ANALYSIS & PDE

Volume 6 No. 1 2013

msp

Analysis & PDE

msp.org/apde

See inside back cover or msp.org/apde for submission instructions.

The subscription price for 2013 is US $160/year for the electronic version, and $310/year (+$35, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues from the last three years and changes of subscribers address should be sent to MSP.

Analysis & PDE (ISSN 1948-206X electronic, 2157-5045 printed) at Mathematical Sciences Publishers, 798 Evans Hall #3840, c/o University of California, Berkeley, CA 94720-3840, is published continuously online. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices.

APDE peer review and production are managed by EditFLOW® from Mathematical Sciences Publishers.

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

ANALYSIS AND PDE
Vol. 6, No. 1, 2013
dx.doi.org/10.2140/apde.2013.6.1

SOME RESULTS ON SCATTERING FOR LOG-SUBCRITICAL AND LOG-SUPERCRITICAL NONLINEAR WAVE EQUATIONS

HSI-WEI SHIH

We consider two problems in the asymptotic behavior of semilinear second order wave equations. First, we consider the $\dot{H}^1_x \times L^2_x$ scattering theory for the energy log-subcritical wave equation

$$\Box u = |u|^4 u g(|u|)$$

in $\mathbb{R}^{1+3}$, where g has logarithmic growth at 0. We discuss the solution with general (respectively spherically symmetric) initial data in the logarithmically weighted (respectively lower regularity) Sobolev space. We also include some observation about scattering in the energy subcritical case. The second problem studied involves the energy log-supercritical wave equation

$$\Box u = |u|^4 u \log^\alpha(2+|u|^2) \quad \text{for } 0 < \alpha \le \tfrac{4}{3}$$

in $\mathbb{R}^{1+3}$. We prove the same results of global existence and $(\dot{H}^1_x \cap \dot{H}^2_x) \times H^1_x$ scattering for this equation with a slightly higher power of the logarithm factor in the nonlinearity than that allowed in previous work by Tao (*J. Hyperbolic Differ. Equ.*, **4**:2 (2007), 259–265).

1. Introduction

Consider the semilinear wave equation

$$\begin{aligned} \Box u := -\partial_t^2 u + \triangle u &= f(u) \quad \text{on } \mathbb{R}\times\mathbb{R}^3, \\ u(0,x) &= u_0(x), \\ \partial_t u(0,x) &= u_1(x), \end{aligned} \tag{1}$$

where f is a complex-valued function. Let the potential function $F:\mathbb{C}\to\mathbb{R}$ be a real-valued function such that

$$2F_{\bar z}(z) = f(z), \tag{2}$$

with $F(0)=0$ and u being the solution to (1) with initial data $u_0 \in \dot{H}^1_x \cap \{\phi : \int_{\mathbb{R}^3} F(\phi)\,dx < \infty\}$ and $u_1 \in L^2_x$. We can easily verify that the equation has conserved energy

$$E(u)(t) := \int_{\mathbb{R}^3} \tfrac{1}{2}|u_t(t,x)|^2 + \tfrac{1}{2}|\nabla u(t,x)|^2 + F(u(t,x))\,dx. \tag{3}$$

MSC2010: 35L15.
Keywords: scattering, log-subcritical, radial Sobolev inequality.

The main goal of the paper is to study the $\dot{H}^1_x \times L^2_x$ scattering theory for log-subcritical wave equations with finite energy initial data, where the energy is defined by (3). In this paper, the term *log-subcritical wave equation* refers to (1) with f defined by

$$f(z) := \begin{cases} |z|^4 z g(|z|), & |z| \neq 0, \\ 0, & |z| = 0, \end{cases} \tag{4}$$

where $g : (0, \infty) \to \mathbb{R}$ is smooth, nonincreasing, and satisfies

$$g(x) := \begin{cases} -\log x, & 0 < x < \frac{1}{3}, \\ \sim 1, & \frac{1}{3} \leq x < 1, \\ 1, & x \geq 1. \end{cases} \tag{5}$$

We also prove global existence in the case of spherical symmetry for *log-supercritical wave equations*, by which we mean equations of the form

$$\Box u = |u|^4 u \log^\alpha (2 + |u|^2) \tag{6}$$

In this paper, we will allow $0 < \alpha \leq \frac{4}{3}$, extending the range $0 < \alpha \leq 1$ allowed in [Tao 2007]. We also assume that the initial data is in the energy space, the set of data for which the energy (3) is finite.

Remark 1.1. We can easily compute that the potential function of log-subcritical wave equations (1), (4), and (5) is

$$F_{\text{sub}}(z) = \begin{cases} -\frac{1}{6}|z|^6\big(\log(|z|) - \frac{1}{6}\big), & 0 < |z| < \frac{1}{3}, \\ \sim \frac{1}{6}|z|^6, & \frac{1}{3} \leq |z| < 1, \\ \frac{1}{6}|z|^6, & |z| \geq 1, \end{cases} \tag{7}$$

and the potential function of the log-supercritical wave equations (6) is

$$F_{\text{sup}}(z) \sim |z|^6 \log^\alpha (2 + |z|^2). \tag{8}$$

We quickly recall some common terminology associated to the scaling properties of (1). Consider $f(z) = |z|^{p-1}z$ and let u be the solution of (1). By scaling, $\lambda^{2/(1-p)}u(t/\lambda, x/\lambda)$ is also a solution with initial data $\lambda^{2/(1-p)}u_0(t_0/\lambda, x/\lambda)$ and $\lambda^{(1+p)/(1-p)}u_1(t_0/\lambda, x/\lambda)$. Hence the scaling of u preserves the homogeneous Sobolev norm $\|u_0\|_{\dot{H}^{s_c}(\mathbb{R}^3)} + \|u_1\|_{\dot{H}^{s_c-1}(\mathbb{R}^3)}$ if

$$s_c := \frac{3}{2} - \frac{2}{p-1}, \qquad \text{or equivalently} \qquad p = 1 + \frac{4}{3 - 2s_c}.$$

Definition 1.2. For $f(z) = |z|^{p-1}z$ and a given value s, we call (1) an $\dot{H}^s_x$-critical (subcritical, supercritical) nonlinear wave equation if p equals (is less than, is greater than) $1 + 4/(3-2s)$. In particular, when $s = 1$, we call (1) an energy critical (subcritical, supercritical) nonlinear wave equation if $p = 5$ ($p < 5$, $p > 5$).

The results of global existence and uniqueness for the energy-critical ($\Box u = |u|^4 u$) and energy-subcritical ($\Box u = |u|^{p-1}u$, where $p < 5$) wave equations are already established by [Brenner and von Wahl 1981; Struwe 1988; Grillakis 1990; 1992; Shatah and Struwe 1993; 1994; Kapitanski 1994; Ginibre and Velo 1985]. It is natural to consider the decay of the solution, which we expect to behave linearly as $t \to \pm\infty$. The decay estimate and scattering theory (see section 2 for definition) of critical wave equations are shown in [Bahouri and Shatah 1998]; see also [Bahouri and Gérard 1999; Ginibre and Velo 1989; Nakanishi 1999]. Hidano [2001] (see also [Ginibre and Velo 1987]), by the property of conformal invariance, proved that the solutions for certain subcritical wave equations ($\frac{5}{2} < p \le 3$) scatter in the weighted Sobolev space $\Sigma := X \times Y$, where

$$X := H_x^1(\mathbb{R}^3) \cap \{\phi : \ |x|\nabla\phi \in L_x^2(\mathbb{R}^3)\}, \quad Y := L_x^2(\mathbb{R}^3) \cap \{\phi : \ |x|\phi \in L_x^2(\mathbb{R}^3)\}.$$

However, for energy subcritical equations, the $\dot{H}_x^1 \times L_x^2$ scattering theory[1] still remains open. In this paper, we consider the solutions to the log-subcritical wave equations (1), (4), and (5) with finite energy initial data. The global existence result is established in [Grillakis 1990; 1992; Kapitanski 1994; Nakanishi 1999]. We will prove that the solutions with a class of initial data scatter in $\dot{H}_x^1 \times L_x^2$. This class of data is contained in logarithmically weighted Sobolev spaces $X_1 \times Y_1$, where

$$\begin{aligned} X_1 &:= \dot{H}_x^1(\mathbb{R}^3) \cap \{\phi : \ \log^\gamma(1+|x|)\nabla\phi \in L_x^2(\mathbb{R}^3)\}, \\ Y_1 &:= L_x^2(\mathbb{R}^3) \cap \{\phi : \ \log^\gamma(1+|x|)\phi \in L_x^2(\mathbb{R}^3)\} \end{aligned} \tag{9}$$

for some $\gamma > \frac{1}{2}$. For initial data in these spaces, we show that the potential energy of the solution decays logarithmically for all large times. After dividing the time interval suitably, this decay helps us to control the key spacetime norm $\|f(u)\|_{L_t^1 L_x^2}$. This spacetime bound implies scattering (we will sketch the proof in Section 2; see also [Bahouri and Shatah 1998]). Our proof of the spacetime bound involves establishing a decay rate for certain constant-time norms of the solution and a bootstrap scheme motivated by that in [Tao 2007]. We rely heavily on ideas from [Bahouri and Shatah 1998].

The second part of this paper considers the solution of log-subcritical wave equations with spherically symmetric data. We prove that the solution u with initial data in $X_2 \times Y_2$ scatters in $\dot{H}_x^1 \times L_x^2$, where

$$X_2 := \dot{H}_x^1(\mathbb{R}^3) \cap \Big(\bigcup_{\delta>0} \dot{H}_x^{1-\delta}(\mathbb{R}^3)\Big), \quad Y_2 := L_x^2(\mathbb{R}^3) \cap \Big(\bigcup_{\delta>0} \dot{H}_x^{-\delta}(\mathbb{R}^3)\Big). \tag{10}$$

Our proof again uses the ideas from [Tao 2007] and the classical Morawetz inequality; see [Morawetz 1968]. However, we need a slightly sharpened version of the bootstrap argument. We also give remarks for some specific energy subcritical wave equations (see page 15 and following).

The third part of this paper studies global existence for log-supercritical wave equations. The global regularity of energy supercritical wave equations ($\Box u = |u|^{p-1}u$, where $p > 5$) is still open. In [Tao 2007], the author considered the log-supercritical wave equation

$$\Box u = u^5 \log^\alpha(2+u^2) \tag{11}$$

with spherically symmetric initial data and established a global regularity result for $0 < \alpha \le 1$. For general initial data, the same result for loglog-supercritical wave equations

$$\Box u = u^5 \log^c(\log(10+u^2))$$

[1] $\dot{H}_x^1 \times L_x^2$ scattering is defined in Definition 2.1.

with $0 < c < \frac{8}{225}$ is obtained in [Roy 2009]. In the present paper, we extend the result in [Tao 2007] to the range $0 < \alpha \le \frac{4}{3}$, again for spherically symmetric data. This improvement is attained by employing the potential energy bound in place of the kinetic energy bound used in [Tao 2007] for pointwise control.

2. Definitions, notation, and preliminaries

Throughout this paper, we use $M \lesssim N$ to denote the estimate $M \le CN$ for some absolute constant C (which can vary from line to line).

We use $L_t^q L_x^r$ to denote the spacetime norm

$$\|u\|_{L_t^q L_x^r(I\times\mathbb{R}^3)} := \left(\int_I \left(\int_{\mathbb{R}^3} |u(t,x)|^r\,dx\right)^{q/r} dt\right)^{1/q}$$

with the usual modifications when q or r is equal to infinity.

Definition 2.1. We say that a global solution $u : \mathbb{R}\times\mathbb{R}^3 \to \mathbb{C}$ to (1) scatters in $\dot{H}_x^1 \times L_x^2$ (or $\dot{H}_x^1 \times L_x^2$ scattering) as $t \to +\infty$ $(-\infty)$ if there exists a linear solution v^+ (v^-) with initial data in $\dot{H}_x^1 \times L_x^2$ such that

$$\|u(t,x) - v^+(t,x)\|_{\dot{H}_x^1\times L_x^2} \to 0 \quad \text{as } t \to +\infty$$
$$(\|u(t,x) - v^-(t,x)\|_{\dot{H}_x^1\times L_x^2} \to 0 \quad \text{as } t \to -\infty).$$

Remark 2.2. We will sketch here that the *spacetime bound*,

$$\|f(u)\|_{L_t^1 L_x^2([t_0,\infty)\times\mathbb{R}^3)} < \infty \tag{12}$$

for some $t_0 > 0$, of the solution u to (1) implies the $\dot{H}_x^1 \times L_x^2$ scattering (as $t \to \infty$). Let

$$u \in C_t^1(\mathbb{R}, \dot{H}_x^1(\mathbb{R}^3)) \cap C_t^0(\mathbb{R}, L_x^2(\mathbb{R}^3))$$

be the solution to (1) and let v satisfy $\Box v = 0$ with initial data $v_0 \in \dot{H}_x^1(\mathbb{R}^3)$, $v_1 \in L_x^2(\mathbb{R}^3)$ (to be chosen shortly). By Duhamel's formula,

$$u(t,x) = \cos\big(t\sqrt{-\Delta}\big)u_0(x) + \frac{\sin\big(t\sqrt{-\Delta}\big)}{\sqrt{-\Delta}}u_1(x) - \int_0^t \frac{\sin\big((t-\tau)\sqrt{-\Delta}\big)}{\sqrt{-\Delta}} f(u(\tau))\,d\tau \tag{13}$$

and

$$v(t,x) = \cos\big(t\sqrt{-\Delta}\big)v_0(x) + \frac{\sin\big(t\sqrt{-\Delta}\big)}{\sqrt{-\Delta}}v_1(x), \tag{14}$$

where the operators $\cos\big(t\sqrt{-\Delta}\big)$ and $\sin\big(t\sqrt{-\Delta}\big)/\sqrt{-\Delta}$ are defined by

$$\big(\cos\big(t\sqrt{-\Delta}\big)\phi\big)\widehat{\ }(\xi) = \cos(t|\xi|)\hat{\phi}(\xi)$$

and

$$\left(\frac{\sin\big(t\sqrt{-\Delta}\big)}{\sqrt{-\Delta}}\phi\right)\widehat{\ }(\xi) = \frac{\sin(t|\xi|)}{|\xi|}\hat{\phi}(\xi).$$

Hence, that the solution u scatters and asymptotically approaches v in $\dot{H}_x^1 \times L_x^2$ means that

$$\left\| \cos\left(t\sqrt{-\Delta}\right)(u_0 - v_0) + \frac{\sin\left(t\sqrt{-\Delta}\right)}{\sqrt{-\Delta}}(u_1 - v_1) - \int_0^t \frac{\sin\left((t-\tau)\sqrt{-\Delta}\right)}{\sqrt{-\Delta}} f(u(\tau))\, d\tau \right\|_{\dot{H}_x^1 \times L_x^2} \to 0 \quad (15)$$

as $t \to \infty$. From basic trigonometric identities, we can verify that (15) is implied by

$$\left\| (u_0 - v_0) + \int_0^t \frac{\sin\left(-\tau\sqrt{-\Delta}\right)}{\sqrt{-\Delta}} f(u(\tau))\, d\tau \right\|_{\dot{H}_x^1} \to 0$$

and

$$\left\| (u_1 - v_1) + \int_0^t \cos\left(-\tau\sqrt{-\Delta}\right) f(u(\tau))\, d\tau \right\|_{L_x^2} \to 0$$

as $t \to \infty$. Therefore, if

$$\left(\int_0^t \frac{\sin\left(-\tau\sqrt{-\Delta}\right)}{\sqrt{-\Delta}} f(u(\tau))\, d\tau,\ \int_0^t \cos\left(-\tau\sqrt{-\Delta}\right) f(u(\tau))\, d\tau \right) \quad (16)$$

converges in $\dot{H}_x^1 \times L_x^2$ as $t \to \infty$, and we take

$$\begin{aligned} v_0(x) &:= u_0(x) - \int_0^\infty \frac{\sin\left(-\tau\sqrt{-\Delta}\right)}{\sqrt{-\Delta}} f(u(\tau))\, d\tau, \\ v_1(x) &:= u_1(x) - \int_0^\infty \cos\left(-\tau\sqrt{-\Delta}\right) f(u(\tau))\, d\tau, \end{aligned}$$

in (14), we then have, by (13), (14), and elementary trigonometric formulas,

$$\begin{aligned} \|u - v\|_{\dot{H}_x^1 \times L_x^2} &= \Bigg\| -\int_0^t \frac{\sin\left((t-\tau)\sqrt{-\Delta}\right)}{\sqrt{-\Delta}} f(u(\tau))\, d\tau + \int_0^\infty \frac{\cos\left(t\sqrt{-\Delta}\right)\sin\left(-\tau\sqrt{-\Delta}\right)}{\sqrt{-\Delta}} f(u(\tau))\, d\tau \\ &\qquad + \int_0^\infty \frac{\sin\left(t\sqrt{-\Delta}\right)\cos\left(-\tau\sqrt{-\Delta}\right)}{\sqrt{-\Delta}} f(u(\tau))\, d\tau \Bigg\|_{\dot{H}_x^1 \times L_x^2} \\ &= \left\| \int_t^\infty \frac{\sin(t-\tau)\sqrt{-\Delta}}{\sqrt{-\Delta}} f(u(\tau))\, d\tau \right\|_{\dot{H}_x^1 \times L_x^2}. \end{aligned} \quad (17)$$

It remains to show two things:

(i) Our initial data v_0, v_1 are well-defined, that is, that (16) does indeed converge in $\dot{H}_x^1 \times L_x^2$.

(ii) The right side of (17) converges to 0 as $t \to \infty$.

The claim (i) can be shown in several ways, for example, by showing that

$$\lim_{N\to\infty} \left\| \int_N^\infty \frac{\sin\left(-\tau\sqrt{-\Delta}\right)}{\sqrt{-\Delta}} f(u(\tau))\, d\tau \right\|_{\dot{H}_x^1} = 0 \quad \text{and} \quad \lim_{N\to\infty} \left\| \int_N^\infty \cos\left(-\tau\sqrt{-\Delta}\right) f(u(\tau))\, d\tau \right\|_{L_x^2} = 0,$$

where $N \in \mathbb{N}$. These two equalities follow from the dominated convergence theorem once we show that

$$\int_0^\infty \left\| \frac{\sin\left(-\tau\sqrt{-\Delta}\right)}{\sqrt{-\Delta}} f(u(\tau)) \right\|_{\dot{H}_x^1} (\tau)\, d\tau < \infty \quad \text{and} \quad \int_0^\infty \left\| \cos\left(-\tau\sqrt{-\Delta}\right) f(u(\tau)) \right\|_{L_x^2} (\tau)\, d\tau < \infty.$$

But this follows quickly from (12) and the Plancherel theorem.

Claim (ii) has already been established in the discussion of claim (i). This concludes the argument that the finiteness of (12) implies scattering.

Definition 2.3. We say that the pair (q, r) is admissible if $2 \le q, r \le \infty$, $(q, r) \neq (2, \infty)$ and

$$\frac{1}{q} + \frac{1}{r} \le \frac{1}{2}. \tag{18}$$

Theorem 2.4 (Strichartz estimates for wave equation [Strichartz 1977; Kapitanski 1989; Ginibre and Velo 1995; Lindblad and Sogge 1995; Keel and Tao 1998]). *Let I be a time interval and let $u : I \times \mathbb{R}^3 \to \mathbb{C}$ be a Schwartz solution to the wave equation $\Box u = G$ with initial data $u(t_0) = u_0$, $\partial_t u(t_0) = u_1$ for some $t_0 \in I$. Then we have the estimates*

$$\|u\|_{L_t^q L_x^r(I\times\mathbb{R}^3)} + \|u\|_{C_t^0 \dot{H}_x^\sigma(I\times\mathbb{R}^3)} + \|\partial_t u\|_{C_t^0 \dot{H}_x^{\sigma-1}(I\times\mathbb{R}^3)} \lesssim \|u_0\|_{\dot{H}_x^\sigma(\mathbb{R}^3)} + \|u_1\|_{\dot{H}_x^{\sigma-1}(\mathbb{R}^3)} + \|G\|_{L_t^{\tilde{q}'} L_x^{\tilde{r}'}(I\times\mathbb{R}^3)}, \tag{19}$$

where (q, r) and $(\tilde{q}, \tilde{r})$ are admissible pairs and obey the scaling condition

$$\frac{1}{q} + \frac{3}{r} = \frac{3}{2} - \sigma = \frac{1}{\tilde{q}'} + \frac{3}{\tilde{r}'} - 2, \tag{20}$$

and where $\tilde{q}'$ and $\tilde{r}'$ are conjugate to $\tilde{q}$ and $\tilde{r}$, respectively. In addition, if u is a spherically symmetric solution, we allow $(q, r) = (2, \infty)$.

We define the Strichartz space $S_\sigma(I)$ for any time interval I, as the closure of the Schwartz function on $I \times \mathbb{R}^3$ under the norm

$$\|u\|_{S_\sigma(I)} := \sup_{(q,r)\text{ admissible}} \|u\|_{L_t^q L_x^r(I\times\mathbb{R}^3)}, \tag{21}$$

where (q, r) satisfies (20).

Morawetz inequality [Morawetz 1968]. *Let I be any time interval and $u : I \times \mathbb{R}^3 \to \mathbb{C}$ be the solution to* (1) *with finite energy E. Let F be the potential function as in* (2). *Then*

$$\int_I \int_{\mathbb{R}^3} \frac{F(u)}{|x|}\, dx\, dt \lesssim E. \tag{22}$$

Spherically symmetric solutions. In the last part of this section, we assume that u is the spherically symmetric solution to the log-subcritical wave equations (4), (5) (or log-supercritical wave equation (6)) and F is the corresponding potential function. We obtain the following a priori estimate for the solution.

Lemma 2.5 (pointwise estimate for spherically symmetric solution [Ginibre et al. 1992; Tao 2007]). *Let I be any time interval and let $u : I \times \mathbb{R}^3 \to \mathbb{C}$ be the spherically symmetric solution to the log-subcritical wave equations* (4), (5) (*or log-supercritical wave equation* (6)) *with finite energy E and vanishing at ∞. Let F be the potential function. Then, for any $t \in I$,*

$$|x|^2 (F(u)^{1/2} |u|)(t, x) \lesssim E \tag{23}$$

Proof. We tackle the log-subcritical case; the proof for the log-supercritical case is similar and easier. Define $\phi(z) := (F(z))^{1/2}z$ and $r := |x|$. From (7), we can compute that, for fixed t,

$$\left|\partial_r(\phi(u(t,x)))\right| \lesssim |u|^3|\partial_r u|(t,x)\chi_{\{|u|\ge 1/3\}}(x) + |u|^3(-\log|u|)^{1/2}|\partial_r u|(t,x)\chi_{\{|u|<1/3\}}(x),$$

where χ is the characteristic function on $\mathbb{R}^3$. Then, by the fundamental theorem of calculus, Hölder's inequality, and energy conservation,

$$\begin{aligned}
|\phi(u(t,x))| &\lesssim \left|\int_r^\infty \big[|u|^3|\partial_r u|\chi_{\{|u|\ge 1/3\}} + |u|^3(-\log|u|)^{1/2}|\partial_r u|\chi_{\{|u|<1/3\}}\big](t,s)\,ds\right| \\
&\lesssim \left(\int_r^\infty \frac{|u|^6}{s^2}s^2\chi^2_{\{|u|\ge 1/3\}}\,ds\right)^{1/2}\left(\int_r^\infty \frac{|\partial_r u|^2}{s^2}\chi_{\{|u|\ge 1/3\}}s^2\,ds\right)^{1/2} \\
&\quad + \left(\int_r^\infty \frac{|u|^6(-\log|u|)}{s^2}s^2\chi^2_{\{|u|<1/3\}}\,ds\right)^{1/2}\left(\int_r^\infty \frac{|\partial_r u|^2}{s^2}\chi_{\{|u|<1/3\}}s^2\,ds\right)^{1/2} \\
&\lesssim \frac{1}{r^2}\left(\int_{\mathbb{R}^3} F(u)\,dx\right)^{1/2} E^{1/2} \lesssim \frac{1}{r^2}E. \qquad \square
\end{aligned}$$

Inserting (23) into (22), we obtain that, for any time interval I,

$$\int_I\int_{\mathbb{R}^3} F^{5/4}(u)|u|^{\frac12}\,dx\,dt \le \int_I\int_{\mathbb{R}^3} \frac{F(u)}{|x|}\cdot \sup_{x\in\mathbb{R}^3}(|x|F^{1/4}(u)|u|^{1/2})\,dx\,dt \lesssim E^{3/2}. \tag{24}$$

This implies

$$\int_I\int_{\{|u|\le 1/3\}} |u|^8(-\log|u|)^{5/4}\,dx\,dt + \int_I\int_{\{|u|>1/3\}} |u|^8\,dx\,dt \lesssim E^{3/2} \qquad \text{(log-subcritical case)} \tag{25}$$

and

$$\int_I\int_{\mathbb{R}^3} |u|^8\log^{5\alpha/4}(2+|u|^2)\,dx\,dt \lesssim E^{3/2} \qquad \text{(log-supercritical case).} \tag{26}$$

3. Log-subcritical wave equations

In this section, we consider the scattering theory for log-subcritical wave equations. We can take advantage of time reversal symmetry, and it suffices to prove that the solution u scatters in $\dot H^1_x \times L^2_x$ as $t\to\infty$.

Throughout this section, we use the notation

$$A = \left\{(t,x)\in(0,\infty)\times\mathbb{R}^3:\ |u|<\tfrac13\right\}, \quad B = \left\{(t,x)\in(0,\infty)\times\mathbb{R}^3:\ |u|\ge\tfrac13\right\},$$

and for any interval I,

$$A_I = A\cap(I\times\mathbb{R}^3), \quad B_I = B\cap(I\times\mathbb{R}^3). \tag{27}$$

General initial data in log-weighted Sobolev spaces.

Theorem 3.1. *Let $\gamma > \frac12$ and let u be the solution to the log-subcritical wave equations* (1), (4), *and* (5) *with initial data*

$$u_0(x)\in X_1, \quad u_1(x)\in Y_1, \tag{28}$$

where X_1 and Y_1 are defined by (9). *Then u scatters in $\dot H^1_x\times L^2_x$.*

Proof. We need some decay estimates for the equation with initial data satisfying (28).

Lemma 3.2. *Let γ and u be as in Theorem 3.1. There exists $T = T(\|u_0\|_{X_1}, \|u_1\|_{Y_1}, \gamma) \gg 1$ such that, for $\tau > T$,*

$$\int_{\mathbb{R}^3} F(u(\tau, x))\, dx \lesssim \frac{1}{\log^{2\gamma} \tau}, \tag{29}$$

where $F(z) = F_{\mathrm{sub}}(z)$ is defined by (7).

Proof. We essentially follow the proof of Lemma 2.1 in [Bahouri and Shatah 1998], with some changes. Define

$$e[u](t, x) := \tfrac{1}{2}|\partial_t u(t, x)|^2 + \tfrac{1}{2}|\nabla u(t, x)|^2 + F(u(t, x)).$$

We claim that there exists $C_\gamma = C_\gamma(\|u_0\|_{X_1}, \|u_1\|_{Y_1}, \gamma) \gg 1$ such that for $s > C_\gamma$,

$$\int_{|x|>s} e[u](0, x)\, dx \lesssim \frac{1}{\log^{2\gamma} s}. \tag{30}$$

We prove this claim in the Appendix and continue the proof of this lemma here. Choose T such that $T > \max(C_\gamma^2, \log^{4\gamma} T)$. We aim to show that (29) holds for all $\tau > T$.

Define the truncated forward light cone by

$$K_a^b(c) := \{(t, x) : a \le t \le b,\ |x| \le t + c,\ 0 \le a < b \le \infty\}$$

and the boundary of the truncated cone by

$$M_a^b(c) := \partial K_a^b(c) = \{(t, x) : a \le t \le b,\ |x| = t + c,\ 0 \le a < b \le \infty\}.$$

Fix $\tau > T$ and let $s = \sqrt{\tau} > C_\gamma$. For any $t_1 > 0$, the energy conservation law on the exterior of the truncated forward light cone $K_0^{t_1}(s)$ implies that

$$\int_{|x|>s+t_1} e[u](t_1)\, dx + \frac{1}{\sqrt{2}} \operatorname{flux}(0, t_1, s) = \int_{|x|>s} e[u](0)\, dx \lesssim \frac{1}{\log^{2\gamma} s}, \tag{31}$$

where

$$\operatorname{flux}(a, b, c) := \int_{M_a^b(c)} \left\{ \frac{1}{2}\left| u_t + \frac{x \cdot \nabla u}{|x|} \right|^2 + F(u) \right\} d\sigma.$$

Hence

$$\int_{|x|>s+\tau} F(u(\tau))\, dx \le \int_{|x|>s+\tau} e[u](\tau)\, dx \lesssim \frac{1}{\log^{2\gamma} s} \lesssim \frac{1}{\log^{2\gamma} \tau}, \tag{32}$$

and it suffices to show that

$$\int_{|x|\le s+\tau} F(u(\tau))\, dx \lesssim \frac{1}{\log^{2\gamma} \tau}. \tag{33}$$

Define $w(t, x) = u(t - s, x)$. The bound (33) is equivalent to

$$\int_{|x|\le s+\tau} F(w(s + \tau))\, dx \lesssim \frac{1}{\log^{2\gamma} \tau}.$$

Set $w_t := \partial_t w$. Multiplying the equation $f(w) - \Box w = 0$ by $t w_t + x \cdot \nabla w + w$, we get

$$\partial_t(tQ_0 + w_t w) - \operatorname{div}(tP_0) + R_0 = 0, \tag{34}$$

where

$$\begin{aligned}
Q_0 &= e[w] + w_t\Big(\frac{x}{t}\cdot\nabla w\Big),\\
P_0 &= \frac{x}{t}\bigg(\frac{w_t^2 - |\nabla w|^2}{2} - F(w)\bigg) + \nabla w\Big(w_t + \frac{x}{t}\cdot\nabla w + \frac{w}{t}\Big),\\
R_0 &= |w|^6 g(|w|) - 4F(w),
\end{aligned}$$

with g defined by (5). Define the horizontal sections of the forward solid cone by

$$D(t) := \{|x| \in \mathbb{R}^3 : |x| \le t\}.$$

Fix $0 < T_1 < T_2$ and integrate (34) on $K_{T_1}^{T_2}(0)$. By the divergence theorem, we have

$$\begin{aligned}
\int_{D(T_2)}(T_2Q_0+w_tw)\,dx-\int_{D(T_1)}(T_1Q_0+w_tw)\,dx-\frac{1}{\sqrt{2}}\int_{M_{T_1}^{T_2}(0)}\Big(tQ_0+w_tw+tP_0\frac{x}{|x|}\Big)\,d\sigma+\int_{K_{T_1}^{T_2}(0)}R_0\,dx\,dt\\
=: L_1+L_2+L_3+L_4=0. \quad (35)
\end{aligned}$$

Now, following the same steps as in [Bahouri and Shatah 1998], we define $v(y) := w(|y|, y)$. Since L_3 is the integral on $M_{T_1}^{T_2}(0)$, using spherical coordinates, we obtain that

$$\begin{aligned}
L_3 &= -\int_{T_1}^{T_2}\int_{S^2} r\Big(v_r+\frac{v}{r}\Big)^2 r^2\,dr\,d\omega + \frac12\int_{S^2}T_2^2v^2(T_2\omega)\,d\omega - \frac12\int_{S^2}T_1^2v^2(T_1)\omega\,d\omega, \quad (36)\\
L_1 &= \int_{D(T_2)}\bigg\{T_2\bigg(\frac{|w_t|^2}{2}+\frac12\Big(w_r+\frac1r w\Big)^2+\frac{1}{2r^2}|\nabla_\omega w|^2+F(w)\bigg)+r\Big(w_r+\frac1r w\Big)w_t\bigg\}\,dx\\
&\qquad - \frac12\int_{S^2}T_2^2v^2(T_2\omega)\,d\omega, \quad (37)
\end{aligned}$$

and

$$\begin{aligned}
L_2 = -\int_{D(T_1)}\bigg\{T_1\bigg(\frac{|w_t|^2}{2}+\tfrac12\Big(w_r+\frac1r w\Big)^2+\frac{1}{2r^2}|\nabla_\omega w|^2+F(w)\bigg)+r\Big(w_r+\frac1r w\Big)w_t\bigg\}\,dx\\
+\frac12\int_{S^2}T_1^2v^2(T_1\omega)\,d\omega. \quad (38)
\end{aligned}$$

Since $L_4 \ge 0$, plugging (36), (37) and (38) into (35), we deduce that

$$T_2\int_{D(T_2)}F(w)\,dx \le CT_1E + \int_{T_1}^{T_2}\int_{S^2}T_2\Big(v_r+\frac{v}{r}\Big)^2 r^2\,dr\,d\omega,$$

where C is a constant and E is the energy. Therefore,

$$\int_{D(T_2)}F(w(T_2))\,dx \le C\frac{T_1}{T_2}E + \int_{T_1}^{T_2}\int_{S^2}\Big(v_r+\frac{v}{r}\Big)^2 r^2\,dr\,d\omega. \tag{39}$$

For any $T_1 \geq s$, by (31), the second term in the right-hand side of (39) is controlled by

$$\int_{T_1}^{T_2}\int_{S^2}\Big(v_r+\frac{v}{r}\Big)^2 r^2\,dr\,d\omega \lesssim \int_{M_{T_1}^{T_2}(0)}\Big\{\frac{1}{2}\Big|w_t+\frac{x\cdot\nabla w}{|x|}\Big|^2\Big\}\,d\sigma \lesssim \frac{1}{\log^{2\gamma} s} \lesssim \frac{1}{\log^{2\gamma}\tau}.$$

Now, choosing $T_2 = \tau + s$ and $T_1 = (\tau+s)/\log^{2\gamma}\tau > \sqrt{\tau} = s$, (39) implies

$$\int_{D(\tau+s)} F(w(\tau+s,x))\,dx \lesssim \frac{1}{\log^{2\gamma}\tau}. \tag{40}$$

Combining (32) and (40), the lemma is proved. □

Before we prove Theorem 3.1, let's observe the following fact. Let I be any time interval with length $3 < |I| < \infty$. By Hölder's inequality, we have that, for $0 < \delta < 2$,

$$\begin{aligned}
&\| |u|^4 u(-\log|u|)\|_{L_t^1 L_x^2(I\times\mathbb{R}^3)} \\
&\quad \leq \|u^{3-\delta}(-\log|u|)^{(3-\delta)/6}\|_{L_t^\infty L_x^{6/(3-\delta)}(I\times\mathbb{R}^3)} \|u^2\|_{L_t^{2/(2-\delta)} L_x^{6/\delta}(I\times\mathbb{R}^3)} \|u^\delta(-\log|u|)^{(3+\delta)/6}\|_{L_t^{2/\delta} L_x^\infty(I\times\mathbb{R}^3)} \\
&\quad = \|u(-\log|u|)^{1/6}\|^{3-\delta}_{L_t^\infty L_x^6(I\times\mathbb{R}^3)} \|u\|^2_{L_t^{4/(2-\delta)} L_x^{12/\delta}(I\times\mathbb{R}^3)} \|u^\delta(-\log|u|)^{(3+\delta)/6}\|_{L_t^{2/\delta} L_x^\infty(I\times\mathbb{R}^3)} \\
&\quad \leq \|u(-\log|u|)^{1/6}\|^{3-\delta}_{L_t^\infty L_x^6(I\times\mathbb{R}^3)} \|u\|^2_{L_t^{4/(2-\delta)} L_x^{12/\delta}(I\times\mathbb{R}^3)} \|u^\delta(-\log|u|)^{(3+\delta)/6}\|_{L_t^\infty L_x^\infty(I\times\mathbb{R}^3)} |I|^{\delta/2}.
\end{aligned}$$

If $|u| \leq \frac{1}{3}$, we can estimate that

$$\|u^\delta(-\log|u|)^{(3+\delta)/6}\|_{L_t^\infty L_x^\infty(I\times\mathbb{R}^3)} \lesssim \Big(\frac{1}{\delta}\Big)^{1/2+\delta/6}.$$

Letting $\delta = 2/\log|I|$, we obtain

$$\| |u|^4 u(-\log|u|)\|_{L_t^1 L_x^2(I\times\mathbb{R}^3)} \lesssim \|u(-\log|u|)^{1/6}\|^{3-\delta}_{L_t^\infty L_x^6(I\times\mathbb{R}^3)} \|u\|^2_{L_t^{4/(2-\delta)} L_x^{12/\delta}(I\times\mathbb{R}^3)} \log^{1/2}|I|. \tag{41}$$

To complete the proof of Theorem 3.1, by Remark 2.2, it suffices to show that

$$\|f(u)\|_{L_t^1 L_x^2([T,\infty)\times\mathbb{R}^3)} < \infty \quad \text{for some } T < \infty.$$

Let $J = (3^i, \infty)$, where i is sufficiently large and to be determined later. Then

$$\|f(u)\|_{L_t^1 L_x^2(J\times\mathbb{R}^3)} \lesssim \||u|^4 u(-\log|u|)\|_{L_t^1 L_x^2(A_J)} + \| |u|^4 u\|_{L_t^1 L_x^2(B_J)} =: M_1 + M_2.$$

Since $(2+\delta, 6(2+\delta)/\delta)$ is an admissible pair satisfying (20) for $\sigma = 1$, from Hölder's inequality and Lemma 3.2,

$$M_2 \leq \|u\|^{3-\delta}_{L_t^\infty L_x^6(B_J)} \|u\|^{2+\delta}_{L_t^{2+\delta} L_x^{6(2+\delta)/\delta}(B_J)} \lesssim \frac{1}{(\log(3^i))^{(3-\delta)/3\gamma}} \|u\|^{2+\delta}_{S_1(J)}. \tag{42}$$

On the other hand, define interval J_k by subdividing J according to $J = \bigcup_{k=1}^\infty (3^{2^{k-1}i}, 3^{2^k i}) =: \bigcup_{k=1}^\infty J_k$. Define $\delta_k := 2/\log|J_k|$. By (41), Lemma 3.2, and the fact that the admissible pairs $(4/(2-\delta_k), 12/\delta_k)$

satisfy (20) for $\sigma = 1$, we have

$$M_1 \le \sum_{k=1}^{\infty} \|u^5(-\log|u|)\|_{L_t^1 L_x^2(J_k\times\mathbb{R}^3)} \lesssim \sum_{k=1}^{\infty}\left(\frac{1}{(\log 3^{2^{k-1}i})^{(3-\delta_k)/3\gamma}}(\log 3^{2^k i})^{1/2}\right)\|u\|^2_{L_t^{4/(2-\delta_k)} L_x^{12/\delta_k}(J_k\times\mathbb{R}^3)}$$
$$\lesssim \left(\sum_{k=1}^{\infty} i^{1/2-(3-\delta_k)/3\gamma}\cdot 2^{(k-1)(1/2-(3-\delta_k)/3)\gamma}\right)\|u\|^2_{S_1(J)}.$$

Since $\gamma > \frac{1}{2}$, we can choose i sufficiently large such that $\big(((3-\delta_k)/3)\gamma - \frac{1}{2}\big) > c > 0$ for all k. Hence

$$M_1 \lesssim i^{-c}\sum_{k=1}^{\infty} 2^{-(k-1)c}\|u\|^2_{S_1(J)}. \tag{43}$$

Combining (42) and (43), for $\epsilon_0 > 0$ sufficiently small, we can choose i sufficiently large such that

$$\|f(u)\|_{L_t^1 L_x^2(J\times\mathbb{R}^3)} \le \epsilon_0(\|u\|^2_{S_1(J)} + \|u\|^{2+\delta}_{S_1(J)}).$$

By the Strichartz estimate (19), we have

$$\|u\|_{S_1(J)} \le CE^{1/2} + \epsilon_0(\|u\|^2_{S_1(J)} + \|u\|^{2+\delta}_{S_1(J)}).$$

From a continuity argument, we conclude that

$$\|u\|_{S_1(J)} \le 2CE^{1/2}.$$

This implies that

$$\|f(u)\|_{L_t^1 L_x^2(J\times\mathbb{R}^3)} < \infty. \qquad \square$$

Spherically symmetric initial data in lower regularity Sobolev spaces. In this subsection, we consider the solutions to the log-subcritical wave equations with spherically symmetric initial data. If the finite energy initial data are in any lower regularity Sobolev spaces, we obtain the $\dot{H}_x^1 \times L_x^2$ scattering. The spirit of the proof follows from [Tao 2007] and a slightly sharpened bootstrap argument in Lemmas 3.5 and 3.6.

Throughout this subsection, for given $\delta > 0$, we denote

$$Z(t) := \|u(t,x)\|_{\dot{H}_x^{1-\delta}(\mathbb{R}^3)} + \|\partial_t u(t,x)\|_{\dot{H}_x^{-\delta}(\mathbb{R}^3)}. \tag{44}$$

It is easy to show that $Z(t) > 0$ for any time t.[2]

Theorem 3.3. *Let u be the solution to the log-subcritical wave equations* (1), (4), (5) *with spherically symmetric initial data*

$$u_0(x) \in X_2, \quad u_1(x) \in Y_2, \tag{45}$$

where X_2 and Y_2 are defined by (10). *Then u scatters in $\dot{H}_x^1 \times L_x^2$.*

To prove Theorem 3.3, we need some intermediate lemmas.

[2] If $Z(t_0) = 0$ for some t_0, it is easy to prove that the solution u has energy $E(t_0) = 0$ and, hence, $E(t) = 0$ for any time t, by energy conservation. This implies the solution $u(t,x) \equiv 0$ for all t.

Lemma 3.4. *Let $I=[a,b]$ be any interval where $0\le a<b\le\infty$ and let u be the solution to the log-subcritical wave equations* (1), (4), (5) *with spherically symmetric initial data*

$$u(a,x)=u_0(x)\in \dot H^1_x\cap \dot H^{1-\delta}_x,\quad \partial_t u(a,x)=u_1(x)\in L^2_x\cap \dot H^{-\delta}_x$$

for some fixed $0<\delta<\frac{1}{2}$. Then there exists $0<\epsilon(\delta)\ll 1$ such that for $0<\epsilon<\epsilon(\delta)$,

$$\|u\|_{S_{1-\delta}(I)}\lesssim Z(a)+(\|u\|^{1+\epsilon/(2\delta)}_{S_{1-\delta}(I)}+\|u\|_{S_{1-\delta}(I)})\big(\||u|(-\log|u|)^{5/32}\|^{4-\epsilon/(2\delta)-\epsilon}_{L^8_{t,x}(A_I)}+\|u\|^4_{L^8_{t,x}(B_I)}\big)\Big(\frac{1}{\epsilon}\Big)^{7/16},\tag{46}$$

where the constant hidden in (46) *is independent of the interval I and ϵ.*

Proof. By the Strichartz estimate (19),

$$\|u\|_{S_{1-\delta}(I)}\lesssim Z(a)+\|f(u)\|_{L_t^{2/(2-\delta)}L_x^{2/(1+\delta)}(I\times\mathbb{R}^3)}.\tag{47}$$

Consider that

$$\|f(u)\|_{L_t^{2/(2-\delta)}L_x^{2/(1+\delta)}(I\times\mathbb{R}^3)}\lesssim\|-|u|^4u(\log(|u|))\|_{L_t^{2/(2-\delta)}L_x^{2/(1+\delta)}(A_I)}+\||u|^4u\|_{L_t^{2/(2-\delta)}L_x^{2/(1+\delta)}(B_I)}=:N_1+N_2$$

with A_I and B_I as in (27). By Hölder's inequality,

$$N_2\le\|u\|_{L_t^{2/(1-\delta)}L_x^{2/\delta}(B_I)}\|u\|^4_{L^8_{t,x}(B_I)}\le\|u\|_{S_{1-\delta}(I)}\|u\|^4_{L^8_{t,x}(B_I)}.\tag{48}$$

On the other hand, choosing $\epsilon(\delta)$ sufficiently small such that for $0<\epsilon<\epsilon(\delta)$,

$$0<\frac{1}{p}:=\frac{8\delta+\epsilon-8\delta^2+2\epsilon\delta}{8(2\delta+\epsilon)}\le\frac{1}{2},\quad 0<\frac{1}{q}:=\frac{\delta}{2}+\frac{\epsilon(1-2\delta)}{8(2\delta+\epsilon)}\le\frac{1}{2},\quad \frac{3}{8}\approx\frac{12+5\epsilon/(2\delta)+5\epsilon}{32}<\frac{7}{16}.$$

It is clear that (p,q) is an admissible pair satisfying (20) for $\sigma=1-\delta$. By Hölder's inequality and interpolation theory, we can estimate that

$$\begin{aligned}N_1&\le\||u|^{5-\epsilon}(-\log|u|)^{5(4-\epsilon/(2\delta)-\epsilon)/32}\|_{L_t^{2/(2-\delta)}L_x^{2/(1+\delta)}(A_I)}\||u|^{\epsilon}(-\log|u|)^{(12+5\epsilon/(2\delta)+5\epsilon)/32}\|_{L^\infty_{t,x}(A_I)}\\&\le\|u\|^{1+\epsilon/(2\delta)}_{L_t^pL_x^q(A_I)}\|u(-\log|u|)^{5/32}\|^{4-\epsilon/(2\delta)-\epsilon}_{L^8_{t,x}(A_I)}\||u|^{\epsilon}(-\log|u|)^{(12+5\epsilon/(2\delta)+5\epsilon)/32}\|_{L^\infty_{t,x}(A_I)}\qquad(49)\\&\lesssim\|u\|^{1+\epsilon/(2\delta)}_{L_t^pL_x^q(A_I)}\||u|(-\log|u|)^{5/32}\|^{4-\epsilon/(2\delta)-\epsilon}_{L^8_{t,x}(A_I)}\Big(\frac{1}{\epsilon}\Big)^{upnfrac12+5\epsilon/(2\delta)+5\epsilon32}.\qquad(50)\end{aligned}$$

The last factor of (50) comes from maximizing the last factor on the right of (49) using calculus. We note that the constant hidden in the last inequality is independent of ϵ. By (48) and (50), we have

$$\|f(u)\|_{L_t^{2/(2-\delta)}L_x^{2/(1+\delta)}(I\times\mathbb{R}^3)}\lesssim\|u\|^{1+\epsilon/(2\delta)}_{S_{1-\delta}(I)}\||u|(-\log|u|)^{5/32}\|^{4-\epsilon/(2\delta)-\epsilon}_{L^8_{t,x}(A_I)}\Big(\frac{1}{\epsilon}\Big)^{7/16}+\|u\|_{S_{1-\delta}(I)}\|u\|^4_{L^8_{t,x}(B_I)}.$$

From (47),

$$\|u\|_{S_{1-\delta}(I)}\lesssim Z(a)+\|u\|^{1+\epsilon/(2\delta)}_{S_{1-\delta}(I)}\||u|(-\log|u|)^{5/32}\|^{4-\epsilon/(2\delta)-\epsilon}_{L^8_{t,x}(A_I)}\Big(\frac{1}{\epsilon}\Big)^{\frac{7}{16}}+\|u\|_{S_{1-\delta}(I)}\|u\|^4_{L^8_{t,x}(B_I)}\lesssim\text{RHS of (46)}.$$

One can check that all constants hidden in the inequalities above are independent of the interval I and ϵ. Hence, Lemma 3.4 is proved. ∎

Lemma 3.5 (continuity argument). *Let I $(=[a,b])$ and u satisfy the assumptions of Lemma 3.4, C be the constant hidden in* (46) *and $0<\epsilon(\delta)$ be chosen in Lemma 3.4. Let $\epsilon_0 = 1/(100C)$ and $0<\epsilon<\epsilon(\delta)$ such that $Z(a)^{\epsilon/(2\delta)} \geq \frac{1}{2}$ and $(2C)^{\epsilon/(2\delta)} \leq 2$. We define*

$$Q(I) := \left(\||u|(-\log|u|)^{5/32}\|_{L^8_{t,x}(A_I)}^{4-\epsilon/(2\delta)-\epsilon} + \|u\|_{L^8_{t,x}(B_I)}^{4-\epsilon/(2\delta)-\epsilon}\right).$$

If $\|u\|_{L^8_{t,x}(B_I)} \leq 1$ and $Q(I) \leq \epsilon_0(\epsilon^{7/16}/(Z(a)^{\epsilon/(2\delta)}))$, we have

$$\|u\|_{S_{1-\delta}(I)} \leq 2CZ(a).$$

Proof. We prove this lemma by contradiction. For $0 \leq t \leq b-a$, from the dominated convergence theorem, we have that the function $\Phi(t) := \|u\|_{S_{1-\delta}([a,a+t])}$ is nondecreasing and continuous in $[0, b-a]$ and $\Phi(0)=0$. By the hypothesis and (46), we have

$$\Phi(t) \leq CZ(a) + \tfrac{1}{100}(\Phi(t)^{1+\epsilon/(2\delta)} + \Phi(t))\left(\frac{1}{Z(a)^{\epsilon/(2\delta)}}\right) \tag{51}$$

for all $t \in [0, b-a]$. Assume for contradiction that there exists $t_0 \in [0, b-a)$ such that $\Phi(t_0) = 2CZ(a)$. If $2CZ(a) < 1$, (51) implies that

$$2CZ(a) = \Phi(t_0) \leq CZ(a) + \frac{1}{50}(2CZ(a))\left(\frac{1}{Z(a)^{\epsilon/(2\delta)}}\right) \leq \frac{11}{10}CZ(a).$$

On the other hand, if $2CZ(a) \geq 1$, (51) implies that

$$2CZ(a) = \Phi(t_0) \leq CZ(a) + \frac{1}{50}(2CZ(a))^{1+\epsilon/(2\delta)}\left(\frac{1}{Z(a)^{\epsilon/(2\delta)}}\right) \leq \frac{11}{10}CZ(a).$$

We get contradictions in both situations, and the lemma is proved. □

Lemma 3.6 (finite division). *Let I $(=[a,b])$ and u satisfy the assumptions of Lemma 3.4 and C be the constant hidden in* (46). *We denote $Z_i = (2C)^i Z(a)$, where $i = 0, 1, 2, \ldots$. For any $\epsilon_0 > 0$, we can choose $\epsilon \ll 1$ and finitely many numbers $a = T_0 < T_1 < T_2 < \cdots < T_N < T_{N+1} = b$, where $N = N(\epsilon_0, \epsilon, \delta, E, Z_0, C)$, such that for $I_j := [T_j, T_{j+1}]$,*

$$Q(I_j) = \epsilon_0\left(\frac{\epsilon^{7/16}}{Z_j^{\epsilon/(2\delta)}}\right) \tag{52}$$

for $0 \leq j \leq N-1$ and $Q(I_N) \leq \epsilon_0(\epsilon^{7/16}/Z_N^{\epsilon/(2\delta)})$.

Proof. We observe that

$$\sum_{i=0}^{\infty}\left[\epsilon_0\left(\frac{\epsilon^{7/16}}{Z_i^{\epsilon/(2\delta)}}\right)\right]^{8/(4-\epsilon/(2\delta)-\epsilon)} \gtrsim_{\epsilon_0, Z_0} \left\{\epsilon^{7/(8-\epsilon/(\delta)-2\epsilon)}\sum_{i=0}^{\infty}\frac{1}{(2C)^{8i\epsilon/(8\delta-\epsilon-2\delta\epsilon)}}\right\} \to \infty \quad \text{as } \epsilon \to 0.$$

Therefore, by (25), we can choose ϵ sufficiently small such that

$$3\left(\int\int_A |u|^8(-\log|u|)^{5/4}\,dx\,dt + \int\int_B |u|^8\,dx\,dt\right) < \sum_{i=0}^{K}\left[\epsilon_0\left(\frac{\epsilon^{7/16}}{Z_i^{\epsilon/(2\delta)}}\right)\right]^{8/(4-\epsilon/(2\delta)-\epsilon)} \tag{53}$$

for some $K = K(\epsilon_0, \epsilon, \delta, E, Z_0, C)$.

Fix this ϵ. If $Q(I) < \epsilon_0(\epsilon^{7/16}/(Z_0^{\epsilon/(2\delta)}))$, we say $T_1 = b$ and the lemma is proved. Otherwise, we can choose $0 < T_1 < b$ such that (52) holds for $j = 0$. Again, if $Q([T_1, b) < \epsilon_0(\epsilon^{7/16}/(Z_1^{\epsilon/(2\delta)}))$, we say $T_2 = b$. Otherwise, we can choose $T_1 < T_2 < b$ such that (52) holds for $j = 1$. By continuing this process, we can choose $a < T_1 < T_2 < \cdots$ such that (52) holds for $j = 0, 1, \ldots$. It suffices to show that this process will stop in at most $K + 1$ steps. Indeed, assume that there are more than $K + 1$ subintervals satisfying (52). Since

$$Q(I_j)^{8/(4-\epsilon/(2\delta)-\epsilon)} \leq 3\bigg(\iint_{A_{I_j}} |u|^8(-\log|u|)^{5/4}\,dx\,dt + \iint_{B_{I_j}} |u|^8\,dx\,dt\bigg),$$

for $j = 0, 1, \ldots$, by our construction of I_j, we have

$$\begin{aligned}\sum_{j=0}^{K+1} \epsilon_0\bigg(\frac{\epsilon^{7/16}}{Z_j^{\epsilon/(2\delta)}}\bigg) &= \sum_{j=0}^{K+1} Q(I_j)^{8/(4-\epsilon/(2\delta)-\epsilon)}\\ &\leq \sum_{i=0}^{K+1} 3\bigg(\iint_{A_{I_j}} |u|^8(-\log|u|)^{5/4}\,dx\,dt + \iint_{B_{I_j}} |u|^8\,dx\,dt\bigg)\\ &\leq 3\bigg(\iint_{A} |u|^8(-\log|u|)^{5/4}\,dx\,dt + \iint_{B} |u|^8\,dx\,dt\bigg).\end{aligned}$$

This contradicts (53), and the lemma is proved. □

Corollary 3.7. *Let I and u satisfy the assumptions of Lemma 3.4 and C be the constant hidden in* (46)*. If $\|u\|_{L^8_{t,x}(B_I)} \leq 1$, $u \in L_{t,x}^{8/(1+2\delta)}(I \times \mathbb{R}^3)$.*

Proof. Let $\epsilon(\delta)$ be chosen in Lemma 3.4 and $0 < \epsilon < \epsilon(\delta)$ satisfy Lemma 3.6, $Z(a)^{\epsilon/(2\delta)} \geq \frac{1}{2}$ and $(2C)^{\epsilon/(2\delta)} \leq 2$. Let $\{I_j\}_{j=0}^N$ be the subintervals constructed by Lemma 3.6 such that (52) holds for $0 \leq j \leq N$.

We claim that

$$\|u\|_{S_{1-\delta}(I_j)} \leq 2CZ_j \quad \text{for } 0 \leq j \leq N, \tag{54}$$

where $Z_j = (2C)^j Z(a)$. Indeed, by Lemma 3.5, (54) holds for $j = 0$. Again, if (54) holds for $j = k-1$, we have $Z(T_k) \leq \|u\|_{S_{1-\delta}(I_{k-1})} \leq Z_k$. Since $Z_k^{\epsilon/(2\delta)} \geq Z(a)^{\epsilon/(2\delta)} \geq \frac{1}{2}$, applying Lemma 3.5 on the interval I_k, we obtain (54) for $j = k$. By induction on j, the claim is proved and this implies

$$\|u\|_{L_{t,x}^{8/(1+2\delta)}(I\times\mathbb{R}^3)} \leq \sum_{j=0}^{N+1} \|u\|_{S_{1-\delta}(I_j)} \leq \sum_{j=0}^{N+1} (2C)^j Z_0 < \infty.$$ □

Corollary 3.8. *Let u be the solution to the log-subcritical wave equations* (1), (4), (5) *with spherically symmetric initial data*

$$u(0,x) = u_0(x) \in \dot{H}_x^1 \cap \dot{H}_x^{1-\delta}, \quad \partial_t u(0,x) = u_1(x) \in L_x^2 \cap \dot{H}_x^{-\delta}$$

for some fixed $0 < \delta < \frac{1}{2}$. Then $u \in L_{t,x}^{8/(1+2\delta)}(\mathbb{R}_+ \times \mathbb{R}^3)$.

Proof. By (25), we can choose finitely many numbers $0 = S_0 < S_1 < \cdots < S_{M-1} < S_M = \infty$ such that $\|u\|_{L^8_{t,x}(B_{[S_k,S_{k+1}]})} \le 1$ for $0 \le k \le M$. By Corollary 3.7 and energy conservation, we have

$$(u(S_k, x), \partial_t u(S_k, x)) \in (\dot{H}^1_x \cap \dot{H}^{1-\delta}_x) \times (L^2_x \cap \dot{H}^{-\delta}_x)$$

and $\|u\|_{L^{8/1+2\delta}_{t,x}([S_k,S_{k+1}]\times\mathbb{R}^3)} < \infty$ for $0 \le k \le M$. Hence

$$\|u\|_{L^{8/(1+2\delta)}_{t,x}(\mathbb{R}_+\times\mathbb{R}^3)} \le \sum_{k=0}^{M} \|u\|_{L^{8/(1+2\delta)}_{t,x}([S_k,S_{k+1}]\times\mathbb{R}^3)} < \infty. \qquad \square$$

To finish the proof of Theorem 3.3, by Remark 2.2, it suffices to show that $\|f(u)\|_{L^1_t L^2_x((T,\infty)\times\mathbb{R}^3)} < \infty$ for some $0 < T < \infty$. Since the initial data satisfy (45), we can choose some $0 < \delta < \frac{1}{2}$ such that $u_0 \in \dot{H}^1_x(\mathbb{R}^3) \cap \dot{H}^{1-\delta}_x(\mathbb{R}^3)$ and $u_1 \in L^2_x(\mathbb{R}^3) \cap \dot{H}^{-\delta}_x(\mathbb{R}^3)$. Observe that

$$\begin{aligned}
&\|f(u)\|_{L^1_t L^2_x((T,\infty)\times\mathbb{R}^3)} \\
&\qquad\lesssim \| |u|^5(\log(|u|))\|_{L^1_t L^2_x(A_T)} + \| |u|^5\|_{L^1_t L^2_x(B_T)} \\
&\qquad\lesssim \|u\|^{4/(1+2\delta)}_{L^{8}_{t,x}/(1+2\delta)(A_T)} \|u\|_{L^2_t L^\infty_x(A_T)} \|u^{8\delta/(1+2\delta)}(\log(|u|))\|_{L^\infty_{t,x}(A_T)} + \|u\|^4_{L^8_{t,x}(B_T)} \|u\|_{L^2_t L^\infty_x(B_T)} \\
&\qquad\lesssim \|u\|_{L^2_t L^\infty_x((T,\infty)\times\mathbb{R}^3)} \Big[\Big(\frac{1+2\delta}{8\delta}\Big) \|u\|^{4/(1+2\delta)}_{L^{8}_{t,x}/(1+2\delta)(A_T)} + \|u\|^4_{L^8_{t,x}(B_T)}\Big],
\end{aligned}$$

where $A_T := A \cap ((T, \infty) \times \mathbb{R}^3)$ and $B_T := B \cap ((T, \infty) \times \mathbb{R}^3)$. The last inequality above is from the fact that $|u^{8\delta/(1+2\delta)}(\log(|u|))| \lesssim (1 + 2\delta)/(8\delta)$ for $|u| \le \frac{1}{3}$. By Corollary 3.8 and (25), for sufficiently small $\epsilon > 0$, we can choose $T = T(\epsilon)$ sufficiently large such that

$$\Big(\frac{1+2\delta}{8\delta}\Big) \|u\|^{4/(1+2\delta)}_{L^{8}_{t,x}/(1+2\delta)(A_T)} + \|u\|^4_{L^8_{t,x}(B_T)} < \epsilon.$$

Hence, by the Strichartz inequality [Klainerman and Machedon 1993],

$$\|u\|_{L^2_t L^\infty_x((T,\infty)\times\mathbb{R}^3)} \le C E^{1/2} + \epsilon C \|u\|_{L^2_t L^\infty_x((T,\infty)\times\mathbb{R}^3)}.$$

Again for $\epsilon < 1/(2C)$, we have $\|u\|_{L^2_t L^\infty_x((T,\infty)\times\mathbb{R}^3)} < 2CE^{1/2}$ and this implies $\|f(u)\|_{L^1_t L^2_x((T,\infty)\times\mathbb{R}^3)} < \infty$.

Energy subcritical nonlinear wave equations with specific spherically symmetric initial data. In the last part of this section, we will discuss an observation, for energy subcritical nonlinear wave equations, inspired by the proof of Theorem 3.3. For given $0 < \delta < \frac{1}{2}$, let $(u_0, u_1) \in (\dot{H}^1_x(\mathbb{R}^3) \cap \dot{H}^{1-\delta}_x(\mathbb{R}^3)) \times (L^2_x(\mathbb{R}^3) \cap \dot{H}^{-\delta}_x(\mathbb{R}^3))$ be spherically symmetric functions. In this subsection, we consider the energy-subcritical nonlinear wave equation

$$\Box u = |u|^{4-\epsilon} u, \quad u(0, x) = u_0(x), \quad \partial_t u(0, x) = u_1(x), \tag{55}$$

where we allow ϵ to depend on the given data (u_0, u_1). That is, we find a relation (R) (see Definition 3.10) among ϵ, the energy E, and $Z(0)$ as in (44), the lower regularity norm of the initial data, for which the solution scatters. We remark that relation (R) holds for data large in both the energy and $\dot{H}^{1-\delta}$ norms provided that ϵ is taken sufficiently small (depending on the size of these norms). In [Lindblad and Sogge

1995], scattering was established in $\dot{H}^{1-\delta}$ for the $\dot{H}^{1-\delta}$ critical nonlinear wave equation from small data. Our remarks here are related to that work, for example, relation (R) quantifies the extent to which large data can be allowed. Also, we will prove scattering in $\dot{H}^1$, rather than $\dot{H}^{1-\delta}$.

In order to prove that u scatters in $\dot{H}^1_x \times L^2_x$, It suffices to show that $\|u^{5-\epsilon}\|_{L^1_t L^2_x([T,\infty)\times\mathbb{R}^3)} < \infty$ for some $T < \infty$. By the Strichartz estimate and Hölder's inequality,

$$\begin{aligned}\|u\|_{L^2_t L^\infty_x([T,\infty)\times\mathbb{R}^3)} &\le CE^{1/2} + C\|u^{5-\epsilon}\|_{L^1_t L^2_x([T,\infty)\times\mathbb{R}^3)}\\ &\le CE^{1/2} + C\|u\|_{L^2_t L^\infty_x([T,\infty)\times\mathbb{R}^3)}\|u\|^{4-\epsilon}_{L^{8-2\epsilon}_{t,x}([T,\infty)\times\mathbb{R}^3)}.\end{aligned}$$

Following similar arguments as in the proof of Theorem 3.3, we only need to show that

$$\|u\|_{L^{8-2\epsilon}_{t,x}([T,\infty)\times\mathbb{R}^3)} < \infty \quad \text{for some } T < \infty.$$

Let $\epsilon_0(\delta) := 8\delta/(1+2\delta)$ (so that $\dot{H}^{1-\delta}$ is the scale invariant norm for (55) with $\epsilon = \epsilon_0(\delta)$). We restrict to the case $0 < \epsilon < \epsilon_0(\delta)$.

In this case, (55) is $\dot{H}^{1-\delta}$-supercritical nonlinear wave equation. We denote

$$\gamma_\epsilon = \frac{3\epsilon}{16\delta - \frac{5}{2}\delta\epsilon - \frac{5}{4}\epsilon}, \quad \kappa_\epsilon = \frac{8-\frac{5}{4}\epsilon}{4-\gamma_\epsilon-\epsilon}, \quad \frac{1}{\alpha_\epsilon} = \frac{1+2\delta}{8} + \frac{3(1-2\delta)}{8(1+\gamma_\epsilon)}, \quad \frac{1}{\beta_\epsilon} = \frac{1+2\delta}{8} - \frac{1-2\delta}{8(1+\gamma_\epsilon)}.$$

Note that

(i) as $\epsilon \to \epsilon_0(\delta)$, $\gamma_\epsilon \to 4-\epsilon$ and $\kappa_\epsilon \to \infty$;

(ii) $(\alpha_\epsilon, \beta_\epsilon)$ is an admissible pair satisfying (20) for $\sigma = 1-\delta$.

Remark 3.9. Let u be the spherically symmetric solution to the energy-subcritical nonlinear wave equation (55) with energy E. We observe that Lemma 2.5 holds for u. Hence, for any interval $I = [a, b]$ where $0 \le a < b \le \infty$, (24) implies

$$\int_I \int_{\mathbb{R}^3} |u(t,x)|^{8-5\epsilon/4}\, dx\, dt \le C_1 E^{3/2}, \tag{56}$$

where we can choose the constant C_1 to be independent of ϵ. Moreover, by the Strichartz estimate,

$$\|u\|_{S_{1-\delta}(I)} \le CZ(a) + C\|u^{5-\epsilon}\|_{L^{2/(2-\delta)}_t L^{2/(1+\delta)}_x(I\times\mathbb{R}^3)}. \tag{57}$$

Definition 3.10. Given $0 < \delta < \frac{1}{2}$, let $0 < \epsilon < \epsilon_0(\delta)$, u be the solution to (55) with energy E and lower regularity norm $Z(0) > 0$. We say that the triple $(E, Z(0), \epsilon)$ satisfies the relation (R) if

$$C_1 E^{3/2} \le \left(\frac{1}{2(2C)^{1+\gamma_\epsilon} Z(0)^{\gamma_\epsilon}}\right)^{\kappa_\epsilon} \frac{1}{1-(2C)^{-\gamma_\epsilon\kappa_\epsilon}}$$

Lemma 3.11. *Given $0 < \delta < \frac{1}{2}$ and $0 < \epsilon < \epsilon_0(\delta)$, let*

$$(u_0, u_1) \in (\dot{H}^1_x(\mathbb{R}^3) \cap \dot{H}^{1-\delta}_x(\mathbb{R}^3)) \times (L^2_x(\mathbb{R}^3) \cap \dot{H}^{-\delta}_x(\mathbb{R}^3))$$

be spherically symmetric functions such that the triple $(E, Z(0), \epsilon)$ satisfies (R) and u is the solution to (55). Then $u \in L^{8/(1+2\delta)}_{t,x}(\mathbb{R}_+ \times \mathbb{R}^3)$.

Proof. Since $(E, Z(0), \epsilon)$ satisfies (R), by (56) and an argument similar to that in proof of Lemma 3.6, we can choose finitely many numbers $0 = T_0 < T_1 < \cdots < T_N < T_{N+1} = \infty$ such that

$$\|u\|^{4-\gamma_\epsilon-\epsilon}_{L^{8-5\epsilon/4}_{t,x}([T_i,T_{i+1}]\times\mathbb{R}^3)} = \frac{1}{2(2C)^{1+\gamma_\epsilon}((2C)^i Z(0))^{\gamma_\epsilon}} \tag{58}$$

for $0 \le i \le N-1$ and

$$\|u\|^{4-\gamma_\epsilon-\epsilon}_{L^{8-5\epsilon/4}_{t,x}([T_N,T_{N+1}]\times\mathbb{R}^3)} \le \frac{1}{2(2C)^{1+\gamma_\epsilon}((2C)^N Z(0))^{\gamma_\epsilon}}.$$

We claim that

$$Z(T_i) < (2C)^i Z(0) \tag{59}$$

and

$$\|u\|_{S_{1-\delta}([T_i,T_{i+1}])} < (2C)^{i+1} Z(0) \tag{60}$$

for $0 \le i \le N$.

Observe that (59) is clearly true for $i = 0$ and $Z(T_i) \le \|u\|_{S_{1-\delta}([T_{i-1},T_i])}$ for $1 \le i \le N$. Hence it suffices to show that (60) holds and then (59) is automatically true.

A similar proof to that of Lemma 3.5 applies here. Assume (60) is true for $i \le j-1$. We aim to prove (60) for $i = j$. (Note that (59) follows from our assumption when $i = j$.) Let $\phi(t) = \|u\|_{S_{1-\delta}([T_j,T_j+t])}$. Then ϕ is a continuous and nondecreasing function on $[0, T_{j+1}-T_j]$ and $\phi(0) = 0$. Assume for contradiction that there exists $t_0 \in [0, T_{j+1}-T_j]$ such that $\phi(t_0) = (2C)^{j+1}Z(0)$. By Hölder's inequality, (57), (58), and (59), we have

$$\begin{aligned}
(2C)^{j+1}Z(0) = \phi(t_0) &\le CZ(T_j) + C\|u^{5-\epsilon}\|_{L^2_t/(2-\delta)L^{2/(1+\delta)}_x([T_j,T_j+t_0]\times\mathbb{R}^3)} \\
&\le CZ(T_j) + C\|u\|^{1+\gamma_\epsilon}_{L^{\alpha_\epsilon}_t L^{\beta_\epsilon}_x([T_j,T_j+t_0]\times\mathbb{R}^3)} \|u\|^{4-\gamma_\epsilon-\epsilon}_{L^{8-5\epsilon/4}_{t,x}([T_j,T_j+t_0]\times\mathbb{R}^3)} \\
&\le CZ(T_j) + C\|u\|^{1+\gamma_\epsilon}_{S_{1-\delta}([T_j,T_j+t_0])} \|u\|^{4-\gamma_\epsilon-\epsilon}_{L^{8-5\epsilon/4}_{t,x}([T_j,T_j+t_0]\times\mathbb{R}^3)} \\
&\le C(2C)^j Z(0) + \frac{1}{4[(2C)^{j+1}Z(0)]^{\gamma_\epsilon}} \|u\|^{1+\gamma_\epsilon}_{S_{1-\delta}([T_j,T_j+t_0])} \\
&< \tfrac{1}{2}(2C)^{j+1}Z(0) + \frac{1}{4[(2C)^{j+1}Z(0)]^{\gamma_\epsilon}} \times [(2C)^{j+1}Z(0)]^{1+\gamma_\epsilon} \\
&= \tfrac{3}{4}(2C)^{j+1}Z(0).
\end{aligned}$$

The contradiction implies that (60) holds for $i = j$. By an inductive argument on i, the claim is proved. To finish proving this lemma, we have

$$\|u\|_{L^{8/(1+2\delta)}_{t,x}(\mathbb{R}_+\times\mathbb{R}^3)} \le \sum_{i=0}^{N+1} \|u\|_{S_{1-\delta}([T_i,T_{i+1}])} \le \sum_{i=0}^{N+1} (2C)^i Z(0) < \infty \qquad \square$$

Corollary 3.12. *Let δ, ϵ, u_0, u_1 and u satisfy the assumptions of Lemma 3.11. Then u scatters in $\dot{H}^1_x \times L^2_x$.*

Proof. By the above discussion, it suffices to show

$$\|u\|_{L^{8-2\epsilon}_{t,x}([T,\infty)\times\mathbb{R}^3)} < \infty$$

for some $T < \infty$, since $0 < \epsilon < \epsilon_0(\delta)$ is equivalent to $8/(1+2\delta) < 8-2\epsilon$. The proof of $L^{8-2\epsilon}_{t,x}$ spacetime bound is straightforward by (56), Lemma 3.11, and interpolation theory. $\square$

4. Log-supercritical wave equation

For spherically symmetric log-supercritical nonlinear wave equation (1), (6) with finite energy E, we observe that the potential energy bound provides slightly better pointwise control, (26), of the solution than the one from the kinetic energy bound[3]; see [Ginibre et al. 1992; Tao 2007]. In this section, we consider a slightly more supercritical wave equation than the equation in [Tao 2007] and prove the same global regularity result by using (26).

Theorem 4.1. *Define*

$$\widetilde{H}^2_x(\mathbb{R}^3) := \dot{H}^1_x(\mathbb{R}^3) \cap \dot{H}^2_x(\mathbb{R}^3).$$

Let $0 < \alpha \le \frac{4}{3}$ *and* (u_0, u_1) *be smooth, compactly supported, and spherically symmetric initial data with energy* E. *Then there exists a global smooth solution to*

$$\Box u = |u|^4 u \log^\alpha(2+|u|^2), \quad u(0,x) = u_0(x), \quad \partial_t u(0,x) = u_1(x). \tag{61}$$

Furthermore, we have the universal bound of $\widetilde{H}^2_x \times H^1_x$ *norm, which depends on both the energy* E *and* $\widetilde{H}^2_x \times H^1_x$ *norm of the initial data, of the solution* u; *this implies that the solution* u *scatters in* $\widetilde{H}^2_x(\mathbb{R}^3) \times H^1_x(\mathbb{R}^3)$.[4]

Remark 4.2. This theorem was proved in [Tao 2007] for $\alpha = 1$, and it is easy to get the same result for $\alpha < 1$ from that argument. We take advantage of (26) to extend the range of α up to $\frac{4}{3}$. In the remainder of this section, we will essentially follow Tao's argument to prove Theorem 4.1 using (26) and sketch the proof of $\widetilde{H}^2_x \times H^1_x$ scattering. We will skip the argument providing an explicit $\widetilde{H}^2_x \times H^1_x$ universal bound here; see [Tao 2007] for details.

We will use a well-known global continuation result (for a proof see [Sogge 1995], for example).

Theorem 4.3 (classical existence theory). *Let* $u : [0,T] \times \mathbb{R}^3 \to \mathbb{C}$ *be a classical solution*[5] *to* (61) *satisfying*

$$\|u\|_{L^\infty_t L^\infty_x([0,T]\times\mathbb{R}^3)} < \infty.$$

Then there is $\delta > 0$ *such that one can extend the solution* u *to* $[0, T+\delta] \times \mathbb{R}^3$.

Proof of Theorem 4.1. By time reversal symmetry, it suffices to consider the global existence and scattering theory of u on $\mathbb{R}_+ \times \mathbb{R}^3$.

---

[3] The kinetic energy bound can only provide $\int_I \int_{\mathbb{R}^3} |u|^8 \log^\alpha(2+|u|^2)\,dx\,dt \lesssim E^{3/2}$.

[4] The definition of $\widetilde{H}^2_x \times H^1_x$ scattering for the solution u is similar to Definition 2.1, but the $\dot{H}^1_x \times L^2_x$-norm is replaced by the $\widetilde{H}^2_x \times H^1_x$-norm.

[5] We call u a classical solution to (1) if u solves (1) and is smooth and compactly supported for each time.

By the Sobolev embedding theorem, for a classical solution u to (61) on $[0, T] \times \mathbb{R}^3$, we have

$$\|u\|_{L_t^\infty L_x^\infty([0,T]\times\mathbb{R}^3)} \lesssim \sum_{j=1}^{2} \|\nabla_x^j u\|_{L_t^\infty L_x^2([0,T]\times\mathbb{R}^3)}. \tag{62}$$

Hence, applying classical existence theory (Theorem 4.3), in order to show global existence, it suffices to prove that for any fixed $0 < T \leq \infty$, we have

$$\sum_{j=1}^{2} \|\nabla_x^j u\|_{L_t^\infty L_x^2([0,T]\times\mathbb{R}^3)} < \infty,$$

provided that u is the classical solution to (61) on $[0, T] \times \mathbb{R}^3$.

Let $I = [a, b] \subseteq [0, T]$ be any interval. We define

$$\begin{aligned}
M_I &:= \int_I \int_{\mathbb{R}^3} |u(t,x)|^8 \log^{5\alpha/4}(2+|u(t,x)|^2)\, dx\, dt,\\
N_I &:= \sum_{j=0}^{1} \|\nabla_x^j u\|_{L_t^2 L_x^\infty(I\times\mathbb{R}^3)} + \|\nabla_{t,x}\nabla_x^j u\|_{L_t^\infty L_x^2(I\times\mathbb{R}^3)},\\
D_I &:= \|\nabla_{t,x} u(a)\|_{H_x^1(\mathbb{R}^3)}.
\end{aligned}$$

In addition, we set $D = \|\nabla_{t,x} u(0)\|_{H_x^1(\mathbb{R}^3)}$.

From the Strichartz inequality, Hölder's inequality, and (62), we have

$$\begin{aligned}
N_I &\leq C\|\nabla_{t,x} u(a)\|_{H_x^1(\mathbb{R}^3)} + C\sum_{j=0}^{1} \|\nabla_x^j(|u|^4 u \log^\alpha(2+|u|^2))\|_{L_t^1 L_x^2(I\times\mathbb{R}^3)}\\
&\leq CD_I + C\sum_{j=0}^{1} \|\,|u|^4 |\nabla_x^j u| \log(2+|u|^2)\|_{L_t^1 L_x^2(I\times\mathbb{R}^3)}\\
&\leq CD_I + C\|\,|u|^4 \log^{5\alpha/8}(2+|u|^2)\|_{L_t^2 L_x^2(I\times\mathbb{R}^3)} \left(\sum_{j=0}^{1} \|\nabla_x^j u\|_{L_t^2 L_x^\infty(I\times\mathbb{R}^3)} \|\log^{3\alpha/8}(2+|u|^2)\|_{L_t^\infty L_x^\infty(I\times\mathbb{R}^3)}\right)\\
&\leq CD_I + C\|u \log^{5\alpha/32}(2+|u|^2)\|_{L_t^8 L_x^8(I\times\mathbb{R}^3)}^4 \left(\sum_{j=0}^{1} \|\nabla_x^j u\|_{L_t^2 L_x^\infty(I\times\mathbb{R}^3)} \|\log(2+|u|^2)\|_{L_t^\infty L_x^\infty(I\times\mathbb{R}^3)}^{3\alpha/8}\right)\\
&\leq CD_I + CM_I^{1/2} N_I \log^{3\alpha/8}(2+\|u\|_{L_t^\infty L_x^2(I\times\mathbb{R}^3)}^2)\\
&\leq CD_I + CM_I^{1/2} N_I \log^{1/2}(2+N_I^2).
\end{aligned}$$

From the result in [Tao 2007, Corollary 3.2], for any $\epsilon_0 > 0$,

$$\sum_{i=0}^{k} \frac{\epsilon_0}{\log(2+(2C)^i D)} \to \infty \quad \text{as } k \to \infty.$$

Hence, for any fixed ϵ_0, the finiteness of $M_{[0,T]}$ from (26) implies that we can choose finitely many numbers $0 = T_0 < T_1 < \cdots < T_K < T_{K+1} = T$, with K depending on D, E, and ϵ_0, such that

$$M_i := \int_{T_i}^{T_{i+1}} \int_{\mathbb{R}^3} |u(t,x)|^8 \log^{5\alpha/4}(2+|u(t,x)|^2)\, dx\, dt = \frac{\epsilon_0}{\log(2+(2C)^i D)}$$

for $0 \le i \le K-1$ and $M_K \le \epsilon_0 / \log(2+(2C)^K D)$.

Choosing $\epsilon_0 = 1/(100C)^2$, by iteration and continuity arguments, we claim that $N_{[T_i,T_{i+1}]} < (2C)^{i+1} D$ for $0 \le i \le K$.[6] Indeed, assume that this claim is false for some $i = j$. Then there exists $t_0 \in (T_j, T_{j+1})$ such that $N_{[T_j,t_0]} = (2C)^{j+1} D$. We have

$$\begin{aligned}(2C)^{j+1} D &\le C(2C)^j D + C M_j^{1/2} N_{[T_j,t_0]} \log^{1/2}(2+N_{[T_j,t_0]}^2)\\ &\le \tfrac{1}{2}(2C)^{j+1} D + \frac{\log^{1/2}(2+(2C)^{j+1} D)}{100 \log^{1/2}(2+(2C)^j D)} \times (2C)^{j+1} D\\ &\le \tfrac{3}{4}(2C)^{j+1} D.\end{aligned}$$

Thus the claim is proved by contradiction. This implies

$$\sum_{j=1}^{2} \|\nabla_x^j u\|_{L_t^\infty L_x^2([0,T]\times\mathbb{R}^3)} \le N_{[0,T]} \le \sum_{i=0}^{K} N_{[T_i,T_{i+1}]} < \sum_{i=0}^{K} (2C)^{i+1} D < \infty.$$

The universal bound only depends on D and E[7], indicating the global existence.

Now we sketch the proof of $\widetilde{H}_x^2 \times H_x^1$ scattering. From a similar argument as the one discussed in Remark 2.2, in order to prove $\widetilde{H}_x^2(\mathbb{R}^3) \times H_x^1(\mathbb{R}^3)$ scattering, it suffices to show that

$$\| |u|^4 u \log^\alpha(2+|u|^2) \|_{L_t^1 H_x^1(\mathbb{R}_+\times\mathbb{R}^3)} < \infty. \tag{63}$$

By the above discussion, the universal bound is independent of T. Hence we have $N_{\mathbb{R}_+} < \infty$. By Hölder's inequality,

$$\| |u|^4 u \log^\alpha(2+|u|^2) \|_{L_t^1 H_x^1(\mathbb{R}_+\times\mathbb{R}^3)} \lesssim M_{\mathbb{R}_+}^{1/2} N_{\mathbb{R}_+} \log^{1/2}(2+N_{\mathbb{R}_+}^2) < \infty. \qquad \square$$

Appendix: Proof of (30)

Since (u_0, u_1) lies in $X_1 \times Y_1$, defined in (9), we have

$$\|u_0\|_{X_1}^2 \ge \int_{\mathbb{R}^3} |\nabla u_0|^2 \log^{2\gamma}(1+|x|)\, dx \gtrsim (\log^{2\gamma} s) \int_{|x|>s} |\nabla u_0|^2\, dx.$$

Hence

$$\int_{|x|>s} |\nabla u_0|^2\, dx \lesssim \frac{\|u_0\|_{X_1}^2}{\log^{2\gamma} s}. \tag{64}$$

Similarly,

$$\int_{|x|>s} |u_1|^2\, dx \lesssim \frac{\|u_1\|_{Y_1}^2}{\log^{2\gamma} s}. \tag{65}$$

[6] See the similar arguments in Lemma 3.5 and Corollary 3.8 or Proposition 3.1 in [Tao 2007].

[7] In fact, from corollary 3.2 in [Tao 2007], we have $N_{\mathbb{R}_+} \lesssim (2+D)^{(2+D)^{O(E)}}$.

Now, consider

$$\begin{aligned}\int_{|x|>s} F(u_0(x))\,dx &= \int_{\{|x|>s\}\cap\{|u_0|<1/3\}} F(u_0(x))\,dx + \int_{\{|x|>s\}\cap\{|u_0|\geq 1/3\}} F(u_0(x))\,dx \\ &\lesssim \int_{\{|x|>s\}\cap\{|u_0|<1/3\}} |u_0|^6(-\log|u_0|)\,dx + \int_{\{|x|>s\}\cap\{|u_0|\geq 1/3\}} |u_0|^6\,dx =: I + II.\end{aligned}$$

Let

$$\begin{aligned}I &= \int_{\{|x|>s\}\cap\{|u_0(x)|<1/|x|^{2/3}\}} |u_0|^6(-\log|u_0|)\,dx + \int_{\{|x|>s\}\cap\{1/|x|^{2/3}\leq|u_0(x)|\leq 1/3\}} |u_0|^6(-\log|u_0|)\,dx \\ &=: I_1 + I_2.\end{aligned}$$

When s is sufficiently large,

$$\begin{aligned}I_1 &\lesssim \int_{\{|x|>s\}\cap\{|u_0|<1/|x|^{2/3}\}} |u_0|^{11/2}\Big(\sup_{|u_0|<s^{-2/3}} |u_0|^{1/2}(-\log|u_0|)\Big)\,dx \\ &\lesssim \int_{|x|>s} |x|^{-11/3}\,dx \lesssim s^{-2/3} \lesssim \frac{1}{\log^{2\gamma} s}.\end{aligned}\tag{66}$$

Now we aim to prove that $I_2 + II \lesssim \dfrac{1}{\log^{2\gamma} s}$ for s sufficiently large. For $\alpha\in\mathbb{R}$, define

$$Q(\alpha) := \int_{\mathbb{R}^3}\left|\frac{u_0\log^{\alpha}(2+|x|)}{2+|x|}\right|^2 dx.$$

We claim that

$$Q(\alpha) \leq C(\|u_0\|_{X_1}, E, \alpha) \quad \text{for } \alpha\leq\gamma, \tag{67}$$

where E is the energy. Indeed, if $\alpha\leq 0$, by Hölder's inequality and Hardy's inequality,

$$\begin{aligned}Q(\alpha) &= \int_{|x|<3}\left|\frac{u_0\log^{\alpha}(2+|x|)}{2+|x|}\right|^2 dx + \int_{|x|\geq 3}\left|\frac{u_0\log^{\alpha}(2+|x|)}{2+|x|}\right|^2 dx \\ &\lesssim_\alpha \int_{|x|<3}|u_0|^2\,dx + \int_{\mathbb{R}^3}\left|\frac{u_0}{|x|}\right|^2 dx \lesssim \|u_0\|^2_{\dot H^1_x(\mathbb{R}^3)} + \left(\int_{\mathbb{R}^3} F(u_0)\,dx\right)^{1/3} \leq C(E,\alpha).\end{aligned}\tag{68}$$

Again, if $0<\alpha\leq\gamma$,

$$\begin{aligned}Q(\alpha) &= \int_{\mathbb{R}^3}\left|\frac{u_0\log^{\alpha}(2+|x|)}{2+|x|}\right|^2 dx \lesssim_\alpha \int_{|x|<3}|u_0|^2\,dx + \int_{\mathbb{R}^3}\left|\frac{u_0\log^{\alpha}(2+|x|)}{|x|}\right|^2 dx \\ &\lesssim \left(\int_{|x|<3}|u_0|^6\,dx\right)^{1/3} + \int_{\mathbb{R}^3}|\nabla(u_0\log^{\alpha}(2+|x|))|^2\,dx \\ &\lesssim_\alpha \left(\int_{\mathbb{R}^3} F(u)\,dx\right)^{1/3} + \int_{\mathbb{R}^3}|\nabla u_0\log^{\alpha}(2+|x|)|^2\,dx + \int_{\mathbb{R}^3}\left|\frac{u_0\log^{\alpha-1}(2+|x|)}{2+|x|}\right|^2 dx \\ &\lesssim E^{1/3} + \int_{|x|<3}|\nabla u_0|^2\,dx + \int_{|x|\geq 3}|\nabla u_0\log^{\gamma}(1+|x|)|^2\,dx + Q(\alpha-1) \\ &\lesssim E^{1/3} + E + \|u_0\|_{X_1} + Q(\alpha-1).\end{aligned}$$

By an inductive argument and (68), the claim is proved.

Fix $s \gg 1$. Let χ be the smooth radial function which equals 1 on $\{|x| > s\}$, 0 on $\{|x| < s/2\}$, $0 \le \chi \le 1$ and $|\nabla\chi| \lesssim 1/s$. Then we have $|\nabla\chi| \lesssim 1/|x|$. By the Sobolev embedding theorem and Hardy's inequality,

$$\begin{aligned}
&\log^{6\gamma} s \int_{|x|>s} |u_0|^6\,dx \\
&\le \int_{|x|>s} |u_0|^6 \log^{6\gamma}(|x|)\,dx \le \int_{\mathbb{R}^3} (\chi |u_0| \log^{\gamma}(2+|x|))^6\,dx \lesssim \left(\int_{\mathbb{R}^3} \left|\nabla\big(\chi u_0 \log^{\gamma}(2+|x|)\big)\right|^2 dx \right)^3 \\
&\lesssim_\gamma \left(\int_{\mathbb{R}^3} |\nabla\chi u_0 \log^{\gamma}(2+|x|)|^2\,dx + \int_{\mathbb{R}^3} |\chi \nabla u_0 \log^{\gamma}(2+|x|)|^2\,dx + \int_{\mathbb{R}^3} \left| \frac{\chi u_0 \log^{\gamma-1}(2+|x|)}{2+|x|} \right|^2 dx \right)^3 \\
&=: (J_1 + J_2 + J_3)^3.
\end{aligned}$$

We can compute that

$$\begin{aligned}
&J_2 \lesssim_\gamma \int_{\mathbb{R}^3} \left|\nabla u_0 \log^{\gamma}(1+|x|)\right|^2 dx + \int_{|x|<3} |\nabla u_0|^2\,dx \le \|u_0\|_{X_1}^2 + E, \\
&J_3 \lesssim C(\|u_0\|_{X_1}, E, \gamma), \qquad \text{by (67)}.
\end{aligned}$$

Since $\nabla\chi \lesssim 1/|x|$,

$$\begin{aligned}
J_1 &\lesssim \int_{|x|>s/2} \left| \frac{u_0 \log^{\gamma}(2+|x|)}{|x|} \right|^2 dx \lesssim \int_{|x|>s/2} \left| \frac{u_0 \log^{\gamma}(2+|x|)}{2+|x|} \right|^2 dx \\
&\lesssim C(\|u_0\|_{X_1}, E, \gamma).
\end{aligned}$$

Hence $\log^{6\gamma} s \int_{|x|>s} |u_0|^6\,dx \le C(\|u_0\|_{X_1}, E, \gamma)$ for sufficiently large s. Then we deduce

$$II \le \int_{|x|>s} |u_0|^6\,dx \lesssim \frac{1}{\log^{6\gamma} s} \le \frac{1}{\log^{2\gamma} s}. \tag{69}$$

Similarly,

$$\begin{aligned}
\log^{6\gamma-1} s \int_{\{|x|>s\}\cap\{1/|x|^{2/3} \le |u_0| \le 1/3\}} |u_0|^6(-\log|u_0|)\,dx &\lesssim \log^{6\gamma-1} s \int_{|x|>s} |u_0|^6 \log(|x|)\,dx \\
&\lesssim \int_{|x|>s} |u_0|^6 \log^{6\gamma}(|x|)\,dx \lesssim C(\|u_0\|_{X_1}, E, \gamma).
\end{aligned}$$

Therefore,

$$I_2 \lesssim \frac{1}{\log^{6\gamma-1} s} \le \frac{1}{\log^{2\gamma} s}. \tag{70}$$

Combining (64), (65), (66), (69), and (70), we obtain (30).

Acknowledgements

The author thanks Markus Keel and Terence Tao for illuminating conversations about the problems considered here.

References

[Bahouri and Gérard 1999] H. Bahouri and P. Gérard, "High frequency approximation of solutions to critical nonlinear wave equations", *Amer. J. Math.* **121**:1 (1999), 131–175. MR 2000i:35123 Zbl 0919.35089

[Bahouri and Shatah 1998] H. Bahouri and J. Shatah, "Decay estimates for the critical semilinear wave equation", *Ann. Inst. H. Poincaré Anal. Non Linéaire* **15**:6 (1998), 783–789. MR 99h:35136 Zbl 0924.35084

[Brenner and von Wahl 1981] P. Brenner and W. von Wahl, "Global classical solutions of nonlinear wave equations", *Math. Z.* **176**:1 (1981), 87–121. MR 82m:35103 Zbl 0457.35059

[Ginibre and Velo 1985] J. Ginibre and G. Velo, "The global Cauchy problem for the nonlinear Klein–Gordon equation", *Math. Z.* **189**:4 (1985), 487–505. MR 86f:35149 Zbl 0549.35108

[Ginibre and Velo 1987] J. Ginibre and G. Velo, "Conformal invariance and time decay for nonlinear wave equations. I, II", *Ann. Inst. H. Poincaré Phys. Théor.* **47**:3 (1987), 221–261, 263–276. MR 89e:35021 Zbl 0644.35067

[Ginibre and Velo 1989] J. Ginibre and G. Velo, "Scattering theory in the energy space for a class of nonlinear wave equations", *Comm. Math. Phys.* **123**:4 (1989), 535–573. MR 90i:35172 Zbl 0698.35112

[Ginibre and Velo 1995] J. Ginibre and G. Velo, "Generalized Strichartz inequalities for the wave equation", *J. Funct. Anal.* **133**:1 (1995), 50–68. MR 97a:46047 Zbl 0849.35064

[Ginibre et al. 1992] J. Ginibre, A. Soffer, and G. Velo, "The global Cauchy problem for the critical nonlinear wave equation", *J. Funct. Anal.* **110**:1 (1992), 96–130. MR 94d:35105 Zbl 0813.35054

[Grillakis 1990] M. G. Grillakis, "Regularity and asymptotic behaviour of the wave equation with a critical nonlinearity", *Ann. of Math.* (2) **132**:3 (1990), 485–509. MR 92c:35080 Zbl 0736.35067

[Grillakis 1992] M. G. Grillakis, "Regularity for the wave equation with a critical nonlinearity", *Comm. Pure Appl. Math.* **45**:6 (1992), 749–774. MR 93e:35073 Zbl 0785.35065

[Hidano 2001] K. Hidano, "Scattering problem for the nonlinear wave equation in the finite energy and conformal charge space", *J. Funct. Anal.* **187**:2 (2001), 274–307. MR 2002k:35242 Zbl 1005.35066

[Kapitanski 1989] L. V. Kapitanski, "Some generalizations of the Strichartz–Brenner inequality", *Algebra i Analiz* **1**:3 (1989), 127–159. In Russian. translated in *Leningr. Math. J.* **1**:3 (1990), 693–726. MR 90h:46063 Zbl 0732.35118

[Kapitanski 1994] L. Kapitanski, "Global and unique weak solutions of nonlinear wave equations", *Math. Res. Lett.* **1**:2 (1994), 211–223. MR 95f:35158 Zbl 0841.35067

[Keel and Tao 1998] M. Keel and T. Tao, "Endpoint Strichartz estimates", *Amer. J. Math.* **120**:5 (1998), 955–980. MR 2000d:35018 Zbl 0922.35028

[Klainerman and Machedon 1993] S. Klainerman and M. Machedon, "Space-time estimates for null forms and the local existence theorem", *Comm. Pure Appl. Math.* **46**:9 (1993), 1221–1268. MR 94h:35137 Zbl 0803.35095

[Lindblad and Sogge 1995] H. Lindblad and C. D. Sogge, "On existence and scattering with minimal regularity for semilinear wave equations", *J. Funct. Anal.* **130**:2 (1995), 357–426. MR 96i:35087 Zbl 0846.35085

[Morawetz 1968] C. S. Morawetz, "Time decay for the nonlinear Klein–Gordon equations", *Proc. Roy. Soc. Ser. A* **306** (1968), 291–296. MR 38 #2455 Zbl 0157.41502

[Nakanishi 1999] K. Nakanishi, "Unique global existence and asymptotic behaviour of solutions for wave equations with non-coercive critical nonlinearity", *Comm. Partial Differential Equations* **24**:1-2 (1999), 185–221. MR 2000f:35104 Zbl 0926.35022

[Roy 2009] T. Roy, "Global existence of smooth solutions of a 3D log-log energy-supercritical wave equation", *Anal. PDE* **2**:3 (2009), 261–280. MR 2603799 Zbl 1195.35222

[Shatah and Struwe 1993] J. Shatah and M. Struwe, "Regularity results for nonlinear wave equations", *Ann. of Math.* (2) **138**:3 (1993), 503–518. MR 95f:35164 Zbl 0836.35096

[Shatah and Struwe 1994] J. Shatah and M. Struwe, "Well-posedness in the energy space for semilinear wave equations with critical growth", *Internat. Math. Res. Notices* 7 (1994), 303–309. MR 95e:35132 Zbl 0830.35086

[Sogge 1995] C. D. Sogge, *Lectures on nonlinear wave equations*, Monographs in Analysis **2**, International Press, Boston, 1995. MR 2000g:35153 Zbl 1089.35500

[Strichartz 1977] R. S. Strichartz, "Restrictions of Fourier transforms to quadratic surfaces and decay of solutions of wave equations", *Duke Math. J.* **44**:3 (1977), 705–714. MR 58 #23577 Zbl 0372.35001

[Struwe 1988] M. Struwe, "Globally regular solutions to the u^5 Klein–Gordon equation", *Ann. Scuola Norm. Sup. Pisa Cl. Sci.* (4) **15**:3 (1988), 495–513 (1989). MR 90j:35142 Zbl 0728.35072

[Tao 2007] T. Tao, "Global regularity for a logarithmically supercritical defocusing nonlinear wave equation for spherically symmetric data", *J. Hyperbolic Differ. Equ.* **4**:2 (2007), 259–265. MR 2009b:35294 Zbl 1124.35043

Received 27 May 2011. Revised 22 Nov 2011. Accepted 20 Mar 2012.

