measure spaces*, J. Funct. Anal., 259 (2010), pp. 28–56.

[13] D. Bakry, *Transformations de Riesz pour les semi-groupes symétriques. II. Étude sous la condition $\Gamma_2 \geq 0$*, in Séminaire de probabilités, XIX, 1983/84, vol. 1123 of Lecture Notes in Math., Springer, Berlin, 1985, pp. 145–174.

[14] D. Bakry and M. Émery, *Diffusions hypercontractives*, in Séminaire de probabilités, XIX, 1983/84, vol. 1123 of Lecture Notes in Math., Springer, Berlin, 1985, pp. 177–206.

[15] D. Bakry, I. Gentil, and M. Ledoux, *Analysis and geometry of Markov diffusion operators*, vol. 348 of Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], Springer, Cham, 2014.

[16] D. Bao, S.-S. Chern, and Z. Shen, *An introduction to Riemann-Finsler geometry*, vol. 200 of Graduate Texts in Mathematics, Springer-Verlag, New York, 2000.

[17] J.-D. Benamou, G. Carlier, M. Cuturi, L. Nenna, and G. Peyré, *Iterative Bregman projections for regularized transportation problems*, SIAM J. Sci. Comput., 37 (2015), pp. A1111–A1138.

[18] J.-D. Benamou, G. Carlier, and L. Nenna, *A numerical method to solve multi-marginal optimal transport problems with Coulomb cost*, in Splitting methods in communication, imaging, science, and engineering, Sci. Comput., Springer, Cham, 2016, pp. 577–601.

[19] S. Bernstein, *Sur les liaisons entre les grandeurs aléatoires*, Verh. Internat. Math.-Kongr., Zurich, (1932), pp. 288–309.

[20] A. Beurling, *An automorphism of product measures*, Ann. of Math. (2), 72 (1960), pp. 189–200.

[21] M. Born, *Quantenmechanik der stoßvorgänge*, Zeitschrift für Physik A Hadrons and Nuclei, 38 (1926), pp. 803–827.

[22] J. M. Borwein, A. S. Lewis, and R. D. Nussbaum, *Entropy minimization, DAD problems, and doubly stochastic kernels*, J. Funct. Anal., 123 (1994), pp. 264–307.

[23] G. Carlier, V. Duval, G. Peyré, and B. Schmitzer, *Convergence of entropic schemes for optimal transport and gradient flows*, SIAM J. Math. Anal., 49 (2017), pp. 1385–1418.

[24] P. Cattiaux and C. Léonard, *Large deviations and Nelson processes*, Forum Math., 7 (1995), pp. 95–115.

[25] F. Cavalletti and E. Milman, *The globalization theorem for the curvature dimension condition.* Preprint, arXiv:1612.07623, 2016.

[26] F. Cavalletti and A. Mondino, *Sharp and rigid isoperimetric inequalities in metric-measure spaces with lower ricci curvature bounds.* Accepted at Inv. Math., arXiv:1502.06465, 2015.

[27] L. Chaumont and G. Uribe Bravo, *Markovian bridges: weak continuity and pathwise constructions*, Ann. Probab., 39 (2011), pp. 609–647.

[28] J. Cheeger, *Differentiability of Lipschitz functions on metric measure spaces*, Geom. Funct. Anal., 9 (1999), pp. 428–517.

[29] J. Cheeger and T. H. Colding, *Lower bounds on Ricci curvature and the almost rigidity of warped products*, Ann. of Math. (2), 144 (1996), pp. 189–237.

[30] J. Cheeger and D. Gromoll, *The splitting theorem for manifolds of nonnegative Ricci curvature*, J. Differential Geometry, 6 (1971/72), pp. 119–128.

[31] K. L. Chung and J.-C. Zambrini, *Introduction to random time and quantum randomness*, vol. 1 of Monographs of the Portuguese Mathematical Society, World Scientific Publishing Co., Inc., River Edge, NJ, new ed., 2003.

[32] T. H. Colding and A. Naber, *Sharp Hölder continuity of tangent cones for spaces with a lower Ricci curvature bound and applications*, Ann. of Math. (2), 176 (2012), pp. 1173–1229.