HSI-WEI SHIH: shihx029@umn.edu
School of Mathematics, University of Minnesota, 127 Vincent Hall, 206 Church St. SE, Minneapolis, MN 55455, United States
http://math.umn.edu/~shihx029/

ANALYSIS AND PDE
Vol. 6, No. 1, 2013
dx.doi.org/10.2140/apde.2013.6.25

LOCALISATION AND COMPACTNESS PROPERTIES OF THE NAVIER–STOKES GLOBAL REGULARITY PROBLEM

TERENCE TAO

In this paper we establish a number of implications between various qualitative and quantitative versions of the global regularity problem for the Navier–Stokes equations in the periodic, smooth finite energy, smooth H^1, Schwartz, and mild H^1 categories, and with or without a forcing term. In particular, we show that if one has global well-posedness in H^1 for the periodic Navier–Stokes problem with a forcing term, then one can obtain global regularity both for periodic and for Schwartz initial data (thus yielding a positive answer to both official formulations of the problem for the Clay Millennium Prize), and can also obtain global almost smooth solutions from smooth H^1 data or smooth finite energy data, although we show in this category that fully smooth solutions are not always possible. Our main new tools are localised energy and enstrophy estimates to the Navier–Stokes equation that are applicable for large data or long times, and which may be of independent interest.

1. Introduction

The purpose of this paper is to establish some implications between various formulations of the global regularity problem (either with or without a forcing term) for the Navier–Stokes system of equations, including the four formulations appearing in the Clay Millennium Prize formulation [Fefferman 2006] of the problem, and in particular to isolate a single formulation that implies these four formulations, as well as several other natural versions of the problem. In the course of doing so, we also establish some new local energy and local enstrophy estimates which seem to be of independent interest.

To describe these various formulations, we must first define properly the concept of a solution to the Navier–Stokes problem. We will need to study a number of different types of solutions, including periodic solutions, finite energy solutions, H^1 solutions, and smooth solutions; we will also consider a forcing term f in addition to the initial data u_0. We begin in the classical regime of smooth solutions. Note that even within the category of smooth solutions, there is some choice in what decay hypotheses to place on the initial data and solution; for instance, one can require that the initial velocity u_0 be Schwartz class, or merely smooth with finite energy. Intermediate between these two will be data which is smooth and in H^1.

More precisely, we define:

The author is supported by NSF Research Award CCF-0649473, the NSF Waterman Award and a grant from the MacArthur Foundation.
MSC2010: 35Q30, 76D05, 76N10.
Keywords: Navier–Stokes equation, global regularity.

Definition 1.1 (Smooth solutions to the Navier–Stokes system). A *smooth set of data* for the Navier–Stokes system up to time T is a triplet (u_0, f, T), where $0 < T < \infty$ is a time, the initial velocity vector field $u_0 : \mathbb{R}^3 \to \mathbb{R}^3$ and the forcing term $f : [0, T] \times \mathbb{R}^3 \to \mathbb{R}^3$ are assumed to be smooth on $\mathbb{R}^3$ and $[0, T] \times \mathbb{R}^3$ respectively (thus, u_0 is infinitely differentiable in space, and f is infinitely differentiable in space-time), and u_0 is furthermore required to be divergence-free:

$$\nabla \cdot u_0 = 0. \tag{1}$$

If $f = 0$, we say that the data is *homogeneous*.

The *total energy* $E(u_0, f, T)$ of a smooth set of data (u_0, f, T) is defined by the quantity[1]

$$E(u_0, f, T) := \tfrac{1}{2}\big(\|u_0\|_{L^2_x(\mathbb{R}^3)} + \|f\|_{L^1_t L^2_x([0,T]\times\mathbb{R}^3)}\big)^2, \tag{2}$$

and (u_0, f, T) is said to have *finite energy* if $E(u_0, f, T) < \infty$. We define the H^1 *norm* $\mathcal{H}^1(u_0, f, T)$ of the data to be the quantity

$$\mathcal{H}^1(u_0, f, T) := \|u_0\|_{H^1_x(\mathbb{R}^3)} + \|f\|_{L^\infty_t H^1_x(\mathbb{R}^3)} < \infty,$$

and say that (u_0, f, T) is H^1 if $\mathcal{H}^1(u_0, f, T) < \infty$; note that the H^1 regularity is essentially one derivative higher than the energy regularity, which is at the level of L^2, and instead matches the regularity of the *initial enstrophy*

$$\tfrac{1}{2}\int_{\mathbb{R}^3} |\omega_0(t, x)|^2 \, dx,$$

where $\omega_0 := \nabla \times u_0$ is the initial vorticity. We say that a smooth set of data (u_0, f, T) is *Schwartz* if, for all integers $\alpha, m, k \geq 0$, one has

$$\sup_{x\in\mathbb{R}^3} (1 + |x|)^k \big|\nabla_x^\alpha u_0(x)\big| < \infty$$

and

$$\sup_{(t,x)\in[0,T]\times\mathbb{R}^3} (1 + |x|)^k \big|\nabla_x^\alpha \partial_t^m f(x)\big| < \infty.$$

Thus, for instance, the Schwartz property implies H^1, which in turn implies finite energy. We also say that (u_0, f, T) is *periodic* with some period $L > 0$ if one has $u_0(x + Lk) = u_0(x)$ and $f(t, x + Lk) = f(t, x)$ for all $t \in [0, T]$, $x \in \mathbb{R}^3$, and $k \in \mathbb{Z}^3$. Of course, periodicity is incompatible with the Schwartz, H^1, and finite energy properties, unless the data is zero. To emphasise the periodicity, we will sometimes write a periodic set of data (u_0, f, T) as (u_0, f, T, L).

A *smooth solution to the Navier–Stokes system*, or a *smooth solution*, is a quintuplet (u, p, u_0, f, T), where (u_0, f, T) is a smooth set of data, and the velocity vector field $u : [0, T] \times \mathbb{R}^3 \to \mathbb{R}^3$ and pressure field $p : [0, T] \times \mathbb{R}^3 \to \mathbb{R}$ are smooth functions on $[0, T] \times \mathbb{R}^3$ that obey the Navier–Stokes equation

$$\partial_t u + (u \cdot \nabla)u = \Delta u - \nabla p + f \tag{3}$$

and the incompressibility property

$$\nabla \cdot u = 0 \tag{4}$$

[1] We will review our notation for space-time norms such as $L^p_t L^q_x$, together with sundry other notation, in Section 2.

on all of $[0, T] \times \mathbb{R}^3$, and also the initial condition

$$u(0, x) = u_0(x) \tag{5}$$

for all $x \in \mathbb{R}^3$. We say that a smooth solution (u, p, u_0, f, T) has *finite energy* if the associated data (u_0, f, T) has finite energy, and in addition one has[2]

$$\|u\|_{L^\infty_t L^2_x([0,T]\times\mathbb{R}^3)} < \infty. \tag{6}$$

Similarly, we say that (u, p, u_0, f, T) is H^1 if the associated data (u_0, f, T) is H^1, and in addition one has

$$\|u\|_{L^\infty_t H^1_x([0,T]\times\mathbb{R}^3)} + \|u\|_{L^2_t H^2_x([0,T]\times\mathbb{R}^3)} < \infty. \tag{7}$$

We say instead that a smooth solution (u, p, u_0, f, T) is *periodic* with period $L > 0$ if the associated data $(u_0, f, T) = (u_0, f, T, L)$ is periodic with period L, and if $u(t, x + Lk) = u(t, x)$ for all $t \in [0, T]$, $x \in \mathbb{R}^3$, and $k \in \mathbb{Z}^3$. (Following [Fefferman 2006], however, we will not initially directly require any periodicity properties on the pressure.) As before, we will sometimes write a periodic solution (u, p, u_0, f, T) as (u, p, u_0, f, T, L) to emphasise the periodicity.

We will sometimes abuse notation and refer to a solution (u, p, u_0, f, T) simply as (u, p) or even u. Similarly, we will sometimes abbreviate a set of data (u_0, f, T) as (u_0, f) or even u_0 (in the homogeneous case $f = 0$).

Remark 1.2. In [Fefferman 2006], one considered[3] smooth finite energy solutions associated to Schwartz data, as well as periodic smooth solutions associated to periodic smooth data. In the latter case, one can of course normalise the period L to equal 1 by a simple scaling argument. In this paper we will be focussed on the case when the data (u_0, f, T) is large, although we will not study the asymptotic regime when $T \to \infty$.

We recall the two standard *global regularity* conjectures for the Navier–Stokes equation, using the formulation in [Fefferman 2006]:

Conjecture 1.3 (Global regularity for homogeneous Schwartz data). Let $(u_0, 0, T)$ be a homogeneous Schwartz set of data. Then there exists a smooth finite energy solution $(u, p, u_0, 0, T)$ with the indicated data.

Conjecture 1.4 (Global regularity for homogeneous periodic data). Let $(u_0, 0, T)$ be a smooth homogeneous periodic set of data. Then there exists a smooth periodic solution $(u, p, u_0, 0, T)$ with the indicated data.

---

[2]Following [Fefferman 2006], we omit the finite energy dissipation condition $\nabla u \in L^2_t L^2_x([0, T] \times \mathbb{R}^3)$ that often appears in the literature, particularly when discussing Leray–Hopf weak solutions. However, it turns out that this condition is actually automatic from (6) and smoothness; see Lemma 8.1. Similarly, from Corollary 11.1 we shall see that the $L^2_t H^2_x$ condition in (7) is in fact redundant.

[3]The viscosity parameter ν was not normalised in [Fefferman 2006] to equal 1, as we are doing here, but one can easily reduce to the $\nu = 1$ case by a simple rescaling.

In view of these conjectures, one can naturally try to extend them to the inhomogeneous case as follows:

Conjecture 1.5 (Global regularity for Schwartz data). Let (u_0, f, T) be a Schwartz set of data. Then there exists a smooth finite energy solution (u, p, u_0, f, T) with the indicated data.

Conjecture 1.6 (Global regularity for periodic data). Let (u_0, f, T) be a smooth periodic set of data. Then there exists a smooth periodic solution (u, p, u_0, f, T) with the indicated data.

As described in [Fefferman 2006], a positive answer to either Conjecture 1.3 or Conjecture 1.4, or a negative answer to Conjecture 1.5 or Conjecture 1.6, would qualify for the Clay Millennium Prize.

However, Conjecture 1.6 is not quite the "right" extension of Conjecture 1.4 to the inhomogeneous setting, and needs to be corrected slightly. This is because there is a technical quirk in the inhomogeneous periodic problem as formulated in Conjecture 1.6, due to the fact that the pressure p is not required to be periodic. This opens up a Galilean invariance in the problem which allows one to homogenise away the role of the forcing term. More precisely, we have:

Proposition 1.7 (Elimination of forcing term). *Conjecture 1.6 is equivalent to Conjecture 1.4.*

We establish this fact in Section 6. We remark that this is the only implication we know of that can deduce a global regularity result for the inhomogeneous Navier–Stokes problem from a global regularity result for the homogeneous Navier–Stokes problem.

Proposition 1.7 exploits the technical loophole of nonperiodic pressure. The same loophole can also be used to easily demonstrate failure of uniqueness for the periodic Navier–Stokes problem (although this can also be done by the much simpler expedient of noting that one can adjust the pressure by an arbitrary constant without affecting (3)). This suggests that in the nonhomogeneous case $f \neq 0$, one needs an additional normalisation to "fix" the periodic Navier–Stokes problem to avoid such loopholes. This can be done in a standard way, as follows. If one takes the divergence of (3) and uses the incompressibility (4), one sees that

$$\Delta p = -\partial_i \partial_j (u_i u_j) + \nabla \cdot f, \tag{8}$$

where we use the usual summation conventions. If (u, p, u_0, f, T) is a smooth periodic solution, then the right-hand side of (8) is smooth and periodic and has mean zero. From Fourier analysis, we see that given any smooth periodic mean-zero function F, there is a unique smooth periodic mean-zero function $\Delta^{-1} F$ with Laplacian equal to F. We then say that the periodic smooth solution (u, p, u_0, f, T) has *normalised pressure* if one has[4]

$$p = -\Delta^{-1} \partial_i \partial_j (u_i u_j) + \Delta^{-1} \nabla \cdot f. \tag{9}$$

We remark that this normalised pressure condition can also be imposed for smooth finite energy solutions (because $\partial_i \partial_j (u_i u_j)$ is a second derivative of an $L^1_x(\mathbb{R}^3)$ function, and $\nabla \cdot f$ is the first derivative of an $L^2_x(\mathbb{R}^3)$ function), but it will turn out that normalised pressure is essentially automatic in that setting anyway; see Lemma 4.1.

[4]Up to the harmless freedom to add a constant to p, this normalisation is equivalent to requiring that the pressure be periodic with the same period as the solution u.

It is well known that once one imposes the normalised pressure condition, the periodic Navier–Stokes problem becomes locally well-posed in the smooth category (in particular, smooth solutions are now unique, and exist for sufficiently short times from any given smooth data); see Theorem 5.1. Related to this, the Galilean invariance trick that allows one to artificially homogenise the forcing term f is no longer available. We can then pose a "repaired" version of Conjecture 1.6:

Conjecture 1.8 (Global regularity for periodic data with normalised pressure). Let (u_0, f, T) be a smooth periodic set of data. Then there exists a smooth periodic solution (u, p, u_0, f, T) with the indicated data and with normalised pressure.

It is easy to see that the homogeneous case $f = 0$ of Conjecture 1.8 is equivalent to Conjecture 1.4; see, for example, Lemma 4.1 below.

We now leave the category of classical (smooth) solutions for now, and turn instead to the category of *periodic H^1 mild solutions* (u, p, u_0, f, T, L). By definition, these are functions

$$u, f : [0, T] \times \mathbb{R}^3/L\mathbb{Z}^3 \to \mathbb{R}^3, \quad p : [0, T] \times \mathbb{R}^3/L\mathbb{Z}^3 \to \mathbb{R}, \quad u_0 : \mathbb{R}^3/L\mathbb{Z}^3 \to \mathbb{R}^3,$$

with $0 < T, L < \infty$, obeying the regularity hypotheses

$$\begin{aligned} &u_0 \in H^1_x(\mathbb{R}^3/L\mathbb{Z}^3), \\ &f \in L^\infty_t H^1_x([0, T] \times (\mathbb{R}^3/L\mathbb{Z}^3)), \\ &u \in L^\infty_t H^1_x \cap L^2_t H^2_x([0, T] \times (\mathbb{R}^3/L\mathbb{Z}^3)), \end{aligned}$$

with p being given by (9), which obey the divergence-free conditions (4), (1) and obey the integral form

$$u(t) = e^{t\Delta} u_0 + \int_0^t e^{(t-t')\Delta}\big(-(u \cdot \nabla)u - \nabla p + f\big)(t')\, dt' \tag{10}$$

of the Navier–Stokes equation (3) with initial condition (5); using the Leray projection P onto divergence-free vector fields, we may also express (19) equivalently as

$$u(t) = e^{t\Delta} u_0 + \int_0^t e^{(t-t')\Delta}(PB(u, u) + Pf)(t')\, dt', \tag{11}$$

where $B(u, v)$ is the symmetric bilinear form

$$B(u, v)_i := -\tfrac{1}{2}\partial_j(u_i v_j + u_j v_i). \tag{12}$$

Similarly, we define *periodic H^1 data* to be a quadruplet (u_0, f, T, L) whose H^1 norm

$$\mathcal{H}^1(u_0, f, T, L) := \|u_0\|_{H^1_x((\mathbb{R}^3/L\mathbb{Z}^3))} + \|f\|_{L^\infty_t H^1_x((\mathbb{R}^3/L\mathbb{Z}^3))}$$

is finite, with u_0 divergence-free.

Note from Duhamel's formula (20) that every smooth periodic solution with normalised pressure is automatically a periodic H^1 mild solution.

As we will recall in Theorem 5.1 below, the Navier–Stokes equation is locally well-posed in the periodic H^1 category. We can then formulate a global well-posedness conjecture in this category:

Conjecture 1.9 (Global well-posedness in periodic H^1). Let (u_0, f, T, L) be a periodic H^1 set of data. Then there exists a periodic H^1 mild solution (u, p, u_0, f, T, L) with the indicated data.

We may also phrase a quantitative variant of this conjecture:

Conjecture 1.10 (*A priori* periodic H^1 bound). There exists a function $F : \mathbb{R}^+ \times \mathbb{R}^+ \times \mathbb{R}^+ \to \mathbb{R}^+$ with the property that whenever (u, p, u_0, f, T, L) is a smooth periodic normalised-pressure solution with $0 < T < T_0 < \infty$ and

$$\mathcal{H}^1(u_0, f, T, L) \le A < \infty,$$

we have

$$\|u\|_{L_t^\infty H_x^1([0,T]\times\mathbb{R}^3/L\mathbb{Z}^3)} \le F(A, L, T_0).$$

Remark 1.11. By rescaling, one may set $L = 1$ in this conjecture without any loss of generality; by partitioning the time interval $[0, T_0]$ into smaller subintervals, we may also simultaneously set $T_0 = 1$ if desired. Thus, the key point is that the size of the data A is allowed to be large (for small A the conjecture follows from the local well-posedness theory; see Theorem 5.1).

As we shall soon see, Conjecture 1.9 and Conjecture 1.10 are actually equivalent.

We now turn to the nonperiodic setting. In Conjecture 1.5, the hypothesis that the initial data be Schwartz may seem unnecessarily restrictive, given that the incompressible nature of the fluid implies that the Schwartz property need not be preserved over time; also, there are many interesting examples of initial data that are smooth and finite energy (or H^1) but not Schwartz. In particular, one can consider generalising Conjecture 1.5 to data that is merely smooth and H^1, or even smooth and finite energy, rather than Schwartz[5] of Conjecture 1.5. Unfortunately, the naive generalisation of Conjecture 1.5 (or even Conjecture 1.3) fails instantaneously in this case:

Theorem 1.12 (No smooth solutions from smooth H^1 data). *There exists smooth $u_0 \in H_x^1(\mathbb{R}^3)$ such that there does not exist any smooth finite energy solution $(u, p, u_0, 0, T)$ with the indicated data for any $T > 0$.*

We prove this proposition in Section 15. At first glance, this proposition looks close to being a negative answer to either Conjecture 1.5 or Conjecture 1.3, but it relies on a technicality; for smooth H^1 data, the second derivatives of u_0 need not be square-integrable, and this can cause enough oscillation in the pressure to prevent the pressure from being C_t^2 (or the velocity field from being C_t^3) at the initial time[6] $t = 0$. This theorem should be compared with the classical local existence theorem of Heywood [1980], which obtains smooth solutions for small *positive* times from smooth data with finite enstrophy, but merely obtains continuity at the initial time $t = 0$.

The situation is even worse in the inhomogeneous setting; the argument in Theorem 1.12 can be used to construct inhomogeneous smooth H^1 data whose solutions will now be nonsmooth in time at all times,

[5] We are indebted to Andrea Bertozzi for suggesting these formulations of the Navier–Stokes global regularity problem.

[6] For most evolutionary PDEs, one can gain unlimited time differentiability at $t = 0$ assuming smooth initial data by differentiating the PDE in time (see the proof of the Cauchy–Kowalesky theorem). However, the problem here is that the pressure p in the Navier–Stokes equation does not obey an evolutionary PDE, but is instead determined in a nonlocal fashion from the initial data u (see (9)), which prevents one from obtaining much time regularity of the pressure initially.

not just at the initial time $t = 0$. Because of this, we will not attempt to formulate a global regularity problem in the inhomogeneous smooth H^1 or inhomogeneous smooth finite energy categories.

In the homogeneous setting, though, we can get around this technical obstruction by introducing the notion of an *almost smooth finite energy solution* (u, p, u_0, f, T), which is the same concept as a smooth finite energy solution, but instead of requiring u, p to be smooth on $[0, T] \times \mathbb{R}^3$, we instead require that u, p are smooth on $(0, T] \times \mathbb{R}^3$, and for each $k \geq 0$, the functions $\nabla_x^k u$, $\partial_t \nabla_x^k u$, $\nabla_x^k p$ exist and are continuous on $[0, T] \times \mathbb{R}^3$. Thus, the only thing that almost smooth solutions lack when compared to smooth solutions is a limited amount of time differentiability at the starting time $t = 0$; informally, u is only $C_t^1 C_x^\infty$ at $t = 0$, and p is only $C_t^0 C_x^\infty$ at $t = 0$. This is still enough regularity to interpret the Navier–Stokes equation (3) in the classical sense, but is not a completely smooth solution.

The "corrected" conjectures for global regularity in the homogeneous smooth H^1 and smooth finite energy categories are then:

Conjecture 1.13 (Global almost regularity for homogeneous H^1). Let $(u_0, 0, T)$ be a smooth homogeneous H^1 set of data. Then there exists an almost smooth finite energy solution $(u, p, u_0, 0, T)$ with the indicated data.

Conjecture 1.14 (Global almost regularity for homogeneous finite energy data). Let $(u_0, 0, T)$ be a smooth homogeneous finite energy set of data. Then there exists an almost smooth finite energy solution $(u, p, u_0, 0, T)$ with the indicated data.

We carefully note that these conjectures only concern *existence* of smooth solutions, and not uniqueness; we will comment on some of the uniqueness issues later in this paper.

Another way to repair the global regularity conjectures in these settings is to abandon smoothness altogether, and work instead with the notion of mild solutions. More precisely, define a H^1 *mild solution* (u, p, u_0, f, T) to be fields $u, f : [0, T] \times \mathbb{R}^3 \to \mathbb{R}^3$, $p : [0, T] \times \mathbb{R}^3 \to \mathbb{R}$, $u_0 : \mathbb{R}^3 \to \mathbb{R}^3$ with $0 < T < \infty$, obeying the regularity hypotheses

$$\begin{aligned}
&u_0 \in H_x^1(\mathbb{R}^3),\\
&f \in L_t^\infty H_x^1([0, T] \times \mathbb{R}^3),\\
&u \in L_t^\infty H_x^1 \cap L_t^2 H_x^2([0, T] \times \mathbb{R}^3),
\end{aligned}$$

with p being given by (9), which obey (4), (1), and (10) (and thus (11)). Similarly, define the concept of H^1 data (u_0, f, T).

We then have the following conjectures in the homogeneous setting:

Conjecture 1.15 (Global well-posedness in homogeneous H^1). Let $(u_0, 0, T)$ be a homogeneous H^1 set of data. Then there exists an H^1 mild solution $(u, p, u_0, 0, T)$ with the indicated data.

Conjecture 1.16 (*A priori* homogeneous H^1 bound). There exists a function $F : \mathbb{R}^+ \times \mathbb{R}^+ \to \mathbb{R}^+$ with the property that whenever $(u, p, u_0, 0, T)$ is a smooth H^1 solution with $0 < T < T_0 < \infty$ and

$$\|u_0\|_{H_x^1(\mathbb{R}^3)} \leq A < \infty,$$

we have

$$\|u\|_{L_t^\infty H_x^1([0,T]\times\mathbb{R}^3)} \leq F(A, T_0).$$

We also phrase a global-in-time variant:

Conjecture 1.17 (*A priori* global homogeneous H^1 bound). There exists a function $F : \mathbb{R}^+ \to \mathbb{R}^+$ with the property that whenever $(u, p, u_0, 0, T)$ is a smooth H^1 solution with

$$\|u_0\|_{H_x^1(\mathbb{R}^3)} \leq A < \infty,$$

then

$$\|u\|_{L_t^\infty H_x^1([0,T]\times\mathbb{R}^3)} \leq F(A).$$

In the inhomogeneous setting, we will state two slightly technical conjectures:

Conjecture 1.18 (Global well-posedness from spatially smooth Schwartz data). Let (u_0, f, T) be data obeying the bounds

$$\sup_{x\in\mathbb{R}^3} (1+|x|)^k |\nabla_x^\alpha u_0(x)| < \infty$$

and

$$\sup_{(t,x)\in[0,T]\times\mathbb{R}^3} (1+|x|)^k |\nabla_x^\alpha f(x)| < \infty$$

for all $k, \alpha \geq 0$. Then there exists an H^1 mild solution (u, p, u_0, f, T) with the indicated data.

Conjecture 1.19 (Global well-posedness from spatially smooth H^1 data). Let (u_0, f, T) be an H^1 set of data, such that

$$\sup_{x\in K} |\nabla_x^\alpha u_0(x)| < \infty$$

and

$$\sup_{(t,x)\in[0,T]\times K} |\nabla_x^\alpha f(x)| < \infty$$

for all $\alpha \geq 0$ and all compact K. Then there exists an H^1 mild solution (u, p, u_0, f, T) with the indicated data.

Needless to say, we do not establish[7] any of these conjectures unconditionally in this paper. However, as the main result of this paper, we are able to establish the following implications:

Theorem 1.20 (Implications). (i) *Conjectures 1.9 and 1.10 are equivalent.*

(ii) *Conjecture 1.9 implies Conjecture 1.8* (*and hence also Conjectures 1.6 and 1.4*).

(iii) *Conjecture 1.9 implies Conjecture 1.19*, *which is equivalent to Conjecture 1.18.*

(iv) *Conjecture 1.19 implies Conjectures 1.13 and 1.5* (*and hence also Conjecture 1.3*).

(v) *Conjecture 1.13 is equivalent to Conjecture 1.14.*

(vi) *Conjectures 1.13*, *1.15*, *1.16*, *and 1.17 are all equivalent.*

[7] Indeed, the arguments here do not begin to address the main issue in any of these conjectures, namely the analysis of fine-scale (and turbulent) behaviour. The results in this paper do not prevent singularities from occurring in the Navier–Stokes flow; but they can largely localise the impact of such singularities to a bounded region of space.

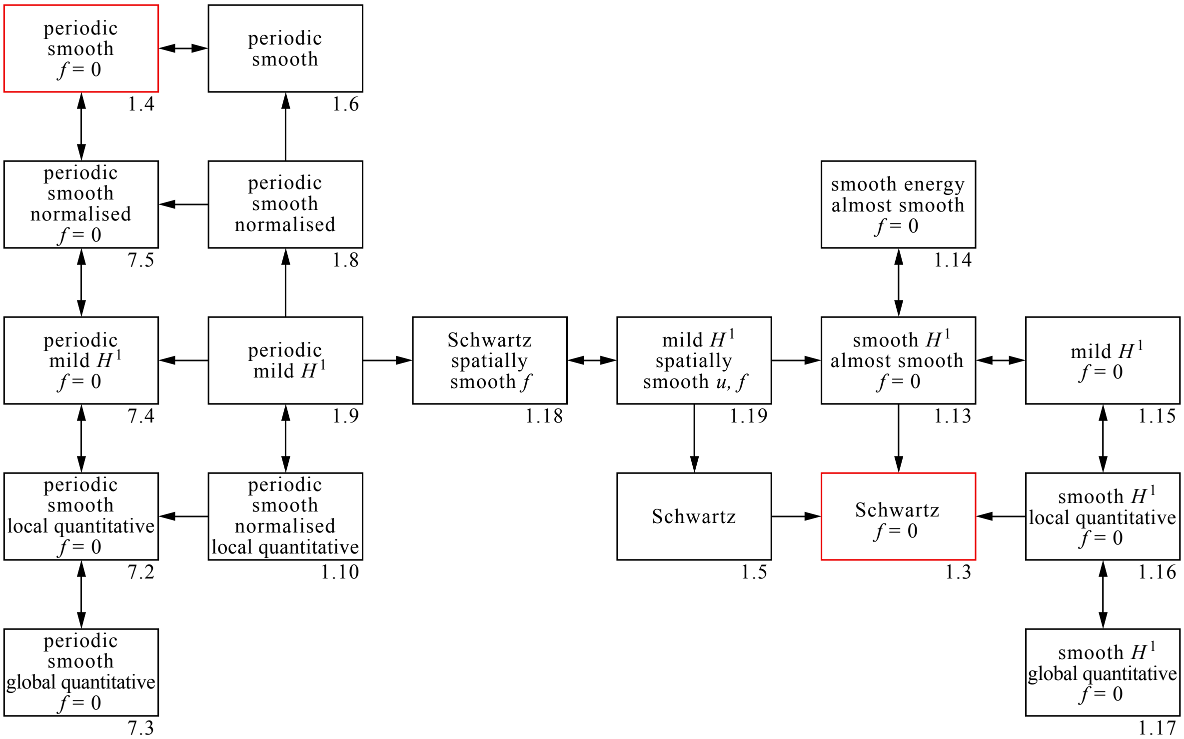

Figure 1. Known implications between the various conjectures described here (existence of smooth or mild solutions, or local or global quantitative bounds in the periodic, Schwartz, H^1, or finite energy categories, with or without normalised pressure, and with or without the $f = 0$ condition) and also in [Tao 2007] (the latter conjectures and implications occupy the far left column). A positive solution to the red problems, or a negative solution to the blue problems, qualify for the Clay Millennium prize, as stated in [Fefferman 2006].

The logical relationship between these conjectures, given by the implications above (as well as some trivial implications, and the equivalences in [Tao 2007]), is displayed in Figure 1.

Among other things, these results essentially show that in order to solve the Navier–Stokes global regularity problem, it suffices to study the periodic setting (but with the caveat that one now has to consider forcing terms with the regularity of $L^\infty_t H^1_x$).

Theorem 1.20(i) is a variant of the compactness arguments used in [Tao 2007] (see also [Gallagher 2001; Rusin and Šverák 2011]), and is proven in Section 7. Part (ii) of this theorem is a standard consequence of the periodic H^1 local well-posedness theory, which we review in Section 5. In the homogeneous $f = 0$ case it is possible to reverse this implication by the compactness arguments mentioned previously; see [Tao 2007]. However, we were unable to obtain this converse implication in the inhomogeneous case. Part (iv) is similarly a consequence of the nonperiodic H^1 local well-posedness theory, and is also proven in Section 5.

Part (vi) is also a variant of the results in [Tao 2007], with the main new ingredient being a use of concentration compactness instead of compactness in order to deal with the unboundedness of the spatial domain $\mathbb{R}^3$, using the methods from [Bahouri and Gérard 1999; Gérard 1998; Gallagher 2001]. We establish these results in Section 14.

The more novel aspects of this theorem are parts (iii) and (v), which we establish in Sections 12 and 13 respectively. These results rely primarily on a new localised enstrophy inequality (Theorem 10.1) which can be viewed as a weak version of finite speed of propagation[8] for the enstrophy $\frac{1}{2}\int_{\mathbb{R}^3} |\omega(t,x)|^2\,dx$, where $\omega := \nabla \times u$ is the vorticity. We will also obtain a similar localised energy inequality for the energy $\frac{1}{2}\int_{\mathbb{R}^3} |u(t,x)|^2\,dx$, but it will be the enstrophy inequality that is of primary importance to us, as the enstrophy is a subcritical quantity and can be used to obtain regularity (and local control on enstrophy can similarly be used to obtain local regularity). Remarkably, one is able to obtain local enstrophy inequalities even though the only *a priori* controlled quantity, namely the energy, is supercritical; the main difficulty is a harmonic analysis one, namely to control nonlinear effects primarily in terms of the local enstrophy and only secondarily in terms of the energy.

Remark 1.21. As one can see from Figure 1, the precise relationship between all the conjectures discussed here is rather complicated. However, if one is willing to ignore the distinction between homogeneous and inhomogeneous data, as well as the (rather technical) distinction between smooth and almost smooth solutions, then the main implications can then be informally summarised as follows:

- (Homogenisation) Without pressure normalisation, the inhomogeneity in the periodic global regularity conjecture is irrelevant: the inhomogeneous regularity conjecture is equivalent to the homogeneous one.
- (Localisation) The global regularity problem in the Schwartz, H^1, and finite energy categories are "essentially" equivalent to each other.
- (More localisation) The global regularity problem in any of the above three categories is "essentially" a consequence of the global regularity problem in the periodic category.
- (Concentration compactness) Quantitative and qualitative versions of the global regularity problem (in a variety of categories) are "essentially" equivalent to each other.

The qualifier "essentially" here though needs to be taken with a grain of salt; again, one should consult Figure 1 for an accurate depiction of the implications.

The local enstrophy inequality has a number of other consequences, for instance allowing one to construct Leray–Hopf weak solutions whose (spatial) singularities are compactly supported in space; see Proposition 11.9.

Remark 1.22. Since the submission of this manuscript, the referee pointed out that the partial regularity theory of Caffarelli, Kohn, and Nirenberg [1982] also allows one to partially reverse the implication in

[8] Actually, in our setting, "finite distance of propagation" would be more accurate; we obtain an L^1_t bound for the propagation velocity (see Proposition 9.1) rather than an L^∞_t bound.

Theorem 1.20(iii), and more specifically to deduce Conjecture 1.8 from Conjecture 1.19. We sketch the referee's argument in Remark 12.3.

2. Notation and basic estimates

We use $X \lesssim Y$, $Y \gtrsim X$, or $X = O(Y)$ to denote the estimate $X \leq CY$ for an absolute constant C. If we need C to depend on a parameter, we shall indicate this by subscripts; thus for instance $X \lesssim_s Y$ denotes the estimate $X \leq C_s Y$ for some C_s depending on s. We use $X \sim Y$ as shorthand for $X \lesssim Y \lesssim X$.

We will occasionally use the Einstein summation conventions, using Roman indices i, j to range over the three spatial dimensions 1, 2, 3, though we will not bother to raise and lower these indices; for instance, the components of a vector field u will be u_i. We use ∂_i to denote the derivative with respect to the i-th spatial coordinate x_i. Unless otherwise specified, the Laplacian $\Delta = \partial_i \partial_i$ will denote the spatial Laplacian. (In Lemma 12.1, though, we will briefly need to deal with the Laplace–Beltrami operator Δ_{S^2} on the sphere S^2.) Similarly, ∇ will refer to the spatial gradient $\nabla = \nabla_x$ unless otherwise stated. We use the usual notations ∇f, $\nabla \cdot u$, $\nabla \times u$, for the gradient, divergence, or curl of a scalar field f or a vector field u.

It will be convenient (particularly when dealing with nonlinear error terms) to use *schematic notation*, in which an expression such as $\mathbb{O}(uvw)$ involving some vector- or tensor-valued quantities u, v, w denotes some constant-coefficient combination of products of the components of u, v, w respectively, and similarly for other expressions of this type. Thus, for instance, $\nabla \times \nabla \times u$ could be written schematically as $\mathbb{O}(\nabla^2 u)$, $|u \times v|^2$ could be written schematically as $\mathbb{O}(uuvv)$, and so forth.

For any centre $x_0 \in \mathbb{R}^3$ and radius $R > 0$, we use $B(x_0, R) := \{x \in \mathbb{R}^3 : |x - x_0| \leq R\}$ to denote the (closed) Euclidean ball. Much of our analysis will be localised to a ball $B(x_0, R)$, an annulus $B(x_0, R)\backslash B(x_0, r)$, or an exterior region $\mathbb{R}^3 \backslash B(x_0, R)$ (and often x_0 will be normalised to the origin 0).

We define the absolute value of a tensor in the usual Euclidean sense. Thus, for instance, if $u = u_i$ is a vector field, then $|u|^2 = u_i u_i$, $|\nabla u|^2 = (\partial_i u_j)(\partial_i u_j)$, $|\nabla^2 u|^2 = (\partial_i \partial_j u_k)(\partial_i \partial_j u_k)$, and so forth.

If E is a set, we use 1_E to denote the associated indicator function; thus $1_E(x) = 1$ when $x \in E$ and $1_E(x) = 0$ otherwise. We sometimes also use a statement in place of E; thus for instance $1_{k \neq 0}$ would equal 1 if $k \neq 0$ and 0 when $k = 0$.

We use the usual Lebesgue spaces $L^p(\Omega)$ for various domains Ω (usually subsets of Euclidean space $\mathbb{R}^3$ or a torus $\mathbb{R}^3/L\mathbb{Z}^3$) and various exponents $1 \leq p \leq \infty$, which will always be equipped with an obvious Lebesgue measure. We often write $L^p(\Omega)$ as $L^p_x(\Omega)$ to emphasise the spatial nature of the domain Ω. Given an absolutely integrable function $f \in L^1_x(\mathbb{R}^3)$, we define the Fourier transform $\hat{f} : \mathbb{R}^3 \to \mathbb{C}$ by the formula

$$\hat{f}(\xi) := \int_{\mathbb{R}^3} e^{-2\pi i x \cdot \xi} f(x)\, dx;$$

we then extend this Fourier transform to tempered distributions in the usual manner. For a function f which is periodic with period 1, and thus representable as a function on the torus $\mathbb{R}^3/\mathbb{Z}^3$, we define the discrete Fourier transform $\hat{f} : \mathbb{Z}^3 \to \mathbb{C}$ by the formula

$$\hat{f}(k) := \int_{\mathbb{R}^3/\mathbb{Z}^3} e^{-2\pi i k\cdot x} f(x)\, dx$$

when f is absolutely integrable on $\mathbb{R}^3/\mathbb{Z}^3$, and extend this to more general distributions on $\mathbb{R}^3/\mathbb{Z}^3$ in the usual fashion. Strictly speaking, these two notations are not compatible with each other, but it will always be clear in context whether we are using the nonperiodic or the periodic Fourier transform.

For any spatial domain Ω (contained in either $\mathbb{R}^3$ or $\mathbb{R}^3/L\mathbb{Z}^3$) and any natural number $k \geq 0$, we define the classical Sobolev norms $\|u\|_{H^k_x(\Omega)}$ of a smooth function $u : \Omega \to \mathbb{R}$ by the formula

$$\|u\|_{H^k_x(\Omega)} := \left(\sum_{j=0}^{k} \|\nabla^j u\|^2_{L^2_x(\Omega)} \right)^{1/2},$$

and say that $u \in H^k_x(\Omega)$ when $\|u\|_{H^k_x(\Omega)}$ is finite. Note that we do not impose any vanishing conditions at the boundary of Ω, and to avoid technical issues we will not attempt to define these norms for nonsmooth functions u in the event that Ω has a nontrivial boundary. In the domain $\mathbb{R}^3$ and for $s \in \mathbb{R}$, we define the Sobolev norm $\|u\|_{H^s_x(\mathbb{R}^3)}$ of a tempered distribution $u : \mathbb{R}^3 \to \mathbb{R}$ by the formula

$$\|u\|_{H^s_x(\mathbb{R}^3)} := \left(\int_{\mathbb{R}^3} (1 + |\xi|^2)^s |\hat{u}(\xi)|^2\, d\xi \right)^{1/2}.$$

Strictly speaking, this conflicts slightly with the previous notation when k is a nonnegative integer, but the two norms are equivalent up to constants (and both norms define a Hilbert space structure), so the distinction will not be relevant for our purposes. For $s > -\frac{3}{2}$, we also define the homogeneous Sobolev norm

$$\|u\|_{\dot{H}^s_x(\mathbb{R}^3)} := \left(\int_{\mathbb{R}^3} |\xi|^{2s} |\hat{u}(\xi)|^2\, d\xi \right)^{1/2},$$

and let $H^s_x(\mathbb{R}^3)$, $\dot{H}^s_x(\mathbb{R}^3)$ be the space of tempered distributions with finite $H^s_x(\mathbb{R}^3)$ or $\dot{H}^s_x(\mathbb{R}^3)$ norm respectively. Similarly, on the torus $\mathbb{R}^3/\mathbb{Z}^3$ and $s \in \mathbb{R}$, we define the Sobolev norm $\|u\|_{H^s_x(\mathbb{R}^3/\mathbb{Z}^3)}$ of a distribution $u : \mathbb{R}^3/\mathbb{Z}^3 \to \mathbb{R}$ by the formula

$$\|u\|_{H^s_x(\mathbb{R}^3/\mathbb{Z}^3)} := \left(\sum_{k\in\mathbb{Z}^3} (1 + |k|^2)^s |\hat{u}(k)|^2 \right)^{1/2};$$

again, this conflicts slightly with the classical Sobolev norms $H^k_x(\mathbb{R}^3/\mathbb{Z}^3)$, but this will not be a serious issue in this paper. We define $H^s_x(\mathbb{R}^3/\mathbb{Z}^3)$ to be the space of all distributions u with finite $H^s_x(\mathbb{R}^3/\mathbb{Z}^3)$ norm, and $H^s_x(\mathbb{R}^3/\mathbb{Z}^3)_0$ to be the codimension-one subspace of functions or distributions u which are mean-zero in the sense that $\hat{u}(0) = 0$.

In a similar vein, given a spatial domain Ω and a natural number $k \geq 0$, we define $C^k_x(\Omega)$ to be the space of all k times continuously differentiable functions $u : \Omega \to \mathbb{R}$ whose norm

$$\|u\|_{C^k_x(\Omega)} := \sum_{j=0}^{k} \|\nabla^j u\|_{L^\infty_x(\Omega)}$$

is finite[9]. Given any spatial norm $\| \|_{X_x(\Omega)}$ associated to a function space X_x defined on a spatial domain Ω, and a time interval I, we can define mixed-norms $\|u\|_{L^p_t X_x(I\times\Omega)}$ on functions $u : I \times \Omega \to \mathbb{R}$ by the formula

$$\|u\|_{L^p_t X_x(I\times\Omega)} := \left(\int_I \|u(t)\|^p_{X_x(\Omega)}\, dt \right)^{1/p}$$

when $1 \leq p < \infty$, and

$$\|u\|_{L^\infty_t X_x(I\times\Omega)} := \operatorname{ess\,sup}_{t\in I} \|u(t)\|_{X_x(\Omega)},$$

assuming in both cases that $u(t)$ lies in $X(\Omega)$ for almost every Ω, and then let $L^p_t X_x(I \times \Omega)$ be the space of functions (or, in some cases, distributions) whose $L^p_t X_x(I \times \Omega)$ is finite. Thus, for instance, $L^\infty_t C^2_x(I\times\Omega)$ would be the space of functions $u : I\times\Omega \to \mathbb{R}$ such that for almost every $x \in I$, $u(t) : \Omega \to \mathbb{R}$ is in $C^2_x(\Omega)$, and the norm

$$\|u\|_{L^\infty_t C^2_x(I\times\Omega)} := \operatorname{ess\,sup}_{t\in I} \|u(t)\|_{C^2_x(\Omega)}$$

is finite.

Similarly, for any natural number $k \geq 0$, we define $C^k_t X_x(I \times \Omega)$ to be the space of all functions $u : I \times \Omega \to \mathbb{R}$ such that the curve $t \mapsto u(t)$ from I to $X_x(\Omega)$ is k times continuously differentiable, and such that the norm

$$\|u\|_{C^k_t X_x(I\times\Omega)} := \sum_{j=0}^{k} \|\nabla^j u\|_{L^\infty_t X_x(I\times\Omega)}$$

is finite.

Given two normed function spaces X, Y on the same domain (in either space or space-time), we can endow their intersection $X \cap Y$ with the norm

$$\|u\|_{X\cap Y} := \|u\|_X + \|u\|_Y.$$

For us, the most common example of such hybrid norms will be the spaces

$$X^s(I \times \Omega) := L^\infty_t H^s_x(I \times \Omega) \cap L^2_x H^{s+1}_x(I \times \Omega), \tag{13}$$

defined whenever I is a time interval, s is a natural number, and Ω is a spatial domain, or whenever I is a time interval, s is real, and Ω is either $\mathbb{R}^3$ or $\mathbb{R}^3/\mathbb{Z}^3$. The X^s spaces (particularly X^1) will play a prominent role in the (subcritical) local well-posedness theory for the Navier–Stokes equations; see

[9]Note that if Ω is noncompact, then it is possible for a smooth function to fail to lie in $C^k(\Omega)$ if it becomes unbounded or excessively oscillatory at infinity. One could use a notation such as $C^k_{x,\mathrm{loc}}(\Omega)$ to describe the space of functions that are k times continuously differentiable with no bounds on derivatives, but we will not need such notation here.

Section 5. The space X^0 will also be naturally associated with energy estimates, and the space X^1 with enstrophy estimates.

All of these above function spaces can of course be extended to functions that are vector or tensor-valued without difficulty (there are multiple ways to define the norms in these cases, but all such definitions will be equivalent up to constants).

We use the Fourier transform to define a number of useful multipliers on $\mathbb{R}^3$ or $\mathbb{R}^3/\mathbb{Z}^3$. On $\mathbb{R}^3$, we formally define the inverse Laplacian operator Δ^{-1} by the formula

$$\widehat{\Delta^{-1} f}(\xi) := \frac{-1}{4\pi^2|\xi|^2}\hat{f}(\xi), \tag{14}$$

which is well-defined for any tempered distribution $f : \mathbb{R}^3 \to \mathbb{R}$ for which the right-hand side of (14) is locally integrable. This is for instance the case if f lies in the k-th derivative of a function in $L^1_x(\mathbb{R}^3)$ for some $k \geq 0$, or the k-th derivative of a function in $L^2_x(\mathbb{R}^3)$ for some $k \geq 1$. If $f \in L^1_x(\mathbb{R}^3)$, then as is well known, one has the Newton potential representation

$$\Delta^{-1} f(x) = \frac{-1}{4\pi}\int_{\mathbb{R}^3} \frac{f(y)}{|x-y|}\, dy. \tag{15}$$

Note in particular that (15) implies that if $f \in L^1_x(\mathbb{R}^3)$ is supported on some closed set K, then $\Delta^{-1} f$ will be smooth away from K. Also observe from Fourier analysis (and decomposition into local and global components) that if f is smooth and is either the k-th derivative of a function in $L^1_x(\mathbb{R}^3)$ for some $k \geq 0$, or the k-th derivative of a function in $L^2_x(\mathbb{R}^3)$ for some $k \geq 1$, then $\Delta^{-1} f$ will be smooth also.

We also note that the Newton potential $-1/(4\pi|x-y|)$ is smooth away from the diagonal $x = y$. Because of this, we will often be able to obtain large amounts of regularity in space in the "far field" region when $|x|$ is large, for fields such as the velocity field u. However, it will often be significantly more challenging to gain significant amounts of regularity in *time*, because the inverse Laplacian Δ^{-1} has no smoothing properties in the time variable.

On $\mathbb{R}^3/\mathbb{Z}^3$, we similarly define the inverse Laplacian operator Δ^{-1} for distributions $f : \mathbb{R}^3/\mathbb{Z}^3 \to \mathbb{R}$ with $\hat{f}(0) = 0$ by the formula

$$\widehat{\Delta^{-1} f}(k) := \frac{-1_{k\neq 0}}{4\pi^2|k|^2}\hat{f}(k). \tag{16}$$

We define the *Leray projection* Pu of a (tempered distributional) vector field $u : \mathbb{R}^3 \to \mathbb{R}^3$ by the formula

$$Pu := \Delta^{-1}(\nabla \times \nabla \times u).$$

If u is square-integrable, then Pu is the orthogonal projection of u onto the space of square-integrable divergence-free vector fields; from Calderón–Zygmund theory, we know that the projection P is bounded on $L^p_x(\mathbb{R}^3)$ for every $1 < p < \infty$, and from Fourier analysis we see that P is also $H^s_x(\mathbb{R}^3)$ for every $s \in \mathbb{R}$. Note that if u is square-integrable and divergence-free, then $Pu = u$, and we thus have the *Biot–Savart law*

$$u = \Delta^{-1}(\nabla \times \omega), \tag{17}$$

where $\omega := \nabla \times u$.

In either $\mathbb{R}^3$ or $\mathbb{R}^3/L\mathbb{Z}^3$, we let $e^{t\Delta}$ for $t > 0$ be the usual heat semigroup associated to the heat equation $u_t = \Delta u$. On $\mathbb{R}^3$, this takes the explicit form

$$e^{t\Delta} f(x) = \frac{1}{(4\pi t)^{3/2}} \int_{\mathbb{R}^3} e^{-|x-y|^2/4t} f(y)\, dy$$

for $f \in L^p_x(\mathbb{R}^3)$ for some $1 \le p \le \infty$. From Young's inequality, we thus record the dispersive inequality

$$\|e^{t\Delta} f\|_{L^q(\mathbb{R}^3)} \lesssim t^{3/2q-3/2p} \|f\|_{L^p(\mathbb{R}^3)} \tag{18}$$

whenever $1 \le p \le q \le \infty$ and $t > 0$.

We recall *Duhamel's formula*

$$u(t) = e^{(t-t_0)\Delta} u(t_0) + \int_{t_0}^{t} e^{(t-t')\Delta} (\partial_t u - \Delta u)(t')\, dt' \tag{19}$$

whenever $u : [t_0, t] \times \Omega \to \mathbb{R}$ is a smooth tempered distribution, with Ω equal to either $\mathbb{R}^3$ or $\mathbb{R}^3/\mathbb{Z}^3$.

We record some linear and bilinear estimates involving Duhamel-type integrals and the spaces X^s defined in (13), which are useful in the local H^1 theory for the Navier–Stokes equation:

Lemma 2.1 (Linear and bilinear estimates). *Let $[t_0, t_1]$ be a time interval, let Ω be either $\mathbb{R}^3$ or $\mathbb{R}^3/\mathbb{Z}^3$, and suppose that $u : [t_0, t_1] \times \Omega \to \mathbb{R}$ and $F : [t_0, t_1] \times \Omega \to \mathbb{R}$ are tempered distributions such that*

$$u(t) = e^{(t-t_0)\Delta} u(t_0) + \int_{t_0}^{t} e^{(t-t')\Delta} F(t')\, dt'. \tag{20}$$

Then we have the standard energy estimate[10]

$$\|u\|_{X^s([t_0,t_1]\times\Omega)} \lesssim_s \|u(t_0)\|_{H^s_x(\Omega)} + \|F\|_{L^1_t H^s_x([t_0,t_1]\times\Omega)} \tag{21}$$

for any $s \ge 0$, as well as the variant

$$\|u\|_{X^s([t_0,t_1]\times\Omega)} \lesssim_s \|u(t_0)\|_{H^s_x(\Omega)} + \|F\|_{L^2_t H^{s-1}_x([t_0,t_1]\times\Omega)} \tag{22}$$

for any $s \ge 1$. We also note the further variant

$$\|u\|_{X^s([t_0,t_1]\times\Omega)} \lesssim_s \|u(t_0)\|_{H^s_x(\Omega)} + \|F\|_{L^4_t L^2_x([t_0,t_1]\times\Omega)} \tag{23}$$

for any $s < 3/2$.

We also have the bilinear estimate

$$\|\nabla(uv)\|_{L^4_t L^2_x([t_0,t_1]\times\Omega)} \lesssim \|u\|_{X^1([t_0,t_1]\times\Omega)} \|v\|_{X^1([t_0,t_1]\times\Omega)} \tag{24}$$

for any $u, v : [t_0, t_1] \times \mathbb{R}^3 \to \mathbb{R}$, which in particular implies (by a Hölder in time) that

$$\|\nabla(uv)\|_{L^2_t L^2_x([t_0,t_1]\times\mathbb{R}^3)} \lesssim (t_1 - t_0)^{1/4} \|u\|_{X^1([t_0,t_1]\times\mathbb{R}^3)} \|v\|_{X^1([t_0,t_1]\times\mathbb{R}^3)}. \tag{25}$$

[10] We adopt the convention that an estimate is vacuously true if the right-hand side is infinite or undefined.

Proof. The estimates[11] (22), (23), (24) are established in [Tao 2007, Lemma 2.1, Proposition 2.2]. The estimate (21) follows from the $F = 0$ case of (21) and Minkowski's inequality. $\square$

Finally, we define the Littlewood–Paley projection operators on $\mathbb{R}^3$. Let $\varphi(\xi)$ be a fixed bump function supported in the ball $\{\xi \in \mathbb{R}^3 : |\xi| \leq 2\}$ and equal to 1 on the ball $\{\xi \in \mathbb{R}^3 : |\xi| \leq 1\}$. Define a *dyadic number* to be a number N of the form $N = 2^k$ for some integer k. For each dyadic number N, we define the Fourier multipliers

$$\begin{aligned}
\widehat{P_{\leq N} f}(\xi) &:= \varphi(\xi/N)\hat{f}(\xi), \\
\widehat{P_{>N} f}(\xi) &:= (1 - \varphi(\xi/N))\hat{f}(\xi), \\
\widehat{P_N f}(\xi) &:= \psi(\xi/N)\hat{f}(\xi) := (\varphi(\xi/N) - \varphi(2\xi/N))\hat{f}(\xi).
\end{aligned}$$

We similarly define $P_{<N}$ and $P_{\geq N}$. Thus for any tempered distribution, we have $f = \sum_N P_N f$ in a weakly convergent sense at least, where the sum ranges over dyadic numbers. We recall the usual *Bernstein estimates*

$$\begin{aligned}
\|D^s P_N f\|_{L^p_x(\mathbb{R}^3)} &\lesssim_{p,s,D^s} N^s \|P_N f\|_{L^p_x(\mathbb{R}^3)}, \\
\|\nabla^k P_N f\|_{L^p_x(\mathbb{R}^3)} &\sim_{k,s} N^k \|P_N f\|_{L^p_x(\mathbb{R}^3)}, \\
\|P_{\leq N} f\|_{L^q_x(\mathbb{R}^3)} &\lesssim_{p,q} N^{3/p-3/q} \|P_{\leq N} f\|_{L^p_x(\mathbb{R}^3)}, \\
\|P_N f\|_{L^q_x(\mathbb{R}^3)} &\lesssim_{p,q} N^{3/p-3/q} \|P_N f\|_{L^p_x(\mathbb{R}^3)},
\end{aligned} \tag{26}$$

for all $1 \leq p \leq q \leq \infty$, $s \in \mathbb{R}$, $k \geq 0$, and pseudodifferential operators D^s of order s; see, for example, [Tao 2006, Appendix A].

We recall the *Littlewood–Paley trichotomy*: an expression of the form $P_N((P_{N_1} f_1)(P_{N_2} f_2))$ vanishes unless one of the following three scenarios holds:

- (Low-high interaction) $N_2 \lesssim N_1 \sim N$.
- (High-low interaction) $N_1 \lesssim N_2 \sim N$.
- (High-high interaction) $N \lesssim N_1 \sim N_2$.

This trichotomy is useful for obtaining estimates on bilinear expressions, as we shall see in Section 9.

We have the following frequency-localised variant of (18):

Lemma 2.2. *If N is a dyadic number and $f : \mathbb{R}^3 \to \mathbb{R}$ has Fourier transform supported on an annulus $\{\xi : |\xi| \sim N\}$, then we have*

$$\|e^{t\Delta} f\|_{L^q(\mathbb{R}^3)} \lesssim t^{3/2q-3/2p} \exp(-ctN^2) \|f\|_{L^p(\mathbb{R}^3)} \tag{27}$$

for some absolute constant $c > 0$ and all $1 \leq p \leq q \leq \infty$.

[11] Strictly speaking, the result in [Tao 2007] was stated for the torus rather than $\mathbb{R}^3$, but the argument works without modification in either domain, after first truncating $u(t_0)$, F to be Schwartz to avoid technicalities at infinity, and using a standard density argument.

Proof. By Littlewood–Paley projection, it suffices to show that

$$\|e^{t\Delta} P_N f\|_{L^q(\mathbb{R}^3)} \lesssim t^{3/2q-3/2p} \exp(-ctN^2) \|f\|_{L^p(\mathbb{R}^3)}$$

for all test functions f. By rescaling, we may set $t = 1$; in view of (18) we may then set $N \geq 1$. One then verifies from Fourier analysis that $e^{t\Delta} P_N$ is a convolution operator whose kernel has an $L^\infty_x(\mathbb{R}^3)$ and an $L^1_x(\mathbb{R}^3)$ norm that are both $O(\exp(-cN^2))$ for some absolute constant $c > 0$, and the claim follows from Young's inequality. □

From the uniform smoothness of the heat kernel, we also observe the estimate

$$\|e^{t\Delta} f\|_{C^k_x(K)} \lesssim_{k,K,T,p} \exp(-c_T r^2) \|f\|_{L^p_x(\mathbb{R}^3)} \tag{28}$$

whenever $0 \leq t \leq T$, $1 \leq p \leq \infty$, $k \geq 0$, K is a compact subset of $\mathbb{R}^3$, $r \geq 1$, f is supported on the set $\{x \in \mathbb{R}^3 : \operatorname{dist}(x, K) \geq r\}$, and some quantity $c_T > 0$ depending only on T. In practice, this estimate will be an effective substitute for finite speed of propagation for the heat equation.

3. Symmetries of the equation

In this section we review some well known symmetries of the Navier–Stokes flow that transform a given smooth solution (u, p, u_0, f, T) to another smooth solution $(\tilde{u}, \tilde{p}, \tilde{u}_0, \tilde{f}, \tilde{T})$, as these symmetries will be useful at various points in the paper.

The simplest symmetry is the *spatial translation symmetry*

$$\begin{aligned}
\tilde{u}(t, x) &:= u(t, x - x_0), \\
\tilde{p}(t, x) &:= p(t, x - x_0), \\
\tilde{u}_0(x) &:= u_0(x - x_0), \\
\tilde{f}(t, x) &:= f(t, x - x_0), \\
\tilde{T} &:= T,
\end{aligned} \tag{29}$$

valid for any $x_0 \in \mathbb{R}^3$; this transformation clearly maps mild, smooth, or almost smooth solutions to solutions of the same type, and also preserves conditions such as finite energy, H^1, periodicity, pressure normalisation, or the Schwartz property. In a similar vein, we have the *time translation symmetry*

$$\begin{aligned}
\tilde{u}(t, x) &:= u(t + t_0, x), \\
\tilde{p}(t, x) &:= p(t + t_0, x), \\
\tilde{u}_0(x) &:= u(t_0, x), \\
\tilde{f}(t, x) &:= f(t + t_0, x), \\
\tilde{T} &:= T - t_0,
\end{aligned} \tag{30}$$

valid for any $t_0 \in [0, T]$. Again, this maps mild, smooth, or almost smooth solutions to solutions of the same type (and if $t_0 > 0$, then almost smooth solutions are even upgraded to smooth solutions). If the original solution is finite energy or H^1, then the transformed solution will be finite energy or H^1

also. Note however that if it is only the original *data* that is assumed to be finite energy or H^1, as opposed to the *solution*, it is not immediately obvious that the time-translated solution remains finite energy or H^1, especially in view of the fact that the H^1 norm (or the enstrophy) is not a conserved quantity of the Navier–Stokes flow. (See however Lemma 8.1 and Corollary 11.1 below.) The situation is particularly dramatic in the case of Schwartz data; as remarked earlier, time translation can instantly convert[12] Schwartz data to non-Schwartz data, due to the slow decay of the Newton potential appearing in (9) (or of its derivatives, such as the Biot–Savart kernel in (17)).

Next, we record the *scaling symmetry*

$$\begin{aligned} \tilde{u}(t,x) &:= \frac{1}{\lambda} u\Big(\frac{t}{\lambda^2}, \frac{x}{\lambda}\Big), \\ \tilde{p}(t,x) &:= \frac{1}{\lambda^2} p\Big(\frac{t}{\lambda^2}, \frac{x}{\lambda}\Big), \\ \tilde{u}_0(x) &:= \frac{1}{\lambda} u\Big(\frac{x}{\lambda}\Big), \\ \tilde{f}(t,x) &:= \frac{1}{\lambda^3} f\Big(\frac{t}{\lambda^2}, \frac{x}{\lambda}\Big), \\ \tilde{T} &:= T\lambda^2, \end{aligned} \tag{31}$$

valid for any $\lambda > 0$; it also maps mild, smooth, or almost smooth solutions to solutions of the same type, and preserves properties such as finite energy, finite enstrophy, pressure normalisation, periodicity, or the Schwartz property, though note in the case of periodicity that a solution of period L will map to a solution of period λL. We will only use scaling symmetry occasionally in this paper, mainly because most of the quantities we will be manipulating will be supercritical with respect to this symmetry. Nevertheless, this scaling symmetry serves a fundamentally important conceptual purpose, by making the key distinction between subcritical, critical (or dimensionless), and supercritical quantities, which can help illuminate many of the results in this paper (and was also crucial in allowing the author to discover[13] these results in the first place).

We record three further symmetries that impact upon the issue of pressure normalisation. The first is the *pressure shifting symmetry*

$$\begin{aligned} \tilde{u}(t,x) &:= u(t,x), \\ \tilde{p}(t,x) &:= p(t,x) + C(t), \\ \tilde{u}_0(x) &:= u_0(x), \\ \tilde{f}(t,x) &:= f(t,x), \\ \tilde{T} &:= T, \end{aligned} \tag{32}$$

[12]This can be seen for instance by noting that moments such as $\int_{\mathbb{R}^3} \omega_1(t,x)(x_2^2 - x_3^2)\, dx$ are not conserved in time, but must equal zero whenever $u(t)$ is Schwartz.

[13]The author also found dimensional analysis to be invaluable in checking the calculations for errors. One *could*, if one wished, exploit the scaling symmetry to normalise a key parameter (for example, the energy E, or a radius parameter r) to equal one, which would simplify the numerology slightly, but then one would lose the use of dimensional analysis to check for errors, and so we have elected to largely avoid the use of scaling normalisations in this paper.

valid for any smooth function $C : \mathbb{R} \to \mathbb{R}$. This clearly maps smooth or almost smooth solutions to solutions of the same type, and preserves properties such as finite energy, H^1, periodicity, and the Schwartz property; however, it destroys pressure normalisation (and thus the notion of a mild solution). A slightly more sophisticated symmetry in the same spirit is the *Galilean symmetry*

$$\begin{aligned}
\tilde{u}(t,x) &:= u\left(t, x - \int_0^t v(s)\, ds\right) + v(t),\\
\tilde{p}(t,x) &:= p\left(t, x - \int_0^t v(s)\, ds\right) - x \cdot v'(t),\\
\tilde{u}_0(x) &:= u_0(x) + v(0),\\
\tilde{f}(t,x) &:= f\left(t, x - \int_0^t v(s)\, ds\right),\\
\tilde{T} &:= T,
\end{aligned} \tag{33}$$

valid for any smooth function $v : \mathbb{R} \to \mathbb{R}^3$. One can carefully check that this symmetry indeed maps mild, smooth solutions to smooth solutions and preserves periodicity (recall here that in our definition of a periodic solution, the pressure was *not* required to be periodic). On the other hand, this symmetry does not preserve finite energy, H^1, or the Schwartz property. It also clearly destroys the pressure normalisation property.

Finally, we observe that one can absorb divergences into the forcing term via the *forcing symmetry*

$$\begin{aligned}
\tilde{u}(t,x) &:= u(t,x),\\
\tilde{p}(t,x) &:= p(t,x) + q(t,x),\\
\tilde{u}_0(x) &:= u_0(x),\\
\tilde{f}(t,x) &:= f(t,x) + \nabla \cdot q(t,x),\\
\tilde{T} &:= T,
\end{aligned} \tag{34}$$

valid for any smooth function $P : [0,T] \times \mathbb{R}^3 \to \mathbb{R}^3$. If the new forcing term $\tilde{f}$ still has finite energy or is still periodic, then the normalisation of pressure is preserved. In the periodic setting, we will apply (34) with a linear term $q(t,x) := x \cdot a(t)$, allowing one to alter f by an arbitrary constant $a(t)$. In the finite energy or H^1 setting, one can use (34) and the Leray projection P to reduce to the divergence-free case $\nabla \cdot f = 0$; note, though, that this projection can destroy the Schwartz nature of f. This divergence-free reduction is particularly useful in the case of normalised pressure, since (9) then simplifies to

$$p = -\Delta^{-1} \partial_i \partial_j (u_i u_j). \tag{35}$$

One can of course compose these symmetries together to obtain a larger (semi)group of symmetries. For instance, by combining (33) and (34), we observe the symmetry

$$\begin{aligned}
\tilde{u}(t,x) &:= u\left(t, x - \int_0^t v(s)\,ds\right) + v(t),\\
\tilde{p}(t,x) &:= p\left(t, x - \int_0^t v(s)\,ds\right),\\
\tilde{u}_0(x) &:= u_0(x) + v(0),\\
\tilde{f}(t,x) &:= f\left(t, x - \int_0^t v(s)\,ds\right) + v'(t),\\
\tilde{T} &:= T,
\end{aligned} \tag{36}$$

for any smooth function $v : \mathbb{R} \to \mathbb{R}^3$. This symmetry is particularly useful for periodic solutions; note that it preserves both the periodicity property and the normalised pressure property. By choosing $v(t)$ appropriately, we see that we can use this symmetry to normalise periodic data (u_0, f, T, L) to be *mean-zero* in the sense that

$$\int_{\mathbb{R}^3/L\mathbb{Z}^3} u_0(x)\,dx = 0 \tag{37}$$

and

$$\int_{\mathbb{R}^3/L\mathbb{Z}^3} f(t,x)\,dx = 0 \tag{38}$$

for all $0 \le t \le T$. By integrating (3) over the torus $\mathbb{R}^3/L\mathbb{Z}^3$, we then conclude with this normalisation that u remains mean-zero for all times $0 \le t \le T$:

$$\int_{\mathbb{R}^3/L\mathbb{Z}^3} u(t,x)\,dx = 0. \tag{39}$$

The same conclusion also holds for periodic H^1 mild solutions.

4. Pressure normalisation

The symmetries in (32), (34) can alter the velocity field u and pressure p without affecting the data (u_0, f, T), thus leading to a breakdown of uniqueness for the Navier–Stokes equation. In this section we investigate this loss of uniqueness, and show that (in the smooth category, at least) one can "quotient out" these symmetries by reducing to the situation (9) of normalised pressure, at which point uniqueness can be recovered (at least in the H^1 category).

More precisely, we show:

Lemma 4.1 (Reduction to normalised pressure). (i) *If (u, p, u_0, f, T) is an almost smooth finite energy solution, then for almost every time $t \in [0, T]$, one has*

$$p(t,x) = -\Delta^{-1}\partial_i\partial_j(u_i u_j)(t,x) + \Delta^{-1}\nabla\cdot f(t,x) + C(t), \tag{40}$$

for some bounded measurable function $C : [0, T] \to \mathbb{R}$.

(ii) *If* (u, p, u_0, f, T) *is a periodic smooth solution, then there exist smooth functions* $C : [0, T] \to \mathbb{R}$ *and* $a : [0, T] \to \mathbb{R}^3$ *such that*

$$p(t, x) = -\Delta^{-1}\partial_i \partial_j (u_i u_j)(t, x) + \Delta^{-1}\nabla \cdot f(t, x) + x \cdot a(t) + C(t). \tag{41}$$

In particular, after applying a Galilean transformation (33) *followed by a pressure-shifting transformation* (32), *one can transform* (u, p, u_0, f, T) *into a periodic smooth solution with normalised pressure.*

Remark 4.2. Morally, in (i) the function C should be smooth (at least for times $t > 0$), which would then imply that one can apply a pressure-shifting transformation (32) to convert (u, p, u_0, f, T) into a smooth solution with normalised pressure. However, there is the technical difficulty that in our definition of a finite energy smooth solution, we do not *a priori* have any control of time derivatives of u in any $L^p_x(\mathbb{R}^3)$ norms, and as such we do not have time regularity on the component $\Delta^{-1}\partial_i \partial_j (u_i u_j)$ of (40). In practice, though, this possible irregularity of $C(t)$ will not bother us, as we only need to understand the gradient ∇p of the pressure, rather than the pressure itself, in order to solve the Navier–Stokes equations (3).

Proof. We begin with the periodic case, which is particularly easy due to Liouville's theorem (which, among other things, implies that the only harmonic periodic functions are the constants). We may normalise the period L to equal 1. Fix an almost smooth periodic solution (u, p, u_0, f, T). Define the normalised pressure $p_0 : [0, T] \times \mathbb{R}^3 \to \mathbb{R}$ by the formula

$$p_0 := -\Delta^{-1}\partial_i \partial_j (u_i u_j) + \Delta^{-1}\nabla \cdot f. \tag{42}$$

As u, f are smooth and periodic, p_0 is smooth also, and from (8) one has $\Delta p = \Delta p_0$. Thus one has

$$p = p_0 + h,$$

where $h : [0, T] \times \mathbb{R}^3 \to \mathbb{R}$ is a smooth function with $h(t)$ harmonic in space for each time t. The function h need not be periodic; however, from (3) we have

$$\partial_t u + (u \cdot \nabla)u = \Delta u - \nabla p_0 - \nabla h + f.$$

Every term aside from ∇h is periodic, and so ∇h is periodic also. Since ∇h is also harmonic, it must therefore be constant in space by Liouville's theorem. We therefore may write

$$h(t, x) = x \cdot a(t) + C(t)$$

for some $a(t) \in \mathbb{R}^3$ and $C(t) \in \mathbb{R}$; since h is smooth, a, C are smooth also, and the claim follows.

Now we turn to the finite energy case; thus (u, p, u_0, f, T) is now an almost smooth finite energy solution. By the time translation symmetry (30) with an arbitrarily small time shift parameter t_0, we may assume without loss of generality that (u, p, u_0, f, T) is smooth (and not just almost smooth). We define the normalised pressure p_0 by (42) as before; then for each time $t \in [0, T]$, one sees from (8) that

$$p(t) = p_0(t) + h(t)$$

for some harmonic function $h(t) : \mathbb{R}^3 \to \mathbb{R}$. As u, f are smooth and finite energy, one sees from (42) that p_0 is bounded on compact subsets of space-time; since p is smooth, we conclude that h is bounded on compact subsets of space-time also. From harmonicity, this implies that all spatial derivatives $\nabla^k h$ are also bounded on compact subsets of space time. However, as noted previously, we cannot impose any time regularity on p_0 or h because we do not have decay estimates on time derivatives of u.

It is easy to see that h is measurable. To obtain the lemma, it suffices to show that $h(t)$ is a constant function of x for almost every time t.

Let $[t_1, t_2]$ be any interval in $[0, T]$. Integrating (3) in time on this interval, we see that

$$u(t_2, x) - u(t_1, x) + \int_{t_1}^{t_2} (u \cdot \nabla) u(t, x)\, dt = \int_{t_1}^{t_2} \Delta u(t, x) - \nabla p(t, x) + f(t, x)\, dt.$$

Next, let $\chi : \mathbb{R}^3 \to \mathbb{R}$ be a smooth compactly supported spherically symmetric function of total mass 1. We integrate the above formula against $(1/R^3)\chi(x/R)$ for some large parameter R, and conclude after some integration by parts (which is justified by the compact support of χ and the smooth (and hence C^1) nature of all functions involved) that

$$\begin{aligned} R^{-3}\int_{\mathbb{R}^3} u(t_2, x)\chi\Big(\frac{x}{R}\Big)\, dx - R^{-3}\int_{\mathbb{R}^3} u(t_1, x)\chi\Big(\frac{x}{R}\Big)\, dx - R^{-4}\int_{t_1}^{t_2}\int_{\mathbb{R}^3} u(t, x)\big(u(t, x)\cdot\nabla\chi\big)\Big(\frac{x}{R}\Big)\, dx\, dt \\ = R^{-5}\int_{t_1}^{t_2}\int_{\mathbb{R}^3} u(t, x)(\Delta\chi)\Big(\frac{x}{R}\Big)\, dx\, dt + R^{-3}\int_{t_1}^{t_2}\int_{\mathbb{R}^3} \nabla p(t, x)\chi\Big(\frac{x}{R}\Big)\, dx\, dt \\ + R^{-3}\int_{t_1}^{t_2}\int_{\mathbb{R}^3} f(t, x)\chi\Big(\frac{x}{R}\Big)\, dx\, dt. \end{aligned}$$

From the finite energy hypothesis and the Cauchy–Schwarz inequality, one easily verifies that

$$\begin{aligned} \lim_{R\to\infty} R^{-3}\int_{\mathbb{R}^3} u(t_i, x)\chi\Big(\frac{x}{R}\Big)\, dx = 0, \\ \lim_{R\to\infty} R^{-4}\int_{t_1}^{t_2}\int_{\mathbb{R}^3} u(t, x)\big(u(t, x)\cdot\nabla\chi\big)\Big(\frac{x}{R}\Big)\, dx\, dt = 0, \\ \lim_{R\to\infty} R^{-5}\int_{t_1}^{t_2}\int_{\mathbb{R}^3} u(t, x)(\Delta\chi)\Big(\frac{x}{R}\Big)\, dx\, dt = 0, \\ \lim_{R\to\infty} R^{-3}\int_{t_1}^{t_2}\int_{\mathbb{R}^3} f(t, x)\chi\Big(\frac{x}{R}\Big)\, dx\, dt = 0, \end{aligned}$$

and thus

$$\lim_{R\to\infty} R^{-3}\int_{t_1}^{t_2}\int_{\mathbb{R}^3} \nabla p(t, x)\chi\Big(\frac{x}{R}\Big)\, dx\, dt = 0. \tag{43}$$

Next, by an integration by parts and (42), we can express

$$R^{-3}\int_{t_1}^{t_2}\int_{\mathbb{R}^3} \nabla p_0(t, x)\chi\Big(\frac{x}{R}\Big)\, dx\, dt$$

as

$$R^{-4}\int_{t_1}^{t_2}\int_{\mathbb{R}^3} u_i u_j(t,x)\big(\nabla\Delta^{-1}\partial_i\partial_j\chi\big)\Big(\frac{x}{R}\Big)\,dx\,dt + R^{-3}\int_{t_1}^{t_2}\int_{\mathbb{R}^3} f_i(t,x)\big(\nabla\Delta^{-1}\partial_i\chi\big)\Big(\frac{x}{R}\Big)\,dx\,dt.$$

From the finite energy nature of (u, p, u_0, f, T) we see that this expression goes to zero as $R \to \infty$. Subtracting this from (43), we conclude that

$$\lim_{R\to\infty} R^{-3}\int_{t_1}^{t_2}\int_{\mathbb{R}^3} \nabla h(t,x)\chi\Big(\frac{x}{R}\Big)\,dx\,dt = 0. \tag{44}$$

The function $x \mapsto \int_{t_1}^{t_2} \nabla h(t,x)$ is weakly harmonic, and hence harmonic. By the mean-value property of harmonic functions (and our choice of χ), we thus have

$$R^{-3}\int_{t_1}^{t_2}\int_{\mathbb{R}^3} \nabla h(t,x)\chi\Big(\frac{x}{R}\Big)\,dx\,dt = \int_{t_1}^{t_2} \nabla h(t,0)\,dt,$$

and thus

$$\int_{t_1}^{t_2} \nabla h(t,0)\,dt = 0.$$

Since t_1, t_2 were arbitrary, we conclude from the Lebesgue differentiation theorem that $\nabla h(t,0) = 0$ for almost every $t \in [0,T]$. Using spatial translation invariance (29) to replace the spatial origin by an element of a countable dense subset of $\mathbb{R}^3$, and using the fact that harmonic functions are continuous, we conclude that $\nabla h(t)$ is identically zero for almost every $t \in [0,T]$, and so $h(t)$ is constant for almost every t as desired. □

We note a useful corollary of Lemma 4.1(i):

Corollary 4.3 (Almost smooth H^1 solutions are essentially mild). *Let (u, p, u_0, f, T) be an almost smooth H^1 solution. Then $(u, \tilde{p}, u_0, f, T)$ is a mild H^1 solution, where*

$$\tilde{p}(t,x) := -\Delta^{-1}\partial_i\partial_j(u_i u_j)(t,x) + \Delta^{-1}\nabla\cdot f(t,x).$$

Furthermore, for almost every $t \in [0,T]$, $p(t)$ and $\tilde{p}(t)$ differ by a constant (and thus $\nabla p = \nabla\tilde{p}$).

Proof. By Lemma 4.1(i), ∇p is equal to $\nabla\tilde{p}$ almost everywhere; in particular, $\nabla p = \nabla\tilde{p}$ is a smooth tempered distribution. The claim then follows from (3) and the Duhamel formula (19). □

5. Local well-posedness theory in H^1

In this section we review the (subcritical) local well-posedness theory for both periodic and nonperiodic H^1 mild solutions. The material here is largely standard (and in most cases has been superseded by the more powerful critical well-posedness theory); for instance the uniqueness theory already follows from [Prodi 1959] and [Serrin 1963], the blowup criterion already is present in [Leray 1934], the local existence theory follows from [Kato and Ponce 1988], regularity of mild solutions follows from [Ladyzhenskaya 1967], the stability results given here follow from the stronger stability results of [Chemin and Gallagher 2009], and the compactness results were already essentially present in [Tao 2007]. However, for the convenience of the reader (and because we want to use the X^s function spaces defined in (13) as the basis for the theory) we shall present all this theory in a self-contained manner. There are now a number of advanced local well-posedness results at critical regularity, most notably that of [Koch and Tataru 2001], but we will not need such powerful results here.

We begin with the periodic theory. By taking advantage of the scaling symmetry (31), we may set the period L equal to 1. Using the symmetry (36), we may also restrict attention to data obeying the mean zero conditions (37), (38), and thus $u_0 \in H^1_x(\mathbb{R}^3/\mathbb{Z}^3)_0$ and $f \in L^\infty_t H^1_x([0,T] \times \mathbb{R}^3/\mathbb{Z}^3)_0$.

Theorem 5.1 (Local well-posedness in periodic H^1). *Let $(u_0, f, T, 1)$ be periodic H^1 data obeying the mean-zero conditions* (37), (38).

(i) (Strong solution). *If $(u, p, u_0, f, T, 1)$ is a periodic H^1 mild solution, then*

$$u \in C^0_t H^1_x([0,T] \times \mathbb{R}^3/\mathbb{Z}^3).$$

In particular, one can unambiguously define $u(t)$ in $H^1_x(\mathbb{R}^3/\mathbb{Z}^3)$ for each $t \in [0,T]$.

(ii) (Local existence). *If*

$$\left(\|u_0\|_{H^1_x(\mathbb{R}^3/\mathbb{Z}^3)} + \|f\|_{L^1_t H^1_x(\mathbb{R}^3/\mathbb{Z}^3)}\right)^4 T \le c \tag{45}$$

for a sufficiently small absolute constant $c > 0$, then there exists a periodic H^1 mild solution $(u, p, u_0, f, T, 1)$ with the indicated data with

$$\|u\|_{X^1([0,T]\times\mathbb{R}^3/\mathbb{Z}^3)} \lesssim \|u_0\|_{H^1_x(\mathbb{R}^3/\mathbb{Z}^3)} + \|f\|_{L^1_t H^1_x(\mathbb{R}^3/\mathbb{Z}^3)}$$

and more generally

$$\|u\|_{X^k([0,T]\times\mathbb{R}^3/\mathbb{Z}^3)} \lesssim_{k, T, \|u_0\|_{H^k_x(\mathbb{R}^3/\mathbb{Z}^3)}, \|f\|_{L^1_t H^k_x(\mathbb{R}^3/\mathbb{Z}^3)}} 1$$

for each $k \ge 1$. In particular, one has local existence whenever T is sufficiently small depending on $\mathcal{H}^1(u_0, f, T, 1)$.

(iii) (Uniqueness). *There is at most one periodic H^1 mild solution $(u, p, u_0, f, T, 1)$ with the indicated data.*

(iv) (Regularity). *If $(u, p, u_0, f, T, 1)$ is a periodic H^1 mild solution, and $(u_0, f, T, 1)$ is smooth, then $(u, p, u_0, f, T, 1)$ is smooth.*

(v) (Lipschitz stability). *Let $(u, p, u_0, f, T, 1)$ be a periodic H^1 mild solution with the bounds $0 < T \le T_0$ and*

$$\|u\|_{X^1([0,T]\times\mathbb{R}^3/\mathbb{Z}^3)} \le M.$$

Let $(u_0', f', T, 1)$ be another set of periodic H^1 data, and define the function

$$F(t) := e^{t\Delta}(u_0' - u_0) + \int_0^t e^{(t-t')\Delta}(f'(t') - f(t'))\, dt'.$$

If the quantity $\|F\|_{X^1([0,T]\times\mathbb{R}^3/\mathbb{Z}^3)}$ is sufficiently small depending on T, M, then there exists a periodic *mild solution* $(u', p', u'_0, f', T, 1)$ *with*

$$\|u - u'\|_{X^1([0,T]\times\mathbb{R}^3/\mathbb{Z}^3)} \lesssim_{T,M} \|F\|_{X^1([0,T]\times\mathbb{R}^3/\mathbb{Z}^3)}.$$

Proof. We first prove the strong solution claim (i). The linear solution

$$e^{t\Delta}u_0 + \int_0^t e^{(t-t')\Delta} Pf(t')\,dt'$$

is easily verified to lie in $C^0_t H^1_x([0, T] \times \mathbb{R}^3/\mathbb{Z}^3)$, so in view of (11), it suffices to show that

$$\int_0^t e^{(t-t')\Delta} PB(u(t'), u(t'))\,dt'$$

also lies in $C^0_t H^1_x([0, T] \times \mathbb{R}^3/\mathbb{Z}^3)$. But as u is an H^1 mild solution, u lies in $X^1([0, T] \times \mathbb{R}^3/\mathbb{Z}^3)$, so by (24), $PB(u, u)$ lies in $L^4_t L^2_x([0, T] \times \mathbb{R}^3/\mathbb{Z}^3)$. The claim (i) then follows easily from (22).

Now we establish local existence (ii). Let $\delta := \|u_0\|_{H^1_x(\mathbb{R}^3/\mathbb{Z}^3)} + \|f\|_{L^1_t H^1_x(\mathbb{R}^3/\mathbb{Z}^3)}$; thus by (45) we have $\delta^4 T \le c$. Using this and (25), (22), one easily establishes that the nonlinear map $u \mapsto \Phi(u)$ defined by

$$\Phi(u)(t) := e^{t\Delta}u_0 + \int_0^t e^{(t-t')\Delta} PB\big(u(t'), u(t') + Pf(t')\big)\,dt'$$

is a contraction on the ball

$$\big\{u \in X^1([0, T] \times \mathbb{R}^3/\mathbb{Z}^3) : \|u\|_{X^1([0,T]\times\mathbb{R}^3/\mathbb{Z}^3)} \le C\delta\big\}$$

if C is large enough. From the contraction mapping principle, we may then find a fixed point of Φ in this ball, and the claim (ii) follows (the estimates for higher k follow from variants of the above argument and an induction on k, and are left to the reader).

Now we establish uniqueness (iii). Suppose, in order to get a contradiction, that we have distinct solutions $(u, p, u_0, f, T, 1)$ and $(u', p', u_0, f, T, 1)$ for the same data. Then we have

$$\|u\|_{X^1([0,T]\times\mathbb{R}^3/\mathbb{Z}^3)},\ \|u'\|_{X^1([0,T]\times\mathbb{R}^3/\mathbb{Z}^3)} \le M.$$

To show uniqueness, it suffices to do so assuming that T is sufficiently small depending on M, as the general case then follows by subdividing $[0, T]$ into small enough time intervals and using induction. Subtracting (11) for u, u' and writing $v := u' - u$, we see that

$$v(t) = \int_0^t e^{(t-t')\Delta} P\big(2B(u(t'), v(t')) + B(v(t'), v(t'))\big)\,dt',$$

and thus by (22),

$$\|v\|_{X^1([0,T]\times\mathbb{R}^3/\mathbb{Z}^3)} \lesssim MT^{1/4}\|v\|_{X^1([0,T]\times\mathbb{R}^3/\mathbb{Z}^3)}.$$

If T is sufficiently small depending on M, this forces $\|v\|_{X^1([0,T]\times\mathbb{R}^3/\mathbb{Z}^3)} = 0$, giving uniqueness up to time T; iterating this argument gives the claim (iii).

Now we establish regularity (iv). To abbreviate the notation, all norms will be on $[0, T] \times \mathbb{R}^3/\mathbb{Z}^3$. As u is an H^1 mild solution, it lies in X^1, and hence by (25), $PB(u, u)$ lies in $L^4_t L^2_x$. Applying (11), (23), and the smoothness of u_0, f, we conclude that $u \in X^s$ for all $s < \frac{3}{2}$. In particular, by Sobolev embedding we see that

$u \in L_t^\infty L_x^{12}$, $\nabla u \in L_t^2 L_x^{12} \cap L_t^\infty L_x^{12/5}$, and $\nabla^2 u \in L_t^2 L_x^{12/5}$, and hence $PB(u, u) \in L_t^2 H_x^1([0, T] \times \mathbb{R}^3/\mathbb{Z}^3)$. Returning to (11), (23), we now conclude that $u \in X^2([0, T] \times \mathbb{R}^3/\mathbb{Z}^3)$. One can then repeat these arguments iteratively to conclude that $u \in X^k([0, T] \times \mathbb{R}^3/\mathbb{Z}^3)$ for all $k \geq 1$, and thus $u \in L_t^\infty C^k([0, T] \times \mathbb{R}^3/\mathbb{Z}^3)$ for all $k \geq 0$. From (9) we then have $p \in L^\infty C^k([0, T] \times \mathbb{R}^3/\mathbb{Z}^3)$ for all $k \geq 0$, and then from (3) we have $\partial_t u \in L_t^\infty C^k([0, T] \times \mathbb{R}^3/\mathbb{Z}^3)$ for all $k \geq 0$. One can then obtain bounds on $\partial_t p$ and then on higher time derivatives of u and t, giving the desired smoothness, and the claim (iv) follows.

Now we establish stability (v). It suffices to establish the claim in the short-time case when T is sufficiently small depending only on M (more precisely, we take $M^4 T \leq c$ for some sufficiently small absolute constant $c > 0$), as the long-time case then follows by subdividing time and using induction. The existence of the solution $(u', p', u_0', f_0', T, 1)$ is then guaranteed by (ii). Evaluating (11) for u, u' and subtracting, and setting $v := u' - u$, we see that

$$v(t) = F + \int_0^t e^{(t-t')\Delta} P(2B(u, v) + B(v, v))(t')\, dt'$$

for all $t \in [0, T]$. Applying (22), (25), we conclude that

$$\|v\|_{X^1} \lesssim \|F\|_{X^1} + T^{1/4}(\|u\|_{X^1} + \|v\|_{X^1})\|v\|_{X^1},$$

where all norms are over $[t_0, t_1] \times \mathbb{R}^3$. Since $\|u\|_{X^1} + \|v\|_{X^1}$ is finite, we conclude (if T is small enough) that $\|v\|_{X^1([0,T]\times\mathbb{R}^3/\mathbb{Z}^3)} \lesssim \|F\|_{X^1([0,T]\times\mathbb{R}^3/\mathbb{Z}^3)}$, and the claim follows. □

We may iterate the local well-posedness theory to obtain a dichotomy between existence and blowup. Define an *incomplete periodic mild H^1 solution* $(u, p, u_0, f, T_*^-, 1)$ from periodic H^1 data $(u_0, f, T_*, 1)$ to be fields $u : [0, T_*) \times \mathbb{R}^3/\mathbb{Z}^3 \to \mathbb{R}^3$ and $v : [0, T_*) \times \mathbb{R}^3/\mathbb{Z}^3 \to \mathbb{R}$ such that for any $0 < T < T_*$, the restriction $(u, p, u_0, f, T, 1)$ of $(u, p, u_0, f, T_*^-, 1)$ to the slab $[0, T] \times \mathbb{R}^3/\mathbb{Z}^3$ is a periodic mild H^1 solution. We similarly define the notion of an incomplete periodic smooth solution.

Corollary 5.2 (Maximal Cauchy development). *Let $(u_0, f, T, 1)$ be periodic H^1 data. Then at least one of the following two statements holds*:

- *There exists a periodic H^1 mild solution $(u, p, u_0, f, T, 1)$ with the given data.*
- *There exist a blowup time $0 < T_* < T$ and an incomplete periodic H^1 mild solution*

$$(u, p, u_0, f, T_*^-, 1)$$

up to time T_*^-, which blows up in H^1 in the sense that

$$\lim_{t \to T_*^-} \|u(t)\|_{H_x^1(\mathbb{R}^3/\mathbb{Z}^3)} = +\infty.$$

We refer to such solutions as maximal Cauchy developments.

A similar statement holds with "H^1 data" and "H^1 mild solution" replaced by "smooth data" and "smooth solution" respectively.

Next we establish a compactness property of the periodic H^1 flow.

Proposition 5.3 (Compactness). *If $(u_0^{(n)}, f^{(n)}, T, 1)$ is a sequence of periodic H^1 data obeying* (37), (38) *which is uniformly bounded in $H^1_x(\mathbb{R}^3/\mathbb{Z}^3)_0 \times L^\infty_t H^1_x([0,T]\times\mathbb{R}^3/\mathbb{Z}^3)_0$ and converges weakly*[14] *to $(u_0, f, T, 1)$, and $(u, p, u_0, f, T, 1)$ is a periodic H^1 mild solution with the indicated data, then for n sufficiently large, there exist periodic H^1 mild solutions $(u^{(n)}, p^{(n)}, u_0^{(n)}, f^{(n)}, T, 1)$ with the indicated data, with $u^{(n)}$ converging weakly in $X^1([0,T]\times\mathbb{R}^3/\mathbb{Z}^3)$ to u. Furthermore, for any $0 < \tau < T$, $u^{(n)}$ converges strongly in $X^1([\tau,T]\times\mathbb{R}^3/\mathbb{Z}^3)$ to u.*

If $u_0^{(n)}$ converges strongly in $H^1_x(\mathbb{R}^3/\mathbb{Z}^3)_0$ to u_0, then one can set $\tau = 0$ in the previous claim.

Proof. This result is essentially in [Tao 2007, Proposition 2.2], but for the convenience of the reader we give a full proof here.

To begin with, we assume that $u^{(n)}$ converges strongly in $H^1_x(\mathbb{R}^3/\mathbb{Z}^3)_0$ to u_0, and relax this to weak convergence later. In view of the stability component of Theorem 5.1, it suffices to show that $F^{(n)}$ converges strongly in $X^1([0,T]\times\mathbb{R}^3/\mathbb{Z}^3)$ to zero, where

$$F^{(n)}(t) := e^{t\Delta}(u_0^{(n)} - u_0) + \int_0^t e^{(t-t')\Delta} P(f^{(n)}(t') - f(t'))\,dt'.$$

We have that $u_0^{(n)} - u_0$ converges strongly in $H^1_x(\mathbb{R}^3/\mathbb{Z}^3)$ to zero, while $f^{(n)} - f$ converges weakly in $L^\infty_t H^1_x([0,T]\times\mathbb{R}^3/\mathbb{Z}^3) \to 0$, and hence strongly in $L^2_t L^2_x([0,T]\times\mathbb{R}^3/\mathbb{Z}^3)$. The claim then follows from (22).

Now we only assume that $u^{(n)}$ converges weakly in $H^1_x(\mathbb{R}^3/\mathbb{Z}^3)_0$ to u_0. Let $0 < \tau < T$ be a sufficiently small time; then from local existence (Theorem 5.1(ii)) we see that $u^{(n)}$ and u are bounded in $X^1([0,\tau]\times\mathbb{R}^3/\mathbb{Z}^3)$ uniformly in n by some finite quantity M. Writing $v^{(n)} := u^{(n)} - u$, we obtain from (11) the difference equation

$$v^{(n)}(t) = F^{(n)}(t) + \int_0^t e^{(t-t')\Delta} P\big(B(u, v^{(n)}) + B(u^{(n)}, v^{(n)})\big)(t')\,dt'.$$

Since $u_0^{(n)} - u_0$ converges weakly in $H^1_x(\mathbb{R}^3/\mathbb{Z}^3)$ to zero, it converges strongly in $L^2_x(\mathbb{R}^3/\mathbb{Z}^3)$ to zero too. Using (21) as before, we see that $F^{(n)}$ converges strongly in $X^0([0,\tau]\times\mathbb{R}^3/\mathbb{Z}^3)$ to zero. From (22) we thus have

$$\|v^{(n)}\|_{X^0} \lesssim o(1) + \|B(u, v^{(n)})\|_{L^2_t H^{-1}_x} + \|B(u^{(n)}, v^{(n)})\|_{L^2_t H^{-1}_x},$$

where $o(1)$ goes to zero as $n \to \infty$, and all space-time norms are over $[0,\tau]\times\mathbb{R}^3/\mathbb{Z}^3$. From the form of B and Hölder's inequality, we have

$$\|B(u^{(n)}, v^{(n)})\|_{L^2_t H^{-1}_x} \lesssim \|\mathcal{O}(u^{(n)} v^{(n)})\|_{L^2_t L^2_x} \lesssim \tau^{1/4} \|u^{(n)}\|_{L^\infty_t L^6_x} \|v^{(n)}\|^{1/2}_{L^\infty_t L^2_x} \|v^{(n)}\|^{1/2}_{L^2_t L^6_x} \lesssim M\tau^{1/4}\|v^{(n)}\|_{X^0},$$

and similarly for $B(u, v^{(n)})$, and thus

$$\|v^{(n)}\|_{X^0} \lesssim o(1) + M\tau^{1/4}\|v^{(n)}\|_{X^0}.$$

[14] Strictly speaking, we should use "converges in the weak-* sense" or "converges in the sense of distributions" here, in order to avoid the pathological (and irrelevant) elements of the dual space of $L^\infty_t H^1_x$ that can be constructed from the axiom of choice.

Thus, for τ small enough, one has

$$\|v^{(n)}\|_{X^0} = o(1),$$

which among other things gives weak convergence of $u^{(n)}$ to u in $[0, \tau] \times \mathbb{R}^3/\mathbb{Z}^3$. Also, by the pigeonhole principle, one can find times $\tau^{(n)}$ in $[0, \tau]$ such that

$$\|v^{(n)}(\tau^{(n)})\|_{H^1_x(\mathbb{R}^3/\mathbb{Z}^3)} = o(1).$$

Using the stability theory, and recalling that τ is small, this implies that

$$\|v^{(n)}(\tau)\|_{H^1_x(\mathbb{R}^3/\mathbb{Z}^3)} = o(1);$$

thus $u^{(n)}(\tau)$ converges strongly to $u(\tau)$. Now we can use our previous arguments to extend $u^{(n)}$ to all of $[0, T] \times \mathbb{R}^3/\mathbb{Z}^3$ and obtain strong convergence in $X^1([\tau, T] \times \mathbb{R}^3/\mathbb{Z}^3)$, as desired. □

Now we turn to the nonperiodic setting. We have the following analogue of Theorem 5.1:

Theorem 5.4 (Local well-posedness in H^1). *Let (u_0, f, T) be H^1 data.*

(i) (Strong solution). *If $(u, p, u_0, f, T, 1)$ is an H^1 mild solution, then*

$$u \in C^0_t H^1_x([0, T] \times \mathbb{R}^3).$$

(ii) (Local existence and regularity). *If*

$$\left(\|u_0\|_{H^1_x(\mathbb{R}^3)} + \|f\|_{L^1_t H^1_x(\mathbb{R}^3)}\right)^4 T \leq c \tag{46}$$

for a sufficiently small absolute constant $c > 0$, then there exists a H^1 mild solution (u, p, u_0, f, T) with the indicated data, with

$$\|u\|_{X^1([0,T]\times\mathbb{R}^3)} \lesssim \|u_0\|_{H^1_x(\mathbb{R}^3)} + \|f\|_{L^1_t H^1_x(\mathbb{R}^3)},$$

and more generally

$$\|u\|_{X^k([0,T]\times\mathbb{R}^3)} \lesssim_{k, \|u_0\|_{H^k_x(\mathbb{R}^3)}, \|f\|_{L^1_t H^k_x(\mathbb{R}^3)}, 1}$$

for each $k \geq 1$. In particular, one has local existence whenever T is sufficiently small depending on $\mathcal{H}^1(u_0, f, T)$.

(iii) (Uniqueness). *There is at most one H^1 mild solution (u, p, u_0, f, T) with the indicated data.*

(iv) (Regularity). *If $(u, p, u_0, f, T, 1)$ is an H^1 mild solution, and (u_0, f, T) is Schwartz, then u and p are smooth; in fact, one has $\partial_t^j u, \partial_t^j p \in L^\infty_t H^k([0, T] \times \mathbb{R}^3)$ for all $j, k \geq 0$.*

(v) (Lipschitz stability). *Let (u, p, u_0, f, T), (u', p', u'_0, f', T) be H^1 mild solutions with the bounds $0 < T \leq T_0$ and*

$$\|u\|_{X^1([0,T]\times\mathbb{R}^3)}, \|u'\|_{X^1([0,T]\times\mathbb{R}^3)} \leq M.$$

Define the function

$$F(t) := e^{t\Delta}(u'_0 - u_0) + \int_0^t e^{(t-t')\Delta}(f'(t') - f(t'))\, dt'.$$

If the quantity $\|F\|_{L^2_t L^2_x([0,T]\times\mathbb{R}^3)}$ *is sufficiently small depending on* T, M, *then*

$$\|u - u'\|_{X^1([0,T]\times\mathbb{R}^3)} \lesssim_{T,M} \|F\|_{L^2_t L^2_x([0,T]\times\mathbb{R}^3)}.$$

Proof. This proceeds by repeating the proof of Theorem 5.1 verbatim. The one item which perhaps requires some care is the regularity item (iv). The arguments from Theorem 5.1 yield the regularity

$$u \in X^k([0, T] \times \mathbb{R}^3)$$

for all $k \geq 0$ without difficulty. In particular, $u \in L^\infty_t H^k_x([0, T] \times \mathbb{R}^3)$ for all $k \geq 0$. From (9) and Sobolev embedding, one then has $p \in L^\infty_t H^k_x([0, T] \times \mathbb{R}^3)$ for all $k \geq 0$, and then from (3) and more Sobolev embedding, one has $\partial_t u \in L^\infty_t H^k_x([0, T] \times \mathbb{R}^3)$ for all $k \geq 0$. One can then obtain bounds on $\partial_t p$ and then on higher time derivatives of u and t, giving the desired smoothness, and the claim (iv) follows. (Note that these arguments did not require the full power of the hypothesis that (u_0, f, T) was Schwartz; it would have sufficed to have $u_0 \in H^k_x(\mathbb{R}^3)$ and $f \in C^j_t H^k_x(\mathbb{R}^3)$ for all $j, k \geq 0$.) $\square$

From the regularity component of the above theorem, we immediately conclude that Conjecture 1.19 implies Conjecture 1.5, which is one half of Theorem 1.20(iv).

We will also need a more quantitative version of the regularity statement in Theorem 5.4.

Lemma 5.5 (Quantitative regularity). *Let* (u, p, u_0, f, T) *be an* H^1 *mild solution obeying* (46) *for a sufficiently small absolute constant* $c > 0$, *and such that*

$$\|u_0\|_{H^1_x(\mathbb{R}^3)} + \|f\|_{L^1_t H^k_x(\mathbb{R}^3)} \leq M < \infty.$$

Then one has

$$\|u\|_{L^\infty_t H^k_x([\tau,T]\times\mathbb{R}^3)} \lesssim_{k,\tau,T,M} 1$$

for all natural numbers $k \geq 1$ *and all* $0 < \tau < T$.

Proof. We allow all implied constants to depend on k, T, M. From Theorem 5.1 we have

$$\|u\|_{X^1([0,T]\times\mathbb{R}^3)} \lesssim 1,$$

which already gives the $k = 1$ case. Now we turn to the $k \geq 2$ case. From (25) we have

$$\|PB(u, u)\|_{L^4_t L^2_x([0,T]\times\mathbb{R}^3)} \lesssim 1,$$

while from Fourier analysis one has

$$\|e^{t\Delta} u_0\|_{L^\infty_t H^k_x([\tau,T]\times\mathbb{R}^3)} \lesssim_\tau 1.$$

From this and (11), (21) we see that

$$\|u\|_{X^s([\tau,T]\times\mathbb{R}^3)} \lesssim_{s,\tau} 1$$

for all $s < \frac{3}{2}$. From Sobolev embedding we conclude

$$\|u\|_{L_t^\infty L_x^{12}([\tau,T]\times\mathbb{R}^3)} \lesssim_\tau 1,$$
$$\|\nabla u\|_{L_t^2 L_x^{12}([\tau,T]\times\mathbb{R}^3)} \lesssim_\tau 1,$$
$$\|\nabla u\|_{L_t^\infty L_x^{12/5}([\tau,T]\times\mathbb{R}^3)} \lesssim_\tau 1,$$
$$\|\nabla^2 u\|_{L_t^2 L_x^{12}([\tau,T]\times\mathbb{R}^3)} \lesssim_\tau 1,$$

and hence

$$\|PB(u,u)\|_{L_t^2 H_x^1([\tau,T]\times\mathbb{R}^3)} \lesssim 1.$$

Returning to (11), (23), we now conclude that

$$\|u\|_{X^2([\tau,T]\times\mathbb{R}^3)} \lesssim_\tau 1,$$

which gives the $k = 2$ case. One can repeat these arguments iteratively to then give the higher k cases. □

We extract a particular consequence of the above lemma:

Proposition 5.6 (Almost regularity). *Let $(u, p, u_0, 0, T)$ be a homogeneous H^1 mild solution obeying* (46) *for a sufficiently small absolute constant $c > 0$. Then u, p are smooth on $[\tau, T] \times \mathbb{R}^3$ for all $0 < \tau < T$; in fact, all derivatives of u, p lie in $L_t^\infty L_x^2([\tau, T] \times \mathbb{R}^3)$. If furthermore u_0 is smooth, then $(u, p, u_0, 0, T)$ is an almost smooth solution.*

Proof. From Lemma 5.5 we see that

$$u \in L_t^\infty H_x^k([\tau, T] \times \mathbb{R}^3)$$

for all $k \geq 0$ and $0 < \tau < T$. Arguing as in the proof of Theorem 5.4(iv), we conclude that u, p are smooth on $[\tau, T] \times \mathbb{R}^3$.

Now suppose that u_0 is smooth. Then (since u_0 is also in $H_x^1(\mathbb{R}^3)$) $e^{t\Delta}u_0$ is smooth[15] on $[0, T] \times \mathbb{R}^3$, and in particular one has

$$\eta e^{t\Delta}u_0 \in L_t^\infty H_x^k([0, T] \times \mathbb{R}^3)$$

for any smooth, compactly supported cutoff function $\eta : \mathbb{R}^3 \to \mathbb{R}$. Meanwhile, by arguing as in Lemma 5.5 one has

$$PB(u, u) \in L_t^4 L_x^2([0, T] \times \mathbb{R}^3). \tag{47}$$

Using (11), (21), one concludes that

$$\eta u \in X^s([0, T] \times \mathbb{R}^3)$$

for all cutoff functions η and all $s < \frac{3}{2}$. Continuing the arguments from Lemma 5.5, we conclude that

$$\eta PB(u, u) \in L_t^2 H_x^1([0, T] \times \mathbb{R}^3)$$

[15] To obtain smoothness at a point (t_0, x_0), one can for instance split u_0 into a smooth compactly supported component and a component that vanishes near x_0 but lies in $H_x^1(\mathbb{R}^3)$, and verify that the contribution of each component to $e^{t\Delta}u_0$ is smooth at (t_0, x_0).

for all cutoffs η. Using (11), (23) (and using (28), (47) to deal with the far field contribution of $PB(u, u)$, and shrinking η as necessary), one then concludes that

$$\eta u \in X^2([0, T] \times \mathbb{R}^3)$$

for all cutoffs η. Repeating these arguments iteratively, one eventually concludes that

$$\eta u \in X^k([0, T] \times \mathbb{R}^3)$$

for all cutoffs η, and in particular

$$u \in L_t^\infty H_x^k([0, T] \times K)$$

for all $k \geq 0$ and all compact sets K. By Sobolev embedding, this implies that

$$u \in L_t^\infty C_x^k([0, T] \times K)$$

for all $k \geq 0$ and all compact sets K.

We also have $u \in X^1([0, T] \times \mathbb{R}^3)$, and hence

$$u \in L_t^\infty H_x^1([0, T] \times \mathbb{R}^3).$$

In particular,

$$u_i u_j \in L_t^\infty L_x^1([0, T] \times \mathbb{R}^3) \tag{48}$$

and

$$u_i u_j \in L_t^\infty C_x^k([0, T] \times K)$$

for all $k \geq 0$ and compact K. From this and (9) (splitting the inverse Laplacian Δ^{-1} smoothly into local and global components), one has

$$p \in L_t^\infty C_x^k([0, T] \times K);$$

inserting this into (3), we then see that

$$\partial_t u \in L_t^\infty C_x^k([0, T] \times K) \tag{49}$$

for all $k \geq 0$ and compact K.

This is a little weaker than what we need for an almost smooth solution, because we want $\nabla^k u$, $\nabla^k p$, $\partial_t \nabla^k p$ to extend continuously down to $t = 0$, and the above estimates merely give $L_t^\infty C_x^\infty$ control on these quantities. To upgrade the L_t^∞ control to continuity in time, we first observe[16] from (49) and integration in time that we can at least make $\nabla^k u$ extend continuously to $t = 0$:

$$u \in C_t^0 C_x^k([0, T] \times K).$$

In particular,

$$u_i u_j \in C_t^0 C_x^k([0, T] \times K) \tag{50}$$

[16] An alternate argument here would be to approximate the initial data u_0 by Schwartz divergence-free data (using Lemma 12.1) and to use a limiting argument and the stability and regularity theory in Theorem 5.1; we omit the details.

for all $k \geq 0$ and compact K.

Now we consider $\nabla^k p$ in a compact region $[0, T] \times K$. From (9) we have

$$\nabla^k p(t, x) = \nabla^k \partial_i \partial_j \int_{\mathbb{R}^3} \frac{1}{4\pi |x-y|} u_i u_j(t, y)\, dy.$$

Using a smooth cutoff, we split the Newton potential $1/(4\pi|x-y|)$ into a "local" portion supported on $B(0, 2R)$ and a "global" portion supported outside of $B(0, R)$, where R is a large radius. From (50) one can verify that the contribution of the local portion is continuous on $[0, T] \times K$, while from (48) the contribution of the global portion is $O_u(1/R^3)$. Sending $R \to \infty$, we conclude that $\nabla^k p$ is continuous on $[0, T] \times K$, and thus

$$p \in C^0_t C^k_x([0, T] \times K)$$

for all $k \geq 0$ and compact K. Inserting this into (3), we then conclude that

$$\partial_t u \in C^0_t C^k_x([0, T] \times K)$$

for all $k \geq 0$ and compact K, and so we have an almost smooth solution as required. □

Remark 5.7. Because u has the regularity of $L^\infty_t H^1_x$, we can continue iterating the above argument a little more, and eventually get $u \in C^2_t C^k_x([0, T] \times K)$ and $p \in C^1_t C^k_x([0, T] \times K)$ for all $k \geq 0$ and compact K. Using the vorticity equation (see (84) below), one can then also get $\omega \in C^3_t C^k_x([0, T] \times K)$ as well. But without further decay conditions on higher derivatives of u (or of ω), one cannot gain infinite regularity on u, p, ω in time; see Section 15.

On the other hand, it is possible to use energy methods and the vorticity equation (84) to show (working in the homogeneous case $f = 0$ for simplicity) that if u_0 is smooth and the initial vorticity $\omega_0 := \nabla \times u_0$ is Schwartz, then the solution in Proposition 5.6 is in fact smooth, with ω remaining Schwartz throughout the lifespan of that solution; we omit the details.

As a corollary of the above proposition we see that Conjecture 1.19 implies Conjecture 1.13, thus completing the proof of Theorem 1.20(iv).

As before, we obtain a dichotomy between existence and blowup. Define an *incomplete mild H^1 solution* (u, p, u_0, f, T_*^-) from H^1 data (u_0, f, T_*) to be fields $u : [0, T_*) \times \mathbb{R}^3 \to \mathbb{R}^3$ and $v : [0, T_*) \times \mathbb{R}^3 \to \mathbb{R}$ such that for any $0 < T < T_*$, the restriction $(u, p, u_0, f, T, 1)$ of $(u, p, u_0, f, T_*^-, 1)$ to the slab $[0, T] \times \mathbb{R}^3$ is a mild H^1 solution. We similarly define the notion of an incomplete smooth H^1 solution.

Corollary 5.8 (Maximal Cauchy development). *Let (u_0, f, T) be H^1 data. Then at least one of the following two statements holds*:

- *There exists a mild H^1 solution (u, p, u_0, f, T) with the given data.*
- *There exist a blowup time $0 < T_* < T$ and an incomplete mild H^1 solution (u, p, u_0, f, T_*^-) up to time T_*^- that blows up in the enstrophy norm in the sense that*

$$\lim_{t \to T_*^-} \|u(t)\|_{H^1_x(\mathbb{R}^3)} = +\infty.$$

Remark 5.9. In the second conclusion of Corollary 5.8, more information about the blowup is known. For instance, in [Iskauriaza et al. 2003] it was demonstrated that the $L^3_x(\mathbb{R}^3)$ norm must also blow up (in the homogeneous case $f = 0$, at least).

6. Homogenisation

In this section we prove Proposition 1.7.

Fix smooth periodic data (u_0, f, T, L); our objective is to find a smooth periodic solution

$$(u, p, u_0, f, T, L)$$

(without pressure normalisation) with this data. By the scaling symmetry (31), we may normalise the period L to equal 1. Using the symmetry (36), we may impose the mean-zero conditions (37), (38) on this data.

By hypothesis, one can find a smooth periodic solution $(\tilde{u}, \tilde{p}, u_0, 0, T, 1)$ with data $(u_0, 0, T, 1)$. By Lemma 4.1, and applying a Galilean transform (33) if necessary, we may assume the pressure is normalised, which in particular makes $(\tilde{u}, \tilde{p}, u_0, 0, T, 1)$ a periodic H^1 mild solution.

By the Galilean invariance (33) (with a linearly growing velocity $v(t) := 2wt$), it suffices to find a smooth periodic solution (u, p, u_0, f_w, T) (this time *with* pressure normalisation) for the Galilean-shifted data (u_0, f_w, T), where

$$f_w(t, x) := f(t, x - wt^2),$$

and $w \in \mathbb{R}^3$ is arbitrary. Note that the data (u_0, f_w, T) continues to obey the mean-zero conditions (37), (38) and is bounded in $H^1_x(\mathbb{R}^3/\mathbb{Z}^3)_0 \times L^\infty_t H^1_x([0, T] \times \mathbb{R}^3/\mathbb{Z}^3)_0$ uniformly in w. We now make a key observation:

Lemma 6.1. *If $\alpha \in \mathbb{R}^3/\mathbb{Z}^3$ is* irrational *in the sense that $k \cdot \alpha \neq 0$ in $\mathbb{R}/\mathbb{Z}$ for all $k \in \mathbb{Z}^3 \backslash \{0\}$, then $f_{\lambda\alpha}$ converges weakly (or more precisely, converges in the sense of space-time distributions) to zero in $L^\infty_t H^1_x([0, T] \times \mathbb{R}^3/\mathbb{Z}^3)_0$.*

Proof. It suffices to show that

$$\int_0^T \int_{\mathbb{R}^3/\mathbb{Z}^3} f_{\lambda\alpha}(t, x)\phi(t, x)\, dx\, dt \to 0$$

for all smooth functions $\phi : [0, T] \times \mathbb{R}^3/\mathbb{Z}^3 \to \mathbb{R}$. Taking the Fourier transform, the left-hand side becomes

$$\sum_{k \in \mathbb{Z}^3} \int_0^T e^{-2\pi i \lambda k t^2 \cdot \alpha}\, \widehat{f(t)}(k)\widehat{\phi(t)}(-k)\, dt,$$

with the sum being absolutely convergent due to the rapid decrease of the Fourier transform of $\phi(t)$. Because f has mean zero, we can delete the $k = 0$ term from the sum. This makes $k \cdot \alpha$ nonzero by irrationality, and so by the Riemann–Lebesgue lemma, each summand goes to zero as $\lambda \to \infty$. The claim then follows from the dominated convergence theorem. $\square$

Let $\alpha \in \mathbb{R}^3/\mathbb{Z}^3$ be irrational. By the above lemma, $(u_0, f_{\lambda\alpha}, T, 1)$ converges weakly to $(u_0, 0, T, 1)$ while being bounded in $H^1_x(\mathbb{R}^3/\mathbb{Z}^3)_0 \times L^\infty_t H^1_x(\mathbb{R}^3/\mathbb{Z}^3)_0$. As $(u_0, 0, T, 1)$ has a periodic mild H^1 solution $(\tilde{u}, \tilde{p}, u_0, 0, T, 1)$, we conclude from Proposition 5.3 that for λ sufficiently large, $(u_0, f_{\lambda\alpha}, T, 1)$ also has a periodic mild H^1 solution, which is necessarily smooth since u_0 and $f_{\lambda\alpha}$ are smooth. The claim follows.

Remark 6.2. Suppose that $(u_0, f, \infty, 1)$ is periodic H^1 data extending over the half-infinite time interval $[0, +\infty)$. The above argument shows (assuming Conjecture 1.4) that one can, for each $0 < T < \infty$, construct a smooth periodic (but not pressure-normalised) solution $(u^{(T)}, p^{(T)}, u_0, f, T, 1)$ up to time T with the above data, by choosing a sufficiently rapidly growing linear velocity $v^{(T)} = 2w^{(T)}t$, applying a Galilean transform, and then using the compactness properties of the H^1 local well-posedness theory. As stated, this argument gives a different solution $(u^{(T)}, p^{(T)}, u_0, f, T, 1)$ for each time T (note that we do not have uniqueness once we abandon pressure normalisation). However, it is possible to modify the argument to obtain a single global smooth periodic solution $(u, p, u_0, f, \infty, 1)$ (which is still not pressure-normalised, of course), by using the ability in (33) to choose a *nonlinear* velocity $v(t)$ rather than a linear one. By reworking the above argument, and taking $v(t)$ to be a sufficiently rapidly growing function of t, it is then possible to obtain a global smooth periodic solution $(u, p, u_0, f, \infty, 1)$ to the indicated data; we omit the details.

7. Compactness

In this section we prove Theorem 1.20(i) by following the compactness arguments of [Tao 2007]. By the scaling symmetry (31), we may normalise $L = 1$.

We first assume that Conjecture 1.10 holds, and deduce Conjecture 1.9. Suppose for contradiction that Conjecture 1.9 failed. By Corollary 5.2, there thus exists an incomplete periodic pressure-normalised mild H^1 solution $(u, p, u_0, f, T_*^-, 1)$ such that

$$\lim_{t \to T_*^-} \|u(t)\|_{H^1_x(\mathbb{R}^3/\mathbb{Z}^3)} = \infty. \tag{51}$$

By Galilean invariance (36), we may assume that u_0 and f (and hence u) have mean zero.

Let $(u_0^{(n)}, f^{(n)}, T_*, 1)$ be a sequence of periodic smooth mean-zero data converging strongly in

$$H^1_x(\mathbb{R}^3/\mathbb{Z}^3)_0 \times L^\infty_t H^1_x([0, T_*] \times \mathbb{R}^3/\mathbb{Z}^3)_0$$

to the periodic H^1 data $(u, f, T_*, 1)$. For each time $0 < T < T_*$, we see from Theorem 5.1 that for n sufficiently large, we may find a smooth solution $(u^{(n)}, p^{(n)}, u_0^{(n)}, T, 1)$ with this data, with $u^{(n)}$ converging strongly in $L^\infty_t H^1_x([0, T] \times \mathbb{R}^3/\mathbb{Z}^3)$ to u. By Conjecture 1.10, the $L^\infty_t H^1_x([0, T] \times \mathbb{R}^3/\mathbb{Z}^3)$ norm of $u^{(n)}$ is bounded uniformly in both T and n, so by taking limits as $n \to \infty$, we conclude that $\|u(t)\|_{H^1_x(\mathbb{R}^3/\mathbb{Z}^3)}$ is bounded uniformly for $0 \le t < T_*$, contradicting (51) as desired.

Conversely, suppose that Conjecture 1.9 held, but Conjecture 1.10 failed. Carefully negating all the quantifiers, we conclude that there exists a time $0 < T_0 < \infty$ and a sequence $(u^{(n)}, p^{(n)}, u_0^{(n)}, f^{(n)}, T^{(n)}, 1)$ of smooth periodic data with $0 < T^{(n)} < T_0$ and $\mathcal{H}^1(u_0^{(n)}, f^{(n)}, T^{(n)}, 1)$ uniformly bounded in n, such that

$$\lim_{n \to \infty} \|u\|_{L^\infty_t H^1_x([0,T^{(n)}] \times \mathbb{R}^3/\mathbb{Z}^3)} = \infty. \tag{52}$$

Using Galilean transforms (36), we may assume that $u_0^{(n)}$, $f^{(n)}$ (and hence $u^{(n)}$) have mean zero. From the short-time local existence (and uniqueness) theory in Theorem 5.1, we see that $T^{(n)}$ is bounded uniformly away from zero. Thus by passing to a subsequence, we may assume that $T^{(n)}$ converges to a limit T_* with $0 < T_* \le T_0$.

By sequential weak compactness, we may pass to a further subsequence and assume that for each $0 < T < T_*$, $(u_0^{(n)}, f^{(n)}, T, 1)$ converges weakly (or more precisely, in the sense of distributions) to a periodic H^1 limit $(u_0, f, T, 1)$; gluing these limits together, one obtains periodic H^1 data $(u_0, f, T_*, 1)$, which still has mean zero. By Conjecture 1.9, we can then find a periodic H^1 mild solution $(u, p, u_0, f, T_*, 1)$ with this data, which then necessarily also has mean zero.

By Theorem 5.1 and Proposition 5.3, we see that for every $0 < \tau < T < T_*$, $u^{(n)}$ converges strongly in $L_t^\infty H_x^1([\tau, T] \times \mathbb{R}^3/\mathbb{Z}^3)$ to u. In particular, for any $0 < T < T_*$, one has

$$\limsup_{n\to\infty} \|u^{(n)}(T)\|_{H_x^1(\mathbb{R}^3/\mathbb{Z}^3)} \le \|u\|_{L_t^\infty H_x^1([0,T_*]\times\mathbb{R}^3/\mathbb{Z}^3)} < \infty.$$

Taking T sufficiently close to T_* and then taking n sufficiently large, we conclude from Theorem 5.1 that

$$\limsup_{n\to\infty} \|u^{(n)}\|_{L_t^\infty H_x^1([T,T^{(n)}]\times\mathbb{R}^3/\mathbb{Z}^3)} < \infty;$$

also, from the strong convergence in $L_t^\infty H_x^1([\tau, T] \times \mathbb{R}^3/\mathbb{Z}^3)$, we have

$$\limsup_{n\to\infty} \|u^{(n)}\|_{L_t^\infty H_x^1([\tau,T]\times\mathbb{R}^3/\mathbb{Z}^3)} < \infty$$

for any $0 < \tau < T$, and finally from the local existence (and uniqueness) theory in Theorem 5.1, one has

$$\limsup_{n\to\infty} \|u^{(n)}\|_{L_t^\infty H_x^1([0,\tau]\times\mathbb{R}^3/\mathbb{Z}^3)} < \infty$$

for sufficiently small τ. Putting these bounds together, we contradict (52), and the claim follows.

Remark 7.1. It should be clear to the experts that one could have replaced the H^1 regularity in the above conjectures by other subcritical regularities, such as H^k for $k > 1$, and obtained a similar result to Theorem 1.20(i).

As remarked previously, the homogeneous case $f = 0$ of Theorem 1.20(i) was established in [Tao 2007]. We recall the main results of that paper. We introduce the following homogeneous periodic conjectures:

Conjecture 7.2 (*A priori* homogeneous periodic H^1 bound). There exists a function $F : \mathbb{R}^+ \times \mathbb{R}^+ \times \mathbb{R}^+ \to \mathbb{R}^+$ with the property that whenever $(u, p, u_0, 0, T, L)$ is a smooth periodic homogeneous normalised-pressure solution with $0 < T < T_0 < \infty$ and

$$\mathcal{H}^1(u_0, 0, T, L) \le A < \infty,$$

then

$$\|u\|_{L_t^\infty H_x^1([0,T]\times\mathbb{R}^3/L\mathbb{Z}^3)} \le F(A, L, T_0).$$

Conjecture 7.3 (*A priori* homogeneous global periodic H^1 bound). There exists a function

$$F : \mathbb{R}^+ \times \mathbb{R}^+ \to \mathbb{R}^+$$

with the property that whenever $(u, p, u_0, 0, T, L)$ is a smooth periodic homogeneous normalised-pressure solution with

$$\mathcal{H}^1(u_0, 0, T, L) \leq A < \infty,$$

then

$$\|u\|_{L_t^\infty H_x^1([0,T]\times\mathbb{R}^3/L\mathbb{Z}^3)} \leq F(A, L).$$

Conjecture 7.4 (Global well-posedness in periodic homogeneous H^1). Let $(u_0, 0, T, L)$ be a homogeneous periodic H^1 set of data. Then there exists a periodic H^1 mild solution $(u, p, u_0, 0, T, L)$ with the indicated data.

Conjecture 7.5 (Global regularity for homogeneous periodic data with normalised pressure). Suppose $(u_0, 0, T)$ is a smooth periodic set of data. Then there exists a smooth periodic solution $(u, p, u_0, 0, T)$ with the indicated data and with normalised pressure.

In [Tao 2007, Theorem 1.4] it was shown that Conjectures 1.4, 7.2, 7.3 are equivalent. As implicitly observed in that paper also, Conjecture 1.4 is equivalent to Conjecture 7.5 (this can be seen from Lemma 4.1), and from the local well-posedness and regularity theory (Theorem 5.1 or [Tao 2007, Proposition 2.2]), we also see that Conjecture 7.5 is equivalent to Conjecture 7.4.

8. Energy localisation

In this section we establish the energy inequality for the Navier–Stokes equation in the smooth finite energy setting. This energy inequality is utterly standard (see for example [Scheffer 1976]) for weaker notions of solutions, so long as one has regularity of $L_t^2 H_x^1$, but (somewhat ironically) requires more care in the smooth finite energy setting, because we do *not* assume *a priori* that smooth finite energy solutions lie in the space $L_t^2 H_x^1$. The methods used here are local in nature, and will also provide an energy localisation estimate for the Navier–Stokes equation (see Theorem 8.2).

We begin with the global energy inequality.

Lemma 8.1 (Global energy inequality). *Let (u, p, u_0, f, T) be a finite energy almost smooth solution. Then*

$$\|u\|_{L_t^\infty L_x^2([0,T]\times\mathbb{R}^3)} + \|\nabla u\|_{L_t^2 L_x^2([0,T]\times\mathbb{R}^3)} \lesssim E(u_0, f, T)^{1/2}. \tag{53}$$

In particular, u lies in the space $X^1([0, T] \times \mathbb{R}^3)$.

Proof. To abbreviate the notation, all spatial norms here will be over $\mathbb{R}^3$.

Using the forcing symmetry (34), we may set f to be divergence-free, so in particular by Corollary 4.3 we have

$$\nabla p(t) = \nabla \tilde{p}(t) \tag{54}$$

for almost all times t, where

$$\tilde{p} = -\Delta^{-1}\partial_i\partial_j(u_iu_j). \tag{55}$$

As (u, p, u_0, f, T) is finite energy, we have the *a priori* hypothesis

$$\|u\|_{L_t^\infty L_x^2([0,T]\times\mathbb{R}^3)} \leq A$$

for some $A < \infty$, though recall that our final bounds are not allowed to depend on this quantity A. Because u is smooth, we see in particular from Fatou's lemma that

$$\|u(t)\|_{L_x^2} \leq A \tag{56}$$

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

Taking the inner product of the Navier–Stokes equation (3) with u and rearranging, we obtain the energy density identity

$$\partial_t\big(\tfrac{1}{2}|u|^2\big) + u\cdot\nabla\big(\tfrac{1}{2}|u|^2\big) = \Delta\big(\tfrac{1}{2}|u|^2\big) - |\nabla u|^2 - u\cdot\nabla p + u\cdot f. \tag{57}$$

We would like to integrate this identity over all of $\mathbb{R}^3$, but we do not yet have enough decay in space to achieve this, even with the normalised pressure. Instead, we will localise by integrating the identity against a cutoff η^4, where $\eta(x) := \chi((|x| - R)/r)$, $\chi : \mathbb{R} \to \mathbb{R}^+$ is a fixed smooth function that equals 0 on $[0, +\infty]$ and 1 on $[-\infty, -1]$, and $0 < r < R/2$ are parameters to be chosen later. (The exponent 4 is convenient for technical reasons, in that η^4 and $\nabla(\eta^4)$ share a large common factor η^3, but it should be ignored on a first reading.) Thus we see that η^4 is supported on the ball $B(0, R)$ and equals 1 on $B(0, R - r)$, with the derivative bounds

$$\nabla^j\eta = O(r^{-j}) \tag{58}$$

for $j = 0, 1, 2$. We define the localised energy

$$E_{\eta^4}(t) := \int_{\mathbb{R}^3} \tfrac{1}{2}|u|^2(t, x)\eta^4(x)\,dx. \tag{59}$$

Clearly we have the initial condition

$$E_{\eta^4}(0) \lesssim E(u_0, f). \tag{60}$$

Because η^4 is compactly supported and u is almost smooth, E_{η^4} is C_t^1, and we may differentiate under the integral sign and integrate by parts without difficulty; using (54), we see for almost every time t that

$$\partial_t E_{\eta^4} = -X_1 + X_2 + X_3 + X_4 + X_5, \tag{61}$$

where X_1 is the dissipation term

$$X_1 := \int_{\mathbb{R}^3} |\nabla u|^2\eta^4\,dx = \|\eta^2\nabla u\|_{L_x^2}^2, \tag{62}$$

X_2 is the heat flux term

$$X_2 := \frac{1}{2}\int_{\mathbb{R}^3} |u|^2\Delta(\eta^4)\,dx,$$

X_3 is the transport term

$$X_3 := 4 \int_{\mathbb{R}^3} |u|^2 u \cdot \eta^3 \nabla \eta \, dx,$$

X_4 is the forcing term

$$X_4 := \int_{\mathbb{R}^3} u \cdot f \eta^4 \, dx,$$

and X_5 is the pressure term

$$X_5 := 4 \int_{\mathbb{R}^3} u \tilde{p} \eta^3 \nabla \eta \, dx.$$

The dissipation term X_1 is nonnegative, and will be useful in controlling some of the other terms present here. The heat flux term X_2 can be bounded using (56) and (58) by

$$X_2 \lesssim \frac{A^2}{r^2},$$

so we turn now to the transport term X_3. Using Hölder's inequality and (58), we may bound

$$X_3 \lesssim \frac{1}{r} \|u\eta^2\|_{L^6_x}^{3/2} \|u\|_{L^2_x}^{3/2}, \tag{63}$$

and thus by (56) and Sobolev embedding

$$X_3 \lesssim \frac{A^{3/2}}{r} \|\nabla(u\eta^2)\|_{L^2_x}^{3/2}.$$

By the Leibniz rule and (62), (56), (58), one has

$$\|\nabla(u\eta^2)\|_{L^2_x} \lesssim X_1^{1/2} + \frac{A}{r},$$

and thus

$$X_3 \lesssim \frac{A^{3/2}}{r} X_1^{3/4} + \frac{A^3}{r^{5/2}}.$$

Now we move on to the forcing term X_4. By Cauchy–Schwarz, we can bound this term by

$$X_4 \lesssim E_{\eta^4}^{1/2} a(t),$$

where $a(t) := \|f(t)\|_{L^2_x(B(0,R))}$. Note from (2) that

$$\int_0^T a(t) \, dt \lesssim E(u_0, f, T)^{1/2}. \tag{64}$$

Now we turn to the pressure term X_5. From (55) we have

$$X_5 = \int_{\mathbb{R}^3} \mathbb{O}\big(u(\Delta^{-1}\nabla^2(uu))\eta^3 \nabla\eta\big).$$

We will argue as in the estimation of X_4, but we will first need to move the η^3 weight past the singular integral $\Delta^{-1}\nabla^2$. We therefore bound $X_5 = X_{5,1} + X_{5,2}$, where

$$X_{5,1} = \int_{\mathbb{R}^3} \mathbb{O}\big(u(\Delta^{-1}\nabla^2(uu\eta^3))\nabla\eta\big)$$

and

$$X_{5,2} = \int_{\mathbb{R}^3} \mathbb{O}\big(u[\Delta^{-1}\nabla^2, \eta^3](uu)\nabla\eta\big),$$

where $[A, B] := AB - BA$ is the commutator and η^3 is interpreted as the multiplication operator $\eta^3 : u \mapsto \eta^3 u$. For $X_{6,1}$, we apply Hölder's inequality and (58) to obtain

$$X_{5,1} \lesssim \frac{1}{r}\|u\|_{L^2_x}\|\Delta^{-1}\nabla^2(uu\eta^3)\|_{L^2_x}.$$

The singular integral $\Delta^{-1}\nabla^2$ is bounded on L^2, so it may be discarded; applying Hölder's inequality again, we conclude that

$$X_{5,1} \lesssim \frac{1}{r}\|u\|^{3/2}_{L^2_x}\|u\eta^2\|^{3/2}_{L^6_x}.$$

This is the same bound (63) used to bound X_3, and so by repeating the X_3 analysis, we conclude that

$$X_{5,1} \lesssim \frac{A^{3/2}}{r} X_1^{3/4} + \frac{A^3}{r^{5/2}}.$$

As for $X_{5,2}$, we observe from direct computation of the integral kernel that when $r = 1$, $[\Delta^{-1}\nabla^2, \chi^3]$ is a smoothing operator of infinite order (see [Kato and Ponce 1988]), and in particular

$$\|[\Delta^{-1}\nabla^2, \eta^3]f\|_{L^2_x} \lesssim \|f\|_{L^1_x}$$

in the $r = 1$ case. In the general case, a rescaling argument then gives

$$\|[\Delta^{-1}\nabla^2, \eta^3]f\|_{L^2_x} \lesssim \frac{1}{r^{3/2}}\|f\|_{L^1_x}.$$

Applying Hölder's inequality and (56), we conclude that

$$X_{5,2} \lesssim \frac{A^3}{r^{5/2}}.$$

Putting all the estimates together, we conclude that

$$\partial_t E_{\eta^4} \le -X_1 + O\bigg(\frac{A^2}{r^2} + \frac{A^{3/2}}{r}X_1^{3/4} + \frac{A^3}{r^{5/2}} + E^{1/2}_{\eta^4}a(t)\bigg).$$

By Young's inequality, we have

$$-\tfrac{1}{2}X_1 + O\bigg(\frac{A^{3/2}}{r}X_1^{3/4}\bigg) \lesssim \frac{A^6}{r^4}$$

and

$$\frac{A^3}{r^{5/2}} \lesssim \frac{A^2}{r^2} + \frac{A^6}{r^4},$$

and so we obtain

$$\partial_t E_{\eta^4} + X_1 \lesssim \frac{A^2}{r^2} + \frac{A^6}{r^4} + E^{1/2}_{\eta^4}a(t), \tag{65}$$

and hence for almost every time t,

$$\partial_t \big(E_{\eta^4} + E(u_0, f, T)\big)^{1/2} \lesssim E(u_0, f, T)^{-1/2}\bigg(\frac{A^2}{r^2} + \frac{A^6}{r^4}\bigg) + a(t).$$

By the fundamental theorem of calculus, (64), and (60), we conclude that

$$E_{\eta^4}(t)^{1/2} \lesssim E(u_0, f, T)^{1/2} + E(u_0, f, T)^{-1/2}\bigg(\frac{A^2}{r^2} + \frac{A^6}{r^4}\bigg)T$$

for all $t \in [0, T]$ and all sufficiently large R; sending $r, R \to \infty$ and using the monotone convergence theorem, we conclude that

$$\|u\|_{L^\infty_t L^2_x([0,T]\times\mathbb{R}^3)} \lesssim E(u_0, f, T)^{1/2}.$$

In particular, we have

$$E_{\eta^4}(t) \lesssim E(u_0, f, T)$$

for all r, R; inserting this back into (65) and integrating, we obtain that

$$\int_0^T X_1(t)\,dt \lesssim \bigg(\frac{A^2}{r^2} + \frac{A^6}{r^4}\bigg)T + E(u_0, f, T).$$

Sending $r, R \to \infty$ and using monotone convergence again, we conclude that

$$\|\nabla u\|_{L^2_t L^2_x([0,T]\times\mathbb{R}^3)} \lesssim E(u_0, f, T)^{1/2},$$

and Lemma 8.1 follows. □

We can bootstrap the proof of Lemma 8.1 as follows. *A posteriori*, we see that we may take $A \lesssim E(u_0, f, T)^{1/2}$. If we return to (65), we may then obtain

$$\partial_t (E_{\eta^4} + e)^{1/2} \lesssim e^{-1/2}\bigg(\frac{E(u_0, f, T)}{r^2} + \frac{E(u_0, f, T)^3}{r^4}\bigg) + a(t),$$

where $e > 0$ is an arbitrary parameter which we will optimise later. From the fundamental theorem of calculus, we then have

$$E_{\eta^4}^{1/2} \lesssim E_{\eta^4}(0)^{1/2} + e^{1/2} + e^{-1/2}\bigg(\frac{E(u_0, f, T)}{r^2} + \frac{E(u_0, f, T)^3}{r^4}\bigg)T + \|f\|_{L^1_t L^2_x},$$

where the $L^1_t L^2_x$ norm is over $[0, T] \times B(0, R)$; optimising in e, we conclude that

$$E_{\eta^4}^{1/2} \lesssim E_{\eta^4}(0)^{1/2} + \bigg(\frac{E(u_0, f, T)}{r^2} + \frac{E(u_0, f, T)^3}{r^4}\bigg)^{1/2} T^{1/2} + \|f\|_{L^1_t L^2_x}.$$

Inserting this back into (65) and integrating, we also conclude that

$$\int_0^T X_1(t)\,dt \lesssim \bigg(E_{\eta^4}(0)^{1/2} + \bigg(\frac{E(u_0, f, T)}{r^2} + \frac{E(u_0, f, T)^3}{r^4}\bigg)^{1/2} T^{1/2} + \|f\|_{L^1_t L^2_x}\bigg)^2.$$

Applying spatial translation invariance (29) to move the origin from 0 to an arbitrary point x_0, we deduce an energy localisation result:

Theorem 8.2 (Local energy estimate). *Let (u, p, u_0, f, T) be a finite energy almost smooth solution with f divergence-free. Then for any $x_0 \in \mathbb{R}^3$ and any $0 < r < R/2$, one has*

$$\|u\|_{L^\infty_t L^2_x([0,T]\times B(x_0,R-r))} + \|\nabla u\|_{L^2_t L^2_x([0,T]\times B(x_0,R-r))} \lesssim \|u_0\|_{L^2_x(B(x_0,R))} + \|f\|_{L^1_t L^2_x([0,T]\times B(x_0,R))} + \frac{E(u_0, f, T)^{1/2} T^{1/2}}{r} + \frac{E(u_0, f, T)^{3/2} T^{1/2}}{r^2}. \quad (66)$$

Remark 8.3. One can verify that the estimate (66) is dimensionally consistent. Indeed, if L denotes a length scale, then r, R, $E(u_0, f)$ have the units of L, T has the units of L^2, u has the units of L^{-1}, and all terms in (66) have the scaling of $L^{1/2}$. Note also that the global energy estimate 8.1 can be viewed as the limiting case of (66) when one sends r, R to infinity.

Remark 8.4. A minor modification of the proof of Theorem 8.2 allows one to replace the ball $B(x_0, R)$ by an annulus

$$B(x_0, R')\backslash B(x_0, R)$$

for some $0 < R < R'$ with $r < (R' - R)/2$, $R/2$, with the smaller ball $B(x_0, R - r)$ being replaced by the smaller annulus

$$B(x_0, R' - r)\backslash B(x_0, R + r).$$

The proof is essentially the same, except that the cutoff η has to be adapted to the two indicated annuli rather than to the two indicated balls; we omit the details. Sending $R' \to \infty$ using the monotone convergence theorem, we deduce in particular an external local energy estimate of the form

$$\|u\|_{L^\infty_t L^2_x([0,T]\times(\mathbb{R}^3\backslash B(x_0,R+r)))} + \|\nabla u\|_{L^2_t L^2_x([0,T]\times(\mathbb{R}^3\backslash B(x_0,R+r)))} \lesssim \|u_0\|_{L^2_x(\mathbb{R}^3\backslash B(x_0,R))} + \|f\|_{L^1_t L^2_x([0,T]\times(\mathbb{R}^3\backslash B(x_0,R)))} + \frac{E(u_0, f, T)^{1/2} T^{1/2}}{r} + \frac{E(u_0, f, T)^{3/2} T^{1/2}}{r^2}, \quad (67)$$

whenever $0 < r < R/2$.

Remark 8.5. The hypothesis that f is divergence-free can easily be removed using the symmetry (34), but then f needs to be replaced by Pf on the right-hand side of (66).

Remark 8.6. Theorem 8.2 can be extended without difficulty to the periodic setting, with the energy $E(u_0, f, T)$ being replaced by the periodic energy

$$E_L(u_0, f, T) := \tfrac{1}{2}\big(\|u_0\|_{L^2_x(\mathbb{R}^3/L\mathbb{Z}^3)} + \|f\|_{L^1_t L^2_x([0,T]\times\mathbb{R}^3/L\mathbb{Z}^3)}\big)^2$$

as long as the radius R of the ball is significantly smaller than the period L of the solution, for example, $R < L/100$. The reason for this is that the analysis used to prove Theorem 8.2 takes place almost entirely inside the ball $B(x_0, R)$, and so there is almost no distinction between the finite energy and the periodic cases. The only place where there is any "leakage" outside of $B(x_0, R)$ is in the estimation of the term $X_{5,2}$, which involves the nonlocal commutator $[\Delta^{-1}\nabla^2, \eta^3]$. However, in the regime $R < L/100$, one easily verifies that the commutator essentially obeys the same sort of kernel bounds in the periodic setting as it does in the nonperiodic setting, and so the argument goes through as before. We omit the details.

Remark 8.7. Theorem 8.2 asserts, roughly speaking, that if the energy of the data is small in a large ball, then the energy will remain small in a slightly smaller ball for future times T; similarly, (67) asserts that if the energy of the data is small outside a ball, then the energy will remain small outside a slightly larger ball for future times T. Unfortunately, this estimate is not of major use for the purposes of establishing Theorem 1.20, because energy is a supercritical quantity for the Navier–Stokes equation, and so smallness of energy (local or global) is not a particularly powerful conclusion. To achieve this goal, we will need a variant of Theorem 8.2 in which the energy $\frac{1}{2}\int |u|^2$ is replaced by the *enstrophy* $\frac{1}{2}\int |\omega|^2$, which is subcritical and thus able to control the regularity of solutions effectively.

Remark 8.8. It should be possible to extend Theorem 8.2 to certain classes of weak solutions, such as mild solutions or Leray–Hopf solutions, perhaps after assuming some additional regularity on the solution u. We will not pursue these matters here.

9. Bounded total speed

Let (u, p, u_0, f, T) be an almost smooth finite energy solution. Applying the Leray projection P to (3) (and using Corollary 4.3), we see that

$$\partial_t u = \Delta u + PB(u, u) + Pf \tag{68}$$

for almost all times t, where $B(u, v) = \mathbb{O}(\nabla(uv))$ was defined in (12). As all expressions here are tempered distributions, we thus have the Duhamel formula (11), which we rewrite here as

$$u(t) = e^{t\Delta}u_0 + \int_0^t e^{(t-t')\Delta}\big(P\mathbb{O}(\nabla(uu)) + Pf\big)(t')\,dt'. \tag{69}$$

One can then insert the *a priori* bounds from Lemma 8.1 into (69) to obtain further *a priori* bounds on u in terms of the energy $E(u_0, f, T)$ (although, given that (53) was supercritical with respect to scaling, any further bounds obtained by this scheme must be similarly supercritical).

Many such bounds of this type already exist in the literature. For instance:[17]

- One can bound the vorticity $\omega := \nabla \times u$ in $L^\infty_t L^1_x$ norm [Constantin 1990; Qian 2009].
- One can bound $\nabla^2 u$ in $L^{4/3,\infty}_{t,x}$ [Constantin 1990; Lions 1996].
- More generally, for any $\alpha \geq 1$, one can bound $\nabla^\alpha u$ in $L_t^{4/(\alpha+1),\infty} L_x^{4/(\alpha+1),\infty}$ [Vasseur 2010; Choi and Vasseur 2011].
- For any $k \geq 0$, one can bound $t^k \partial_t^k u$ in $L^2_{t,x}$ [Chae 1992].
- One can bound ∇u in $L_t^{1/2} L_x^\infty$ [Foiaş et al. 1981].
- For any $r \geq 0$ and $k \geq 1$, one can bound $D_t^r \nabla_x^s u$ in $L_t^{2/(4r+2k-1)} L^2_x$ [Foiaş et al. 1981; Doering and Foias 2002; Duff 1990].

[17] These bounds are usually localised in both time and space, or are restricted to the periodic setting, and some bounds were only established in the model case $f = 0$; some of these bounds also apply to weaker notions of solution than classical solutions. For the purposes of this exposition we will not detail these technicalities.

- For any $1 \leq m \leq \infty$, one can bound ω in $L_t^{2m/(4m-3)} L_x^{2m}$ [Gibbon 2012].
- One can bound moments of wave-number like quantities [Doering and Gibbon 2002; Cheskidov and Shvydkoy 2011].

In this section we present another *a priori* bound which will be absolutely crucial for our localisation arguments, and which (somewhat surprisingly) does not appear to be previously in the literature:

Proposition 9.1 (Bounded total speed). *Let (u, p, u_0, f, T) be a finite energy almost smooth solution. Then we have*

$$\|u\|_{L_t^1 L_x^\infty([0,T]\times\mathbb{R}^3)} \lesssim E(u_0, f, T)^{1/2} T^{1/4} + E(u_0, f, T). \tag{70}$$

We observe that the estimate (70) is dimensionally consistent with respect to the scaling (31). Indeed, if L denotes a length scale, then T scales like L^2, u scales like L^{-1}, and E_0 scales like L, so both sides of (70) have the scaling of L.

Before we prove this proposition rigorously, let us first analyse equation (68) heuristically, using Littlewood–Paley projections, to get some feel of what kind of *a priori* estimates one can hope to establish purely from (68) and (53). For the simplicity of this exposition we shall assume $f = 0$. We consider a high-frequency component $u_N := P_N u$ of the velocity field u for some $N \gg 1$. Applying P_N to (68), and using the ellipticity of Δ to adopt the heuristic[18] $P_N \Delta \sim -N^2 P_N$ and $P_N P \nabla \sim N P_N$, we arrive at the heuristic equation

$$\partial_t u_N = -N^2 u_N + \mathbb{O}(N P_N(u^2)).$$

Let us cheat even further and pretend that $P_N(u^2)$ is analogous to $u_N u_N$ (in practice, there will be more terms than this, but let us assume this oversimplification for the sake of discussion). Then we have

$$\partial_t u_N = -N^2 u_N + \mathbb{O}(N u_N^2).$$

Heuristically, this suggests that the high-frequency component u_N should quickly damp itself out into nothingness if $|u_N| \ll N$, but can exhibit nonlinear behaviour when $|u_N| \gg N$. Thus, as a heuristic, one can pretend that u_N has magnitude $\gg N$ on the regions where it is nonnegligible.

This heuristic, coupled with the energy bound (53), already can be used to informally justify many of the known *a priori* bounds on Navier–Stokes solutions. In particular, projecting (53) to the u_N component, one expects that

$$\|u_N\|_{L_t^2 L_x^2} \lesssim N^{-1} \tag{71}$$

(dropping the dependencies of constants on parameters such as E_0 and being vague about the space-time region on which the norms are being evaluated), which by Bernstein's inequality implies that

$$\|u_N\|_{L_t^2 L_x^\infty} \lesssim N^{1/2}.$$

However, with the heuristic that $|u_N| \gg N$ on the support of u_N, we expect that

$$\|u_N\|_{L_t^1 L_x^\infty} \lesssim \frac{1}{N} \|u_N\|_{L_t^2 L_x^\infty}^2 \lesssim 1;$$

[18] One can informally justify this heuristic by inspecting the symbols of the Fourier multipliers appearing in these expressions.

summing in N (and ignoring the logarithmic divergence that results, which can in principle be recovered by using Bessel's inequality to improve upon (71)), we obtain a nonrigorous derivation of Proposition 9.1.

We now turn to the formal proof of Proposition 9.1. All space-time norms are understood to be over the region $[0, T] \times \mathbb{R}^3$ (and all spatial norms over $\mathbb{R}^3$) unless otherwise indicated. We abbreviate $E_0 := E(u_0, f, T)$. From (53) and (2) we have the bounds

$$\|u\|_{L_t^\infty L_x^2} \lesssim E_0^{1/2}, \tag{72}$$

$$\|\nabla u\|_{L_t^2 L_x^2} \lesssim E_0^{1/2}, \tag{73}$$

$$\|u_0\|_{L_x^2} + \|f\|_{L_t^1 L_x^2} \lesssim E_0^{1/2}. \tag{74}$$

We expand out u using (69). For the free term $e^{t\Delta}u_0$, one has by (18)

$$\|e^{t\Delta}u_0\|_{L_x^\infty} \lesssim t^{-3/4}\|u_0\|_{L_x^2}$$

for $t \in [0, T]$, so this contribution to (70) is acceptable by (74). In a similar spirit, we have

$$\|e^{(t-t')\Delta} Pf(t')\|_{L_x^\infty} \lesssim (t-t')^{-3/4}\|Pf(t')\|_{L_x^2} \lesssim (t-t')^{-3/4}\|f(t')\|_{L_x^2},$$

and so this contribution is also acceptable by the Minkowski and Young inequalities and (74).

It remains to show that

$$\left\| \int_0^t e^{(t-t')\Delta} \mathbb{O}(P\nabla(uu)(t'))\, dt' \right\|_{L_t^1 L_x^\infty} \lesssim E_0.$$

By Littlewood–Paley decomposition, the triangle inequality, and Minkowski's inequality, we can bound the left-hand side by

$$\lesssim \sum_N \int_0^T \int_0^t \left\| P_N e^{(t-t')\Delta} \mathbb{O}(P\nabla(uu)(t')) \right\|_{L_x^\infty} dt'\, dt.$$

Using (27) and bounding the first-order operator $P\nabla$ by N on the range of P_N, we may bound this by

$$\lesssim \sum_N \int_0^T \int_0^t \exp\bigl(-c(t-t')N^2\bigr) N \|P_N \mathbb{O}(uu)(t')\|_{L_x^\infty}\, dt'\, dt$$

for some $c > 0$; interchanging integrals and evaluating the t integral, this becomes

$$\lesssim \sum_N \int_0^T N^{-1} \|P_N \mathbb{O}(uu)(t')\|_{L_x^\infty}\, dt'. \tag{75}$$

We now apply the Littlewood–Paley trichotomy (see Section 2) and symmetry to write

$$P_N \mathbb{O}(uu) = \sum_{N_1 \sim N} \sum_{N_2 \lesssim N} P_N \mathbb{O}(u_{N_1} u_{N_2}) + \sum_{N_1 \gtrsim N} \sum_{N_2 \sim N_1} P_N \mathbb{O}(u_{N_1} u_{N_2}),$$

where $u_N := P_N u$. For N_1, N_2 in the first sum, we use Bernstein's inequality to estimate

$$\begin{aligned}\|P_N \mathbb{O}(u_{N_1} u_{N_2})\|_{L^\infty_x} &\lesssim \|u_{N_1}\|_{L^\infty_x} \|u_{N_2}\|_{L^\infty_x} \\ &\lesssim N_1^{3/2} \|u_{N_1}\|_{L^2_x} N_2^{3/2} \|u_{N_2}\|_{L^2_x} \\ &\lesssim N (N_2/N_1)^{1/2} \|\nabla u_{N_1}\|_{L^2_x} \|\nabla u_{N_2}\|_{L^2_x}.\end{aligned}$$

For N_1, N_2 in the second sum, we use Bernstein's inequality in a slightly different way to estimate

$$\begin{aligned}\|P_N \mathbb{O}(u_{N_1} u_{N_2})\|_{L^\infty_x} &\lesssim N^3 \|\mathbb{O}(u_{N_1} u_{N_2})\|_{L^1_x} \\ &\lesssim N^3 \|u_{N_1}\|_{L^2_x} \|u_{N_2}\|_{L^2_x} \\ &\lesssim N (N/N_1)^2 \|\nabla u_{N_1}\|_{L^2_x} \|\nabla u_{N_2}\|_{L^2_x}.\end{aligned}$$

Applying these bounds, we can estimate (75) by

$$\begin{aligned}&\lesssim \sum_N \sum_{N_1 \sim N} \sum_{N_2 \lesssim N} (N_2/N_1)^{1/2} \int_0^T \|\nabla u_{N_1}(t')\|_{L^2_x} \|\nabla u_{N_2}(t')\|_{L^2_x}\, dt' \\ &\quad + \sum_N \sum_{N_1 \gtrsim N} \sum_{N_2 \sim N_1} (N/N_1)^2 \int_0^T \|\nabla u_{N_1}(t')\|_{L^2_x} \|\nabla u_{N_2}(t')\|_{L^2_x}\, dt'.\end{aligned}$$

Performing the N summation first and then using Cauchy–Schwarz, one can bound this by

$$\lesssim \sum_{N_1 \gtrsim 1} \sum_{N_2 \lesssim N_1} (N_2/N_1)^{1/2} a_{N_1} a_{N_2} + \sum_{N_1 \gtrsim 1} \sum_{N_2 \sim N_1} a_{N_1} a_{N_2},$$

where

$$a_N := \|\nabla u_N\|_{L^2_t L^2_x}.$$

But from (73) and Bessel's inequality (or the Plancherel theorem), one has

$$\sum_N a_N^2 \lesssim E_0,$$

and the claim (70) then follows from Schur's test (or Young's inequality).

Remark 9.2. An inspection of this argument reveals that the L^∞_x norm in (70) can be strengthened to a Besov norm $(\dot{B}^{0,\infty}_1)_x$, defined by

$$\|u\|_{(\dot{B}^{0,\infty}_1)_x} := \sum_N \|P_N u\|_{L^\infty_x}.$$

Remark 9.3. An inspection of the proof of Proposition 9.1 reveals that the time-dependent factor $T^{1/4}$ on the right-hand side of Proposition 9.1 was only necessary in order to bound the linear components

$$e^{t\Delta} u_0 + \int_0^t e^{(t-t')\Delta} (Pf)(t')\, dt'$$

of the Duhamel formula (69). If one had some other means to bound these components in $L^1_t L^\infty_x$ by a bound independent of T (for instance, if one had some further control on the decay of u_0 and f, such as

L^1_x and $L^1_t L^1_x$ bounds), then this would lead to a similarly time-independent bound in Proposition 9.1, which could be useful for analysis of the long-time asymptotics of Navier–Stokes solutions (which is not our primary concern here).

Remark 9.4. It is worth comparing the (supercritical) control given by Proposition 9.1 with the well-known (critical) Prodi–Serrin–Ladyzhenskaya regularity condition [Prodi 1959; Serrin 1963; Ladyzhenskaya 1967; Fabes et al. 1972; Struwe 1988], a special case of which (roughly speaking) asserts that smooth solutions to the Navier–Stokes system can be continued as long as u is bounded in $L^2_t L^\infty_x$, and the equally well known (and also critical) regularity condition of Beale, Kato, and Majda [1984], which asserts that smooth solutions can be continued as long as the *vorticity*

$$\omega := \nabla \times u \tag{76}$$

stays bounded in $L^1_t L^\infty_x$.

Remark 9.5. As pointed out by the anonymous referee, one can also obtain $L^1_t L^\infty_x$ bounds on the velocity field u by a Gagliardo–Nirenberg type interpolation between the $L^{1/2}_t L^\infty_x$ bound on ∇u from [Foiaş et al. 1981] with the $L^2_t L^6_x$ bound on u arising from the energy inequality and Sobolev embedding.

Although we will not need it in this paper, Proposition 9.1 when combined with the Picard well-posedness theorem for ODE yields the following immediate corollary, which may be of use in future applications:

Corollary 9.6 (Existence of material coordinates). *Let (u, p, u_0, f, T) be a finite energy smooth solution. Then there exists a unique smooth map $\Phi : [0, T] \times \mathbb{R}^3 \to \mathbb{R}^3$ such that*

$$\Phi(0, x) = x$$

for all $x \in \mathbb{R}^3$, and

$$\partial_t \Phi(t, x) = u(\Phi(t, x))$$

for all $(t, x) \in [0, T] \times \mathbb{R}^3$, and furthermore $\Phi(t) : \mathbb{R}^3 \to \mathbb{R}^3$ is a diffeomorphism for all $t \in [0, T]$. Finally, one has

$$|\Phi(t, x) - x| \lesssim E(u_0, f, T)^{1/2} T^{1/4} + E(u_0, f, T)$$

for all $(t, x) \in [0, T] \times \mathbb{R}^3$.

Remark 9.7. One can extend the results in this section to the periodic case, as long as one assumes normalised pressure and imposes the additional condition $T \leq L^2$, which roughly speaking ensures that the periodic heat kernel behaves enough like its nonperiodic counterpart that estimates such as (18) are maintained; we omit the details. (Without normalised pressure, the Galilean invariance (33) shows that one cannot hope to bound the $L^1_t L^\infty_x$ norm of u by the initial data, and even energy estimates do not work any more.) When the inequality $T \leq L^2$ fails, one can still obtain estimates (but with weaker bounds) by using the crude observation that a solution which is periodic with period L is also periodic with period kL for any positive integer k, and choosing k to be the first integer such that $T \leq (kL)^2$.

10. Enstrophy localisation

The purpose of this section is to establish a subcritical analogue of Theorem 8.2, in which the energy $\frac{1}{2}\int |u|^2$ is replaced by the enstrophy $\frac{1}{2}\int |\omega|^2$. Because the latter quantity is not conserved, we will need a smallness condition on the initial local enstrophy; however, the initial *global* enstrophy is allowed to be arbitrarily large (or even infinite).

Theorem 10.1 (Enstrophy localisation). *Let (u, p, u_0, f, T) be a finite energy almost smooth solution. Let $B(x_0, R)$ be a ball such that*

$$\|\omega_0\|_{L^2_x(B(x_0,R))} + \|\nabla \times f\|_{L^1_t L^2_x([0,T]\times B(x_0,R))} \leq \delta \tag{77}$$

for some $\delta > 0$, where $\omega_0 := \nabla \times u_0$ is the initial vorticity. Assume the smallness condition

$$\delta^4 T + \delta^5 E(u_0, f, T)^{1/2} T \leq c \tag{78}$$

for some sufficiently small absolute constant $c > 0$ (independent of all parameters). Let $0 < r < R/2$ be a quantity such that

$$r > C\big(E(u_0, f, T) + E(u_0, f, T)^{1/2} T^{1/4} + \delta^{-2}\big) \tag{79}$$

for some sufficiently large absolute constant C (again independent of all parameters). Then

$$\|\omega\|_{L^\infty_x L^2_x([0,T]\times B(x_0,R-r))} + \|\nabla\omega\|_{L^2_t L^2_x([0,T]\times B(x_0,R-r))} \lesssim \delta.$$

Remark 10.2. Once again, this theorem is dimensionally consistent (and so one could use (31) to normalise one of the nondimensionless parameters above to equal 1 if desired). Indeed, if L is a unit of length, then u has the units of L^{-1}, ω has the units of L^{-2}, $E(u_0, f, T)$, r, R have the units of L, T has the units of L^2, and δ has the units of $L^{-1/2}$ (so in particular $\delta^4 T$ and $\delta^5 E(u_0, f, T)^{1/2} T$ are dimensionless).

Remark 10.3. The smallness of $\delta^4 T$ also comes up, not coincidentally, as a condition in the local well-posedness theory for the Navier–Stokes at the level of H^1; see (46). The smallness of $\delta^5 E(u_0, f, T)^{1/2} T$ is a more artificial condition, and it is possible that a more careful argument would eliminate it, but we will not need to do so for our applications. For future reference, it will be important to note the fact that δ is permitted to be large in the above theorem, so long as the time T is small.

Remark 10.4. A variant to Theorem 10.1 can also be deduced from the result[19] in [Caffarelli et al. 1982, Theorem D]. Here, instead of assuming a small L^2 condition on the enstrophy, one needs to assume smallness of quantities such as $\int_{\mathbb{R}^3}(|u_0(x)|^2/|x-x_0|)dx$ for all sufficiently large x_0, and then regularity results are obtained outside of a sufficiently large ball in space-time.

We now prove the theorem. Let (u, p, u_0, f, T), $B(x_0, R)$, δ, r be as in the theorem. We may use spatial translation symmetry (29) to normalise $x_0 = 0$. We assume $c > 0$ is a sufficiently small absolute constant, and then assume $C > 0$ is a sufficiently large constant (depending on c). We abbreviate $E_0 := E(u_0, f, T)$.

In principle, this is a subcritical problem, because the local enstrophy $\frac{1}{2}\int_{B(x_0,R)} |\omega|^2$ (or regularised versions thereof) is subcritical with respect to scaling (31). As such, standard energy methods should

[19] We thank the anonymous referee for this observation.

in principle suffice to keep the enstrophy small for small times (using the smallness condition (78), of course). The main difficulty is that the local enstrophy is not fully *coercive*: it controls ω (and, to a lesser extent, u) inside $B(x_0, R)$, but not outside $B(x_0, R)$; while we do have some global control of the solution thanks to the energy estimate (Lemma 8.1), this is supercritical and thus needs to be used sparingly. We will therefore expend a fair amount of effort to prevent our estimates from "leaking" outside $B(x_0, R)$; in particular, one has to avoid the use of nonlocal singular integrals (such as the Leray projection or the Biot–Savart law) and work instead with more local techniques such as integration by parts. This will inevitably lead to some factors that blow up as one approaches the boundary of $B(x_0, R)$ (actually, for technical reasons, we will be using a slightly smaller ball $B(x_0, R'(t))$ as our domain). It turns out, however, that thanks to a moderate amount of harmonic analysis, these boundary factors can (barely) be controlled if one chooses exactly the right type of weight function to define the local enstrophy (it has to be Lipschitz continuous, but no better).

We turn to the details. We will need an auxiliary initial radius $R' = R'(0)$ in the interval $[R - r/4, R]$ which we will choose later (by a pigeonholing argument). Given this R', we then define a time-dependent radius function

$$R'(t) := R' - \frac{1}{c} \int_0^t \|u(s)\|_{L^\infty_x(\mathbb{R}^3)}\, ds.$$

From Proposition 9.1 one has

$$R'(t) \geq R' - O_c(E_0 + E_0^{1/2} T^{1/4}),$$

and thus (by (79)) one has

$$R'(t) \geq R - r/2$$

if the constant C in (79) is sufficiently large depending on c. The reason we introduce this rapidly shrinking radius is that we intend to "outrun" all difficulties caused by the transport component of the Navier–Stokes equation when we deploy the energy method. Note that the bounded total speed property (Proposition 9.1) prevents us from running the radius down to zero when we do this.

We introduce a time-varying Lipschitz continuous cutoff function

$$\eta(t, x) = \min\bigl(\max(0, c^{-0.1}\delta^2(R'(t) - |x|)), 1\bigr).$$

This function is supported on the ball $B(0, R'(t))$ and equals one on $B(0, R'(t) - c^{0.1}\delta^{-2})$, and is radially decreasing; in particular, from (79), we see that η is supported on $B(0, R)$ and equals 1 on $B(0, R - r)$ if C is large enough. As t increases, this cutoff shrinks at speed $(1/c)\|u(t)\|_{L^\infty_x(\mathbb{R}^3)}$, leading to the useful pointwise estimate

$$\partial_t \eta(t, x) \leq -\frac{1}{c}\|u(t)\|_{L^\infty_x(\mathbb{R}^3)} |\nabla_x \eta(t, x)|, \tag{80}$$

which we will use later in this argument to control transport-like terms in the energy estimate (or more precisely, the enstrophy estimate).

Remark 10.5. It will be important that η is Lipschitz continuous but no better; Lipschitz is the minimal regularity for which one can still control the heat flux term (see Y_3 below), but is also the maximal regularity for which there is enough coercivity to control the nonlinear term (see Y_6 below). The argument is in fact remarkably delicate, necessitating a careful application of harmonic analysis techniques (and in particular, a Whitney decomposition of the ball).

We introduce the localised enstrophy

$$W(t) := \tfrac{1}{2} \int_{\mathbb{R}^3} |\omega(t,x)|^2 \eta(t,x)\, dx. \tag{81}$$

From the hypothesis (77) one has the initial condition

$$W(0) \lesssim \delta^2, \tag{82}$$

and to obtain the proposition, it will suffice to show that

$$W(t) \lesssim_c \delta^2 \tag{83}$$

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

As u is almost smooth, W is C^1_t. As in Section 8, we will compute the derivative $\partial_t W$. We first take the curl of (3) to obtain the well-known *vorticity equation*

$$\partial_t \omega + (u \cdot \nabla)\omega = \Delta\omega + \mathbb{O}(\omega \nabla u) + \nabla \times f. \tag{84}$$

This leads to the *enstrophy equation*

$$\partial_t \tfrac{1}{2}|\omega|^2 + (u \cdot \nabla)\tfrac{1}{2}|\omega|^2 = \Delta(\tfrac{1}{2}|\omega|^2) - |\nabla\omega|^2 + \mathbb{O}(\omega\omega\nabla u) + \omega \cdot (\nabla \times f).$$

All terms in this equation are smooth. Integrating this equation against the Lipschitz, compactly supported η and integrating by parts as in Section 8 (interpreting derivatives of η in a distributional sense), we conclude that

$$\partial_t W = -Y_1 - Y_2 + Y_3 + Y_4 + Y_5 + Y_6, \tag{85}$$

where Y_1 is the dissipation term

$$Y_1 := \int_{\mathbb{R}^3} |\nabla\omega|^2 \eta,$$

Y_2 is the recession term

$$Y_2 := -\tfrac{1}{2} \int_{\mathbb{R}^3} |\omega|^2 \partial_t \eta,$$

Y_3 is the heat flux term

$$Y_3 := \tfrac{1}{2} \int_{\mathbb{R}^3} |\omega|^2 \Delta\eta,$$

Y_4 is the transport term

$$Y_4 := \tfrac{1}{2} \int_{\mathbb{R}^3} |\omega|^2 u \cdot \nabla\eta,$$

Y_5 is the forcing term

$$Y_5 := \int_{\mathbb{R}^3} \omega \cdot (\nabla \times f)\eta,$$

and Y_6 is the nonlinear term

$$Y_6 := \int_{\mathbb{R}^3} \mathbb{O}(\omega\omega\nabla u)\eta.$$

The term Y_1 is nonnegative, and will be needed to control some of the other terms. The term Y_2 is also nonnegative; by (80) we see that

$$\int_{\mathbb{R}^3} |\omega|^2 |\nabla \eta| \lesssim c \|u(t)\|_{L^\infty_x(\mathbb{R}^3)} Y_2. \tag{86}$$

We skip the heat flux term Y_3 for now and use (86) to bound the transport term Y_4 by

$$|Y_4| \lesssim c Y_2. \tag{87}$$

Now we turn to the forcing term Y_5. By Cauchy–Schwarz and (81), we have

$$|Y_5| \lesssim W^{1/2} a(t),$$

where

$$a(t) := \|\nabla \times f\|_{L^2_x(B(0,R))}.$$

Note from (77) that

$$\int_0^T a(t)\, dt \lesssim \delta. \tag{88}$$

We return now to the heat flux term Y_3. Computing the distributional Laplacian[20] of η in polar coordinates, we see that

$$Y_3 \lesssim b(t),$$

where $b(t) = b_{R'}(t)$ is the quantity

$$b(t) := c^{-0.1} \delta^2 R^2 \int_{S^2} |\omega(t, R'(t)\alpha)|^2 \, d\alpha + c^{-0.2} \delta^4 \int_{R'(t) - c^{0.1}\delta^{-2} \leq |x| \leq R'(t)} |\omega(t,x)|^2 \, dx,$$

and $d\alpha$ is surface measure on the unit sphere S^2. (Note that while $\Delta\eta$ also has a component on the sphere $|x| = R'(t) - c^{0.1}\delta^{-2}$, this component is negative and thus can be discarded.)

To control $b(t)$, we take advantage of the freedom to choose R'. From Fubini's theorem and a change of variables, we see that

$$\int_{R-\frac{r}{4}}^R \int_0^T b_{R'}(t)\, dt\, dR' \lesssim c^{-0.1} \delta^2 \int_0^T \int_{\mathbb{R}^3} |\omega(t,x)|^2 \, dx.$$

From Lemma 8.1, the right-hand side is $O(\delta^2 E_0 / c^{0.1})$. Thus, by the pigeonhole principle, we may select a radius R' such that

$$\int_0^T b(t)\, dt \lesssim \frac{\delta^2 E_0}{c^{0.1} r},$$

and in particular, by (79),

$$\int_0^T b(t)\, dt \lesssim \delta^2 \tag{89}$$

if C is large enough.

[20] Alternatively, if one wishes to avoid distributions, one can regularise η by a small epsilon parameter to become smooth, compute the Laplacian of the regularised term, and take limits as epsilon goes to zero. One can also rescale either R or δ (but not both) to equal 1 to simplify the computations.

Henceforth we fix R' so that (89) holds. We now turn to the most difficult term, namely the nonlinear term Y_6. Morally speaking, the ∇u term in Y_6 has the "same strength" as ω, and so Y_6 is heuristically as strong as

$$\int_{\mathbb{R}^3} \mathbb{O}(\omega^3)\eta.$$

A standard Whitney decomposition of the support of η, followed by rescaled versions of the Sobolev inequality, bounds this latter expression by

$$O\Big(\Big(\int_{\mathbb{R}^3} |\omega|^2\eta\Big)^{1/2}\Big(\int_{\mathbb{R}^3} |\nabla\omega|^2\eta\Big)\Big).$$

If we could similarly bound Y_6 by this expression by an analogous argument, this would greatly simplify the argument below. Unfortunately, the relationship between ∇u and ω is rather delicate (especially when working relative to the weight η), and we have to perform a much more involved analysis (though still ultimately one which is inspired by the preceding argument).

We turn to the details. We fix t and work in the domain

$$\Omega := B(0, R'(t)).$$

We apply a Whitney-type decomposition, covering Ω by a boundedly overlapping collection of balls $B_i = B(x_i, r_i)$ with radius

$$r_i := \tfrac{1}{100}\min(\operatorname{dist}(x_i, \partial\Omega), c^{0.1}/\delta^2).$$

In particular, we have

$$\eta \sim c^{-0.1}\delta^2 r_i \tag{90}$$

on $B(x_i, 10r_i)$. We can then bound

$$|Y_6| \lesssim c^{-0.1}\delta^2 \sum_i r_i \int_{B_i} |\omega|^2|\nabla u|.$$

The first step is to convert ∇u into an expression that only involves ω (modulo lower-order terms), while staying inside the domain Ω. To do this, we first observe from the divergence-free nature of u that

$$\Delta u = \nabla\times\nabla\times u = \nabla\times\omega.$$

Let ψ_i be a smooth cutoff to the ball $3B_i := B(x_i, 3r_i)$ that equals 1 on $2B_i := B(x_i, 2r_i)$. On $2B_i$, we thus have the local Biot–Savart law

$$u = \mathbb{O}(\Delta^{-1}\nabla(\psi_i\omega)) + v,$$

where v is harmonic on $2B_i$. In particular, from Sobolev embedding one has

$$\|v\|_{L^2_x(2B_i)} \lesssim \|\psi_i\omega\|_{L^{6/5}_x(\mathbb{R}^3)} + \|u\|_{L^2_x(2B_i)}.$$

From Hölder's inequality one has

$$\|\psi_i\omega\|_{L^{6/5}_x(\mathbb{R}^3)} \lesssim r_i\|\omega\|_{L^2_x(2B_i)},$$

while from the mean value principle for harmonic functions one has

$$\|\nabla v\|_{L^\infty_x(B_i)} \lesssim r_i^{-5/2}\|v\|_{L^2_x(2B_i)}.$$

We conclude that

$$\|\nabla v\|_{L^\infty_x(B_i)} \lesssim r_i^{-3/2}\|\omega\|_{L^2_x(2B_i)} + r_i^{-5/2}\|u\|_{L^2_x(2B_i)},$$

and we thus have the pointwise estimate

$$|\nabla u| \lesssim |\nabla\Delta^{-1}\nabla(\psi_i\omega)| + r_i^{-3/2}\|\omega\|_{L^2_x(2B_i)} + r_i^{-5/2}\|u\|_{L^2_x(2B_i)}$$

on B_i. We can thus bound $|Y_6| \leq Y_{6,1} + Y_{6,2}$, where

$$Y_{6,1} \lesssim c^{-0.1}\delta^2 \sum_i r_i \int_{B_i} |\omega|^2 F_i \tag{91}$$

and

$$F_i := |\nabla\Delta^{-1}\nabla(\psi_i\omega)| + r_i^{-3/2}\|\omega\|_{L^2_x(2B_i)}$$

and

$$Y_{6,2} \lesssim c^{-0.1}\delta^2 \sum_i r_i^{-3/2}\|u\|_{L^2_x(2B_i)} \int_{B_i} |\omega|^2.$$

Let us first deal with $Y_{6,2}$, which is the only term not locally controlled by the vorticity alone. If the ball B_i is contained in the annular region

$$\{x \in \Omega : |x| \geq R'(t) - c^{0.1}\delta^{-2}\},$$

which is the region where η is not constant, then we use Hölder to get the bound

$$r^{-3/2}\|u\|_{L^2_x(2B_i)} \lesssim \|u\|_{L^\infty_x(\mathbb{R}^3)}$$

and observe that $c^{-0.1}\delta^2 = |\nabla\eta|$ on B_i. Thus, by (86), the contribution of this term to $Y_{6,2}$ is $O(c^{0.9}Y_2)$. If instead the ball B_i intersects the ball $B(0, R'(t) - c^{0.1}\delta^{-2})$, then $r_i \sim c^{0.1}\delta^{-2}$ and $\eta \sim 1$ on B_i, and we use Lemma 8.1 to obtain the bound

$$r_i^{-3/2}\|u\|_{L^2_x(2B_i)} \lesssim c^{-0.15}\delta^3 E_0^{1/2},$$

and then by (81), (78) the contribution of this case is

$$O(c^{-0.25}\delta^5 E_0^{1/2} W) = O(c^{0.75} W/T);$$

and thus

$$Y_{6,2} \lesssim c^{0.9}Y_2 + c^{0.75}W/T.$$

Now we turn to $Y_{6,1}$. From Plancherel's theorem we have

$$\|\nabla\Delta^{-1}\nabla(\psi_i\omega)\|_{L^2_x(\mathbb{R}^3)} \lesssim \|\psi_i\omega\|_{L^2_x(\mathbb{R}^3)} \lesssim \|\omega\|_{L^2_x(2B_i)},$$

and thus

$$\|F_i\|_{L^2_x(B_i)} \lesssim \|\omega\|_{L^2_x(2B_i)}.$$

From Hölder's inequality we thus have

$$Y_{6,1} \lesssim c^{-0.1}\delta^2 \sum_i r_i^{3/2} \|\omega\|_{L^6_x(B_i)}^2 \|\omega\|_{L^2_x(2B_i)}.$$

To deal with this, we let w_i denote the averages

$$w_i := \left(\frac{1}{|3B_i|}\int_{3B_i} |\omega|^2\right)^{1/2};$$

then

$$\|\omega\|_{L^2_x(2B_i)} \lesssim r_i^{3/2} w_i.$$

Also, from the Sobolev inequality one has

$$\|\omega\|_{L^6_x(B_i)} \lesssim \|\omega\psi_i\|_{L^6_x(\mathbb{R}^3)} \lesssim \|\nabla(\omega\psi_i)\|_{L^2_x(\mathbb{R}^3)} \lesssim \|\nabla\omega\|_{L^2_x(3B_i)} + r_i^{-1}\|\omega\|_{L^2_x(3B_i)} \lesssim \|\nabla\omega\|_{L^2_x(3B_i)} + r_i^{1/2} w_i,$$

and thus

$$Y_{6,1} \lesssim c^{-0.1}\delta^2 \sum_i r_i^3 w_i \|\nabla\omega\|_{L^2_x(3B_i)}^2 + c^{-0.1}\delta^2 \sum_i r_i^4 w_i^3. \tag{92}$$

To deal with the first term of (92), observe from (81) and (90) that

$$\sum_i r_i^4 w_i^2 \lesssim c^{0.1}\delta^{-2} W, \tag{93}$$

and in particular

$$w_i \lesssim c^{0.05}\delta^{-1} W^{1/2} r_i^{-2} \tag{94}$$

for all i. We may thus bound

$$c^{-0.1}\delta^2 \sum_i r_i^3 w_i \|\nabla\omega\|_{L^2_x(3B_i)}^2 \lesssim c^{-0.05}\delta W^{1/2} \sum_i r_i \|\nabla\omega\|_{L^2_x(3B_i)}^2,$$

which by (90) and the bounded overlap of the B_i is

$$\lesssim c^{0.05}\delta^{-1} W^{1/2} \int_\Omega |\nabla\omega|^2\eta \lesssim c^{0.05}\delta^{-1} W^{1/2} Y_1.$$

The second term of (92), $c^{-0.1}\delta^2 \sum_i r_i^4 w_i^3$, is trickier to handle. Call a ball "large" if its radius is at least $10^{-4}c^{-0.1}\delta^{-2}$ (say), and "small" otherwise. To deal with the small balls we use the Poincaré inequality. From this inequality, we see in particular that

$$\left|\left(\frac{1}{|3B_i|}\int_{3B_i} |\omega|^2\right)^{1/2} - \left(\frac{1}{|3B_j|}\int_{3B_j} |\omega|^2\right)^{1/2}\right| \lesssim \left(r_i^{-1}\int_{10B_i} |\nabla\omega|^2\right)^{1/2}$$

whenever B_i, B_j intersect. (Indeed, the Poincaré inequality implies that both terms in the left-hand side are within $O((r_i^{-1}\int_{10B_i}|\nabla\omega|^2)^{1/2})$ of $|(1/|10B_i|)\int_{10B_i}\omega|$.) In other words, we have

$$|w_i - w_j| \lesssim r_i^{-1/2}\left(\int_{10B_i} |\nabla\omega|^2\right)^{1/2} \tag{95}$$

whenever B_i, B_j intersect.

Now for any small ball B_i, we may assign a "parent" ball $B_{p(i)}$ which touches the ball but has radius at least 1.001 (say) times as large as that of B_i. We may iterate this until we reach a large ball $B_{a(i)}$, and write

$$w_i \leq w_{a(i)} + \sum_{k \geq 0} |w_{p^k(i)} - w_{p^{k+1}(i)}|,$$

where the sum is over all k for which $p^{k+1}(i)$ is well-defined; note that this inequality also holds for large balls if we set $a(i) = i$. Taking cubes and using Hölder's inequality, we obtain

$$w_i^3 \lesssim w_{a(i)}^3 + \sum_{k \geq 0} (1+k)^{10} |w_{p^k(i)} - w_{p^{k+1}(i)}|^3,$$

and so we can bound $c^{-0.1}\delta^2 \sum_i r_i^4 w_i^3$ by

$$\lesssim c^{-0.1}\delta^2 \sum_i r_i^4 w_{a(i)}^3 + c^{-0.1}\delta^2 \sum_{k \geq 0} (1+k)^{10} \sum_i r_i^4 |w_{p^k(i)} - w_{p^{k+1}(i)}|^3.$$

If one fixes a large ball B_j, one easily checks that $\sum_{i:a(i)=j} r_i^4 \lesssim r_j^4$, and thus

$$c^{-0.1}\delta^2 \sum_i r_i^4 w_{a(i)}^3 \lesssim c^{-0.1}\delta^2 \sum_{j:r_j > 10^{-4}c^{0.1}\delta^{-2}} r_j^4 w_j^3;$$

applying (94) and (93), we thus have

$$c^{-0.1}\delta^2 \sum_i r_i^4 w_{a(i)}^3 \lesssim c^{-0.25}\delta^5 W^{1/2} \sum_j r_j^4 w_j^2 \lesssim c^{-0.15}\delta^3 W^{3/2}.$$

Similarly, if one fixes a small ball B_j, one verifies that

$$\sum_{k \geq 0} (1+k)^{10} \sum_{i:p^k(i)=j} r_i^4 \lesssim r_j^4,$$

and thus

$$c^{-0.1}\delta^2 \sum_{k \geq 0} (1+k)^{10} \sum_i r_i^4 |w_{p^k(i)} - w_{p^{k+1}(i)}|^3 \lesssim c^{-0.1}\delta^2 \sum_{j:r_j \leq 10^{-4}c^{0.1}\delta^{-2}} r_j^4 |w_j - w_{p(j)}|^3.$$

From (94) (once) and (95) (twice) one has

$$|w_j - w_{p(j)}|^3 \lesssim c^{0.05}\delta^{-1} W^{1/2} r_j^{-3} \int_{10B_j} |\nabla\omega|^2,$$

and so we may bound the preceding expression by

$$\lesssim c^{-0.05}\delta W^{1/2} \sum_j r_j \int_{10B_j} |\nabla\omega|^2,$$

which by (90) and the bounded overlap of the B_j can be bounded by

$$\lesssim c^{0.05}\delta^{-1} W^{1/2} \int_\Omega |\nabla\omega|^2 \eta \lesssim c^{0.05}\delta^{-1} W^{1/2} Y_1.$$

Putting the $Y_{6,1}$ bounds together, we conclude that

$$Y_{6,1} \lesssim c^{-0.15}\delta^3 W^{3/2} + c^{0.05}\delta^{-1}W^{1/2}Y_1;$$

collecting the bounds for $Y_1, \ldots, Y_6$, we thus have

$$\partial_t W \leq -Y_1 + O\big(c^{0.05}\delta^{-1}W^{1/2}Y_1 + c^{-0.15}\delta^3 W^{3/2} + c^{0.75}W/T + a(t)W^{1/2} + b(t)\big).$$

To solve this differential inequality we use the continuity method. Suppose that $0 \leq T' \leq T$ is a time for which

$$\sup_{t\in[0,T']} W(t) \leq c^{-0.01}\delta^2. \tag{96}$$

Then, if c is small enough, we can absorb the $O(c^{0.05}\delta^{-1}W^{1/2}Y_1)$ term by the $-Y_1$ term, and can also use this bound and (78) to obtain

$$c^{-0.15}\delta^3 W^{3/2} \lesssim c^{-0.155}\delta^4 W \lesssim c^{0.75}W/T$$

and

$$a(t)W^{1/2} \lesssim c^{-0.005}\delta a(t).$$

We thus have

$$\partial_t W \lesssim c^{0.75}W/T + c^{-0.005}\delta a(t) + b(t).$$

From Gronwall's inequality and (82), (88), (89), we thus have

$$\sup_{t\in[0,T']} W(t) \lesssim c^{-0.005}\delta^2.$$

For c a small enough absolute constant, this is (slightly) better than the hypothesis (96), and so from the continuity method (and (82)), we conclude that

$$\sup_{t\in[0,T]} W(t) \lesssim c^{-0.005}\delta^2,$$

and the claim (83) follows. The proof of Theorem 10.1 is now complete.

Remark 10.6. As with Remark 8.4, we may adapt the proof of Theorem 10.1 to an annulus, replacing the ball $B(x_0, R)$ with an annulus $B(x_0, R')\backslash B(x_0, R)$ for some $0 < R < R'$ with $0 < r < R/2, (R'-R)/2$, and replacing the smaller ball $B(x_0, R-r)$ with the smaller annulus $B(x_0, R'-r)\backslash B(x_0, R+r)$. To do this, one has to replace the cutoff η (which was shrinking inside the ball $B(x_0, R)$ towards $B(x_0, R-r)$) with a slightly more complicated cutoff (which is shrinking inside the annulus $B(x_0, R')\backslash B(x_0, R)$ towards the smaller annulus $B(x_0, R'-r)\backslash B(x_0, R+r)$). However, aside from this detail, the proof method is essentially identical and is omitted. Sending R' to infinity and using the monotone convergence theorem, we may in fact replace the annulus $B(x_0, R')\backslash B(x_0, R)$ with the exterior region $\mathbb{R}^3\backslash B(x_0, R)$, and the annulus $B(x_0, R'-r)\backslash B(x_0, R+r)$ with $\mathbb{R}^3\backslash B(x_0, R+r)$.

Theorem 10.1 asserts, roughly speaking, that if the H^1_x norm of the data is small on a ball, then for a quantitative amount of later time, the H^1_x norm of the solution remains small on a slightly smaller ball.

As the H^1 norm is subcritical, we expect this sort of result to persist to higher regularities, in the spirit of [Serrin 1962]. It is therefore unsurprising that this is indeed the case:

Proposition 10.7 (Higher regularity). *Let* (u, p, u_0, f, T) *be a finite energy almost smooth solution with* $T \leq T_*$. *Let* $B(x_0, R)$, η, δ, r *obey the conditions* (77), (78), (79) *from Theorem 10.1. Then for any compact subset* K *in the interior of* $B(x_0, R-r)$ *and any* $k \geq 1$, *one can bound*

$$\|\nabla^k u\|_{L^\infty_t L^2_x([0,T]\times K)} + \|\nabla^{k+1} u\|_{L^2_t L^2_x([0,T]\times K)} \lesssim_{k,K,E(u_0,f,T),\delta,T_*,R,A_k} 1,$$

where

$$A_k := \sum_{j=0}^{k} \|\nabla^j u_0\|_{L^2_x(B(x_0,R))} + \|\nabla^j f\|_{L^\infty_t L^2_x([0,T]\times B(x_0,R))}.$$

In particular, one has

$$\|u\|_{X^k([0,T]\times K)} \lesssim_{k,K,E(u_0,f,T),\delta,T_*,R,A_k} 1.$$

Proof. We allow all implied constants to depend on k, K, $E(u_0, f, T)$, δ, T_*, R, A_k. We introduce a compact set

$$K \subset K_1 \subset K_2 \subset K_3 \subset K_4 \subset K_5 \subset B(x_0, R-r),$$

with each set lying in the interior of the next set. Let η be a smooth function supported on K_2 that equals 1 on K_1; we allow implied constants to depend on η.

We begin with the $k = 1$ case. From Theorem 10.1 one already has

$$\|\omega\|_{L^\infty_t L^2_x([0,T]\times K_1)} + \|\nabla\omega\|_{L^2_t L^2_x([0,T]\times K_1)} \lesssim 1.$$

To pass from ω to u, we use integration by parts. Since $\omega = \nabla \times u$ and u is divergence-free, a standard integration by parts shows that

$$\frac{1}{2}\int_{\mathbb{R}^3} |\omega|^2 \eta = \int_{\mathbb{R}^3} |\nabla u|^2 \eta + \int_{\mathbb{R}^3} \mathbb{O}(|u|^2 \nabla^2 \eta).$$

By Lemma 8.1, the error term is $O(1)$, and so we have

$$\int_K |\nabla u|^2 \lesssim 1.$$

Similarly, by replacing ω and u by their derivatives, we also see that

$$\frac{1}{2}\int_{\mathbb{R}^3} |\nabla\omega|^2 \eta = \int_{\mathbb{R}^3} |\nabla^2 u|^2 \eta + \int_{\mathbb{R}^3} \mathbb{O}(|\nabla u|^2 \nabla^2 \eta).$$

By Lemma 8.1, the error term is $O(1)$ after integration in time, and so we also have

$$\int_0^T \int_K |\nabla^2 u|^2 \, dx \, dt \lesssim 1$$

as desired.

We now turn to the $k = 2$ case. This is the most difficult, as we currently only control regularities that are half a derivative better than the critical regularity (which would place u in $H^{1/2}_x$), and wish to boost this to three halves of a derivative above critical; this requires at least two iterations of the Duhamel formula. The arguments will be analogous to the regularity arguments in Theorem 5.1 or Lemma 5.5. By (68) we see that $u\eta$ obeys the truncated equation

$$\partial_t(\eta u) - \Delta(\eta u) = \eta \mathbb{O}(P\nabla(uu)) + \eta P f + \mathbb{O}(\nabla u \nabla \eta) + \mathbb{O}(u\nabla^2\eta) \tag{97}$$

for almost all t. Meanwhile, from the $k = 1$ case and Lemma 8.1, we already have the estimates

$$\|u\|_{L^\infty_t L^2_x([0,T]\times\mathbb{R}^3)} + \|\nabla u\|_{L^\infty_t L^2_x([0,T]\times K_4)} + \|\nabla^2 u\|_{L^2_t L^2_x([0,T]\times K_4)} \lesssim 1, \tag{98}$$

and from the definition of A_2, we have

$$\|\nabla^j u_0\|_{L^2_x(B(x_0,R))} + \|\nabla^j f\|_{L^\infty_t L^2_x([0,T]\times B(x_0,R))} \lesssim 1 \tag{99}$$

for $j = 0, 1, 2$.

We claim that all terms on the right-hand side of (97) have an $L^4_t L^2_x([0, T] \times \mathbb{R}^3)$ norm of $O(1)$. The only difficult term here is $\eta P\mathbb{O}(\nabla(uu))$; the other three terms on the right-hand side are easily estimated in $L^4_t L^2_x$ (and even in $L^2_t L^2_x$) using (98) and (99). We now estimate

$$\|\eta\mathbb{O}(P\nabla(uu))\|_{L^4_t L^2_x([0,T]\times\mathbb{R}^3)}.$$

We split $uu = \tilde{\eta}uu + (1 - \tilde{\eta})uu$, where $\tilde{\eta}$ is a smooth cutoff supported on K_4 that equals 1 on K_3. For the contribution of the nonlocal portion $(1 - \tilde{\eta})$, one can use the smoothness of the kernel of the operator P away from the origin to bound this contribution by $\lesssim \|\mathbb{O}(uu)\|_{L^4_t L^1_x([0,T]\times\mathbb{R}^3)}$, which is acceptable by (98); for future reference, we note that this argument bounds this contribution in $L^2_t L^2_x$ norm as well as in $L^4_t L^2_x$ norm. For the local portion $\tilde{\eta}uu$, we discard the η and P projections and bound this by

$$\lesssim \|\mathbb{O}(\nabla(\tilde{\eta}uu))\|_{L^4_t L^2_x([0,T]\times\mathbb{R}^3)}.$$

But this is acceptable by (24).

We have now placed the right-hand side of (97) in $L^4_t L^2_x([0, T] \times \mathbb{R}^3)$ with norm $O(1)$. Meanwhile, from (99), the initial data $u_0\eta$ is in $H^2_x(\mathbb{R}^3)$ with norm $O(1)$. Applying the energy estimate (23), we conclude that

$$\|u\eta\|_{L^\infty_t H^{3/2-\sigma}_x([0,T]\times\mathbb{R}^3)} + \|u\eta\|_{L^2_t H^{5/2-\sigma}_x([0,T]\times\mathbb{R}^3)} \lesssim_\sigma 1$$

for any $\sigma > 0$. A similar argument (shifting the compact sets) also gives

$$\|u\eta'\|_{L^\infty_t H^{3/2-\sigma}_x([0,T]\times\mathbb{R}^3)} + \|u\eta'\|_{L^2_t H^{5/2-\sigma}_x([0,T]\times\mathbb{R}^3)} \lesssim_\sigma 1,$$

where η' is a smooth function supported on K_5 that equals 1 on K_4. In particular, by Sobolev embedding, on $[0, T] \times K_4$, u is in $L^\infty_t L^{12}_x$, ∇u is in $L^2_t L^{12}_x \cap L^\infty_t L^{12/5}_x$, and $\nabla^2 u$ is in $L^2_t L^{12/5}_x$, which together with (98) and the Hölder inequality now allows one to conclude that $\mathbb{O}(\nabla(\tilde{\eta}uu))$ has an $L^2_t H^1_x([0, T] \times \mathbb{R}^3)$ norm of $O(1)$. Repeating the previous arguments, we now conclude that the right-hand side of (97) lies in $L^2_t H^1_x([0,T]\times\mathbb{R}^3)$ with norm $O(1)$, and hence by (22),

$$\|\eta u\|_{L^\infty_t H^2_x([0,T]\times\mathbb{R}^3)} + \|\eta u\|_{L^2_t H^3_x([0,T]\times\mathbb{R}^3)},$$

which gives the $k=2$ case.

The higher k cases are proven by similar arguments, but are easier as we now have enough regularity to place u in $L^\infty_t L^\infty_x([0,T]\times K_5)$ with norm $O(1)$; we leave the details to the reader. (For instance, to establish the $k=3$ case, one can verify using the estimates already obtained from the $k=2$ case that the right-hand side of (97) has an $L^2_t H^1_x([0,T]\times\mathbb{R}^3)$ norm of $O(1)$. □

Remark 10.8. As in Remark 9.7, one can extend the results here to the periodic setting so long as one has $T\le L^2$ and $R\le L$; we omit the details.

For our application to constructing Leray–Hopf weak solutions, we will need a generalisation of Theorem 10.1 to the case when one has hyperdissipation. More precisely, we introduce a small hyperdissipation parameter $\varepsilon>0$, and consider solutions $(u^{(\varepsilon)}, p^{(\varepsilon)}, u_0, f, T)$ to the regularised Navier–Stokes equation, which are defined precisely as with the usual concept of a Navier–Stokes solution, but with (3) replaced by the regularised variant

$$\partial_t u^{(\varepsilon)} + (u^{(\varepsilon)}\cdot\nabla)u^{(\varepsilon)} = \Delta u^{(\varepsilon)} - \varepsilon\Delta^2 u^{(\varepsilon)} - \nabla p^{(\varepsilon)} + f. \tag{100}$$

With hyperdissipation, the global regularity problem becomes much easier (the energy is now subcritical rather than supercritical), and indeed it is not difficult to use energy methods (see, for example, [Lions 1969]) to show the existence of a unique almost smooth finite energy solution to this regularised equation $(u^{(\varepsilon)}, p^{(\varepsilon)}, u_0, f, T)$ from any given smooth finite energy data (u_0, f, T). The energy estimate in Lemma 8.1 remains true in this case (uniformly in ε), and one easily verifies that one obtains an additional estimate

$$\varepsilon\int_0^T\int_{\mathbb{R}^3} |\nabla^2 u(t,x)|^2\,dt\,dx \lesssim E(u_0, f, T) \tag{101}$$

in this hyperdissipative setting. One can also verify (with a some tedious effort) that Proposition 9.1 also holds in this hyperdissipative setting as long as ε is sufficiently small, basically because the hyperdissipative heat operators $e^{t(\Delta-\varepsilon\Delta^2)}$ obey essentially the same estimates (18), (27) as $e^{t\Delta}$ if $0\le t\le T$ and ε is sufficiently small depending on T; we omit the details.

One can define the vorticity $\omega^{(\varepsilon)} := \nabla\times u^{(\varepsilon)}$ of a regularised solution as before. This vorticity obeys an equation almost identical to (84), but with an additional hyperdissipative term $-\varepsilon\nabla^2\omega^{(\varepsilon)}$ on the right-hand side. One can then repeat the proof of Theorem 10.1 with this additional term. Integrating by parts a large number of times, one obtains a similar decomposition to (85) for the derivative of the localised enstrophy, but with the addition of a negative term $-\varepsilon\int_{\mathbb{R}^3}|\nabla^2\omega|^2\eta$ on the right-hand side, plus some boundary terms which are bounded by $\tilde b(t)$, where

$$\tilde b(t) := \sum_{r=R'(t),\,R'(t)-c^{0.1}\delta^{-2}} \varepsilon c^{-0.1}\delta^2 R^2\int_{S^2}|\nabla\omega(t,r\alpha)|^2\,d\alpha + \varepsilon c^{-0.2}\delta^4\int_{R'(t)-c^{0.1}\delta^{-2}\le|x|\le R'(t)}|\nabla\omega(t,x)|^2\,dx$$

is a hyperdissipative analogue of $b(t)$. By using the same averaging argument used to bound $\int_0^T b(t)dt$

for typical R', one can also simultaneously obtain a comparable bound for $\int_0^T \tilde{b}(t)dt$ (taking advantage of the additional estimate (101)). The rest of the argument in Theorem 10.1 works with essentially no changes; we omit the details. The proof of Proposition 10.7 is also essentially identical, after one notes that energy estimates such as (22) continue to hold in the hyperdissipative setting. Summarising, we obtain:

Proposition 10.9. *Theorem 10.1 and Proposition 10.7 continue to hold in the presence of hyperdissipation, uniformly in the limit $\varepsilon \to 0$.*

11. Consequences of enstrophy localisation

We now give a number of applications of the enstrophy localisation result, Theorem 10.1. Many of these applications resemble existing results in the literature, but with weaker decay hypotheses on the initial data and solution (in particular, we will usually only assume either finite energy or finite H^1 norm); the main point is that the localisation afforded by Theorem 10.1 can significantly reduce the need to assume any stronger decay hypotheses.