[33] D. Cordero-Erausquin, R. J. McCann, and M. Schmuckenschläger, *A Riemannian interpolation inequality à la Borell, Brascamp and Lieb*, Invent. Math., 146 (2001), pp. 219–257.

[34] A. B. Cruzeiro, L. Wu, and J. C. Zambrini, *Bernstein processes associated with a Markov process*, in Stochastic analysis and mathematical physics (Santiago, 1998), Trends Math., Birkhäuser Boston, Boston, MA, 2000, pp. 41–72.

[35] A. B. Cruzeiro and J.-C. Zambrini, *Malliavin calculus and Euclidean quantum mechanics. I. Functional calculus*, J. Funct. Anal., 96 (1991), pp. 62–95.

[36] I. Csiszár, *I-divergence geometry of probability distributions and minimization problems*, Ann. Probability, 3 (1975), pp. 146–158.

[37] M. Cuturi, *Sinkhorn distances: Lightspeed computation of optimal transport*, in Advances in Neural Information Processing Systems, 2013, pp. 2292–2300.

[38] P. Dai Pra, *A stochastic control approach to reciprocal diffusion processes*, Appl. Math. Optim., 23 (1991), pp. 313–329.

[39] G. Dal Maso, *An introduction to* Γ*-convergence*, vol. 8, Springer Science & Business Media, 2012.

[40] D. A. Dawson and J. Gärtner, *Large deviations from the McKean-Vlasov limit for weakly interacting diffusions*, Stochastics, 20 (1987), pp. 247–308.

[41] J. L. Doob, *Conditional Brownian motion and the boundary limits of harmonic functions*, Bull. Soc. Math. France, 85 (1957), pp. 431–458.

[42] J. L. Doob, *Classical potential theory and its probabilistic counterpart*, Classics in Mathematics, Springer-Verlag, Berlin, 2001. Reprint of the 1984 edition.

[43] M. H. Duong, V. Laschos, and M. Renger, *Wasserstein gradient flows from large deviations of many-particle limits*, ESAIM: Control, Optimisation and Calculus of Variations, 19 (2013), pp. 1166–1188.

[44] A. Einstein, B. Podolsky, and N. Rosen, *Can quantum-mechanical description of physical reality be considered complete?*, Physical review, 47 (1935), p. 777.

[45] M. Erbar, K. Kuwada, and K.-T. Sturm, *On the equivalence of the entropic curvature-dimension condition and Bochner's inequality on metric measure spaces*, Inventiones mathematicae, 201 (2014), pp. 1–79.

[46] M. Erbar, J. Maas, M. Renger, et al., *From large deviations to wasserstein gradient flows in multiple dimensions*, Electronic Communications in Probability, 20 (2015).

[47] H. Föllmer, *Random fields and diffusion processes*, in École d'Été de Probabilités de Saint-Flour XV–XVII, 1985–87, vol. 1362 of Lecture Notes in Math., Springer, Berlin, 1988, pp. 101–203.

[48] H. Föllmer and N. Gantert, *Entropy minimization and Schrödinger processes in infinite dimensions*, Ann. Probab., 25 (1997), pp. 901–926.

[49] R. Fortet, *Résolution d'un système d'équations de M. Schrödinger*, J. Math. Pures Appl., 19 (1940), pp. 83–105.

[50] W. Gangbo and A. Świech, *Metric viscosity solutions of Hamilton-Jacobi equations depending on local slopes*, Calc. Var. Partial Differential Equations, 54 (2015), pp. 1183–1218.

[51] N. Garofalo and A. Mondino, *Li-Yau and Harnack type inequalities in* $\mathrm{RCD}^*(K,N)$ *metric measure spaces*, Nonlinear Anal., 95 (2014), pp. 721–734.

[52] I. Gentil, C. Léonard, and L. Ripani, *About the analogy between optimal transport and minimal entropy*, arXiv preprint arXiv:1510.08230, (2015).