We begin with the observation that finite energy smooth solutions automatically have bounded enstrophy if the initial data has bounded enstrophy:

Corollary 11.1 (Bounded enstrophy). *Let (u, p, u_0, f, T) be an almost smooth, finite energy solution, such that the initial data (u_0, f, T) has finite H^1 norm. Then $u \in X^1([0, T] \times \mathbb{R}^3)$; in particular, (u, p, u_0, f, T) is an H^1 solution.*

Proof. Let $\delta > 0$ be small enough (depending on $E(u_0, f, T), T$) that the condition (78) holds. As (u_0, f, T) has finite H^1 norm, we have

$$\|\omega_0\|_{L^2_x(\mathbb{R}^3)} + \|\nabla \times f\|_{L^1_t L^2_x([0,T]\times\mathbb{R}^3)} < \infty.$$

By the monotone convergence theorem, we thus have for R sufficiently large that

$$\|\omega_0\|_{L^2_x(\mathbb{R}^3\backslash B(0,R))} + \|\nabla \times f\|_{L^1_t L^2_x([0,T]\times(\mathbb{R}^3\backslash B(0,R)))} \leq \delta.$$

Applying Theorem 10.1 (inverted as in Remark 10.6), we conclude that

$$\|\omega\|_{L^\infty_x L^2_x([0,T]\times(\mathbb{R}^3\backslash B(0,R+r)))} + \|\nabla\omega\|_{L^2_t L^2_x([0,T]\times(\mathbb{R}^3\backslash B(0,R+r)))} \lesssim \delta$$

for some finite radius r, if R is sufficiently large; in particular, ω lies in $L^\infty_t L^2_x \cap L^2_t H^1_x$ in the exterior region $[0, T] \times (\mathbb{R}^3 \backslash B(0, R+r))$. On the other hand, as u is almost smooth, ω also lies in $L^\infty_t L^2_x \cap L^2_t H^1_x$ in the interior region $[0, T] \times B(0, R+r+1)$ (say). Gluing these two bounds together, we conclude that

$$\omega \in L^\infty_t L^2_x \cap L^2_t H^1_x([0, T] \times \mathbb{R}^3);$$

meanwhile, from Lemma 8.1 one has

$$u \in L^\infty_t L^2_x \cap L^2_t H^1_x([0, T] \times \mathbb{R}^3).$$

Since u is divergence-free and $\omega = \nabla \times u$, the claim then follows from Fourier analysis. $\square$

Remark 11.2. From Corollary 5.8 we know that smooth solutions to the Navier–Stokes solutions can be continued in time as long as the H^1 norm remains bounded. However, Corollary 11.1 certainly does not allow one to solve the global regularity problem for Navier–Stokes, because the proof heavily relies on the solution u being *complete* rather than *incomplete*, and thus it is (almost) smooth all the way up to the final time T, and not just smooth on $[0, T)$. Instead, what Corollary 11.1 does is to show that the solution from H^1 data is well-behaved when one is sufficiently close to spatial infinity; in particular, it does not prevent turbulent behaviour in bounded regions of space-time.

Remark 11.3. If $(u, p, u_0, 0, T)$ is an almost smooth homogeneous finite energy solution, then by Lemma 8.1 we see that $u(t) \in H^1_x(\mathbb{R}^3)$ for almost every time $t \in [0, T]$. Applying the time translation symmetry (30) for a small time shift t_0, we can then convert the finite energy data to H^1 data, and then by Corollary 11.1, we conclude that in fact $u(t) \in H^1_x(\mathbb{R}^3)$ for *all* nonzero times $t \in (0, T]$, and furthermore that $u(t)$ is bounded in H^1_x as soon as t is bounded away from zero.

Since H^1 almost smooth solutions with normalised pressure are automatically H^1 mild solutions, for which uniqueness was established in Theorem 5.4, we thus have uniqueness in the almost smooth finite energy category from smooth H^1 data:

Corollary 11.4 (Unconditional uniqueness). *Let (u_0, f, T) be smooth H^1 data. Then there is at most one almost smooth finite energy solution (u, p, u_0, f, T) with this data and with normalised pressure.*

This result resembles the standard "weak-strong uniqueness" results in the literature, such as those in [Prodi 1959; Serrin 1963; Germain 2006; 2008]. The main novelty here is the lack of decay hypotheses beyond the finite energy hypothesis; note that the almost smoothness of the solution gives plenty of integrability on compact regions of space, but does not imply any global integrability in space.

Remark 11.5. We conjecture that one still retains uniqueness even if the data (u_0, f, T_*) is merely smooth and finite energy, rather than smooth and H^1. Note from Lemma 8.1 that $u(t)$ has finite $H^1_x(\mathbb{R}^3)$ norm for almost every time t, which in principle allows one to enforce uniqueness after any given positive time (in the homogeneous case $f = 0$, at least), but it is not clear to the author how to prevent instantaneous failure of uniqueness at the initial time $t = 0$ with only a smooth finite energy hypothesis on the initial data. It may however be possible to adapt the "weak-strong" uniqueness results of Germain [2006; 2008] to this category, perhaps in combination with the local H^1 control given by Theorem 10.1.

We now use the enstrophy localisation result to study solutions as they approach a (potential) blowup time T_*.

Proposition 11.6 (Uniform smoothness outside a ball). *Let (u, p, u_0, f, T_*^-) be an incomplete almost smooth H^1 solution with normalised pressure for all times $0 < T < T_*$. Then there exists a ball $B(0, R)$ such that*

$$u, p, f, \partial_t u \in L^\infty_t C^k_x([0, T_*) \times K) \tag{102}$$

for all $k \geq 0$ and all compact subsets K of $\mathbb{R}^3 \backslash B(0, R)$.

We remark that similar results were obtained by Caffarelli, Kohn, and Nirenberg [1982] assuming additional spatial decay hypotheses on the data at infinity, and in particular that $\int_{\mathbb{R}^3} |u_0(0,x)|^2 |x| dx < \infty$. The main novelty in this proposition is that one only assumes square-integrability of u_0 and its first derivatives, without any further decay assumption.

Proof. From the argument in the proof of Corollary 11.1 (noting that the bounds are uniform for all times T in a compact set), one can already find a ball $B(0, R_0)$ for which

$$u \in X^1\big([0, T_*) \times (\mathbb{R}^3 \backslash B(0, R_0))\big).$$

Using Proposition 10.7, we then conclude the existence of a larger ball $B(0, R)$ such that

$$u \in X^k([0, T_*) \times K)$$

for all $k \geq 1$ and all compact subsets K of $\mathbb{R}^3 \backslash B(0, R)$. From this, Sobolev embedding, and (9) (using the smoothness of the kernel of $\nabla^k \Delta^{-1}$ away from the origin), we obtain (102) for u, p, f as desired. If one then applies (3) and solves for $\partial_t u$, one obtains the bound for $\partial_t u$ also. □

Remark 11.7. From (102) one can *continuously* extend u up to the portion $\{T_*\} \times (\mathbb{R}^3 \backslash B(0, R))$ of the boundary (compare the partial regularity theory in [Caffarelli et al. 1982]). However, we were unable to demonstrate that u could be extended *smoothly* up to the boundary (or even that $\partial_t u$ is continuous in time at the boundary). The problem is due to the nonlocal effects of pressure; the solution u could be blowing up at time T_* in the interior of $B(0, R)$, leading (via (9)) to time oscillations of the pressure in K (which cannot be directly damped out by the smoothness of the Δ^{-1} kernel, which only attenuates *spatial* oscillations), which by (3) could lead to time oscillations of the solution u in K. Indeed, as Theorem 1.12 shows, these time oscillations can have a nontrivial effect on the regularity of the solution.

Remark 11.8. For future reference, we observe that Proposition 11.6 did not require the full space-time smoothness on f; it would suffice to have $f \in L^\infty_t C^k_x([0, T_*) \times K)$ for all $k \geq 0$ and compact K in order to obtain the conclusion (102). This is because at no stage in the argument was it necessary to differentiate f in time.

In a similar spirit, we may construct Leray–Hopf weak solutions that are spatially smooth outside of a ball for any fixed time T. More precisely, define a *Leray–Hopf weak solution* (u, p, u_0, f, T) to smooth finite energy data (u_0, f, T) to be a distributional solution $u \in X^0([0, T] \times \mathbb{R}^3)$ to (3) (after expressing this equation in divergence form) which is continuous in time in the weak topology of $L^2_x(\mathbb{R}^3)$, and which obeys the energy inequality

$$\tfrac{1}{2}\|u(t)\|^2_{L^2_x(\mathbb{R}^3)} + \int_0^t \|\nabla u(t)\|^2_{L^2_x(\mathbb{R}^3)}\, dx \leq E(u_0, f, T). \tag{103}$$

The existence of such solutions was famously demonstrated in [Leray 1934] for arbitrary finite energy data (u_0, f, T); the singularities of these solutions were analysed in a vast number of papers, which are too numerous to cite here, but we will point out in particular the seminal work [Caffarelli et al. 1982].

Our main regularity result for Leray–Hopf solutions is as follows.

Proposition 11.9 (Existence of partially smooth Leray–Hopf weak solutions). *Let (u_0, f, T) be smooth H^1 data. Then there exists a Leray–Hopf weak solution (u, p, u_0, f, T) to the given data and a ball $B(0, R)$ such that u is spatially smooth in $[0, T] \times (\mathbb{R}^3 \backslash B(0, R))$ (that is, for each $t \in [0, T]$, $u(t)$ is smooth outside of $B(0, R)$).*

Again, similar results were obtained in [Caffarelli et al. 1982] under stronger decay hypotheses on the initial data. We also remark that weak solutions which were only locally of finite energy, from data of uniformly locally finite energy, were constructed in [Lemarié-Rieusset 1999]; the ability to localise the weak solution construction in this fashion is similar in spirit to the results in the proposition.

Proof. (Sketch) We use a standard hyperdissipation[21] regularisation argument. Let $\varepsilon > 0$ be a small parameter, and consider the almost smooth finite-energy solution $(u^{(\varepsilon)}, p^{(\varepsilon)}, u_0, f, T)$ to the regularised Navier–Stokes system (100), which can be shown to exist by energy methods. By Proposition 10.9, we can extend Theorem 10.1 and Proposition 10.7 (and thence Proposition 11.6) to these regularised solutions $u^{(\varepsilon)}$, with bounds that are uniform in ε as $\varepsilon \to 0$. As a consequence, we can find a ball $B(0, R)$ *independent of* ε such that for every compact set K outside of $B(0, R)$ and every $k \geq 0$, $\nabla^k u^{(\varepsilon)}$ lies in $L^\infty_t L^\infty_x([0, T_*] \times K)$ uniformly in N. If we then extract a weak limit point u of the $u^{(\varepsilon)}$, then by standard arguments one verifies that u is a Leray–Hopf weak solution which is spatially smooth outside of $B(0, R)$. □

Remark 11.10. As before, we are unable to demonstrate regularity of u in time due to potential nonlocal effects caused by the pressure, which could in principle cause singularities inside $B(0, R)$ to create time singularities outside of $B(0, R)$.

Remark 11.11. Uniqueness of Leray–Hopf solutions remains a major unsolved problem, for which we have nothing new to contribute; in particular, we do not assert that *all* Leray–Hopf solutions from smooth data obey the conclusions of Proposition 11.9. However, if (u_0, f, ∞) is globally defined smooth H^1 data, the argument above gives a single global Leray–Hopf weak solution (u, p, u_0, f, ∞) with the property that, for each finite time $T < \infty$, there exists a radius $R_T < \infty$ such that u is smooth in $[0, T] \times (\mathbb{R}^3 \backslash B(0, R))$. If we restrict to the case $f = 0$, then from (103) we see that $\|\nabla u(t)\|_{L^2_x(\mathbb{R}^3)}$ must become arbitrarily small along some sequence of times $t = t_n$ going to infinity. If $\|\nabla u(t)\|_{L^2_x(\mathbb{R}^3)}$ is small enough depending on $E(u_0, 0, \infty)$, then standard perturbation theory arguments (see, for example, [Kato 1984]) allow one to obtain a smooth, bounded enstrophy solution from the data $u(t)$ on $(t, +\infty)$, which by the uniqueness theory of Serrin [1963] must match the Leray–Hopf weak solution u on $(t, +\infty)$. As such, we conclude in the homogeneous smooth H^1 case that one can construct a global Leray–Hopf weak solution which is spatially smooth outside of a compact subset of space-time $[0, +\infty) \times \mathbb{R}^3$. Again, we emphasise that this global weak solution need not be unique.

[21] It may also be possible to use other regularisation methods here, such as velocity regularisation, to construct the Leray–Hopf weak solution; however, due to the delicate nature of the proof of the localised enstrophy estimate (Theorem 10.1), we were not able to verify that this estimate remained true in the velocity-regularised setting, uniformly in the regularisation parameter, due to the less favourable vorticity equation in this setting.

12. Smooth H^1 solutions

The purpose of this section is to establish Theorem 1.20(iii). To do this, we will need the ability to localise smooth divergence-free vector fields, as follows.

Lemma 12.1 (Localisation of divergence-free vector fields). *Let $T > 0$, $0 < R_1 < R_2 < R_3 < R_4$, and let $u : [0, T) \times (B(0, R_4)\backslash B(0, R_1)) \to \mathbb{R}^3$ be spatially smooth and divergence-free, such that*

$$u, \partial_t u \in L_t^\infty C_x^k\big([0, T) \times (B(0, R_4)\backslash B(0, R_1))\big)$$

for all $k \geq 0$ and

$$\int_{|x|=r} u(t, x) \cdot n \, d\alpha(x) = 0 \tag{104}$$

for all $R_1 < r < R_4$ and $t \in [0, T)$, where n is the outward normal and $d\alpha$ is surface measure. Then there exists a spatially smooth and divergence-free vector field $\tilde{u} : [0, T) \times (B(0, R_4)\backslash B(0, R_1)) \to \mathbb{R}^3$ which agrees with u on $[0, T) \times (B(0, R_2)\backslash B(0, R_1))$ but vanishes on $[0, T) \times (B(0, R_4)\backslash B(0, R_3))$. Furthermore, we have

$$\tilde{u}, \partial_t u \in L_t^\infty C_x^k\big([0, T) \times (B(0, R_4)\backslash B(0, R_1))\big)$$

for all $k \geq 0$.

Finally, if we have

$$1 \leq 2R_2 \leq R_3 \lesssim R_2,$$

then we have the more quantitative bound

$$\|\tilde{u}\|_{L_t^\infty H^k([0,T)\times(B(0,R_4)\backslash B(0,R_1)))} \lesssim_k \|u\|_{L_t^\infty H^{k+1}([0,T)\times(B(0,R_4)\backslash B(0,R_1)))} \tag{105}$$

for any k. (This latter property will come in handy in the next section.)

Note that the hypothesis (104) is necessary, as can be seen from Stokes' theorem. Lemmas of this type first appear in [Bogovskii 1980].

Proof. One can obtain this lemma as a consequence of the machinery of compactly supported divergence-free wavelets [Lemarie-Rieusset 1992], but for the convenience of the reader we give a self-contained proof here.

Let X denote the vector space of all divergence-free smooth functions $u : B(0, R_4)\backslash B(0, R_1) \to \mathbb{R}^3$ obeying the mean zero condition

$$\int_{|x|=r} u(x) \cdot n \, d\alpha(x) = 0 \tag{106}$$

for all $R_1 < r < R_4$, and such that $\|u\|_{C^k((0,R_4)\backslash B(0,R_1))} < \infty$ for all k. It will suffice to construct a linear transformation $P : X \to X$ that is bounded[22] from C^{k+2} to C^k, that is,

$$\|Pu\|_{C^k((0,R_4)\backslash B(0,R_1))} \lesssim_{R_1,R_2,R_3,R_4,k} \|u\|_{C^{k+2}((0,R_4)\backslash B(0,R_1))}$$

[22] One can reduce this loss of regularity by working in more robust spaces than the classical C^k spaces, such as Sobolev spaces H^s or Hölder spaces $C^{k,\alpha}$, but we will not need to do so here.

for all $k \geq 0$, and such that Pu equals u on $B(0, R_2)\backslash B(0, R_1)$ and vanishes on $B(0, R_4)\backslash B(0, R_3)$, as one can then simply define $\tilde{u}(t) := P\tilde{u}(t)$ for each $t \in [0, T)$.

We now construct P. We work in polar coordinates $x = r\alpha$ with $R_1 \leq r \leq R_4$ and $\alpha \in S^2$ (thus avoiding the coordinate singularity at the origin), and decompose $u(r, \alpha)$ as the sum of a radial vector field $u_r(r, \alpha)\alpha$ for some scalar field u_r and an angular vector field $u_\alpha(r, \alpha)$ which is orthogonal to α; thus, for fixed r, $u_\alpha(r)$ can be viewed as a smooth vector field on the unit sphere S^2 (that is, a smooth section of the tangent bundle of S^2). The divergence-free condition on u in these coordinates then reads

$$\partial_r u_r(r) + \frac{1}{r}\nabla_\alpha \cdot u_\alpha(r) = 0, \tag{107}$$

while the mean-zero condition (106) reads

$$\int_{S^2} u_r(r, \alpha)\, d\alpha = 0.$$

Note that either of these conditions implies that $\partial_r u_r(r)$ has mean zero on S^2 for each r. From (107) and Hodge theory, we see that

$$u_\alpha(r) = r\Delta_\alpha^{-1}\nabla_\alpha \partial_r u_r(r) + v(r),$$

where Δ_α^{-1} inverts the Laplace–Beltrami operator Δ_α on smooth mean-zero functions on S^2 and $v(r)$ is a smooth divergence-free vector field on S^2 that varies smoothly with r.

Let $\eta : [R_1, R_4] \to \mathbb{R}^+$ be a smooth function that equals 1 on $[R_1, R_2]$ and vanishes on $[R_3, R_4]$. Set

$$\tilde{u}_r := \eta(r)u_r$$

and

$$\tilde{u}_\alpha(r) = r\Delta_\alpha^{-1}\nabla_\alpha \partial_r \tilde{u}_r(r) + \eta(r)v(r)$$

and

$$Tu := \tilde{u} := \tilde{u}_r\alpha + \tilde{u}_\alpha.$$

One then easily verifies that $\tilde{u}$ is smooth and divergence-free and obeys (106), depends linearly on u, equals u on $B(0, R_2)\backslash B(0, R_1)$, and vanishes on $B(0, R_4)\backslash B(0, R_4)$. It is also not difficult (using the fundamental solution of Δ_α^{-1}) to see that T maps C^{k+2} to C^k (with some room to spare). The claim follows.

Finally, we prove (105). It suffices to show that

$$\|Tu\|_{H^k(B(0,R_3)\backslash B(0,R_2))} \lesssim_k 1$$

whenever $k \geq 0$, and $u \in X$ is such that

$$\|u\|_{H^{k+2}(B(0,R_4)\backslash B(0,R_1))} \lesssim 1.$$

Henceforth all spatial norms will be on $B(0, R_3)\backslash B(0, R_2)$, and all implied constants may depend on k. As u has an H^{k+1} norm of $O(1)$, u_r and hence $\tilde{u}_r$ has an H^{k+1} norm of $O(1)$ also. As for $\tilde{u}_\alpha$, we observe from the Leibniz rule that

$$\tilde{u}_\alpha = \eta u_\alpha + (r\partial_r \eta(r))\Delta_\alpha^{-1}\nabla_\alpha u_r(r).$$

As u has an H^{k+1} norm of $O(1)$, we know $r^{-i}\nabla_\alpha^i \partial_r^j u_\alpha$ has an L^2 norm of $O(1)$ whenever $i+j \le k+1$, which (using elliptic regularity in the angular variable) implies that $r^{-i}\nabla_\alpha^i \partial_r^j \tilde{u}_\alpha$ has an L^2 norm of $O(1)$ whenever $i+j \le k$. This gives $\tilde{u} = \tilde{u}_r + \tilde{u}_\alpha$, an H^k norm of $O(1)$, as claimed. □

We can now establish Theorem 1.20(iii):

Theorem 12.2. *Suppose Conjecture 1.9 is true. Then Conjecture 1.19 is true.*

Proof. In view of Corollary 5.8, it suffices to show that if (u, p, u_0, f, T_*^-) is an incomplete H^1 mild solution up to time T_*, with u_0, f spatially smooth in the sense of Conjecture 1.19, then u does not blow up in enstrophy norm; thus

$$\limsup_{t \to T_*^-} \|u(t)\|_{H_x^1(\mathbb{R}^3)} < \infty.$$

Let $R > 0$ be a sufficiently large radius. By arguing as in Corollary 11.1, we have

$$u \in L_t^\infty H_x^1(\mathbb{R}^3 \backslash B(0, R)),$$

and thus the blowup must be localised in space:

$$\limsup_{t \to T_*^-} \|u(t)\|_{H_x^1(B(0,R))} < \infty. \tag{108}$$

By Proposition 11.6 and Remark 11.8 (and increasing R if necessary), we also have

$$u, p, f, \partial_t u \in L_t^\infty C_x^k\big([0, T_*) \times (B(0, 5R) \backslash B(0, 2R))\big) \tag{109}$$

for all $k \ge 0$. From Stokes' theorem and the divergence-free nature of u, we also have

$$\int_{|x|=r} u(t, x) \cdot n \, d\alpha(x) = 0$$

for all $r > 0$ and $t \in [0, T)$. Applying Lemma 12.1, we can then find a spatially smooth divergence-free vector field $\tilde{u} : [0, T) \times (B(0, 5R) \backslash B(0, 2R)) \to \mathbb{R}^3$ which agrees with u on $B(0, 3R) \backslash B(0, 2R)$ and vanishes outside of $B(0, 4R)$, with

$$\tilde{u}, \partial_t \tilde{u} \in L_t^\infty C_x^k(B(0, 5R) \backslash B(0, 2R)) \tag{110}$$

for all $k \ge 0$. We then extend $\tilde{u}$ by zero outside of $B(0, 5R)$ and by u inside of $B(0, 2R)$; then $\tilde{u}$ is now smooth on all of $[0, T) \times \mathbb{R}^3$.

Let η be a smooth function supported on $B(0, 5R)$ that equals 1 on $B(0, 4R)$. We define a new forcing term $\tilde{f} : [0, T) \times \mathbb{R}^3 \to \mathbb{R}$ by the formula

$$\tilde{f} := \partial_t \tilde{u} + (\tilde{u} \cdot \nabla)\tilde{u} - \Delta \tilde{u} + \nabla(p\eta); \tag{111}$$

then $\tilde{f}$ is spatially smooth and supported on $B(0, 5R)$ and agrees with f on $B(0, 3R)$. From this and (110), (109) we easily verify that

$$\tilde{f} \in L_t^\infty H_x^1([0, T_*) \times \mathbb{R}^3).$$

Note from taking divergences in (111) and using the compact support of $p\eta$, $\tilde{u}$, $\tilde{f}$ that

$$p\eta = -\Delta^{-1}((\tilde{u}\cdot\nabla)\tilde{u}) + \Delta^{-1}\nabla\cdot\tilde{f}.$$

Thus, $(\tilde{u}, p\eta, \tilde{u}(0), \tilde{f}, T_*^-)$ is an incomplete H^1 pressure-normalised (and hence mild) solution with all components supported in $B(0, 5R)$. If we then choose a period L larger than $10R$, we may embed $B(0, 5R)$ inside $\mathbb{R}^3/L\mathbb{Z}^3$ and obtain an incomplete periodic smooth solution

$$\big(\iota(\tilde{u}), \iota(p\eta), \iota(\tilde{u}(0)), \iota(\tilde{f}), T_*^-, L\big),$$

where we use $\iota(f)$ to denote the extension by zero of a function f supported in $B(0, 5R)$, after embedding the latter in $\mathbb{R}^3/L\mathbb{Z}^3$. By construction, we then have

$$\iota(\tilde{f}) \in L_t^\infty H_x^1([0, T_*) \times \mathbb{R}^3/L\mathbb{Z}^3).$$

As $\{T_*\}$ has measure zero, we may arbitrarily extend $\tilde{f}$ to $[0, T_*] \times \mathbb{R}^3/L\mathbb{Z}^3$ while staying in $L_t^\infty H_x^1$. Applying either Conjecture 1.9 (and the uniqueness component to Theorem 5.1) or Conjecture 1.10, we conclude that

$$\iota(\tilde{u}) \in L_t^\infty H_x^1([0, T_*) \times \mathbb{R}^3/L\mathbb{Z}^3),$$

which implies (since u and $\tilde{u}$ agree on $B(0, R)$) that

$$u \in L_t^\infty H_x^1([0, T_*) \times B(0, R)),$$

which contradicts (108). The claim follows. □

Observe that if we omit the embedding of $B(0, 5R)$ in $\mathbb{R}^3/L\mathbb{Z}^3$ in the preceding argument, we can also deduce Conjecture 1.19 from Conjecture 1.18. Since Conjecture 1.19 clearly implies Conjecture 1.18 as a special case, we obtain Theorem 1.20(iii).

Remark 12.3. The referee has pointed out a variant of the argument above using the partial regularity theory of Caffarelli, Kohn, and Nirenberg [1982], which allows one to partially reverse the above implications, and in particular deduce Conjecture 1.8 from Conjecture 1.19. We sketch the argument as follows. Assume Conjecture 1.19, and assume for contradiction that Conjecture 1.8 fails; thus, there is a periodic solution with smooth inhomogeneous data which first develops singularities at some finite time T, and in particular at some location (T, x_0). We may extend the solution beyond this time as a weak solution. Applying a periodic version of the theory in [Caffarelli et al. 1982], we see that the set of singularities has zero one-dimensional parabolic measure, which among other things implies that the set of radii $r > 0$ such that the solution is singular at (T, x) for some x with $|x - x_0| = r$ has measure zero. Because of this, one can find radii $r_2 > r_1 > 0$ such that the solution is smooth in the annular region $\{(t, x) : 0 \le t \le T;\ r_1 \le |x - x_0| \le r_2\}$. By smoothly truncating the solution u to this annulus as in the proof of Theorem 12.2, one can then create a nonperiodic H^1 mild solution to the inhomogeneous Navier–Stokes equation with spatially smooth data which develops a singularity at (T, x_0) while remaining smooth up to time T, contradicting Conjecture 1.19 (when combined with standard uniqueness and regularity results, such as those in Theorem 5.4).

13. Smooth finite energy solutions

In this section we establish Theorem 1.20(v). It is trivial that Conjecture 1.14 implies Conjecture 1.13, so it suffices to establish:

Theorem 13.1. *Suppose that Conjecture 1.13 is true. Then Conjecture 1.14 is true.*

Proof. Let $(u_0, 0, T)$ be smooth homogeneous finite energy data. Our task is to obtain an almost smooth finite energy solution $(u, p, u_0, 0, T)$ with this data. We allow all implied constants to depend on u_0.

We use a regularisation argument. Let N_n be a sequence of frequencies going to infinity, and set $u_0^{(n)} := P_{\leq N_n} u_0$; then $u_0^{(n)}$ converges to u_0 strongly in $L^2_x(\mathbb{R}^3)$, and $(u_0^{(n)}, 0, T)$ is smooth H^1 data for each n. Thus, by hypothesis, we may find a sequence of almost smooth finite energy solutions $(u^{(n)}, p^{(n)}, u_0^{(n)}, 0, T)$ with this data.

One could try invoking weak compactness right now to extract a solution, but as is well known, one only obtains a Leray–Hopf weak solution by doing so, which need not be smooth. So we will first work to establish some additional regularity on the sequence (after passing to a subsequence as necessary) before extracting a weakly convergent limit.

Since the $(u_0^{(n)}, 0, T)$ are uniformly bounded in energy, we see from Lemma 8.1 that

$$\|u^{(n)}\|_{X^0([0,T]\times\mathbb{R}^3)} \lesssim 1. \tag{112}$$

Now let $0 < \tau_0 < T/2$ be a small time. From (112) and the pigeonhole principle, we may find a sequence of times $\tau^{(n)} \in [0, \tau_0]$ such that

$$\|u^{(n)}(\tau^{(n)})\|_{H^1_x(\mathbb{R}^3)} \lesssim \tau_0^{-1}.$$

Passing to a subsequence, we may assume that $\tau^{(n)}$ converges to a limit $\tau \in [0, \tau_0]$. If we then take $\tau' \in [\tau, 2\tau_0]$ sufficiently close to τ, we may apply Lemma 5.5 and conclude that

$$\|u^{(n)}(\tau')\|_{H^{10}_x(\mathbb{R}^3)} \lesssim_{\tau,\tau',\tau_0} 1$$

(say) for all sufficiently large n. Passing to a further subsequence, we may then assume that $u^{(n)}(\tau')$ converges weakly in $H^{10}_x(\mathbb{R}^3)$ (and thus locally strongly in H^9_x) to a limit $u_0' \in H^{10}_x(\mathbb{R}^3)$. By hypothesis, we may thus find an almost smooth H^1 solution $(u', p', u_0', 0, T-\tau')$ with this data.

Meanwhile, by time translation symmetry (30), $(u^{(n)}(\cdot + \tau'), p^{(n)}(\cdot + \tau'), u^{(n)}(\tau'), 0, T-\tau')$ is also a sequence of almost smooth H^1 solutions. Since $u^{(n)}(\tau')$ converges locally strongly in $H^9_x(\mathbb{R}^3)$ to u_0', we would like to conclude that $u^{(n)}(t+\tau')$ also converges locally strongly to $u(t)$ in $H^1_x(\mathbb{R}^3)$, uniformly in $t \in [0, T-\tau']$. This does not quite follow from the standard local well-posedness theory in Theorem 5.4, because this theory requires strong convergence in the *global* $H^1_x(\mathbb{R}^3)$ norm. However, we may take advantage of the local enstrophy estimates to spatially localise the local well-posedness theory, as follows.

Let $\varepsilon > 0$ be a small quantity (depending on the solution $u' = (u', p', u_0', 0, T-\tau')$) to be chosen later, and let $R > 0$ be a sufficiently large radius (depending on ε and $(u', p', u_0', 0, T-\tau')$) to be chosen later. Since u_0' is in $H^{10}_x(\mathbb{R}^3)$, we see from monotone convergence that

$$\|u_0'\|_{H^{10}_x(\mathbb{R}^3\backslash B(0,R))} \lesssim \varepsilon, \tag{113}$$

if R is sufficiently large depending on ε. Since the $u^{(n)}(\tau')$ converge locally strongly in $H^1_x(\mathbb{R}^3)$ to u'_0, we conclude that

$$\|u^{(n)}(\tau')\|_{H^{10}_x(B(0,10R)\backslash B(0,R))} \lesssim \varepsilon,$$

if n is sufficiently large depending on R, ε. Applying Theorem 10.1, we conclude (if R is large enough depending on u'_0 and $T-\tau'$) that

$$\|u^{(n)}(\cdot+\tau')\|_{X^1([0,T-\tau']\times(B(0,9R)\backslash B(0,2R)))} \lesssim \varepsilon,$$

for n sufficiently large depending on R, ε. Using Duhamel's formula (and Corollary 4.3) repeatedly as in the proof of Proposition 10.7, we may in fact conclude that

$$\|\partial_t^i u^{(n)}(\cdot+\tau')\|_{L^\infty_t H^6_x([0,T-\tau']\times(B(0,8R)\backslash B(0,3R)))} \lesssim_{u',T} \varepsilon \tag{114}$$

(say) for $i=0,1$, taking R large enough depending on u', T, ε to ensure that the contributions to the Duhamel formula coming outside $B(0,9R)$ or inside $B(0,2R)$ are negligible, and taking n sufficiently large as always.

We let $\tilde{p}^{(n)}$ be the normalised pressure, defined by (9); by Corollary 4.3, $\tilde{p}^{(n)}(t)$ and $p^{(n)}(t)$ differ by a constant $C(t)$ for almost every t. Using (9), (114) and Lemma 8.1, we see that

$$\|\tilde{p}^{(n)}\|_{L^\infty_t H^2_x([0,T-\tau']\times(B(0,7R)\backslash B(0,4R)))} \lesssim_{u',T} \varepsilon,$$

if R is large enough depending on u', T, ε.

Applying Lemma 12.1, we may find divergence-free smooth vector fields $\tilde{u}^{(n)} : [\tau', T] \times \mathbb{R}^3 \to \mathbb{R}^3$ which agree with $u^{(n)}$ on $[\tau', T] \times B(0,5R)$ but vanish outside of $[\tau', T] \times B(0,6R)$, with

$$\|\partial_t^i \tilde{u}^{(n)}(\cdot+\tau')\|_{L^\infty_t H^5_x([0,T-\tau']\times(B(0,8R)\backslash B(0,3R)))} \lesssim_{u',T} \varepsilon \tag{115}$$

(say) for n sufficiently large and $i=0,1$.

Let η be a smooth function that equals 1 on $B(0,6R)$, is supported on $B(0,7R)$, and obeys the usual derivative bounds in between. We then consider the smooth solutions

$$\big(\tilde{u}^{(n)}(\cdot+\tau'), \eta\tilde{p}^{(n)}(\cdot+\tau'), \tilde{u}^{(n)}(\tau'), \tilde{f}^{(n)}, T-\tau'\big), \tag{116}$$

where

$$\tilde{f}^{(n)} := \big(\partial_t \tilde{u}^{(n)} + \tilde{u}^{(n)} \cdot \nabla \tilde{u}^{(n)} - \Delta \tilde{u}^{(n)} + \nabla(\eta p^{(n)})\big)(\cdot+\tau').$$

By construction, $\tilde{f}'$ and $\tilde{f}^{(n)}$ are smooth and supported on $[0, T-\tau'] \times (B(0,7R)\backslash B(0,5R))$, and the (116) are smooth, compactly supported solutions. From the preceding bounds on $\tilde{u}^{(n)}$, $\tilde{p}^{(n)}$, we see that

$$\|\tilde{f}^{(n)}\|_{L^\infty_t H^1_x([0,T-\tau]\times\mathbb{R}^3)} \lesssim_{u',T} \varepsilon$$

for n sufficiently large.

Also, using (113), (115) we have

$$\|\tilde{u}^{(n)}(\tau') - u'_0\|_{H^1_x(\mathbb{R}^3)} \lesssim_{u',T} \varepsilon$$

for n sufficiently large. If ε is sufficiently small, we conclude from the local H^1 well-posedness theory (Theorem 5.4) that

$$\|\tilde{u}^{(n)}(\cdot + \tau') - u'\|_{X^1([0,T-\tau']\times\mathbb{R}^3)} \lesssim_{u',T} \varepsilon,$$

and in particular

$$\|u^{(n)}(\cdot + \tau') - u'\|_{X^1([0,T-\tau']\times B(0,R))} \lesssim_{u',T} \varepsilon$$

for n large enough. Sending ε to zero (and R to infinity), we conclude that $u^{(n)}(\cdot + \tau')$ converges weakly to u'. In particular, we see that any weak limit of the $u^{(n)}$ is smooth on $[\tau', T] \times \mathbb{R}^3$ (and furthermore, the weak limit is unique in this space-time region).

The above analysis was for a single choice of τ. Choosing τ to be a sequence of times going to zero (and repeatedly taking subsequences of the $u^{(n)}$ and diagonalising as necessary), we may thus arrive at a subsequence $u^{(n)}$ with the property that there is a unique weak limit u of the $u^{(n)}$, which is smooth on $(0, T] \times \mathbb{R}^3$. If we then set p by (9), we see on taking distributional limits that $(u, p, u_0, 0, T)$ is a Leray–Hopf weak solution to the initial data $(u_0, 0, T)$.

To finish the argument, we need to show that $(u, p, u_0, 0, T)$ is almost smooth at $(0, x_0)$ for every $x_0 \in \mathbb{R}^3$. Fix x_0, and let $R > 0$ be a large radius. As u_0 is smooth, $\|u_0\|_{H^1(B(x_0,5R))}$ is finite, and hence $\|u_0^{(n)}\|_{H^1(B(x_0,5R))}$ is uniformly bounded. Applying Theorem 10.1 (recalling that the $u^{(n)}$ have uniformly bounded energy), we conclude (for R large enough) that there exists $0 < \tau < T$ such that $\|u^{(n)}\|_{X^1([0,\tau]\times B(x_0,4R))}$ is uniformly bounded in n. Using Duhamel's formula as in Proposition 11.6, and noting that $u^{(n)}$ is uniformly smooth on $B(x_0, 4R)$, we conclude that $\|u^{(n)}\|_{L^\infty_t C^k((0,\tau]\times B(x_0,3R))}$ is uniformly bounded for all $k \geq 0$. Taking weak limits, we conclude that

$$u \in L^\infty_t C^k((0, \tau] \times B(x_0, 3R))$$

for all $k \geq 0$. From this and (9) (and Lemma 8.1), we also see that

$$p \in L^\infty_t C^k((0, \tau] \times B(x_0, 2R))$$

for all $k \geq 0$. Using (3), we conclude that

$$\partial_t u \in L^\infty_t C^k((0, \tau] \times B(x_0, 2R))$$

for all $k \geq 0$. A similar argument also shows that

$$\partial_t u^{(n)} \in L^\infty_t C^k((0, \tau] \times B(x_0, 2R))$$

uniformly in n. From this, we see that the $\nabla_x^k u^{(n)}$ are uniformly Lipschitz in a neighbourhood of $(0, x_0)$. Since $\nabla_x^k u^{(n)}$ converges weakly to the smooth function $\nabla_x^k u$ in $(0, T] \times \mathbb{R}^3$, and also converges strongly at time zero in $H^1_x(\mathbb{R}^3)$ to the smooth function $\nabla_x^k u_0$, we conclude that $\nabla_x^k u$ can be extended in a locally Lipschitz continuous manner from $(0, T] \times \mathbb{R}^3$ to $[0, T] \times \mathbb{R}^3$ in such a way that it agrees with $\nabla_x^k u_0$ at time zero.

Now we consider derivatives $\nabla^k p$ of the pressure near $(0, x_0)$. Let $\varepsilon > 0$ be arbitrary. Then by the monotone convergence theorem, we see that if $R' > 0$ is a sufficiently large radius, then

$$\|u_0\|_{L^2_x(\mathbb{R}^3 \backslash B(x_0, R'))} \leq \varepsilon,$$

and thus

$$\|u_0^{(n)}\|_{L^2_x(\mathbb{R}^3 \backslash B(x_0, R'))} \lesssim \varepsilon$$

for n large enough.

By Theorem 8.2, we conclude that if R' is large enough, there exists a time $0 < \tau < T$ such that

$$\|u^{(n)}\|_{L^\infty_t L^2_x([0,\tau] \times (\mathbb{R}^3 \backslash B(x_0, 2R')))} \lesssim \varepsilon,$$

and hence on taking weak limits,

$$\|u\|_{L^\infty_t L^2_x([0,\tau] \times (\mathbb{R}^3 \backslash B(x_0, 2R')))} \lesssim \varepsilon.$$

On the other hand, as $\nabla^k u$ is continuous at $t = 0$, $u(t)$ converges in $C^k(B(x_0, 2R'))$ to u_0 as $t \to 0$ for any $k \geq 0$. From this and (9) (and the decay of derivatives of the kernel of Δ^{-1} away from the origin), we see that

$$\limsup_{(t,x) \to (0,x_0); t > 0} |\nabla^k p(t, x) - \nabla^k p_0(x_0)| \lesssim_k \varepsilon$$

for any $k \geq 0$, where p_0 is defined from u_0 using (9). Sending $\varepsilon \to 0$ and $R' \to \infty$, we conclude that $\nabla^k p$ extends continuously to $\nabla^k p_0(x_0)$ at $(0, x_0)$, and thus extends continuously to $\nabla^k p_0$ on all of the initial slice $\{0\} \times \mathbb{R}^3$. By (3) we conclude that $\partial_t \nabla^k u$ also extends continuously to the initial slice, with the Navier–Stokes equation (3) being obeyed both for times $t > 0$ and times $t = 0$. We have thus constructed an almost smooth finite energy solution $(u, p, u_0, 0, T)$ as desired. □

Remark 13.2. We emphasise that Theorem 13.1 only establishes *existence* of a smooth finite energy solution (assuming Conjecture 1.13), and not uniqueness; see Remark 11.5. However, it is not difficult to see from the argument that one can at least ensure that the solution constructed is independent of the choice of time T, and can thus be extended to a single global smooth finite energy solution. (Alternatively, from Lemma 8.1 we see that the enstrophy of the solution will become arbitrarily small for a sequence of times going to infinity, so for a sufficiently large time one can in fact construct a global smooth solution by standard perturbation theory techniques.)

Remark 13.3. One can modify the above argument to also establish Conjecture 1.14 with a nonzero Schwartz forcing term f, provided of course that one also assumes Conjecture 1.13 can be extended to the same class of f. We have not, however, investigated the weakest class of forcing terms f for which the argument works, though certainly finite energy seems insufficient.

14. Quantitative H^1 bounds

In this section we prove Theorem 1.20(vi). We begin with some easy implications. Firstly, it is trivial that Conjecture 1.17 implies Conjecture 1.16, and from the local well-posedness and regularity theory in Theorem 5.4 (or Corollary 5.8), we see that Conjecture 1.16 implies Conjecture 1.15, which in turn implies Conjecture 1.13 (thanks to Proposition 5.6).

Next, we observe from Theorem 5.1 and Lemma 5.5 that given any H^1 data $(u_0, 0, T)$, there exists a time $0 < \tau < T$ such that one has an H^1 mild solution $(u, p, u_0, 0, \tau)$ with $u(\tau)$ smooth. If Conjecture 1.13 holds, then one can then continue the solution in an almost smooth finite energy manner (and hence in an almost smooth H^1 manner, thanks to Corollary 11.1) from τ up to T. Normalising the pressure of this latter solution using Lemma 4.1 and gluing the two solutions together, we obtain an H^1 mild solution up to time T. From this we see that Conjecture 1.13 implies Conjecture 1.15.

Now we show that Conjecture 1.16 implies Conjecture 1.17. Suppose that one has homogeneous H^1 data $(u_0, 0, T)$ with

$$\|u_0\|_{H^1_x(\mathbb{R}^3)} \leq A < \infty.$$

By Conjecture 1.16 (which implies Conjecture 1.15), we may obtain a mild H^1 solution $(u, p, u_0, 0, T)$, which is smooth for positive times. Our objective is to show that

$$\|u\|_{L^\infty_t H^1_x([0,T]\times\mathbb{R}^3)} \lesssim_A 1.$$

Let $\varepsilon > 0$ be a quantity depending on A to be chosen later. We may assume that T is sufficiently large depending on ε, A; otherwise the claim will follow immediately from Conjecture 1.16. Using Lemma 8.1 and the pigeonhole principle, we may then find a time $0 < T_1 < T$ with $T_1 \lesssim_A 1$ such that

$$\|\nabla u(T_1)\|_{L^2_x(\mathbb{R}^3)} \leq \varepsilon.$$

Meanwhile, from energy estimates, one has

$$\|u(T_1)\|_{L^2_x(\mathbb{R}^3)} \lesssim_A 1.$$

On $[T_1, T]$, we split $u = u_1 + v$, where u_1 is the linear solution $u_1(t) := e^{(t-T_1)}u(T_1)$ and $v := u - u_1$. From (21), one thus has

$$\|u_1\|_{X^0} \lesssim_A 1$$

and

$$\|\nabla u_1\|_{X^0} \lesssim \varepsilon.$$

From (11), (22) one has

$$\|v\|_{X^1([T_1,T]\times\mathbb{R}^3)} \lesssim \left\|\mathbb{O}(u_1\nabla u_1 + u_1\nabla v + v\nabla u_1 + v\nabla v)\right\|_{L^2_t L^2_x([T_1,T]\times\mathbb{R}^3)}.$$

We now estimate various contributions to the right-hand side. We begin with the nonlinear term $\mathbb{O}(v\nabla v)$. By Hölder (and dropping the domain $[T_1, T] \times \mathbb{R}^3$ for brevity) followed by Lemma 8.1, we have

$$\|\mathbb{O}(v\nabla v)\|_{L^2_t L^2_x} \lesssim \|\nabla v\|^{1/2}_{L^2_t L^6_x} \|\nabla v\|^{1/2}_{L^\infty_t L^2_x} \|v\|^{1/2}_{L^\infty_t L^6_x} \|v\|^{1/2}_{L^2_t L^6_x} \lesssim \|v\|^{3/2}_{X^1} \|v\|_{X_0^{1/2}} \lesssim_A \|v\|^{3/2}_{X^1}.$$

A similar argument gives

$$\|\mathbb{O}(v\nabla u_1)\|_{L^2_t L^2_x} \lesssim \|\nabla u_1\|_{X^0} \|v\|^{1/2}_{X^1} \|v\|^{1/2}_{X_0} \lesssim \varepsilon \|v\|_{X^1}$$

and

$$\|\mathbb{O}(u_1 \nabla u_1)\|_{L^2_t L^2_x} \lesssim \|\nabla u_1\|_{X^0} \|\nabla u_1\|_{X^0}^{1/2} \|u_1\|_{X_0}^{1/2} \| \lesssim_A \varepsilon^{3/2}$$

and

$$\|\mathbb{O}(u_1 \nabla v)\|_{L^2_t L^2_x} \lesssim \|\nabla v\|_{X^0} \|\nabla u_1\|_{X^0}^{1/2} \|u_1\|_{X_0}^{1/2} \lesssim_A \varepsilon^{1/2} \|v\|_{X^1}$$

and thus

$$\|v\|_{X^1} \lesssim_A \varepsilon^{3/2} + \varepsilon^{1/2} \|v\|_{X^1} + \|v\|_{X^1}^{3/2}.$$

If ε is small enough depending on A, a continuity argument in the T variable then gives

$$\|v\|_{X^1} \lesssim_A \varepsilon^{3/2}$$

and thus

$$\|u\|_{X^1([T_1,T])} \lesssim_A 1.$$

Using this and the triangle inequality, we conclude that Conjecture 1.16 implies Conjecture 1.17.

We now turn to the most difficult implication:

Proposition 14.1 (Concentration compactness). *If Conjecture 1.15 is true*, *so is Conjecture 1.16.*

We now prove this proposition. The methods are essentially those of [Gallagher 2001] (which are in turn based in [Bahouri and Gérard 1999; Gérard 1998]), which treated the (more difficult) critical analogue of this implication; indeed, one can view Proposition 14.1 as a subcritical analogue of the critical result [Gallagher 2001, Corollary 1]. For the convenience of the reader, though, we give a self-contained proof here, which does not need the full power of the machinery in the previously cited papers because we are now working in a subcritical regularity H^1 rather than a critical regularity such as $\dot H^{1/2}$, and as such one does not need to consider the role of the scaling symmetry (31).

We first make the remark that to prove Conjecture 1.16, it suffices to do so with the condition

$$\|u_0\|_{H^1_x(\mathbb{R}^3)} \le A \tag{117}$$

replaced by (say)

$$\|u_0\|_{H^{100}_x(\mathbb{R}^3)} \le A. \tag{118}$$

To see this, observe that if we take data u_0 in $H^1_x(\mathbb{R}^3)$, then from Theorem 5.4 and Lemma 5.5 there exists a time $T_1 > 0$ depending only on A such that

$$\|u\|_{L^\infty_t H^1_x([0,\min(T,T_1)]\times\mathbb{R}^3)} \lesssim_A 1,$$

and such that

$$\|u(T_1)\|_{H^{100}_x(\mathbb{R}^3)} \lesssim_A 1$$

if $T > T_1$. From this and time translation symmetry (30), we see that we can deduce the $H^1_x(\mathbb{R}^3)$ version of Conjecture 1.16 from the $H^{100}_x(\mathbb{R}^3)$ version.

Now suppose for contradiction that the $H^{100}_x(\mathbb{R}^3)$ version of Conjecture 1.16 failed. Carefully negating the quantifiers, we can find a sequence $(u^{(n)}, p^{(n)}, u_0^{(n)}, 0, T^{(n)})$ of smooth homogeneous H^1 solutions, with $T^{(n)}$ uniformly bounded, and $u_0^{(n)}$ uniformly bounded in $H^{100}_x(\mathbb{R}^3)$, such that

$$\lim_{n\to\infty} \|u^{(n)}\|_{L^\infty_t H^1_x([0,T^{(n)}]\times\mathbb{R}^3)} = \infty. \tag{119}$$

By Lemma 4.1 we may assume that these solutions have normalised pressure.

If we were working on a compact domain, such as $\mathbb{R}^3/\mathbb{Z}^3$, we could now extract a subsequence of the $u_0^{(n)}$ that converged strongly in a lower regularity space, such as $H^{99}_x(\mathbb{R}^3/\mathbb{Z}^3)$. But our domain $\mathbb{R}^3$ is noncompact, and in particular has the action of a noncompact symmetry group, namely the translation group $\tau_{x_0}u(x) := u(x-x_0)$. However, as is well known, we have a substitute for compactness in this setting, namely *concentration compactness*. Specifically:

Proposition 14.2 (Profile decomposition). *Let $u_0^{(n)} \in H^{100}_x(\mathbb{R}^3)$ be a sequence with*

$$\limsup_{n\to\infty} \|u_0^{(n)}\|_{H^{100}_x(\mathbb{R}^3)} \leq A,$$

and let $\varepsilon > 0$. Then, after passing to a subsequence, there exists a decomposition

$$u_0^{(n)} = \sum_{j=1}^{J} \tau_{x_j^{(n)}} w_{j,0} + r_0^{(n)},$$

where $|J| \lesssim_{A,\varepsilon} 1$, $w_{1,0}, \dots, w_{J,0} \in H^{100}_x(\mathbb{R}^3)$, $x_j^{(n)} \in \mathbb{R}^3$, and the remainder $r_0^{(n)}$ obeys the estimates

$$\limsup_{n\to\infty} \|r_0^{(n)}\|_{H^{100}_x(\mathbb{R}^3)} \leq A$$

and

$$\limsup_{n\to\infty} \|r_0^{(n)}\|_{L^\infty_x(\mathbb{R}^3)} \leq \varepsilon. \tag{120}$$

Furthermore, for any $1 \leq j < j' \leq J$, one has

$$|x_j^{(n)} - x_{j'}^{(n)}| \to \infty, \tag{121}$$

and for any $1 \leq j \leq J$, the sequence $\tau_{-x_j^{(n)}} r_0^{(n)}$ converges weakly in $H^{100}_x(\mathbb{R}^3)$ to zero.

Finally, if the $u_0^{(n)}$ are divergence-free, then the $w_{j,0}$ and $r_0^{(n)}$ are also divergence-free.

Proof. See, for example, [Gérard 1998]. We sketch the (standard) proof as follows. If

$$\|u_0^{(n)}\|_{L^\infty_x(\mathbb{R}^3)} \leq \varepsilon$$

for all sufficiently large n, then there is nothing to prove (just take $J = 0$ and $r_0^{(n)} := u_0^{(n)}$). Otherwise, after passing to a subsequence, we can find a sequence $x_1^{(n)} \in \mathbb{R}^3$ such that $|u_0^{(n)}(x_1^{(n)})| \geq \varepsilon/2$ (say). The sequence $\tau_{-x_1^{(n)}} u_0^{(n)}$ is then bounded in $H^{100}_x(\mathbb{R}^3)$ and bounded away from zero at the origin; by passing to a further subsequence, we may assume that it converges weakly in $H^{100}_x(\mathbb{R}^3)$ to a limit w_1, which then has an $H^{100}_x(\mathbb{R}^3)$ norm of $\gtrsim_{A,\varepsilon} 1$ and is asymptotically orthogonal in the Hilbert space $H^{100}_x(\mathbb{R}^3)$ to $\tau_{-x_1^{(n)}} u_0^{(n)}$. We then have the decomposition

$$u_0^{(n)} = \tau_{x_1^{(n)}} w_{1,0} + u_0^{(n),1},$$

and from an application of the cosine rule in the Hilbert space $H^{100}_x(\mathbb{R}^3)$, one can verify that

$$\limsup_{n\to\infty} \|u_0^{(n),1}\|^2_{H^{100}_x(\mathbb{R}^3)} \le A^2 - c$$

for some $c > 0$ depending only on ε, A. We can then iterate this procedure $O_{J,\varepsilon}(1)$ times to obtain the desired decomposition. □

We apply this proposition with a value of $\varepsilon > 0$ depending on A, T to be chosen later. The $w_{j,0}$ lie in $H^{100}_x(\mathbb{R}^3)$, and thus by the assumption that Conjecture 1.15 is true, we can find mild H^1 solutions $(w_j, p_j, w_{j,0}, 0, T)$ with this data. By Theorem 5.1, we have

$$\|w_j\|_{X^{100}} < \infty$$

for each $1 \le j \le J$, and to abbreviate the notation, we adopt the convention that the space-time domain is understood to be $[0, T] \times \mathbb{R}^3$.

Next, we consider the remainder term $r_0^{(n)}$. From (21) one has

$$\|e^{t\Delta} r_0^{(n)}\|_{X^{100}} \lesssim A,$$

while from (120) one has

$$\|e^{t\Delta} r_0^{(n)}\|_{L^\infty_t L^\infty_x} \lesssim \varepsilon$$

for n sufficiently large. Interpolating between the two, we soon conclude that

$$\|e^{t\Delta} r_0^{(n)}\|_{X^1} \lesssim_{A,T} \varepsilon^c$$

for some absolute constant $c > 0$. If we take ε sufficiently small depending on A, T, we can use stability of the zero solution (see Theorem 5.1; one could also have used here the results from [Chemin and Gallagher 2009]) to conclude the existence of a mild H^1 solution $(r^{(n)}, p_*^{(n)}, r_0^{(n)}, 0, T)$ with this data, with the estimates

$$\|r^{(n)}\|_{X^1} \lesssim_{A,T} \varepsilon^c; \tag{122}$$

from Theorem 5.1, we then also have

$$\|r^{(n)}\|_{X^{100}} \lesssim_{A,T} 1.$$

We now form the solution

$$(\tilde{u}^{(n)}, \tilde{p}^{(n)}, u_0^{(n)}, \tilde{f}^{(n)}, T),$$

where the velocity field $\tilde{u}^{(n)}$ is given by

$$\tilde{u}^{(n)} := \sum_{j=1}^{J} \tau_{x_j^{(n)}} w_j + r^{(n)},$$

the pressure field $\tilde{p}^{(n)}$ is given by (9), and the forcing term $\tilde{f}^{(n)}$ is given by the formula

$$\tilde{f}^{(n)} := \partial_t \tilde{u}^{(n)} - \Delta \tilde{u}^{(n)} - PB(\tilde{u}^{(n)}, \tilde{u}^{(n)}).$$

This is clearly a mild H^1 solution, with

$$\|\tilde{u}^{(n)}\|_{X^{100}} \lesssim_{A,T,\varepsilon} 1.$$

We now estimate $\tilde{f}^{(n)}$. From (68) for the solutions $\tau_{x_j^{(n)}} w_j + r^{(n)}$, we have an expansion of $\tilde{f}^{(n)}$ purely involving nonlinear interaction terms:

$$\tilde{f}^{(n)} = \sum_{1 \le j < j' \le J} P\mathbb{O}(\nabla(\tau_{x_j^{(n)}} w_j, \tau_{x_{j'}^{(n)}} w_{j'})) + \sum_{1 \le j < J} P\mathbb{O}(\nabla(\tau_{x_j^{(n)}} w_j, r^{(n)})).$$

In particular, from the triangle inequality and translation invariance we have

$$\|\tilde{f}^{(n)}\|_{L_t^2 L_x^2} \lesssim \sum_{1 \le j < j' \le J} \left\| \mathbb{O}(\nabla(w_j, \tau_{x_{j'}^{(n)} - x_j^{(n)}} w_{j'})) \right\|_{L_t^2 L_x^2} + \sum_{1 \le j < J} \left\| \mathbb{O}(\nabla(w_j, \tau_{-x_j^{(n)}} r^{(n)})) \right\|_{L_t^2 L_x^2}.$$

But by (121) and Sobolev embedding,

$$\tau_{x_{j'}^{(n)} - x_j^{(n)}} w_{j'} \quad \text{and} \quad \tau_{-x_j^{(n)}} r^{(n)}$$

are bounded in $L_t^\infty L_x^\infty$ and converge locally uniformly to zero, and so we conclude that

$$\lim_{n \to \infty} \|\tilde{f}^{(n)}\|_{L_t^2 L_x^2} = 0.$$

From this and the stability theory in Theorem 5.4, we conclude that for n large enough, there is an H^1 mild solution $(u^{(n)}, p^{(n)}, u_0^{(n)}, 0, T)$ with

$$\lim_{n \to \infty} \|\tilde{u}^{(n)} - u^{(n)}\|_{X^1} = 0,$$

and in particular

$$\limsup_{n \to \infty} \|u^{(n)}\|_{L_t^\infty H_x^1([0,T] \times \mathbb{R}^3)} < \infty.$$

By the uniqueness theory in Theorem 5.4, this solution must agree with the original solutions

$$(u^{(n)}, p^{(n)}, u_0^{(n)}, 0, T^{(n)})$$

on $[0, T^{(n)}] \times \mathbb{R}^3$; but then we contradict (119). Proposition 14.1 follows.

15. Nonexistence of smooth solutions

In this section we establish Theorem 1.12. Informally, the reason for the irregularity is as follows. Assuming normalised pressure, one concludes from (9) that

$$p = \mathbb{O}(\Delta^{-1} \nabla^2 (uu)).$$

If one then differentiates this twice in time, using (3) to convert time derivatives of u into Δu plus lower-order terms and using integration by parts to redistribute derivatives, we eventually obtain (formally, at least) a formula of the form

$$\partial_t^2 p = \mathcal{O}(\Delta^{-1}\nabla^2(\Delta u \Delta u)) + \text{lower-order terms.}$$

But if u_0 is merely assumed to be smooth and in H^1, then Δu can grow arbitrarily fast at infinity at time $t = 0$, and this should cause p to fail to be C_t^2 at time zero.

We turn to the details. To eliminate the normalised pressure assumption, we will work with ∇p instead of p, and thus we will seek to establish bad behaviour for $\nabla \partial_t^2 p$ at time $t = 0$. For technical reasons it is convenient to work in the weak topology in space. The key quantitative step is the following:

Proposition 15.1 (Quantitative failure of regularity). *Let $u_0 : \mathbb{R}^3 \to \mathbb{R}^3$ be smooth, divergence-free, and compactly supported, and let $\psi : \mathbb{R}^3 \to \mathbb{R}$ be smooth, compactly supported, and have total mass $\int_{\mathbb{R}^3} \psi = 1$. Let $R, M, \varepsilon > 0$. Then there exists a smooth divergence-free compactly supported function u_1 which vanishes on $B(0, R)$ with*

$$\|u_1\|_{H_x^1(\mathbb{R}^3)} \lesssim \varepsilon$$

and such that if $(u, p, u_0 + u_1, 0, T)$ is a mild H^1 (and hence smooth, by Proposition 5.6) solution with data $(u_0 + u_1, 0, T)$, then

$$\left| \int_{\mathbb{R}^3} \nabla \partial_t^2 p(0, x) \psi(x)\, dx \right| > M. \tag{123}$$

Let us assume this proposition for now and conclude Theorem 1.12. We will use an argument reminiscent of that used to establish the Baire category theorem or the uniform boundedness principle. Let $\psi : \mathbb{R}^3 \to \mathbb{R}$ be a fixed smooth, compactly supported function with total mass 1. We will need a rapidly decreasing sequence

$$\varepsilon^{(1)} > \varepsilon^{(2)} > \cdots > 0$$

of small quantities to be chosen later, with each $\varepsilon^{(n)}$ sufficiently small depending on the previous $\varepsilon^{(1)}, \dots, \varepsilon^{(n-1)}$. Applying Proposition 15.1 recursively starting with $u_0 = 0$, one can then find a sequence of smooth, divergence-free, and compactly supported functions $u_1^{(n)}$ for $n = 1, 2, \dots$ such that

$$\|u_1^{(n)}\|_{H_x^1(\mathbb{R}^3)} \lesssim \varepsilon^{(n)},$$

with $u_1^{(n)}$ vanishing on $B(0, 1/\varepsilon^{(n)})$, such that if $(u^{(n)}, p^{(n)}, u_0^{(n)}, 0, T^{(n)})$ is a mild H^1 (and hence smooth) solution with data

$$u_0^{(n)} := u_1^{(1)} + \cdots + u_1^{(n)},$$

then

$$\left| \int_{\mathbb{R}^3} \nabla \partial_t^2 p^{(n)}(0, x) \psi(x)\, dx \right| > 1/\varepsilon^{(n)}. \tag{124}$$

Furthermore, each $u_1^{(n)}$ depends only on $\varepsilon^{(1)}, \dots, \varepsilon^{(n)}$, and in particular is independent of $\varepsilon^{(n+1)}$.

By the triangle inequality (and assuming the $\varepsilon^{(n)}$ decay fast enough), the data $u_0^{(n)}$ is strongly convergent in $H^1_x(\mathbb{R}^3)$ to a limit $u_0 = \sum_{n=1}^\infty u_1^{(n)} \in H^1_x(\mathbb{R}^3)$, with

$$\|u_0 - u_0^{(n)}\|_{H^1_x(\mathbb{R}^3)} \lesssim \varepsilon^{(n+1)}.$$

If we make each $\varepsilon^{(n+1)}$ sufficiently small depending on $u_0^{(n)}$, and hence on $\varepsilon^{(1)}, \dots, \varepsilon^{(n)}$, then the $u_1^{(n)}$ will have disjoint supports; as each $u_1^{(n)}$ is smooth and divergence-free, this implies that

$$u_0 = \sum_{n=1}^\infty u_1^{(n)}$$

is also smooth and divergence-free.

Applying Theorem 5.1, we may then take the times $T^{(n)} = 1$ (if the $\varepsilon^{(n)}$ are small enough), and $(u^{(n)}, p^{(n)}, u_0^{(n)}, 0, 1)$ will converge to a mild H^1 solution $(u, p, u_0, 0, 1)$ in the sense that $u^{(n)}$ converges strongly in $X^1([0, 1] \times \mathbb{R}^3)$ to u. Indeed, from the Lipschitz stability property, we see (if the $\varepsilon^{(n)}$ decay fast enough) that

$$\|u - u^{(n)}\|_{X^1([0,1]\times\mathbb{R}^3)} \lesssim \varepsilon^{(n+1)}.$$

Also, $u, u^{(n)}$ are bounded in $X^1([0, 1] \times \mathbb{R}^3)$ by $O(1)$. Using (9) and Sobolev embedding, this implies

$$\|p - p^{(n)}\|_{L^\infty_t L^3_x([0,1]\times\mathbb{R}^3)} \lesssim \varepsilon^{(n+1)},$$

and so if one sets

$$F^{(n)}(t) := \int_{\mathbb{R}^3} \nabla p^{(n)}(t, x)\psi(x)\, dx$$

and

$$F(t) := \int_{\mathbb{R}^3} \nabla p(t, x)\psi(x)\, dx,$$

then from integration by parts, we have

$$\|F - F^{(n)}\|_{L^\infty_t([0,1])} \lesssim \varepsilon^{(n+1)}. \tag{125}$$

Meanwhile, each $F^{(n)}$ is smooth, and F continuous, from Proposition 5.6, and from (124) one has

$$|\partial_t^2 F^{(n)}(0)| \geq 1/\varepsilon^{(n)}.$$

In particular, if $\varepsilon^{(n+1)}$ is sufficiently small depending on $F^{(n)}$ (which in turn depends on $\varepsilon^{(1)}, \dots, \varepsilon^{(n)}$), one has from Taylor's theorem with remainder that

$$\frac{\left|F^{(n)}(2(\varepsilon^{(n+1)})^{0.1}) - 2F^{(n)}((\varepsilon^{(n+1)})^{0.1}) + F^{(n)}(0)\right|}{(\varepsilon^{(n+1)})^{0.2}} \gtrsim \frac{1}{\varepsilon^{(n)}}.$$

Applying (125), we conclude that

$$\frac{\left|F(2(\varepsilon^{(n+1)})^{0.1}) - 2F((\varepsilon^{(n+1)})^{0.1}) + F(0)\right|}{(\varepsilon^{(n+1)})^{0.2}} \gtrsim \frac{1}{\varepsilon^{(n)}},$$

if $\varepsilon^{(n+1)}$ is sufficiently small depending on $\varepsilon^{(1)}, \dots, \varepsilon^{(n)}$. In particular,

$$\limsup_{h \to 0^+} \frac{|F(2h) - 2F(h) + F(0)|}{h^2} = +\infty,$$

which by Taylor's theorem with remainder implies that F is not smooth at 0.

We claim that the data u_0 gives the desired counterexample to Theorem 1.12. Indeed, suppose for contradiction that there was a smooth solution $(\tilde{u}, \tilde{p}, u_0, 0, T)$ for some $T > 0$. By shrinking T, we may assume $T \leq 1$. By Lemma 4.1, we see that $\tilde{p}(t)$ has normalised pressure up to a constant for almost every t, and thus after adjusting $\tilde{p}(t)$ by that constant, $(\tilde{u}, \tilde{p}, u_0, 0, T)$ is a mild H^1 solution. Using the uniqueness property in Theorem 5.1, we conclude that $u = \tilde{u}$, and $p(t)$ and $\tilde{p}(t)$ differ by a constant for almost every t, and hence (by continuity of both p and $\tilde{p}$) for every t. In particular, $\nabla p = \nabla \tilde{p}$, and so

$$F(t) = \int_{\mathbb{R}^3} \nabla \tilde{p}(t, x) \psi(x)\, dx.$$

But as $\tilde{p}$ is smooth on $[0, T] \times \mathbb{R}^3$, F is smooth at 0, a contradiction.

Remark 15.2. The above argument showed that ∇p failed to be smooth at $t = 0$; by using (3), we conclude that the velocity field u must then also be nonsmooth at $t = 0$ (though the velocity u has one more degree of time regularity than the pressure p). Thus the failure of regularity is not just an artefact of pressure normalisation. Using the vorticity equation (84), one can then show a similar failure of time regularity for the vorticity, although again one gains an additional degree of time differentiability over the velocity u.

The irregularities in time stem from the unbounded growth of high derivatives of the initial data. If one assumes that *all* spatial derivatives of u_0 are in $L^2_x(\mathbb{R}^3)$, that is, that $u_0 \in H^\infty(\mathbb{R}^3)$, then one can prove iteratively[23] that all time derivatives of u and p at time zero are bounded, and also have first spatial derivatives in $H^\infty(\mathbb{R}^3)$ (basically because the first derivative of the kernel of the Leray projection is integrable at infinity). In particular, u and p now remain smooth at time 0.

It remains to establish Proposition 15.1. Fix $u_0, \psi, R, M, \varepsilon$, and let u_1 be a smooth divergence-free compactly supported function u_1 vanishing on $B(0, R)$ with $H^1_x(\mathbb{R}^3)$ norm $O(\varepsilon)$ to be chosen later. Let $(u, p, u_0 + u_1, 0, T)$ be a mild H^1 solution with this given data. By Theorem 5.1, this is a smooth solution, with all derivatives of u, p lying in $L^\infty_t L^2_x$. From Lemma 4.1 we thus have

$$\nabla p = -\nabla \Delta^{-1} \partial_i \partial_j (u_i u_j) \tag{126}$$

for almost all times t. But both sides are smooth in $[0, T] \times \mathbb{R}^3$, so this formula is valid for all times t (and in particular at $t = 0$). In particular, we may apply a Leray projection P to (3) and conclude that

$$\partial_t u = \Delta u + PB(u, u). \tag{127}$$

We differentiate (126) once in time to obtain

$$\nabla \partial_t p = -2 \nabla \Delta^{-1} \partial_i \partial_j (u_i \partial_t u_j).$$

[23] We thank Richard Melrose for this observation.

Expanding out $\partial_t u_j$ using (3), we obtain

$$\nabla \partial_t p = -2\nabla\Delta^{-1}\partial_i\partial_j(u_i \Delta u_j) + \mathcal{O}\big(\Delta^{-1}\nabla^3(uPB(u,u))\big).$$

Writing

$$\partial_i\partial_j(u_i\Delta u_j) = -2\partial_i\partial_j((\partial_k u_i)(\partial_k u_j)) + \mathcal{O}(\nabla^4(uu)),$$

we thus have

$$\nabla\partial_t p = 2\nabla\Delta^{-1}\partial_i\partial_j(\partial_k u_i \partial_k u_j) + \mathcal{O}(\Delta^{-1}\nabla^5(uu)) + \mathcal{O}\big(\Delta^{-1}\nabla^3(uPB(u,u))\big).$$

We differentiate this in time again and use (127) to obtain

$$\begin{aligned}\nabla\partial_t^2 p = {} & 4\nabla\Delta^{-1}\partial_i\partial_j(\partial_k u_i \partial_k \Delta u_j)\\ & + \mathcal{O}\big(\Delta^{-1}\nabla^3((\nabla u)\nabla PB(u,u))\big)\\ & + \mathcal{O}(\Delta^{-1}\nabla^5(u\partial_t u))\\ & + \mathcal{O}\big(\Delta^{-1}\nabla^3((\partial_t u)PB(u,u))\big)\\ & + \mathcal{O}\big(\Delta^{-1}\nabla^3(uPB(u,\partial_t u))\big).\end{aligned}$$

We can write $\partial_k u_i \partial_k \Delta u_j = -(\Delta u_i)(\Delta u_j) + \mathcal{O}(\nabla(\nabla u \Delta u))$, so that

$$\begin{aligned}\nabla\partial_t^2 p = {} & -4\nabla\Delta^{-1}\partial_i\partial_j(\Delta u_i \Delta u_j)\\ & + \mathcal{O}(\Delta^{-1}\nabla^4(\nabla u\Delta u))\\ & + \mathcal{O}\big(\Delta^{-1}\nabla^3((\nabla u)\nabla PB(u,u))\big)\\ & + \mathcal{O}(\Delta^{-1}\nabla^5(u\partial_t u))\\ & + \mathcal{O}\big(\Delta^{-1}\nabla^3((\partial_t u)PB(u,u))\big)\\ & + \mathcal{O}\big(\Delta^{-1}\nabla^3(uPB(u,\partial_t u))\big).\end{aligned}$$

Integrating this against ψ, we may thus expand

$$\int_{\mathbb{R}^3} \nabla\partial_t^2 p(0,x)\psi(x)\,dx = 4X_0 + \sum_{i=1}^{5}\mathcal{O}(X_i),$$

where

$$\begin{aligned} X_0 &:= \int_{\mathbb{R}^3}(\partial_i\partial_j\nabla\Delta^{-1}\psi)\Delta u_i\Delta u_j, & X_1 &:= \int_{\mathbb{R}^3}(\nabla^4\Delta^{-1}\psi)\nabla u\Delta u,\\ X_2 &:= \int_{\mathbb{R}^3}(\nabla^3\Delta^{-1}\psi)\nabla u\nabla PB(u,u), & X_3 &:= \int_{\mathbb{R}^3}(\nabla^5\Delta^{-1}\psi)u\partial_t u,\\ X_4 &:= \int_{\mathbb{R}^3}(\nabla^3\Delta^{-1}\psi)(\partial_t u)PB(u,u), & X_5 &:= \int_{\mathbb{R}^3}(\nabla^3\Delta^{-1}\psi)uPB(u,\partial_t u),\end{aligned}$$

with all expressions being evaluated at time 0.

From (127) and Sobolev embedding, one has

$$\|\partial_t u(0)\|_{L^2_x(\mathbb{R}^3)} \lesssim_{u_0} 1 + \|u_1\|_{H^2_x(\mathbb{R}^3)}.$$

Meanwhile, if ε is small enough, we see that

$$\|u(0)\|_{H^1_x(\mathbb{R}^3)} \lesssim_{u_0} 1,$$

and thus from the Gagliardo–Nirenberg inequality,

$$\|u(0)\|_{L^\infty_x(\mathbb{R}^3)} \lesssim_{u_0} (1+\|u_1\|_{H^2_x(\mathbb{R}^3)})^{1/2}.$$

From many applications of the Sobolev and Hölder inequalities (and, in the case of X_5, an integration by parts to move the derivative off of $\partial_t u$), we conclude that

$$|X_i| \lesssim_{u_0,\psi} (1+\|u_1\|_{H^2_x(\mathbb{R}^3)})^{3/2},$$

for $i = 1, 2, 3, 4, 5$. In a similar spirit, one has

$$X_0 = \int_{\mathbb{R}^3} (\partial_i \partial_j \nabla \Delta^{-1} \psi) \Delta u_{1,i} \Delta u_{1,j} + O_{u_0,\psi}(1+\|u_1\|_{H^2_x(\mathbb{R}^3)}).$$

To demonstrate (123), it thus suffices to exhibit a sequence $u_1^{(n)} : \mathbb{R}^3 \to \mathbb{R}^3$ of smooth divergence-free compactly supported vector fields supported outside of $B(0, R)$ such that

$$\left| \int_{\mathbb{R}^3} (\partial_i \partial_j \nabla \Delta^{-1} \psi) \Delta u^{(n)}_{1,i} \Delta u^{(n)}_{1,j} \right| \gtrsim_{R,\psi} \|u_1^{(n)}\|^2_{H^2_x(\mathbb{R}^3)},$$

with

$$\|u_1^{(n)}\|_{H^1_x(\mathbb{R}^3)} \to 0 \quad \text{and} \quad \|u_1^{(n)}\|_{H^2_x(\mathbb{R}^3)} \to \infty.$$

We construct $u_1^{(n)}$ explicitly as the "wave packet"

$$u_1^{(n)}(x) := n^{-5/2} \nabla \times \Psi^{(n)}(x_0),$$

where e_1, e_2, e_3 is the standard basis, $x_0 \in \mathbb{R}^3$ is a point (independent of n) outside of $B(0, R+1)$ to be chosen later, and

$$\Psi^{(n)}(x) = \chi(x) \sin(n\xi \cdot x)\eta,$$

where $\xi \in \mathbb{R}^3$ is a nonzero frequency (independent of n) to be chosen later, $\eta \in \mathbb{R}^3$ is a nonzero direction, and $\chi : \mathbb{R}^3 \to \mathbb{R}$ is a smooth bump function supported on $B(0, 1)$ to be chosen later. Note from construction that $u_1^{(n)}$ is smooth, divergence-free, and supported on $B(x_0, 1)$, and thus vanishing on $B(0, R)$ for $R_0 > R+1$. One can compute that

$$\|u_1^{(n)}\|_{H^1_x(\mathbb{R}^3)} \ll_\chi n^{-1/2} \quad \text{and} \quad \|u_1^{(n)}\|_{H^2_x(\mathbb{R}^3)} \gg_\chi n^{1/2},$$

as long as χ is not identically zero. To conclude the theorem, it thus suffices to show that

$$\left| \int_{\mathbb{R}^3} (\partial_i \partial_j \nabla \Delta^{-1} \psi) \Delta u^{(n)}_{1,i} \Delta u^{(n)}_{1,j} \right| \gg_{R_0,\psi,\chi} n$$

if R_0 and n are large enough.

Observe that

$$u_1^{(n)}(x) := n^{-3/2} \sin(n\xi \cdot (x - x_0)) \chi(x - x_0)(\xi \times \eta) + O(n^{-5/2})$$

and similarly

$$\Delta u_1^{(n)}(x) := -n^{1/2}|\xi|^2 \sin(n\xi \cdot (x - x_0))\chi(x - x_0)(\xi \times \eta) + O(n^{-1/2}),$$

and so by choosing χ appropriately and using the Riemann–Lebesgue lemma, it suffices to find $x_0, \xi, \eta \in \mathbb{R}^3$ such that

$$(\partial_i \partial_j \nabla \Delta^{-1} \psi)(\xi \times \eta)_i (\xi \times \eta)_j (x_0) \neq 0.$$

But as ψ has mean one, we see that $\nabla^3 \Delta^{-1} \psi(x_0)$ is not identically zero for x_0 large enough, and the claim follows.

Acknowledgements

The author is also indebted to Andrea Bertozzi for valuable discussions, and in particular for raising the question of whether the answer to the global regularity problem of the Navier–Stokes equation is sensitive to the decay hypotheses on the initial velocity field. We thank Xiaoyi Zhang and the anonymous referee for corrections and suggestions.

References

[Bahouri and Gérard 1999] H. Bahouri and P. Gérard, “High frequency approximation of solutions to critical nonlinear wave equations”, *Amer. J. Math.* **121**:1 (1999), 131–175. MR 2000i:35123 Zbl 0919.35089

[Beale et al. 1984] J. T. Beale, T. Kato, and A. Majda, “Remarks on the breakdown of smooth solutions for the 3-D Euler equations”, *Comm. Math. Phys.* **94**:1 (1984), 61–66. MR 85j:35154 Zbl 0573.76029

[Bogovskii 1980] M. E. Bogovskii, “Solutions of some problems of vector analysis, associated with the operators div and grad”, pp. 5–40, 149 in *Theory of cubature formulas and the application of functional analysis to problems of mathematical physics*, Trudy Sem. S. L. Soboleva **1**, Akad. Nauk SSSR Sibirsk. Otdel. Inst. Mat., Novosibirsk, 1980. In Russian. MR 82m:26014

[Caffarelli et al. 1982] L. Caffarelli, R. Kohn, and L. Nirenberg, “Partial regularity of suitable weak solutions of the Navier–Stokes equations”, *Comm. Pure Appl. Math.* **35**:6 (1982), 771–831. MR 84m:35097 Zbl 0509.35067

[Chae 1992] D. Chae, “Some a priori estimates for weak solutions of the 3-D Navier–Stokes equations”, *J. Math. Anal. Appl.* **167**:1 (1992), 236–244. MR 93d:35112 Zbl 0755.35085

[Chemin and Gallagher 2009] J.-Y. Chemin and I. Gallagher, “Wellposedness and stability results for the Navier–Stokes equations in $\mathbf{R}^3$”, *Ann. Inst. H. Poincaré Anal. Non Linéaire* **26**:2 (2009), 599–624. MR 2010i:35278 Zbl 1165.35038

[Cheskidov and Shvydkoy 2011] A. Cheskidov and R. Shvydkoy, “Regularity problem for the 3D Navier–Stokes equations: the use of Kolmogorov’s dissipation range”, preprint, 2011. arXiv 1102.1944v1

[Choi and Vasseur 2011] K. Choi and A. Vasseur, “Estimates on fractional higher derivatives of weak solutions for the Navier–Stokes equations”, preprint, Oxford Centre for Nonlinear PDE, 2011, Available at http://goo.gl/wDKRz. arXiv 1105.1526v1

[Constantin 1990] P. Constantin, “Navier–Stokes equations and area of interfaces”, *Comm. Math. Phys.* **129**:2 (1990), 241–266. MR 91i:35148 Zbl 0725.35080

[Doering and Foias 2002] C. R. Doering and C. Foias, “Energy dissipation in body-forced turbulence”, *J. Fluid Mech.* **467** (2002), 289–306. MR 2003i:76048 Zbl 1029.76025

[Doering and Gibbon 2002] C. R. Doering and J. D. Gibbon, “Bounds on moments of the energy spectrum for weak solutions of the three-dimensional Navier–Stokes equations”, *Phys. D* **165**:3-4 (2002), 163–175. MR 2003j:76022 Zbl 0997.35046

[Duff 1990] G. F. D. Duff, “Derivative estimates for the Navier–Stokes equations in a three-dimensional region”, *Acta Math.* **164**:3-4 (1990), 145–210. MR 91j:35218 Zbl 0728.35083

[Fabes et al. 1972] E. B. Fabes, B. F. Jones, and N. M. Rivière, “The initial value problem for the Navier–Stokes equations with data in L^p”, *Arch. Rational Mech. Anal.* **45** (1972), 222–240. MR 47 #5463 Zbl 0254.35097

[Fefferman 2006] C. L. Fefferman, "Existence and smoothness of the Navier–Stokes equation", pp. 57–67 in *The millennium prize problems*, edited by J. Carlson et al., Clay Math. Inst., 2006. MR 2238274 Zbl 1194.35002

[Foiaş et al. 1981] C. Foiaş, C. Guillopé, and R. Temam, "New a priori estimates for Navier–Stokes equations in dimension 3", *Comm. Partial Differential Equations* **6**:3 (1981), 329–359. MR 82f:35153 Zbl 0472.35070

[Gallagher 2001] I. Gallagher, "Profile decomposition for solutions of the Navier–Stokes equations", *Bull. Soc. Math. France* **129**:2 (2001), 285–316. MR 2002h:35235 Zbl 0987.35120

[Gérard 1998] P. Gérard, "Description du défaut de compacité de l'injection de Sobolev", *ESAIM Control Optim. Calc. Var.* **3** (1998), 213–233. MR 99h:46051

[Germain 2006] P. Germain, "Multipliers, paramultipliers, and weak-strong uniqueness for the Navier–Stokes equations", *J. Differential Equations* **226**:2 (2006), 373–428. MR 2007c:35128 Zbl 1104.35033

[Germain 2008] P. Germain, "Strong solutions and weak-strong uniqueness for the nonhomogeneous Navier–Stokes system", *J. Anal. Math.* **105** (2008), 169–196. MR 2009j:35268 Zbl 1198.35179

[Gibbon 2012] J. D. Gibbon, "A hierarchy of length scales for weak solutions of the three-dimensional Navier–Stokes equations", *Commun. Math. Sci.* **10**:1 (2012), 131–136. MR 2901304

[Heywood 1980] J. G. Heywood, "The Navier–Stokes equations: on the existence, regularity and decay of solutions", *Indiana Univ. Math. J.* **29**:5 (1980), 639–681. MR 81k:35131 Zbl 0494.35077

[Iskauriaza et al. 2003] L. Iskauriaza, G. A. Serëgin, and V. Shverak, "$L_{3,\infty}$-solutions of Navier–Stokes equations and backward uniqueness", *Uspekhi Mat. Nauk* **58**:2(350) (2003), 3–44. In Russian. Translated in *Russian Math. Surveys*, **58**:2 (2003), 211–250. MR 2004m:35204 Zbl 1064.35134

[Kato 1984] T. Kato, "Strong L^p-solutions of the Navier–Stokes equation in $\mathbf{R}^m$, with applications to weak solutions", *Math. Z.* **187**:4 (1984), 471–480. MR 86b:35171 Zbl 0545.35073

[Kato and Ponce 1988] T. Kato and G. Ponce, "Commutator estimates and the Euler and Navier–Stokes equations", *Comm. Pure Appl. Math.* **41**:7 (1988), 891–907. MR 90f:35162 Zbl 0671.35066

[Koch and Tataru 2001] H. Koch and D. Tataru, "Well-posedness for the Navier–Stokes equations", *Adv. Math.* **157**:1 (2001), 22–35. MR 2001m:35257 Zbl 0972.35084

[Ladyzhenskaya 1967] O. A. Ladyzhenskaya, "Uniqueness and smoothness of generalized solutions of Navier–Stokes equations", *Zap. Naučn. Sem. Leningrad. Otdel. Mat. Inst. Steklov.* (*LOMI*) **5** (1967), 169–185. In Russian. Translated in *Sem. Math. Steklov Math. Inst., Leningrad* **5** (1969), 60–67. MR 38 #4836

[Lemarie-Rieusset 1992] P. G. Lemarie-Rieusset, "Analyses multi-résolutions non orthogonales, commutation entre projecteurs et dérivation et ondelettes vecteurs à divergence nulle", *Rev. Mat. Iberoamericana* **8**:2 (1992), 221–237. MR 94d:42044

[Lemarié-Rieusset 1999] P. G. Lemarié-Rieusset, "Solutions faibles d'énergie infinie pour les équations de Navier–Stokes dans $\mathbb{R}^3$", *C. R. Acad. Sci. Paris Sér. I Math.* **328**:12 (1999), 1133–1138. MR 2000d:35179 Zbl 0930.35125

[Leray 1934] J. Leray, "Sur le mouvement d'un liquide visqueux emplissant l'espace", *Acta. Math.* **63** (1934), 183–248. JFM 60.0726.05

[Lions 1969] J.-L. Lions, *Quelques méthodes de résolution des problèmes aux limites non linéaires*, Dunod, 1969. MR 41 #4326 Zbl 0189.40603

[Lions 1996] P.-L. Lions, *Mathematical topics in fluid mechanics, I: Incompressible models*, Oxford Lecture Series in Mathematics and its Applications **3**, Oxford University Press, New York, 1996. MR 98b:76001 Zbl 0866.76002

[Prodi 1959] G. Prodi, "Un teorema di unicità per le equazioni di Navier–Stokes", *Ann. Mat. Pura Appl.* (4) **48** (1959), 173–182. MR 23 #A3384 Zbl 0148.08202

[Qian 2009] Z. Qian, "An estimate for the vorticity of the Navier–Stokes equation", *C. R. Math. Acad. Sci. Paris* **347**:1-2 (2009), 89–92. MR 2010h:35296 Zbl 1155.35070

[Rusin and Šverák 2011] W. Rusin and V. Šverák, "Minimal initial data for potential Navier–Stokes singularities", *J. Funct. Anal.* **260**:3 (2011), 879–891. MR 2011m:35276 Zbl 1206.35199

[Scheffer 1976] V. Scheffer, "Partial regularity of solutions to the Navier–Stokes equations", *Pacific J. Math.* **66**:2 (1976), 535–552. MR 56 #12677 Zbl 0325.35064

[Serrin 1962] J. Serrin, "On the interior regularity of weak solutions of the Navier–Stokes equations", *Arch. Rational Mech. Anal.* **9** (1962), 187–195. MR 25 #346 Zbl 0106.18302

[Serrin 1963] J. Serrin, "The initial value problem for the Navier–Stokes equations", pp. 69–98 in *Nonlinear problems* (Madison, WI, 1962), edited by R. E. Langer, Univ. of Wisconsin Press, Madison, WI, 1963. MR 27 #442 Zbl 0115.08502

[Struwe 1988] M. Struwe, "On partial regularity results for the Navier–Stokes equations", *Comm. Pure Appl. Math.* **41**:4 (1988), 437–458. MR 89h:35270 Zbl 0632.76034

[Tao 2006] T. Tao, *Nonlinear dispersive equations: local and global analysis*, CBMS Regional Conference Series in Mathematics **106**, Conference Board of the Mathematical Sciences, Washington, DC, 2006. MR 2008i:35211 Zbl 1106.35001

[Tao 2007] T. Tao, "A quantitative formulation of the global regularity problem for the periodic Navier–Stokes equation", *Dyn. Partial Differ. Equ.* **4**:4 (2007), 293–302. MR 2009d:35255 Zbl 1137.35055

[Vasseur 2010] A. Vasseur, "Higher derivatives estimate for the 3D Navier–Stokes equation", *Ann. Inst. H. Poincaré Anal. Non Linéaire* **27**:5 (2010), 1189–1204. MR 2011k:35172 Zbl 05800075

Received 4 Aug 2011. Revised 7 Aug 2011. Accepted 23 May 2012.

TERENCE TAO: tao@math.ucla.edu
Department of Mathematics, University of California, Los Angeles, Los Angeles, CA 90095-1555, United States

ANALYSIS AND PDE
Vol. 6, No. 1, 2013
dx.doi.org/10.2140/apde.2013.6.109

A VARIATIONAL PRINCIPLE FOR CORRELATION FUNCTIONS FOR UNITARY ENSEMBLES, WITH APPLICATIONS

DORON S. LUBINSKY

In the theory of random matrices for unitary ensembles associated with Hermitian matrices, m-point correlation functions play an important role. We show that they possess a useful variational principle. Let μ be a measure with support in the real line, and K_n be the n-th reproducing kernel for the associated orthonormal polynomials. We prove that, for $m \geq 1$,

$$\det\left[K_n(\mu, x_i, x_j)\right]_{1 \leq i,j \leq m} = m! \sup_P \frac{P^2(\underline{x})}{\int P^2(\underline{t})\, d\mu^{\times m}(\underline{t})}$$

where the supremum is taken over all alternating polynomials P of degree at most $n-1$ in m variables $\underline{x} = (x_1, x_2, \dots, x_m)$. Moreover, $\mu^{\times m}$ is the m-fold Cartesian product of μ. As a consequence, the suitably normalized m-point correlation functions are *monotone decreasing in the underlying measure* μ. We deduce pointwise one-sided universality for arbitrary compactly supported measures, and other limits.

1. Introduction

Let μ be a positive measure on the real line with infinitely many points in its support, and $\int x^j \, d\mu(x)$ finite for $j = 0, 1, 2, \dots$. Then we may define orthonormal polynomials

$$p_n(x) = \gamma_n x^n + \cdots, \quad \gamma_n > 0,$$

satisfying

$$\int p_n p_m \, d\mu = \delta_{mn}.$$

The n-th *reproducing kernel* is

$$K_n(\mu, x, t) = \sum_{j=0}^{n-1} p_j(x) p_j(t)$$

and the n-th *Christoffel function* is

$$\lambda_n(\mu, x) = 1/K_n(\mu, x, x) = 1 \Big/ \sum_{j=0}^{n-1} p_j^2(x). \tag{1-1}$$

Research supported by NSF grant DMS1001182 and US-Israel BSF grant 2008399.
MSC2010: 15B52, 60B20, 60F99, 42C05, 33C50.
Keywords: orthogonal polynomials, random matrices, unitary ensembles, correlation functions, Christoffel functions.

It admits an extremal property that is very useful in investigating asymptotics of orthogonal polynomials [Nevai 1986; Simon 2011]:

$$\lambda_n(\mu, x) = \inf_{\deg(P)<n} \frac{\int P(t)^2 \, d\mu(t)}{P^2(x)}.$$

Equivalently,

$$K_n(\mu, x, x) = \sup_{\deg(P)<n} \frac{P^2(x)}{\int P(t)^2 \, d\mu(t)}. \tag{1-2}$$

We shall prove a direct generalization for $\det[K_n(\mu, x_i, x_j)]_{1\le i,j\le m}$, a determinant that plays a key role in analysis of random matrices.

Random Hermitian matrices rose to prominence with the work of Eugene Wigner, who used their eigenvalues as a model for scattering theory of heavy nuclei. One places a probability distribution on the entries of an n by n Hermitian matrix. When expressed in "spectral form", that is, as a probability distribution on the (real) eigenvalues $x_1, x_2, \ldots, x_n$, it has the form

$$\mathcal{P}^{(n)}(x_1, x_2, \ldots, x_n) = \frac{\left(\prod_{1\le j<k\le n}(x_k - x_j)^2\right) d\mu(x_1)\, d\mu(x_2) \cdots d\mu(x_n)}{\int \cdots \int \left(\prod_{1\le j<k\le n}(t_k - t_j)^2\right) d\mu(t_1) \cdots d\mu(t_n)};$$

see [Deift 1999, p. 102]. Given $1 \le m \le n$, we define the m-point correlation function

$$R_m^n(\mu; x_1, \ldots, x_m) = \frac{n!}{(n-m)!} \frac{\int \cdots \int \left(\prod_{1\le j<k\le n}(x_k - x_j)^2\right) d\mu(x_{m+1}) \cdots d\mu(x_n)}{\int \cdots \int \left(\prod_{1\le j<k\le n}(t_k - t_j)^2\right) d\mu(t_1) \cdots d\mu(t_n)}. \tag{1-3}$$

Thus R_m^n is, up to normalization, a marginal distribution, where we integrate out $x_{m+1}, x_{m+2}, \ldots, x_n$. Note that we exclude from R_m^n a factor of $\mu'(x_1)\mu'(x_2)\cdots\mu'(x_m)$, which is used by Deift. It is a well established fact [Deift 1999, p. 112] that

$$R_m^n(\mu; x_1, x_2, \ldots, x_m) = \det\left[K_n(\mu, x_i, x_j)\right]_{1\le i,j\le m}. \tag{1-4}$$

Again, we emphasize that in [Deift 1999], as distinct from this paper, μ' is absorbed into K_n. Since much of the interest lies in asymptotics as $n \to \infty$, for fixed m, it is obviously easier to handle asymptotics of this fixed size determinant, than to deal with the $(n-m)$-fold integral in (1-3).

R_m^n can be used to describe the local spacing of m-tuples of eigenvalues. For example, if $m = 2$, and $B \subset \mathbb{R}$ is measurable, then [Deift 1999, p. 117]

$$\int_B \int_B R_2^n(\mu; t_1, t_2)\, d\mu(t_1)\, d\mu(t_2)$$

is the expected number of pairs (t_1, t_2) of eigenvalues, with both $t_1, t_2 \in B$.

Of course there are other settings for random matrices that do not involve orthogonal polynomials. There one considers a class of matrices (such as normal matrices or symmetric matrices) where the elements of the matrix are independently distributed, or there are appropriate bounds on the dependence. The methods are quite different, but remarkably, similar limiting results arise [Erdős 2011; Erdős et al. 2010; 2011; Forrester 2010; Tao and Vu 2011].

The formulation of our main result involves $\mathcal{AL}_n^m$, the alternating polynomials of degree at most n in m variables. We say that $P \in \mathcal{AL}_n^m$ if

$$P(x_1, x_2, \dots, x_m) = \sum_{0 \le j_1, j_2, \dots, j_m \le n} c_{j_1 j_2 \cdots j_m} x_1^{j_1} x_2^{j_2} \cdots x_m^{j_m}, \tag{1-5}$$

so that P is a polynomial of degree less than or equal to n in each of its m variables, and in addition is *alternating*, so that for every pair (i, j) with $1 \le i < j \le m$,

$$P(x_1, \dots, x_i, \dots, x_j, \dots, x_m) = -P(x_1, \dots, x_j, \dots, x_i, \dots, x_m). \tag{1-6}$$

Thus swapping variables changes the sign. Sometimes, these are called *skew-symmetric* polynomials.

Observe that if P_i is a univariate polynomial of degree less than or equal to n for each $i = 1, 2, \dots, m$, then

$$P(t_1, t_2, \dots, t_m) = \det\big[P_i(t_j)\big]_{1 \le i,j \le m} \in \mathcal{AL}_n^m. \tag{1-7}$$

The set of such determinants of polynomials is a proper subset of $\mathcal{AL}_n^m$. It is well known, and easy to see, that every alternating polynomial is the product of a Vandermonde determinant and a symmetric polynomial. Thus $P \in \mathcal{AL}_n^m$ if and only if

$$P(t_1, t_2, \dots, t_m) = \Bigg(\prod_{1 \le i < j \le m} (t_j - t_i)\Bigg) S(t_1, t_2, \dots, t_m),$$

where S is symmetric, and of degree less than or equal to $n - m + 1$ in each variable.

Given a fixed m, we shall use the notation

$$\underline{x} = (x_1, x_2, \dots, x_m), \quad \underline{t} = (t_1, t_2, \dots, t_m)$$

while $\mu^{\times m}$ denotes the m-fold Cartesian product of μ, so that

$$d\mu^{\times m}(\underline{t}) = d\mu(t_1)\, d\mu(t_2) \cdots d\mu(t_m). \tag{1-8}$$

We prove:

Theorem 1.1. *Let $m \ge 1$, $n \ge m + 1$. Let $\underline{x} = (x_1, x_2, \dots, x_m)$ be an m-tuple of real numbers. Then*

$$\det\big[K_n(\mu, x_i, x_j)\big]_{1 \le i,j \le m} = m! \sup_{P \in \mathcal{AL}_{n-1}^m} \frac{(P(\underline{x}))^2}{\int (P(\underline{t}))^2 \, d\mu^{\times m}(\underline{t})}. \tag{1-9}$$

The supremum is attained for

$$P(\underline{t}) = \det\big[K_n(\mu, x_i, t_j)\big]_{1 \le i,j \le m}. \tag{1-10}$$

We could also just take the supremum in (1-9) over the strictly smaller class of determinants of the form (1-7). An immediate, but important, consequence is:

Corollary 1.2. *$R_m^n(\mu; x_1, x_2, \dots, x_m)$ is a monotone decreasing function of μ, and a monotone increasing function of n.*

Despite an extensive literature search, I have not found Theorem 1.1 or Corollary 1.2 in the rich literature on random matrices. At the very least, they must be new to those interested in universality limits, because of the applications they have there. We shall present some in Section 2.

The proof of Theorem 1.1 is based on multivariate orthogonal polynomials built from μ. Given $m \geq 1$, and nonnegative integers $j_1, j_2, \ldots, j_m$, we define

$$T_{j_1,j_2,\ldots,j_m}(x_1, x_2, \ldots, x_m) = \det(p_{j_i}(x_k))_{1 \leq i,k \leq m} = \det \begin{bmatrix} p_{j_1}(x_1) & p_{j_1}(x_2) & \ldots & p_{j_1}(x_m) \\ p_{j_2}(x_1) & p_{j_2}(x_2) & \ldots & p_{j_2}(x_m) \\ \vdots & \vdots & \ddots & \vdots \\ p_{j_m}(x_1) & p_{j_m}(x_2) & \ldots & p_{j_m}(x_m) \end{bmatrix}. \tag{1-11}$$

We show that the $\{T_{j_1,j_2,\ldots,j_m}\}_{j_1<j_2<\cdots<j_m}$ form an orthogonal family with respect to $\mu^{\times m}$, and moreover, the m-point correlation function admits an expansion as a sum of squares of $\{T_{j_1,j_2,\ldots,j_m}\}$, just as does K_n in terms of squares of the orthonormal polynomials. We shall need an associated reproducing kernel,

$$K_n^m(\mu, \underline{x}, \underline{t}) = \frac{1}{m!} \sum_{1 \leq j_1 < j_2 < \cdots < j_m \leq n} T_{j_1,j_2,\ldots,j_m}(\underline{x}) T_{j_1,j_2,\ldots,j_m}(\underline{t}). \tag{1-12}$$

Theorem 1.3. (a) *Let* $0 \leq j_1 < j_2 < \cdots < j_m$ *and* $0 \leq k_1 < k_2 < \cdots < k_m$. *Then*

$$\int T_{j_1,j_2,\ldots,j_m}(\underline{t}) T_{k_1,k_2,\ldots,k_m}(\underline{t})\, d\mu^{\times m}(\underline{t}) = m!\, \delta_{j_1k_1}\delta_{j_2k_2}\cdots\delta_{j_mk_m}. \tag{1-13}$$

(b) *For* $P \in \mathcal{A}\mathcal{L}_{n-1}^m$, *and* $\underline{x} \in \mathbb{R}^n$,

$$P(\underline{x}) = \int P(\underline{t}) K_n^m(\mu, \underline{x}, \underline{t})\, d\mu^{\times m}(\underline{t}). \tag{1-14}$$

(c) *For* $\underline{x}, \underline{t} \in \mathbb{R}^n$,

$$\det\big[K_n(\mu, x_i, t_j)\big]_{1 \leq i,j \leq m} = m!\, K_n^m(\mu, \underline{x}, \underline{t}). \tag{1-15}$$

In particular,

$$\det\big[K_n(\mu, x_i, x_j)\big]_{1 \leq i,j \leq m} = \sum_{1 \leq j_1 < j_2 < \cdots < j_m \leq n} (T_{j_1,j_2,\ldots,j_m}(\underline{x}))^2. \tag{1-16}$$

Remarks. (a) In the case $m = 1$, (1-16) reduces to (1-1) for $K_n(\mu, x, x)$. After an extensive literature search, we found that (1-16) already appears for general m in [Erdős 2011, Section 1.5.3]. We may also express it as

$$\det\big[K_n(\mu, x_i, x_j)\big]_{1 \leq i,j \leq m} = \frac{1}{m!} \sum_{1 \leq j_1, j_2, \ldots, j_m \leq n} (T_{j_1,j_2,\ldots,j_m}(\underline{x}))^2, \tag{1-17}$$

as $T_{j_1,j_2,\ldots,j_m}$ vanishes if any two indices j_i are equal.

(b) The expression (1-15) may also be thought of as a Christoffel–Darboux formula, for it expresses the sum (1-12) in a compact form involving an $m \times m$ determinant.

One consequence of the variational principle is a lower bound for ratios of correlation functions:

Theorem 1.4. *Let $m \geq 2$, $n \geq m+1$, and $x_1, x_2, \ldots, x_m$ be distinct real numbers. Define a measure ν by*

$$d\nu(t) = d\mu(t) \prod_{j=2}^{m} (t - x_j)^2.$$

Then

$$K_n(\mu, x_1, x_1) \geq \frac{\det\left[K_n(\mu, x_i, x_j)\right]_{1\leq i,j\leq m}}{\det\left[K_n(\mu, x_i, x_j)\right]_{2\leq i,j\leq m}} \geq \frac{1}{m} K_{n-m+1}(\nu, x_1, x_1) \prod_{j=2}^{m} (x_1 - x_j)^2. \tag{1-18}$$

The upper bound is a well known consequence of inequalities for positive definite matrices. It is the lower bound that is new.

This paper is organized as follows: in Section 2, we state some applications of Theorem 1.1 to asymptotics and universality limits. In Section 3, we first prove Theorem 1.3, and then deduce Theorem 1.1 and Corollary 1.2, followed by Theorem 1.4. Theorems 2.1, 2.2, and 2.3 are proved in Section 4. Theorem 2.4 is proved in Section 5, and Theorem 2.5 and Corollary 2.6 in Section 6.

2. Applications to asymptotics and universality limits

The extremal property (1-2) is essential in proving the following: if μ is any measure with support in $[-1, 1]$, then at every Lebesgue point x of μ in $(-1, 1)$,

$$\liminf_{n\to\infty} \frac{1}{n} K_n(\mu, x, x)\mu'(x) \geq \frac{1}{\pi\sqrt{1-x^2}}. \tag{2-1}$$

Here μ' is understood as the Radon–Nikodym derivative of the absolutely continuous part of μ. This is more commonly formulated for Christoffel functions as

$$\limsup_{n\to\infty} n\lambda_n(\mu, x) \leq \mu'(x)\pi\sqrt{1-x^2}.$$

Barry Simon calls this the *Máté–Nevai–Totik upper bound*. See, for example, [Máté et al. 1991; Simon 2011, Theorem 5.11.1, p. 334; Totik 2000].

Under additional conditions, including regularity of μ, there is equality in (2-1), with a full limit. We say that μ is *regular in the sense of Stahl, Totik, and Ullman*, or just *regular*, if the leading coefficients $\{\gamma_n\}$ of its orthonormal polynomials satisfy

$$\lim_{n\to\infty} {\gamma_n}^{1/n} = \frac{1}{\operatorname{cap}(\operatorname{supp}[\mu])}. \tag{2-2}$$

Here $\operatorname{cap}(\operatorname{supp}[\mu])$ is the logarithmic capacity of the support of μ. We shall need only a very simple criterion for regularity, namely a version of the Erdős–Turán criterion: if the support of μ consists of finitely many intervals, and $\mu' > 0$ a.e. with respect to Lebesgue measure in that support, then μ is regular [Stahl and Totik 1992, p. 102].

Máté, Nevai and Totik [Máté et al. 1991] showed that if μ is a regular measure with support $[-1, 1]$, and in some subinterval I of $(-1, 1)$, we have

$$\int_I \log \mu' > -\infty, \tag{2-3}$$

then for a.e. $x \in I$,

$$\lim_{n\to\infty} \frac{1}{n} K_n(\mu, x, x)\mu'(x) = \frac{1}{\pi\sqrt{1-x^2}}. \tag{2-4}$$

Totik gave a far-reaching extension of this to measures with compact support J [Totik 2000; 2009]. Here one needs the equilibrium measure ν_J for the compact set J, as well as its Radon–Nikodym derivative, which we denote by ω_J. Thus ν_J is the unique probability measure that minimizes the energy integral

$$\iint \log \frac{1}{|s-t|}\, d\nu(s)\, d\nu(t)$$

amongst all probability measures ν with support in J [Ransford 1995; Saff and Totik 1997]. If I is some subinterval of J, then ν_J is absolutely continuous in I, and moreover, $\omega_J > 0$ in the interior I^o of I. In the special case $J = [-1, 1]$, we have

$$d\nu_J(x) = \omega_J(x)\, dx = \frac{dx}{\pi\sqrt{1-x^2}}.$$

Totik showed that if μ is regular, and in some subinterval I of J, we have (2-3), then

$$\lim_{n\to\infty} \frac{1}{n} K_n(\mu, x, x)\mu'(x) = \omega_J(x) \quad \text{ for a.e. } x \in I. \tag{2-5}$$

Further developments are explored in [Simon 2011].

It is a fairly straightforward consequence of this last relation, and the Christoffel–Darboux formula, that, for $m \geq 2$ and a.e. $(x_1, x_2, \dots, x_m) \in I^m$,

$$\lim_{n\to\infty} \frac{1}{n^m} \det\bigl[K_n(\mu, x_i, x_j)\bigr]_{1\leq i,j\leq m} = \prod_{j=1}^{m} \frac{\omega_J(x_j)}{\mu'(x_j)}. \tag{2-6}$$

The right-hand side is interpreted as ∞ if any $\mu'(x_j) = 0$. Thus, the matrix $[K_n(\mu, x_i, x_j)]_{1\leq i,j\leq m}$ behaves essentially like its diagonal. We shall prove this in Section 4. Without having to assume regularity, or (2-3), we can use Theorem 1.1 to prove one-sided versions of (2-6).

For measures μ with compact support J, and $x \in J$, we let

$$\omega_\mu(x) = \inf\{\omega_L(x) : L \subset J \text{ is compact, } \mu_{|L} \text{ is regular, } x \in L\}. \tag{2-7}$$

Since ν_L decreases as L increases, one can roughly think of ω_μ as the density of the equilibrium measure of the largest set to whose restriction μ is regular. In the sequel, J^o denotes the interior of J.

Theorem 2.1. *Let μ have compact support J, of positive Lebesgue measure, and let ω_J denote the equilibrium density of J. Let $m \geq 1$.*

(a) *For Lebesgue a.e.* $(x_1, x_2, \dots, x_m) \in (J^o)^m$,

$$\liminf_{n\to\infty} \frac{1}{n^m} \det\left[K_n(\mu, x_i, x_j)\right]_{1\le i,j\le m} \ge \prod_{j=1}^{m} \frac{\omega_J(x_j)}{\mu'(x_j)}. \tag{2-8}$$

The right-hand side is interpreted as ∞ *if any* $\mu'(x_j) = 0$.

(b) *Suppose that* I *is a compact subset of* J *consisting of finitely many intervals*, *for which* (2-3) *holds. Then*, *for Lebesgue a.e.* $(x_1, x_2, \dots, x_m) \in I^m$,

$$\limsup_{m\to\infty} \frac{1}{n^m} \det\left[K_n(\mu, x_i, x_j)\right]_{1\le i,j\le m} \le \prod_{j=1}^{m} \frac{\omega_\mu(x_j)}{\mu'(x_j)}. \tag{2-9}$$

A perhaps more impressive application of Theorem 1.1 is to universality limits in the bulk, which describe local spacing of eigenvalues of random Hermitian matrices [Deift 1999; Deift and Gioev 2009; Forrester 2010; Mehta 1991]. One of the more standard formulations, for a measure μ supported on $[-1, 1]$, is

$$\begin{aligned}\lim_{n\to\infty}&\left(\frac{\mu'(x)\pi\sqrt{1-x^2}}{n}\right)^m R_m^n\left(\mu; x + a_1\frac{\pi\sqrt{1-x^2}}{n}, \dots, x + a_m\frac{\pi\sqrt{1-x^2}}{n}\right)\\ &= \lim_{n\to\infty}\left(\frac{\mu'(x)\pi\sqrt{1-x^2}}{n}\right)^m \det\left[K_n\left(\mu; x + a_i\frac{\pi\sqrt{1-x^2}}{n}, x + a_j\frac{\pi\sqrt{1-x^2}}{n}\right)\right]_{1\le i,j\le m}\\ &= \det(S(a_i - a_j))_{1\le i,j\le m},\end{aligned}$$

where

$$S(t) = \frac{\sin \pi t}{\pi t} \tag{2-10}$$

is the sine (or sinc) kernel. There is a vast literature for universality limits, especially in the case where μ is replaced by varying weights. A great many methods have been applied, including classical asymptotics for orthonormal polynomials, Riemann Hilbert techniques, and theory of entire functions of exponential type [Baik et al. 2003; 2008; Deift 1999; Deift and Gioev 2009; Deift et al. 1999; Findley 2008; Forrester 2010; Levin and Lubinsky 2008; Lubinsky 2009a; Simon 2008a; 2011; Totik 2009].

For fixed measures μ with compact support J, the most general pointwise result is due to Totik [2009]. It asserts that if μ is regular, while (2-3) holds in some interval I in the support, then, for a.e. $x \in I$, and all real $a_1, a_2, \dots, a_m$, there are limits for the scaled reproducing kernels that immediately yield

$$\lim_{n\to\infty}\left(\frac{\mu'(x)}{n\omega_J(x)}\right)^m R_m^n\left(\mu; x + \frac{a_1}{n\omega_J(x)}, \dots, x + \frac{a_m}{n\omega_J(x)}\right) = \det(S(a_i - a_j))_{1\le i,j\le m}.$$

Simon [2008a; 2008b] had a similar result, proved using Jost functions. Totik used the comparison method of [Lubinsky 2009a], together with "polynomial pullbacks". Without any local or global restrictions on μ, we showed in [Lubinsky 2012] that universality holds in measure in $\{\mu' > 0\} = \{x : \mu'(x) > 0\}$.

We prove pointwise, almost everywhere, one-sided universality, without any local or global restrictions on μ:

Theorem 2.2. *Let μ have compact support J, and let ω_J denote the equilibrium density of J. Let $m \geq 1$.*

(a) *For a.e. $x \in J^o \cap \{\mu' > 0\}$, and for all real $a_1, a_2, \dots, a_m$,*

$$\liminf_{n\to\infty}\left(\frac{\mu'(x)}{n\omega_J(x)}\right)^m R_m^n\left(\mu; x+\frac{a_1}{n\omega_J(x)}, \dots, x+\frac{a_m}{n\omega_J(x)}\right) \geq \det(S(a_i-a_j))_{1\leq i,j\leq m}. \tag{2-11}$$

(b) *Suppose that I is a compact subset of J consisting of finitely many intervals, for which* (2-3) *holds. Then for a.e. $x \in I$, and for all real $a_1, a_2, \dots, a_m$,*

$$\limsup_{n\to\infty}\left(\frac{\mu'(x)}{n\omega_\mu(x)}\right)^m R_m^n\left(\mu; x+\frac{a_1}{n\omega_\mu(x)}, \dots, x+\frac{a_m}{n\omega_\mu(x)}\right) \leq \det(S(a_i-a_j))_{1\leq i,j\leq m}. \tag{2-12}$$

Pointwise universality at a given point x seems to usually require at least something like μ' being continuous at x, or x being a Lebesgue point of μ. Indeed, when μ' has a jump discontinuity, the universality limit is different from the sine kernel [Foulquié Moreno et al. 2011], and involves de Branges spaces [Lubinsky 2009b]. In our next result, we show that one can still bound the behavior of the correlation function above and below near such a given x. It is noteworthy, though, that pure singularly continuous measures can exhibit sine kernel behavior [Breuer 2011].

Theorem 2.3. *Let μ have compact support J, be regular, and let ω_J denote the equilibrium density of J. Assume that the singular part μ_s of μ satisfies, at a given x in the interior of J,*

$$\lim_{h\to 0+} \mu_s[x-h, x+h]/h = 0. \tag{2-13}$$

Assume moreover that the derivative μ' of the absolutely continuous part of μ satisfies

$$0 < C_1 = \liminf_{t\to x} \mu'(t) \leq \limsup_{t\to x} \mu'(t) = C_2 < \infty. \tag{2-14}$$

Then, for all real $a_1, a_2, \dots, a_m$,

$$\begin{aligned} C_2^{-m}\det(S(a_i-a_j))_{1\leq i,j\leq m} &\leq \liminf_{n\to\infty}\left(\frac{1}{n\omega_J(x)}\right)^m R_m^n\left(\mu; x+\frac{a_1}{n\omega_J(x)}, \dots, x+\frac{a_m}{n\omega_J(x)}\right) \\ &\leq \limsup_{n\to\infty}\left(\frac{1}{n\omega_J(x)}\right)^m R_m^n\left(\mu; x+\frac{a_1}{n\omega_J(x)}, \dots, x+\frac{a_m}{n\omega_J(x)}\right) \\ &\leq C_1^{-m}\det(S(a_i-a_j))_{1\leq i,j\leq m}. \end{aligned} \tag{2-15}$$

At the boundary of the support of the measure (referred to as the edge of the spectrum in random matrix theory), the universality limit takes a different form [Forrester 2010; Kuijlaars and Vanlessen 2002]. For fixed measures that behave like Jacobi weights near the endpoints, they involve the Bessel kernel of order $\alpha > -1$:

$$\mathbb{J}_\alpha(u,v) = \frac{J_\alpha(\sqrt{u})\sqrt{v}J'_\alpha(\sqrt{v}) - J_\alpha(\sqrt{v})\sqrt{u}J'_\alpha(\sqrt{u})}{2(u-v)}.$$

Here J_α is the usual Bessel function of the first kind and order α. Using a comparison method, the author proved [Lubinsky 2008] that if μ is a regular measure on $[-1,1]$, and μ is absolutely continuous in some left neighborhood $(1-\eta, 1]$ of 1, and there $\mu'(t) = h(t)(1-t)^\alpha$, where $h(1) > 0$ and h is continuous at 1, then

$$\lim_{n\to\infty} \frac{1}{2n^2} \tilde{K}_n\left(\mu, 1-\frac{a}{2n^2}, 1-\frac{b}{2n^2}\right) = \mathbb{J}_\alpha(a,b), \tag{2-16}$$

uniformly for a, b in compact subsets of $(0,\infty)$. Here, and in the sequel,

$$\tilde{K}_n(\mu,x,y) = \mu'(x)^{1/2}\mu'(y)^{1/2}K_n(\mu,x,y).$$

When $\alpha \geq 0$, we may allow also $a, b = 0$. This has the immediate consequence that, for $m \geq 2$, and $a_1, a_2, \dots, a_m > 0$,

$$\lim_{n\to\infty}\left(\frac{1}{2n^2}\right)^m R_m^n\left(\mu; 1-\frac{a_1}{2n^2}, \dots, 1-\frac{a_m}{2n^2}\right)\left(\prod_{j=1}^m \mu'\left(1-\frac{a_j}{2n^2}\right)\right) = \det(\mathbb{J}_\alpha(a_i,a_j))_{1\leq i,j\leq m}. \tag{2-17}$$

Under weak conditions at the edge, we can prove one-sided universality:

Theorem 2.4. *Let μ have support contained in $[-1,1]$ and let 1 be the right endpoint of that support. Assume that μ is absolutely continuous near 1, and, for some $\alpha > -1$,*

$$0 < C_1 = \liminf_{t\to 1-} \mu'(t)(1-t)^{-\alpha} \leq \limsup_{t\to 1-} \mu'(t)(1-t)^{-\alpha} = C_2 < \infty. \tag{2-18}$$

Then, for $a_1, a_2, \dots, a_m > 0$,

$$\liminf_{n\to\infty}\left(\frac{1}{2n^2}\right)^m R_m^n\left(\mu; 1-\frac{a_1}{2n^2}, \dots, 1-\frac{a_m}{2n^2}\right)\prod_{j=1}^m \mu'\left(1-\frac{a_j}{2n^2}\right) \geq \left(\frac{C_1}{C_2}\right)^m \det(\mathbb{J}_\alpha(a_i,a_j))_{1\leq i,j\leq m}. \tag{2-19}$$

If $\alpha \geq 0$, we may also allow $a_1, a_2, \dots, a_m \geq 0$.

We note that if, in addition, μ has support $[-1,1]$ and is regular, then we may replace the lim inf by lim sup, the asymptotic lower bound by an upper bound, provided we replace $(C_1/C_2)^m$ by $(C_2/C_1)^m$.

Our final result has a comparison or "localization" flavor, generalizing similar results for Christoffel functions. Recall that a set $J \subset \mathbb{R}$ is said to be regular for the Dirichlet problem [Ransford 1995; Stahl and Totik 1992] if, for every function f continuous on J, there exists a function harmonic in $\bar{\mathbb{C}}\backslash J$, continuous on $\mathbb{C}$, whose restriction to J is f. Of course, this is confusing when juxtaposed with the notion of a regular measure!

Theorem 2.5. *Let μ, ν have compact support J and both be regular. Assume that J is regular with respect to the Dirichlet problem. Let $\xi \in J$ and $\mu'(\xi)$, $\nu'(\xi)$ be finite and positive, with*

$$\lim_{\operatorname{dist}(I,\xi)\to 0} \frac{\mu(I)}{\nu(I)} = \frac{\mu'(\xi)}{\nu'(\xi)}, \tag{2-20}$$

where the limit is taken over intervals I of length $|I|$, and $\operatorname{dist}(I,\xi)=\sup\{|x-\xi|:x\in I\}$. *Let $m\geq 1$. Assume that, for $n\geq 1$,*

$$\underline{y}_n=(y_{1n},y_{2n},\ldots,y_{mn})$$

is a vector of real numbers satisfying

$$\lim_{n\to\infty}\Big(\max_{1\leq j\leq m}|y_{mj}-\xi|\Big)=0, \tag{2-21}$$

and

$$\lim_{\varepsilon\to 0+}\left(\limsup_{n\to\infty}\left|\frac{K^m_{[n(1\pm\varepsilon)]}(\nu,\underline{y}_n,\underline{y}_n)}{K^m_n(\nu,\underline{y}_n,\underline{y}_n)}-1\right|\right)=0. \tag{2-22}$$

Then

$$\lim_{n\to\infty}\frac{K^m_n(\mu,\underline{y}_n,\underline{y}_n)}{K^m_n(\nu,\underline{y}_n,\underline{y}_n)}=\left(\frac{\nu'(\xi)}{\mu'(\xi)}\right)^m. \tag{2-23}$$

Of course, in (2-22), $[n(1\pm\varepsilon)]$ denotes the integer part of $n(1\pm\varepsilon)$. As an immediate consequence, we obtain:

Corollary 2.6. *Let μ, ν have compact support J and be regular. Assume that J is regular with respect to the Dirichlet problem. Let $x\in J$ and $\mu'(x)$, $\nu'(x)$ be finite and positive, with* (2-20) *holding at $\xi=x$. Assume that, for given $m\geq 2$ and all real $a_1,a_2,\ldots,a_m$,*

$$\lim_{n\to\infty}\left(\frac{\nu'(x)}{n\omega_J(x)}\right)^m R^n_m\left(\nu;x+\frac{a_1}{n\omega_J(x)},\ldots,x+\frac{a_m}{n\omega_J(x)}\right)=\det(S(a_i-a_j))_{1\leq i,j\leq m}. \tag{2-24}$$

Then, for all real $a_1,a_2,\ldots,a_m$,

$$\lim_{n\to\infty}\left(\frac{\mu'(x)}{n\omega_J(x)}\right)^m R^n_m\left(\mu;x+\frac{a_1}{n\omega_J(x)},\ldots,x+\frac{a_m}{n\omega_J(x)}\right)=\det(S(a_i-a_j))_{1\leq i,j\leq m}. \tag{2-25}$$

3. Proofs of Theorems 1.1, 1.3, 1.4 and Corollary 1.2

Proof of Theorem 1.3(a). We use σ and η to denote permutations of $(1,2,\ldots,m)$ with respective signs ε_σ and ε_η. We see that

$$\begin{aligned}
I&=\int\cdots\int T_{j_1,j_2,\ldots,j_m}(t_1,t_2,\ldots,t_m)T_{k_1,k_2,\ldots,k_m}(t_1,t_2,\ldots,t_m)\,d\mu(t_1)\cdots d\mu(t_m)\\
&=\sum_{\sigma,\eta}\varepsilon_\sigma\varepsilon_\eta\int\cdots\int\Big(\prod_{i=1}^m p_{j_{\sigma(i)}}(t_i)\Big)\Big(\prod_{i=1}^m p_{k_{\eta(i)}}(t_i)\Big)\,d\mu(t_1)\cdots d\mu(t_m)\\
&=\sum_{\sigma,\eta}\varepsilon_\sigma\varepsilon_\eta\prod_{i=1}^m\delta_{j_{\sigma(i)}k_{\eta(i)}}=\sum_{\sigma,\eta}\varepsilon_\sigma\varepsilon_\eta\prod_{\ell=1}^m\delta_{j_\ell k_{\eta(\sigma^{-1}(\ell))}},
\end{aligned} \tag{3-1}$$

where σ^{-1} is the inverse of the permutation σ. For a term in this last sum to be nonzero, we need

$$j_\ell=k_{\eta(\sigma^{-1}(\ell))}\quad\text{for all } 1\leq\ell\leq m. \tag{3-2}$$

Since $j_1 < j_2 < \cdots < j_m$ and $k_1 < k_2 < \cdots < k_m$, we see that this will fail unless

$$\eta(\sigma^{-1}(\ell)) = \ell \quad \text{for all } 1 \le \ell \le m.$$

Indeed, if $\eta(\sigma^{-1}(i)) \neq i$ for some smallest i, then $j_{i-1} = k_{i-1}$ but either $j_i = k_{\eta(\sigma^{-1}(i))} \ge k_{i+1}$ or $j_i = k_{\eta(\sigma^{-1}(i))} \le k_{i-1}$. In the former case, all of $j_i, j_{i+1}, \ldots, j_m > k_i$, and k_i is omitted from the equalities in (3-2), a contradiction. In the latter case, we obtain $j_i \le j_{i-1}$, contradicting the strict monotonicity of the j's. Thus necessarily $\eta = \sigma$, so (3-1) becomes, under (3-2),

$$I = \sum_{\sigma} \varepsilon_\sigma^2 = m!. \qquad \square$$

Proof of Theorem 1.3(b). We first show that every $P \in \mathscr{AL}_{n-1}^m$ is a linear combination of the T polynomials. We can write

$$P(x_1, x_2, \ldots, x_m) = \sum_{0 \le j_1, j_2, \ldots, j_m < n} c_{j_1 j_2 \cdots j_m} p_{j_1}(x_1) p_{j_2}(x_2) \cdots p_{j_m}(x_m).$$

Because of the alternating property (1-6), and the linear independence of

$$\{p_{j_1}(x_1) p_{j_2}(x_2) \cdots p_{j_m}(x_m)\}_{1 \le j_1, j_2, \ldots, j_m \le n},$$

necessarily, when we swap indices j_k and j_ℓ, the coefficients change sign; that is,

$$c_{j_1 \cdots j_k \cdots j_\ell \cdots j_m} = -c_{j_1 \cdots j_\ell \cdots j_k \cdots j_m}.$$

In particular, coefficients vanish if any two subscripts coincide. More generally, this implies that if σ is a permutation of $\{1, 2, \ldots, m\}$ with sign ε_σ, then

$$c_{j_{\sigma(1)} j_{\sigma(2)} \cdots j_{\sigma(m)}} = \varepsilon_\sigma c_{j_1 j_2 \cdots j_m}.$$

Next, given distinct $0 \le j_1, j_2, \ldots, j_m < n$, let $\tilde{j}_1 < \tilde{j}_2 < \cdots < \tilde{j}_m$ denote these indices in increasing order. We can write, for some permutation σ,

$$j_i = \tilde{j}_{\sigma(i)}, \quad 1 \le i \le m.$$

Conversely, for the given $\{\tilde{j}_i\}$, every such permutation σ defines indices $\{j_i\}$ with $0 \le j_1, j_2, \ldots, j_m < n$. Thus

$$\begin{aligned}
P(x_1, x_2, \ldots, x_m) &= \sum_{0 \le \tilde{j}_1 < \tilde{j}_2 < \cdots < \tilde{j}_m < n} c_{\tilde{j}_1 \tilde{j}_2 \cdots \tilde{j}_m} \sum_{\sigma} \varepsilon_\sigma p_{\tilde{j}_{\sigma(1)}}(x_1) p_{\tilde{j}_{\sigma(2)}}(x_2) \cdots p_{\tilde{j}_{\sigma(m)}}(x_m) \\
&= \sum_{0 \le \tilde{j}_1 < \tilde{j}_2 < \cdots < \tilde{j}_m < n} c_{\tilde{j}_1 \tilde{j}_2 \cdots \tilde{j}_m} \det\bigl[p_{\tilde{j}_i}(x_k)\bigr]_{1 \le i,k \le m} \\
&= \sum_{0 \le \tilde{j}_1 < \tilde{j}_2 < \cdots < \tilde{j}_m < n} c_{\tilde{j}_1 \tilde{j}_2 \cdots \tilde{j}_m} T_{\tilde{j}_1 \tilde{j}_2 \cdots \tilde{j}_m}(x_1, x_2, \ldots, x_m). \qquad (3\text{-}3)
\end{aligned}$$

Inasmuch as each $T_{\tilde{j}_1 \tilde{j}_2 \cdots \tilde{j}_m}$ lies in $\mathscr{AL}_{n-1}^m$, we have shown that $\mathscr{AL}_{n-1}^m$ is the linear span of the T

polynomials, and (3-3) is an orthogonal expansion. Orthogonality in the form (1-13) gives

$$c_{\tilde{j}_1\tilde{j}_2\cdots\tilde{j}_m} = \frac{1}{m!}\int P(\underline{t})T_{\tilde{j}_1\tilde{j}_2\cdots\tilde{j}_m}(\underline{t})\,d\mu^{\times m}(\underline{t}).$$

Now our definition (1-12) of the reproducing kernel gives (1-14). □

Proof of Theorem 1.3(c). Fix $\underline{x} = (x_1, x_2, \ldots, x_m)$. Let

$$P(\underline{t}) = P(t_1, t_2, \ldots, t_m) = \det\big[K_n(\mu, x_i, t_j)\big]_{1\le i,j\le m}. \tag{3-4}$$

By successively extracting the sums from the 1st, 2nd, …, m-th rows, we see that

$$\begin{aligned} P(\underline{t}) &= \det\begin{bmatrix} \sum_{j_1=0}^{n-1} p_{j_1}(x_1)p_{j_1}(t_1) & \ldots & \sum_{j_1=0}^{n-1} p_{j_1}(x_1)p_{j_1}(t_m) \\ \vdots & \ddots & \vdots \\ \sum_{j_m=0}^{n-1} p_{j_m}(x_m)p_{j_m}(t_1) & \ldots & \sum_{j_m=0}^{n-1} p_{j_m}(x_m)p_{j_1}(t_m) \end{bmatrix} \\ &= \sum_{j_1=0}^{n-1}\cdots\sum_{j_m=0}^{n-1}\big(p_{j_1}(x_1)\cdots p_{j_m}(x_m)\big)T_{j_1j_2\cdots j_m}(t_1, t_2, \ldots, t_m). \end{aligned}$$

When $j_i = j_k$ for distinct i, k, then $T_{j_1j_2\cdots j_m} = 0$. Thus only terms with $j_1, j_2, \ldots, j_m$ distinct are nonzero. As in the proof of Theorem 1.3(b), given distinct $0 \le j_1, j_2, \ldots, j_m < n$, we can write, for some permutation σ uniquely determined by these indices,

$$j_i = \tilde{j}_{\sigma(i)}$$

where $0 \le \tilde{j}_1 < \tilde{j}_2 < \cdots < \tilde{j}_m < n$. As there, this yields

$$\begin{aligned} P(\underline{t}) &= \sum_{0\le\tilde{j}_1<\tilde{j}_2<\cdots<\tilde{j}_m<n}\ \sum_{\sigma}\varepsilon_\sigma\big(p_{\tilde{j}_{\sigma(1)}}(x_1)\cdots p_{\tilde{j}_{\sigma(m)}}(x_m)\big)T_{\tilde{j}_1\tilde{j}_2\cdots\tilde{j}_m}(t_1, t_2, \ldots, t_m) \\ &= \sum_{0\le\tilde{j}_1<\tilde{j}_2<\cdots<\tilde{j}_m<n} T_{\tilde{j}_1\tilde{j}_2\cdots\tilde{j}_m}(x_1, x_2, \ldots, x)T_{\tilde{j}_1\tilde{j}_2\cdots\tilde{j}_m}(t_1, t_2, \ldots, t_m). \end{aligned}$$

So

$$\det\big[K_n(\mu, x_i, t_j)\big]_{1\le i,j\le m} = P(\underline{t}) = m!\,K_n^m(\mu, \underline{x}, \underline{t}),$$

and we have (1-15). Then (1-16) follows from (1-12). □

Proof of Theorem 1.1. By the reproducing kernel relation (1-14), and Cauchy–Schwarz, for all $P \in \mathcal{AL}_{n-1}^m$,

$$P(\underline{x})^2 \le \left(\int P(\underline{t})^2\,d\mu^{\times m}(\underline{t})\right)\left(\int K_n^m(\mu, \underline{x}, \underline{t})^2\,d\mu^{\times m}(\underline{t})\right) = \left(\int P(\underline{t})^2\,d\mu^{\times m}(\underline{t})\right)K_n^m(\mu, \underline{x}, \underline{x}).$$

Thus

$$K_n^m(\mu, \underline{x}, \underline{x}) \ge \sup_{P\in\mathcal{AL}_{n-1}^m}\frac{(P(\underline{x}))^2}{\int (P(\underline{t}))^2\,d\mu^{\times m}(\underline{t})}. \tag{3-5}$$

By choosing P as in (3-4), we obtain equality in (3-5). Now (1-9) follows from (1-15). □

Proof of Corollary 1.2. This follows immediately from (1-9) and the positivity of all the terms there. □

Proof of Theorem 1.4. The upper bound in (1-18) is a standard inequality for determinants involving symmetric positive definite matrices. See, for example, [Beckenbach and Bellman 1961, Theorem 7, p. 63]. For the lower bound, let $R(t_2, t_3, \ldots, t_m) \in \mathcal{AL}_{m-1}^{n-1}$. Let P be a univariate polynomial of degree less than or equal to $n-1$ satisfying $P(x_j) = 0$, $2 \le j \le m$. Let

$$S(t_1, t_2, \ldots, t_m) = \sum_{j=1}^{m} P(t_j)(-1)^j R(t_1, t_2, \ldots, t_{j-1}, t_{j+1}, \ldots, t_m).$$

We claim that $S \in \mathcal{AL}_m^{n-1}$. Suppose we swap the variables t_k and t_ℓ, where $1 \le k < \ell \le m$. The terms involving $P(t_k)$ and $P(t_\ell)$ before the variable swap are

$$\begin{aligned} &P(t_k)(-1)^k R(t_1, \ldots, t_{k-1}, t_{k+1}, \ldots, t_{\ell-1}, t_\ell, t_{\ell+1}, \ldots, t_m) \\ &\qquad + P(t_\ell)(-1)^\ell R(t_1, \ldots, t_{k-1}, t_k, t_{k+1}, \ldots, t_{\ell-1}, t_{\ell+1}, \ldots, t_m) \end{aligned}$$

and become, after swapping t_k, t_ℓ,

$$\begin{aligned} &P(t_\ell)(-1)^k R(t_1, \ldots, t_{k-1}, t_{k+1}, \ldots, t_{\ell-1}, t_k, t_{\ell+1}, \ldots, t_m) \\ &\qquad + P(t_k)(-1)^\ell R(t_1, \ldots, t_{k-1}, t_\ell, t_{k+1}, \ldots, t_{\ell-1}, t_{\ell+1}, \ldots, t_m). \end{aligned}$$

Using $\ell - k - 1$ swaps of adjacent variables in each R term, the alternating property of R gives

$$\begin{aligned} &-\bigl\{P(t_\ell)(-1)^\ell R(t_1, \ldots, t_{k-1}, t_k, t_{k+1}, \ldots, t_{\ell-1}, t_{\ell+1}, \ldots, t_m) \\ &\qquad + P(t_k)(-1)^k R(t_1, \ldots, t_{k-1}, t_{k+1}, \ldots, t_{\ell-1}, t_\ell, t_{\ell+1}, \ldots, t_m)\bigr\}. \end{aligned}$$

In the remaining terms $P(t_j)(-1)^j R(t_1, t_2, \ldots, t_{j-1}, t_{j+1}, \ldots, t_m)$ with $j \ne k, \ell$, we swap t_k and t_ℓ, and use the alternating property to obtain $-P(t_j)(-1)^j R(t_1, t_2, \ldots, t_{j-1}, t_{j+1}, \ldots, t_m)$. So we have proved that $S \in \mathcal{AL}_m^n$. Moreover, as P has zeros at $x_2, x_3, \ldots, x_m$, we have

$$S(x_1, x_2, \ldots, x_m) = -P(x_1) R(x_2, x_3, \ldots, x_m).$$

Next, by Cauchy–Schwarz,

$$\begin{aligned} \int S^2 \, d\mu^{\times m} &\le m \int \sum_{j=1}^{m} P^2(t_j) R^2(t_1, \ldots, t_{j-1}, t_{j+1}, \ldots, t_m) \, d\mu(t_1) \cdots d\mu(t_m) \\ &= m^2 \Bigl(\int P^2 \, d\mu\Bigr)\Bigl(\int R^2 \, d\mu^{\times(m-1)}\Bigr). \end{aligned}$$

Then (1-9) gives

$$\det\bigl[K_n(\mu, x_i, x_j)\bigr]_{1 \le i,j \le m} \ge m! \frac{S^2(x_1, x_2, \ldots, x_m)}{\int S^2 \, d\mu^{\times m}} \ge \frac{m!}{m^2} \frac{P^2(x_1)}{\int P^2 \, d\mu} \frac{R^2(x_2, \ldots, x_m)}{\int R^2 \, d\mu^{\times(m-1)}}.$$

Write

$$P(t) = P_1(t) \prod_{j=2}^{m} (t - x_j),$$

where P_1 is any polynomial of degree at most $n-m$. Next, take the supremum over P_1 of degree at most $n-m$ and $R \in \mathcal{AL}_{m-1}^{n-1}$. Recalling the definition of ν and (1-2) gives

$$\det\big[K_n(\mu,x_i,x_j)\big]_{1\le i,j\le m} \ge \frac{m!}{m^2} K_{n-m+1}(\nu,x_1,x_1)\left(\prod_{j=2}^{m}(x_1-x_j)^2\right)\frac{1}{(m-1)!}\det\big[K_n(\mu,x_i,x_j)\big]_{2\le i,j\le m}.$$

This gives the lower bound in (1-18). □

4. Proofs of Theorems 2.1, 2.2, and 2.3

Lemma 4.1. *Let μ have compact support J, let μ be regular, and assume that I is a subset of the support consisting of finitely many intervals in which* (2-3) *holds. Let $m \ge 2$. Then, for Lebesgue a.e. $(x_1, x_2, \ldots, x_m) \in I^m$,*

$$\lim_{n\to\infty} \frac{1}{n^m} \det\big[K_n(\mu, x_i, x_j)\big]_{1\le i,j\le m} = \prod_{j=1}^{m} \frac{\omega_J(x_j)}{\mu'(x_j)}. \tag{4-1}$$

Proof. We already know that, for a.e. $x \in I$,

$$\lim_{n\to\infty} \frac{1}{n} K_n(\mu, x, x) \frac{\mu'(x)}{\omega_J(x)} = 1, \tag{4-2}$$

by Totik's result (2-5). (Formally, the integral condition (2-3) follows in each of the intervals whose union is I, and hence (2-5) does.) We next show that there is a set $\mathscr{E}$ of Lebesgue measure 0 such that for distinct $x, y \in I\backslash\mathscr{E}$, both (4-2) holds, and

$$\lim_{n\to\infty} \frac{1}{n} K_n(\mu, x, y)\left(\frac{\mu'(x)\mu'(y)}{\omega_J(x)\omega_J(y)}\right)^{1/2} = 0. \tag{4-3}$$

These last two assertions give the result. Indeed for distinct $x_1, x_2 \cdots x_m \in I\backslash\mathscr{E}$, we have

$$\begin{aligned}\frac{1}{n^m} \det\big[K_n(\mu, x_i, x_j)\big]_{1\le i,j\le m} \prod_{j=1}^{m} \frac{\mu'(x_j)}{\omega_J(x_j)} &= \sum_{\sigma} \varepsilon_\sigma \prod_{i=1}^{m}\left(\frac{1}{n} K_n(\mu, x_i, x_{\sigma(i)})\left(\frac{\mu'(x_i)\mu'(x_{\sigma(i)})}{\omega_J(x_i)\omega_J(x_{\sigma(i)})}\right)^{1/2}\right)\\ &= \prod_{i=1}^{m}\left(\frac{1}{n} K_n(\mu, x_i, x_i)\frac{\mu'(x_i)}{\omega_J(x_i)}\right) + o(1) = 1 + o(1),\end{aligned}$$

by (4-2) and (4-3). Of course the set of $x_1, x_2, \ldots, x_m$ where any two $x_i = x_j$ with $i \ne j$ has Lebesgue measure 0 in I^m.

We turn to the proof of (4-3). It follows from (4-2) that there is a set $\mathscr{E}$ of measure 0 such that, for $x \in I\backslash\mathscr{E}$, we have

$$\lim_{n\to\infty} \frac{1}{n} p_n^2(x) = \lim_{n\to\infty} \frac{1}{n}(K_{n+1}(\mu, x, x) - K_n(\mu, x, x)) = 0.$$

Then, for distinct x, y, the Christoffel–Darboux formula gives, for $x, y \in I\backslash\mathscr{E}$,

$$\frac{1}{n} K_n(\mu, x, y) = \frac{1}{n}\frac{\gamma_{n-1}}{\gamma_n}\frac{p_n(x)p_{n-1}(y) - p_{n-1}(x)p_n(y)}{x-y} = o(1).$$

Here we are also using the fact that $\{\gamma_{n-1}/\gamma_n\}$ is bounded as μ has compact support. □

Proof of Theorem 2.1(a). Since $J = \operatorname{supp}[\mu]$ is compact, we can find a decreasing sequence of compact sets $\{J_\ell\}_{\ell=1}^{\infty}$ such that each J_ℓ consists of finitely many disjoint closed intervals, and

$$J = \bigcap_{\ell=1}^{\infty} J_\ell.$$

(This follows by a straightforward covering of J by open intervals, and using compactness, then closing them up; at the $(\ell+1)$-st stage, we ensure that $J_{\ell+1} \subset J_\ell$ by intersecting those intervals in $J_{\ell+1}$ with those in J_ℓ.) For $\ell \geq 1$, let

$$d\mu_\ell(x) = d\mu(x) + \frac{1}{\ell}\omega_{J_\ell}(x)\,dx, \tag{4-4}$$

so that we are adding a (small) multiple of the equilibrium measure for J_ℓ to μ. Because $\omega_{J_\ell} > 0$ in the interior of each J_ℓ, we have $\mu_\ell' > 0$ a.e. in J_ℓ, so μ_ℓ is a regular measure [Stahl and Totik 1992, p. 102]. Moreover, ω_{J_ℓ} is positive and continuous in each compact subinterval I of the interior of J_ℓ, so

$$\int_I \log \mu_\ell' > -\infty. \tag{4-5}$$

By Lemma 4.1, for a.e. $(x_1, x_2, \dots, x_m) \in I^m$,

$$\lim_{n\to\infty} \frac{1}{n^m} \det\bigl[K_n(\mu_\ell, x_i, x_j)\bigr]_{1\leq i,j\leq m} = \prod_{j=1}^{m} \frac{\omega_{J_\ell}(x_j)}{\mu_\ell'(x_j)}.$$

As $\mu_\ell \geq \mu$, Corollary 1.2 gives

$$\liminf_{n\to\infty} \frac{1}{n^m} \det\bigl[K_n(\mu, x_i, x_j)\bigr]_{1\leq i,j\leq m} \geq \prod_{j=1}^{m} \frac{\omega_{J_\ell}(x_j)}{\mu_\ell'(x_j)}. \tag{4-6}$$

Since a countable union of sets of the form I^m exhausts J_ℓ^m, this last relation actually holds for a.e. $(x_1, x_2, \dots, x_m) \in J_\ell^m$. Now, by [Totik 2009, Lemma 4.2], uniformly for x in compact subsets of an open set contained in J,

$$\lim_{\ell\to\infty} \omega_{J_\ell}(x) = \omega_J(x). \tag{4-7}$$

Moreover, ω_J is positive and continuous in that open set. We can now let $\ell \to \infty$ in (4-6) and use the fact that the left-hand side in (4-6) is independent of ℓ to obtain (2-8). □

Proof of Theorem 2.1(b). Let L be a compact subset of $\operatorname{supp}[\mu]$ such that $\mu_{|L}$ is regular. $L = I$ is one such choice, because of the Szegő condition (2-3). We may assume that $I \subset L$, since ω_L decreases as L increases. Let

$$d\nu(x) = \mu'(x)_{|L}\,dx, \tag{4-8}$$

so that $d\nu$ is the restriction to L of the absolutely continuous part of μ. Here $\int_I \log \nu' > -\infty$, so ν satisfies the hypotheses of Lemma 4.1, while $\mu \geq \nu$, so Corollary 1.2, followed by Lemma 4.1, gives, for a.e. $(x_1, x_2, \dots, x_m) \in I^m$,

$$\limsup_{n\to\infty} \frac{1}{n^m} \det\bigl[K_n(\mu, x_i, x_j)\bigr]_{1\le i,j\le m} \le \limsup_{n\to\infty} \frac{1}{n^m} \det\bigl[K_n(\nu, x_i, x_j)\bigr]_{1\le i,j\le m} = \prod_{j=1}^{m} \frac{\omega_L(x_j)}{\mu'(x_j)};$$

recall that $\nu' = \mu'$ in $I \subset L$. Now take the infimum over all such L and use the fact that the left-hand side is independent of L. □

We turn to:

Proof of Theorem 2.2(a). Let μ_ℓ and J_ℓ be as in the proof of Theorem 2.1(a). It then follows from results of Totik [2009, Theorem 2.3] and/or Simon [2011, Theorem 5.11.13, p. 344] that, for a.e. $x \in J_\ell$, and all real $a_1, a_2, \dots a_m$, and $1 \le i, j \le m$,

$$\lim_{n\to\infty} \frac{1}{n} K_n\left(\mu_\ell, x + \frac{a_i}{n}, x + \frac{a_j}{n}\right) = \frac{\omega_{J_\ell}(x)}{\mu_\ell'(x)} S((a_i - a_j)\omega_{J_\ell}(x)).$$

Consequently,

$$\lim_{n\to\infty} \frac{1}{n^m} R_m^n\left(\mu_\ell; x + \frac{a_1}{n}, \dots, x + \frac{a_m}{n}\right) = \left(\frac{\omega_{J_\ell}(x)}{\mu_\ell'(x)}\right)^m \det\bigl(S((a_i - a_j)\omega_{J_\ell}(x))\bigr)_{1\le i,j\le m}.$$

Now we use the fact that $\mu \le \mu_\ell$, and Corollary 1.2: for a.e. $x \in J$, and all $a_1, a_2, \dots, a_m$,

$$\liminf_{n\to\infty} \frac{1}{n^m} R_m^n\left(\mu; x + \frac{a_1}{n}, \dots, x + \frac{a_m}{n}\right) \ge \left(\frac{\omega_{J_\ell}(x)}{\mu_\ell'(x)}\right)^m \det\bigl(S((a_i - a_j)\omega_{J_\ell}(x))\bigr)_{1\le i,j\le m}. \tag{4-9}$$

Moreover we have (4-7). We can now let $\ell \to \infty$ in (4-9), and use the fact that the left-hand side in (4-9) is independent of ℓ to obtain (2-11), with a scale change. □

Proof of Theorem 2.2(b). Let L and ν be as in the proof of Theorem 2.1(b). We can use the aforementioned results of Totik applied to ν, to obtain, for a.e. $x \in I$, and real $a_1, a_2, \dots, a_m$,

$$\lim_{n\to\infty} \frac{1}{n^m} R_m^n\left(\nu; x + \frac{a_1}{n}, \dots, x + \frac{a_m}{n}\right) = \left(\frac{\omega_L(x)}{\nu'(x)}\right)^m \det\bigl(S((a_i - a_j)\omega_L(x))\bigr)_{1\le i,j\le m}. \tag{4-10}$$

Now we use the fact that $\mu \ge \nu$, and that $\mu' = \nu'$ in $I \subset L$ and Corollary 1.2: for a.e. $x \in I$, and real $a_1, a_2, \dots, a_m$,

$$\limsup_{n\to\infty} \frac{1}{n^m} R_m^n\left(\mu; x + \frac{a_1}{n}, \dots, x + \frac{a_m}{n}\right) \le \left(\frac{\omega_L(x)}{\mu'(x)}\right)^m \det\bigl(S((a_i - a_j)\omega_L(x))\bigr)_{1\le i,j\le m}.$$

Now choose a sequence of compact subsets L of $\operatorname{supp}[\mu]$ such that $\omega_L(x)$ converges to the infimum $\omega_\mu(x)$. □

Proof of Theorem 2.3. Let $\eta \in (0, C_1)$, and choose $\delta > 0$ such that, in $(x - \delta, x + \delta)$,

$$C_1 - \eta \le \mu' \le C_2 + \eta.$$

Here μ' denotes the derivative of the absolutely continuous component of μ. Define

$$d\nu = d\mu \quad \text{in } J\backslash(x-\delta, x+\delta)$$

and

$$d\nu(t) = d\mu_s(t) + (C_1 - \eta)\, dt \quad \text{in } (x-\delta, x+\delta).$$

Then $d\nu \leq d\mu$, and ν is regular on J (see [Stahl and Totik 1992, Theorem 5.3.3, p. 148]). Moreover, the derivative ν' of the absolutely continuous part of ν exists and equals $C_1 - \eta$ in $(x-\delta, x+\delta)$, while (2-13) implies that

$$\lim_{h\to 0} \nu_s[x-h, x+h]/h = 0.$$

By a theorem of Totik [2009, Theorem 2.3], we obtain, for the given x and real $a_1, a_2, \ldots, a_m$, that

$$\lim_{n\to\infty} \frac{1}{n^m} R_m^n\left(\nu; x + \frac{a_1}{n}, \ldots, x + \frac{a_m}{n}\right) = \left(\frac{\omega_J(x)}{C_1 - \eta}\right)^m \det\bigl(S((a_i - a_j)\omega_J(x))\bigr)_{1\leq i,j\leq m}. \tag{4-11}$$

Note that the Lebesgue condition for the local Szegő function required by Totik is satisfied because ν' is smooth (even constant) near x. Then Corollary 1.2 gives

$$\limsup_{n\to\infty} \frac{1}{n^m} R_m^n\left(\mu; x + \frac{a_1}{n}, \ldots, x + \frac{a_m}{n}\right) \leq \left(\frac{\omega_J(x)}{C_1 - \eta}\right)^m \det\bigl(S((a_i - a_j)\omega_J(x))\bigr)_{1\leq i,j\leq m}.$$

As the left-hand side is independent of η, we obtain

$$\limsup_{n\to\infty} \frac{1}{n^m} R_m^n\left(\mu; x + \frac{a_1}{n}, \ldots, x + \frac{a_m}{n}\right) \leq \left(\frac{\omega_J(x)}{C_1}\right)^m \det\bigl(S((a_i - a_j)\omega_J(x))\bigr)_{1\leq i,j\leq m}.$$

The lower bound is similar. □

5. Proof of Theorem 2.4

Let

$$w(t) = (1-t)^\alpha, \quad t \in (-1, 1).$$

Choose $\delta > 0$ such that μ is absolutely continuous in $(1-\delta, 1)$, satisfying there

$$(C_1 - \delta)w(t) \leq \mu'(t) \leq (C_2 + \delta)w(t).$$

Here C_1, C_2 are as in (2-18). Let

$$d\nu(t) = d\mu(t) + (C_2 + \delta)w(t)\, dt \quad \text{in } (-1, 1-\delta]$$

and

$$d\nu(t) = (C_2 + \delta)w(t)\, dt \quad \text{in } (1-\delta, 1].$$

Then

$$d\nu \geq d\mu \quad \text{in } [-1, 1].$$

Note too that, in $(1-\delta, 1)$, the derivative μ' of the absolutely continuous component of μ satisfies

$$\frac{\mu'(t)}{\nu'(t)} \geq \frac{C_1-\delta}{C_2+\delta}. \tag{5-1}$$

Inasmuch as $w > 0$ in $(-1, 1)$, ν is a regular measure in the sense of Stahl, Totik and Ullman, while $\nu'(t)(1-t)^{-\alpha}$ is continuous and positive at 1. By a result of the author [Lubinsky 2008, Theorem 1.2],

$$\lim_{n\to\infty} \frac{1}{2n^2} \tilde{K}_n\left(\nu, 1-\frac{a}{2n^2}, 1-\frac{b}{2n^2}\right) = \mathbb{J}_\alpha(a, b),$$

uniformly for a, b in compact subsets of $(0, \infty)$. If $\alpha \geq 0$, we may also allow a, b to lie in compact subsets of $[0, \infty)$. Then, for $m \geq 2$, Corollary 1.2 and (5-1) give, for $a_1, a_2, \dots, a_m > 0$,

$$\begin{aligned}
\liminf_{n\to\infty} &\left(\frac{1}{2n^2}\right)^m R_m^n\left(\mu; 1-\frac{a_1}{2n^2}, \dots, 1-\frac{a_m}{2n^2}\right) \prod_{j=1}^m \mu'\left(1-\frac{a_j}{2n^2}\right) \\
&\geq \left(\frac{C_1-\delta}{C_2+\delta}\right)^m \liminf_{n\to\infty}\left(\frac{1}{2n^2}\right)^m R_m^n\left(\nu; 1-\frac{a_1}{2n^2}, \dots, 1-\frac{a_m}{2n^2}\right) \prod_{j=1}^m \nu'\left(1-\frac{a_j}{2n^2}\right) \\
&= \left(\frac{C_1-\delta}{C_2+\delta}\right)^m \det(\mathbb{J}_\alpha(a_i, a_j))_{1\leq i,j\leq m}.
\end{aligned}$$

Now let $\delta \to 0+$. □

6. Proofs of Theorem 2.5 and Corollary 2.6

We begin with a lemma that uses the by now classical technique of Totik involving fast decreasing polynomials:

Lemma 6.1. *Assume the hypotheses of Theorem 2.5, except that we do not assume* (2-22)*, nor that μ is regular. Let $\varepsilon \in (0, 1)$. Then*

$$\liminf_{n\to\infty} \frac{K_n^m(\mu, \underline{y}_n, \underline{y}_n)}{K_{[n(1-\varepsilon)]}^m(\nu, \underline{y}_n, \underline{y}_n)} \geq \left(\frac{\nu'(\xi)}{\mu'(\xi)}\right)^m. \tag{6-1}$$

Proof. We may assume that the common support J of μ and ν is contained in $[-1, 1]$, as a linear transformation of the variable changes the limits in a trivial way. Let $\eta > 0$, and

$$c = \frac{\mu'(\xi)}{\nu'(\xi)}.$$

Our hypothesis (2-20) ensures that we can choose $\delta > 0$ such that

$$\frac{\mu(I)}{\nu(I)} \leq (c+\eta) \quad \text{for } I \subset [\xi-\delta, \xi+\delta]. \tag{6-2}$$

Let $n \geq 4/\varepsilon$ and $\ell = \ell(n) = \left[\frac{1}{2}\varepsilon n\right]$, so that $n - \ell \geq [n(1-\varepsilon)]$. We may choose a polynomial R_ℓ of degree less than or equal to ℓ and $\kappa \in (0,1)$ such that

$$0 \leq R_\ell \leq 1 \quad \text{in } [-2,2],$$

$$|R_\ell(t) - 1| \leq \kappa^\ell \quad \text{in } [-\delta/2, \delta/2], \tag{6-3}$$

$$|R_\ell(t)| \leq \kappa^\ell \quad \text{in } [-2,-\delta] \cup [\delta, 2]. \tag{6-4}$$

The crucial thing here is that κ is independent of ℓ, depending only on δ. These polynomials are easily constructed from the approximations to the sign function of Ivanov and Totik [1990, Theorem 3, p. 3]. For the given ξ and n, we let

$$\Psi_n(\underline{t}) = \Psi_n(t_1, t_2, \ldots, t_m) = \prod_{j=1}^{m} R_\ell(\xi - t_j).$$

Observe that this is a symmetric polynomial in $t_1, t_2, \ldots, t_m$. Moreover, for large enough n, we have from (2-21), (6-3), and (6-4),

$$\Psi_n(\underline{y}_n) \geq (1 - \kappa^\ell)^m; \tag{6-5}$$

$$|\Psi_n(\underline{t})| \leq \kappa^l \quad \text{in } [-1,1]^m \backslash \mathbb{Q}, \tag{6-6}$$

where

$$\mathbb{Q} = \left\{(t_1, t_2, \ldots, t_m) : \max_{1 \leq j \leq m} |\xi - t_j| \leq \delta\right\}.$$

Next, let $P_1 \in \mathcal{AL}^m_{n-\ell-1}$, and set $P = P_1 \Psi_n$. We see that $P \in \mathcal{AL}^m_{n-1}$. Using (6-2), (6-6), we see that

$$\int P^2 \, d\mu^{\times m} \leq (c+\eta)^m \int_{\mathbb{Q}} P_1^2 \, d\nu^{\times m} + \|P_1\|^2_{L_\infty(J^m)} \kappa^{2\ell} \int_{J^m \backslash \mathbb{Q}} d\mu^{\times m}. \tag{6-7}$$

Now we use the regularity of ν, and the fact that J is regular for the Dirichlet problem. These properties imply that [Stahl and Totik 1992, Theorem 3.2.3(v), p. 68]

$$\lim_{n \to \infty} \left(\sup_{\deg(T) \leq n} \frac{\|T\|^2_{L_\infty(J)}}{\int |T^2| \, d\nu} \right)^{1/n} = 1.$$

The supremum is taken over all univariate polynomials T of degree at most n. By successively applying this in each of the m variables, we see that

$$\|P_1\|^2_{L_\infty(J^m)} \leq (1 + o(1))^n \int P_1^2 \, d\nu^{\times m},$$

where the $o(1)$ term is crucially independent of P_1. Thus we may continue (6-7) as

$$\int P^2 \, d\mu^{\times m} \leq (c+\eta)^m \left(\int P_1^2 \, d\nu^{\times m} \right) \left(1 + (1+o(1))^n \kappa^{n\varepsilon}\right).$$

Since also

$$P^2(\underline{y}_n) \geq P_1^2(\underline{y}_n)(1 + O(\kappa^{\varepsilon n})),$$

we see from (3-5), with an appropriate choice of P_1, that

$$K_n^m(\mu, \underline{y}_n, \underline{y}_n) \geq \frac{P^2(\underline{y}_n)}{\int P^2\, d\mu^{\times m}} \geq \sup_{P_1 \in \mathcal{AL}_{n-\ell-1}^m} \frac{P_1^2(\underline{y}_n)(1+O(\kappa^{\varepsilon n}))}{(c+\eta)^m(\int P_1^2\, d\nu^{\times m})(1+(1+o(1))^n\kappa^{n\varepsilon})}$$
$$= \frac{1+o(1)}{(c+\eta)^m} K_{n-\ell}^m(\nu, \underline{y}_n, \underline{y}_n).$$

Thus

$$\liminf_{n\to\infty} \frac{K_n^m(\mu, \underline{y}_n, \underline{y}_n)}{K_{[n(1-\varepsilon)]}^m(\nu, \underline{y}_n, \underline{y}_n)} \geq (c+\eta)^{-m}.$$

As the left-hand side is independent of η, we obtain (6-1). □

Proof of Theorem 2.5. Lemma 6.1 asserts that

$$\liminf_{n\to\infty} \frac{K_n^m(\mu, \underline{y}_n, \underline{y}_n)}{K_{[n(1-\varepsilon)]}^m(\nu, \underline{y}_n, \underline{y}_n)} \geq \left(\frac{\nu'(\xi)}{\mu'(\xi)}\right)^m.$$

Swapping the roles of μ and ν, Lemma 6.1 also gives

$$\liminf_{n\to\infty} \frac{K_{[n(1+\varepsilon)]}^m(\nu, \underline{y}_n, \underline{y}_n)}{K_n^m(\mu, \underline{y}_n, \underline{y}_n)} \geq \left(\frac{\mu'(\xi)}{\nu'(\xi)}\right)^m.$$

Now we apply our hypothesis (2-22) and let $\varepsilon \to 0+$. □

Proof of Corollary 2.6. We apply Theorem 2.5 with $\xi = x$ and, for $n \geq 1$,

$$\underline{y}_n = \left(x + \frac{a_1}{n\omega_J(x)}, \dots, x + \frac{a_m}{n\omega_J(x)}\right).$$

This satisfies (2-21) with $\xi = x$. Now $\det[S(a_i - a_j)]_{1\leq i,j\leq m} > 0$, so our hypothesis (2-24) easily implies (2-22). Then (1-4) and Theorem 2.5 give the result. □

References

[Baik et al. 2003] J. Baik, T. Kriecherbauer, K. T.-R. McLaughlin, and P. D. Miller, "Uniform asymptotics for polynomials orthogonal with respect to a general class of discrete weights and universality results for associated ensembles: Announcement of results", *Int. Math. Res. Not.* **2003**:15 (2003), 821–858. MR 2004e:41038 Zbl 1036.42023

[Baik et al. 2008] J. Baik, T. Kriecherbauer, L.-C. Li, K. D. T.-R. McLaughlin, and C. Tomei (editors), *Integrable systems and random matrices* (New York, 2006), Contemporary Mathematics **458**, American Mathematical Society, Providence, RI, 2008. MR 2009i:00012 Zbl 1139.37001

[Beckenbach and Bellman 1961] E. F. Beckenbach and R. Bellman, *Inequalities*, Ergeb. Math. Grenzgeb. **30**, Springer, Berlin, 1961. MR 28 #1266 Zbl 0097.26502

[Breuer 2011] J. Breuer, "Sine kernel asymptotics for a class of singular measures", *J. Approx. Theory* **163**:10 (2011), 1478–1491. MR 2012j:28003 Zbl 1228.42028

[Deift 1999] P. A. Deift, *Orthogonal polynomials and random matrices: a Riemann–Hilbert approach*, Courant Lecture Notes in Mathematics **3**, American Mathematical Society, Providence, RI, 1999. MR 2000g:47048

[Deift and Gioev 2009] P. Deift and D. Gioev, *Random matrix theory: invariant ensembles and universality*, Courant Lecture Notes in Mathematics **18**, American Mathematical Society, Providence, RI, 2009. MR 2011f:60008 Zbl 1171.15023

[Deift et al. 1999] P. Deift, T. Kriecherbauer, K. T.-R. McLaughlin, S. Venakides, and X. Zhou, "Uniform asymptotics for polynomials orthogonal with respect to varying exponential weights and applications to universality questions in random matrix theory", *Comm. Pure Appl. Math.* **52**:11 (1999), 1335–1425. MR 2001g:42050 Zbl 0944.42013

[Erdős 2011] L. Erdős, "Universality of Wigner random matrices: a survey of recent results", *Russia Math. Surveys* **66**:3 (2011), 507–626. MR 2859190 Zbl 1230.82032

[Erdős et al. 2010] L. Erdős, J. Ramírez, B. Schlein, T. Tao, V. Vu, and H.-T. Yau, "Bulk universality for Wigner Hermitian matrices with subexponential decay", *Math. Res. Lett.* **17**:4 (2010), 667–674. MR 2011j:60018 Zbl 05937399

[Erdős et al. 2011] L. Erdős, B. Schlein, and H.-T. Yau, "Universality of random matrices and local relaxation flow", *Invent. Math.* **185**:1 (2011), 75–119. MR 2012f:60020 Zbl 1225.15033

[Findley 2008] E. Findley, "Universality for locally Szegő measures", *J. Approx. Theory* **155**:2 (2008), 136–154. MR 2011c:42068 Zbl 1225.15033

[Forrester 2010] P. J. Forrester, *Log-gases and random matrices*, London Math. Soc. Monogr. Ser. **34**, Princeton University Press, 2010. MR 2011d:82001 Zbl 1217.82003

[Foulquié Moreno et al. 2011] A. Foulquié Moreno, A. Martínez-Finkelshtein, and V. L. Sousa, "Asymptotics of orthogonal polynomials for a weight with a jump on $[-1, 1]$", *Constr. Approx.* **33**:2 (2011), 219–263. MR 2012b:42044 Zbl 1213.42090

[Ivanov and Totik 1990] K. G. Ivanov and V. Totik, "Fast decreasing polynomials", *Constr. Approx.* **6**:1 (1990), 1–20. MR 90k:26023 Zbl 0682.41014

[Kuijlaars and Vanlessen 2002] A. B. J. Kuijlaars and M. Vanlessen, "Universality for eigenvalue correlations from the modified Jacobi unitary ensemble", *Int. Math. Res. Not.* **2002**:30 (2002), 1575–1600. MR 2003g:30043 Zbl 1122.30303

[Levin and Lubinsky 2008] E. Levin and D. S. Lubinsky, "Universality limits in the bulk for varying measures", *Adv. Math.* **219**:3 (2008), 743–779. MR 2010a:60009 Zbl 1176.28014

[Lubinsky 2008] D. S. Lubinsky, "A new approach to universality limits at the edge of the spectrum", pp. 281–290 in *Integrable systems and random matrices* (New York, 2006), edited by J. Baik et al., Contemp. Math. **458**, American Mathematical Society, Providence, RI, 2008. MR 2010a:42097 Zbl 1147.15306

[Lubinsky 2009a] D. S. Lubinsky, "A new approach to universality limits involving orthogonal polynomials", *Ann. of Math.* (2) **170**:2 (2009), 915–939. MR 2011a:42042 Zbl 1176.42022

[Lubinsky 2009b] D. S. Lubinsky, "Universality limits for random matrices and de Branges spaces of entire functions", *J. Funct. Anal.* **256**:11 (2009), 3688–3729. MR 2012b:46057 Zbl 1184.46029

[Lubinsky 2012] D. S. Lubinsky, "Bulk universality holds in measure for compactly supported measures", *J. Anal. Math.* **116** (2012), 219–253. MR 2892620

[Máté et al. 1991] A. Máté, P. Nevai, and V. Totik, "Szegő's extremum problem on the unit circle", *Ann. of Math.* (2) **134**:2 (1991), 433–453. MR 92i:42014 Zbl 0752.42015

[Mehta 1991] M. L. Mehta, *Random matrices*, 2nd ed., Academic Press, Boston, MA, 1991. MR 92f:82002 Zbl 0780.60014

[Nevai 1986] P. Nevai, "Géza Freud, orthogonal polynomials and Christoffel functions: A case study", *J. Approx. Theory* **48**:1 (1986), 167. MR 88b:42032 Zbl 0606.42020

[Ransford 1995] T. Ransford, *Potential theory in the complex plane*, London Mathematical Society Student Texts **28**, Cambridge University Press, 1995. MR 96e:31001 Zbl 0828.31001

[Saff and Totik 1997] E. B. Saff and V. Totik, *Logarithmic potentials with external fields*, Grundlehren Math. Wiss. **316**, Springer, Berlin, 1997. MR 99h:31001 Zbl 0881.31001

[Simon 2008a] B. Simon, "The Christoffel–Darboux kernel", pp. 295–335 in *Perspectives in partial differential equations, harmonic analysis and applications*, edited by D. Mitrea and M. Mitrea, Proc. Sympos. Pure Math. **79**, American Mathematical Society, Providence, RI, 2008. MR 2010d:42045 Zbl 1159.42020

[Simon 2008b] B. Simon, "Two extensions of Lubinsky's universality theorem", *J. Anal. Math.* **105** (2008), 345–362. MR 2010c:42054 Zbl 1168.42304

[Simon 2011] B. Simon, *Szegő's theorem and its descendants: Spectral theory for L^2 perturbations of orthogonal polynomials*, Princeton University Press, 2011. MR 2012b:47080 Zbl 1230.33001

[Stahl and Totik 1992] H. Stahl and V. Totik, *General orthogonal polynomials*, Encyclopedia of Mathematics and its Applications **43**, Cambridge University Press, 1992. MR 93d:42029 Zbl 0791.33009

[Tao and Vu 2011] T. Tao and V. Vu, "The Wigner–Dyson–Mehta bulk universality conjecture for Wigner matrices", preprint, 2011. arXiv 1101.5707

[Totik 2000] V. Totik, "Asymptotics for Christoffel functions for general measures on the real line", *J. Anal. Math.* **81** (2000), 283–303. MR 2001j:42021 Zbl 0966.42017

[Totik 2009] V. Totik, "Universality and fine zero spacing on general sets", *Ark. Mat.* **47**:2 (2009), 361–391. MR 2010f:42055 Zbl 1180.42017

Received 12 Aug 2011. Revised 16 Aug 2011. Accepted 13 Feb 2012.

DORON S. LUBINSKY: lubinsky@math.gatech.edu
School of Mathematics, Georgia Institute of Technology, 686 Cherry Street, Atlanta, GA 30332, United States

ANALYSIS AND PDE
Vol. 6, No. 1, 2013
dx.doi.org/10.2140/apde.2013.6.131

RELATIVE KÄHLER–RICCI FLOWS AND THEIR QUANTIZATION

ROBERT J. BERMAN

Let $\pi : \mathscr{X} \to S$ be a holomorphic fibration and let $\mathscr{L}$ be a relatively ample line bundle over $\mathscr{X}$. We define relative Kähler–Ricci flows on the space of all Hermitian metrics on $\mathscr{L}$ with relatively positive curvature and study their convergence properties. Mainly three different settings are investigated: the case when the fibers are Calabi–Yau manifolds and the case when $\mathscr{L} = \pm K_{\mathscr{X}/S}$ is the relative (anti)canonical line bundle. The main theme studied is whether "positivity in families" is preserved under the flows and its relation to the variation of the moduli of the complex structures of the fibers. The "quantization" of this setting is also studied, where the role of the Kähler–Ricci flow is played by Donaldson's iteration on the space of all Hermitian metrics on the finite rank vector bundle $\pi_*\mathscr{L} \to S$. Applications to the construction of canonical metrics on the relative canonical bundles of canonically polarized families and Weil–Petersson geometry are given. Some of the main results are a parabolic analogue of a recent elliptic equation of Schumacher and the convergence towards the Kähler–Ricci flow of Donaldson's iteration in a certain double scaling limit.

1. Introduction

1A. *Background.* On an n-dimensional Kähler manifold (X, ω_0) Hamilton's Ricci flow [Hamilton 1982] on the space of Riemannian metrics on X preserves the Kähler condition of the initial metric and may be written as the *Kähler–Ricci flow*

$$\frac{\partial \omega_t}{\partial t} = -\mathrm{Ric}\, \omega_t. \tag{1-1}$$

When X is a Calabi–Yau manifold (which here will mean that the canonical line bundle K_X is holomorphically trivial) it was shown by Cao [1985] that the corresponding flow in the space of Kähler metrics in $[\omega_0] \in H^2(X, \mathbb{R})$ has a large time limit. The limit is thus a fixed point of the flow which coincides with the unique Ricci flat Kähler metric in $[\omega_0]$, whose existence was first established by Yau [1978] in his celebrated proof of the Calabi conjecture. The non-Calabi–Yau cases when $[\omega_0]$ is the first Chern class $c_1(L)$ of $L = rK_X$, where $r = \pm 1$, have also been studied extensively (where $-r\omega$ is added to the right side in (1-1)). In general the fixed points of the corresponding Kähler–Ricci flows are hence

MSC2010: 14J32, 32G05, 32Q20, 53C55.
Keywords: Kähler–Ricci flow, positivity, Kähler–Einstein metric, balanced metric, Weil–Petersson metric.

Kähler–Einstein metrics of negative ($r = 1$) and positive ($r = -1$) scalar curvature. The convergence towards a fixed point — when it exists — in the latter positive case (i.e., X is a Fano manifold) was only established very recently by Perelman (unpublished) and by Tian and Zhu [2007].

A distinctive feature of Kähler geometry is that a Kähler metric ω may be locally described in terms of a local function ϕ, such that $\omega = dd^c\phi$. In the integral case, that is, when $[\omega_0] = c_1(L)$ is the first Chern class of an ample line bundle $L \to X$, this just amounts to the global fact that the space of Kähler metrics ω in $c_1(L)$ may be identified with the space $\mathcal{H}_L$ of smooth metrics h on the line bundle L with positive curvature form ω, modulo the action of $\mathbb{R}$ on $\mathcal{H}_L$ by scalings. Locally, $h = e^{-\phi}$ and we will refer to the additive object ϕ as a *weight* on L (see Section 2A). In this notation the Kähler–Einstein equations may be expressed as Monge–Ampère equations on $\mathcal{H}_L$. For example, on a Calabi–Yau manifold $\omega_\phi := dd^c\phi$ is Ricci flat precisely when

$$(dd^c\phi)^n/n! = \mu, \tag{1-2}$$

where μ is the canonical probability measure on X such that $\mu = i^{n^2}\Omega \wedge \bar{\Omega}$, for Ω a suitable global holomorphic n-form trivializing K_X (to simplify the notion we will in the following always assume that the volume of the given class $[\omega_0]$ is equal to one, so that $\omega_0^n/n!$ defines a probability measure on X for any $\omega \in [\omega_0]$). By letting μ depend on ϕ in a suitable way general Kähler–Einstein metrics are obtained.

As emphasized by Yau [1987] one can expect to obtain approximations to Kähler–Einstein metrics by using holomorphic sections of high powers of a line bundle. In this direction Donaldson [2009] introduced certain iterations on the "quantization" (at level k) of the space $\mathcal{H}_L$ of Kähler metrics in $c_1(L)$. Geometrically, this quantized space, denoted by $\mathcal{H}^{(k)}$, is the space of all Hermitian metrics on the finite-dimensional vector space $H^0(X, kL)$ of global holomorphic sections of kL, where kL denotes the k-th tensor power of L, in our additive notation (for the definition see Section 2D). In other words $\mathcal{H}^{(k)}$ can be identified with the symmetric space $\mathrm{GL}(N_k, \mathbb{C})/U(N_k)$ of $N_k \times N_k$ Hermitian matrices which in turn, using projective embeddings, corresponds to the space of level k Bergman metrics on L. The fixed points of Donaldson's iteration are called *balanced* metrics at level k (with respect to μ) and they first appeared in the previous work of Bourguignon, Li, and Yau [Bourguignon et al. 1994]. Again, in the $\pm K_X$-setting one lets μ depend on ϕ in a suitable way leading to different settings (see below). In the limit when L is replaced by a large tensor power it has very recently been shown that balanced metrics in the different settings indeed converge to Kähler–Einstein metrics [Wang 2005; Keller 2009; Berman et al. 2009]. It was pointed out in [Donaldson 2009] that it seems likely that these iterations can be viewed as discrete approximations of the Ricci flow. This will be made precise and confirmed in the present paper (Theorem 3.15 and Theorem 4.18).

1B. ***Outline of the present setting and the main results.*** The aim of the present paper is to study *relative* versions of the Kähler–Ricci flow and Donaldson's iteration (in the various settings). More precisely, the geometric setting is that of a holomorphic fibration $\pi : \mathcal{X} \to S$ of relative dimension n and a relatively ample line bundle $\mathcal{L} \to \mathcal{X}$. The fibration will mainly be assumed to be a proper submersion over a connected base, so that all fibers are diffeomorphic (for general quasiprojective morphisms see Section 4E). Note that in applications S typically arises as a moduli space or Teichmüller space and $\mathscr{X}$ as the corresponding universal family.

The main points that will be considered are

- the question whether "positivity in families" is preserved under the flows;
- the convergence of the "quantized" (finite-dimensional) setting of Donaldson's iteration towards the Kähler–Ricci flow setting in the "semiclassical" limit (i.e., the large k-limit).

More precisely, denote by $\mathscr{H}_{\mathscr{L}/S}$ the space of all metrics on $\mathscr{L}$ which are fiberwise of positive curvature. In other words, $\mathscr{H}_{\mathscr{L}/S}$ is an infinite-dimensional fiber bundle over S whose fibers are of the form $\mathscr{H}_L$, as in the previous section. The *relative Kähler–Ricci flows* are defined as suitable flows on $\mathscr{H}_{\mathscr{L}/S}$ such that the induced flow of curvature forms restricts to the usual Kähler–Ricci flow fiberwise: we will say that "positivity is preserved under the flow" if, for any initial metric with positive curvature (in *all* directions on $\mathscr{X}$), the evolved metric also has positive curvature for all times; that is, the flow induces a flow of Kähler forms on the total space $\mathscr{X}$ of the fibration (and not only along the fibers).

As will be explained below, the two points above are closely related. For example, the preservation of positivity in the relative Kähler–Ricci flow setting can be seen as a limiting version of the well-known positivity of direct image bundles in the quantized setting (the latter positivity is a fundamental tool in complex geometry; see [Kawamata 1982; Berndtsson 2009a], for example). As another application of the convergence in the second point above (in the absolute case when S is a point) we will deduce the uniform convergence of Donaldson's canonically balanced metrics from the well-known convergence of the Kähler–Ricci flow (Theorem 4.20).

The Calabi–Yau setting. Let us first summarize the main results in the setting when the fibers are Calabi–Yau. It should however be stressed that the setting when the fibers are canonically polarized appears to be the one most suited for geometric applications (see below). In the Calabi–Yau setting flow ϕ_t in $\mathscr{H}_{\mathscr{L}/S}$ is defined fiberwise by

$$\frac{\partial \phi_t}{\partial t} = \log \frac{(dd^c\phi_t)^n/n!}{\mu}, \tag{1-3}$$

with μ a measure as in (1-2). Of course, adding the pull-back of a time-dependent function on the base S to the right side of the previous equation does not alter the induced flows of the *fiberwise restricted* Kähler forms $d_X d_X^c \phi_t$, but it certainly effects the flow of $dd^c\phi_t$ on $\mathscr{X}$ which will typically *not* preserve the initial Kähler property.

One of the main results of the present paper is a parabolic evolution equation along the flow (1-3) for the function

$$c(\phi) := \frac{1}{n}(dd^c\phi)^{n+1}/(d_X d_X^c\phi)^n \wedge i\,ds \wedge d\bar{s}$$

on $\mathscr{X}$ which is well-defined when S is embedded in $\mathbb{C}$. The point is that $c(\phi) > 0$ precisely when $dd^c\phi > 0$ on $\mathscr{X}$. The evolution equation for $c(\phi_t)$ reads (Theorem 3.3)

$$\left(\frac{\partial}{\partial t} - \Delta_{\omega_t^X}\right) c(\phi_t) = |A_{\omega_t}|^2_{\omega_t^X} - \omega_{\mathrm{WP}}, \tag{1-4}$$

where ω_t^X denotes the flow of the fiberwise restricted curvature forms, A_{ω_t} is a certain representative of the Kodaira–Spencer class of the fiber $\mathscr{X}_s$ and ω_{WP} is the pull-back to $\mathscr{X}$ of the (generalized) Weil–Petersson form on the base S; by a result of Tian [1987] and Todorov [1989], which we will reprove, ω_{WP} can be represented by the global squared L^2-norm of $A_{\omega_{\mathrm{KE}}}$ for ω_{KE} the unique Ricci flat metric in $c_1(L)$. Applying the maximum principle then gives (Corollary 3.4) that the initial condition $dd^c\phi_0 > 0$ implies that

$$dd^c\phi_t > -t\omega_{\mathrm{WP}} \tag{1-5}$$

(and similarly when the initial curvature is *semi*positive). By its very definition ω_{WP} vanishes at s precisely when the infinitesimal deformation of the complex structure on the fibers $\mathscr{X}_s$ (i.e., the Kodaira–Spencer class) vanishes at s. In particular, if the fibration $\pi : \mathscr{X} \to S$ is holomorphically trivial, then, by inequality (1-5), positivity is indeed preserved along the flow. This latter situation appears naturally in Kähler geometry. Indeed, if the base S is an annulus in $\mathbb{C}$ and ϕ_s is rotationally invariant, then ϕ_s corresponds to a curve in $\mathscr{H}_L$ and $c(\phi_s)$ is then the geodesic curvature of the curve ϕ_s when $\mathscr{H}_L$ is equipped with its symmetric space Riemannian metric (see [Chen 2000] and references therein). In the nonnormalized K_X-setting (see Section 4) the equation (1-4) can be seen as a parabolic generalization of a very recent elliptic equation of Schumacher [2008].

Similarly, the "quantized" version of the previous setting is studied, that is, the relative version of Donaldson's iteration. It gives an iteration on the space of all Hermitian metrics H on the finite rank vector bundle $\pi_* k\mathscr{L} \to S$ for any positive integer k (recall that the fiber of $\pi_*\mathscr{L}$ over s is, by definition, the space $H^0(\mathscr{X}_s, \mathscr{L}_s)$ of all global holomorphic sections on the fiber $\mathscr{X}_s$ with values in $\mathscr{L}_{|\mathscr{X}_s}$). More precisely, we will study the equivalent fiberwise iteration $\phi_m^{(k)}$ in $\mathscr{H}_{\mathscr{L}/S}$ obtained by applying the (scaled) Fubini–Study map to Donaldson's iteration. It will be called the *relative Bergman iteration at level k*. When the discrete time m tends to infinity it is shown (Theorem 3.9) that the iteration converges to a fiberwise balanced weight:

$$\phi_m^{(k)} \to \phi_\infty^{(k)}$$

in the $\mathscr{C}^\infty$-topology on $\mathscr{X}_s$, uniformly with respect to s. It is also observed that an analogue of the inequality (1-5) holds; that is,

$$dd^c\phi_m^{(k)} \geq -\frac{k}{m}\omega_{\mathrm{WP}}. \tag{1-6}$$

This turns out to be a simple consequence of a recent theorem of Berndtsson [2009a] about the curvature of vector bundles of the form $\pi_*(\mathscr{L} + K_{\mathscr{X}/S})$. We also confirm Donaldson's expectation about the semiclassical limit when the level k tends to infinity. More precisely, it is shown that, in the double scaling limit where $m/k \to t$, the (relative) Bergman iteration at level k approaches the (relative) Kähler–Ricci flow (1-3):

$$\phi_m^{(k)} \to \phi_t \tag{1-7}$$

uniformly on $\mathscr{X}$. In particular, combining this convergence with (1-6) gives an alternative proof of the semipositivity in the inequality (1-5). Moreover, by taking $m = m_k$ such that $m/k \to \infty$ this gives a dynamical construction of solutions to the inhomogeneous Monge–Ampère equation (1-2) in the setting where μ is any fixed volume form (Corollary 3.16).

The (anti)canonical setting. The previous results also have analogues in the setting when the ample line bundle $\mathscr{L}$ is either the relative canonical line bundle $K_{\mathscr{X}/S}$ over $\mathscr{X}$ or its dual, which we write as $\mathscr{L} = \pm K_{\mathscr{X}/S}$ in our additive notation. The starting point is the fact that any metric $h = e^{-\phi}$ on $\pm K_X$ induces, by the very definition of K_X, a volume form on X which may be written suggestively as $e^{\pm\phi}$. The previous constructions, that is, the relative Kähler–Ricci flows and the Donaldson iteration, can then be repeated word for word for these ϕ-dependent measures $\mu = \mu(\phi)$. For example, the relative Kähler–Ricci flows are defined by

$$\frac{\partial \phi_t}{\partial t} = \log\left(\frac{(dd^c\phi_t)^n/n!}{e^{\pm\phi_t}}\right), \tag{1-8}$$

and we obtain (Theorem 4.7) a corresponding parabolic equation for $c(\phi_t)$:

$$\left(\frac{\partial}{\partial t} - (\Delta_{\omega_t^X} - \pm 1)\right)c(\phi_t) = |A_{\omega_t}|^2_{\omega_t^X},$$

and as a consequence *the flows always preserve positivity* (Corollary 4.9) in these settings. In fact, in the case of infinitesimally nontrivial fibration the flows will even *improve* the positivity; that is, any initial weight which is merely semipositively curved instantly becomes positively curved under the flows. In the $+K_X$-setting the unique fixed point of the flow (1-8) is the (fiberwise) *normalized* Kähler–Einstein weight uniquely determined by

$$e^{-\phi_{\mathrm{KE}}} = (\omega_{\mathrm{KE}})^n/n!,$$

where ω_{KE} is the unique Kähler–Einstein metric on X (Corollary 4.3). The corresponding elliptic equation for $c(\phi_{\mathrm{KE}})$ was first obtained by Schumacher [2008] who used it to deduce the following interesting result: ϕ_{KE} is always semipositively curved on the total space of $\mathscr{X}$ and strictly positively curved for an infinitesimally nontrivial fibration. As a consequence he obtained several applications to the geometry of moduli spaces. For example, applied to the case when $\mathscr{X} \to S$ is the universal curve over the Teichmüller space of Riemann surfaces of genus $g \geq 2$ it gives, when combined with Berndtsson's theorem (Theorem 3.10), a new proof of the hyperbolicity result of Liu, Sun, and Yau [Liu et al. 2008] saying that the curvature of the Weil–Petersson metric on the Teichmüller space is dual Nakano positive.

In the $-K_X$-setting the relative Kähler–Ricci flow will diverge for generic initial data. But using the convergence on the level of Kähler forms, established by Perelman and Tian and Zhu, will show that, if the Fano manifold X admits a unique positively curved Kähler–Einstein metric ω_{KE}, the flow does converge to a weight for ω_{KE} in the *normalized* $\pm K_X$*-setting*. This latter setting is simply obtained by normalizing the volume forms $e^{\pm\phi}$ used above.

We will also use the relative Bergman iteration to obtained a "quantized" version of Schumacher's result: the canonical "semibalanced" metric at level k on $K_{\mathscr{X}/S}$, which by definition is fiberwise normalized and balanced, is smooth with semipositive curvature on $\mathscr{X}$ (Corollary 4.16) and strictly positively curved in the case of an infinitesimally nontrivial fibration. As a consequence the semibalanced metric gives an alternative to the canonical metric on $kK_{\mathscr{X}/S}$ introduced in [Narasimhan and Simha 1968] (see also Kawamata 1982; Tsuji 2011; Berndtsson and Păun 2008a for positivity properties of this latter metric).

In Section 4E some of the results concerning the setting when K_X is ample are generalized to projective fibrations of varieties of general type (i.e., K_X is merely big) and the corresponding canonical semibalanced metric is shown to have a positive curvature current (Theorem 4.21). Relations to deformation invariance of plurigenera [Siu 1998] are also briefly discussed.

1C. ***Further relations to previous results.*** A variant of Donaldson iteration (but with a single parameter k) in the K_X-setting was introduced by Tsuji [2006]. He proved convergence in the L^1-topology towards the normalized Kähler–Einstein weight ϕ_{KE} in the large k-limit (see [Song and Weinkove 2010] for a proof of uniform convergence) and deduced the *semi*positivity result for ϕ_{KE} of Schumacher referred to above. These works of Tsuji and Schumacher provided an important motivation for the present one. Steve Zelditch has also informed the author of a joint work in progress with Jian Song, where they show that the linearization of Tsuji's iteration at the fixed point coincides with the linearization of the Kähler–Ricci flow. It should also be pointed out that another discretization of the Kähler–Ricci flow on a Fano manifold was studied by Rubinstein [2008] and Keller [2009].

The C^0-convergence of the Bergman iteration at a fixed level k in the Calabi–Yau setting (or more generally in the setting of a fixed measure ϕ) was pointed out by Donaldson [2009] and the proof was sketched. Sano [2006] provided an explicit proof in the constant scalar curvature setting (see Section 4F).

It is also interesting to compare with the very recent work of Fine [2010] concerning the constant scalar curvature setting. He shows that a continuous version of Donaldson's iteration in this latter setting, called balancing flows, converges to the Calabi flow, when the latter flow exists. Julien Keller and Huai-Dong Cao have informed the author of a joint work in progress where an analogue of Fine's balancing flows in the Calabi–Yau setting (or more generally in the setting of a fixed volume form μ) is shown to converge to a flow on metrics, which however is different than the Kähler–Ricci flow.

There are also, at last tangential, relations to the work of Gross and Wilson [2000], where fibrations with Calabi–Yau fibers are considered. In particular, they construct certain semiflat Kähler metrics ω on the fibration $\mathscr{X}$; that is, ω is fiberwise Ricci flat. Such metrics first appeared in the string theory literature [Greene et al. 1990]. In this terminology the inequality (1-5) shows that the relative Kähler–Ricci flow deforms any given Kähler metric to a semiflat one, when there is no variation of the moduli of the complex structure of the fibers. More generally, this latter statement holds in a double scaling limit when the variation of the complex structure is very small in the sense that $\omega_{\mathrm{FS}}(s_t)t \to 0$ as $t \to \infty$.

A Kähler–Ricci flow on *compact* fibrations $\mathscr{X}$ with Calabi–Yau fibers was also considered recently by Song and Tian [2012]. But they consider the usual (i.e., nonrelative) Kähler–Ricci flow (with $r = 1$) when the canonical line bundle is only semiample and relatively trivial (i.e., the base S is the canonical model of $\mathscr{X}$). They prove that the flow collapses the fibers so that the limit is the pull-back of metric on the base S solving a "twisted" Kähler–Einstein equation where the twist is described by the (generalized) Weil–Petersson form ω_{FS}.

1D. ***Organization of the paper.*** In Section 2 a general setting is introduced and the associated relative Kähler–Ricci flow and its quantization are defined. General convergence criteria for the flows are given. In the following two sections the general setting is applied to get convergence results in particular settings of geometric relevance: the Calabi–Yau setting (Section 3) and the (anti)canonical setting (Section 4). The new feature of these convergence results for the Kähler–Ricci flows is that the convergence takes place on the level of weights, that is, for the *potentials* of the evolving Kähler metrics. Furthermore, the main question whether "positivity in families" is preserved under the flows is studied in these two sections and relations to Weil–Petersson geometry are also discussed. It is also shown that the quantized flows converge to Kähler–Ricci flows in the large tensor power limit. Applications to canonical metrics on relative canonical bundles are also given.

2. The general setting

In this section we will consider a general setup that will subsequently be applied to particular settings in Sections 3 and 4.

We assume we are given a holomorphic submersion $\pi : \mathcal{X} \to S$ of relative dimension n over a connected base and a relatively ample line bundle $\mathcal{L} \to \mathcal{X}$. In the absolute case when S is a point we will often use the notation $L \to X$ for the corresponding ample line bundle. In this latter case we will write $\mathcal{H}_L$ for the space of all smooth Hermitian metrics on L with positive curvature form. In the relative case we will denote by $\mathcal{H}_{\mathcal{L}/S}$ the space of all metrics on $\mathcal{L}$ which are fiberwise of positive curvature. We will denote by $c_1(L)$ the first Chern class of L, normalized so that it lies in $H^{1,1}(X) \cap H^2(X, \mathbb{Z})$. To simplify the formulas to be discussed we will also assume that the relative volume of $\mathcal{L}$ is equal to one; that is,

$$V := \int_X c_1(L)^n/n! = 1$$

for some (and hence any) fiber X. The general formulas may then be obtained by trivial scalings by V at appropriate places. When considering tensor powers of L, written as kL in additive notation, we will always assume that kL is very ample (which is true for k sufficiently large).

2A. ***The weight notation for*** $\mathcal{H}_L$**.** It will be convenient to use the "weight" representation of a metric h on L: locally, any metric h on L may be represented as $h = e^{-\phi}$, where h is the pointwise norm of a local trivializing section s of L. We will call the additive object ϕ a "weight" on L. One basic feature of this formalism is that even though the functions representing ϕ are merely locally defined the normalized curvature form of the metric h may be expressed as

$$\omega_\phi := dd^c\phi := \frac{i}{2\pi}\partial\bar{\partial}\phi$$

which is hence globally well-defined (but it does not imply that ω is exact!). The normalizations are made so that $[\omega_\phi] = c_1(L) \in H^{1,1}(X) \cap H^2(X, \mathbb{Z})$. In the absolute setting we will denote by $\mathcal{H}_L$ the space of all weights such that $\omega_\phi > 0$. In other words, the map $\phi \mapsto \omega_\phi$ establishes an isomorphism between $\mathcal{H}_L/\mathbb{R}$ and the space of Kähler metrics in $c_1(L)$. In the relative setting we will denote by $\mathcal{H}_{\mathcal{L}/S}$ the space of all smooth weights on $\mathcal{L}$ such that the restriction to each fiber is of positive curvature.

After fixing a reference weight ϕ_0 in $\mathcal{H}_L$ the map $\phi \mapsto u := \phi - \phi_0$ identifies the affine space of all smooth weights on L with the vector space $\mathcal{C}^\infty(X)$. Moreover, the subspace $\mathcal{H}_L$ of all positively curved smooth weights gets identified with the open convex subspace $\mathcal{H}_\omega := \{u : \ dd^c u + \omega_0 > 0\}$

of $\mathscr{C}^\infty(X)$, where ω_0 denotes the Kähler form $dd^c\phi_0$. The L^1-closure of $\mathscr{H}_\omega$ is usually called the space of all ω_0*-plurisubharmonic functions* in the literature [Guedj and Zeriahi 2005]. In fact, all the results in the present paper whose formulation does not use that the given class $[\omega_0]$ is integral are valid in the more general setting when $\mathscr{H}_L$ is replaced by $\mathscr{H}_\omega$ (with essentially the same proofs). However, since the quantized setting (Section 2D) only makes sense for integral classes we will stick to the weight notation in the following.

2B. ***The measure μ_ϕ and associated functionals on $\mathscr{H}_L$.*** First consider the absolute case when S is a point. In each particular setting studied in Sections 3 and 4 we will assume given a function μ on $\mathscr{H}_L$, $\phi \mapsto \mu(\phi)$ (also denoted by μ_ϕ), taking values in the space of volume forms on X, which is *exact* in the following sense. First observe that we may identify $\mu(\phi)$ with a one-form on the affine space $\mathscr{H}_L$ by letting its action on a tangent vector $v \in \mathscr{C}^\infty(X)$ at the point $\phi \in \mathscr{H}_L$ be defined by

$$\langle \mu(\phi), v\rangle := \int_X v\,\mu(\phi).$$

The assumption on $\mu(\phi)$ is then simply that this one-form is closed and hence exact; that is, there is a functional I_μ on $\mathscr{H}_L$ such that $dI_\mu = \mu$:

$$\frac{dI_\mu(\phi_t)}{dt} = \int_X \frac{\partial\phi_t}{\partial t}\,\mu_{\phi_t} \tag{2-1}$$

for any path ϕ_t in $\mathscr{H}_L$. The functional is determined up to a constant which will be fixed in each particular setting to be studied. We will also assume that for any fixed $v \in \mathscr{C}^\infty(X)$ the functional $\phi \mapsto \langle\mu(\phi), v\rangle$ is continuous with respect to the L^∞-topology on $\mathscr{H}_L$.

Two particular examples of such exact one-forms and their antiderivatives that will be used repeatedly are as follows:

- The Monge–Ampère measure $\phi \mapsto (dd^c\phi)^n/n! := \mathrm{MA}(\phi)$. Its antiderivative [Mabuchi 1986] will be denoted by $\mathscr{E}(\phi)$, normalized so that $\mathscr{E}(\phi_0) = 0$ for a fixed reference weight ϕ_0 in $\mathscr{H}_L$. Integrating along line segments in $\mathscr{H}_L$ gives an explicit expression for $\mathscr{E}$, but it will not be used here.
- $\phi \mapsto \mu_0$ for a volume form μ_0 on X, fixed once and for all with $I_{\mu_0}(\phi) := \int_X (\phi - \phi_0)\,\mu_0$. Since we have already fixed a reference weight ϕ_0 it will be convenient to take $\mu_0 := (dd^c\phi_0)^n/n!$.

Given $\mu = \mu(\phi)$ we define the associated functional

$$\mathscr{F}_\mu := \mathscr{E} - I_\mu.$$

By construction its critical points in $\mathscr{H}_L$ are precisely the solutions to the Monge–Ampère equation

$$(dd^c\phi)^n/n! = \mu(\phi). \tag{2-2}$$

We will say that $\mu(\phi)$ is *normalized* if it is a probability measure for all ϕ. Equivalently, this means that I_μ is *equivariant* under scalings; that is, $I_\mu(\phi + c) = I_\mu(\phi) + c$ which in turn is equivalent to $\mathscr{F}_\mu$ being *invariant* under scalings.

In the relative setting we assume that $\mu_s(\phi)$ is a smooth family of measures on the fibers $\mathscr{X}_s$ as above, parametrized by $s \in S$.

Properness and coercivity. We first recall the definition of the well-known J-functional, defined with respect to a fixed reference weight ϕ_0 (see [Berman et al. 2009] for a general setting and references). It is the natural higher-dimensional generalization of the (squared) Dirichlet norm on a Riemann surface and it will play the role of an exhaustion function of $\mathscr{H}_L/\mathbb{R}$ (but without specifying any topology!). In our notation J is simply given by the scale-invariant function

$$J = -\mathscr{F}_{\mu_0}.$$

We will then say that a functional $\mathscr{G}$ is *proper* if

$$J \to \infty \implies \mathscr{G} \to \infty$$

and *coercive* if there exists $\delta > 0$ and C_δ such that

$$\mathscr{G} \geq \delta J - C_\delta.$$

Note that δ may be taken arbitrarily small at the expense of increasing C_δ. In many geometric applications properness (and coercivity) of suitable functionals can be thought as analytic versions of algebro-geometric stability (compare Remark 4.2).