[53] N. Gigli, *On the heat flow on metric measure spaces: existence, uniqueness and stability*, Calc. Var. PDE, 39 (2010), pp. 101–120.

[54] ——, *Optimal maps in non branching spaces with Ricci curvature bounded from below*, Geom. Funct. Anal., 22 (2012), pp. 990–999.

[55] ——, *The splitting theorem in non-smooth context.* Preprint, arXiv:1302.5555, 2013.

[56] ——, *Nonsmooth differential geometry - an approach tailored for spaces with Ricci curvature bounded from below.* Accepted at Mem. Amer. Math. Soc., arXiv:1407.0809, 2014.

[57] ——, *An Overview of the Proof of the Splitting Theorem in Spaces with Non-Negative Ricci Curvature*, Analysis and Geometry in Metric Spaces, 2 (2014), pp. 169–213.

[58] ——, *On the differential structure of metric measure spaces and applications*, Mem. Amer. Math. Soc., 236 (2015), pp. vi+91.

[59] N. Gigli and B. Han, *The continuity equation on metric measure spaces*, Calc. Var. Partial Differential Equations, 53 (2013), pp. 149–177.

[60] N. Gigli, K. Kuwada, and S.-i. Ohta, *Heat flow on Alexandrov spaces*, Communications on Pure and Applied Mathematics, 66 (2013), pp. 307–331.

[61] N. Gigli and S. Mosconi, *The Abresch-Gromoll inequality in a non-smooth setting*, Discrete Contin. Dyn. Syst., 34 (2014), pp. 1481–1509.

[62] N. Gigli, T. Rajala, and K.-T. Sturm, *Optimal Maps and Exponentiation on Finite-Dimensional Spaces with Ricci Curvature Bounded from Below*, J. Geom. Anal., 26 (2016), pp. 2914–2929.

[63] N. Gigli and L. Tamanini, *Second order differentiation formula on compact $rcd^*(k,n)$ spaces.* Preprint, arXiv:1701.03932, 2017.

[64] R. S. Hamilton, *A matrix Harnack estimate for the heat equation*, Comm. Anal. Geom., 1 (1993), pp. 113–126.

[65] B. Han, *Ricci tensor on* $\mathrm{RCD}^*(K,N)$ *spaces.* Preprint, arXiv: 1412.0441.

[66] J. Heinonen, *Nonsmooth calculus*, Bull. Amer. Math. Soc. (N.S.), 44 (2007), pp. 163–232.

[67] B. Jamison, *Reciprocal processes: The stationary Gaussian case*, Ann. Math. Statist., 41 (1970), pp. 1624–1630.

[68] ——, *Reciprocal processes*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete, 30 (1974), pp. 65–86.

[69] ——, *The Markov processes of Schrödinger*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete, 32 (1975), pp. 323–331.

[70] R. Jiang, *The Li-Yau inequality and heat kernels on metric measure spaces*, J. Math. Pures Appl. (9), 104 (2015), pp. 29–57.

[71] R. Jiang, H. Li, and H.-C. Zhang, *Heat Kernel Bounds on Metric Measure Spaces and Some Applications*, Potential Analysis, 44 (2016), pp. 601–627.

[72] R. Jiang and H. Zhang, *Hamilton's gradient estimates and a monotonicity formula for heat flows on metric measure spaces*, Nonlinear Anal., 131 (2016), pp. 32–47.

[73] C. Ketterer, *Cones over metric measure spaces and the maximal diameter theorem*, J. Math. Pures Appl. (9), 103 (2015), pp. 1228–1275.

[74] B. L. Kotschwar, *Hamilton's gradient estimate for the heat kernel on complete manifolds*, Proc. Amer. Math. Soc., 135 (2007), pp. 3013–3019.

[75] M. Ledoux, *The geometry of Markov diffusion generators*, Ann. Fac. Sci. Toulouse Math. (6), 9 (2000), pp. 305–366. Probability theory.

[76] C. Léonard, *On the convexity of the entropy along entropic interpolations.* To appear in "Measure Theory in Non-Smooth Spaces", Partial Differential Equations and Measure Theory. De Gruyter Open, 2017. ArXiv:1310.1274v.