2C. ***The relative Kähler–Ricci flow with respect to μ_ϕ.*** Given an initial weight $\phi_0 \in \mathscr{H}_{\mathscr{L}/S}$ the relative Kähler–Ricci flow in $\mathscr{H}_{\mathscr{L}/S}$ is defined by the fiberwise parabolic Monge–Ampère equation

$$\frac{\partial \phi_t}{\partial t} = \log \frac{(dd^c\phi_t)^n/n!}{\mu(\phi_t)} \tag{2-3}$$

for ϕ_t smooth over $\mathscr{X} \times [0, T]$, where $T \geq 0$. We will make the following assumptions on the flow which will all be satisfied in the particular settings studied in Sections 3 and 4.

Analytical assumptions on the flow.

- Existence: The flow exists and is smooth over $\mathscr{X} \times [0, \infty[$.
- Uniqueness: Any fixed point in $\mathscr{H}_L$ of the flow is unique mod $\mathbb{R}$.
- Stability: For any $l > 0$ and $M > 0$ there is a constant $B_{l,M}$ only depending on the upper bound on the $\mathscr{C}^l$-norm of the initial weight ϕ_0 (with respect to a fixed reference weight) and a lower bound on the absolute value of $dd^c\phi_0$ such that

$$\|\phi_t - \phi_0\|_{\mathscr{C}^l(X\times[0,M])} \leq B_{l,M} \tag{2-4}$$

(locally uniformly with respect to s in the relative setting).

It follows immediately that ϕ is fixed under the flow if and only if it solves the Monge–Ampère equation (2-2). Note that since we have assumed that $\mathrm{Vol}(L) = 1$, a necessary condition to be stationary is that

$\int_X \mu_\phi = 1$. For any solution ϕ_t and fixed fiber $X = \mathscr{X}_s$ the Kähler metrics ω_t on X obtained as the restricted curvature forms of ϕ_t hence evolve according to

$$\frac{\partial \omega_t}{\partial t} = -\mathrm{Ric}\,\omega_t - \eta_\mu, \tag{2-5}$$

where $\mathrm{Ric}\,\omega_t$ in the Ricci curvature of the Kähler metric ω_t and $\eta_\mu = dd^c \log \mu(\phi)$.

Thanks to the following simple lemma the Kähler–Ricci flow is "gradient-like" for the functional $\mathscr{F}_\mu$. For the Fano case, see [Chen and Tian 2002].

Lemma 2.1. *The functional $\mathscr{F}_\mu$ is increasing along the Kähler–Ricci flow on $\mathscr{H}_L$ (defined with respect to μ_ϕ). Moreover, it is* strictly *increasing at ϕ_t unless ϕ_t is stationary.*

Proof. Differentiating along the flow gives

$$\frac{d\mathscr{F}(\phi_t)}{dt} = \int_X \log \frac{\mathrm{MA}(\phi_t)}{\mu(\phi_t)} \big(\mathrm{MA}(\phi_t) - \mu(\phi_t)\big) = \int_X \log \frac{\mathrm{MA}(\phi_t)}{\mu(\phi_t)} \left(\frac{\mathrm{MA}(\phi_t)}{\mu(\phi_t)} - 1\right) \mu(\phi_t) \geq 0$$

where the last inequality follows since both factors in the last integrand clearly have the same sign. □

If, moreover, $\mu(\phi)$ is normalized then both terms appearing in the definition of $\mathscr{F}_\mu$ are monotone:

Lemma 2.2. *Assume that $\mu(\phi)$ is normalized. Then the functionals $-I_\mu$ and $\mathscr{E}$ are both increasing along the Kähler–Ricci flow on $\mathscr{H}_L$ with respect to $\mu(\phi)$. Moreover, they are* strictly *increasing at ϕ_t unless ϕ_t is stationary.*

Proof. Differentiating along the flow gives

$$-\frac{dI(\phi_t)}{dt} = -\int_X \log \frac{\mathrm{MA}(\phi_t)}{\mu(\phi_t)}\, \mu(\phi_t) \geq 0$$

using Jensen's inequality applied to the concave function $f(t) = \log t$ on $\mathbb{R}_+$ in the last step (recall that $\mathrm{MA}(\phi_t)$, $\mu(\phi_t)$ are both probability measures). Similarly,

$$\frac{d\mathscr{E}(\phi_t)}{dt} = \int_X \log \frac{\mathrm{MA}(\phi_t)}{\mu(\phi_t)}\, \mathrm{MA}(\phi_t) = -\int_X \log \frac{\mu(\phi_t)}{\mathrm{MA}(\phi_t)}\, \mathrm{MA}(\phi_t) \geq 0,$$

again using Jensen's inequality, but with the roles of $\mathrm{MA}(\phi_t)$, $\mu(\phi_t)$ reversed. The statement about strict monotonicity also follows from Jensen's inequality since $f(t) = \log t$ is strictly concave. □

From the previous lemma we deduce the following compactness property of the flow.

Lemma 2.3. *Assume that $\mu(\phi)$ is normalized and that the associated functional $-\mathscr{F}_\mu$ is coercive. Then there is a constant C such that $J(\phi_t) \leq C$ and $\int |\phi_t - \phi_0|\, \mu_0 \leq C$ along the Kähler–Ricci flow for ϕ_t (with respect to $\mu(\phi)$).*

Proof. Combining the monotonicity of $\mathscr{F}_\mu$ and the assumption that $\mathscr{F}_\mu$ be coercive (and in particular proper) immediately gives the first inequality $J(\phi_t) \leq C$. Next, by the definition of coercivity there are $\delta \in]0, 1[$ and $C_\delta > 0$ such that $I_\mu - \mathscr{E} \geq \delta I_{\mu_0} - \delta\mathscr{E} - C_\delta$; that is,

$$\delta I_{\mu_0} \leq (-1 + \delta)\mathscr{E} + I_\mu + C_\delta$$

along the flow. Since by the previous lemma $-\mathcal{E}$ and I_μ are both bounded from above along the flow it follows that there is a constant A such that $I_{\mu_0} \leq A$ along the flow. Finally, by basic pluripotential theory the set $\{\phi \in \mathcal{H}_L : J(\phi) \leq C, I_{\mu_0}(\phi) \leq C\}$ is relatively compact in the L^1-topology [Berman et al. 2009]. This proves the last inequality in the statement of the lemma. □

The next proposition shows that, under suitable assumptions, the Kähler–Ricci flow with respect to a normalized measure μ_ϕ converges on the level of weights precisely when it converges on the level of Kähler metrics. In Sections 3 and 4 the proposition will be applied to the usual geometric Kähler–Ricci flows, where the convergence is already known to hold on the level of Kähler metrics. To simplify the notation we will only state the result in the absolute case, the extension to the relative case being immediate.

Proposition 2.4. *Assume that $\mu(\phi)$ is normalized and that the associated functional $-\mathcal{F}_\mu$ is coercive. Let ϕ_t evolve according to the Kähler–Ricci flow defined with respect to μ_ϕ and write $\omega_t = dd^c\phi_t$. Then the following three statements are all equivalent*:

- *The sequence of Kähler metrics ω_t is relatively compact in the $\mathcal{C}^\infty$-topology on X; that is, for any positive integer l the sequence ω_t is uniformly bounded in the $\mathcal{C}^l$-norm on X.*
- *The weights converge: $\phi_t \to \phi_\infty \in \mathcal{H}_L$ in the $\mathcal{C}^\infty$-topology on X as $t \to \infty$.*
- *The Kähler metrics $\omega_t \to \omega_\infty$ in the $\mathcal{C}^\infty$-topology on X, where ω_∞ is a Kähler form.*

Proof. Assume that the first point of the proposition holds. Then it is a basic fact that the sequence of normalized weights $\tilde{\phi}_t := \phi_t - C_t$, where $C_t := I_{\mu_0}(\phi_t)$, is relatively compact in the $\mathcal{C}^\infty$-topology on X and converges to $\tilde{\phi}_\infty \in \mathcal{H}_L$ (as is seen by inverting the associated Laplacians). By the previous lemma $|C_t| \leq D$ for some positive constant D and hence $\{\phi_t\}$ is also relatively compact in the $\mathcal{C}^\infty$-topology on X.

In the rest of the argument we will use the $\mathcal{C}^l$-topology on $\mathcal{H}_L$ for l a large fixed integer. Let $\mathcal{K} := \overline{\{\phi_t\}}$ be the closure of $\{\phi_t\}$ which is relatively compact in $\mathcal{H}_L$ by the previous argument. Denote by ψ_0 an accumulation point in $\mathcal{K}$:

$$\lim_j \phi_{t_j} = \psi_0.$$

By continuity of the "time s flow map" (which follows immediately from the stability assumption on the flow) and the semigroup structure of the flow we deduce that

$$\lim_j \phi_{t_j+s} = \psi_s$$

for any fixed $s > 0$. In other words, $\mathcal{K}$ is in fact *compact* and *invariant* under the "time s flow map". Note also that by monotonicity

$$\lim_t \mathcal{E}(\phi_t) = \mathcal{E}(\psi_0) = \sup_{\mathcal{K}} \mathcal{E}. \tag{2-6}$$

Assume now to get a contradiction that $\psi_s \neq \psi_0$. By the strict monotonicity in Lemma 3.8 we have that $\mathcal{E}(\psi_s) > \mathcal{E}(\psi_0)$, contradicting (2-6) (since $\psi_s \in \mathcal{K}$ as explained above). Hence, ψ_0 is a fixed point of the flow and hence, by the uniqueness assumption on the flow, it is determined up to an additive constant. This means that for any two limit points ψ_0 and ψ_0' of the flow there is a constant C such that

$$\psi_0 - \psi_0' = C.$$

But as explained above $\mathcal{E}(\psi_0) = \mathcal{E}(\psi_0')$ and hence, by the scaling equivariance of $\mathcal{E}$, it follows that $C = 0$. All in all this means that we have shown that the flow ϕ_t converges, in the $\mathcal{C}^\infty$-topology on X, to a limit ϕ_∞ in $\mathcal{H}_L$, that is, that the second point of the proposition holds. The rest of the implications are trivial. □

Remark 2.5. The coercivity is used to make sure that the compactness property of the flow ϕ_t holds without normalizing ϕ_t (say, by subtracting $I_{\mu_0}(\phi_t)$). If one only assumes properness then the same proof shows that the statement still holds upon replacing ϕ_t by $\phi_t - I_{\mu_0}(\phi_t)$ (which, of course, does not effect the curvature forms). The same remark applies to Proposition 2.9 below.

2D. ***Quantization: The Bergman iteration on $\mathcal{H}_L$.*** Proceeding fiberwise it will be enough to consider the absolute case when S is a point and we are given an ample line bundle $L \to X$. For any positive integer k such that kL is very ample the quantization at level k of the space $\mathcal{H}_L$ is defined as the space $\mathcal{H}^{(k)}$ of all Hermitian metrics on the N_k-dimensional complex vector space $H^0(X, kL)$. Hence, $\mathcal{H}^{(k)}$ may be identified with the symmetric space $\mathrm{GL}(N_k, \mathbb{C})/U(N_k)$. In the relative setting $\mathcal{H}^{(k)}$ is replaced by the space of all Hermitian metrics on the rank-N_k vector bundle $\pi_*(k\mathcal{L})$ over the base S (compare the discussion at the bottom of page 156).

Fix a volume form μ_ϕ on X depending on ϕ as above. Then any given $\phi \in \mathcal{H}_L$ induces a Hermitian metric $\mathrm{Hilb}^{(k)}(\phi)$ defined by

$$\mathrm{Hilb}^{(k)}(\phi)(f, f) := \int_X |f|^2 e^{-k\phi}\, d\mu_\phi,$$

giving a map

$$\mathrm{Hilb}^{(k)} : \mathcal{H}_L \to \mathcal{H}^{(k)}.$$

There is also a natural injective map (independent of μ_ϕ) in the reverse direction, called the (scaled) *Fubini–Study map* $\mathrm{FS}^{(k)}$:

$$\mathrm{FS}^{(k)}(H) := \log\left(\frac{1}{N_k}\sum_{i=1}^{N_k} |f_i^H|^2\right)$$

where (f_i^H) is any basis in $H^0(X, kL)$ that is orthonormal with respect to H.

Donaldson's iteration (*with respect to* μ_ϕ) on the space $\mathcal{H}^{(k)}$ is then obtained by iterating the composed map

$$T^{(k)} := \mathrm{Hilb}^{(k)} \circ \mathrm{FS}^{(k)} : \mathcal{H}^{(k)} \to \mathcal{H}^{(k)},$$

and its fixed points are called *balanced metrics at level* k (*with respect to* μ).

In order to facilitate the comparison with the Kähler–Ricci flow it will be convenient to consider the (essentially equivalent) iteration on the space $\mathcal{H}_L$ obtained by iterating the map $\mathrm{FS}^{(k)} \circ \mathrm{Hilb}^{(k)}$. This latter iteration will be called the *Bergman iteration at level* k (*with respect to* μ_ϕ) and we will denote the m-th iterate by $\phi_m^{(k)}$ and call the parameter m *discrete time*. Hence, the iteration immediately enters the *finite-dimensional* submanifold $\mathrm{FS}(\mathcal{H}^{(k)}) \subset \mathcal{H}_L$ of Bergman metrics at level k and stays there forever. By the very definition of the Bergman iteration it may be written as the difference equation

$$\phi_{m+1}^{(k)} - \phi_m^{(k)} = \frac{1}{k} \log \rho^{(k)}(\phi_m^{(k)}),$$

where $\rho^{(k)}(\phi)$ is the *Bergman function at level* k associated to (μ_ϕ, ϕ); that is,

$$\rho^{(k)}(\phi) = \frac{1}{N_k} \sum_{i=1} |f_i|^2 e^{-k\phi},$$

where f_i is an orthonormal basis with respect to the Hermitian metric $\mathrm{Hilb}^{(k)}(\mu, \phi)$. Note that the *Bergman measure* $\rho^{(k)}(\phi)\mu_\phi$ is a probability measure on X and independent of the choice of orthonormal bases. It plays the role of the Monge–Ampère measure in the quantized setting.

It will also be convenient, following [Donaldson 2009], to study functionals defined directly on the space $\mathcal{H}^{(k)}$. Fixing the reference metric $H_0^{(k)} := \mathrm{Hilb}^{(k)}(\phi_0) \in \mathcal{H}^{(k)}$ we may identify $\mathcal{H}^{(k)}$ with the space of all rank N_k Hermitian matrices. We define

$$\mathcal{F}_\mu^{(k)}(H) := -\frac{1}{N_k k} \log\det H - I_\mu \circ \mathrm{FS}^{(k)}(H),$$

whose critical points in $\mathcal{H}^{(k)}$ are precisely the balanced metrics (with respect to μ_ϕ); this is proved exactly as in the particular cases considered in [Donaldson 2005; Berman et al. 2009]. We will also consider the following functional on $\mathcal{H}_L$:

$$\mathcal{L}^{(k)}(\phi) := -\frac{1}{N_k k} \log\det \mathrm{Hilb}^{(k)}(\mu_\phi, \phi),$$

normalized so that $\mathcal{L}^{(k)}(\phi + c) = \mathcal{L}^{(k)}(\phi) + c$. Equivalently, we could have defined $\mathcal{L}^{(k)}$ as the antiderivative of the one-form on $\mathcal{H}_L$ defined by integration against the Bergman measure $\rho^{(k)}(\phi)\mu_\phi$.

Monotonicity. The following monotonicity properties were shown in [Donaldson 2009] in the particular setting considered there (where μ_ϕ is independent of ϕ). See also [Donaldson 2005] for the setting when $\mu(\phi) = \mathrm{MA}(\phi)$ (compare Section 4F). The main new observation here is that concavity of I_μ implies monotonicity.

Lemma 2.6. *Assume that* μ_ϕ *is normalized. Then the following monotonicity with respect to the discrete time* m *holds along the Bergman iteration* $\phi_m^{(k)}$ *on* $\mathcal{H}_L$ (*defined with respect to* μ_ϕ):

- *The functional* $\mathcal{L}^{(k)}$ *is increasing along the Bergman iteration and strictly increasing at* $\phi_m^{(k)}$ *unless* $\phi_m^{(k)}$ *is stationary. Equivalently, the functional* $-\log\det$ *is strictly increasing along the Donaldson iteration in* $\mathcal{H}^{(k)}$ *away from balanced metrics.*

- *If I_μ is concave on the space $\mathcal{H}_L$ with respect to the affine structure then it is decreasing along the iteration and strictly decreasing at $\phi_m^{(k)}$ unless $\phi_m^{(k)}$ is stationary. Equivalently, the functional $I_\mu \circ \mathrm{FS}^{(k)}$ is strictly decreasing along the Donaldson iteration in $\mathcal{H}^{(k)}$ away from balanced metrics.*

Proof. The proof of the first point is essentially the same as in Donaldson's setting [2009], but for completeness we repeat it here. By definition

$$\mathcal{L}^{(k)}(\phi_{m+1}) - \mathcal{L}^{(k)}(\phi_m) = -\frac{1}{N_k k}\log\frac{\det \mathrm{Hilb}^{(k)}(\phi_{m+1})}{\det \mathrm{Hilb}^{(k)}(\phi_m)}.$$

By the concavity of log and Jensen's inequality we hence get

$$\mathcal{L}^{(k)}(\phi_{m+1}) - \mathcal{L}^{(k)}(\phi_m) \geq -\frac{1}{k}\log\frac{1}{N_k}\sum_{i=1}^{N_k}\|f_i\|^2_{T(\mathrm{Hilb}^{(k)}(\phi_m))},$$

where f_i is an orthonormal basis with respect to the Hermitian metric $\mathrm{Hilb}^{(k)}(\phi_m)$ and where by definition $T(\mathrm{Hilb}^{(k)}(\phi_m)) = \mathrm{Hilb}^{(k)}(\mathrm{FS}(\mathrm{Hilb}^{(k)}(\phi_m)))$. Writing out the norms explicitly shows that the right-hand side above may be written as

$$-\frac{1}{k}\log\biggl(\frac{1}{N_k}\biggl(\sum_{i=1}^{N_k}|f_i|^2\Big/\sum_{i=1}|f_i|^2\biggr)\mu_{\mathrm{FS}(\mathrm{Hilb}^{(k)}(\phi_m))}\biggr) = -\frac{1}{k}\log(1) = 0,$$

using that μ_ϕ is normalized. This proves the first point.

To prove the second point we use that I_μ is assumed concave and that, by definition, $\mu_\phi = dI_\mu$ as a differential, to get

$$\begin{aligned} I_\mu(\phi_{m+1}^{(k)}) - I_\mu(\phi_m^{(k)}) &\leq \int (\phi_{m+1}^{(k)} - \phi_m^{(k)})\,\mu_{\phi_m^{(k)}} = \frac{1}{k}\int \log\rho^{(k)}(\phi_m^{(k)})\,\mu_{\phi_m^{(k)}} \\ &\leq \frac{1}{k}\log\int \rho^{(k)}(\phi_m^{(k)})\,\mu_{\phi_m^{(k)}} = 0, \end{aligned}$$

using the definition of the iteration and Jensen's inequality in the last step (and the fact that $\rho^{(k)}(\phi)\mu_\phi$ and μ_ϕ are both probability measures). This proves the monotonicity of I_μ. The statement about strict monotonicity follows immediately from the fact that $\log t$ is strictly concave. ☐

Properness and coercivity. Properness and coercivity of functionals on $\mathcal{H}^{(k)}$ are defined as in Section 2C, but with the functional J replaced by its quantized version on the space $\mathcal{H}^{(k)}$:

$$J^{(k)}(H) := -\mathcal{F}^{(k)}_{\mu_0} := I_{\mu_0}\circ \mathrm{FS}^{(k)} + \frac{1}{kN_k}\log\det H.$$

The content of the following lemma is essentially contained in the proof of Proposition 3 in [Donaldson 2009]. We will fix a metric $H_0 \in \mathcal{H}^{(k)}$. For any given H_0-orthonormal basis (f_i) we can then identify a Hermitian metric H with a matrix and we will denote by H_λ the diagonal matrix with entries $e^{-\lambda_i}$ on the diagonal.

Lemma 2.7.

- *For* $\lambda \in \mathbb{C}^{N_k}$ *let* $\phi_\lambda = \mathrm{FS}^{(k)}(H_\lambda) := \frac{1}{k}\log\Big(\frac{1}{N_k}\sum_i e^{k\lambda_i}|f_i|^2\Big)$. *There is a constant* C *such that*
$$\max_i \lambda_i \leq I_{\mu_0}(\phi_\lambda) + C.$$
- *The functional* $J^{(k)}$ *is an exhaustion function on* $\mathcal{H}^{(k)}/\mathbb{R}^*$ *with respect to its usual topology.*
- *In particular, the set of all* $H \in \mathcal{H}^{(k)}$ *such that*
$$-\log\det H \geq -C, \quad (I_{\mu_0}\circ \mathrm{FS})(H) \leq C \tag{2-7}$$
is relatively compact.

Proof. For the benefit of the reader we repeat Donaldson's simple proof: let $i_{\max}$ be an index such that $\max_i \lambda_i = \lambda_{i_{\max}}$. Clearly,
$$\max_i \lambda_i + \frac{1}{k}\log\Big(\frac{1}{N_k}|f_{i_{\max}}|^2\Big) \leq \phi_\lambda \leq \max_i \lambda_i + \frac{1}{k}\log\Big(\frac{1}{N_k}\sum_i |f_i|^2\Big), \tag{2-8}$$
and hence integrating over X and using the first inequality above gives
$$\max_i \lambda_i + \int_X (\log(|f_{i_{\max}}|^2) - \phi_0)\, d\mu_0 \leq I_{\mu_0}(\phi_\lambda),$$
which proves the lemma since it is well-known that $I_{\mu_0}(\psi) > -\infty$ for any psh (plurisubharmonic) weight ψ if μ_0 is a smooth volume form (as follows from the local fact that any psh function is in L^1) and in particular $-C := I_\mu(\log(|f_{i_{\max}}|^2)) > -\infty$. This proves the first point. As for the second and third one we first note that any Hermitian metric H can be represented by a diagonal matrix (which we write in the form H_λ) after perhaps changing the basis (f_i) above. Moreover, by the compactness of $U(N)$ the constant C in the previous point can be taken to be independent of the base (f_i).

Next, it will be enough to prove the last point of the lemma (the second point then follows since we may by scaling invariance assume that $\det(H_\lambda) = 1$). We may assume that $\inf_i \lambda_i = \lambda_0$ and since, by assumption,
$$-\log\det H = \sum_i \lambda_i \geq -C$$
we get
$$-\inf_i \lambda_i \leq C + \sum_{i\neq 0} \lambda_i \leq C + (N-1)\max_i \lambda_i.$$
By the assumption $(I_{\mu_0}\circ \mathrm{FS})(H) \leq C$ and the first point of the lemma the right-hand side above is bounded from above and hence we conclude that so is $-\inf_i \lambda_i$. All in all this means that $\max_i |\lambda_i|$ is uniformly bounded from above by a constant; that is, H stays in a relatively compact subset of $\mathcal{H}^{(k)}$. □

Remark 2.8. The proof of the previous lemma shows that the conclusion of the lemma remains valid for any choice of a fixed reference weight ϕ_0 and probability measure μ_0 (which are used in the definition of $J^{(k)}$) such that $\int_X \log(|f| - \phi_0)\,\mu_0$ is finite for any section $f \in H^0(X, kL)$.

Criteria for convergence in the large time limit.

Proposition 2.9. *Assume that μ_ϕ is normalized, that I_μ is decreasing along the Bergman iteration, that $\mathcal{F}_\mu^{(k)}$ is coercive and that there is at most one balanced metric (modulo scaling). Then, for any given positive integer k the following holds: in the large time limit, that is, when $m \to \infty$, the weights $\phi_m^{(k)} \to \phi_\infty^{(k)}$ in the $\mathscr{C}^\infty$-topology on X. Moreover, in the relative setting the convergence is uniform with respect to the base parameter s.*

Proof. (a) *Uniform convergence*. We equip $\mathrm{FS}(\mathscr{H}^{(k)})$, that is, the space of all Bergman weights at level k, with the topology induced by the sup norm. It is not hard to see that this is the same topology as the one induced from the finite-dimensional symmetric space $\mathscr{H}^{(k)} = \mathrm{GL}(N_k, \mathbb{C})/U(N_k)$ with its usual Riemannian metric, or with respect to the operator norm on $\mathrm{GL}(N_k, \mathbb{C})$. Hence, it will be enough to prove the convergence of Donaldson's iteration in $\mathscr{H}^{(k)}$.

Since μ_ϕ is assumed normalized, Lemma 2.6 shows that $-\log\det H$ is uniformly bounded from below along the Donaldson iteration in $\mathscr{H}^{(k)}$. Moreover, by assumption $I_{\mu_\phi} \circ \mathrm{FS}^{(k)}$ is uniformly bounded from above along the Donaldson iteration. Hence, just as in the proof of Lemma 2.3 it follows from the coercivity assumption that $I_{\mu_0} \circ \mathrm{FS}^{(k)}$ is also uniformly bounded from above along the Donaldson iteration. But then it follows from Lemma 2.7 that the iteration $H_m^{(k)}$ stays in a compact subset of $\mathscr{H}^{(k)}$.

Now let $\mathcal{K} := \overline{\{H_m^{(k)}\}}$ be the closure of the orbit of $T^{(k)}$ which is relatively compact in $\mathscr{H}^{(k)}$ by the previous argument. Denote by G an accumulation point

$$\lim_j H_{m_j}^{(k)} = G$$

in $\mathscr{H}^{(k)}$. By the continuity of $H \mapsto T^{(k)}(H)$ on $\mathscr{H}^{(k)}$ we deduce that

$$\lim_j T^{(k)}(H_{m_j}^{(k)}) = T^{(k)}(G).$$

In other words, $\mathcal{K}$ is in fact *compact* and *invariant under* $T^{(k)}$. Note also that by monotonicity

$$\lim_j (-\log\det H_{m_j}^{(k)}) = -\log\det G = \sup_{\mathcal{K}} (-\log\det).$$

Assume now to get a contradiction that $T^{(k)}(G) \neq G$. By the strict monotonicity in Lemma 3.8 we have $\log\det(T^{(k)}G) > \log\det G$, contradicting (2-6) (since $T^{(k)}(G) \in \mathcal{K}$). All in all this means that we have shown that the subsequence $(H_{m_j}^{(k)})$ of Donaldson iterations converges to a fixed point, that is, a balanced metric. By the assumption on uniqueness up to scaling it follows, again using monotonicity (just like in the proof of Proposition 2.4), that all accumulation points coincide; that is, the iteration converges.

(b) *Higher order convergence*. To simplify the notation we set $k = 1$ and write $\phi_m^{(k)} = \phi_m$. First note that the L^∞-estimate above is uniform over S, as follows by combining the monotonicity of the functionals with the uniform boundedness of the initial weight ϕ_0. By the uniform convergence of ϕ_m it will hence be enough to prove that

$$\left\| \partial_X^\alpha (h_0/h_{m+1}) \right\|_{L^\infty(X)} \leq C_\alpha \left\| (h_m/h_0) \right\|_{L^\infty(X)} \tag{2-9}$$

where $h_m = e^{-\phi_m}$ and ∂_X^α denotes a real linear differential operator on X of order α (note that while h_m globally corresponds to a metric on L the quotient h_0/h_{m+1} defines a global *function* on X). Accepting this estimate for the moment the uniform convergence of (h_m) hence gives that $\|\partial_X^\alpha(h_0/h_m)\|_{L^\infty(X)}$ is uniformly bounded in m and since $h_m/h_0 \to h_\infty/h_0$ it then follows that $\|\partial_X^\alpha(\phi_m-\phi_0)\|_{L^\infty(X)}$ is also uniformly bounded in m. Hence, standard compactness arguments show the $\mathscr{C}^\infty$-convergence of (ϕ_m).

Finally, the estimate (2-9) is a consequence of the following quasiexplicit integral formula for the Bergman function familiar from the theory of determinantal random point processes (see [Berman 2008] and references therein):

$$\rho(\phi)(x)=\int_{y\in X^{N-1}} f(x,y)e^{-(\phi-\phi_0)(x)}e^{-(\phi-\phi_0)(y)}\,d\mu_\phi(y)^{\otimes N-1}/Z_\phi,\quad Z_\phi:=\int_{X^N} f_0e^{-(\phi-\phi_0)}\,d\mu_\phi^{\otimes N}$$

where $f(x_1,x_2,\dots,x_N)=|\det_{1\le i,j\le N}(f_i(x_i))_{i,j}|^2e^{-\phi_0(x_1)}\cdots e^{-\phi_0(x_N)}$ and (f_i) is any given orthonormal base with respect to the Hermitian metric $\mathrm{Hilb}^{(1)}(\phi_0)$ on $H^0(X,L)$ (note that Z_ϕ appears as the normalizing constant). We have used the notation $\phi(x,\dots,x_m)=\phi(x_1)+\cdots+\phi(x_m)$. In particular,

$$(h_0/h_{m+1})(x)=\int_{y\in X^{N-1}} f(x,y)e^{-(\phi_m-\phi_0)(y)}\,d\mu_\phi(y)^{\otimes N-1}/Z_{\phi_m}$$

and hence differentiating with respect to x by applying ∂_X^α gives

$$|\partial_X^\alpha(h_0/h_{m+1})(x)|=\left|\int(\partial_X^\alpha f(x,y))e^{-(\phi-\phi_0)(y)}\,d\mu_\phi(y)^{\otimes N-1}/Z_\phi\right|\le\frac{A_\alpha}{Z_{\phi_m}}\|e^{-(\phi_m-\phi_0)}\|_{L^\infty(X)},$$

where A_α is a constant independent of m. Since, by the uniform convergence of ϕ_m, we have $Z_{\phi_m}>C>0$ for some positive constant C, this concludes the proof of the estimate (2-9). ☐

The following basic lemma gives a natural criterion for the assumptions (apart from the monotonicity of I_μ) in the previous theorem to be satisfied.

Lemma 2.10. *Suppose that $\mathscr{G}$ is a functional on $\mathscr{H}^{(k)}$ which is geodesically strictly convex with respect to the symmetric Riemann structure and strictly convex modulo scaling. Then $\mathscr{G}$ has at most one critical point (modulo scaling). Moreover, if it has some critical point then $\mathscr{G}$ is coercive.*

Proof. Uniqueness follows immediately from strict convexity and hence we turn to the proof of coercivity. By a simple compactness argument it will be clear that, after fixing a reference metric $H_0\in\mathscr{H}^{(k)}$, which we take to be a critical point of $\mathscr{G}$, it is enough to prove coercivity along any fixed geodesic passing through H_0. To this end let H_t be a geodesic in $\mathscr{H}^{(k)}$ starting at H_0, that is, the orbit of the action of a one-parameter subgroup of $\mathrm{GL}(N_k)$. In the notation of Lemma 2.7 this means that $H_t=H_{t\lambda}$ for $\lambda\in\mathbb{C}^N$ fixed. By scaling invariance we may assume that the determinant of H_t vanishes along the geodesic. Integrating the upper bound in (2-8) over X gives

$$J(H_t)=0+(I_{\mu_0}\circ\mathrm{FS})(H_t)\le Ct+D.$$

Now, let $f(t)=\mathcal{G}(H_t)$. Since by assumption f is convex and 0 is a critical point, we have $df/dt \geq 0$ for all t. Hence, if we fix some number $\epsilon > 0$, then

$$f(t) \geq f(0) + \int_\epsilon^t (df/ds)\, ds.$$

But by the assumption on strict convexity the latter integrand is bounded from below by some $\delta > 0$. All in all this shows that

$$\mathcal{G}(H_t) \geq \delta t - A \geq \frac{\delta}{C} J(H_t) - A',$$

which finishes the proof. □

Large k asymptotics. Next, we will recall the following proposition, which is the link between the Bergman iteration and the Kähler–Ricci flow. It is essentially due to Bouche and Tian, apart from the uniformity with respect to ϕ. In fact, a complete asymptotic expansion in powers of k holds as was proved by Catlin and Zelditch and the uniformity can be obtained by tracing through the same arguments (as remarked in connection to Proposition 6 in [Donaldson 2001]). For references see the recent survey [Zelditch 2009].

Proposition 2.11. *Assume that the volume form μ_ϕ depends smoothly on ϕ. Then the following uniform convergence for the corresponding Bergman function $\rho_{(k)}(\phi)$ holds*: *there is an integer l such that*

$$\sup_X \left| \rho_{(k)}(\phi) - \frac{(dd^c\phi)^n/n!}{\mu_\phi} \right| \leq C/k$$

for all weights ϕ such that $dd^c\phi$ is uniformly bounded from above in $\mathcal{C}^l$-norm with $dd^c\phi$ uniformly bounded from below by some fixed Kähler form.

3. The Calabi–Yau setting

First consider the absolute case where we assume given an ample line bundle $L \to X$. In this section we will the apply the general setting introduced in the previous section to the case when the measure μ is independent of ϕ. We will assume that it is normalized, that is, a probability measure. We will mainly be interested in the case when X is a Calabi–Yau manifold, which induces a canonical probability measure μ on X defined by

$$\mu = c_n \Omega \wedge \bar{\Omega}$$

where Ω is any given holomorphic n-form trivializing the canonical line bundle K_X and c_n is a normalizing constant. In the relative Calabi–Yau setting, where each fiber is assumed to be a Calabi–Yau manifold, this hence yields a canonical smooth family of measures on the fibers.

For a fixed reference element $\phi_0 \in \mathcal{H}_L$ we set

$$I_\mu(\phi) := \int_X (\phi - \phi_0)\, \mu,$$

which is *equivariant* under the usual actions of the additive group $\mathbb{R}$: $I_\mu(\phi + c) = I_\mu(\phi) + c$. Moreover, by definition the associated functional $-\mathcal{F}_\mu$ is coercive.

3A. ***The relative Kähler–Ricci flow.*** The convergence on the level of Kähler forms in the following theorem is due to Cao (apart from the uniqueness, which was first shown by Calabi). We just observe that, since μ is normalized, the convergence of the flow also holds on the level of weights.

Theorem 3.1. *The Kähler–Ricci flow on $\mathcal{H}_L$ with respect to μ exists for all times $t \in [0, \infty[$ and the solution ϕ_t is smooth on $X \times [0, \infty[$. Moreover, $\phi_t \to \phi_\infty$ uniformly in the $\mathcal{C}^\infty$- topology on X when $t \to \infty$, where ϕ_∞ is the unique (modulo scaling) solution to the inhomogeneous Monge–Ampère equation* (1-2). *More precisely, all the analytical assumptions in Section 2C are satisfied. In the Calabi–Yau case ω_∞ is Ricci flat.*

Proof. As shown in [Cao 1985], $\omega_t \to \omega_\infty$ in the $\mathcal{C}^\infty$-topology. But then it follows from Proposition 2.4 that $\phi_t \to \phi_\infty$ uniformly in the $\mathcal{C}^\infty$-topology on X. The smoothness in the relative case was not stated explicitly in [Cao 1985] but follows from basic maximum principle arguments. □

Preliminaries: Kodaira–Spencer classes and Weil–Petersson geometry. In this section we will assume that the base S is one-dimensional and embedded as a domain in $\mathbb{C}$. Recall that the infinitesimal deformation of the complex structures on the smooth manifold $\mathcal{X}_s$ as s varies is captured by the Kodaira–Spencer class $\rho(\frac{\partial}{\partial s}) \in H^{0,1}(T^{1,0}\mathcal{X}_s)$ [Voisin 2007]. When the fibers are Calabi–Yau manifolds the "size" of the deformation is measured by the (generalized) Weil–Petersson form ω_{WP} (see [Fujiki and Schumacher 1990]) on the base S. It was extensively studied by Tian [1987] and Todorov [1989] when the base S is a moduli space of Calabi–Yau manifolds and $\mathcal{X}$ is the corresponding Kuranishi family. The form ω_{WP} is defined by

$$\omega_{\mathrm{WP}}\left(\frac{\partial}{\partial s}, \frac{\partial}{\partial s}\right) := \|A_{CY}\|^2_{\omega_{CY}}, \tag{3-1}$$

where A_{CY} denotes the unique representative in the Kodaira–Spencer class $\rho(\partial/\partial s) \in H^{1,0}(T^{1,0}\mathcal{X}_s)$ that is harmonic with respect to a given Ricci flat metric ω_{CY} on $\mathcal{X}_s$ and the L^2-norm is computed with respect to this latter metric. Moreover, as shown in [Todorov 1989] the following formula holds:

$$\|A_{CY}\|^2_{\omega_{CY}} = \frac{\partial^2 \psi_\Omega}{\partial s \partial \bar{s}}, \quad \psi_\Omega(s) := \log i^{n^2} \int_{\mathcal{X}_s} \Omega_s \wedge \bar{\Omega}_s, \tag{3-2}$$

where Ω_s denotes a holomorphic family of nontrivial holomorphic n-forms on $\mathcal{X}_s$ for $s \in U$, a neighborhood of a fixed point s in S. More generally, for an arbitrary smooth base S the $(1, 1)$-form ω_{WP} on S may be defined as the curvature of the holomorphic line bundle $\pi_*(K_{\chi/S})$ on S. It is in the latter form that ω_{WP} will appear in the proof of Theorem 3.3 below. In fact, the formula (3-1) may then be deduced from Theorem 3.3 (see Remark 3.7).

Next we will explain how, for a fixed base parameter s, a weight ϕ on the line bundle $\mathcal{L} \to \mathcal{X} \to S$ induces the following two objects:

- a $(0, 1)$-form A_ϕ with values in $T^{1,0}\mathcal{X}_s$ representing the Kodaira–Spencer class $\rho(\frac{\partial}{\partial s})$ in $H^{0,1}(T^{1,0}\mathcal{X}_s)$;
- a function $c(\phi)$ on $\mathcal{X}$ measuring the positivity (or lack of positivity) of $dd^c\phi$ on $\mathcal{X}$ in terms of the positivity of the restrictions of $dd^c\phi$ to the fibers $\mathcal{X}_s$.

In fact A_ϕ will only depend on the family, parametrized by s, of two-forms ω_s obtained as the *restrictions* of the curvature form ω_ϕ on $\mathscr{X}$ to all fibers $\mathscr{X}_s$, while $c(\phi)$ will depend on the whole form ω_ϕ.

Trivial fibrations. Assume that $\pi : \mathscr{X} \to S$ is a holomorphically trivial fibration, so that $\mathscr{X}$ is embedded in $\mathbb{C} \times X$ and that $\mathscr{L} = \pi^* L$ where $L \to X$ is an ample line bundle. Given a smooth family of weights $\phi(s, \cdot)$ on $L \to X$ with strictly positive curvature form $\omega_\phi^X := d_X d_X^c \phi$ (for s fixed) one obtains a smooth vector field V_ϕ of type $(1,0)$ as the "complex gradient" of $\partial_s \phi$:

$$\delta_{V_\phi} \omega^X_{\phi(s,\cdot)} = \bar{\partial}_X(\partial_s \phi), \tag{3-3}$$

where δ_{V_ϕ} denotes interior multiplication (i.e., contraction) with V_ϕ. Now the $(0,1)$-form A_ϕ with values in $T^{1,0}X$ (for s fixed) is simply defined by

$$A_\phi := -\bar{\partial}_X V_\phi \tag{3-4}$$

Denote by ω_t^X the curvature forms on X evolving with respect to the time parameter t according to the Kähler–Ricci flow (for s fixed). The Laplacian on X with respect to ω_t^X will be denoted by $\Delta_{\omega_t^X}$. Given $\phi(s, \cdot)$ we define the following function on $\mathscr{X}$:

$$c(\phi) := \frac{1}{n}(dd^c\phi)^{n+1}/(d_X d_X^c \phi)^n \wedge i\, ds \wedge d\bar{s}. \tag{3-5}$$

Note that, since $\omega_\phi^X > 0$ on X, we have that $c(\phi) > 0$ at $(s,x) \in \mathscr{X}$ if and only if $dd^c\phi > 0$ at (s,x).

General submersions. Next we turn to the case of a general holomorphic submersion $\pi : \mathscr{X} \to S$. Any given point in $\mathscr{X}$ has a neighborhood $\mathscr{U}$ such that the fibration $\pi : \mathscr{U} \to S$ is holomorphically trivial and the restriction $\mathscr{L}_{\mathscr{U}}$ is isomorphic to $\pi^* L$ over $\mathscr{U}$. We introduce local holomorphic coordinates (z,s) on $\mathscr{U}$ such that s defines a local holomorphic coordinate on S and the projection $\pi : \mathscr{U} \to S$ corresponds to $(z,s) \mapsto s$. Hence, the vector field V_ϕ defined above is *locally* defined, but in general not *globally* well-defined on $\mathscr{X}$. However, the expression (3-4) turns out to still be globally well-defined. For completeness we will give a proof of this well-known fact [Schumacher 2008; Fujiki and Schumacher 1990]:

Proposition 3.2. *The* $(0,1)$*-form* A_ϕ *with values in* $T^{0,1}\mathscr{X}_s$*, locally defined by formula* (3-4)*, is globally well-defined. It represents the Kodaira–Spencer class in* $H^{0,1}(T^{1,0}\mathscr{X}_s)$.

Proof. Step 1. The locally defined expression

$$W_\phi := \frac{\partial}{\partial s} - V_\phi$$

defines a global vector field on $\mathscr{X}$ of type $(1,0)$.

Indeed W_ϕ may be characterized as the *horizontal lift of* $\partial/\partial s$ with respect to the $(1,1)$-form $dd^c\phi$ on $\mathscr{X}$, which is nondegenerate along fibers. To see this first note that

$$d\pi(W_\phi) = \frac{\partial}{\partial s} \quad \text{and} \quad dd^c\phi(W_\phi, \ker d\pi) = 0. \tag{3-6}$$

The first point is trivial and the second one follows from a direct calculation: locally we may decompose

$$dd^c\phi = d_z d_z^c \phi + \phi_{s\bar{s}} ds \wedge d\bar{s} + (\bar{\partial}_z \phi_s) \wedge ds + (\partial_z \phi_s) \wedge d\bar{s}.$$

Hence, for any fixed index i,

$$dd^c\phi\left(W_\phi, \frac{\partial}{\partial\bar{z}_i}\right) = -d_z d_z^c\phi\left(V_\phi, \frac{\partial}{\partial\bar{z}_i}\right) + 0 + \left(\frac{\partial}{\partial\bar{z}_i}\phi_s\right) + 0 = 0,$$

using the definition (3-3) of V_ϕ in the last step. Finally, note that the properties (3-6) determine W_ϕ uniquely: if W' is another local vector field satisfying (3-6) then clearly $Z := W_\phi - W'$ satisfies

$$d\pi(Z) = 0 \quad \text{and} \quad dd^c\phi(Z, \ker d\pi) = 0.$$

In particular, Z is tangential to the fibers and $dd^c\phi(Z, \bar{Z}) = 0$. But since $dd^c\phi$ is assumed to be nondegenerate along the fibers it follows that $Z = 0$.

Step 2. $A_\phi(s) = (\bar\partial W_\phi)_{\mathscr{X}_s}$ and $A_\phi(s)$ represents the Kodaira–Spencer class in $H^{0,1}(T^{1,0}\mathscr{X}_s)$.

The first formula above follows immediately from a local computation and the second one then follows directly from the definition of the Kodaira–Spencer class (where W_ϕ may be taken as *any* smooth lift to $T^{1,0}\mathscr{X}$ of the vector field $\partial/\partial s$ [Voisin 2007]). □

As for the function $c(\phi)$ defined by formula (3-5) it is still well-defined as we have fixed an embedding of S in $\mathbb{C}$.

Conservation of positivity along the relative Kähler–Ricci flow. Next comes one of the main results of the present paper:

Theorem 3.3. *Let $\pi : \mathscr{X} \to S$ be a proper holomorphic submersion with Calabi–Yau fibers and let $\mathscr{L}$ be a relatively ample line bundle over $\mathscr{X}$. Assume that the base S is a domain in $\mathbb{C}$. The following equation holds along the corresponding relative Kähler–Ricci flow*:

$$\left(\frac{\partial}{\partial t} - \Delta_{\omega_t^X}\right)c(\phi) = |A_\phi|^2_{\omega_t^X} - \|A_{CY}\|^2_{\omega_{CY}^X}\,. \tag{3-7}$$

Proof. Since it will be enough to prove the identity at a fixed point x in X in some local holomorphic coordinates and trivializations we may as well assume that ω_ϕ is the Euclidean metric at the point x, that is, that the complex Hessian matrix $(\partial^2\phi/\partial z_i\bar\partial z_j)$ is the identity for $z = 0$ (corresponding to the fixed point x in X). Moreover, we may assume that locally the holomorphic n-form Ω may be expressed as $\Omega = dz_1 \wedge \cdots \wedge dz_n$. Partial derivatives with respect to s will be indicated by a subscript s and partial derivatives with respect to z_i and $\bar{z}_j$ by subscripts i and $\bar{j}$ respectively. If $h = (h_{ij})$ is a Hermitian matrix we will write (h^{ij}) for the matrix $\overline{H}^{-1}$. The summation convention according to which repeated indices are to be summed over will be used. Next, we turn to the proof of the theorem which is based on a direct and completely elementary calculation.

Step 1. The following formula holds in the case of a holomorphically trivial fibration:

$$\frac{\partial}{\partial t}c(\phi) = \phi_{i\bar{i}s\bar{s}} + \phi_{s\bar{i}}\overline{\phi_{s\bar{j}}}\phi_{i\bar{j}k\bar{k}} - \phi_{i\bar{j}s}\overline{\phi_{i\bar{j}s}} - \phi_{s\bar{i}}\overline{\phi_{s\bar{j}}}\phi_{ik\bar{l}}\phi_{j\bar{k}l} - 2\Re(\phi_{k\bar{k}s\bar{i}}\overline{\phi_{si}}) + 2\Re(\phi_{k\bar{l}s}\phi_{k\bar{l}\bar{i}}\overline{\phi_{s\bar{i}}}).$$

To see this first recall that

$$c(\phi) = \phi_{s\bar{s}} - \Re\big(\phi_{s\bar{i}}\overline{\phi_{s\bar{j}}}\phi^{i\bar{j}}\big)$$

and hence (using that $\phi_{i\bar{j}}=\delta_{ij}$ at $z=0$, so that $\partial\phi^{i\bar{j}}/\partial t=-\partial\phi_{\bar{j}i}/\partial t$ at $z=0$)

$$\frac{\partial}{\partial t}c(\phi)=\frac{\partial}{\partial t}\phi_{s\bar{s}}-2\Re\left(\frac{\partial}{\partial t}\phi_{s\bar{i}}\right)\overline{\phi_{s\bar{i}}}+(\phi_{s\bar{i}}\overline{\phi_{s\bar{j}}})\frac{\partial}{\partial t}\phi_{\bar{j}i} \tag{3-8}$$

Using the definition of the relative Kähler–Ricci flow in the Calabi–Yau case and the simple fact that the linearization of $\psi\mapsto\log\det(\psi_{k\bar{l}})$ at ψ is given by $u\mapsto\Delta_{\omega_\psi}u$, where $\Delta_{\omega_\psi}u=\psi^{k\bar{l}}u_{k\bar{l}}$ is the Laplacian with respect to the Kähler metric ω_ψ, hence gives

$$\begin{aligned}\frac{\partial}{\partial t}c(\phi)&=(\log(\det\phi_{i\bar{j}}))_{s\bar{s}}-2\Re((\log\det(\phi_{k\bar{l}}))_{s\bar{i}}\overline{\phi_{si}})+(\phi_{si}\overline{\phi_{s\bar{j}}})(\log(\det\phi_{k\bar{l}}))_{i\bar{j}}\\&=(\phi_{i\bar{j}s}\phi^{i\bar{j}})_{\bar{s}}-2\Re(\phi_{k\bar{l}s}\phi^{k\bar{l}})_{\bar{i}}\overline{\phi_{s\bar{i}}}+(\phi_{s\bar{i}}\overline{\phi_{s\bar{j}}})(\phi_{ik\bar{l}}\phi^{k\bar{l}})_{\bar{j}}\\&=\phi_{i\bar{i}s\bar{s}}-\phi_{i\bar{j}s}\phi_{j\bar{i}\bar{s}}-2\Re(\phi_{k\bar{k}s\bar{i}}\overline{\phi_{s\bar{i}}}-\phi_{k\bar{l}s}\phi_{l\bar{k}\bar{i}}\overline{\phi_{s\bar{i}}})+(\phi_{s\bar{i}}\overline{\phi_{s\bar{j}}})(\phi_{i\bar{j}k\bar{k}}-\phi_{ik\bar{l}}\phi_{\bar{j}\bar{k}l})\end{aligned}$$

(again using $\phi_{i\bar{j}}=\delta_{ij}$ at $z=0$), finishing the proof of Step 1.

Step 2. The following formula holds in the case of a trivial fibration:

$$\begin{aligned}c(\phi)_{k\bar{k}}=\phi_{k\bar{k}s\bar{s}}+(\phi_{s\bar{i}}\overline{\phi_{s\bar{j}}})(\phi_{k\bar{k}j\bar{i}})-2(\phi_{s\bar{i}}\overline{\phi_{s\bar{j}}})\phi_{\bar{k}\bar{i}m}\phi_{k\bar{m}j}-\phi_{ks\bar{i}}\overline{\phi_{ks\bar{i}}}-\overline{\phi_{\bar{k}s\bar{i}}}\phi_{\bar{k}s\bar{i}}+2\Re(\phi_{ks\bar{i}}\overline{\phi_{s\bar{j}}}\phi_{\bar{k}j\bar{i}})\\+2\Re(\overline{\phi_{\bar{k}s\bar{i}}}\phi_{s\bar{j}})\phi_{\bar{k}j\bar{i}}-2\Re\phi_{\bar{k}ks\bar{i}}\overline{\phi_{s\bar{i}}}.\end{aligned}$$

To see this we first differentiate $c(\phi)$ with respect to z_k to get

$$c(\phi)_k=\phi_{ks\bar{s}}-[(\phi_{s\bar{i}}\overline{\phi_{s\bar{j}}})_k\phi^{i\bar{j}}+(\phi_{s\bar{i}}\overline{\phi_{s\bar{j}}})(\phi^{i\bar{j}})_k]=\phi_{ks\bar{s}}-(\phi_{ks\bar{i}}\overline{\phi_{s\bar{j}}}+\overline{\phi_{\bar{k}s\bar{i}}}\phi_{s\bar{j}})\phi^{i\bar{j}}-(\phi_{s\bar{i}}\overline{\phi_{s\bar{j}}})(\phi^{i\bar{j}})_k.$$

Next, note that if h is a function with values in the space of Hermitian matrices and ∂ a derivation satisfying the Leibniz rule, then

$$\partial(h^{-1})=-h^{-1}(\partial h)h^{-1}.$$

In particular, if $h(0)=I$ then the following holds at 0:

$$(\bar{h}^{-1})_{k\bar{k}}=-\bar{h}_{k\bar{k}}+(\bar{h}_{\bar{k}}\bar{h}_k+\bar{h}_k\bar{h}_{\bar{k}}).$$

Applying this to $h=(\phi_{i\bar{j}})$ (when expanding the term A below) gives

$$\begin{aligned}c(\phi)_{k\bar{k}}&=\phi_{k\bar{k}s\bar{s}}-\big([(\phi_{ks\bar{i}}\overline{\phi_{s\bar{i}}})_{\bar{k}}+(\overline{\phi_{\bar{k}s\bar{i}}}\phi_{s\bar{i}})_{\bar{k}}]-(\phi_{ks\bar{i}}\overline{\phi_{s\bar{j}}}+\overline{\phi_{\bar{k}s\bar{i}}}\phi_{sj})\phi_{\bar{k}i\bar{j}}\big)-A\\&=\phi_{k\bar{k}s\bar{s}}-\big([\phi_{\bar{k}ks\bar{i}}\overline{\phi_{s\bar{i}}}+\phi_{ks\bar{i}}\overline{\phi_{ks\bar{i}}}+\overline{\phi_{k\bar{k}s\bar{i}}}\phi_{s\bar{i}}+\overline{\phi_{\bar{k}s\bar{i}}}\phi_{\bar{k}s\bar{i}}]-(\phi_{ks\bar{i}}\overline{\phi_{s\bar{j}}}+\overline{\phi_{\bar{k}s\bar{i}}}\phi_{s\bar{j}})\phi_{\bar{k}i\bar{j}}\big)-A,\end{aligned}$$

where

$$\begin{aligned}A&:=(\phi_{s\bar{i}}\overline{\phi_{s\bar{j}}})_{\bar{k}}(\phi^{i\bar{j}})_k+(\phi_{s\bar{i}}\overline{\phi_{s\bar{j}}})(\phi^{i\bar{j}})_{k\bar{k}}=-(\phi_{s\bar{i}}\overline{\phi_{s\bar{j}}})_{\bar{k}}\phi_{kj\bar{i}}+(\phi_{s\bar{i}}\overline{\phi_{s\bar{j}}})(-\phi_{k\bar{k}j\bar{i}}+2\Re(\phi_{\bar{k}\bar{i}m}\phi_{k\bar{m}j}))\\&=-(\phi_{s\bar{i}\bar{k}}\overline{\phi_{s\bar{j}}}+\phi_{s\bar{i}}\overline{\phi_{s\bar{j}k}})\phi_{kj\bar{i}}-(\phi_{s\bar{i}}\overline{\phi_{s\bar{j}}})(\phi_{k\bar{k}j\bar{i}}+2\Re(\phi_{\bar{k}\bar{i}m}\phi_{k\bar{m}j})).\end{aligned}$$

Hence,

$$\begin{aligned}c(\phi)_{k\bar{k}}=\phi_{k\bar{k}s\bar{s}}-[\phi_{\bar{k}ks\bar{i}}\overline{\phi_{s\bar{i}}}+\phi_{ks\bar{i}}\overline{\phi_{ks\bar{i}}}+\overline{\phi_{k\bar{k}s\bar{i}}}\phi_{s\bar{i}}+\overline{\phi_{\bar{k}s\bar{i}}}\phi_{\bar{k}s\bar{i}}]+(\phi_{ks\bar{i}}\overline{\phi_{s\bar{j}}}+\overline{\phi_{\bar{k}s\bar{i}}}\phi_{s\bar{j}})\phi_{\bar{k}j\bar{i}}\\+(\phi_{s\bar{i}\bar{k}}\overline{\phi_{s\bar{i}}}+\phi_{s\bar{i}}\overline{\phi_{s\bar{i}k}})\phi_{kj\bar{i}}+(\phi_{s\bar{i}}\overline{\phi_{s\bar{j}}})\phi_{k\bar{k}j\bar{i}}-2\Re(\phi_{s\bar{i}}\overline{\phi_{s\bar{j}}}\phi_{\bar{k}\bar{i}m}\phi_{k\bar{m}j}),\end{aligned}$$

which finishes the proof of Step 2.

Step 3: End of proof of the theorem for a trivial fibration. Subtracting the formulas from the previous steps gives, due to the cancellation of several terms,

$$\frac{\partial}{\partial t}c(\phi)-c(\phi)_{k\bar{k}}=\phi_{s\bar{m}\bar{k}}\overline{\phi_{s\bar{m}\bar{k}}}+(\phi_{s\bar{i}}\overline{\phi_{s\bar{j}}})\phi_{\bar{k}\bar{i}m}\phi_{k\bar{m}j}-2\Re(\overline{\phi_{\bar{s}k\bar{m}}}\phi_{s\bar{l}}\phi_{\bar{k}l\bar{m}})=\sum_{m,k}\left|\phi_{s\bar{m}\bar{k}}-\sum_{l}\phi_{s\bar{l}}\phi_{\bar{k}\bar{m}l}\right|^2.$$

Finally, note that

$$\phi_{s\bar{m}\bar{k}}-\sum_{l}\phi_{s\bar{l}}\phi_{\bar{k}\bar{m}l}=(\phi_{s\bar{m}})_{\bar{k}}-(\phi_{\bar{m}l})_{\bar{k}}\sum_{l}\phi_{s\bar{l}}=(\phi_{s\bar{l}}\phi^{m\bar{l}})_{\bar{k}}=(V_m)_{\bar{k}}$$

(using $\phi_{i\bar{j}}=\delta_{ij}$ at $z=0$), where $V=(V_1,\dots,V_n)$ is the $(0,1)$-vector field (3-3) expressed in local normal coordinates. This hence finishes the proof of the theorem in the case of a trivial fibration.

Step 4. We show that (3-7) holds for a general holomorphic submersion. As recalled above, any given point $P=(x,s_0)$ in $\mathscr{X}$ has a neighborhood $\mathscr{U}$ such that the fibration $\pi:\mathscr{U}\to\pi(\mathscr{U})$ (where we after shrinking S may assume that $\pi(\mathscr{U})=S$) is holomorphically trivial and the restriction $\mathscr{L}_{|\mathscr{U}}$ is isomorphic to π^*L over $\mathscr{U}$. We denote by (z,s) a choice of holomorphic coordinates on $\mathscr{U}$ trivializing the fibration. Moreover, when $\mathscr{X}\to S$ is a relative Calabi–Yau manifold we may furthermore choose (z,s) with the property that there is a family Ω_s of nowhere vanishing holomorphic n-forms on the fibers $\mathscr{X}_s$ such that the restriction of Ω_s to $\mathscr{U}_s$ $(:=\mathscr{U}\cap\mathscr{X}_s)$ coincides with the restriction of $dz:=dz_1\wedge\cdots\wedge dz_n$ to $\mathscr{U}_s$. Indeed, first observe that we may choose Ω_s so that $\Omega_s=f_s(z)\,dz$ on $\mathscr{U}$, where $f(z,s):=f_s(z)$ is holomorphic in (z,s) and invertible, with respect to any given holomorphic coordinates (z,s) as above. This amounts to the well-known fact that the direct image sheaf $\pi_*(K_{\chi/S})$ naturally defines a holomorphic line bundle on S or equivalently that any Ω_{s_0} may be extended to Ω_s such that $\Omega_s\wedge ds$ is a holomorphic $(n+1)$-form on $\mathscr{X}$ (which for example follows from the Ohsawa–Takegoshi extension theorem; see [Berndtsson 2009a] for a more general setting). We may now (after perhaps shrinking $\mathscr{U}$ again) write $f(z,s)=\partial g(z,s)/\partial z_1$ for some holomorphic functions g on $\mathscr{U}$ and define new holomorphic coordinates (ζ,s) on $\mathscr{U}$ (after perhaps again shrinking $\mathscr{U}$) by letting $\zeta_i:=g$ for $i=1$ and $\zeta_i:=z_i$ for $i>1$. By construction we then have $\Omega_{s|\mathscr{U}_s}=d\zeta_{|\mathscr{U}_s}$, as desired.

We can now repeat the previous local computation; the only new contribution comes from the derivatives on the local function $\psi_\Omega(s)$ defined by formula (3-2), which appear in the definition of the relative Kähler–Ricci flow (1-3) in the Calabi–Yau case. Indeed, locally this latter flow may be written as

$$\frac{\partial\phi}{\partial t}=\log\det(\phi_{k\bar{l}})-\psi_\Omega(s)$$

and the only new contribution to the previous calculations hence come from the term $-(\psi_\Omega(s))_{s\bar{s}}$ which appears in the calculation of $(\partial\phi/\partial t)_{s\bar{s}}$. Combining formulae (3-1), (3-2) hence proves that (3-7) holds locally on $\mathscr{X}$. Since all objects appearing in the formula are globally well-defined, this finishes the proof of Step 4. $\square$

Now the maximum principle for parabolic equations [Protter and Weinberger 1967] implies the following:

Corollary 3.4. *Let $\mathcal{L} \to \mathcal{X} \to S$ be a line bundle over a fibration as in the previous theorem.*

- *If the fibration is holomorphically trivial, then the function $c(t) := \inf_X c(\phi)$ is, for a fixed value on s, increasing along the relative Kähler–Ricci flow and hence the flow preserves (semi)positivity of the curvature of ϕ.*
- *For a holomorphically trivial fibration $\mathcal{X} = X \times S$, with $\mathcal{L}$ the pull-back of an ample line bundle $L \to X$, the flow* improves *the positivity of a generic initial weight in the following sense: if ϕ_0 is a semipositively curved weight on $\mathcal{L}$ over $X \times S$ such that $\partial\phi/\partial s$ does not vanish identically on $X \times \{s\}$ for any s, then ϕ_t is* strictly *positively curved on $X \times S$ for $t > 0$.*
- *In the general case the (semi)positivity of the curvature of the weight on $\phi - t\psi_\Omega$ on the $\mathbb{R}$-line bundle $\mathcal{L} - tK_{\mathcal{X}/S}$ is preserved under the flow; that is,*

$$dd^c\phi_t \geq -t\omega_{\mathrm{WP}}$$

for all t (and similarly in the strict case).

Proof. The first and third points follow from the maximum principle exactly as in the proof of Corollary 4.9 below. The second point is proved as follows: If strict positivity does not hold then one concludes (see the proof of Corollary 4.9 below) that $-A_{\phi_0} = \bar{\partial}_X V_{\phi_0}$ vanishes identically on X for some s_0; that is, the corresponding vector field V_{ϕ_0} defined by (3-3) is holomorphic on X. But, it is a well-known fact that any such holomorphic vector field $V^{1,0}$ vanishes identically when X is a Calabi–Yau manifold and hence $\partial\phi/\partial s$ vanishes identically on $X \times \{s_0\}$, giving a contradiction. The vanishing of $V^{1,0}$ may be proved as follows: by a Bochner–Weitzenbock formula $V^{1,0}$ is covariantly constant with respect to any Ricci flat metric on X. Moreover, the imaginary part V_I satisfies $\omega_{\phi_0}(V_I, \cdot) = df$ for some real smooth function f. But since $\omega^X_{\phi_0} > 0$ on $X \times \{s_0\}$ the latter equation forces the vanishing of V_I at any point where f achieves it maximum and hence $V_I \equiv 0$ on X. Similarly, the real part V_R of $V^{1,0}$ vanishes identically (by replacing df with $d^c f$). □

Of course, in the case of an infinitesimally nontrivial fibration the inequality in the previous corollary is useless for the limit ϕ_∞, but its interest lies in the fact that it gives a lower bound on the (possible) loss of positivity along the relative Kähler–Ricci flow, which is independent of the initial data.

Remark 3.5. Throughout the paper we assume, for simplicity, that the initial weight ϕ_0 has relatively positive curvature, when restricted to the fibers of the $\mathcal{X}$. But, as in the previous corollary, we do allow ϕ_0 to have merely semipositive curvature over the total space $\mathcal{X}$. However, using recent developments for the Kähler–Ricci flow [Song and Tian 2009] the relative Kähler–Ricci flows are actually well-defined for any smooth weight ϕ_0 which has merely relatively *semi*positive curvature and ϕ_t becomes relatively positively curved for any $t > 0$. Using this result the previous corollary can be seen to be valid for a general semipositively curved initial weight ϕ_0. Even more generally, as shown in [Song and Tian 2009], the flow is well-defined for any (possibly singular) ϕ_0 with positive curvature current such that ϕ_0 is locally bounded and the Monge–Ampère measure $(dd^c\phi_0)^n$ has local densities in L^p for some $p > 1$.

Evolution of the curvature of the top Deligne pairing. For a general smooth base S (i.e., not necessarily embedded in $\mathbb{C}$) the weight ϕ on L naturally induces a closed $(1,1)$-form $\Theta_\phi(s)$ on S expressed as

$$\Theta_\phi := \pi_*((dd^c\phi)^{n+1}/(n+1)!).$$

Equivalently, for any local holomorphic curve $C \subset S$ with tangent vector $\partial/\partial s \in TS$,

$$\Theta_\phi\left(\frac{\partial}{\partial s}, \frac{\partial}{\partial \bar{s}}\right) := \int_{X_s} c(\phi)\omega_\phi^n/n!$$

where $s \in C$ and π is the induced map $\pi : \mathscr{X} \to C$. Geometrically, the form Θ_ϕ on S may be described as the curvature of the Hermitian holomorphic line bundle $(\mathscr{L}, \phi)^{n+1}$ over S defined as the top *Deligne pairing* of the Hermitian holomorphic line bundle $(\mathscr{L}, \phi) \to \mathscr{X} \to S$ (see [Deligne 1987]; the relevance of Deligne pairings for Kähler geometry has been emphasized by Phong and Sturm [2004]). The form Θ_ϕ also appears as a multiple of the curvature of the Quillen metric on the determinant of the direct image of a certain virtual vector bundle over $\mathscr{X}$ (see [Fujiki and Schumacher 1990] and references therein).

Similarly, one can define a $(1,1)$-form $\omega_{\mathrm{WP}_\phi}$ on S depending on ϕ by letting

$$\omega_{\mathrm{WP}_\phi}\left(\frac{\partial}{\partial s}, \frac{\partial}{\partial \bar{s}}\right) := \int_{X_s} |A_\phi(s)|^2\omega_\phi^n/n!.$$

It can be checked that this yields a well-defined $(1,1)$-form on $\mathscr{X}$. Anyhow this latter fact is also a consequence of the following corollary of the previous theorem.

Corollary 3.6. *We make the same assumptions as in the previous theorem. Let Θ_{ϕ_t} be the curvature form on S of the top Deligne pairing of $(\mathscr{L}, \phi) \to \mathscr{X} \to S$, where ϕ_t evolves according to the relative Kähler–Ricci flow in the Calabi–Yau case. Then*

$$\frac{\partial}{\partial t}\Theta_\phi(s) = -\pi_*\big(R_{\omega_\phi^X}(dd^c\phi)^{n+1}/(n+1)!\big) + \omega_{\mathrm{WP}_{\phi_t}} - \omega_{\mathrm{WP}},$$

where $R_{\omega_\phi^X}$ denotes the fiberwise scalar curvature of the metric ω_ϕ.

Proof. We may without loss of generality assume that S is embedded in $\mathbb{C}_s$. Then

$$\frac{\partial}{\partial t}\int_{X_s} c(\phi)\omega_\phi^n/n! = \int_{X_s}\frac{\partial}{\partial t}c(\phi)\omega_\phi^n/n! + \int c(\phi)\frac{\omega_\phi^{n-1}}{(n-1)!}\wedge dd^c\frac{\partial}{\partial t}\phi.$$

Now, by the definition of the Kähler–Ricci flow in the Calabi–Yau case,

$$\frac{\omega_\phi^{n-1}}{(n-1)!}\wedge dd^c\frac{\partial}{\partial t}\phi = \frac{\omega_\phi^{n-1}}{(n-1)!}\wedge(-\mathrm{Ric}(\omega_\phi^X)) =: -R_{\omega_\phi^X}\omega_\phi^n/(n+1)!,$$

where we have used the definition of the (normalized) scalar curvature $R_{\omega_\phi^X}$ of the Kähler metric ω_ϕ^X in the last step. Finally, integrating the formula in the previous theorem finishes the proof of the corollary. □

Remark 3.7. If the initial weight ϕ for the Kähler–Ricci flow is taken so that ω_ϕ restricts to a Ricci flat metric on all fibers of $\mathscr{X}$, then ϕ is stationary for the Kähler–Ricci flow and hence the previous corollary (and the proof of the previous theorem) shows that

$$\int_{\chi_s} |A_\phi(s)|^2 \omega_\phi^n / n = \frac{\partial^2 \psi_\Omega}{\partial s \partial \bar{s}};$$

that is, $\omega_{\mathrm{WP}} = d_s d_s^c i^{n^2} \log \int_{\mathscr{X}_s} \Omega \wedge \bar{\Omega}$. Since, by Proposition 4.5 below, $A_\phi(s)$ is harmonic on each fiber $\mathscr{X}_s$ with respect to the Ricci flat restriction ω_ϕ, this implies the equivalence between (3-1) and (3-2).

3B. *Quantization: The Bergman iteration on $\mathscr{H}_L$.* In this section we will specialize and develop the general results in Section 2D to the present setting where we have fixed a family of probability measures μ_s (independent of ϕ) on the fibers χ_s.

Convergence and positivity of the Bergman iteration at a fixed level k. The following monotonicity properties were shown by Donaldson [2009] in the present setting.

Lemma 3.8. *The functionals $-I_\mu$ and $\mathscr{L}^{(k)}$ are both increasing along the Bergman iteration on $\mathscr{H}_L$ with respect to μ. Moreover, they are strictly increasing at $\phi_m^{(k)}$ unless $\phi_m^{(k)}$ is stationary.*

Proof. Since I_μ is affine and in particular concave on the affine space of all smooth weights the lemma follows immediately from Lemma 2.6. □

We can now prove the convergence of the Bergman iteration at a fixed level k in the present setting.

Theorem 3.9. *Let $L \to X$ be an ample line bundle and μ a fixed volume form on X giving unit volume to X. Assume a smooth initial weight ϕ_0 is given. For any given positive integer k the following holds: in the large time limit, that is, when $m \to \infty$, the weights $\phi_m^{(k)}$ converge to $\phi_\infty^{(k)}$ in the $\mathscr{C}^\infty$-topology on X. Moreover, in the relative setting the convergence is locally uniform with respect to the base parameter s.*

Proof. By the previous lemma $-I_\mu$ is increasing and by definition $-\mathscr{F}_\mu^{(k)}$ is coercive. Moreover, as shown in [Berman et al. 2009] balanced weights are unique modulo scaling and hence all the convergence criteria in Proposition 2.9 are satisfied. □

Conservation of positivity. Recall that, given a relatively ample line bundle $\mathscr{L}$ over a fibration $\pi : \mathscr{X} \to S$ as above, the corresponding direct image bundle $\pi_*(\mathscr{L} + K_{\mathscr{X}/S}) \to S$ is the vector bundle such that the fiber over s is naturally identified with the space $H^0(X, L + K_X)$ of all holomorphic n-forms f on $X := \mathscr{X}_s$ with values in $L := \mathscr{L}_X$ (as is well-known this is indeed a vector bundle, as shown using vanishing theorems). Moreover, any given weight ϕ on $\mathscr{L}$ induces a Hermitian metric on $\pi_*(k\mathscr{L} + K_{\mathscr{X}/S})$ whose fiberwise restriction will be denoted by $\mathrm{Hilb}_{L+K_X}(\phi)$:

$$\mathrm{Hilb}_{L+K_X}(\phi)(f, f) := i^{n^2} \int_X f \wedge \overline{f} e^{-\phi}.$$

The point is that there is no need to specify an integration measure μ thanks to the twist by the relative canonical line bundle $K_{\mathscr{X}/S}$. We will have great use for the following recent results of Berndtsson.

Theorem 3.10. *Let $\pi : \mathscr{X} \to S$ be a proper holomorphic submersion and let $\mathscr{L}$ be a relatively ample line bundle over $\mathscr{X}$ equipped with a smooth weight with semipositive curvature. Then*:

- [Berndtsson 2009a] *The curvature of the Hermitian vector bundle over S defined as the direct image bundle $\pi_*(\mathscr{L} + K_{\mathscr{X}/S})$ is semipositive in the sense of Nakano (and in particular in the sense of Griffiths).*
- (See [Berndtsson 2011, Theorem 1.2 and subsequent discussion].) *The vector bundle $\pi_*(\mathscr{L} + K_{\mathscr{X}/S})$ has* strictly *positive curvature in the sense of Griffiths if* either *the curvature form of ϕ is strictly positive over all of $\mathscr{X}$* or *strictly positive along the fibers of $\pi : \mathscr{X} \to S$ and the fibration is infinitesimally nontrivial (i.e., the Kodaira–Spencer classes are nontrivial for all $s \in S$).*

We will only use the following simple consequence of Theorem 3.10 (compare [Berndtsson 2009a; Berndtsson and Păun 2008b]):

Corollary 3.11. *Under the assumptions in the first point of the previous theorem we have*

$$dd^c(\mathrm{FS}^{(k)} \circ \mathrm{Hilb}_{k\mathscr{L}+K_{\mathscr{X}/S}})(\phi) \geq 0 \tag{3-9}$$

and the inequality is strict under the assumptions in the second point of the theorem.

Proof. We will denote the line bundle $k\mathscr{L} + K_{\mathscr{X}/S}$ over $\mathscr{X}$ by $\mathscr{F}$ and the vector bundle $\pi_*(\mathscr{F})$ over S by E (and its dual by E^*). First note that the weight on $\mathscr{F}$ that we are interested in may be written as

$$(\mathrm{FS} \circ \mathrm{Hilb}_{\mathscr{F}})(s, x_s) = \log \sup_{f_s \in E_s} \frac{|f_s(x_s)|^2}{\|f(x_s)\|^2} = \log |\Lambda_{(s,x_s)}|^2, \tag{3-10}$$

where $\Lambda_{(s,x_s)}$ is the element in $E_s^* \otimes \mathscr{F}_s$ defined by

$$(\Lambda_{(s,x_s)} f_s) := f_s(x_s).$$

Now let $t \mapsto (s_t, x_{s_t})$ be a local holomorphic curve in $\mathscr{X}$ with $t \in \Delta$ (the unit-disc). Trivializing $\mathscr{F}$ in a neighborhood of the previous curve we may pull back $\Lambda_{(s,x_s)}$ to a holomorphic section Λ_t of E^* over the unit-disc and identify the weight defined by (3-10) with a function $\log |\Lambda_t|^2$ on Δ. We have to prove that this latter function is (strictly) psh. But this follows from the following well-known fact: a vector bundle $E \to \Delta$ is (strictly) positive in the sense of Griffiths if and only if $\log(\|\Lambda_t\|^2)$ is (strictly) subharmonic on Δ where Λ is any nontrivial holomorphic section of the dual vector bundle E^*. For example, to get the required (strict) subharmonicity one just notes that, after a standard computation,

$$\frac{\partial^2 \log(\|\Lambda_t\|^2)}{\partial t \, \partial \bar{t}} \Big|_{t=0} \geq -\frac{\Theta_{E^*}(\Lambda_0, \Lambda_0)}{\|\Lambda_0\|^2},$$

where Θ_{E^*} at t is the Hermitian endomorphism of E_t^* representing the curvature of E. By the previous theorem Θ_E is (strictly) positive which is equivalent to Θ_{E*} being (strictly) negative and the corollary hence follows from the previous inequality. ☐

We next obtain a "quantized" version of Corollary 3.4.

Corollary 3.12. *Let $\pi : \mathcal{X} \to S$ be a proper holomorphic submersion with Calabi–Yau fibers and let $\mathcal{L}$ be a relatively ample line bundle over $\mathcal{X}$.*

- *When π is holomorphically trivial the relative Bergman iteration preserves semipositivity of the curvature of ϕ.*
- *In the case of a general submersion with Calabi–Yau fibers,*

$$dd^c\phi^{(k)}_{(m)} \geq -\frac{m}{k}\omega_{\mathrm{WP}}$$

for all m.

Proof. For simplicity first consider the case of a trivial fibration. Fix a holomorphic n-form Ω on $X := \mathcal{X}_0$ trivializing K_X. Under the assumption that $\mathcal{X} \to S$ is holomorphically trivial Ω extends to a holomorphic n-form on all of $\mathcal{X}$ such that $\psi_\Omega := \log \int_{\mathcal{X}_s} i^{n^2}\Omega \wedge \bar{\Omega}$ is independent of s. In this notation

$$\mathrm{Hilb}^{(k)}(\phi(s,\cdot))(f,f) := \int_{\mathcal{X}_s} |f|^2 e^{-(k\phi(s,\cdot)-\psi_\Omega(s))} i^{n^2}\Omega \wedge \bar{\Omega}.$$

Now consider the fiberwise isomorphism

$$j : H^0(X, kL) \to H^0(X, kL + K_X), \quad j(f) = f \otimes \Omega,$$

which clearly satisfies $\mathrm{Hilb}^{(k)}(\phi(s,\cdot)) = e^{\psi_\Omega} j^* \mathrm{Hilb}_{L+K_X}(\phi(s,\cdot))$. This means that, up to a multiplicative constant independent of s, the map j is an *isometry* when $H^0(X_s, kL + K_{X_s}) = H^0(\mathcal{X}, k\mathcal{L} + K_{\mathcal{X}/S})_{\mathcal{X}_s}$ is equipped with its natural Hermitian product. In particular, by (3-9),

$$dd^c\phi \geq 0 \implies dd^c(\mathrm{FS}^{(k)} \circ \mathrm{Hilb}^{(k)})(\phi) \geq 0.$$

Iterating hence proves the first point in the statement of the corollary. Finally, for a general submersion the same argument gives, but now taking into account the fact that ψ_Ω depends on s, that

$$dd^c\phi \geq 0 \implies dd^c(\mathrm{FS}^{(k)} \circ \mathrm{Hilb}^{(k)})(\phi) \geq -dd^c\psi_\Omega(s)/k = -\omega_{\mathrm{WP}}(s)/k,$$

using formula (3-2) in the last equality. Replacing ϕ with $\mathrm{FS}^{(k)} \circ \mathrm{Hilb}^{(k)} - \psi_\Omega(s)$ and iterating hence finishes the proof of the corollary. □

Convergence towards the Kähler–Ricci flow. The following very simple proposition will turn out to be very useful:

Proposition 3.13. *The following monotonicity holds for the Bergman iteration at level k (with respect to μ). Assume that $\phi^{(k)}_m \leq \psi^{(k)}_m$. Then $\phi^{(k)}_{m+1} \leq \psi^{(k)}_{m+1}$. In particular, the Bergman iteration decreases the distance in $\mathcal{H}_L$ defined with respect to the sup norm $d(\phi,\psi) := \sup_X |\phi - \psi|$.*

Proof. By definition we have

$$\phi^{(k)}_{m+1} = \phi^{(k)}_m + \frac{1}{k}\log \rho^{(k)}(k\phi^{(k)}_m) = \frac{1}{N_k}\sum_i |f_i|^2.$$

By a well-known identity for Bergman kernels,

$$\sum_{i=1} |f_i|^2(x) = \sup_{f\in H^0(X,kL)} \left(|f(x)|^2 \Big/ \int_X |f|^2 e^{-k\phi_m}\, d\mu \right).$$

But this latter expression is clearly monotone in ϕ_m proving the first statement of the proposition. As for the last statement just let $C := \sup_X |\phi_m^{(k)} - \psi_m^{(k)}|$ so that

$$\phi_m^{(k)} \leq \psi_m^{(k)} + C, \quad \psi_m^{(k)} \leq \phi_m^{(k)} + C.$$

Applying the first statement of the proposition finishes the proof. □

Remark 3.14. The previous proposition can be seen as a "quantum" analog of the corresponding result for the Kähler–Ricci flow (1-3), which follows directly from the maximum principle for the Monge–Ampère operator and its parabolic analogue.

Now we can prove the following theorem, which is one of the main results in this paper.

Theorem 3.15. *Let $L \to X$ be an ample line bundle and μ a volume form on X giving unit volume to X. Fix a smooth weight ϕ_0 on L, whose curvature form is fiberwise strictly positive, and consider the corresponding Bergman iteration $\phi_m^{(k)}$ at level k and discrete time m, as well as the Kähler Ricci flow ϕ_t — both defined with respect to μ. Then there is a constant C such that*

$$\sup_X |\phi_m^{(k)} - \phi_{m/k}| \leq Cm/k^2.$$

In particular, if m_k is a sequence such that $m_k/k \to t$, then

$$\phi_{m_k}^{(k)} \to \phi_t$$

uniformly on X. Moreover, in the relative setting C is locally bounded in the base parameter s if μ depends smoothly on s.

Proof. Write $\psi_{k,m} = \phi_{m/k}$ and $F^{(k)}(\psi) = \frac{1}{k}\log\rho^{(k)}(\psi)$.

Step 1. We have $\psi_{k,m+1} - \psi_{k,m} = F^{(k)}(\psi_{k,m}) + O(1/k^2)$ for all (k,m), where the error term is uniform in (k,m). (In the following we will take that as a definition of $O(1/k)$, etc.)

To prove this we write the left-hand side as

$$\frac{1}{k}\left(\frac{\phi_{m/k+1/k} - \phi_{m/k}}{1/k}\right) = \frac{1}{k}\left(\frac{\partial\phi_t}{\partial t}\Big|_{t=m/k} + O(1/k)\right)$$

using that $\left|\partial^2\phi_t/\partial^2 t\right| \leq C$ on $X\times[0,T]$ by Theorem 3.1. More precisely, by the mean value theorem the error term $O(1/k)$ may be written as

$$\frac{1}{k}\frac{\partial^2\phi_t}{\partial^2 t}(\zeta)/2$$

for some $\zeta \in [0, 1/k]$.

Since ϕ_t evolves according to the Kähler–Ricci flow this means that

$$\psi_{k,m+1}-\psi_{k,m}=\frac{1}{k}\log\left(\frac{(dd^c\phi_{m/k})^n/n!}{\mu}\right)+O(1/k^2).$$

But by Proposition 2.11 we have that

$$F^{(k)}(\phi_{m/k})=\frac{1}{k}\log\left(\frac{(dd^c\phi_{m/k})^n/n!}{\mu}\right)+O(1/k^2),$$

where the error term is uniformly bounded in (m,k) for $m/k\leq T$ by Theorem 3.1. In fact, as is well-known the uniform estimates (2-4) on the "space-derivatives" of ϕ_t in Theorem 3.1 also hold for all time-derivatives $d^r\phi_t/d^rt$ (and in particular for $r=1$ and $r=2$ used above). This is well-known and shown by differentiating the flow equation with respect to time and applying the maximum principle repeatedly. Hence, T may be taken to be equal to infinity, which finishes the proof of Step 1.

Step 2. Given Step 1 and the fact that the Bergman iteration decreases the sup norm, we have

$$\sup_X|\phi_m^{(k)}-\psi_{k,m}|\leq Cm/k^2. \tag{3-11}$$

We will prove this by induction over m (for k fixed), the statement being trivially true for $m=0$. By Step 1 there is a uniform constant C such that

$$\sup_X\left|\psi_{k,m+1}-(\psi_{k,m}+F^{(k)}(\psi_{k,m}))\right|\leq C(1/k^2)$$

for all (m,k). Now we fix the integer k and assume as an induction hypothesis that (3-11) holds for m with C the constant in the previous inequality. By Proposition 3.13,

$$\sup_X\left|(\psi_{k,m}+F^{(k)}(\psi_{k,m}))-(\phi_m^{(k)}+F^{(k)}(\phi_m^{(k)}))\right|\leq\sup_X|\psi_{k,m}-\phi_m^{(k)}|\leq Cm/k^2$$

with the same constant C as above, using the induction hypothesis in the last step. Combining this estimate with the previous inequality gives

$$\sup_X|\psi_{k,m+1}-\phi_{m+1}^{(k)}|\leq Cm/k^2+C/k^2,$$

proving the induction step and hence Step 2. □

Of course, it seems natural to expect that $\mathscr{C}^\infty$-convergence holds but we leave this problem for the future.

Combining the previous corollary with Theorem 3.15 and the variational principle in [Berman et al. 2009] (the $\mathscr{C}^\infty$-convergence rather uses [Keller 2009; Wang 2005]) now gives the following:

Corollary 3.16. *The conservation of semipositivity of the curvature of ϕ_t in Corollary 3.4 holds. For a fixed initial data $\phi_0=\phi_0^{(k)}\in\mathscr{H}_L$ the following convergence results hold for the Bergman iteration $\phi_m^{(k)}$:*

- *For any sequence m_k such that $m_k/k\to\infty$ the convergence $\phi_{m_k}^{(k)}\to\phi_\infty$ holds in the L^1-topology on X. Moreover, if it is also assumed that $m_k/k^2\to 0$ then the convergence holds in the C^0-topology.*

- *The balanced weights* $\phi_\infty^{(k)} := \lim_{m\to\infty}\phi_m^{(k)}$ *at level* k *converge, when* $k\to\infty$, *in the* $\mathscr{C}^\infty$*-topology, to the weight* ϕ_∞ *which is the large time limit of the corresponding Kähler–Ricci flow* (*and in particular a solution to the corresponding inhomogeneous Monge–Ampère equation*).

In the relative case the convergence holds fiberwise locally uniformly with respect to the base parameter s.

Proof. The first statement follows immediately by combining Theorem 3.15 and the previous corollary, since semipositivity is preserved under uniform limits of weights. Hence, we turn to the proof of the first point. It is based on the following inequalities:

$$\limsup_{k\to\infty} I_\mu(\phi_{m_k}^{(k)}) \le I_\mu(\phi_\infty), \quad \liminf_{k\to\infty} \mathscr{E}(\phi_{m_k}^{(k)}) \ge \mathscr{E}(\phi_\infty). \tag{3-12}$$

To prove these inequalities take a sequence m'_k such that $m'_k/k\to t$ and $m'_k\le m_k$. By monotonicity (Lemma 3.8),

$$I_\mu(\phi_{m_k}^{(k)}) \le I_\mu(\phi_{m'_k}^{(k)}).$$

Hence, letting $k\to\infty$ and using that $\phi_{m'_k}^{(k)}\to\phi_t$ uniformly (by Theorem 3.15) gives

$$\limsup_{k\to\infty} I_\mu(\phi_{m_k}^{(k)}) \le I_\mu(\phi_t).$$

Finally, letting $t\to\infty$ and using Theorem 3.1 proves the first inequality in (3-12). As for the second inequality in (3-12), it is similarly proved by noting that, by monotonicity,

$$\mathscr{L}^{(k)}\big(\phi_{m'_k}^{(k)}\big) \le \mathscr{L}^{(k)}(\phi_{m_k}^{(k)}).$$

To proceed we will use that $\psi_k\to\psi$ uniformly in $\mathscr{H}_L$ implies that

$$\mathscr{L}^{(k)}(\psi_k)\to\mathscr{E}(\psi).$$

To see this recall that this is well-known when $\psi_k=\psi$ for all k (as follows for example from Proposition 2.11, saying that the convergence holds for the differentials $d\mathscr{L}^{(k)}$ and $d\mathscr{E}$; for more general convergence results see [Berman and Boucksom 2010]). But then the general case follows easily from the fact that $\mathscr{L}^{(k)}$ is monotone in the argument ψ and scaling equivariant. Hence, letting $k\to\infty$ gives, since $\psi_k := \phi_{m'_k}^{(k)}\to\phi_t$ uniformly, that

$$\mathscr{E}(\phi_t) \le \liminf_{k\to\infty} \mathscr{L}^{(k)}(\phi_{m_k}^{(k)}).$$

The proof of the second inequality in (3-12) is finished by using that (as shown in [Berman et al. 2009]), for any sequence (ψ_k) in $\mathscr{H}_L$,

$$\limsup_{k\to\infty} \mathscr{L}^{(k)}(\psi_k) \le \liminf_{k\to\infty} \mathscr{E}(\psi_k).$$

Now, adding up the two inequalities in (3-12) gives

$$\liminf_{k\to\infty} \mathscr{F}(\phi_{m_k}^{(k)}) \ge \mathscr{F}(\phi_\infty).$$

But then it follows from the variational results in [Berman et al. 2009] that

$$\lim_{k\to\infty} \mathcal{F}(\phi^{(k)}_{m_k}) = \mathcal{F}(\phi_\infty) \tag{3-13}$$

and

$$dd^c\phi^{(k)}_{m_k} \to dd^c\phi_\infty \tag{3-14}$$

in the weak topology of currents. Next, note that by the inequalities (3-12) the sequence $\phi^{(k)}_{m_k}$ is contained in a compact subset of $\mathcal{H}_L$ equipped with the L^1-topology (compare the proof of Lemma 2.3) and hence we may assume (perhaps after passing to a subsequence) that $\phi^{(k)}_{m_k} \to \psi$ in the L^1-topology. But then the convergence in (3-14) forces $\psi = \phi_\infty + C$ for some constant C. Hence, it will be enough to prove that $C = 0$. To this end, note that combining (3-13) and the inequalities (3-12) shows that the latter inequalities are in fact equalities. In particular,

$$\lim_{k\to\infty} \mathcal{E}(\phi^{(k)}_{m_k}) = \mathcal{E}(\phi_\infty).$$

By the scaling equivariance of $\mathcal{E}$ it hence follows that $C = 0$, which finishes the proof of the first point. If one assumes that $m_k/k^2 \to 0$ then it follows immediately from combining Theorem 3.1 and Theorem 3.15 that the convergence holds uniformly on X, that is, in the C^0-topology.

To prove the second point in the statement of the corollary note that replacing $\phi^{(k)}_{m_k}$ by $\phi^{(k)}_\infty$ in the previous argument gives, just as before, that $\phi^{(k)}_\infty \to \phi_\infty$ in the L^1-topology. Moreover, since it was shown in [Keller 2009; Wang 2005] that the convergence of the corresponding curvature forms holds in the $\mathcal{C}^\infty$-topology this proves the second point. ☐

4. The (anti)canonical setting

In this section we will consider another particular case of the general setting in Section 2 arising when the line bundle $L := rK_X$ is ample, where $r = 1$ or $r = -1$ (for any fiber X of the fibration). Hence, X is necessarily of general type in the former "positive" case and a Fano manifold in the latter "negative" setting. We will also refer to these two different settings as the $\pm K_X$-settings.

By the very definition of the canonical line bundle any weight ϕ on $\pm K_X$ determines a canonical scale-invariant probability measure $\mu_\pm(\phi)$ on X, where

$$\mu_\pm(\phi) := e^{\pm\phi} \Big/ \int_X e^{\pm\phi}$$

(with a slight abuse of notation), so that $\mu_\pm(\phi + c) = \mu_\pm(\phi)$. Equivalently, $\mu_\pm(\phi)$ may be identified with the one-form on $\mathcal{H}_{\pm K_X}$ obtained as the differential of the following functional $I_\pm(\phi)$ on $\mathcal{H}_{\pm K_X}$:

$$I_\pm(\phi) := \pm \log \int_X e^{\pm\phi}, \quad \mu_\pm(\phi) = dI_\pm.$$

A characteristic feature of the $\pm K_X$-setting is that the antiderivative $I_\pm$ is canonically defined (i.e., not only up to scaling). As a consequence there is a canonical normalization condition for weights that will occasionally be used below, namely the condition that $I_\pm(\phi) = 0$.

We will also have use, as before, for the equivariant functional

$$\mathscr{F}_{\pm} := \mathscr{E} - I_{\pm},$$

where $\mathscr{E}$ is the functional defined in Section 2B (with respect to a fixed reference weight in $\pm K_X$).[1] Note that the critical points of $\mathscr{F}_{\pm}$ on $\mathscr{H}_{\pm K_X}$ are the *Kähler–Einstein weights* ϕ, that is, the weights such that ω_ϕ is a Kähler–Einstein metric on X (compare Theorem 4.1 below).

It will also be important to consider a *nonnormalized* variant of $\mu_{\pm}(\phi)$ defined by

$$\mu'_{\pm}(\phi) := e^{\pm\phi}$$

(which is the differential of the *non*equivariant functional $\phi \mapsto \int e^{\pm\phi}$). In the sequel we will refer to the two different settings defined by $\mu_{\pm}(\phi)$ and $\mu'_{\pm}(\phi)$ as the *normalized* $\pm K_X$*-setting* and the *nonnormalized* $\pm K_X$*-setting*, respectively. It should be pointed out that it is the latter one which usually appears in the literature on the Kähler–Ricci flow (see for example [Cao 1985; Tian and Zhu 2007; Phong et al. 2007]).

4A. ***The relative Kähler–Ricci flow.*** According to the general construction in Section 2 each particular setting introduced above comes with an associated relative Kähler–Ricci flow. For future reference we will write out the fiberwise flow in the nonnormalized $\pm K_X$-setting in local holomorphic coordinates:

$$\frac{\partial\phi}{\partial t} = \log\det\left(\frac{1}{\pi}\frac{\partial^2\phi}{\partial z_i \partial \bar{z_j}}\right)/n! - (\pm\phi). \tag{4-1}$$

The normalized and nonnormalized settings induce the same evolution of the fiberwise curvature forms ω_t:

$$\frac{\partial\omega_t}{\partial t} = -\mathrm{Ric}\,\omega_t - \pm\omega_t, \tag{4-2}$$

in $c_1(\pm K_X)$.[2]

In particular, if ω_t converges to ω_∞ in the large time limit, then ω_∞ is necessarily a Kähler–Einstein metric, which is of *negative* scalar curvature in the K_X-setting and *positive* scalar curvature in the $-K_X$-setting.

The main virtue of the Kähler–Ricci flow in the normalized setting as compared with the nonnormalized one is that the first one is convergent precisely when the flow of curvature forms ω_t is. On the other hand, as will be seen later the flow in the nonnormalized setting (and its quantized version) has better monotonicity and positivity properties.

Theorem 4.1. *The Kähler–Ricci flow in the* $\pm K_X$*-settings always exists and is smooth on* $X \times [0, \infty[$. *More precisely, all the analytical assumptions in Section 2C are satisfied. In the normalized* K_X*-setting it converges to a Kähler–Einstein metric of negative scalar curvature. In the* $-K_X$*-setting the flow converges to a Kähler–Einstein metric of positive scalar curvature under the assumptions that* $H^0(TX) = 0$ *and* X

[1] Note that $\mathscr{F}_{\pm}$ is minus the functional introduced in [Tian 2000].

[2] In the literature this latter flow of Kähler forms is sometimes referred to as the normalized Kähler–Ricci flow, as opposed to Hamilton's original flow, but our use of the term "normalized" is different and only applies on the level of weights on L.

a priori admits a Kähler–Einstein metric. Furthermore, in the relative case the convergence is locally uniform with respect to the base parameter s.

Apart from the uniqueness statement, the first part of the previous theorem is due to Cao [1985]. The convergence on the level of Kähler metrics in the Fano case, that is, when $-K_X$ is ample, was proved by Perelman (unpublished) and Tian and Zhu [2007]. The convergence on the level of weights then follows directly from Proposition 2.4 and the known coercivity of the functionals $-\mathcal{F}_\pm$; the coercivity of $-\mathcal{F}_+$ follows immediately from Jensen's inequality, while the coercivity of $-\mathcal{F}_-$ was shown in [Phong et al. 2008], confirming a conjecture of Tian. The uniqueness in the difficult case of $-K_X$ is due to Bando and Mabuchi (for a comparatively simple proof see [Berman et al. 2009]).

Remark 4.2. The first key analytical ingredient in the proof of the convergence of the flow of Kähler metric ω_t in the Fano case (i.e., the $-K_X$-setting) is an estimate of Perelman saying that the Ricci potential h_t of ω_t, when suitably normalized, is always bounded along the Kähler–Ricci flow for ω_t (see [Tian and Zhu 2007; Phong et al. 2007]). In fact, in the present notation h_t coincides (modulo signs) with the time derivative of ϕ_t evolving according to the *normalized* Kähler–Ricci flow in the $-K_X$-setting. The second key ingredient is the fact that the existence of a Kähler–Einstein metric implies that $-\mathcal{F}_+$ is proper (and conversely [Tian 2000; Phong et al. 2008]). As is well-known there are, in general, obstructions to existence of Kähler–Einstein metrics in the $-K_X$-setting. According to a conjecture of Yau the existence of a Kähler–Einstein metric should be equivalent to a suitable notion of algebraic stability (in the sense of geometric invariant theory). From this point of view the properness (or coercivity) assumption on the functional $-\mathcal{F}_+$ can be considered as an *analytic* stability [Tian 2000].

Definition. A weight ϕ_{KE} on $\pm K_X$ will be called a *normalized Kähler–Einstein weight* if $I_\pm(\phi_{\mathrm{KE}}) = 0$, or equivalently if $e^{\pm\phi_{\mathrm{KE}}} = \omega_{\mathrm{KE}}^n/n!$.

Hence, there is precisely one normalized Kähler–Einstein weight on $+K_X$ when it is ample. The following simple corollary of Theorem 4.1 and Remark 4.2 illustrates the difference between the normalized and nonnormalized settings.

Corollary 4.3. *In the $+K_X$-setting the nonnormalized flow* (4-1) *always converges to the normalized Kähler–Einstein weight.*

Proof. Write ϕ'_t for the evolution under the Kähler–Ricci flow in the nonnormalized K_X-setting so that

$$\phi'_t = \phi_t + C_t,$$

where C_t is a constant for each t. Since $\phi \mapsto (dd^c\phi)^n$ is invariant under scalings, comparing the two flow equations gives

$$\frac{\partial C_t}{\partial t} = -C_t - I_+(\phi_t). \tag{4-3}$$

Let $D_t := C_t - I(\phi_t)$. Then we get

$$\frac{\partial D_t}{\partial t} = -D_t + \epsilon_t, \quad \text{where } \epsilon_t := \frac{\partial I_+(\phi_t)}{\partial t}.$$

In the $+K_X$-setting Theorem 4.1 implies that $\epsilon_t \to 0$; it follows for elementary reasons that $D_t \to 0$. Indeed, assume for a contradiction that D_t does not converge to 0. Then $\partial \log |D_t|/\partial t \to -1$; that is, $|D_t| \leq C_\delta e^{-t(1-\delta)} \to 0$ for $0 < \delta \ll 1$, giving a contradiction. Finally, in the nonnormalized $-K_X$-setting it was shown in [Phong et al. 2007] (building on [Chen and Tian 2002]) that there is a constant c_0 such that ϕ'_t converges. But then it follows immediately from combining the scaling invariance of $\phi \mapsto (dd^c\phi)^n$ and the scaling equivariance of μ'_- that the flow diverges exponentially for any other choice of constant c_0. □

Remark 4.4. In the nonnormalized $-K_X$-setting (under the assumptions in the previous theorem) it was shown in [Phong et al. 2007] (building on [Chen and Tian 2002]) that the flow converges when the initial weight ϕ_0 is replaced by $\phi_0 + c_0$ for a unique constant c_0. The argument in the proof of the previous corollary then gives that for a *generic* initial weight the flow is divergent.

4B. *Weil–Petersson geometry.* As before we may in the following assume that the base S is embedded in $\mathbb{C}$. In the relative $\pm K_X$-setting the (generalized) Weil–Petersson form ω_{WP} on S was introduced in [Koiso 1983] (see also [Fujiki and Schumacher 1990] for generalizations):

$$\omega_{\mathrm{WP}}\left(\frac{\partial}{\partial s}, \frac{\partial}{\partial \bar{s}}\right) := \|A_{\mathrm{KE}}\|^2_{\omega_{\mathrm{KE}}}, \tag{4-4}$$

where A_{KE} denotes the unique representative in the Kodaira–Spencer class $\rho(\frac{\partial}{\partial s}) \in H^{0,1}(T^{1,0}\mathscr{X}_s)$ which is harmonic with respect to the Kähler–Einstein metric on $\mathscr{X}_s$ and the L^2-norm is computed with respect to this latter metric. In fact, as shown in [Fujiki and Schumacher 1990, Proposition 4.12], $A_{\mathrm{KE}} = -\bar{\partial} V_{\omega_s}$, where V_{ω_s} is the local vector field defined by formula (3-3). This is a consequence of the following proposition proved in [Fujiki and Schumacher 1990].

Proposition 4.5. *Let $\pi : \mathscr{X} \to S$ be a proper holomorphic submersion and ω_s a smooth family of 2-forms on the fibers $\mathscr{X}_s$ such that ω_s is Kähler–Einstein on $\mathscr{X}_s$. Then A_{ω_s} is the unique element in $H^{0,1}(T\mathscr{X}_s)$ which is harmonic with respect to ω_s.*

Note that "harmonic" lifts of vector fields were previously used by Siu [1986] in the context of Weil–Petersson geometry.

Remark 4.6. When the relative dimension is one the space $H^{0,1}(T\mathscr{X}_s)$ is isomorphic to $H^{1,0}((T\mathscr{X}_s)^*) = H^0(2K_{\chi_s})$ under Serre duality. Hence, the Weil–Petersson form as defined in terms of harmonic representatives then coincides with the metric on $\mathscr{X}$ introduced by Weil in the case when $\mathscr{X}$ is the universal family over Teichmüller space. As conjectured by Weil and subsequently proved by Ahlfors this latter $(1,1)$-form is *closed* and hence Kähler. In the higher-dimensional case, it was observed in [Fujiki and Schumacher 1990] that the Kähler property of ω_{WP} as defined by (4-4) follows immediately from Corollary 4.10 below.

By an application of the implicit function theorem (in appropriate Banach spaces) the smoothness of the family ω_s (and of the associated normalized weight) in the previous proposition is automatic in the $+K_X$-case case, as well as in the $-K_X$ case if there are no nontrivial holomorphic vector fields tangential to the fibers of the fibration (see Theorem 6.3 in [Fujiki and Schumacher 1990]).

Now we can prove the following variant of Theorem 3.3.

Theorem 4.7. *Let $\pi : \mathscr{X} \to S$ be a proper holomorphic submersion. Assume that $\pm K_{\mathscr{X}/S}$ is relatively ample and that ϕ_t evolves according to the Kähler–Ricci flow in the nonnormalized setting. Then*

$$\frac{\partial c(\phi_t)}{\partial t} = \Delta_{\omega_t^X} c(\phi_t) - \pm c(\phi_t) + |A_{\phi_t}|^2_{\omega_t^X}. \tag{4-5}$$

In particular, if ϕ_{KE} is a fiberwise normalized Kähler–Einstein weight, then

$$\Delta_{\omega_{\mathrm{KE}}} c(\phi_{\mathrm{KE}}) - \pm c(\phi_{\mathrm{KE}}) + |A_{\omega_{\mathrm{KE}}}|^2_{\omega^X_{\mathrm{KE}}} = 0.$$

Proof. To simplify the notation we will only consider the $+K_X$-setting, but the proof in the $-K_X$ setting is essentially the same. We will just indicate the simple modifications of the proof of Theorem 3.3 which arise in the present setting.

Let us first consider the modifications to the calculation of the t-derivative of $c(\phi)$ that arise from the additional term $-\phi$ appearing in the calculation of the time derivative ϕ_t, since now

$$\frac{\partial}{\partial t}\phi_t = \log\det(\phi_{k\bar{l}}) - \phi$$

in local coordinates. To this end we assume to simplify the notation that X is one-dimensional (but the general argument is essentially the same). First recall that, according to formula (3-8),

$$\frac{\partial}{\partial t} c(\phi) = \frac{\partial}{\partial t}\phi_{s\bar{s}} - \left[(\phi_{s\bar{z}}\overline{\phi_{s\bar{z}}})_t \phi_{z\bar{z}}^{-1} - (\phi_{s\bar{z}}\overline{\phi_{s\bar{z}}})\phi_{z\bar{z}}^{-2}\frac{\partial}{\partial t}\phi_{z\bar{z}} \right].$$

Hence, the additional contribution referred to above is of the form

$$B := (-\phi)_{s\bar{s}} - 2\Re(-\phi_{s\bar{z}}\overline{\phi_{s\bar{z}}})\phi_{z\bar{z}}^{-1} + \phi_{s\bar{z}}\overline{\phi_{s\bar{z}}}\phi_{z\bar{z}}^{-2}(-\phi_{z\bar{z}}) = (-\phi_{s\bar{s}}) + 2|\phi_{s\bar{z}}|^2\phi_{z\bar{z}}^{-1} - \phi_{s\bar{z}}\overline{\phi_{s\bar{z}}}\phi_{z\bar{z}}^{-1} = -c(\phi).$$

Hence, the local calculations in the Calabi–Yau case give that

$$\frac{\partial}{\partial t} c(\phi) = \Delta_{\omega_t} c(\phi) - c(\phi) + |A_\phi|^2_{\omega_t^X}.$$

Finally, since a normalized Kähler–Einstein weight is stationary for the nonnormalized Kähler–Ricci flow this finishes the proof of the theorem. □

The last fiberwise elliptic equation in the previous corollary (in the K_X-setting) was first obtained by Schumacher [2008], who used the maximum principle to deduce an interesting consequence:

Corollary 4.8. *Let $\pi : \mathscr{X} \to S$ be a fibration as in the previous theorem and assume that $K_{\mathscr{X}/S}$ is relatively ample. Then the canonical fiberwise Kähler–Einstein weight ϕ_{KE} on $K_{\mathscr{X}/S}$ is smooth with semipositive curvature form on $\mathscr{X}$. Moreover, if the Kodaira–Spencer classes of the fibration are nontrivial for all s, then the curvature form of ϕ_{KE} is* strictly *positive on $\mathscr{X}$.*

The first part of the corollary was also shown by Tsuji [2006; 2011] using his iteration. Similarly, by a simple application of the parabolic maximum principle we deduce the following corollary from the parabolic equation in the previous theorem.

Corollary 4.9. *We make the same assumptions as in the previous theorem. Let ϕ_t evolve according to the Kähler–Ricci flow in the nonnormalized $\pm K_X$-setting. If the initial weight has (semi)positive curvature form on $\mathscr{X}$ then so has ϕ_t for all t. More precisely, $(dd^c\phi_t)_{x_z} > 0$ (in all $n+1$ directions) at any point x_s in the fiber $\mathscr{X}_s$ unless $(dd^c\phi_0)^{n+1}$ and A_{ϕ_0} vanish identically on $\mathscr{X}_s$.*

Proof. As usual we may assume that S is embedded in $\mathbb{C}$. Let us start with the *semi*positive case where the conclusion follows from the *weak* maximum principle. Indeed, assume to get a contradiction that $c(\phi_t) \geq 0$ on $\mathscr{X}$ for $t = 0$ but that there is (t, s, x) such that at (t, s, x) we have $c(\phi_t)(s, x) < 0$. By optimizing over (x, t) we may also assume that $\partial(e^{at}c(\phi_t))/\partial t \leq 0$, $\Delta_{\omega_t^X} c(\phi_t) \geq 0$. Then (4-5) gives

$$0 \geq e^{at}\left(ac(\phi_t) + \frac{\partial c(\phi_t)}{\partial t}\right) = e^{at}\big(\Delta_{\omega_t^X} c(\phi_t) - (a \pm 1)c(\phi_t) + |A_{\phi_t}|^2_{\omega_t^X}\big).$$

But if a is chosen so that $a \pm 1 > 0$, the right-hand side above is strictly positive, giving a contradiction. To handle the remaining cases we invoke the following well-known *strong* maximum principle for the heat operator (which by standard argument can be reduced to the corresponding local statement in [Protter and Weinberger 1967]): let $h_t \geq 0$ satisfy

$$\frac{\partial h_t}{\partial t} \geq \Delta_{g_t} h_t \quad \text{on } [0, T] \times X$$

for any smooth family g_t of Riemannian metrics. Then either $h_t > 0$ for all $t > 0$ or $h_0 \equiv 0$. In our case we set $h_t = e^{at}c(\phi_t)$ with $a = -\pm 1$ and conclude that if it is not the case that $c(\phi_t) > 0$ for all $t > 0$ then $c(\phi_0) \equiv 0$ and hence

$$\frac{\partial}{\partial t} c(\phi_t)_{t=0} = |A_{\phi_0}|^2_{\omega_0^X}.$$

If we now assume, to get a contradiction, that the right-hand side above is strictly positive at x_0 then it follows that there is an $\epsilon > 0$ such that $c(\phi_t)(x_0) > 0$ for $t \in (0, \epsilon[$; that is, for such t it is not the case that $c(\phi_t) \equiv 0$ on X. Hence, as explained above $c(\phi_t) > 0$ on all of $]0, \infty[\times X$, which yields the desired contradiction. □

In particular, the previous corollary says that if the fibration $\mathscr{X}$ is infinitesimally nontrivial then the nonnormalized Kähler–Ricci flows instantly make any semipositively curved initial weight *strictly* positive.

Next we note that integrating the last formula in the previous theorem immediately gives the following corollary first shown by Fujiki and Schumacher [1990, Theorem 7.9].

Corollary 4.10. *We make the same assumptions as in the previous theorem. Let ϕ_{KE} be the weight of a smooth metric on $\pm K_{\mathscr{X}/S}$ which restricts to a normalized Kähler–Einstein weight on each fiber. Then*

$$\pi_*((dd^c\phi_{\mathrm{KE}})^{n+1}/(n+1)!) = \pm\omega_{\mathrm{WP}}$$

on S, where π_* denotes the fiber integral. In particular, if S is effectively parametrized (i.e., all Kodaira–Spencer classes are nontrivial) then $\pm\pi_*(dd^c\phi_-)^{n+1}$ and hence the Weil–Petersson metric ω_{WP} is a Kähler form on the base S.

Remark 4.11. It follows immediately from the previous corollary that when X is a Fano manifold the normalized Kähler–Einstein weight ϕ_{KE} *never* has semipositive curvature on all of $\mathscr{X}$ if the family is effectively parametrized. Combining this fact with Corollary 4.9 shows that the relative Kähler–Ricci flow in the nonnormalized K_X-setting never converges in the $L^1(\mathscr{X})$-topology for an initial weight ϕ_0 with semipositive curvature form on an effectively parametrized fibration $\mathscr{X}$.

4C. *Quantization: The Bergman iteration.* The (*normalized*) *Bergman iteration* in the $\pm K_X$-setting on $\mathscr{H}_{\pm K_X}$ is defined precisely as in Section 3B, but using the probability measure $\mu_\pm(\phi)$ in the definition of $\mathrm{Hilb}^{(k)}(\phi, \mu_\pm(\phi))$. Similarly, the nonnormalized *Bergman iteration* is defined in terms of the measure $\mu'_\pm$. The virtue of the nonnormalized setting is that the corresponding Hilbert norms correspond to the "adjoint" norms appearing in Berndtsson's Theorem 3.10:

$$\mathrm{Hilb}^{(k)}(\phi, \mu_{\mu'_\pm})(f, f) := i^{n^2} \int_X f \wedge \bar{f} e^{-(k\pm 1)\phi} := \mathrm{Hilb}_{(k-1)L+K_X}(\phi) \tag{4-6}$$

for $L = \pm K_X$. Moreover, they are clearly *decreasing* in ϕ (for $k \geq 1$) and hence the analogue of Proposition 3.13 of the corresponding Bergman iteration holds:

Proposition 4.12. *Consider the Bergman iteration $\phi_m^{(k)}$ in the nonnormalized $\pm K_X$-setting and assume that $\phi_m^{(k)} \leq \psi_m^{(k)}$. Then $\phi_{m+1}^{(k)} \leq \psi_{m+1}^{(k)}$. Moreover, if $d(\phi, \psi)$ denotes the sup norm of $\phi - \psi$ then*

$$d(\psi_{m+1}, \phi_{m+1}) \leq d(\psi_{m+1}, \phi_{m+1}) \left(1 \pm \frac{1}{k}\right).$$

In particular, the Bergman iteration decreases the distance $d(\phi, \psi)$ in the nonnormalized K_X-setting.

Proof. Given the discussion preceding the proposition we just have to prove the claimed property of the distance d. But this follows directly from the monotonicity in the first part combined with the fact that $\log \rho^{(k)}(\phi_m + c)/k = \log \rho^{(k)}(\phi_m)/k - \pm \frac{c}{k}$, which in turn follows from $\mu'_\pm(\phi + c) := e^{\pm(\phi+c)} = \mu'_\pm(\phi) e^{\pm c}$. □

On the other hand, the following monotonicity of functionals holds in the *normalized* setting:

Lemma 4.13. *The functionals $-I_{\mu_\pm}$ and $\mathscr{L}^{(k)}$ are increasing along the normalized Bergman iteration on $\mathscr{H}_{\pm K_X}$. Moreover, they are strictly increasing at $\phi_m^{(k)}$ unless $\phi_m^{(k)}$ is stationary* (*when $k > 1$ in the case of I_{μ_+}*).

Proof. By the general Lemma 2.6 $\mathscr{L}^{(k)}$ is increasing and I_- is decreasing under the iteration, since $\phi \mapsto I_-(\phi)$ is concave with respect to the affine structure by Jensen's inequality. To show that $I_+^{(k)}$ is increasing in the K_X-setting just observe that

$$I_+(\phi_{m+1}^{(k)}) - I_+(\phi_m^{(k)}) := \log \frac{\int e^{(\phi_{m+1}^{(k)} - \phi_m^{(k)})} e^{\phi_m^{(k)}}}{\int e^{\phi_m^{(k)}}} = \log \int (\rho^{(k)})^{\frac{1}{k}} \mu(\phi_m^{(k)}) \leq \log \left(\left(\int \rho^{(k)}(\phi_m^{(k)}) \mu \right)^{\frac{1}{k}} \right) = 0,$$

using Jensen's inequality applied to the concave function $t \mapsto t^{1/k}$, which is strictly concave for $k > 1$. □

Convergence of the Bergman iteration at a fixed level k.

Theorem 4.14. *The Bergman iteration $\phi_m^{(k)}$ at level k converges, when the discrete time $m \to \infty$, to a balanced weight $\phi_\infty^{(k)}$ in the following settings:*

- *the normalized K_X-setting;*
- *the normalized $-K_X$-setting if it is a priori assumed that there exists some balanced metric at level k and $H^0(TX) = 0$;*
- *the normalized $-K_X$-setting for k sufficiently large under the assumption that X admits a Kähler–Einstein metric and $H^0(TX) = 0$;*
- *the nonnormalized $+K_X$-setting, where the limiting balanced weight is the unique normalized one.*

Proof. Proof of the first point: By the previous lemma $-I_\mu$ is increasing and as shown in [Berman et al. 2009] $-\mathcal{F}_\mu^{(k)}$ is coercive (as follows immediately from Jensen's inequality). Moreover, as shown in [Berman et al. 2009] balanced weights are unique modulo scaling and hence all the convergence criteria in Proposition 2.9 are hence satisfied.

Proof of the second point: By the previous lemma $-I_\mu$ is increasing and as shown in [Berman et al. 2009] it follows immediately from Berndtsson's theorem (Theorem 3.10) applied to $L = -K_X$ that $-\mathcal{F}_\mu^{(k)}$ is strictly convex modulo scaling. Hence, the convergence follows by combining Proposition 2.9 and Lemma 2.10.

Proof of the third point: The fact that $-\mathcal{F}_\mu^{(k)}$ is coercive was shown in [Berman et al. 2009] (using the corresponding coercivity of $-\mathcal{F}_\mu$ on $\mathcal{H}_L$). Given this coercivity the convergence follows as in the previous point.

Proof of the fourth point: Let $(\phi')_m^{(k)} = \phi_m^{(k)} + C_m^{(k)}$ denote the nonnormalized Bergman iteration in the K_X-setting. By the definition of the Bergman iteration (compare (4-10) below),

$$(C_{m+1}^{(k)} - C_m^{(k)}) = -C_m^{(k)}/k - I(\phi_m^{(k)})/k$$

where by the first point above $I(\phi_m^{(k)}) \to I_\infty$ when $m \to \infty$. Set $D_m := C_m^{(k)} + I(\phi_m^{(k)})$. Then

$$D_{m+1} = \Big(1 - \frac{1}{k}\Big) D_m + \epsilon_m,$$

where $\epsilon_m = (I(\phi_{m+1}^{(k)}) - I(\phi_m^{(k)})) \to 0$ as $m \to \infty$. It follows for elementary reasons that $D_m \to 0$; that is, $C_m^{(k)} \to -I_\infty$ showing that $(\phi')_m^{(k)}$ indeed converges and $I_+((\phi')_m^{(k)}) \to 0$, proving the second point. For completeness we finally show that $D_m \to 0$. Assume for a contradiction that this is not the case. Then $D_{m+1}/D_m \to 1 - \frac{1}{k}$ and hence $D_m \le C_\delta(1 - \frac{1}{k} + \delta)^m \to 0$ for δ sufficiently small, giving a contradiction. ☐

The convergence in the fourth point above also follows immediately from the contracting property of the corresponding iteration (compare the proof of Theorem 4.20 below). We also note the following direct consequence of Berndtsson's theorem (Theorem 3.10), using formula (4-6) in the nonnormalized setting.

Corollary 4.15. *The Bergman iteration in the nonnormalized $\pm K_X$-setting preserves the* (*semi*)*positivity of the curvature of the initial weight. Moreover*, *if the fibration $\mathscr{X}$ is assumed infinitesimally nontrivial then any initial weight on $\pm K_{\mathscr{X}/S}$ which is semipositively curved and strictly positively curved along the fibers of $\mathscr{X}$ becomes strictly positively curved under the iteration.*

Combining the previous corollary and Theorem 4.14 now gives the following:

Corollary 4.16. *Let $\pi : \mathscr{X} \to S$ be a proper holomorphic submersion with $K_{\mathscr{X}/S}$ relatively ample. Let $\phi^{(k)}$ be the weight on $K_{\mathscr{X}/S}$ obtained by requiring that its restriction to any fiber is the unique normalized balanced weight at level k*; *i.e*, $\int_{\mathscr{X}_s} e^{\phi^{(k)}} = 1$. *Then $\phi^{(k)}$ is smooth with semipositive curvature form. Moreover*, *if the fibration $\mathscr{X}$ is assumed infinitesimally nontrivial then $\phi^{(k)}$ is strictly positively curved.*

Proof. Since positivity and smoothness are local notions it is enough to prove the corollary when S is embedded in $\mathbb{C}$.

Smoothness: By definition $\phi^{(k)} = \mathrm{FS}^{(k)}(H^{(k)})$ where $H^{(k)}$ is an element in the finite-dimensional smooth manifold $\mathscr{H}^{(k)}$ uniquely determined by $G^{(k)}(H^{(k)}, s) = 0$ [Berman et al. 2009], where $G^{(k)}$ is the smooth map defined by

$$G^{(k)}(H^{(k)}, s) := (T^{(k)} - I, I_+ \circ \mathrm{FS}^{(k)}) \in \mathscr{H}^{(k)} \times \mathbb{R}.$$

Moreover, as shown in [Berman et al. 2009] the linearization of $T^{(k)} - I$ is invertible modulo scaling (since it represents the differential of a functional on $H^{(k)}$ which is strictly convex modulo scaling). Hence, the claimed smoothness follows from the implicit function theorem.

Positivity: Since $K_{\mathscr{X}/S}$ is assumed relatively ample it admits a smooth weight ϕ_0, which has fiberwise positive curvature form. After adding a sufficiently large multiple of the pull-back from the base of $|s|^2$ we may assume that ϕ_0 has positive curvature over $\mathscr{X}$. By the last point of the previous theorem the Bergman iteration $\phi_m^{(k)}$ in the nonnormalized K_X-setting with initial weight ϕ_0 yields a sequence of weights on $K_{\mathscr{X}/S}$ converging, when $m \to \infty$, uniformly to the unique normalized balanced weight $\phi^{(k)}$ at level k. As a consequence $dd^c\phi^{(k)} \geq 0$ on $\mathscr{X}$. Moreover, if the fibration $\mathscr{X}$ is assumed infinitesimally nontrivial the previous corollary shows that applying the Bergman iteration to $\phi^{(k)}$ yields a strictly positively curved metric. But since $\phi^{(k)}$ is fixed under the iteration this finishes the proof of the corollary. ☐

Corollary 4.17. *Let $\pi : \mathscr{X} \to S$ be the universal curve of the Teichmüller space of complex curves of a genus $g \geq 2$. Fix a positive integer k* (*for $g = 2$ we assume that $k \geq 2$*). *Under the natural isomorphism*

$$(T^{1,0}S)^* = \pi_*(2K_{\mathscr{X}/S})$$

the fiberwise normalized balanced weight $\phi^{(k)}$ on $K_{\mathscr{X}/S}$ at level k (*appearing in the previous corollary*) *induces a Hermitian metric $\omega^{(k)}$ on S with a curvature which is dually Nakano positive. Moreover*, *when $k \to \infty$ the metric $\omega^{(k)}$ converges towards the Weil–Petersson metric ω_{WP} pointwise on S.*

Proof. As is classical the assumptions on k ensure that $K_{\mathscr{X}/S}$ is very ample. By the previous corollary $\phi^{(k)}$ is a smooth weight on $K_{\mathscr{X}/S} \to \mathscr{X}$ with strictly positive curvature and hence the L^2-metric on the direct image bundle $\pi_*(\mathscr{L}+K_{\mathscr{X}/S})$ (with $\mathscr{L}=K_{\mathscr{X}/S}$) induced by $\phi^{(k)}$ has, according to the first point in Theorem 3.10, a curvature which is positive in the sense of Nakano. Since

$$T^{1,0}S_{|s}=H^1(T^{1,0}\mathscr{X}_s)\cong H^0(2K_{\mathscr{X}_s})^*,$$

this proves the first statement. To prove the pointwise convergence on S of $\omega^{(k)}$ towards ω_{WP} it is enough to prove that

$$e^{-\phi^{(k)}}\to e^{-\phi_{\mathrm{KE}}}$$

in $L^1_{\mathrm{loc}}(X)$ for $X=\mathscr{X}_s$ (since, by definition, it implies the pointwise convergence of the corresponding Hermitian metrics on $\pi_*(\mathscr{L}+K_{\mathscr{X}/S})$). But this convergence follows from the L^1 convergence of $\phi^{(k)}$ towards ϕ_{KE} (Theorem 4.14) combined with the fact that $J(\phi^{(k)})$ is uniformly bounded, as shown in [Berman et al. 2009] (see Lemma 6.4 therein). Alternatively, it follows immediately from the uniform convergence in Theorem 4.20 below. □

The convergence in the previous corollary should be compared with the approximation results for the Weil–Petterson metric for moduli spaces of higher-dimensional manifolds recently obtained in [Keller and Lukic 2009]. The approximating *Kähler* metrics ω_k' in that work are related to different balanced metrics, namely those defined with respect to Donaldson's original setting [2001] (where $\mu(\phi)=\mathrm{MA}(\phi)$).

Convergence towards the Kähler–Ricci flow.

Theorem 4.18. *The following convergence results hold in all settings introduced in the beginning of Section 4 (i.e., in the (non)normalized $\pm K_X$-settings). Fix a smooth and strictly psh initial weight ϕ_0 on $\pm K_X$ and consider the corresponding Bergman iteration $\phi_m^{(k)}$ at level k and discrete time m, as well as the corresponding Kähler Ricci flow ϕ_t. Then there is a constant A such that*

$$\sup_X|\phi_m^{(k)}-\phi_{m/k}|\le Am/k^2 \tag{4-7}$$

uniformly in (m,k) satisfying $m/k\le T$ (in the K_X-setting A is independent of T). In particular, if m_k is a sequence such that $m_k/k\to t$, then $\phi_{m_k}^{(k)}\to\phi(t)$ uniformly on X and

$$dd^c\phi_{m_k}^{(k)}\to\omega_t$$

on X in the sense of currents, where ω_t evolves according to the corresponding Kähler–Ricci flow (4-2). *The corresponding result also holds for the corresponding nonnormalized flows and in the relative setting, where the convergence is locally uniform with respect to the base parameter s.*

Proof. In the case of the nonnormalized K_X-setting (denoted by primed objects) the proof of Theorem 3.15 carries over essentially verbatim, thanks to the last statement in Proposition 4.12 and Corollary 4.3 which gives the uniformity with respect to $T\in[0,\infty]$. To handle the nonnormalized $-K_X$-setting we need to modify the previous argument slightly. More precisely, we will prove that

$$\sup_X|\phi_m^{(k)}-\phi_{m/k}|\le A\left(1+\frac{1}{k}\right)^m m/k^2. \tag{4-8}$$

Accepting this for the moment the claimed convergence when $m_k/k \to t$ follows using that

$$\Big(1+\frac{1}{k}\Big)^m = \Big(\Big(1+\frac{1}{k}\Big)^k\Big)^{m/k} \leq e^{m/k} \leq e^T,$$

when $m/k \leq T$. To prove (4-8) first observe that Step 1 in the proof of Theorem 3.15 still applies for (m,k) such that $m/k \leq T$ (using Proposition 2.11 applied to the nonnormalized $-K_X$-setting). In other words, there is a constant A (depending on T) such that

$$\sup_X \big|\psi_{k,m+1} - (\psi_{k,m} + F^{(k)}(\psi_{k,m}))\big| \leq A(1/k^2)$$

for all (m,k) such that $m/k \leq T$. Now we fix the integer k and assume as an induction hypothesis that (4-7) holds for m with A the constant in the previous inequality. By Proposition 3.13,

$$\begin{aligned}\sup_X \big|(\psi_{k,m} + F^{(k)}(\psi_{k,m})) - (\phi_m^{(k)} + F^{(k)}(\phi_m^{(k)}))\big| &\leq \sup_X |\psi_{k,m} - \phi_m^{(k)}|\Big(1+\frac{1}{k}\Big)\\ &\leq \Big(A\Big(1+\frac{1}{k}\Big)^m m/k^2\Big)\Big(1+\frac{1}{k}\Big)\end{aligned}$$

with the same constant A as above, using the induction hypothesis in the last step. Combining this estimate with the previous inequality gives

$$\sup_X |\psi_{k,m+1} - \phi_{m+1}^{(k)}| \leq A\Big(1+\frac{1}{k}\Big)^{m+1} m/k^2 + A/k^2.$$

But using that $1 \leq (1+\frac{1}{k})^{m+1}$ in the last term above proves the induction step and hence finishes the proof of the estimate (4-8).

To treat the Kähler–Ricci flows ϕ_t in the normalized settings we write

$$\phi_t' = \phi_t + C_t,$$

where C_t is a constant for each t. Then

$$\frac{\partial C_t}{\partial t} = -I_\pm(\phi_t'). \tag{4-9}$$

Indeed, by the definition of the flow ϕ_t' and ϕ_t, we have

$$\frac{\partial \phi_t'}{\partial t} = \log\big(\mathrm{MA}(\phi_t') - \pm\phi_t'\big), \qquad \frac{\partial \phi_t}{\partial t} = \log\big(\mathrm{MA}(\phi_t) - \pm\phi_t\big) + \pm I_\pm(\phi_t).$$

By scale invariance we may as well replace ϕ_t with ϕ_t' on the right side of the second equation above and hence subtracting the second equation from the first one proves (4-9).

Similarly, writing

$$(\phi')_m^{(k)} = \phi_m^{(k)} + C_m^{(k)},$$

we obtain the following difference equation, using that the map $\phi \mapsto \rho^{(k)}(\phi)$, defined with respect to $\mu_\pm$, is scale-invariant:

$$C_{m+1}^{(k)} - C_m^{(k)} = -\frac{1}{k} I_\pm((\phi')_m^{(k)}). \tag{4-10}$$

Now, as explained above, the estimate (4-7) holds for the primed objects and hence by the scaling equivariance of $I_\pm$:

$$\left|I_\pm(\phi'_{m/k}) - I_\pm((\phi')^{(k)}_m)\right| \leq Am/k^2. \tag{4-11}$$

A simple version of the argument given in the proof of Theorem 4.18 now shows, by comparing the differential equation (4-9) with the difference equation (4-10) and using (4-11), that

$$|C^{(k)}_m - C_{m/k}| \leq Bm/k^2$$

for a uniform constant B. All in all this hence finishes the proof of the theorem. □

We also have the following analogue of Corollary 3.16:

Corollary 4.19. *For a fixed initial data $\phi_0 = \phi^{(k)}_0 \in \mathcal{H}_{\pm K_X}$ the following convergence results hold for the Bergman iteration $\phi^{(k)}_m$ in the normalized $\pm K_X$-setting* (*in the $-K_X$-setting it is assumed that $H^0(TX) = 0$ and X a priori admits a Kähler–Einstein metric*):

- *For any sequence m_k such that $m_k/k \to \infty$ the convergence $\phi^{(k)}_{m_k} \to \phi_\infty$ holds in the L^1-topology on X.*
- *The balanced weights $\phi^{(k)}_\infty := \lim_{m\to\infty} \phi^{(k)}_m$ at level k converge, when $k \to \infty$, in the $\mathcal{C}^\infty$-topology, to the weight ϕ_∞ which is the large time limit of the corresponding Kähler–Ricci flow.*

Moreover, the convergence in the second point also holds in the nonnormalized K_X-setting, where the limit ϕ_∞ coincides with the canonical Kähler–Einstein weight ϕ_{KE}. In the relative case all convergence results hold fiberwise locally uniformly with respect to the base parameter s.

Proof. The proof of the first two points proceeds exactly as in the previous setting (again using the variational characterization in [Berman et al. 2009]). As for the claimed convergence in the nonnormalized setting it is obtained by noting that the large m limit $(\phi')^{(k)}_m$ in the nonnormalized setting is the unique balanced weight such that $I_\pm((\phi')^{(k)}_\infty) = 0$. In other words, $(\phi')^{(k)}_\infty = \phi^{(k)}_\infty - I_\pm(\phi^{(k)}_\infty)$, where $\phi^{(k)}_\infty$ is the large m limit of the iteration in the normalized setting. But by the second point above this means that $(\phi')^{(k)}_\infty \to \phi_\infty - I_\pm(\phi_\infty)$ in L^1 (also using the continuity with respect to the L^1-topology of the functional $I_\pm$ on compacts; compare [Berman et al. 2009]). By uniqueness, this means that the limit must be ϕ_{KE}. □

4D. ***Uniform convergence of the balanced weights in the K_X-setting.*** Next we point out that in the K_X-setting the convergence of the balanced weights is actually *uniform* (the proof is independent of the variational proof of a weaker convergence given in [Berman et al. 2009]). The proof simply uses that $\phi^{(k)}$ is close to ϕ_{t_k} where ϕ_t is the corresponding Kähler–Ricci flow and t_k is a suitable sequence tending to infinity.

Theorem 4.20. *Let $\phi^{(k)}$ be the balanced weight at level k on the canonical line bundle K_X* (*in the nonnormalized setting*). *When $k \to \infty$, the weights $\phi^{(k)}$ converge uniformly towards the normalized Kähler–Einstein weight ϕ_{KE}.*

Proof. Fix a smooth and positively curved weight ϕ_0 on K_X and denote by $\phi_m^{(k)}$ the Bergman iteration at level k with initial data $\phi_0^{(k)} = \phi_0$. By Proposition 4.12 the map whose iterations define the Bergman iterations is a contraction mapping with contracting constant $q = (1 - \frac{1}{k}) < 1$ and hence it follows from the Banach fixed point theorem that

$$\|\phi^{(k)} - \phi_m^{(k)}\|_{L^\infty} \leq \frac{q^m}{(1-q)} \|\phi_1^{(k)} - \phi_0\|_{L^\infty}.$$

By definition we have $\phi_1^{(k)} - \phi_0 = \frac{1}{k} \log \rho(k\phi)$, which, according to Proposition 2.11, is uniformly bounded by a constant times $\frac{1}{k} \log k$; hence

$$\|\phi^{(k)} - \phi_m^{(k)}\|_{L^\infty} \leq C \left(\left(1 - \frac{1}{k}\right)^k \right)^{m/k} \log k.$$

Next we take the sequence $m = m_k := [k^{3/2}]$ where $[c]$ denotes the smallest integer which is larger than c. Then $t_k := m_k/k = k^{1/2} \to \infty$ as $k \to \infty$ and since $(1 - \frac{1}{k})^k \to e^{-1} < 1$ we conclude that

$$\|\phi_m^{(k)} - \phi_0\|_{L^\infty} \to 0$$

as $k \to \infty$. If now ϕ_t denotes the Kähler–Ricci flow in the nonnormalized K_X-setting we have, according to Theorem 4.18, that

$$\|\phi_m^{(k)} - \phi_{m_k/k}\|_{L^\infty} \to 0$$

using that $m_k/k^2 \to \infty$. Finally, since $\phi_{t_k} \to \phi_{\mathrm{KE}}$ uniformly as $t_k \to \infty$ this proves the theorem. Of course, the last convergence is not really needed for the proof as we may as well start with $\phi_0 = \phi_{\mathrm{KE}}$ which is trivially fixed under the Kähler–Ricci flow. □

It should be pointed out that the uniform convergence in the previous theorem has been previously obtained by Berndtsson (who also related it to Tsuji's iteration [Tsuji 2006]), using a different approach — see the announcement in [Berndtsson 2009c]. But hopefully the relation to the convergence of the Kähler–Ricci flow above may shed some new light on the convergence.

4E. *Families of varieties of general type and comparison with the NS metric.* The quantized setting concerning the case when K_X is ample admits a straightforward generalization to the case when K_X is merely $\mathbb{Q}$-effective [Lazarsfeld 2004]. For simplicity we will only discuss the case when K_X is big; that is, X is a nonsingular variety of general type. Moreover, we will no longer assume that the map π is a submersion. More precisely, we are given a surjective quasiprojective morphism $\pi : \mathscr{X} \to S$ between nonsingular varieties such that the generic fiber is a variety of general type. We denote by S^0 the maximal Zariski open subset of S such that π restricted to $\mathscr{X}^0 := \pi^{-1}(S^0)$ is a submersion, that is, a smooth morphism (and hence the fibers of S^0 are nonsingular varieties of general type).

Let us first consider the general absolute case, where we are given a line bundle $L \to X$ and an integer k such that kL is effective; that is, $H^0(X, kL) \neq \{0\}$. The main new feature in this more general setting is that any Bergman weight ψ_k at level k, that is, $\psi_k \in \mathrm{FS}^{(k)}(\mathscr{H}^{(k)})$, will usually have singularities; that is, it defines a singular metric on L with positive curvature form. More precisely, the weight $k\psi_k$ on kL is singular precisely along the base locus $Bs(kL)$ of kL, that is, the intersection of the zero sets of all elements in $H^0(kL)$. Anyway, the *difference* of any two Bergman metrics is clearly bounded. Moreover, when $L = K_X$ the measure $\mu_{\psi_k} := e^{\psi_k}$ has a smooth density which vanishes precisely along $Bs(kL)$. As a consequence, we may fix such a reference (singular) weight $\phi_0 := \psi_k$ and the reference measure $\mu_0 := e^{\psi_k}$. Then Lemma 2.7 still applies (as explained in the remark following the lemma). As a consequence the proof of the convergence of the Bergman iteration to a balanced weight at level k in the nonnormalized K_X-setting (Theorem 4.14) is still valid as long as kK_X is effective. Combining this latter convergence with the generalizations [Berndtsson and Păun 2008b; 2008a] of Berndtsson's theorem (Theorem 3.10) and the invariance of plurigenera [Siu 1998] then gives the following generalization of Corollary 4.16:

Theorem 4.21. *Let $\pi : \mathscr{X} \to S$ be a surjective quasiprojective morphism such that the generic fiber is a variety of general type. Then, for k sufficiently large there is a unique singular weight $\phi^{(k)}$ on the relative canonical line bundle $K_{\mathscr{X}/S} \to \mathscr{X}$ with positive curvature current, such that the restriction of $\phi^{(k)}$ to any fiber over S^0 is a normalized and balanced weight at level k. Moreover, the weight $\phi^{(k)}$ is smooth on the Zariski open set defined as the complement in $\mathscr{X}$ of $\bigcup_{s\in S^0} Bs(kK_{\mathscr{X}_s}) \cup \pi^{-1}(S - S^0)$.*

Proof. Let us first prove the positivity statement. As before we may assume that S is a domain in $\mathbb{C}$. First we consider the behavior over the set S^0, that is, where the fibration is a submersion. Fix $s_0 \in S_0$ and write $X = \mathscr{X}_{s_0}$. Let (f_i) be a basis in $H^0(X, kK_X)$. By the invariance of plurigenera [Siu 1998] s_0 has a neighborhood $U \subset S_0$ with holomorphic sections F_i of $kK_{\mathscr{X}/S} \to U$ such that F_i restricts to f_i on X. After perhaps shrinking U we may hence assume that the restrictions of F_i to any fiber give a basis in $H^0(\mathscr{X}_s, kK_{\mathscr{X}_s})$. Now let $\phi_0 := \frac{1}{k}\log\bigl(\frac{1}{N_k}\sum |F_i|^2\bigr)$ so that ϕ_0 is a singular weight on $K_{\mathscr{X}/S}$ over U with positive curvature and such that ϕ_0 restricts to a Bergman weight at level k on each fiber. In particular,

$$\int_{\mathscr{X}_s} |f|^2 e^{-(k-1)\phi_0} < \infty \tag{4-12}$$

for any $f \in H^0(\mathscr{X}_s, kK_{\mathscr{X}_s})$. Decomposing, as before, $kK_X = (k-1)L + K_X$ with $L = K_X$, but now using Theorem 3.5 in [Berndtsson and Păun 2008b], shows that the curvature current of the weight $\phi_1^{(k)} := \mathrm{FS}^{(k)} \circ \mathrm{Hilb}^{(k)}(\phi_0)$ on $K_{\mathscr{X}/S}$ is positive over U. Since, by definition, $\phi_1^{(k)}$ is still fiberwise a Bergman weight at level k we may iterate the same argument and conclude that $\phi_m^{(k)}$ has a positive curvature current for any m. Now, as explained in the discussion before the statement of the theorem,

$$m \to \infty \implies \sup_{\mathscr{X}_s} |\phi_m^{(k)} - \phi^{(k)}| \to 0,$$

locally uniformly with respect to s, where $\phi^{(k)}$ is the unique normalized fiberwise balanced weight at level k. In particular, it follows that $\phi^{(k)}$ has a curvature current which is positive over S^0.

To prove the claimed extension property of $\phi^{(k)}$ over $S - S_0$ first note that, writing $X = \mathscr{X}_s$ for a fixed fiber,

$$\phi^{(k)} \le \phi_{\mathrm{NS}}^{(k)} := \log\Biggl(\sup_{f\in H^0(X,kK_X)} \biggl(|f|^{2/k} \Big/ \int_X (f \wedge \bar{f})^{1/k}\biggr)\Biggr), \tag{4-13}$$

where $k\phi_{\mathrm{NS}}^{(k)}$ is the weight of the Narasimhan–Simha (NS) metric on $kK_{\mathscr{X}/S}$ [Narasimhan and Simha 1968; Kawamata 1982; Tsuji 2011; Berndtsson and Păun 2008a]. Accepting this for the moment we can use the result in [Berndtsson and Păun 2008a] saying that $\phi_{\mathrm{NS}}^{(k)}$ is locally bounded from above, with a constant which does not blow up as s converges to a point in $S - S^0$ (this is proved by an $L^{2/k}$ variant of the local Ohsawa–Takegoshi L^2-extension theorem). By the inequality (4-13) it hence follows that $\phi^{(k)}$ is also locally bounded from above by the same constant and then the claimed extension property follows from basic pluripotential theory.

Finally, to prove the inequality (4-13) fix a point $x \in X$. By the extremal definition of Bergman kernels there are sections f_i (depending on x) such that

$$\phi^{(k)}(x) = \frac{1}{k}\log\left(\frac{1}{N_k}|f_1|^2(x)\right) \quad \text{and} \quad \phi^{(k)} = \frac{1}{k}\log\left(\frac{1}{N_k}\sum_i |f_i|^2\right)$$

on X. Since $\int_X e^{\phi^{(k)}} = 1$ it hence follows that

$$\int_X \left(\frac{1}{N_k} f_1 \wedge \bar{f}_1\right)^{1/k} \leq 1,$$

which finishes the proof of the inequality (4-13), since $f_1/(N_k)^{1/2}$ is a candidate for the supremum defining $\phi_{\mathrm{NS}}^{(k)}$.

As for the last smoothness statement in the theorem it is proved exactly as in Corollary 4.16, using that $\pi_*(kK_{\mathscr{X}/S})$ is a locally trivial vector bundle over S^0. Indeed, it follows as before that the fiberwise normalized balanced metrics $H_s^{(k)}$, which by the local freeness may be identified with a family in $\mathrm{GL}(N_k)$, form a *smooth* family. Applying the Fubini–Study map to get $\phi^{(k)}$ then introduces the singular locus described in the statement of the theorem. □

Remark 4.22. If one does not invoke the invariance of plurigenera in the proof of the previous theorem then the same argument gives the slightly weaker statement where S^0 is replaced by the intersection of S^0 with a Zariski open set where $\pi_*(kK_{\mathscr{X}/S})$ is a locally trivial vector bundle. If one could then prove that the extension of $\phi^{(k)}$ is such that the integrability condition (4-12) holds over all of S, then the invariance of plurigenera would follow from a well-known version of the Ohsawa–Takegoshi extension theorem. It would be interesting to see if this approach is fruitful in the nonprojective Kähler case where the invariance of plurigenera is still open. When $\phi^{(k)}$ is replaced by the weight of the NS-metric $\phi_{\mathrm{NS}}^{(k)}$ (see formula (3-14)) this approach was used in [Tsuji 2011] to give a new proof of the invariance of plurigenera (in the projective case).

It should also be pointed out that (singular) Kähler–Einstein metrics and Kähler–Ricci flows have been studied recently for K_X big. For example, using the deep finite generation of the canonical ring there is a unique Kähler–Einstein weight with minimal singularities which satisfies the Monge–Ampère equation

$$(dd^c\phi_{\mathrm{KE}})^n/n! = e^{\phi_{\mathrm{KE}}}$$

on a Zariski open set in X [Eyssidieux et al. 2009; Boucksom et al. 2010]. It seems likely that the positivity result in Corollary 4.8 can be extended to families of such singular weights ϕ_{KE}. But there are several regularity issues which need to be dealt with. Moreover, it also seems likely that the canonical balanced weights $\phi^{(k)}$ converge to ϕ_{KE}, when K_X is big, but this would require a generalization of the convergence results in [Berman et al. 2009] (which only concern ample line bundles). This latter conjectural convergence should be compared with the convergence of the weight of the NS-metrics proved in [Berman and Demailly 2012], saying that $\phi_{\mathrm{NS}}^{(k)}$ converges in L^1 (and uniformly on compacts of an Zariski open set) to

$$\phi_{\mathrm{can}} := \sup\left\{\psi : \int_X e^{\psi} = 1\right\},$$

where the sup is taken over all singular weights ψ on K_X with positive curvature current. In particular, $\phi_{\mathrm{KE}} \leq \phi_{\mathrm{can}}$, which is consistent with the inequality (4-13).

4F. ***Comparison with the constant scalar curvature and other settings.*** Given an ample line bundle $L \to X$ the absolute setting when $\mu(\phi) := (dd^c\phi)^n/n!$ was studied in depth by Donaldson [2001; 2005]. Of course, in this setting the Kähler–Ricci flow is trivial, but the corresponding quantized setting and the study of its large k limit is highly nontrivial. In fact, it was shown in [Donaldson 2001] that, if it is a priori assumed that $c_1(L)$ contains a Kähler metric ω with constant scalar curvature and if $H^0(TX) = \{0\}$, then the curvature forms of any sequence of balanced weights converge in the $\mathscr{C}^\infty$-topology to ω. Moreover, Donaldson showed that such balanced weights do exist for k sufficiently large. As earlier shown by Zhang this latter fact is equivalent to the polarized variety (X, kL) being stable in the sense of Chow–Mumford (with respect to a certain action of the group $\mathrm{SL}(N_k)$). An explicit proof of the convergence of the Bergman iteration in this setting was given in [Sano 2006] (see also [Donaldson 2005]).

Note that in this setting the functional I_μ is precisely the functional $\mathscr{E}$ (compare the beginning of Section 2). Since $\mathscr{E}$ is well-known to be concave on $\mathscr{H}_L$ with respect to the affine structure and $\mathscr{E} \circ \mathrm{FS}$ is geodesically convex on $\mathscr{H}^{(k)}$ the convergence of the corresponding Bergman iteration is also a consequence of Proposition 2.9.

It should also be pointed out that the role of the Kähler–Ricci flow of Kähler metrics in this setting is played by the Calabi flow. Indeed, as shown in [Fine 2010], the balancing flow, which is a continuous version of Donaldson's iteration, converges, at the level of Kähler metrics, in the large k limit to the Calabi flow. More precisely, the balancing flow $H_t^{(k)}$ is simply the scaled gradient flow on the symmetric space $\mathscr{H}^{(k)}$ of the functional $\mathscr{F}^{(k)}$ in this setting and the convergence holds for the curvature forms of the weights $\mathrm{FS}^{(k)}(H_t)$ in $\mathscr{H}_L$.

Remark 4.23. Another, less studied, setting of geometric relevance (see [Berndtsson 2009b]) appears when we let

$$\mu(\phi) := \frac{1}{N_l} \sum_{i=1}^{N_l} f_i \wedge \bar{f}_i e^{-l\phi}$$

for a fixed integer l where f_i is an orthonormal basis for $H^0(lL + K_X)$ equipped with the Hermitian metric induced by ϕ. When $L = -K_X$ and $l = 1$ this is precisely the normalized $-K_X$-setting. In the general case $I_\mu(\phi)$ is essentially the induced metric on the top exterior power of the Hilbert space $H^0(lL + K_X)$.

Moreover, as soon as the corresponding functional $\mathscr{F}_{\mu}^{(k)}$ has a critical point and $H^0(TX) = \{0\}$ the assumptions for convergence in Proposition 2.9 are satisfied (see [Berndtsson 2009b]).

Acknowledgments

I would like to thank Bo Berndtsson for interesting conversations on the topic of the present paper, in particular in connection to [Berndtsson 2009a; 2011]. In the first arXiv versions (1 and 2) of the present paper the very recent result about *strict* positivity of direct image bundles in [Berndtsson 2011] was not available, but in the present version it is used to improve the positivity properties of the Bergman iteration in the case of infinitesimally nontrivial fibrations. Also thanks to Valentino Tosatti and Yanir Rubinstein for comments on a draft of the present paper (in particular, thanks to Valentino Tosatti for pointing out the reference [Gross and Wilson 2000] in connection to semiflat metrics and the references in Remark 4.4, which was stated inaccurately in the previous version). Finally, thanks to the anonymous referee for helpful comments.

References

[Berman 2008] R. J. Berman, "Determinantal point processes and fermions on complex manifolds: Bulk universality", preprint, 2008. arXiv 0811.3341

[Berman and Boucksom 2010] R. Berman and S. Boucksom, "Growth of balls of holomorphic sections and energy at equilibrium", *Invent. Math.* **181**:2 (2010), 337–394. MR 2011h:32021 Zbl 1208.32020

[Berman and Demailly 2012] R. Berman and J.-P. Demailly, "Regularity of plurisubharmonic upper envelopes in big cohomology classes", pp. 39–66 in *Perspectives in analysis, geometry, and topology*, Progress in Mathematics **296**, Birkhäuser, New York, 2012. MR 2012m:32035 Zbl 06082161

[Berman et al. 2009] R. Berman, S. Boucksom, V. Guedj, and A. Zeriahi, "A variational approach to complex Monge–Ampère equations", preprint, 2009. arXiv 0907.4490

[Berndtsson 2009a] B. Berndtsson, "Curvature of vector bundles associated to holomorphic fibrations", *Ann. of Math.* (2) **169**:2 (2009), 531–560. MR 2010c:32036 Zbl 1195.32012

[Berndtsson 2009b] B. Berndtsson, "Positivity of direct image bundles and convexity on the space of Kähler metrics", *J. Differential Geom.* **81**:3 (2009), 457–482. MR 2010e:32020 Zbl 1187.53076

[Berndtsson 2009c] B. Berndtsson, "Remarks on a theorem by H. Tsuji", pp. 1107–1110 in *Multiplier Ideal Sheaves in Algebraic and Complex Geometry* (Oberwolfach, 2009), edited by S. Kebekus et al., Mathematisches Forschungsinstitut Oberwolfach **6**, European Mathematical Society, Zürich, 2009.

[Berndtsson 2011] B. Berndtsson, "Strict and nonstrict positivity of direct image bundles", *Math. Z.* **269**:3-4 (2011), 1201–1218. Zbl 05996694

[Berndtsson and Păun 2008a] B. Berndtsson and M. Păun, "A Bergman kernel proof of the Kawamata subadjunction theorem", preprint, 2008. arXiv 0804.3884

[Berndtsson and Păun 2008b] B. Berndtsson and M. Păun, "Bergman kernels and the pseudoeffectivity of relative canonical bundles", *Duke Math. J.* **145**:2 (2008), 341–378. MR 2009k:32020 Zbl 1181.32025

[Boucksom et al. 2010] S. Boucksom, P. Eyssidieux, V. Guedj, and A. Zeriahi, "Monge–Ampère equations in big cohomology classes", *Acta Math.* **205**:2 (2010), 199–262. MR 2011k:32049 Zbl 1213.32025

[Bourguignon et al. 1994] J.-P. Bourguignon, P. Li, and S.-T. Yau, "Upper bound for the first eigenvalue of algebraic submanifolds", *Comment. Math. Helv.* **69**:2 (1994), 199–207. MR 95j:58168 Zbl 0814.53040

[Cao 1985] H. D. Cao, "Deformation of Kähler metrics to Kähler–Einstein metrics on compact Kähler manifolds", *Invent. Math.* **81**:2 (1985), 359–372. MR 87d:58051 Zbl 0574.53042

[Chen 2000] X. Chen, "The space of Kähler metrics", *J. Differential Geom.* **56**:2 (2000), 189–234. MR 2003b:32031 Zbl 1041.58003

[Chen and Tian 2002] X. X. Chen and G. Tian, "Ricci flow on Kähler–Einstein surfaces", *Invent. Math.* **147**:3 (2002), 487–544. MR 2003c:53095 Zbl 1047.53043

[Deligne 1987] P. Deligne, "Le déterminant de la cohomologie", pp. 93–177 in *Current trends in arithmetical algebraic geometry* (Arcata, CA, 1985), edited by K. A. Ribet, Contemporary Mathematics **67**, American Mathematical Society, Providence, RI, 1987. MR 89b:32038 Zbl 0629.14008

[Donaldson 2001] S. K. Donaldson, "Scalar curvature and projective embeddings, I", *J. Differential Geom.* **59**:3 (2001), 479–522. MR 2003j:32030 Zbl 1052.32017

[Donaldson 2005] S. K. Donaldson, "Scalar curvature and projective embeddings, II", *Q. J. Math.* **56**:3 (2005), 345–356. MR 2006f:32033 Zbl 1159.32012

[Donaldson 2009] S. K. Donaldson, "Some numerical results in complex differential geometry", *Pure Appl. Math. Q.* **5**:2 (2009), 571–618. MR 2010d:32020 Zbl 1178.32018

[Eyssidieux et al. 2009] P. Eyssidieux, V. Guedj, and A. Zeriahi, "Singular Kähler–Einstein metrics", *J. Amer. Math. Soc.* **22**:3 (2009), 607–639. MR 2010k:32031 Zbl 1215.32017

[Fine 2010] J. Fine, "Calabi flow and projective embeddings", *J. Differential Geom.* **84**:3 (2010), 489–523. MR 2012a:32023 Zbl 1202.32018

[Fujiki and Schumacher 1990] A. Fujiki and G. Schumacher, "The moduli space of extremal compact Kähler manifolds and generalized Weil–Petersson metrics", *Publ. Res. Inst. Math. Sci.* **26**:1 (1990), 101–183. MR 92e:32012 Zbl 0714.32007

[Greene et al. 1990] B. R. Greene, A. Shapere, C. Vafa, and S.-T. Yau, "Stringy cosmic strings and noncompact Calabi–Yau manifolds", *Nuclear Phys. B* **337**:1 (1990), 1–36. MR 91g:83057 Zbl 0744.53045

[Gross and Wilson 2000] M. Gross and P. M. H. Wilson, "Large complex structure limits of $K3$ surfaces", *J. Differential Geom.* **55**:3 (2000), 475–546. MR 2003a:32042 Zbl 1027.32021

[Guedj and Zeriahi 2005] V. Guedj and A. Zeriahi, "Intrinsic capacities on compact Kähler manifolds", *J. Geom. Anal.* **15**:4 (2005), 607–639. MR 2006j:32041 Zbl 1087.32020

[Hamilton 1982] R. S. Hamilton, "Three-manifolds with positive Ricci curvature", *J. Differential Geom.* **17**:2 (1982), 255–306. MR 84a:53050 Zbl 0504.53034

[Kawamata 1982] Y. Kawamata, "Kodaira dimension of algebraic fiber spaces over curves", *Invent. Math.* **66**:1 (1982), 57–71. MR 83h:14025 Zbl 0461.14004

[Keller 2009] J. Keller, "Ricci iterations on Kähler classes", *J. Inst. Math. Jussieu* **8**:4 (2009), 743–768. MR 2010h:32026 Zbl 1178.32020

[Keller and Lukic 2009] J. Keller and S. Lukic, "Numerical Weil–Petersson metrics on moduli spaces of Calabi–Yau manifolds", preprint, 2009. arXiv :0907.1387

[Koiso 1983] N. Koiso, "Einstein metrics and complex structures", *Invent. Math.* **73**:1 (1983), 71–106. MR 85d:58018 Zbl 0515.53040

[Lazarsfeld 2004] R. Lazarsfeld, *Positivity in algebraic geometry, I: Classical setting: Line bundles and linear series*, Ergeb. Math. Grenzgeb. (3) **48**, Springer, Berlin, 2004. MR 2005k:14001a Zbl 1093.14501

[Liu et al. 2008] K. Liu, X. Sun, and S.-T. Yau, "Good geometry on the curve moduli", *Publ. Res. Inst. Math. Sci.* **44**:2 (2008), 699–724. MR 2009h:32020 Zbl 1219.14012

[Mabuchi 1986] T. Mabuchi, "K-energy maps integrating Futaki invariants", *Tohoku Math. J.* (2) **38**:4 (1986), 575–593. MR 88b:53060 Zbl 0619.53040

[Narasimhan and Simha 1968] M. S. Narasimhan and R. R. Simha, "Manifolds with ample canonical class", *Invent. Math.* **5** (1968), 120–128. MR 38 #5253 Zbl 0159.37902

[Phong and Sturm 2004] D. H. Phong and J. Sturm, "Scalar curvature, moment maps, and the Deligne pairing", *Amer. J. Math.* **126**:3 (2004), 693–712. MR 2005b:53137 Zbl 1077.53068

[Phong et al. 2007] D. H. Phong, N. Sesum, and J. Sturm, "Multiplier ideal sheaves and the Kähler–Ricci flow", *Comm. Anal. Geom.* **15**:3 (2007), 613–632. MR 2009a:32037 Zbl 1143.53064

[Phong et al. 2008] D. H. Phong, J. Song, J. Sturm, and B. Weinkove, "The Moser–Trudinger inequality on Kähler–Einstein manifolds", *Amer. J. Math.* **130**:4 (2008), 1067–1085. MR 2009e:32027 Zbl 1158.58005

[Protter and Weinberger 1967] M. H. Protter and H. F. Weinberger, *Maximum principles in differential equations*, Prentice-Hall, Englewood Cliffs, NJ, 1967. MR 36 #2935 Zbl 0153.13602

[Rubinstein 2008] Y. A. Rubinstein, "Some discretizations of geometric evolution equations and the Ricci iteration on the space of Kähler metrics", *Adv. Math.* **218**:5 (2008), 1526–1565. MR 2009m:32045 Zbl 1143.53065

[Sano 2006] Y. Sano, "Numerical algorithm for finding balanced metrics", *Osaka J. Math.* **43**:3 (2006), 679–688. MR 2008i:32031 Zbl 1108.32014

[Schumacher 2008] G. Schumacher, "Positivity of relative canonical bundles for families of canonically polarized manifolds", preprint, 2008. arXiv 0808.3259

[Siu 1986] Y. T. Siu, "Curvature of the Weil–Petersson metric in the moduli space of compact Kähler–Einstein manifolds of negative first Chern class", pp. 261–298 in *Contributions to several complex variables* (Notre Dame, IN, 1984), edited by A. Howard and P.-M. Wong, Aspects Math. **E9**, Vieweg, Braunschweig, 1986. MR 87k:32042 Zbl 0598.32020

[Siu 1998] Y.-T. Siu, "Invariance of plurigenera", *Invent. Math.* **134**:3 (1998), 661–673. MR 99i:32035 Zbl 0955.32017

[Song and Tian 2009] J. Song and G. Tian, "The Kahler–Ricci flow through singularities", preprint, 2009. arXiv 0909.4898

[Song and Tian 2012] J. Song and G. Tian, "Canonical measures and Kähler–Ricci flow", *J. Amer. Math. Soc.* **25**:2 (2012), 303–353. MR 2869020 Zbl 1239.53086

[Song and Weinkove 2010] J. Song and B. Weinkove, "Constructions of Kähler–Einstein metrics with negative scalar curvature", *Math. Ann.* **347**:1 (2010), 59–79. MR 2011b:32037 Zbl 1205.53077

[Tian 1987] G. Tian, "Smoothness of the universal deformation space of compact Calabi–Yau manifolds and its Petersson–Weil metric", pp. 629–646 in *Mathematical aspects of string theory* (San Diego, CA, 1986), edited by S.-T. Yau, Adv. Ser. Math. Phys. **1**, World Science Publishing, Singapore, 1987. MR 915841 Zbl 0696.53040

[Tian 2000] G. Tian, *Canonical metrics in Kähler geometry*, Birkhäuser, Basel, 2000. MR 2001j:32024 Zbl 0978.53002

[Tian and Zhu 2007] G. Tian and X. Zhu, "Convergence of Kähler–Ricci flow", *J. Amer. Math. Soc.* **20**:3 (2007), 675–699. MR 2007k:53107 Zbl 1185.53078

[Todorov 1989] A. N. Todorov, "The Weil–Petersson geometry of the moduli space of SU($n \geq 3$) (Calabi–Yau) manifolds, I", *Comm. Math. Phys.* **126**:2 (1989), 325–346. MR 91f:32022 Zbl 0688.53030

[Tsuji 2006] H. Tsuji, "Dynamical construction of Kähler–Einstein metrics", preprint, 2006. arXiv math/0606626

[Tsuji 2011] H. Tsuji, "Canonical singular Hermitian metrics on relative canonical bundles", *Amer. J. Math.* **133**:6 (2011), 1469–1501. MR 2863368 Zbl 05994418

[Voisin 2007] C. Voisin, *Hodge theory and complex algebraic geometry*, vol. 1, Cambridge Studies in Advanced Mathematics **76**, Cambridge University Press, 2007. MR 2009j:32014 Zbl 1129.14019

[Wang 2005] X. Wang, "Canonical metrics on stable vector bundles", *Comm. Anal. Geom.* **13**:2 (2005), 253–285. MR 2006b:32031 1108.32014 Zbl 1093.32008

[Yau 1978] S. T. Yau, "On the Ricci curvature of a compact Kähler manifold and the complex Monge–Ampère equation, I", *Comm. Pure Appl. Math.* **31**:3 (1978), 339–411. MR 81d:53045 Zbl 0369.53059

[Yau 1987] S.-T. Yau, "Nonlinear analysis in geometry", *Enseign. Math.* (2) **33**:1-2 (1987), 109–158. MR 88g:58003 Zbl 0631.53002

[Zelditch 2009] S. Zelditch, "Book review of *Holomorphic Morse inequalities and Bergman kernels* by Xiaonan Ma and George Marinescu; MR2339952", *Bull. Amer. Math. Soc.* (*N.S.*) **46**:2 (2009), 349–361. MR 2729895

Received 21 Aug 2011. Revised 15 Nov 2011. Accepted 20 Dec 2011.

ROBERT J. BERMAN: robertb@chalmers.se
Department of Mathematical Sciences, Chalmers University of Technology and the University of Gothenburg, SE-SE-412 96 Göteborg, Sweden

ANALYSIS AND PDE
Vol. 6, No. 1, 2013
dx.doi.org/10.2140/apde.2013.6.181

RESOLVENT ESTIMATES FOR ELLIPTIC QUADRATIC DIFFERENTIAL OPERATORS

MICHAEL HITRIK, JOHANNES SJÖSTRAND AND JOE VIOLA

Sharp resolvent bounds for nonselfadjoint semiclassical elliptic quadratic differential operators are established, in the interior of the range of the associated quadratic symbol.

1. Introduction and statement of result

It is well known that the spectrum of a nonselfadjoint operator does not control its resolvent, and that the latter may become very large even far from the spectrum. Understanding the behavior of the norm of the resolvent of a given nonselfadjoint operator is therefore a natural and basic problem, which has recently received considerable attention, in particular, within the circle of questions around the notion of the pseudospectrum [Trefethen and Embree 2005]. Some general upper bounds on resolvents are provided by abstract operator theory, and, restricting our attention to the setting of semiclassical pseudodifferential operators on $\mathbb{R}^n$, relevant for this note, we recall a rough statement of such bounds, following [Dencker et al. 2004; Markus 1988; Viola 2012]. Assume that $P = p^w(x, hD_x)$ is the semiclassical Weyl quantization on $\mathbb{R}^n$ of a complex-valued smooth symbol p with $\operatorname{Re} p \geq 0$, belonging to a suitable symbol class and satisfying an ellipticity condition at infinity, guaranteeing that the spectrum of P is discrete in a small neighborhood of the origin. Then the norm of the L^2-resolvent of P is bounded from above by a quantity of the form $\mathcal{O}(1)\exp(\mathcal{O}(1)h^{-n})$, provided that $z \in \operatorname{neigh}(0, \mathbb{C})$ is not too close to the spectrum of P. On the other hand, the available lower bounds on the resolvent of P, coming from the pseudospectral considerations, are typically of the form $C_N^{-1}h^{-N}$, $N \in \mathbb{N}$, or $(1/C)e^{1/(Ch)}$, provided that p enjoys some analyticity properties [Dencker et al. 2004]. Therefore, there appears to be a substantial gap between the available upper and lower bounds on the resolvent, especially when $n \geq 2$. The purpose of this note is to address the issue of bridging this gap in the particular case of an elliptic quadratic semiclassical differential operator on $\mathbb{R}^n$, and to establish a sharp upper bound on the norm of its resolvent.

Sjöstrand has benefited from support from the Agence Nationale de la Recherche under the references JC05-52556 and ANR-08-BLAN-0228-01, as well as a FABER grant from the Conseil régional de Bourgogne.
MSC2010: 15A63, 35P05, 47A10, 53D22.
Keywords: nonselfadjoint operator, resolvent estimate, spectrum, quadratic differential operator, FBI-Bargmann transform.

Let q be a complex-valued quadratic form:

$$q:\mathbb{R}^n_x\times\mathbb{R}^n_\xi\to\mathbb{C},\qquad (x,\xi)\mapsto q(x,\xi). \tag{1-1}$$

We shall assume throughout the following discussion that the quadratic form q is elliptic on $\mathbb{R}^{2n}$, in the sense that $q(X)=0$, $X\in\mathbb{R}^{2n}$, precisely when $X=0$. In this case, according to Lemma 3.1 of [Sjöstrand 1974], if $n>1$, there exists $\lambda\in\mathbb{C}$, $\lambda\neq 0$, such that $\operatorname{Re}(\lambda q)$ is positive definite. In the case when $n=1$, the same conclusion holds, provided that the range of q on $\mathbb{R}^2$ is not all of $\mathbb{C}$ [Sjöstrand 1974; Hitrik 2004], which we assume in what follows. After a multiplication of q by λ, we may and do henceforth assume that $\lambda=1$, so that

$$\operatorname{Re} q>0. \tag{1-2}$$

It follows that the range $\Sigma(q)=q(\mathbb{R}^{2n})$ of q on $\mathbb{R}^{2n}$ is a closed angular sector with a vertex at zero, contained in the union of $\{0\}$ and the open right half-plane.

Associated to the quadratic form q is the semiclassical Weyl quantization $q^w(x,hD_x)$, $0<h\le 1$, which we view as a closed densely defined operator on $L^2(\mathbb{R}^n)$, equipped with the domain

$$\{u\in L^2(\mathbb{R}^n): q^w(x,hD_x)u\in L^2(\mathbb{R}^n)\}.$$

The spectrum of $q^w(x,hD_x)$ is discrete, and following [Sjöstrand 1974], we shall now recall its explicit description. See also [Boutet de Monvel 1974]. To that end, let us introduce the Hamilton map F of q,

$$F:\mathbb{C}^{2n}\to\mathbb{C}^{2n},$$

defined by the identity

$$q(X,Y)=\sigma(X,FY),\qquad X,Y\in\mathbb{C}^{2n}. \tag{1-3}$$

Here the left-hand side is the polarization of q, viewed as a symmetric bilinear form on $\mathbb{C}^{2n}$, and σ is the complex symplectic form on $\mathbb{C}^{2n}$. We notice that the Hamilton map F is skew-symmetric with respect to σ, and, furthermore,

$$FY=\tfrac{1}{2}H_q(Y), \tag{1-4}$$

where $H_q=q'_\xi\cdot\partial_x-q'_x\cdot\partial_\xi$ is the Hamilton field of q.

The ellipticity condition (1-2) implies that the spectrum of the Hamilton map F avoids the real axis, and, in general, we know from Section 21.5 of [Hörmander 1985] that if λ is an eigenvalue of F, so is $-\lambda$, and the algebraic multiplicities agree. Let $\lambda_1,\dots,\lambda_n$ be the eigenvalues of F, counted according to their multiplicity, such that $\lambda_j/i\in\Sigma(q)$, $j=1,\dots,n$. Then the spectrum of the operator $q^w(x,hD_x)$ is given by the eigenvalues of the form

$$h\sum_{j=1}^n\frac{\lambda_j}{i}(2\nu_{j,\ell}+1),\qquad \nu_{j,\ell}\in\mathbb{N}\cup\{0\}. \tag{1-5}$$

We notice that $\operatorname{Spec}(q^w(x,hD_x))\subset\Sigma(q)$, and from [Pravda-Starov 2007] we also know that

$$\operatorname{Spec}(q^w(x,hD_x))\cap\partial\Sigma(q)=\varnothing,$$

provided that the operator $q^w(x,hD_x)$ is not normal.

Here is the main result of this work.

Theorem 1.1. *Let* $q : \mathbb{R}^n_x \times \mathbb{R}^n_\xi \to \mathbb{C}$ *be a quadratic form such that* $\operatorname{Re} q$ *is positive definite. Let* $\Omega \Subset \mathbb{C}$. *There exists* $h_0 > 0$, *and for every* $C > 0$ *there exists* $A > 0$ *such that*

$$\|(q^w(x, hD_x) - z)^{-1}\|_{\mathscr{L}(L^2(\mathbb{R}^n), L^2(\mathbb{R}^n))} \leq A \exp(Ah^{-1}), \tag{1-6}$$

for all $h \in (0, h_0]$, *and all* $z \in \Omega$, *with* $\operatorname{dist}\big(z, \operatorname{Spec}(q^w(x, hD_x))\big) \geq 1/C$. *Furthermore, for all* $C > 0$, $L \geq 1$, *there exists* $A > 0$ *such that for* $h \in (0, h_0]$, *we have*

$$\|(q^w(x, hD_x) - z)^{-1}\|_{\mathscr{L}(L^2(\mathbb{R}^n), L^2(\mathbb{R}^n))} \leq A \exp\Big(Ah^{-1} \log \frac{1}{h}\Big), \tag{1-7}$$

if the spectral parameter $z \in \Omega$ *is such that*

$$\operatorname{dist}\big(z, \operatorname{Spec}(q^w(x, hD_x))\big) \geq h^L/C.$$

Remark 1.2. Assume that the elliptic quadratic form q, with $\operatorname{Re} q > 0$, is such that the Poisson bracket $\{\operatorname{Re} q, \operatorname{Im} q\}$ does not vanish identically, and let $z \in \Sigma(q)^o$, $z \notin \operatorname{Spec}(q^w(x, hD_x))$. Here $\Sigma(q)^o$ is the interior of $\Sigma(q)$. Then it follows from the results of [Dencker et al. 2004; Pravda-Starov 2008] that we have the following lower bound for $(q^w(x, hD_x) - z)^{-1}$, as $h \to 0$:

$$\|(q^w(x, hD_x) - z)^{-1}\|_{\mathscr{L}(L^2(\mathbb{R}^n), L^2(\mathbb{R}^n))} \geq \frac{1}{C_0} e^{1/(C_0 h)}, \quad C_0 > 0.$$

It follows that the upper bound (1-6) is of the right order of magnitude, when $z \in \Sigma(q)^o \cap \Omega$, $|z| \sim 1$, avoids a closed cone $\subset \Sigma(q) \cup \{0\}$, containing the spectrum of $q^w(x, hD_x)$.

Remark 1.3. In Section 4, we give a simple example of an elliptic quadratic operator on $\mathbb{R}^2$, for which the associated Hamilton map has a nonvanishing nilpotent part in its Jordan decomposition, and whose resolvent exhibits the superexponential growth given by the right-hand side of (1-7), in the region of the complex spectral plane where $|z| \sim 1$, $\operatorname{dist}\big(z, \operatorname{Spec}(q^w(x, hD_x))\big) \sim h$. On the other hand, sharper resolvent estimates can be obtained when the Hamilton map F of q is diagonalizable. In this case, we shall see in Section 4 that the bound (1-7) improves to the following, when $z \in \Omega$ and $h \in (0, h_0]$:

$$\|(q^w(x, hD_x) - z)^{-1}\|_{\mathscr{L}(L^2(\mathbb{R}^n), L^2(\mathbb{R}^n))} \leq \frac{Ae^{A/h}}{\operatorname{dist}\big(z, \operatorname{Spec}(q^w(x, hD_x))\big)}. \tag{1-8}$$

Remark 1.4. Let $z_0 \in \operatorname{Spec}(q^w(x, hD_x)) \cap \Omega$ and let

$$\Pi_{z_0} = \frac{1}{2\pi i} \int_{\partial D} (z - q^w(x, hD_x))^{-1} \, dz$$

be the spectral projection of $q^w(x, hD_x)$, associated to the eigenvalue z_0. Here $D \subset \Omega$ is a small open disc centered at z_0, such that the closure $\overline{D}$ avoids the set $\operatorname{Spec}(q^w(x, hD_x)) \backslash \{z_0\}$, and ∂D is its positively oriented boundary. Assume for simplicity that the quadratic form q is such that its Hamilton map is diagonalizable. Then it follows from (1-8) that

$$\Pi_{z_0} = \mathcal{O}(1) \exp(\mathcal{O}(1) h^{-1}) : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n).$$

In the context of elliptic quadratic differential operators in dimension one, resolvent bounds have been studied in particular, in [Boulton 2002; Davies 2000; Davies and Kuijlaars 2004]. We should also mention the general resolvent estimates in [Dencker et al. 2004; Sjöstrand 2010], valid for h-pseudodifferential operators when the spectral parameter is close to the boundary of the range of the corresponding symbol.

The plan of this note is as follows. In Section 2, we make an essentially well-known reduction of our problem to the setting of a quadratic differential operator acting in a Bargmann space of holomorphic functions, convenient for the subsequent analysis. Section 3 is devoted to suitable a priori elliptic estimates, valid for holomorphic functions vanishing to a high, h-dependent order at the origin. The proof of Theorem 1.1 is completed in Section 4 by some elementary considerations in the space of holomorphic polynomials on $\mathbb{C}^n$, of degree not exceeding $\mathcal{O}(h^{-1})$.

2. The normal form reduction

We shall be concerned here with a quadratic form $q : T^*\mathbb{R}^n \to \mathbb{C}$, such that $\operatorname{Re} q$ is positive definite. Let F be the Hamilton map of q, introduced in (1-3). When $\lambda \in \operatorname{Spec}(F)$, we let

$$V_\lambda = \operatorname{Ker}((F-\lambda)^{2n}) \subset T^*\mathbb{C}^n \tag{2-1}$$

be the generalized eigenspace belonging to the eigenvalue λ. The symplectic form σ is then nondegenerate viewed as a bilinear form on $V_\lambda \times V_{-\lambda}$.

We introduce the stable outgoing manifold for the Hamilton flow of the quadratic form $i^{-1}q$, given by

$$\Lambda^+ := \bigoplus_{\operatorname{Im}\lambda>0} V_\lambda \subset T^*\mathbb{C}^n. \tag{2-2}$$

It is then true that Λ^+ is a complex Lagrangian plane such that q vanishes along Λ^+, and Proposition 3.3 of [Sjöstrand 1974] states that the complex Lagrangian Λ^+ is strictly positive in the sense that

$$\frac{1}{i}\sigma(X,\overline{X}) > 0, \quad 0 \neq X \in \Lambda^+. \tag{2-3}$$

We also define

$$\Lambda^- = \bigoplus_{\operatorname{Im}\lambda<0} V_\lambda \subset T^*\mathbb{C}^n, \tag{2-4}$$

which is a complex Lagrangian plane such that q vanishes along Λ^-, and from the arguments of [Sjöstrand 1974] we also know that Λ^- is strictly negative in the sense that

$$\frac{1}{i}\sigma(X,\overline{X}) < 0, \quad 0 \neq X \in \Lambda^-. \tag{2-5}$$

The complex Lagrangians Λ^+ and Λ^- are transversal, and following [Helffer and Sjöstrand 1984; Sjöstrand 1987], we would like to implement a reduction of the quadratic form q to a normal form by applying a linear complex canonical transformation which reduces Λ^+ to $\{(x,\xi) \in \mathbb{C}^{2n} : \xi = 0\}$ and Λ^- to $\{(x,\xi) \in \mathbb{C}^{2n} : x = 0\}$. We shall then be able to implement the canonical transformation in question by an FBI–Bargmann transform. Let us first simplify q by means of a suitable real linear canonical transformation. When doing so, we observe that the fact that the Lagrangian Λ^- is strictly negative implies that it is of the form

$$\eta = A_- y, \quad y \in \mathbb{C}^n,$$

where the complex symmetric $n \times n$ matrix A_- is such that $\operatorname{Im} A_- < 0$. Here (y, η) are the standard canonical coordinates on $T^*\mathbb{R}^n_y$ that we extend to the complexification $T^*\mathbb{C}^n_y$. Using the real linear canonical transformation $(y, \eta) \mapsto (y, \eta - (\operatorname{Re} A_-)y)$, we reduce Λ^- to the form $\eta = i \operatorname{Im} A_- y$, and by a diagonalization of $\operatorname{Im} A_-$, we obtain the standard form $\eta = -iy$. After this real linear symplectic change of coordinates and the conjugation of the semiclassical Weyl quantization $q^w(x, hD_x)$ of q by means of the corresponding unitary metaplectic operator, we may assume that Λ^- is of the form

$$\eta = -iy, \quad y \in \mathbb{C}^n, \tag{2-6}$$

while the positivity property of the complex Lagrangian Λ^+ is unaffected, so that, in the new real symplectic coordinates, extended to the complexification, Λ^+ is of the form

$$\eta = A_+ y, \quad \operatorname{Im} A_+ > 0. \tag{2-7}$$

Let

$$B = B_+ = (1 - iA_+)^{-1} A_+, \tag{2-8}$$

and notice that the matrix B is symmetric. Let us introduce the FBI–Bargmann transform

$$Tu(x) = Ch^{-3n/4} \int e^{i\varphi(x,y)/h} u(y)\, dy, \quad x \in \mathbb{C}^n, \quad C > 0, \tag{2-9}$$

where

$$\varphi(x, y) = \frac{i}{2}(x - y)^2 - \frac{1}{2}(Bx, x). \tag{2-10}$$

The associated complex linear canonical transformation on $\mathbb{C}^{2n}$

$$\kappa_T : (y, -\varphi'_y(x, y)) \mapsto (x, \varphi'_x(x, y)) \tag{2-11}$$

is of the form

$$\kappa_T : (y, \eta) \mapsto (x, \xi) = (y - i\eta, \eta + iB\eta - By), \tag{2-12}$$

and we see that the image of $\Lambda_- : \eta = -iy$ under κ_T is the fiber $\{(x, \xi) \in \mathbb{C}^{2n} : x = 0\}$, while $\kappa_T(\Lambda^+)$ is given by the equation $\{(x, \xi) \in \mathbb{C}^{2n} : \xi = 0\}$.

We know from [Sjöstrand 1996] that for a suitable choice of $C > 0$ in (2-9), the map T is unitary:

$$T : L^2(\mathbb{R}^n) \to H_{\Phi_0}(\mathbb{C}^n), \tag{2-13}$$

where

$$H_{\Phi_0}(\mathbb{C}^n) = \operatorname{Hol}(\mathbb{C}^n) \cap L^2(\mathbb{C}^n : e^{-2\Phi_0/h} L(dx)),$$

and Φ_0 is a strictly plurisubharmonic quadratic form on $\mathbb{C}^n$, given by

$$\Phi_0(x) = \sup_{y \in \mathbb{R}^n}(-\operatorname{Im} \varphi(x, y)) = \tfrac{1}{2}((\operatorname{Im} x)^2 + \operatorname{Im}(Bx, x)). \tag{2-14}$$

We also recall [Sjöstrand 1996] that the canonical transformation κ_T in (2-11) maps $\mathbb{R}^{2n}$ bijectively onto

$$\Lambda_{\Phi_0} := \left\{ \left(x, \frac{2}{i} \frac{\partial \Phi_0}{\partial x}(x) \right) : x \in \mathbb{C}^n \right\}. \tag{2-15}$$

As explained in Chapter 11 of [Sjöstrand 1982], the strict positivity of $\kappa_T(\Lambda^+) = \{(x, \xi) \in \mathbb{C}^{2n} : \xi = 0\}$ with respect to Λ_{Φ_0} implies that the quadratic weight function Φ_0 is strictly convex, so that

$$\Phi_0(x) \sim |x|^2, \quad x \in \mathbb{C}^n. \tag{2-16}$$

Next we have the exact Egorov property [Sjöstrand 1996],

$$T q^w(y, hD_y) u = \tilde{q}^w(x, hD_x) T u, \quad u \in \mathscr{S}(\mathbb{R}^n), \tag{2-17}$$

where $\tilde{q}$ is a quadratic form on $\mathbb{C}^{2n}$ given by $\tilde{q} = q \circ \kappa_T^{-1}$. Therefore it follows that

$$\tilde{q}(x, \xi) = Mx \cdot \xi, \tag{2-18}$$

where M is a complex $n \times n$ matrix. We have

$$H_{\tilde{q}} = Mx \cdot \partial_x - M^t \xi \cdot \partial_\xi,$$

and using (1-4), we see that the corresponding Hamilton map

$$\widetilde{F} = \tfrac{1}{2} \begin{pmatrix} M & 0 \\ 0 & -M^t \end{pmatrix}$$

maps $(x, 0) \in \kappa_T(\Lambda^+)$ to $(1/2)(Mx, 0)$. Now F and $\widetilde{F}$ are isospectral, and we conclude that, with the agreement of algebraic multiplicities, the following holds:

$$\operatorname{Spec}(M) = \operatorname{Spec}(2F) \cap \{\operatorname{Im} \lambda > 0\}. \tag{2-19}$$

Therefore the problem of estimating the norm of the resolvent of $q^w(x, hD_x)$ on $L^2(\mathbb{R}^n)$ is equivalent to controlling the norm of the resolvent of the quadratic operator $\tilde{q}^w(x, hD_x)$, acting in the space $H_{\Phi_0}(\mathbb{C}^n)$, where the quadratic weight Φ_0 enjoys the property (2-16).

In what follows, it will be convenient to reduce the matrix M in (2-18) to its Jordan normal form. To this end, let us notice that we can implement this reduction by considering a complex canonical transformation of the form

$$\kappa_C : \mathbb{C}^{2n} \ni (x, \xi) \mapsto (C^{-1} x, C^t \xi) \in \mathbb{C}^{2n}, \tag{2-20}$$

where C is a suitable invertible complex $n \times n$ matrix. On the operator level, associated to the transformation in (2-20), we have the operator $u(x) \mapsto |\det C| u(Cx)$, which maps the space $H_{\Phi_0}(\mathbb{C}^n)$ unitarily onto the space $H_{\Phi_1}(\mathbb{C}^n)$, where $\Phi_1(x) = \Phi_0(Cx)$ is a strictly plurisubharmonic quadratic weight such that $\kappa_C(\Lambda_{\Phi_0}) = \Lambda_{\Phi_1}$. We notice that the property

$$\Phi_1(x) \sim |x|^2, \quad x \in \mathbb{C}^n \tag{2-21}$$

remains valid.

We summarize the discussion pursued in this section in the following result.

Proposition 2.1. *Let $q : \mathbb{R}^n_x \times \mathbb{R}^n_\xi \to \mathbb{C}$ be a quadratic form with $\operatorname{Re} q > 0$. The operator*

$$q^w(x, hD_x) : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n),$$

equipped with the domain

$$\mathcal{D}(q^w(x, hD_x)) = \{u \in L^2(\mathbb{R}^n) : (x^2 + (hD_x)^2)u \in L^2(\mathbb{R}^n)\},$$

is unitarily equivalent to the quadratic operator

$$\tilde{q}^w(x, hD_x) : H_{\Phi_1}(\mathbb{C}^n) \to H_{\Phi_1}(\mathbb{C}^n),$$

with the domain

$$\mathcal{D}(\tilde{q}^w(x, hD_x)) = \{u \in H_{\Phi_1}(\mathbb{C}^n) : (1 + |x|^2)u \in L^2_{\Phi_1}(\mathbb{C}^n)\}.$$

Here

$$\tilde{q}(x, \xi) = Mx \cdot \xi,$$

where M is a complex $n \times n$ block-diagonal matrix, each block being a Jordan matrix. The eigenvalues of M are precisely those of $2F$ in the upper half-plane, and the quadratic weight function $\Phi_1(x)$ satisfies

$$\Phi_1(x) \sim |x|^2, \quad x \in \mathbb{C}^n.$$

We have the ellipticity property

$$\operatorname{Re} \tilde{q}\left(x, \frac{2}{i} \frac{\partial \Phi_1}{\partial x}(x)\right) \sim |x|^2, \quad x \in \mathbb{C}^n. \tag{2-22}$$

Remark 2.2. The normal form reduction described in Proposition 2.1 is close to the corresponding discussion of Section 3 in [Sjöstrand 1974]. Here, for future computations, it will be convenient for us to work in the Bargmann space $H_{\Phi_1}(\mathbb{C}^n)$.

3. An elliptic estimate

Following the reduction of Proposition 2.1, here we concern ourselves with the quadratic operator $\tilde{q}^w(x, hD_x)$, acting on $H_{\Phi_1}(\mathbb{C}^n)$. The purpose of this section is to establish a suitable a priori estimate for holomorphic functions, vanishing to a high, h-dependent order at the origin, instrumental in the proof of Theorem 1.1. The starting point is the following observation, which comes directly from Lemma 4.5 in [Gérard and Sjöstrand 1987], and whose proof we give only for the convenience of the reader.

Lemma 3.1. *Let $u \in \operatorname{Hol}(\mathbb{C}^n)$ and assume that $\partial^\alpha u(0) = 0$, $|\alpha| < N$, and that $0 < C_0 < C_1 < \infty$. Then*

$$\|u\|_{L^\infty(B(0,C_0))} \le \left(N \frac{C_1}{C_1 - C_0}\right)\left(\frac{C_0}{C_1}\right)^N \|u\|_{L^\infty(B(0,C_1))}. \tag{3-1}$$

Here $B(0, C_j) = \{x \in \mathbb{C}^n : |x| \le C_j\}$, $j = 0, 1$.

Proof. By Taylor's formula, we have

$$u(x)=\int_0^1 \frac{(1-t)^{N-1}}{(N-1)!}\Big(\frac{d}{dt}\Big)^N u(tx)\,dt.$$

We may assume that $|x|=C_0$ and apply Cauchy's inequalities so that

$$\Big|\Big(\frac{d}{dt}\Big)^N u(tx)\Big| \le \frac{C_0^N N!}{(C_1-C_0t)^N}\|u\|_{L^\infty(B(0,C_1))}.$$

It suffices therefore to remark that the expression

$$N\int_0^1 \frac{(1-t)^{N-1}}{(C_1/C_0-t)^N}\,dt$$

does not exceed

$$\frac{N}{C_1/C_0-1}\Big(\frac{C_0}{C_1}\Big)^{N-1}.$$

□

Let $K>0$ be fixed and assume that $u\in H_{\Phi_1}(\mathbb{C}^n)$ is such that $\partial^\alpha u(0)=0$, when $|\alpha|<N$. Using Lemma 3.1, we write

$$\begin{aligned}\|u\|^2_{H_{\Phi_1}(B(0,K))} &\le \|u\|^2_{L^2(B(0,K))}\\ &\le \mathcal{O}_K(1)\|u\|^2_{L^\infty(B(0,K))} \le \mathcal{O}_K(1)N^2e^{-2N}\|u\|^2_{L^\infty(B(0,Ke))}\\ &\le \mathcal{O}_K(1)N^2e^{-2N}\|u\|^2_{L^2(B(0,(K+1)e))} \le \mathcal{O}_K(1)N^2e^{-2N}e^{(2/h)C_1(K+1)^2e^2}\|u\|^2_{H_{\Phi_1}}.\end{aligned} \tag{3-2}$$

In the last inequality we used that $\Phi_1(x)\le C_1|x|^2$ for some $C_1\ge 1$. It follows that

$$\|u\|_{H_{\Phi_1}(B(0,K))} \le \mathcal{O}_K(1)e^{-1/2h}\|u\|_{H_{\Phi_1}}, \tag{3-3}$$

provided that the integer N satisfies

$$N \ge \frac{2C_1(K+1)^2e^2+1}{h}. \tag{3-4}$$

In what follows, we shall let $N_0=N_0(K)\in\mathbb{N}$, $N_0\sim h^{-1}$, be the least integer which satisfies (3-4).

It is now easy to derive an a priori estimate for functions in $H_{\Phi_1}(\mathbb{C}^n)$, which vanish to a high order at the origin. Let $\chi\in C_0^\infty(\mathbb{C}^n)$, $0\le\chi\le 1$, be such that $\operatorname{supp}(\chi)\subset\{x\in\mathbb{C}^n : |x|\le K\}$, with $\chi(x)=1$ for $|x|\le K/2$. If $u\in H_{\Phi_1}(\mathbb{C}^n)$ is such that $(1+|x|^2)u\in L^2_{\Phi_1}(\mathbb{C}^n)$, we have the quantization-multiplication formula [Sjöstrand 1990], valid for z in a compact subset of $\mathbb{C}$,

$$\begin{aligned}&((1-\chi)(\tilde{q}^w(x,hD_x)-z)u,u)_{L^2_{\Phi_1}}\\ &\qquad = \int (1-\chi(x))\Big(\tilde{q}\Big(x,\frac{2}{i}\frac{\partial\Phi_1}{\partial x}(x)\Big)-z\Big)|u(x)|^2e^{-2\Phi_1(x)/h}\,L(dx)+\mathcal{O}(h)\|u\|^2_{H_{\Phi_1}}.\end{aligned}$$

The ellipticity property

$$\operatorname{Re}\tilde{q}\Big(x,\frac{2}{i}\frac{\partial\Phi_1}{\partial x}(x)\Big)\ge\frac{|x|^2}{C_0},\qquad x\in\mathbb{C}^n, \tag{3-5}$$

valid for some $C_0>1$, implies that, on the support of $1-\chi$, we have

$$\mathrm{Re}\left(\tilde{q}\left(x, \frac{2}{i}\frac{\partial \Phi_1}{\partial x}(x)\right) - z\right) \geq \frac{|x|^2}{2C_0},$$

provided that $|z| \leq K^2/8C_0$. Restricting the attention to this range of z's and using the Cauchy–Schwarz inequality, we obtain that

$$\int (1-\chi(x))|u(x)|^2 e^{-2\Phi_1(x)/h} L(dx) \leq \mathcal{O}_K(1) \|(\tilde{q}^w(x, hD_x) - z)u\|_{H_{\Phi_1}} \|u\|_{H_{\Phi_1}} + \mathcal{O}_K(h)\|u\|^2_{H_{\Phi_1}}. \quad (3\text{-}6)$$

If $u \in H_{\Phi_1}(\mathbb{C}^n)$, $(1+|x|^2)u \in L^2_{\Phi_1}(\mathbb{C}^n)$, is such that $\partial^\alpha u(0) = 0$ for all $\alpha \in \mathbb{N}^n$ with $|\alpha| < N_0$, an application of (3-3) shows that the left-hand side of (3-6) is of the form

$$\|u\|^2_{H_{\Phi_1}} + \mathcal{O}_K(h^\infty)\|u\|^2_{H_{\Phi_1}}.$$

We may summarize the discussion so far in the following proposition.

Proposition 3.2. *Let $K > 0$ be fixed and assume that $u \in H_{\Phi_1}(\mathbb{C}^n)$, $(1+|x|^2)u \in L^2_{\Phi_1}(\mathbb{C}^n)$, is such that $\partial^\alpha u(0) = 0$, $|\alpha| < N_0$, where $N_0 \sim h^{-1}$ is the least integer such that*

$$N_0 \geq \frac{2C_1(K+1)^2 e^2 + 1}{h}.$$

Here $\Phi_1(x) \leq C_1|x|^2$, $C_1 \geq 1$. Assume also that $|z| \leq K^2/8C_0$, where $C_0 > 1$ is the ellipticity constant in (3-5). *Then we have the following a priori estimate, valid for all $h > 0$ sufficiently small*:

$$\|u\|_{H_{\Phi_1}} \leq \mathcal{O}(1)\|(\tilde{q}^w(x, hD_x) - z)u\|_{H_{\Phi_1}}.$$

We finish this section by discussing norm estimates for the linear continuous projection operator

$$\tau_N : H_{\Phi_1}(\mathbb{C}^n) \to H_{\Phi_1}(\mathbb{C}^n),$$

given by

$$\tau_N u(x) = \sum_{|\alpha|<N} (\alpha!)^{-1}(\partial^\alpha u(0))x^\alpha. \quad (3\text{-}7)$$

As in Proposition 3.2, we shall be concerned with the case when $N \in \mathbb{N}$ satisfies $N \sim h^{-1}$. The projection operator τ_N is highly nonorthogonal — nevertheless, using the strict convexity of the quadratic weight Φ_1, establishing an exponential upper bound on its norm will be quite straightforward, as well as sufficient for our purposes. In the following, we shall use the fact that

$$\frac{1}{C_1}|x|^2 \leq \Phi_1(x) \leq C_1|x|^2, \quad C_1 \geq 1. \quad (3\text{-}8)$$

Notice also that $[\tau_N, \tilde{q}^w(x, hD_x)] = 0$.

Proposition 3.3. *Assume that $N \in \mathbb{N}$ is such that $Nh \leq \mathcal{O}(1)$. There exists a constant $C > 0$ such that*

$$\|\tau_N\|_{\mathcal{L}(H_{\Phi_1}(\mathbb{C}^n), H_{\Phi_1}(\mathbb{C}^n))} \leq Ce^{C/h}. \quad (3\text{-}9)$$

Proof. We first observe that when deriving the bound (3-9), it suffices to restrict the attention to the space of holomorphic polynomials, which is dense in $H_{\Phi_1}(\mathbb{C}^n)$. Indeed, the analysis in [Sjöstrand 1974] tells us that the linear span of the generalized eigenfunctions of the quadratic operator $q^w(x, hD_x)$ is dense in $L^2(\mathbb{R}^n)$, which implies the density of the holomorphic polynomials in $H_{\Phi_1}(\mathbb{C}^n)$. Let

$$u(x) = \sum_{|\alpha| \le N_1} a_\alpha x^\alpha \tag{3-10}$$

for some N_1, where we may assume that $N_1 > N$. We have

$$\tau_N u = \sum_{|\alpha| < N} a_\alpha x^\alpha,$$

and therefore, using (3-8), we see that

$$\|\tau_N u\|^2_{H_{\Phi_1}} \le \|\tau_N u\|^2_{H_{\Phi_\ell}}, \tag{3-11}$$

where $\Phi_\ell(x) = |x|^2/C_1$. When computing the expression in the right-hand side of (3-11), we notice that since Φ_ℓ is radial, we have

$$(x^\alpha, x^\beta)_{H_{\Phi_\ell}} = 0, \quad \alpha \ne \beta,$$

while

$$(x^\alpha, x^\alpha)_{H_{\Phi_\ell}} = \prod_{j=1}^n \int |x_j|^{2\alpha_j} e^{-2|x_j|^2/C_1 h}\, L(dx_j),$$

which is immediately seen to be equal to

$$\left(\frac{C_1 h}{2}\right)^{n+|\alpha|} \pi^n \alpha!.$$

It follows that

$$\|\tau_N u\|^2_{H_{\Phi_1}} \le \sum_{|\alpha| < N} |a_\alpha|^2 \left(\frac{C_1 h}{2}\right)^{n+|\alpha|} \pi^n \alpha!. \tag{3-12}$$

On the other hand, (3-8) also gives that

$$\|u\|^2_{H_{\Phi_1}} \ge \|u\|^2_{H_{\Phi_u}}, \tag{3-13}$$

where $\Phi_u(x) = C_1|x|^2$, and arguing as above, it is straightforward to see that the right-hand side of (3-13) is given by the expression

$$\sum_{|\alpha| \le N_1} |a_\alpha|^2 \left(\frac{h}{2C_1}\right)^{n+|\alpha|} \pi^n \alpha!.$$

We conclude that when $u \in H_{\Phi_1}(\mathbb{C}^n)$ is a holomorphic polynomial of the form (3-10),

$$\|u\|^2_{H_{\Phi_1}} \ge \sum_{|\alpha| < N} |a_\alpha|^2 \left(\frac{h}{2C_1}\right)^{n+|\alpha|} \pi^n \alpha!. \tag{3-14}$$

Combining (3-12), (3-14), and recalling the fact that $Nh \le \mathbb{O}(1)$, we obtain the result of the proposition. □

4. The finite-dimensional analysis and end of the proof

In this section we analyze the resolvent of the quadratic operator $\tilde{q}^w(x, hD_x)$ acting on the finite-dimensional space $\operatorname{Im}\tau_N$, where τ_N is the projection operator introduced in (3-7) and $N \sim h^{-1}$. This will allow us to complete the proof of Theorem 1.1. For $m = 0, 1, \ldots$, define the finite-dimensional subspace $E_m \subset H_{\Phi_1}(\mathbb{C}^n)$ as the linear span of the monomials x^α, with $|\alpha| = m$. We have

$$\operatorname{Im}\tau_N = \bigoplus_{m=0}^{N-1} E_m.$$

We may notice here that

$$\nu_m := \dim E_m = \frac{1}{(n-1)!}(m+1)\cdots(m+n-1), \tag{4-1}$$

and also that each space E_m is invariant under $\tilde{q}^w(x, hD_x)$. We shall equip $\operatorname{Im}\tau_N$ with the basis

$$\varphi_\alpha(x) := (\pi^n\alpha!)^{-1/2}h^{-n/2}(h^{-1/2}x)^\alpha, \quad |\alpha| < N, \tag{4-2}$$

which will be particularly convenient in the following computations, since the normalized monomials φ_α form an orthonormal basis in the weighted space $H_\Phi(\mathbb{C}^n)$, where $\Phi(x) = (1/2)|x|^2$. We have

$$\operatorname{Im}\tau_N \subset H_{\Phi_1}(\mathbb{C}^n) \cap H_\Phi(\mathbb{C}^n),$$

in view of the strict convexity of the weights.