[77] C. Léonard, *Minimization of energy functionals applied to some inverse problems*, Applied Mathematics & Optimization, 44 (2001), pp. 273–297.

[78] C. Léonard, *Minimizers of energy functionals*, Acta Math. Hungar., 93 (2001), pp. 281–325.

[79] ——, *From the Schrödinger problem to the Monge-Kantorovich problem*, J. Funct. Anal., 262 (2012), pp. 1879–1920.

[80] ——, *Some properties of path measures*, Séminaire de Probabilités XLVI, (2014), pp. 207–230.

[81] ——, *A survey of the Schrödinger problem and some of its connections with optimal transport*, Discrete Contin. Dyn. Syst., 34 (2014), pp. 1533–1574.

[82] C. Léonard, S. Roelly, and J.-C. Zambrini, *Reciprocal processes. A measure-theoretical point of view*, Probab. Surv., 11 (2014), pp. 237–269.

[83] P. Li and S.-T. Yau, *On the parabolic kernel of the Schrödinger operator*, Acta Math., 156 (1986), pp. 153–201.

[84] S. Lisini, *Characterization of absolutely continuous curves in Wasserstein spaces*, Calc. Var. Partial Differential Equations, 28 (2007), pp. 85–120.

[85] J. Lott and C. Villani, *Hamilton-Jacobi semigroup on length spaces and applications*, J. Math. Pures Appl., 88 (2007), pp. 219–229.

[86] ——, *Weak curvature conditions and functional inequalities*, J. Funct. Anal., 245 (2007), pp. 311–333.

[87] ——, *Ricci curvature for metric-measure spaces via optimal transport*, Ann. of Math. (2), 169 (2009), pp. 903–991.

[88] T. Mikami, *Variational processes from the weak forward equation*, Comm. Math. Phys., 135 (1990), pp. 19–40.

[89] ——, *Dynamical systems in the variational formulation of the Fokker-Planck equation by the Wasserstein metric*, Appl. Math. Optim., 42 (2000), pp. 203–227.

[90] ——, *Optimal control for absolutely continuous stochastic processes and the mass transportation problem*, Electron. Comm. Probab., 7 (2002), pp. 199–213.

[91] ——, *Monge's problem with a quadratic cost by the zero-noise limit of h-path processes*, Probab. Theory Related Fields, 129 (2004), pp. 245–260.

[92] ——, *A simple proof of duality theorem for Monge-Kantorovich problem*, Kodai Math. J., 29 (2006), pp. 1–4.

[93] ——, *Optimal transportation problem on stochastic mechanics*, in Selected papers on probability and statistics, vol. 227 of Amer. Math. Soc. Transl. Ser. 2, Amer. Math. Soc., Providence, RI, 2009, pp. 75–94.

[94] ——, *A characterization of the Knothe-Rosenblatt processes by a convergence result*, SIAM J. Control Optim., 50 (2012), pp. 1903–1920.

[95] ——, *Two end points marginal problem by stochastic optimal transportation*, SIAM J. Control Optim., 53 (2015), pp. 2449–2461.

[96] T. Mikami and M. Thieullen, *Duality theorem for the stochastic optimal control problem*, Stochastic Process. Appl., 116 (2006), pp. 1815–1835.

[97] ——, *Optimal transportation problem by stochastic optimal control*, SIAM J. Control Optim., 47 (2008), pp. 1127–1139.

[98] A. Mondino and A. Naber, *Structure Theory of Metric-Measure Spaces with Lower Ricci Curvature Bounds I.* Preprint, arXiv:1405.2222, 2014.

[99] M. Nagasawa, *Transformations of diffusion and Schrödinger processes*, Probab. Theory Related Fields, 82 (1989), pp. 109–136.

[100] ——, *Schrödinger equations and diffusion theory*, vol. 86 of Monographs in Mathematics, Birkhäuser Verlag, Basel, 1993.

[101] ——, *Stochastic processes in quantum physics*, vol. 94 of Monographs in Mathematics, Birkhäuser Verlag, Basel, 2000.

[102] E. Nelson, *Dynamical theories of Brownian motion*, Princeton University Press, Princeton, N.J., 1967.

[103] L. Nenna, *Numerical methods for multi-marginal optimal transportation*, PhD thesis, PSL Research University, 2016.

[104] F. Otto and C. Villani, *Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality*, J. Funct. Anal., 173 (2000), pp. 361–400.

[105] T. Rajala, *Local Poincaré inequalities from stable curvature conditions on metric spaces*, Calc. Var. Partial Differential Equations, 44 (2012), pp. 477–494.

[106] T. Rajala and K.-T. Sturm, *Non-branching geodesics and optimal maps in strong $CD(K,\infty)$-spaces*, Calc. Var. Partial Differential Equations, 50 (2012), pp. 831–846.

[107] L. Rüschendorf and W. Thomsen, *Note on the Schrödinger equation and I-projections*, Statist. Probab. Lett., 17 (1993), pp. 369–375.

[108] ——, *Closedness of sum spaces and the generalized Schrödinger problem*, Teor. Veroyatnost. i Primenen., 42 (1997), pp. 576–590.

[109] G. Savaré, *Self-improvement of the Bakry-Émery condition and Wasserstein contraction of the heat flow in* $\mathrm{RCD}(K,\infty)$ *metric measure spaces*, Discrete Contin. Dyn. Syst., 34 (2014), pp. 1641–1661.

[110] E. Schrödinger, *Über die Umkehrung der Naturgesetze. Von E. Schrödinger. (Sonderausgabe a. d. Sitz.-Ber. d. Preuss. Akad. d. Wiss., Phys.-math. Klasse, 1931, IX.) Verlag W. de Gruyter, Berlin. Preis RM. 1,-*, Angewandte Chemie, 44 (1931), pp. 636–636.

[111] ——, *Sur la théorie relativiste de l'électron et l'interprétation de la mécanique quantique*, Ann. Inst. H. Poincaré, 2 (1932), pp. 269–310.

[112] N. Shanmugalingam, *Newtonian spaces: an extension of Sobolev spaces to metric measure spaces*, Rev. Mat. Iberoamericana, 16 (2000), pp. 243–279.

[113] K.-T. Sturm, *Heat kernel bounds on manifolds.*, Mathematische Annalen, 292 (1992), pp. 149–162.

[114] ——, *Analysis on local Dirichlet spaces. II. Upper Gaussian estimates for the fundamental solutions of parabolic equations*, Osaka J. Math., 32 (1995), pp. 275–312.

[115] ——, *On the geometry of metric measure spaces. I*, Acta Math., 196 (2006), pp. 65–131.

[116] ——, *On the geometry of metric measure spaces. II*, Acta Math., 196 (2006), pp. 133–177.

[117] ——, *Gradient flows for semiconvex functions on metric measure spaces-existence, uniqueness and lipschitz continuity*, arXiv preprint arXiv:1410.3966, (2014).

[118] ——, *Ricci Tensor for Diffusion Operators and Curvature-Dimension Inequalities under Conformal Transformations and Time Changes.* Preprint, arXiv:1401.0687, 2014.

[119] X. Tan and N. Touzi, *Optimal transportation under controlled stochastic dynamics*, Ann. Probab., 41 (2013), pp. 3201–3240.

[120] D. Trevisan, *Well-posedness of Diffusion Processes in Metric Measure Spaces*, PhD thesis, 2014.

[121] C. Villani, *Optimal transport. Old and new*, vol. 338 of Grundlehren der Mathematischen Wissenschaften, Springer-Verlag, Berlin, 2009.

[122] M.-K. von Renesse and K.-T. Sturm, *Transport inequalities, gradient estimates, entropy, and Ricci curvature*, Comm. Pure Appl. Math., 58 (2005), pp. 923–940.

[123] J.-C. Zambrini, *Variational processes and stochastic versions of mechanics*, J. Math. Phys., 27 (1986), pp. 2307–2330.