Let us first derive an upper bound on the norm of the inverse of the operator

$$z - \tilde{q}^w(x, hD_x) : E_m \to E_m, \quad 0 \le m < N \sim h^{-1},$$

assuming that E_m has been equipped with the H_Φ-norm. Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of the Hamilton map F of q in the upper half-plane, repeated according to their algebraic multiplicity. According to Proposition 2.1, we then have

$$\tilde{q}^w(x, hD_x) = \tilde{q}^w_D(x, hD_x) + \tilde{q}^w_N(x, hD_x),$$

where

$$\tilde{q}^w_D(x, hD_x) = \sum_{j=1}^n 2\lambda_j x_j hD_{x_j} + \frac{h}{i}\sum_{j=1}^n \lambda_j, \tag{4-3}$$

is the diagonal part, while

$$\tilde{q}^w_N(x, hD_x) = \sum_{j=1}^{n-1} \gamma_j x_{j+1} hD_{x_j}, \quad \gamma_j \in \{0, 1\}, \tag{4-4}$$

is the nilpotent one. It is also easily seen that the operators $\tilde{q}^w_D(x, hD_x)$ and $\tilde{q}^w_N(x, hD_x)$ commute. It will be important for us to have an estimate of the order of nilpotency of the operator $\tilde{q}^w_N(x, hD_x)$ acting on the space E_m.

Lemma 4.1. *Let $n \geq 2$, $m \geq 1$, and let $E_m(n)$ be the space of homogeneous polynomials of degree m in the variables $x_1, x_2, \dots, x_n$. The operator*

$$\mathcal{N} := \sum_{j=1}^{n-1} x_{j+1} \partial_{x_j} : E_m(n) \to E_m(n)$$

is nilpotent of order $m(n-1)+1$.

Proof. When $\alpha = (\alpha_1, \dots, \alpha_n)$, $|\alpha| = m$, let us write

$$S(\alpha) = \sum_{j=1}^{n} j\alpha_j,$$

and notice that $m \leq S(\alpha) \leq nm$. We have

$$\mathcal{N} x^\alpha = \sum_{\substack{|\alpha'|=m \\ S(\alpha')=S(\alpha)+1}} c_{\alpha'} x^{\alpha'},$$

and similarly for powers $\mathcal{N}^p x^\alpha$, but with $S(\alpha') = S(\alpha) + p$. It follows that $\mathcal{N}^{m(n-1)+1} x^\alpha$ must vanish, as

$$S(\alpha') = S(\alpha) + m(n-1) + 1 \geq mn + 1$$

is impossible. We also notice that $\mathcal{N}^{m(n-1)} x_1^m = C x_n^m \neq 0$, for some $C \neq 0$. $\square$

In what follows, we shall only use that the operator $\tilde{q}_N^w(x, hD_x) : E_m \to E_m$ is nilpotent of order $\mathbb{O}(m)$, with the implicit constant depending on the dimension n only.

It is now straightforward to derive a bound on the norm of the inverse of the operator

$$z - \tilde{q}^w(x, hD_x) : E_m \to E_m,$$

when the space E_m is equipped with the H_Φ-norm. The matrix $\mathcal{D}(m)$ of the operator $\tilde{q}_D^w(x, hD_x)$ with respect to the basis φ_α, $|\alpha| = m$, is diagonal, with the eigenvalues of $\tilde{q}^w(x, hD_x)$,

$$\mu_\alpha = \frac{h}{i} \sum_{j=1}^{n} \lambda_j (2\alpha_j + 1), \quad |\alpha| = m,$$

along the diagonal. On the other hand, using (4-2), we compute

$$x_{j+1} \partial_{x_j} \varphi_\alpha = \alpha_j^{1/2} (\alpha_{j+1} + 1)^{1/2} \varphi_{\alpha - e_j + e_{j+1}}, \quad 1 \leq j \leq n-1,$$

where $\alpha = (\alpha_1, \dots, \alpha_n)$ and $e_1, \dots, e_n$ is the canonical basis in $\mathbb{R}^n$. It follows that

$$\tilde{q}_N^w(x, hD_x) \varphi_\alpha = \sum_{j=1}^{n-1} -ih\gamma_j \alpha_j^{1/2} (\alpha_{j+1} + 1)^{1/2} \varphi_{\alpha - e_j + e_{j+1}}, \tag{4-5}$$

and hence the entries $(\mathcal{N}(m)_{\alpha,\beta}) = \big((\tilde{q}_N^w(x, hD_x)\varphi_\beta, \varphi_\alpha)\big)$, $|\alpha| = |\beta| = m$, of the matrix $\mathcal{N}(m) : \mathbb{C}^{\nu_m} \to \mathbb{C}^{\nu_m}$ of $\tilde{q}_N^w(x, hD_x) : E_m \to E_m$ with respect to the basis $\{\varphi_\alpha\}$, are bounded in modulus by

$$h\alpha_j^{1/2} (\alpha_{j+1} + 1)^{1/2} \leq h(m+1) \leq \mathbb{O}(1),$$

since $|\alpha| = m$ and m does not exceed $N = \mathbb{O}(h^{-1})$. Furthermore, from (4-5), it follows that the matrix $\mathcal{N}(m)$ has no more than $n-1$ nonzero entries in any column, and a similar reasoning shows that each row of $\mathcal{N}(m)$ also has no more than $n-1$ nonzero entries. Since we have just seen that the entries in $\mathcal{N}(m)$ are $\mathbb{O}(1)$, an application of Schur's lemma shows that that the operator norm of $\mathcal{N}(m)$ on $\mathbb{C}^{\nu_m}$ does not exceed

$$\Big(\sup_{\beta}\sum_{\alpha}|\mathcal{N}(m)_{\alpha,\beta}|\Big)^{1/2}\Big(\sup_{\alpha}\sum_{\beta}|\mathcal{N}(m)_{\alpha,\beta}|\Big)^{1/2} \leq \mathbb{O}(1).$$

Now the inverse of the $\nu_m \times \nu_m$ matrix

$$z - \mathfrak{D}(m) - \mathcal{N}(m) : \mathbb{C}^{\nu_m} \to \mathbb{C}^{\nu_m}$$

is given by

$$(z - \mathfrak{D}(m))^{-1}\sum_{j=0}^{\infty}\big((z - \mathfrak{D}(m))^{-1}\mathcal{N}(m)\big)^{j}, \tag{4-6}$$

and according to Lemma 4.1 and the fact that $\big[\tilde{q}_D^w(x, hD_x), \tilde{q}_N^w(x, hD_x)\big] = 0$, we know that the Neumann series in (4-6) is finite, containing at most $\mathbb{O}(m)$ terms. It follows that

$$(z - \mathfrak{D}(m) - \mathcal{N}(m))^{-1} = \frac{\exp(\mathbb{O}(m))}{d(z, \sigma_m)^{\mathbb{O}(m)}} : \mathbb{C}^{\nu_m} \to \mathbb{C}^{\nu_m}, \tag{4-7}$$

where $d(z, \sigma_m) = \inf_{|\alpha|=m}|z - \mu_\alpha|$ is the distance from $z \in \mathbb{C}$ to the set of eigenvalues $\{\mu_\alpha\}$ of $\tilde{q}^w(x, hD_x)$, restricted to E_m.

Using the fact that $\operatorname{Im}\tau_N$ is the orthogonal direct sum of the spaces E_m, $0 \leq m \leq N-1$, we may summarize the discussion so far in the following result.

Proposition 4.2. *Assume that $N \in \mathbb{N}$ is such that $Nh \leq \mathbb{O}(1)$, and let us equip the finite-dimensional space $\operatorname{Im}\tau_N \subset H_{\Phi_1}(\mathbb{C}^n) \cap H_\Phi(\mathbb{C}^n)$ with the H_Φ-norm, where $\Phi(x) = (1/2)|x|^2$. Assume that $z \in \mathbb{C}$ satisfies $\operatorname{dist}(z, \operatorname{Spec}(\tilde{q}^w(x, hD_x))) \geq h^L/C$, for some $C > 0$, $L \geq 1$. Then we have*

$$(z - \tilde{q}^w(x, hD_x))^{-1} = \mathbb{O}(1)\exp\Big(\mathbb{O}(1)h^{-1}\log\frac{1}{h}\Big) : \operatorname{Im}\tau_N \to \operatorname{Im}\tau_N. \tag{4-8}$$

Assuming that $\operatorname{dist}\big(z, \operatorname{Spec}(\tilde{q}^w(x, hD_x))\big) \geq 1/C$, the bound (4-8) improves to

$$(z - \tilde{q}^w(x, hD_x))^{-1} = \mathbb{O}(1)\exp(\mathbb{O}(1)h^{-1}) : \operatorname{Im}\tau_N \to \operatorname{Im}\tau_N. \tag{4-9}$$

Remark 4.3. Assume that the quadratic form q is such that the nilpotent part in the Jordan decomposition of the Hamilton map F is trivial. The quadratic operator $\tilde{q}^w(x, hD_x)$ acting on $H_\Phi(\mathbb{C}^n)$ is then normal, and therefore, the estimate (4-8) improves to

$$\|(z - \tilde{q}^w(x, hD_x))^{-1}\|_{\mathcal{L}(\operatorname{Im}\tau_N, \operatorname{Im}\tau_N)} \leq \frac{1}{\operatorname{dist}\big(z, \operatorname{Spec}(\tilde{q}^w(x, hD_x))\big)}.$$

Example 4.4. Let $n = 2$. Consider the semiclassical Weyl quantization of the elliptic quadratic form

$$\tilde{q}(x, \xi) = 2\lambda\sum_{j=1}^{2}x_j\xi_j + x_2\xi_1, \quad \lambda = \frac{i}{2},$$

acting on $H_\Phi(\mathbb{C}^2)$. The eigenvalues of $\tilde{q}^w(x, hD_x)$ are of the form $\mu_\alpha = h(|\alpha|+1)$, $|\alpha| \geq 0$, and writing

$$\tilde{q}^w_D(x, hD_x) = 2\lambda \sum_{j=1}^{2} x_j hD_{x_j} + \frac{2\lambda h}{i}, \quad \tilde{q}^w_N(x, hD_x) = x_2 hD_{x_1},$$

we have

$$\tilde{q}^w_D(x, hD_x)\varphi_\alpha = \mu_\alpha \varphi_\alpha,$$

and

$$\tilde{q}^w_N(x, hD_x)\varphi_\alpha = -ih(\alpha_1(\alpha_2+1))^{1/2}\varphi_{\alpha-e_1+e_2}, \tag{4-10}$$

where the φ_α were introduced in (4-2).

Let $|\alpha| = m$, and let us write, following (4-6),

$$(\tilde{q}^w(x, hD_x) - z)^{-1}\varphi_\alpha = (\mu_\alpha - z)^{-1} \sum_{j=0}^{m} (\mu_\alpha - z)^{-j} (\tilde{q}^w_N(x, hD_x))^j \varphi_\alpha. \tag{4-11}$$

It is then natural to take $\alpha = (m, 0)$, and using (4-10), a straightforward computation shows that, for $0 \leq j \leq m$,

$$(\tilde{q}^w_N(x, hD_x))^j \varphi_{(m,0)} = (-ih)^j \sqrt{\frac{j!\,m!}{(m-j)!}}\,\varphi_{(m-j,j)}.$$

Let $z = 1$ and take $m = h^{-1} \in \mathbb{N}$ so that $\mu_\alpha - z = h$. By Parseval's formula,

$$\|(\tilde{q}^w(x, hD_x) - z)^{-1}\varphi_{(m,0)}\|^2_{H_\Phi} = \sum_{j=0}^{m} h^{-2}h^{-2j}h^{2j} \frac{j!\,m!}{(m-j)!}, \tag{4-12}$$

and the right-hand side can be estimated from below simply by discarding all terms except when $j = m$. An application of Stirling's formula shows that

$$\|(\tilde{q}^w(x, hD_x) - z)^{-1}\varphi_{(m,0)}\|_{H_\Phi} \geq m! \geq \exp\Big(\frac{1}{2h}\log\frac{1}{h}\Big),$$

for all $h > 0$ sufficiently small, and therefore, we see that the result of Proposition 4.2 cannot be improved. Let us finally notice that, as can be checked directly, the quadratic operator $\tilde{q}^w(x, hD_x)$ acting on $H_\Phi(\mathbb{C}^2)$ is unitarily equivalent, via an FBI-Bargmann transform, to the quadratic operator

$$q(x, hD_x) : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n),$$

of the form

$$q(x, hD_x) = q_0(x, hD_x) - \frac{i}{2} a_2^* a_1,$$

where

$$q_0(x, hD_x) = -\tfrac{1}{2}h^2\Delta + \tfrac{1}{2}x^2 = \tfrac{1}{2}(a_1^* a_1 + a_2^* a_2) + h$$

is the semiclassical harmonic oscillator, while

$$a_j^* = x_j - h\partial_{x_j}, \quad a_j = x_j + h\partial_{x_j}, \quad j = 1, 2$$

are the creation and annihilation operators, respectively. See also [Caliceti et al. 2007].

We shall now complete the proof of Theorem 1.1 in a straightforward manner, combining our earlier computations and estimates. Elementary considerations analogous to those used in the proof of Proposition 3.3 show that for some constant $C > 0$, we have, when $u \in \operatorname{Im} \tau_N$,

$$\|u\|_{H_{\Phi_1}} \le Ce^{C/h}\|u\|_{H_\Phi}, \quad \|u\|_{H_\Phi} \le Ce^{C/h}\|u\|_{H_{\Phi_1}}. \tag{4-13}$$

Here we recall that $N \sim h^{-1}$. It follows therefore that the result of Proposition 4.2,

$$(z - \tilde{q}^w(x, hD_x))^{-1} = \mathbb{O}(1)\exp\left(\mathbb{O}(1)h^{-1}\log\frac{1}{h}\right) : \operatorname{Im}\tau_N \to \operatorname{Im}\tau_N, \tag{4-14}$$

also holds when the space $\operatorname{Im}\tau_N \subset H_{\Phi_1}(\mathbb{C}^n) \cap H_\Phi(\mathbb{C}^n)$ is equipped with the H_{Φ_1}-norm, at the expense of an $\mathbb{O}(1)$ loss in the exponent. The same conclusion holds for the bound (4-9).

Let $\Omega \Subset \mathbb{C}$ and assume that $z \in \Omega \subset\subset \mathbb{C}$ is such that $\operatorname{dist}\big(z, \operatorname{Spec}(\tilde{q}^w(x, hD_x))\big) \ge h^L/C$ for some $L \ge 1$ and $C > 0$ fixed. Then, according to Proposition 3.2, there exists $N_0 \in \mathbb{N}$, $N_0 \sim h^{-1}$, such that if $u \in H_{\Phi_1}(\mathbb{C}^n)$ is such that $(1 + |x|^2)u \in L^2_{\Phi_1}(\mathbb{C}^n)$, then, using that $[\tilde{q}^w(x, hD_x), \tau_{N_0}] = 0$, we get, for all $h > 0$ small enough,

$$\begin{aligned}\|(1-\tau_{N_0})u\|_{H_{\Phi_1}} &\le \mathbb{O}(1)\|(\tilde{q}^w(x, hD_x) - z)(1-\tau_{N_0})u\|_{H_{\Phi_1}}\\ &\le \mathbb{O}(1)\exp(\mathbb{O}(1)h^{-1})\|(\tilde{q}^w(x, hD_x) - z)u\|_{H_{\Phi_1}}.\end{aligned} \tag{4-15}$$

Here we also used Proposition 3.3. On the other hand, the bound (4-14) and Proposition 3.3 show that

$$\begin{aligned}\|\tau_{N_0}u\|_{H_{\Phi_1}} &\le \mathbb{O}(1)\exp\left(\mathbb{O}(1)h^{-1}\log\frac{1}{h}\right)\|\tau_{N_0}(\tilde{q}^w(x, hD_x) - z)u\|_{H_{\Phi_1}}\\ &\le \mathbb{O}(1)\exp\left(\mathbb{O}(1)h^{-1}\log\frac{1}{h}\right)\|(\tilde{q}^w(x, hD_x) - z)u\|_{H_{\Phi_1}}.\end{aligned} \tag{4-16}$$

Combining (4-15) and (4-16), we obtain the bound (1-7). The estimate (1-6) follows in a similar way, and hence, the proof of Theorem 1.1 is complete. □

References

[Boulton 2002] L. S. Boulton, "Non-self-adjoint harmonic oscillator, compact semigroups and pseudospectra", *J. Operator Theory* **47**:2 (2002), 413–429. MR 2003d:34185 Zbl 1034.34099

[Caliceti et al. 2007] E. Caliceti, S. Graffi, and J. Sjöstrand, "$\mathcal{PT}$ symmetric non-self-adjoint operators, diagonalizable and non-diagonalizable, with a real discrete spectrum", *J. Phys. A* **40**:33 (2007), 10155–10170. MR 2008k:81100 Zbl 1124.81018

[Davies 2000] E. B. Davies, "Wild spectral behaviour of anharmonic oscillators", *Bull. London Math. Soc.* **32**:4 (2000), 432–438. MR 2001e:47060 Zbl 1043.47502

[Davies and Kuijlaars 2004] E. B. Davies and A. B. J. Kuijlaars, "Spectral asymptotics of the non-self-adjoint harmonic oscillator", *J. London Math. Soc.* (2) **70**:2 (2004), 420–426. MR 2005e:34266 Zbl 1073.34093

[Dencker et al. 2004] N. Dencker, J. Sjöstrand, and M. Zworski, "Pseudospectra of semiclassical (pseudo-) differential operators", *Comm. Pure Appl. Math.* **57**:3 (2004), 384–415. MR 2004k:35432 Zbl 1054.35035

[Gérard and Sjöstrand 1987] C. Gérard and J. Sjöstrand, "Semiclassical resonances generated by a closed trajectory of hyperbolic type", *Comm. Math. Phys.* **108**:3 (1987), 391–421. MR 88k:58151 Zbl 0637.35027

[Helffer and Sjöstrand 1984] B. Helffer and J. Sjöstrand, "Multiple wells in the semiclassical limit, I", *Comm. Partial Differential Equations* **9**:4 (1984), 337–408. MR 86c:35113 Zbl 0546.35053

[Hitrik 2004] M. Hitrik, "Boundary spectral behavior for semiclassical operators in dimension one", *Int. Math. Res. Not.* **2004**:64 (2004), 3417–3438. MR 2005m:35206 Zbl 1077.35105

[Hörmander 1985] L. Hörmander, *The analysis of linear partial differential operators, III: Pseudodifferential operators*, Grundlehren Math. Wiss. **274**, Springer, Berlin, 1985. MR 87d:35002a Zbl 0601.35001

[Markus 1988] A. S. Markus, *Introduction to the spectral theory of polynomial operator pencils*, Translations of Mathematical Monographs **71**, American Mathematical Society, Providence, RI, 1988. MR 89h:47023 Zbl 0678.47005

[Boutet de Monvel 1974] L. Boutet de Monvel, "Hypoelliptic operators with double characteristics and related pseudo-differential operators", *Comm. Pure Appl. Math.* **27** (1974), 585–639. MR 51 #6498 Zbl 0294.35020

[Pravda-Starov 2007] K. Pravda-Starov, "Sur le pseudo-spectre de certaines classes d'opérateurs pseudo-différentiels non auto-adjoints", pp. Exposé XV in *Séminaire Équations aux Dérivées Partielles 2006-2007*, École Polytechnique, Palaiseau, 2007. MR 2009g:35355 Zbl 1201.35188

[Pravda-Starov 2008] K. Pravda-Starov, "On the pseudospectrum of elliptic quadratic differential operators", *Duke Math. J.* **145**:2 (2008), 249–279. MR 2010f:35458 Zbl 1157.35129

[Sjöstrand 1974] J. Sjöstrand, "Parametrices for pseudodifferential operators with multiple characteristics", *Ark. Mat.* **12** (1974), 85–130. MR 50 #5236 Zbl 0317.35076

[Sjöstrand 1982] J. Sjöstrand, "Singularités analytiques microlocales", pp. 1–166 in *Astérisque*, Astérisque **95**, Soc. Math. France, Paris, 1982. MR 84m:58151 Zbl 0524.35007

[Sjöstrand 1987] J. Sjöstrand, "Semiclassical resonances generated by nondegenerate critical points", pp. 402–429 in *Pseudodifferential operators* (Oberwolfach, 1986), edited by H. O. Cordes et al., Lecture Notes in Mathematics **1256**, Springer, Berlin, 1987. MR 88m:35003

[Sjöstrand 1990] J. Sjöstrand, "Geometric bounds on the density of resonances for semiclassical problems", *Duke Math. J.* **60**:1 (1990), 1–57. MR 91e:35166 Zbl 0702.35188

[Sjöstrand 1996] J. Sjöstrand, "Function spaces associated to global I-Lagrangian manifolds", pp. 369–423 in *Structure of solutions of differential equations* (Katata,1995), edited by M. Morimoto and T. Kawai, World Scientific Publishing, River Edge, NJ, 1996. MR 98i:58230 Zbl 0889.46027

[Sjöstrand 2010] J. Sjöstrand, "Resolvent estimates for non-selfadjoint operators via semigroups", pp. 359–384 in *Around the research of Vladimir Maz'ya*, vol. 3, edited by A. Laptev, Int. Math. Ser. (N. Y.) **13**, Springer, New York, 2010. MR 2011g:47115 Zbl 1198.47068

[Trefethen and Embree 2005] L. N. Trefethen and M. Embree, *Spectra and pseudospectra: The behavior of nonnormal matrices and operators*, Princeton University Press, 2005. MR 2006d:15001 Zbl 1085.15009

[Viola 2012] J. Viola, "Resolvent estimates for non-selfadjoint operators with double characteristics", *J. Lond. Math. Soc.* (2) **85**:1 (2012), 41–78. MR 2012k:47094 Zbl 1238.47033

Received 23 Sep 2011. Revised 29 Nov 2011. Accepted 16 Jan 2012.

MICHAEL HITRIK: hitrik@math.ucla.edu
Department of Mathematics, UCLA, Los Angeles 90095-1555, United States

JOHANNES SJÖSTRAND: johannes.sjostrand@u-bourgogne.fr
IMB, Université de Bourgogne, 9, Av. A. Savary, BP 47870, 21078 Dijon, France

and

UMR 5584 CNRS

JOE VIOLA: jviola@maths.lth.se
Mathematical Sciences, Lund University, Box 118, SE-22100 Lund, Sweden

ANALYSIS AND PDE
Vol. 6, No. 1, 2013
dx.doi.org/10.2140/apde.2013.6.197

BILINEAR HILBERT TRANSFORMS ALONG CURVES I: THE MONOMIAL CASE

XIAOCHUN LI

We establish an $L^2 \times L^2$ to L^1 estimate for the bilinear Hilbert transform along a curve defined by a monomial. Our proof is closely related to multilinear oscillatory integrals.

1. Introduction

Let $d \geq 2$ be a positive integer. We consider the bilinear Hilbert transform along a curve $\Gamma(t) = (t, t^d)$, defined by

$$H_\Gamma(f, g)(x) = \text{p.v.} \int_{\mathbb{R}} f(x-t)g(x-t^d)\frac{dt}{t}, \tag{1-1}$$

where f, g are Schwartz functions on $\mathbb{R}$.

The main theorem we prove in this paper is:

Theorem 1.1. *The bilinear Hilbert transform along the curve $\Gamma(t) = (t, t^d)$ can be extended to a bounded operator from $L^2 \times L^2$ to L^1.*

Remark 1.2. It can be shown, with a little modification of our method, that the bilinear Hilbert transforms along polynomial curves $(t, P(t))$ are bounded from $L^p \times L^q$ to L^r whenever $(1/p, 1/q, 1/r)$ is in the closed convex hull of $\left(\frac{1}{2}, \frac{1}{2}, 1\right)$, $\left(\frac{1}{2}, 0, \frac{1}{2}\right)$ and $\left(0, \frac{1}{2}, \frac{1}{2}\right)$. The condition $d \in \mathbb{N}$ is not necessary. Indeed, d can be any positive real number that is not equal to 1.

This problem is motivated by the Hilbert transform along a curve $\Gamma = (t, \gamma(t))$, defined by

$$H_\Gamma(f)(x_1, x_2) = \text{p.v.} \int_{\mathbb{R}} f\big(x_1 - t, x_2 - \gamma(t)\big)\frac{dt}{t},$$

and the bilinear Hilbert transform, defined by

$$H(f, g)(x) = \text{p.v.} \int_{\mathbb{R}} f(x-t)g(x+t)\frac{dt}{t}.$$

Among various curves, one simple model case is the parabola (t, t^2) in the two-dimensional plane. This work was initiated by Fabes and Rivière [1966] in order to study the regularity of parabolic differential equations. In the last thirty years, considerable work on this type of problem has been done. A nice survey on this type of operators can be found in [Stein and Wainger 1978]. For curves on homogeneous nilpotent

This research was partially supported by the NSF.
MSC2010: primary 42B20, 42B25; secondary 46B70, 47B38.
Keywords: bilinear Hilbert transform along curves.

Lie groups, the L^p estimates were established in [Christ 1985a]. The work for the Hilbert transform along more general curves with certain geometric conditions, such as the "flat" case, can be found in papers by Christ [1985b], Duoandikoetxea and J. L. Rubio de Francia [1986], and Nagel, Vance, Wainger and Weinberg [Nagel et al. 1983]. The general results were established recently in [Christ et al. 1999] for the singular Radon transforms and their maximal analogues over smooth submanifolds of $\mathbb{R}^n$ with some curvature conditions.

In recent years there has been a very active trend of harmonic analysis using time-frequency analysis to deal with multilinear operators. A breakthrough on the bilinear Hilbert transform was made by Lacey and Thiele [1997; 1999]. Following their work, the field of multilinear operators has been actively developed, to the point that some of the most interesting open questions have a strong connection to analysis on nilpotent groups. For instance, the trilinear Hilbert transform

$$\text{p.v.} \int f_1(x+t) f_2(x+2t) f_3(x+3t) \frac{dt}{t}$$

has a hidden quadratic modulation symmetry which must be accounted for in any proposed method of analysis. This nonabelian character is explicit in the work of B. Host and B. Kra [2005], who characterize the characteristic factor of the corresponding ergodic averages

$$N^{-1} \sum_{n=1}^{N} f_1(T^n) f_2(T^{2n}) f_3(T^{3n}) \longrightarrow \prod_{j=1}^{3} \mathbb{E}(f_j \mid \mathcal{N}).$$

Here, $(X, \mathcal{A}, \mu, T)$ is a measure-preserving system, and $\mathcal{N} \subset \mathcal{A}$ is the sigma-field which describes the characteristic factor, related to certain 2-step nilpotent groups. The limit above is in the sense of L^2-norm convergence, and holds for all bounded f_1, f_2, f_3.

The ergodic analogue of the bilinear Hilbert transform along a parabola is the nonconventional bilinear average

$$N^{-1} \sum_{n=1}^{N} f_1(T^n) f_2(T^{n^2}) \longrightarrow \prod_{j=1}^{2} \mathbb{E}(f_j \mid \mathcal{K}_{\text{profinite}}),$$

where $\mathcal{K}_{\text{profinite}} \subset \mathcal{A}$ is the profinite factor, a sub-σ-field of the maximal abelian factor of $(X, \mathcal{A}, \mu, T)$. The proof of the characteristic factor result above, due to Furstenberg [1990], utilizes the characteristic factor for the three-term result. We are indebted to M. Lacey for bringing Furstenberg's theorems to our attention. However, a notable fact is that our proof for the bilinear Hilbert transform along a monomial curve does not have to go through the trilinear Hilbert transform. The proof provided in this article relies heavily on the concept of a "quadratic uniformity", inspired by [Gowers 1998].

Another prominent theme is the relation of the bilinear Hilbert transforms along curves and the multilinear oscillatory integrals. The bilinear Hilbert transforms along curves are closely associated to the multilinear oscillatory integrals of the type

$$\Lambda_\lambda(f_1, f_2, f_3) = \int_{\boldsymbol{B}} f_1(\boldsymbol{x} \cdot \boldsymbol{v}_1) f_2(\boldsymbol{x} \cdot \boldsymbol{v}_2) f_3(\boldsymbol{x} \cdot \boldsymbol{v}_3) e^{i\lambda\varphi(\boldsymbol{x})}\, d\boldsymbol{x}, \tag{1-2}$$

where $\boldsymbol{B}$ is the unit ball in $\mathbb{R}^3$, $\boldsymbol{v}_1$, $\boldsymbol{v}_2$, $\boldsymbol{v}_3$ are vectors in $\mathbb{R}^3$, and the phase function φ satisfies a nondegenerate condition

$$\left|\prod_{j=1}^{3}(\nabla\cdot\boldsymbol{v}_j^{\perp})\varphi(\boldsymbol{x})\right|\geq 1. \tag{1-3}$$

Here $\boldsymbol{v}_j^{\perp}$ is a unit vector orthogonal to $\boldsymbol{v}_j$, for each j. For a polynomial phase φ with the nondegenerate condition (1-3), it was proved in [Christ et al. 2005] that

$$\left|\Lambda_\lambda(f_1, f_2, f_3)\right| \leq C(1+|\lambda|)^{-\varepsilon}\prod_{j=1}^{3}\|f_j\|_\infty \tag{1-4}$$

holds for some positive number ε. For the particular vectors $\boldsymbol{v}$ and the nondegenerate phase φ encountered in our problem, an estimate similar to (1-4) still holds. However, one of the main difficulties arises from the falsity of L^2 decay estimates for the trilinear form Λ_λ. It is to overcome this difficulty that we introduce the quadratic uniformity, which plays the role of a bridge connecting two spaces L^2 and L^∞.

The method used in this paper essentially works for those curves on nilpotent groups. It is possible to extend Theorem 1.1 to the general setting of nilpotent Lie groups. But we will not pursue this in this article. There are some related questions one can pose. Besides the generalization to the more general curves, it is natural to ask the corresponding problems in higher-dimensional cases and/or in multilinear cases. For instance, in the trilinear case, one can consider

$$T(f_1, f_2, f_3)(x) = \text{p.v.}\int f_1(x+t)f_2(x+p_1(t))f_3(x+p_2(t))\frac{dt}{t}. \tag{1-5}$$

Here p_1, p_2 are polynomials of t. The investigation of such problems will be discussed in subsequent papers.

2. A decomposition

Let ρ_1 be a standard bump function supported on $\left[\frac{1}{2}, 2\right]$, and let

$$\rho(t) = \rho_1(t)\mathbf{1}_{\{t>0\}} - \rho_1(-t)\mathbf{1}_{\{t<0\}}.$$

It is clear that ρ is an odd function. To obtain the L^r estimates for H_Γ, it is sufficient to get L^r estimates for T_Γ defined by $T_\Gamma = \sum_{j\in\mathbb{Z}} T_{\Gamma,j}$, where $T_{\Gamma,j}$ is

$$T_{\Gamma,j}(f,g)(x) = \int f(x-t)g(x-t^d)2^j\rho(2^jt)\,dt. \tag{2-1}$$

Let L be a large positive number (larger than 2^{100}). By Lemma 9.1, we have that if $|j|\leq L$,

$$\left\|T_{\Gamma,j}(f,g)\right\|_r \leq C_L\|f\|_p\|g\|_q$$

for all $p, q > 1$ and $1/p+1/q = 1/r$, where the operator norm C_L depends on the upper bound L. Hence in the following we only need to consider the case when $|j| > L$. In fact we prove the following theorem.

Theorem 2.1. *Let $T_{\Gamma,j}$ be defined as in* (2-1). *Then the bilinear operator $T_L = \sum_{j\in\mathbb{Z}:|j|>L} T_{\Gamma,j}$ is bounded from $L^2 \times L^2$ to L^1.*

Clearly Theorem 1.1 follows from Theorem 2.1 and Lemma 9.1. The rest of the article is devoted to a proof of Theorem 2.1.

We begin the proof of Theorem 2.1 by constructing an appropriate decomposition of the operator $T_{\Gamma,j}$. This is done by an analysis of the bilinear symbol associated with the operator.

Expressing $T_{\Gamma,j}$ in dual frequency variables, we have

$$T_{\Gamma,j}(f,g)(x) = \iint \widehat{f}(\xi)\widehat{g}(\eta)e^{2\pi i(\xi+\eta)x}\mathfrak{m}_j(\xi,\eta)\,d\xi\,d\eta,$$

where the symbol $\mathfrak{m}_j$ is defined by

$$\mathfrak{m}_j(\xi,\eta) = \int \rho(t)\exp\bigl(-2\pi i(2^{-j}\xi t + 2^{-dj}\eta t^d)\bigr)\,dt. \tag{2-2}$$

First we introduce a resolution of the identity. Let Θ be a Schwarz function supported on $(-1,1)$ such that $\Theta(\xi)=1$ if $|\xi|\le\frac{1}{2}$. Set Φ to be a Schwartz function satisfying

$$\widehat{\Phi}(\xi) = \Theta\left(\frac{\xi}{2}\right) - \Theta(\xi)\,.$$

Then Φ is a Schwartz function such that $\widehat{\Phi}$ is supported on $\bigl\{\xi : \frac{1}{2} < |\xi| < 2\bigr\}$ and

$$\sum_{m\in\mathbb{Z}}\widehat{\Phi}\left(\frac{\xi}{2^m}\right) = 1 \quad \text{for all} \quad \xi\in\mathbb{R}\backslash\{0\}, \tag{2-3}$$

and for any $m_0\in\mathbb{Z}$,

$$\widehat{\Phi}_{m_0}(\xi) := \sum_{m=-\infty}^{m_0}\widehat{\Phi}\left(\frac{\xi}{2^m}\right) = \Theta\left(\frac{\xi}{2^{m_0+1}}\right), \tag{2-4}$$

which is a bump function supported on $(-2^{m_0+1}, 2^{m_0+1})$.

From (2-3), we can decompose $T_{\Gamma,j}$ into two parts: $T_{\Gamma,j,1}$ and $T_{\Gamma,j,2}$, where $T_{\Gamma,j,1}$ is given by

$$\sum_{m\in\mathbb{Z}}\sum_{\substack{m'\in\mathbb{Z}:\\|m'-m|>10^d}}\iint \widehat{f}(\xi)\widehat{g}(\eta)e^{2\pi i(\xi+\eta)x}\widehat{\Phi}\left(\frac{2^{-j}\xi}{2^m}\right)\widehat{\Phi}\left(\frac{2^{-dj}\eta}{2^{m'}}\right)\mathfrak{m}_j(\xi,\eta)\,d\xi\,d\eta, \tag{2-5}$$

and $T_{\Gamma,j,2}$ is defined by

$$\sum_{m\in\mathbb{Z}}\sum_{\substack{m'\in\mathbb{Z}:\\|m'-m|\le 10^d}}\iint \widehat{f}(\xi)\widehat{g}(\eta)e^{2\pi i(\xi+\eta)x}\widehat{\Phi}\left(\frac{2^{-j}\xi}{2^m}\right)\widehat{\Phi}\left(\frac{2^{-dj}\eta}{2^{m'}}\right)\mathfrak{m}_j(\xi,\eta)\,d\xi\,d\eta. \tag{2-6}$$

Define $\mathfrak{m}_d$ by

$$\mathfrak{m}_d(\xi,\eta) = \int \rho(t)\exp\bigl(-2\pi i(\xi t + \eta t^d)\bigr)\,dt. \tag{2-7}$$

Clearly $\mathfrak{m}_j(\xi,\eta) = \mathfrak{m}_d(2^{-j}\xi, 2^{-dj}\eta)$. In $T_{\Gamma,j,1}$, the phase function $\phi_{\xi,\eta}(t) = \xi t + \eta t^d$ does not have any critical point in a neighborhood of the support of ρ, and therefore a very rapid decay can be obtained

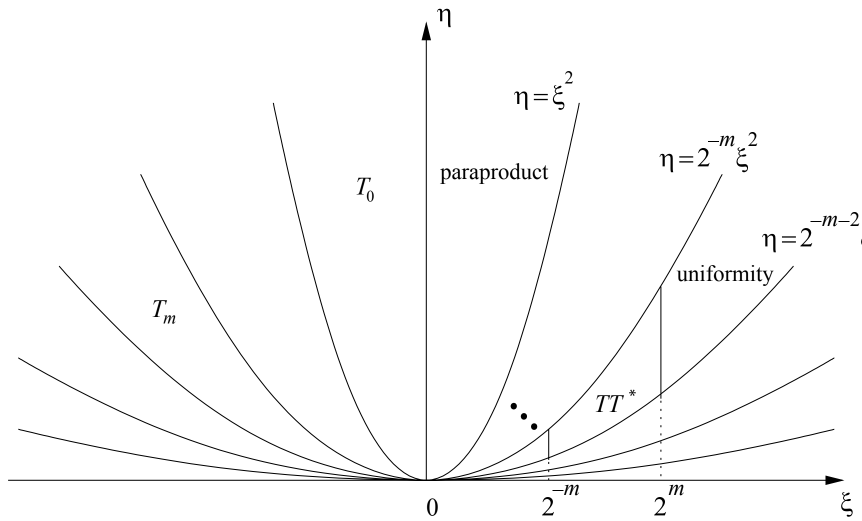

Figure 1. Decomposition of the (ξ, η)-plane for $\sum_m T_m$ when $d = 2$.

by integration by parts so that we can show that $\sum_j T_{\Gamma,j,1}$ is essentially a finite sum of paraproducts (see Section 3). A critical point of the phase function may occur in $T_{\Gamma,j,2}$, and therefore the method of stationary phase must be brought to bear in this case, exploiting in particular the oscillatory term. This case requires the most extensive analysis. Heuristically, the decomposition is made according to the curvature of the curve (t, t^d). For example, for the parabola case, the frequency space is broken into parabolic regions $\{(\xi, \eta) : \eta \sim 2^{-m}\xi^2\}$, as shown in the figure. Naturally, the $2^{-\varepsilon m}$ decay estimate is expected in order to sum up all parabolic regions.

Notice that there are only finitely many m' if m is fixed in (2-6). Without loss of generality, we can assume $m' = m$. Then in order to get the L^r estimates for $\sum_j T_{\Gamma,j,2}$, it suffices to prove the L^r boundedness of $\sum_m T_m$, where the T_m are defined by

$$T_m(f, g)(x) = \sum_{|j|>L} \iint \widehat{f}(\xi)\widehat{g}(\eta)e^{2\pi i(\xi+\eta)x}\,\widehat{\Phi}\left(\frac{2^{-j}\xi}{2^m}\right)\widehat{\Phi}\left(\frac{2^{-dj}\eta}{2^m}\right)\mathfrak{m}_j(\xi, \eta)\, d\xi\, d\eta. \tag{2-8}$$

It can be proved that $T_0 = \sum_{m\le 0} T_m$ is equal to $\sum_{m\le 0} O(2^{m/2})\Pi_m$, where Π_m is a paraproduct studied in Theorem 3.1. This can be done by Fourier series and the cancellation condition of ρ, and thus T_0 is essentially a paraproduct. We omit the details, since they are exactly the same as those in Section 3 for the case $\sum_j T_{\Gamma,j,1}$. Therefore, the most difficult term is $\sum_{m\ge 1} T_m$. For this term, we have the following theorem.

Theorem 2.2. *Let T_m be a bilinear operator defined as in* (2-8). *Then there exists a constant C such that*

$$\Big\| \sum_{m\ge 1} T_m(f, g) \Big\|_1 \le C \|f\|_2\, \|g\|_2 \tag{2-9}$$

holds for all $f, g \in L^2$.

A delicate analysis is required for proving this theorem. We prove it on page 204. Theorem 2.1 follows from Theorem 2.2 and the boundedness of $\sum_j T_{\Gamma,j,1}$. The rest of the article is organized as follows. In Section 3, the L^r-boundedness will be established for $\sum_j T_{\Gamma,j,1}$. Some crucial bilinear restriction estimates will appear in Section 4 and as a consequence Theorem 2.2 follows. Sections 5–8 are devoted to a proof of the bilinear restriction estimates.

3. Paraproducts and uniform estimates

In this section we prove that $\sum_j T_{\Gamma,j,1}$ is essentially a finite sum of certain paraproducts bounded from $L^p \times L^q$ to L^r.

First let us introduce the paraproduct encountered in our problem. Let $j \in \mathbb{Z}$, L_1, L_2 be positive integers and M_1, M_2 be integers. Then

$$\omega_{1,j} = \left[\frac{2^{L_1 j+M_1}}{2}, 2 \cdot 2^{L_1 j+M_1}\right]$$

and

$$\omega_{2,j} = [-2^{L_2 j+M_2}, 2^{L_2 j+M_2}].$$

Let Φ_1 be a Schwartz function whose Fourier transform is a standard bump function supported on a small neighborhood of $\left[\frac{1}{2}, 2\right]$ or $\left[-2, -\frac{1}{2}\right]$, and Φ_2 be a Schwartz function whose Fourier transform is a standard bump function supported on $[-1, 1]$ and $\widehat{\Phi}_2(0) = 1$. For $l \in \{1, 2\}$ and $n_1, n_2 \in \mathbb{Z}$, define Φ_{l,j,n_l} by

$$\widehat{\Phi}_{l,j,n_l}(\xi) = \left(e^{2\pi i n_l(\cdot)}\widehat{\Phi}_l(\cdot)\right)\left(\frac{\xi}{2^{L_l j+M_l}}\right).$$

It is clear that $\widehat{\Phi}_{l,j,n_l}$ is supported on $\omega_{l,j}$. For locally integrable functions f_l, we define $f_{l,j}$ by

$$f_{l,j,n_l}(x) = f_l * \Phi_{l,j,n_l}(x).$$

We now define a paraproduct to be

$$\Pi_{L_1,L_2,M_1,M_2,n_1,n_2}(f_1, f_2)(x) = \sum_{j\in\mathbb{Z}} \prod_{l=1}^{2} f_{l,j,n_l}(x). \tag{3-1}$$

For this paraproduct, we have the following uniform estimates.

Theorem 3.1. *For any $p_1 > 1$, $p_2 > 1$ with $1/p_1 + 1/p_2 = 1/r$, there exists a constant C independent of M_1, M_2, n_1, n_2 such that*

$$\left\|\Pi_{L_1,L_2,M_1,M_2,n_1,n_2}(f_1, f_2)\right\|_r \le C\big(1+|n_1|\big)^{10}\big(1+|n_2|\big)^{10}\|f_1\|_{p_1}\,\|f_2\|_{p_2}, \tag{3-2}$$

for all $f_1 \in L^{p_1}$ and $f_2 \in L^{p_2}$.

The case $r > 1$ can be handled by a telescoping argument. The case $r < 1$ is more complicated and requires a time-frequency analysis. A proof of Theorem 3.1 can be found in [Li 2008]. The constant C in Theorem 3.1 may depend on L_1, L_2. It is easy to see that C is $O(\max\{2^{L_1}, 2^{L_2}\})$. It is possible to get a much better upper bound, such as $O\big(\log(1+\max\{L_2/L_1, L_1/L_2\})\big)$, by tracking the constants carefully in the proof in [Li 2008]. But we do not need the sharp constant in this article. The independence on M_1, M_2 is the most important issue here.

We now return to $\sum_j T_{\Gamma,j,1}$. This sum can be written as $T_{L,1}+T_{L,2}$, where $T_{L,1}$ is a bilinear operator defined by

$$\sum_{|j|>L}\sum_{m\in\mathbb{Z}}\sum_{\substack{m'\in\mathbb{Z}\\ m'<m-10^d}}\iint \widehat{f}(\xi)\widehat{g}(\eta)e^{2\pi i(\xi+\eta)x}\widehat{\Phi}\Big(\frac{2^{-j}\xi}{2^m}\Big)\widehat{\Phi}\Big(\frac{2^{-dj}\eta}{2^{m'}}\Big)\mathfrak{m}_j(\xi,\eta)\,d\xi\,d\eta,$$

and $T_{L,2}$ is a bilinear operator given by

$$\sum_{|j|>L}\sum_{m'\in\mathbb{Z}}\sum_{\substack{m\in\mathbb{Z}\\ m<m'-10^d}}\iint \widehat{f}(\xi)\widehat{g}(\eta)e^{2\pi i(\xi+\eta)x}\widehat{\Phi}\Big(\frac{2^{-j}\xi}{2^m}\Big)\widehat{\Phi}\Big(\frac{2^{-dj}\eta}{2^{m'}}\Big)\mathfrak{m}_j(\xi,\eta)\,d\xi\,d\eta.$$

It is standard to verify that $T_{L,1}$ and $T_{L,2}$ are paraproducts as defined in (3-1). Hence the $L^p\times L^q\to L^r$ estimates of these paraproducts follow from Theorem 3.1, for all $p,q>1$ and $1/p+1/q=1/r$.

4. Bilinear Fourier restriction estimates

Let $d\geq 2$, $m\geq 0$, $j\in\mathbb{Z}$. We define a bilinear Fourier restriction operator of f, g by

$$\mathcal{B}_{j,m}(f,g)(x)=2^{-(d-1)j/2}\int_{\mathbb{R}}R_\Phi f\big(2^{-(d-1)j}x-2^m t\big)R_\Phi g(x-2^m t^d)\rho(t)\,dt \quad \text{if } j\geq 0 \tag{4-1}$$

and

$$\mathcal{B}_{j,m}(f,g)(x)=2^{(d-1)j/2}\int_{\mathbb{R}}R_\Phi f(x-2^m t)R_\Phi g(2^{(d-1)j}x-2^m t^d)\rho(t)\,dt \quad \text{if } j<0, \tag{4-2}$$

where $R_\Phi f$ and $R_\Phi g$ are the Fourier (smooth) restrictions of f, g on the support of $\widehat{\Phi}$ respectively. More precisely, $R_\Phi f$, $R_\Phi g$ are given by

$$\widehat{R_\Phi f}(\xi)=\widehat{f}(\xi)\widehat{\Phi}(\xi), \tag{4-3}$$

$$\widehat{R_\Phi g}(\xi)=\widehat{g}(\xi)\widehat{\Phi}(\xi). \tag{4-4}$$

By inserting absolute values throughout and applying the Cauchy–Schwarz inequality, the boundedness of $\mathcal{B}_{j,m}$ from $L^2\times L^2$ to L^1 follows immediately. Moreover, since the Fourier transforms of f, g are restricted on the support of $\widehat{\Phi}$, we actually can improve the estimate. Let us state the improved estimates in the following theorems, which are of independent interest.

Theorem 4.1. *Let $d\geq 2$ and $\mathcal{B}_{j,m}$ be defined as in* (4-1) *and* (4-2). *If $L\leq |j|\leq m/(d-1)$, then there exists a constant C independent of j, m such that*

$$\big\|\mathcal{B}_{j,m}(f,g)\big\|_1\leq C\,2^{\frac{(d-1)|j|-m}{8}}\|f\|_2\,\|g\|_2 \quad \textit{for all } f,g\in L^2. \tag{4-5}$$

Theorem 4.2. *Let $d\geq 2$ and $\mathcal{B}_{j,m}$ be defined as in* (4-1) *and* (4-2). *If $|j|\geq m/(d-1)$, then there exist a positive number ε_0 and a constant C independent of j, m such that*

$$\big\|\mathcal{B}_{j,m}(f,g)\big\|_1\leq C\max\big\{2^{\frac{m-(d-1)|j|}{3}},2^{-\varepsilon_0 m}\big\}\|f\|_2\,\|g\|_2 \quad \textit{for all } f,g\in L^2. \tag{4-6}$$

The positive number ε_0 in Theorem 4.2 can be chosen to be $1/(8d)$. Theorem 4.1 can be proved by a TT^* method. However, the TT^* method fails when $|j| > m/(d-1)$. To obtain Theorem 4.2, we will employ a method related to the uniformity of functions.

Now we can see that Theorem 2.2 is a consequence of Theorems 4.1 and 4.2.

Proof of Theorem 2.2. Define a bilinear operator $T_{j,m}$ to be

$$T_{j,m}(f,g)(x) = \iint \widehat{f}(\xi)\widehat{g}(\eta)e^{2\pi i(\xi+\eta)x}\widehat{\Phi}\Big(\frac{2^{-j}\xi}{2^m}\Big)\widehat{\Phi}\Big(\frac{2^{-dj}\eta}{2^m}\Big)\mathfrak{m}_j(\xi,\eta)\,d\xi\,d\eta. \tag{4-7}$$

Let $\gamma_{j,m}$ be defined by

$$\gamma_{j,m} = \begin{cases} 2^{\frac{(d-1)|j|-m}{8}} & \text{if } |j| \le \dfrac{m}{d-1}, \\ \max\big\{2^{\frac{m-(d-1)|j|}{3}}, 2^{-\varepsilon_0 m}\big\} & \text{if } |j| \ge \dfrac{m}{d-1}. \end{cases} \tag{4-8}$$

A rescaling argument and Theorems 4.1 and 4.2 yield

$$\big\|T_{j,m}(f,g)\big\|_1 \le C\gamma_{j,m}\|f\|_2\,\|g\|_2. \tag{4-9}$$

Since $\sum_m T_m = \sum_m \sum_{j:|j|\ge L} T_{j,m}$, we obtain

$$\Big\|\sum_{m\ge1} T_m(f,g)\Big\|_1 \le C\sum_{m\ge1}\sum_{j:|j|\ge L}\gamma_{j,m}\|f_{j,m}\|_2\,\|g_{j,m}\|_2, \tag{4-10}$$

where

$$\widehat{f_{j,m}}(\xi) = \widehat{f}(\xi)\widehat{\Phi}\Big(\frac{\xi}{2^{j+m}}\Big),$$

$$\widehat{g_{j,m}}(\eta) = \widehat{g}(\eta)\widehat{\Phi}\Big(\frac{\eta}{2^{dj+m}}\Big).$$

Clearly the right-hand side of (4-10) is bounded by $C\|f\|_2\|g\|_2$. Therefore, we finish the proof of Theorem 2.2.

Since t is localized, it is sufficient to consider $\tilde{\mathcal{B}}_{j,m,n}$ given by

$$\tilde{\mathcal{B}}_{j,m} = \mathcal{B}_{j,m}\mathbf{1}_I^*. \tag{4-11}$$

Here I is an interval whose size is $2^{(d-1)|j|+m}$ and $\mathbf{1}_I^* = \mathbf{1}_I * \phi_k$, where $\phi_k(x)$ equals $2^{-k}\phi(2^{-k}x)$ for a given nonnegative Schwartz function ϕ whose Fourier transform is a standard bump function on $\big[-\frac12, \frac12\big]$. In what follows, we still use $\mathcal{B}_{j,m}$ to denote the localized operator $\tilde{\mathcal{B}}_{j,m}$.

Trilinear forms. Let f_1, f_2, f_3 be measurable functions supported on $\frac{1}{16} \le |\xi| \le \frac{39}{16}$. Define a trilinear form $\Lambda_{j,m}(f_1, f_2, f_3)$ by

$$\Lambda_{j,m}(f_1,f_2,f_3) := \big\langle \mathcal{B}_{j,m}(\check{f}_1, \check{f}_2), \check{f}_3\big\rangle. \tag{4-12}$$

Theorems 4.1 and 4.2 can be reduced to the following theorems respectively.

Theorem 4.3. *Let $d \geq 2$ and $\Lambda_{j,m}(f_1, f_2, f_3)$ be defined as in* (4-12). *If $|j| \leq m/(d-1)$, then there exists a constant C independent of j, m such that*

$$\left|\Lambda_{j,m}(f_1, f_2, f_3)\right| \leq C 2^{\frac{-(d-1)|j|-m}{2}} 2^{-\frac{m-(d-1)|j|}{6}} \|f_1\|_2 \|f_2\|_2 \|f_3\|_2 \tag{4-13}$$

for all $f_1, f_2, f_3 \in L^2$.

Theorem 4.4. *Let $d \geq 2$ and $\Lambda_{j,m}(f_1, f_2, f_3)$ be defined as in* (4-12). *If $|j| \geq m/(d-1)$, then there exist a positive number ε_0 and a constant C independent of j, m such that*

$$\left|\Lambda_{j,m}(f_1, f_2, f_3)\right| \leq C \max\left\{2^{\frac{-(d-1)|j|+m}{2}}, 2^{-\varepsilon_0 m}\right\} \|f_1\|_2 \|f_2\|_2 \|\widehat{f_3}\|_\infty \tag{4-14}$$

holds for all $f_1, f_2 \in L^2$ and $\widehat{f_3} \in L^\infty$ such that f_1, f_2, f_3 are supported on $\frac{1}{16} \leq |\xi| \leq \frac{39}{16}$.

A proof of Theorem 4.3 will be provided in Section 5, and a proof of Theorem 4.4 will be given in Section 7.

5. Stationary phases and trilinear oscillatory integrals

In this section we provide a proof of Theorem 4.3 by utilizing essentially a TT^* method. In this case, one cannot reduce the problem to the standard paraproduct problem because the critical points of the phase function may occur in a neighborhood of $\frac{1}{2} \leq |t| \leq 2$, say $\frac{1}{4} \leq |t| \leq \frac{5}{2}$, which provides a stationary phase for the Fourier integral $\mathfrak{m}_d$. This stationary phase gives a highly oscillatory factor in the integral. We expect a suitable decay from the highly oscillatory factor.

Let $\Lambda_{j,m}(f_1, f_2, f_3) = \langle \mathcal{B}_{j,m}(\check{f}_1, \check{f}_2), \check{f}_3\rangle$. To prove Theorem 4.3, it suffices to prove the following L^2 estimate for the trilinear form $\Lambda_{j,m}(f_1, f_2, f_3)$:

$$\left|\Lambda_{j,m}(f_1, f_2, f_3)\right| \leq C 2^{\frac{-(d-1)|j|-m}{2}} 2^{-\frac{m-(d-1)|j|}{6}} \|f_1\|_2 \|f_2\|_2 \|f_3\|_2 \tag{5-1}$$

holds for all $f_1, f_2, f_3 \in L^2$. Clearly $\Lambda_{j,m}(f_1, f_2, f_3)$ can be expressed as

$$2^{-(d-1)j/2} \iint f_1(\xi)\widehat{\Phi}(\xi) f_2(\eta)\widehat{\Phi}(\eta) f_3\big(2^{-(d-1)j}\xi + \eta\big)\mathfrak{m}_d(2^m\xi, 2^m\eta)\, d\xi\, d\eta$$

if $j > 0$, and as

$$2^{(d-1)j/2} \iint f_1(\xi)\widehat{\Phi}(\xi) f_2(\eta)\widehat{\Phi}(\eta) f_3\big(\xi + 2^{(d-1)j}\eta\big)\mathfrak{m}_d(2^m\xi, 2^m\eta)\, d\xi\, d\eta$$

if $j \leq 0$,

Whenever $\xi, \eta \in \operatorname{supp} \widehat{\Phi}$, the second-order derivative of the phase function $\phi_{m,\xi,\eta}(t) = 2^m(\xi t + \eta t^d)$ is comparable to 2^m. We only need to focus on the worst situation, when there is a critical point of the phase function in a small neighborhood of $\operatorname{supp} \rho$. Thus the method of stationary phase yields

$$\mathfrak{m}_d(2^m\xi, 2^m\eta) \sim 2^{-m/2} \exp\big(i c_d 2^m \xi^{d/(d-1)} \eta^{-1/(d-1)}\big), \tag{5-2}$$

where c_d is a constant depending only on d (see [Sogge 1993; Stein 1993]). Henceforth we reduce Theorem 4.3 to the following lemma.

Proposition 5.1. *Let $\Lambda^*_{j,m}$ be defined by*

$$\Lambda^*_{j,m}(f_1, f_2, f_3) = \iint f_1(\xi)\widehat{\Phi}(\xi) f_2(\eta)\widehat{\Phi}(\eta) f_3\big(2^{-(d-1)j}\xi + \eta\big)\exp\big(i c_d 2^m \xi^{\frac{d}{d-1}}\eta^{-\frac{1}{d-1}}\big)\, d\xi\, d\eta \tag{5-3}$$

if $j > 0$, and by

$$\Lambda^*_{j,m}(f_1, f_2, f_3) = \iint f_1(\xi)\widehat{\Phi}(\xi) f_2(\eta)\widehat{\Phi}(\eta) f_3\big(\xi + 2^{(d-1)j}\eta\big)\exp\big(i c_d 2^m \xi^{\frac{d}{d-1}}\eta^{-\frac{1}{d-1}}\big)\, d\xi\, d\eta. \tag{5-4}$$

if $j \le 0$. There exists a positive constant C such that

$$\big|\Lambda^*_{j,m}(f_1, f_2, f_3)\big| \le C\, 2^{-\frac{m-(d-1)|j|}{6}} \|f_1\|_2 \|f_2\|_2 \|f_3\|_2 \tag{5-5}$$

holds for all $f_1, f_2, f_3 \in L^2$.

Proof. Without loss of generality, we assume that $\widehat{\Phi}$ is supported on $\big[\frac{1}{2}, 2\big]$ or $\big[-2, -\frac{1}{2}\big]$. And we only give a proof for the case $j > 0$, since a similar argument yields the case $j \le 0$. Let $\phi_{d,m}$ be a phase function defined by

$$\phi_{d,m}(\xi, \eta) = c_d \xi^{d/(d-1)} \eta^{-1/(d-1)},$$

and let $b_1 = 1 - 2^{-(d-1)j}$ and $b_2 = 2^{-(d-1)j}$. Changing variables $\xi \mapsto \xi - \eta$ and $\eta \mapsto b_1\xi + b_2\eta$, we have that $\Lambda^*_{j,m}(f_1, f_2, f_3)$ equals

$$\iint f_1(\xi - \eta) f_2(b_1\xi + b_2\eta) f_3(\xi)\widehat{\Phi}(\xi - \eta)\widehat{\Phi}(b_1\xi + b_2\eta) e^{i\,2^m\phi_{d,m}(\xi-\eta, b_1\xi+b_2\eta)}\, d\xi\, d\eta.$$

Thus, by Cauchy–Schwarz, we dominate $|\Lambda^*_{j,m}(f_1, f_2, f_3)|$ by

$$\big\|\boldsymbol{T}_{d,j,m}(f_1, f_2)\big\|_2 \|f_3\|_2,$$

where $\boldsymbol{T}_{d,j,m}$ is defined by

$$\boldsymbol{T}_{d,j,m}(f_1, f_2)(\xi) = \int f_1(\xi - \eta) f_2(b_1\xi + b_2\eta)\widehat{\Phi}(\xi - \eta)\widehat{\Phi}(b_1\xi + b_2\eta) e^{i\,2^m\phi_{d,m}(\xi-\eta, b_1\xi+b_2\eta)}\, d\eta.$$

It is easy to see that $\big\|\boldsymbol{T}_{d,j,m}(f_1, f_2)\big\|_2^2$ equals

$$\int \bigg(\iint F(\xi, \eta_1, \eta_2) G(\xi, \eta_1, \eta_2) e^{i\,2^m(\phi_{d,m}(\xi-\eta_1, b_1\xi+b_2\eta_1) - \phi_{d,m}(\xi-\eta_2, b_1\xi+b_2\eta_2))} d\eta_1\, d\eta_2\bigg)\, d\xi,$$

where

$$F(\xi, \eta_1, \eta_2) = (f_1\widehat{\Phi})(\xi - \eta_1)\overline{(f_1\widehat{\Phi})(\xi - \eta_2)},$$
$$G(\xi, \eta_1, \eta_2) = (f_2\widehat{\Phi})(b_1\xi + b_2\eta_1)\overline{(f_2\widehat{\Phi})(b_1\xi + b_2\eta_2)}.$$

Changing variables $\eta_1 \mapsto \eta$ and $\eta_2 \mapsto \eta + \tau$, we see that $\big\|\boldsymbol{T}_{d,j,m}(f_1, f_2)\big\|_2^2$ equals

$$\int \bigg(\iint F_\tau(\xi - \eta) G_\tau(b_1\xi + b_2\eta) e^{i\,2^m(\phi_{d,m}(\xi-\eta, b_1\xi+b_2\eta) - \phi_{d,m}(\xi-\eta-\tau, b_1\xi+b_2(\eta+\tau)))}\, d\xi\, d\eta\bigg)\, d\tau,$$

where

$$F_\tau(\,\cdot\,) = (f_1\widehat{\Phi})(\,\cdot\,)\overline{(f_1\widehat{\Phi})(\,\cdot - \tau)},$$
$$G_\tau(\,\cdot\,) = (f_2\widehat{\Phi})(\,\cdot\,)\overline{(f_2\widehat{\Phi})(\,\cdot + b_2\tau)}.$$

Changing coordinates to $(u, v) = (\xi - \eta, b_1\xi + b_2\eta)$, the inner integral becomes

$$\iint F_\tau(u) G_\tau(v) \exp\bigl(i\, 2^m\, \tilde{\boldsymbol{Q}}_\tau(u, v)\bigr)\, du\, dv, \tag{5-6}$$

where $\tilde{\boldsymbol{Q}}_\tau$ is defined by

$$\tilde{\boldsymbol{Q}}_\tau(u, v) = \phi_{d,m}(u, v) - \phi_{d,m}(u - \tau, v + b_2\tau).$$

When j is large enough, the mean value theorem yields

$$\bigl|\partial_u \partial_v\, \tilde{\boldsymbol{Q}}_\tau(u, v)\bigr| \geq C\tau, \tag{5-7}$$

if $u, v, u - \tau, v + b_2\tau \in \operatorname{supp}\widehat{\Phi}$.

A well-known theorem of Hörmander on the nondegenerate phase [Hörmander 1973; Phong and Stein 1994] gives for (5-6) the estimate

$$C \min\bigl\{1, 2^{-m/2}|\tau|^{-1/2}\bigr\}\, \|F_\tau\|_2\, \|G_\tau\|_2.$$

Hence, by the Cauchy–Schwarz inequality, $\bigl\|\boldsymbol{T}_{d,j,m}(f_1, f_2)\bigr\|_2^2$ is bounded by

$$\tau_0 \|f_1\|_2^2\, \|f_2\|_2^2 + C \int_{\tau_0 < |\tau| < 10} \min\bigl\{1, 2^{-m/2}|\tau|^{-1/2}\bigr\}\, \|F_\tau\|_2\, \|G_\tau\|_2\, d\tau$$

for any $\tau_0 > 0$. By one more use of the Cauchy–Schwarz inequality, $\bigl\|\boldsymbol{T}_{d,j,m}(f_1, f_2)\bigr\|_2^2$ is dominated by $\bigl(\tau_0 + C\tau_0^{-1/2} 2^{-m/2} 2^{(d-1)j/2}\bigr) \|f_1\|_2^2\, \|f_2\|_2^2$ for any $\tau_0 > 0$. Thus we have

$$\bigl|\Lambda^*_{j,m}(f_1, f_2, f_3)\bigr| \leq C\, 2^{\frac{(d-1)j - m}{6}}\, \|f_1\|_2\, \|f_2\|_2\, \|f_3\|_2. \tag{5-8}$$

This completes the proof of Proposition 5.1. □

It is easy to see that

$$\bigl|\Lambda^*_{j,m}(f_1, f_2, f_3)\bigr| \leq C\, 2^{-\varepsilon m}\, \|f_1\|_2\, \|f_2\|_2\, \|f_3\|_2 \tag{5-9}$$

fails for all $|j| \geq m/(d-1)$. Indeed, let us only consider the case $j > m/(d-1)$. Assume that (5-9) holds for all $j > m/(d-1)$. Let $j \to \infty$; then (5-9) implies

$$\bigl|\Lambda^*_m(f_1, f_2, f_3)\bigr| \leq C\, 2^{-\varepsilon m}\, \|f_1\|_2\, \|f_2\|_2\, \|f_3\|_2, \tag{5-10}$$

where

$$\Lambda^*_m(f_1, f_2, f_3) = \iint f_1(\xi)\widehat{\Phi}(\xi) f_2(\eta)\widehat{\Phi}(\eta) f_3(\eta) \exp\bigl(i c_d 2^m \xi^{d/(d-1)} \eta^{-1/(d-1)}\bigr)\, d\xi\, d\eta.$$

Simply taking $f_2 = f_3$, we obtain

$$\sup_{\eta \sim 1} \left| \int f_1(\xi)\widehat{\Phi}(\xi) \exp\bigl(i c_d 2^m \xi^{d/(d-1)} \eta^{-1/(d-1)}\bigr)\, d\xi \right| \leq C\, 2^{-\varepsilon m}\, \|f_1\|_2. \tag{5-11}$$

This clearly cannot be true, and hence we get a contradiction. Therefore, (5-9) does not hold for $j > m/(d-1)$. Hence the TT^* method cannot work for the case $|j| > m/(d-1)$. In the following sections, we have to introduce a concept of uniformity and employ a "quadratic" Fourier analysis.

6. Uniformity

We introduce a concept related to a notion of uniformity employed by Gowers [1998]. A similar uniformity was utilized in [Christ et al. 2005]. Let $\sigma \in (0, 1]$, let $\mathfrak{Q}$ be a collection of some real-valued measurable functions, and fix a bounded interval $\boldsymbol{I}$ in $\mathbb{R}$.

Definition 6.1. A function $f \in L^2(\boldsymbol{I})$ is σ-uniform in $\mathfrak{Q}$ if

$$\left| \int_{\boldsymbol{I}} f(\xi) e^{-iq(\xi)} d\xi \right| \le \sigma \|f\|_{L^2(\boldsymbol{I})} \tag{6-1}$$

for all $q \in \mathfrak{Q}$. Otherwise, f is said to be σ-nonuniform in $\mathfrak{Q}$.

Theorem 6.2. *Let L be a bounded sublinear functional from $L^2(\boldsymbol{I})$ to $\mathbb{C}$, let $\boldsymbol{S}_\sigma$ be the set of all functions that are σ-uniform in $\mathfrak{Q}$, and let*

$$U_\sigma = \sup_{f \in \boldsymbol{S}_\sigma} \frac{|L(f)|}{\|f\|_{L^2(\boldsymbol{I})}}. \tag{6-2}$$

Then, for all functions in $L^2(\boldsymbol{I})$,

$$|L(f)| \le \max\{U_\sigma, 2\sigma^{-1} Q\} \|f\|_{L^2(\boldsymbol{I})}, \tag{6-3}$$

where

$$Q = \sup_{q \in \mathfrak{Q}} |L(e^{iq})|. \tag{6-4}$$

Proof. Clearly the complement $\boldsymbol{S}_\sigma^c$ is a set of all functions that are σ-nonuniform in $\mathfrak{Q}$. Let us set

$$A := \sup_{f \in L^2(\boldsymbol{I})} \frac{|L(f)|}{\|f\|_{L^2(\boldsymbol{I})}} \quad \text{and} \quad A_1 := \sup_{f \in \boldsymbol{S}_\sigma^c} \frac{|L(f)|}{\|f\|_{L^2(\boldsymbol{I})}}.$$

Clearly $A = \max\{A_1, U_\sigma\}$. In order to obtain (6-3), it suffices to prove that if $U_\sigma < A_1$, then

$$A_1 \le 2\sigma^{-1} Q. \tag{6-5}$$

For any $\varepsilon > 0$, there exists a function $f \in \boldsymbol{S}_\sigma^c$ such that

$$(A_1 - \varepsilon) \|f\|_{L^2(\boldsymbol{I})} \le |L(f)|. \tag{6-6}$$

Let $\langle \cdot, \cdot \rangle_{\boldsymbol{I}}$ be an inner product on $L^2(\boldsymbol{I})$ defined by

$$\langle f, g \rangle_{\boldsymbol{I}} = \int_{\boldsymbol{I}} f(x) \overline{g(x)}\, dx,$$

for all $f, g \in L^2(\boldsymbol{I})$. Since f is σ-nonuniform in $\mathfrak{Q}$, there exists a function q in $\mathfrak{Q}$ such that

$$\left| \langle f, e^{iq} \rangle_{\boldsymbol{I}} \right| \ge \sigma \|f\|_{L^2(\boldsymbol{I})}. \tag{6-7}$$

There exists $g \in L^2(\boldsymbol{I})$ (depending on f) such that $g \perp e^{iq}$, $\|g\|_{L^2(\boldsymbol{I})} = 1$, and

$$f = \langle f, g \rangle_{\boldsymbol{I}}\, g + \frac{\langle f, e^{iq} \rangle_{\boldsymbol{I}}}{|\boldsymbol{I}|} e^{iq}. \tag{6-8}$$

Sublinearity of L and the triangle inequality then yield

$$\left| L(f) \right| \le \left| \langle f, g \rangle_{\boldsymbol{I}} \right| \left| L(g) \right| + |\boldsymbol{I}|^{-1} \left| \langle f, e^{iq} \rangle_{\boldsymbol{I}} \right| \left| L(e^{iq}) \right|. \tag{6-9}$$

Notice that $A = A_1$ if $U_\sigma < A_1$ and

$$\langle f, f\rangle_{\boldsymbol{I}} = \big|\langle f, g\rangle_{\boldsymbol{I}}\big|^2 + |\boldsymbol{I}|^{-1}\big|\langle f, e^{iq}\rangle_{\boldsymbol{I}}\big|^2. \tag{6-10}$$

Then from (6-6) and (6-9), we have

$$(A_1-\varepsilon)\|f\|_{L^2(\boldsymbol{I})} \le A_1\|f\|_{L^2(\boldsymbol{I})}\sqrt{1-\frac{\big|\langle f, e^{iq}\rangle_{\boldsymbol{I}}\big|^2}{|\boldsymbol{I}|\langle f, f\rangle_{\boldsymbol{I}}}} + |\boldsymbol{I}|^{-1}\big|\langle f, e^{iq}\rangle_{\boldsymbol{I}}\big|Q. \tag{6-11}$$

Applying the elementary inequality $\sqrt{1-x} \le 1 - x/2$ if $0 \le x \le 1$, we then get

$$A_1 \le \frac{2\|f\|_{L^2(\boldsymbol{I})}}{\big|\langle f, e^{iq}\rangle_{\boldsymbol{I}}\big|}Q + \varepsilon|\boldsymbol{I}|\frac{2\|f\|^2_{L^2(\boldsymbol{I})}}{\big|\langle f, e^{iq}\rangle_{\boldsymbol{I}}\big|^2}. \tag{6-12}$$

From (6-7), we have

$$A_1 \le 2\sigma^{-1}Q + 2\varepsilon|\boldsymbol{I}|\sigma^{-2}. \tag{6-13}$$

Now let $\varepsilon \to 0$, and we then obtain (6-5). Therefore we complete the proof. □

7. Estimates of the trilinear forms

We now start to prove Theorem 4.4, and we only present the details for the case $j > 0$, since the other case can be done similarly. Without loss of generality, in the following sections we assume that f_i is supported on $\boldsymbol{I}_i$ for $i \in \{1, 2, 3\}$, where $\boldsymbol{I}_i$ is either $\big[\frac{1}{16}, \frac{39}{16}\big]$ or $\big[-\frac{39}{16}, -\frac{1}{16}\big]$. Let $\mathfrak{Q}_1$ be a set of some functions defined by

$$\mathfrak{Q}_1 = \big\{a\xi^{d/d-1} + b\xi : 2^{m-100} \le |a| \le 2^{m+100} \text{ and } a, b \in \mathbb{R}\big\}. \tag{7-1}$$

Proposition 7.1. *Let $f_1\widehat{\Phi_1}$ be σ-uniform in $\mathfrak{Q}_1$, and let $j > 0$ and $\Lambda_{j,m}(f_1, f_2, f_3)$ be defined as in* (4-12). *Then there exists a constant C independent of j, m, n, f_1 such that*

$$\big|\Lambda_{j,m}(f_1, f_2, f_3)\big| \le C2^{-\frac{(d-1)j}{2}-\frac{m}{2}}\max\big\{2^{-100m}, 2^{\frac{-(d-1)j+m}{2}}, \sigma\big\}\prod_{i=1}^{3}\|f_i\|_{L^2(\boldsymbol{I}_i)} \tag{7-2}$$

holds for all $f_2 \in L^2(\boldsymbol{I}_2)$ and $f_3 \in L^2(\boldsymbol{I}_3)$.

Proof. Since $\mathfrak{B}_{j,m}$ is supported in an interval with size $2^{(d-1)j+m}$, without loss of generality, we may assume that it is restricted to the interval $I_0 = [0, 2^{(d-1)j+m}]$. Let $\mathbf{1}_{m,l} = \mathbf{1}_{\boldsymbol{I}_{m,l}}$, where $\boldsymbol{I}_{m,l} = [2^m l, 2^m(l+1)]$. Also let $\mathfrak{B}_{j,m,l}$ be a bilinear operator defined by

$$\mathfrak{B}_{j,m,l}(f, g)(x) = \mathfrak{B}_{j,m}(f, g)(x)\mathbf{1}_{m,l}(x),$$

for all f, g. Decompose $\Lambda_{j,m}(f_1, f_2, f_3)$ into $\sum_l \Lambda_{j,m,l}$, where

$$\Lambda_{j,m,l}(f_1, f_2, f_3) = \big\langle\mathfrak{B}_{j,m,l}(\check{f}_1, \check{f}_2), \check{f}_3\big\rangle.$$

Let $\alpha_{m,l}$ be a fixed point in the interval $\boldsymbol{I}_{m,l}$. And set $F_{\Phi_1,j,m,l}(x, t)$ to be

$$F_{\Phi_1,j,m,l}(x, t) := R_{\Phi_1}\check{f}_1\big(2^{-(d-1)j}x - 2^m t\big) - R_{\Phi_1}\check{f}_1\big(2^{-(d-1)j}\alpha_{m,l} - 2^m t\big).$$

Split $\mathcal{B}_{j,m,l}(\check{f}_1, \check{f}_2)$ into two terms:

$$\mathcal{B}^{(1)}_{j,m,l}(\check{f}_1, \check{f}_2) + \mathcal{B}^{(2)}_{j,m,l}(\check{f}_1, \check{f}_2),$$

where $\mathcal{B}^{(1)}_{j,m,l}(\check{f}_1, \check{f}_2)$ is equal to

$$2^{-(d-1)j/2}\int_{\mathbb{R}} F_{\Phi_1,j,m,l}(x,t) R_{\Phi_1}\check{f}_2(x-2^m t^d)\rho(t)\,dt\big(\mathbf{1}^*_{I_0}(x)\mathbf{1}_{m,l}(x)\big),$$

and $\mathcal{B}^{(2)}_{j,m,l}(\check{f}_1, \check{f}_2)$ equals

$$2^{-(d-1)j/2}\int_{\mathbb{R}} R_{\Phi_1}\check{f}_1\big(2^{-(d-1)j}\alpha_{m,l}-2^m t\big) R_{\Phi_1}\check{f}_2(x-2^m t^d)\rho(t)\,dt\big(\mathbf{1}^*_{I_0}(x)\mathbf{1}_{m,l}(x)\big).$$

For $i=1,2$, let $\Lambda^{(i)}_{j,m}(f_1, f_2, f_3)$ denote

$$\sum_l \big\langle \mathcal{B}^{(i)}_{j,m,l}(\check{f}_1, \check{f}_2), \check{f}_3\big\rangle.$$

We now start to prove that

$$\big|\Lambda^{(1)}_{j,m}(f_1, f_2, f_3)\big| \le 2^{-(d-1)j/2}2^{-(d-1)j+m}\|\check{f}_1\|_\infty\,\|\check{f}_2\|_2\,\|\check{f}_3\|_2. \tag{7-3}$$

The mean value theorem and the smoothness of Φ_1 yield that for $x\in \boldsymbol{I}_{m,l}$,

$$\big|F_{\Phi_1,j,m,l}(x,t)\big| \le C\|\check{f}_1\|_\infty 2^{-(d-1)j}\big|x-\alpha_{m,l}\big| \le C\,2^{-(d-1)j+m}\|\check{f}_1\|_\infty. \tag{7-4}$$

Because $|t|\sim 1$ when $t\in \operatorname{supp}\rho$, $\mathcal{B}^{(1)}_{j,m,l}(\check{f}_1, \check{f}_2)$ can be written as

$$2^{-(d-1)j/2}\int_{\mathbb{R}} F_{\Phi_1,j,m,l}(x,t)\sum_{l_0}\big(\mathbf{1}_{m,l+l_0}R_{\Phi_1}\check{f}_2\big)(x-2^m t^d)\rho(t)\,dt\big(\mathbf{1}^*_{I_0}(x)\mathbf{1}_{m,l}(x)\big), \tag{7-5}$$

where l_0 is an integer between -10 and 10. Taking absolute values throughout and applying (7-4) plus the Cauchy–Schwarz inequality, we then estimate $\big|\Lambda^{(1)}_{j,m}(f_1, f_2, f_3)\big|$ by

$$C\,2^{-(d-1)j/2}2^{-(d-1)j+m}\|\check{f}_1\|_\infty\sum_{l_0=-10}^{10}\sum_l\big\|\mathbf{1}_{m,l+l_0}R_{\Phi_1}\check{f}_2\big\|_2\big\|\mathbf{1}_{m,l}\check{f}_3\big\|_2,$$

which clearly gives (7-3) by one more use of the Cauchy–Schwarz inequality.

We now prove that

$$\big|\Lambda^{(1)}_{j,m}(f_1, f_2, f_3)\big| \le 2^{-(d-1)j/2}2^{-m}\|\check{f}_1\|_1\,\|\check{f}_2\|_2\,\|\check{f}_3\|_2. \tag{7-6}$$

From (7-5), we get that $\Lambda^{(1)}_{j,m}(f_1, f_2, f_3)$ equals

$$2^{-(d-1)j/2}\sum_{l_0=-10}^{10}\sum_l \Lambda_{j,m,l_0,l,1}(f_1, f_2, f_3)-\Lambda_{j,m,l_0,l,2}(f_1, f_2, f_3),$$

where $\Lambda_{j,m,l_0,l,1}(f_1, f_2, f_3)$ is equal to

$$\int_{\mathbb{R}^2} R_{\Phi_1}\check{f}_1\big(2^{-(d-1)j}x-2^m t\big)\big(\mathbf{1}_{m,l+l_0}R_{\Phi_1}\check{f}_2\big)(x-2^m t^d)\rho(t)\big(\mathbf{1}^*_{I_0}\mathbf{1}_{m,l}\check{f}_3\big)(x)\,dt\,dx$$

and $\Lambda_{j,m,l_0,l,2}(f_1, f_2, f_3)$ equals

$$\int_{\mathbb{R}^2} R_{\Phi_1}\check{f}_1\big(2^{-(d-1)j}\alpha_{m,l} - 2^m t\big)\big(\mathbf{1}_{m,l+l_0}R_{\Phi_1}\check{f}_2\big)(x-2^m t^d)\rho(t)\big(\mathbf{1}^*_{I_0}\mathbf{1}_{m,l}\check{f}_3\big)(x)\,dt\,dx.$$

The Cauchy–Schwarz inequality yields that

$$\big|\Lambda_{j,m,l_0,l,2}(f_1, f_2, f_3)\big| \le C2^{-m}\|\check{f}_1\|_1\big\|\mathbf{1}_{m,l+l_0}R_{\Phi_1}\check{f}_2\big\|_2\|\mathbf{1}_{m,l}\check{f}_3\|_2. \tag{7-7}$$

In order to obtain a similar estimate for $\Lambda_{j,m,l_0,l,1}(f_1, f_2, f_3)$, we change variables by $u = 2^{-(d-1)j}x - 2^m t$ and $v = x - 2^m t^d$ to express $\Lambda_{j,m,l_0,l,1}(f_1, f_2, f_3)$ as

$$\iint R_{\Phi_1}\check{f}_1(u)\big(\mathbf{1}_{m,l+l_0}R_{\Phi_1}\check{f}_2\big)(v)\rho(t(u,v))\big(\mathbf{1}^*_{I_0}\mathbf{1}_{m,l}\check{f}_3\big)(x(u,v))\frac{du\,dv}{J(u,v)},$$

where $J(u,v)$ is the Jacobian $\partial(u,v)/\partial(x,t)$. It is easy to see that the Jacobian $\partial(u,v)/\partial(x,t) \sim 2^m$. As for $\Lambda_{j,m,l_0,l,1}$, we dominate the previous integral by

$$C2^{-m}\int\big|R_{\Phi_1}\check{f}_1(u)\big|\big\|\mathbf{1}_{m,l+l_0}R_{\Phi_1}\check{f}_2\big\|_2\bigg(\int\Big|\big(\mathbf{1}_{m,l}\check{f}_3\big)(x(u,v))\rho(t(u,v))\Big|^2 dv\bigg)^{1/2}du.$$

Notice that $|\partial x/\partial v| \sim 1$ whenever $t \in \operatorname{supp}\rho$. We then estimate

$$\big|\Lambda_{j,m,l_0,l,1}(f_1, f_2, f_3)\big| \le C2^{-m}\|\check{f}_1\|_1\big\|\mathbf{1}_{m,l+l_0}R_{\Phi_1}\check{f}_2\big\|_2\,\|\mathbf{1}_{m,l}\check{f}_3\|_2; \tag{7-8}$$

(7-6) follows from (7-7) and (7-8). An interpolation of (7-3) and (7-6) then yields

$$\big|\Lambda^{(1)}_{j,m}(f_1, f_2, f_3)\big| \le C2^{-\frac{(d-1)j}{2}-\frac{m}{2}}2^{\frac{-(d-1)j+m}{2}}\prod_{i=1}^{3}\|f_i\|_{L^2(\boldsymbol{I}_i)}. \tag{7-9}$$

We now turn to prove that

$$\big|\Lambda^{(2)}_{j,m}(f_1, f_2, f_3)\big| \le C_N 2^{-\frac{(d-1)j}{2}-\frac{m}{2}}\max\{2^{-100m}, \sigma\}\prod_{i=1}^{3}\|f_i\|_{L^2(\boldsymbol{I}_i)}. \tag{7-10}$$

In dual frequency variables, $\Lambda^{(2)}_{j,m}(f_1, f_2, f_3)$ can be expressed as

$$\sum_{l_0=-10}^{10}\sum_{l}2^{-\frac{(d-1)j}{2}}\iint f_1(\xi)\widehat{\Phi}_1(\xi)\exp(2\pi i 2^{-(d-1)j}\alpha_{m,l}\xi)\widehat{F_{2,m,l_0,l}}(\eta)\mathfrak{m}(\xi,\eta)\widehat{F_{3,m,l}}(\eta)\,d\xi\,d\eta,$$

where

$$\mathfrak{m}(\xi,\eta) = \int\rho(t)exp\big(-2\pi i(2^m\xi t + 2^m\eta t^d)\big)\,dt, \tag{7-11}$$

$$F_{2,m,l_0,l} = \mathbf{1}_{m,l+l_0}R_{\Phi_1}\check{f}_2, \quad \text{and} \quad F_{3,m,l} = \mathbf{1}^*_{I_0}\mathbf{1}_{m,l}\check{f}_3.$$

If η is not in a small neighborhood of $\widehat{\Phi}_1$, then there is no critical point of the phase function $\phi_{\xi,\eta}(t) = \xi t + \eta t^d$ occurring in a small neighborhood of $\operatorname{supp}\rho$. Integration by parts gives a rapid decay $O(2^{-Nm})$

for $\mathfrak{m}$. Thus in this case, we dominate $\left|\Lambda^{(2)}_{j,m,n}(f_1, f_2, f_3)\right|$ by

$$C_N 2^{-Nm} \prod_{i=1}^{3} \|f_i\|_{L^2(\boldsymbol{I}_i)}, \tag{7-12}$$

for any positive integer N. We now only need to consider the worst case, when there is a critical point of the phase function $\phi_{\xi,\eta}(t) = \xi t + \eta t^d$ in a small neighborhood of $\operatorname{supp} \rho$. In this case, η must be in a small neighborhood of $\widehat{\Phi}_1$, and the stationary phase method gives

$$\mathfrak{m}(\xi, \eta) \sim 2^{-m/2} \exp\bigl(2\pi i c_d 2^m \eta^{-1/(d-1)} \xi^{d/(d-1)}\bigr), \tag{7-13}$$

where c_d is a constant depending on d only. Thus the principal term of $\Lambda^{(2)}_{j,m}(f_1, f_2, f_3)$ is

$$\sum_{l_0=-10}^{10} \sum_{l} 2^{-\frac{(d-1)j}{2} - \frac{m}{2}} \iint f_1(\xi) \widehat{\Phi}_1(\xi) e^{i\phi_{d,m,\eta}(\xi)} \widehat{F_{2,m,l_0,l}}(\eta) \widehat{\Phi}_2(\eta) \widehat{F_{3,m,l}}(\eta)\, d\xi\, d\eta,$$

where $\widehat{\Phi}_2$ is a Schwartz function supported on a small neighborhood of $\widehat{\Phi}_1$, and

$$\phi_{d,m,\eta}(\xi) = 2\pi c_d 2^m \eta^{-1/(d-1)} \xi^{d/(d-1)} + 2\pi\, 2^{-(d-1)j} \alpha_{m,l} \xi.$$

The key point is that the integral in the previous expression can be viewed as an inner product of $F_{3,m,l}$ and $\mathcal{M} F_{2,m,l_0,l}$, where $\mathcal{M}$ is a multiplier operator defined by

$$\widehat{\mathcal{M} f}(\eta) = \mathfrak{m}_{d,j,m}(\eta) \widehat{f}(\eta).$$

Here the multiplier $\mathfrak{m}_{d,j,m}$ is given by

$$\mathfrak{m}_{d,j,m}(\eta) = \int f_1(\xi) \widehat{\Phi}_1(\xi) e^{i\phi_{d,m,\eta}(\xi)} d\xi. \tag{7-14}$$

Observe that $\phi_{d,m,\eta}(\xi) + b\xi$ is in $\mathfrak{Q}_1$ for any $b \in \mathbb{R}$ and $\eta \in \operatorname{supp} \widehat{\Phi}_2$. Thus σ-uniformity in $\mathfrak{Q}_1$ of f_1 yields

$$\|\mathfrak{m}_{d,j,m}\|_\infty \le C\sigma \|f_1\|_{L^2(\boldsymbol{I}_1)}. \tag{7-15}$$

And henceforth we dominate $\Lambda^{(2)}_{j,m}(f_1, f_2, f_3)$ by

$$\sum_{l_0=-10}^{10} \sum_{l} 2^{-\frac{(d-1)j}{2} - \frac{m}{2}} \sigma \|f_1\|_{L^2(\boldsymbol{I}_1)} \bigl\|F_{2,m,l_0,l}\bigr\|_2 \|F_{3,m,l}\|_2,$$

which clearly is bounded by

$$2^{-\frac{(d-1)j}{2} - \frac{m}{2}} \sigma \prod_{i=1}^{3} \|f_i\|_{L^2(\boldsymbol{I}_i)}. \tag{7-16}$$

Now (7-10) follows from (7-12) and (7-16). Combining (7-9) and (7-10), we finish the proof. □

Corollary 7.2. *Let $\Lambda_{j,m}(f_1, f_2, f_3)$ be defined as in* (4-12). *Then there exists a constant C independent of j, m, n such that*

$$\bigl|\Lambda_{j,m}(f_1, f_2, f_3)\bigr| \le C \max\bigl\{2^{-100m}, 2^{\frac{-(d-1)j+m}{2}}, \sigma\bigr\} \|f_1\|_{L^2(\boldsymbol{I}_1)} \|f_2\|_{L^2(\boldsymbol{I}_1)} \|\widehat{f}_3\|_\infty \tag{7-17}$$

holds for all $f_1 \in L^2(\boldsymbol{I}_1)$ *which are* σ*-uniform in* $\mathfrak{Q}_1$, $f_2 \in L^2(\boldsymbol{I}_2)$ *and* $\widehat{f_3} \in L^\infty$.

Proof. Since there is a smooth restriction factor $\mathbf{1}^*_{I_0}$ in the definition of $\mathcal{B}_{j,m}$, the right-hand side of (7-2) can be sharpened to

$$C2^{-\frac{(d-1)j}{2}-\frac{m}{2}} \max\left\{2^{-100m}, 2^{\frac{-(d-1)j+m}{2}}, \sigma\right\} \|f_1\|_{L^2(\boldsymbol{I}_1)} \|f_2\|_{L^2(\boldsymbol{I}_2)} \left\|\mathbf{1}^{**}_{(d-1)j+m,n}\check{f}_3\right\|_2, \tag{7-18}$$

which is clearly bounded by

$$C \max\left\{2^{-100m}, 2^{\frac{-(d-1)j+m}{2}}, \sigma\right\} \|f_1\|_{L^2(\boldsymbol{I}_1)} \|f_2\|_{L^2(\boldsymbol{I}_1)} \|\widehat{f_3}\|_\infty. \qquad \square$$

Proposition 7.3. *Let* $\Lambda_{j,m}(f_1, f_2, f_3)$ *be defined as in* (4-12). *Then there exists a constant C independent of* j, m, n *such that*

$$\left|\Lambda_{j,m}(e^{iq_1}, f_2, f_3)\right| \le C2^{-\mathfrak{D}(d-1)m/2} \|f_2\|_{L^2(\boldsymbol{I}_2)} \|\widehat{f_3}\|_\infty \tag{7-19}$$

holds for all $q_1 \in \mathfrak{Q}_1$, $f_2 \in L^2(\boldsymbol{I}_2)$ *and* $\widehat{f_3} \in L^\infty$, *where* $\mathfrak{D}(d-1)$ *is the positive constant defined in* (8-3).

A proof of Proposition 7.3 will be provided in Section 8.

Proof of Theorem 4.4. Corollary 7.2, Proposition 7.3 and Theorem 6.2 yield that $\left|\Lambda_{j,m}(f_1, f_2, f_3)\right|$ is dominated by

$$C\left(\max\left\{2^{-100m}, 2^{\frac{-(d-1)j+m}{2}}, \sigma\right\} + \frac{2^{-\mathfrak{D}(d-1)m/2}}{\sigma}\right) \|f_1\|_{L^2(\boldsymbol{I}_1)} \|f_2\|_{L^2(\boldsymbol{I}_1)} \|\widehat{f_3}\|_\infty \tag{7-20}$$

for all $f_1 \in L^2(\boldsymbol{I}_1)$, $f_2 \in L^2(\boldsymbol{I}_2)$ and $\widehat{f_3} \in L^\infty$. Take σ to be $2^{-\mathfrak{D}(d-1)m/4}$; then we have

$$\left|\Lambda_{j,m}(f_1, f_2, f_3)\right| \le C \max\left\{2^{\frac{-(d-1)j+m}{2}}, 2^{-\mathfrak{D}(d-1)m/4}\right\} \|f_1\|_{L^2(\boldsymbol{I}_1)} \|f_2\|_{L^2(\boldsymbol{I}_1)} \|\widehat{f_3}\|_\infty. \tag{7-21}$$

This gives the desired estimate for the case $j > 0$. Similarly, for $j \ne 0$, we have

$$\left|\Lambda_{j,m}(f_1, f_2, f_3)\right| \le C \max\left\{2^{\frac{(d-1)j+m}{2}}, 2^{-m/8}\right\} \|f_1\|_{L^2(\boldsymbol{I}_1)} \|f_2\|_{L^2(\boldsymbol{I}_1)} \|\widehat{f_3}\|_\infty. \tag{7-22}$$

Combining (7-21) and (7-22) proves Theorem 4.4. $\square$

8. Proof of Proposition 7.3

Lemma 8.1. *Let* $l \ge 1$. *Let* $\boldsymbol{I}_1$ *and* $\boldsymbol{I}_2$ *be fixed bounded intervals, and let* $\varphi : \boldsymbol{I}_1 \times \boldsymbol{I}_2 : \mathbb{R}$ *satisfy*

$$\left|\partial_x^l \partial_y \varphi(x, y)\right| \ge 1 \quad \textit{for all } (x, y) \in \boldsymbol{I}_1 \times \boldsymbol{I}_2. \tag{8-1}$$

Assume an additional condition holds in the case $l = 1$:

$$\left|\partial_x^2 \partial_y \varphi(x, y)\right| \ne 0 \quad \textit{for all } (x, y) \in \boldsymbol{I}_1 \times \boldsymbol{I}_2. \tag{8-2}$$

Set

$$\mathfrak{D}(l) = \begin{cases} 1/(2l) & \textit{if } l \ge 2, \\ 1/(2+\varepsilon) & \textit{if } l = 1, \end{cases} \tag{8-3}$$

for some $\varepsilon > 0$. *Then there exists a constant depending on the length of* $\boldsymbol{I}_1$ *and* $\boldsymbol{I}_2$ *but independent of* φ, λ

and the locations of $\boldsymbol{I}_1$ and $\boldsymbol{I}_2$ such that

$$\left|\iint_{\boldsymbol{I}_1\times\boldsymbol{I}_2} e^{i\lambda\varphi(x,y)} f(x)g(y)\,dx\,dy\right| \le C(1+|\lambda|)^{-\mathfrak{D}(l)}\|f\|_2\,\|g\|_2, \quad \text{for all } f, g \in L^2. \tag{8-4}$$

This lemma is related to a two-dimensional van der Corput lemma proved in [Carbery et al. 1999]. The case $l \ge 2$ was proved in [Carbery et al. 1999], and a proof of the case $l = 1$ can be found in [Phong and Stein 1994]. The estimates on $\mathfrak{D}(l)$ in (8-3) are not sharp. With some additional convexity conditions on the phase function φ, one can improve $\mathfrak{D}(l)$ to $1/(l+1)$ (see [Carbery et al. 1999] for some such improvements). But in this article we do not need to pursue the sharp estimates.

Lemma 8.2. *Let $c, \tau \in \mathbb{R}$ and φ be a function defined by*

$$\varphi_c(x, y) = \big(x - y^{1/d} + c\big)^d. \tag{8-5}$$

Define $\boldsymbol{Q}_{c,j.\tau}(x, y)$ by

$$\boldsymbol{Q}_{c,j,\tau}(x, y) = \varphi_c(x, y) - \varphi_c\big(x + 2^{-(d-1)j}\tau, y + \tau\big). \tag{8-6}$$

Then there exists a constant C_d depending only on d such that

$$\big|\partial_x^{d-1}\partial_y \boldsymbol{Q}_{c,j,\tau}(x, y)\big| \ge C_d|\tau| \tag{8-7}$$

for all y such that $y + \tau \in [2^{-100}, 2^{100}]$. Moreover, if $d = 2$,

$$\big|\partial_x\partial_y^2 \boldsymbol{Q}_{c,j,\tau}(x, y)\big| \ge C_d|\tau| \tag{8-8}$$

for all y such that $y + \tau \in [2^{-100}, 2^{100}]$.

Proof. A direct computation yields

$$\partial_x^{d-1}\partial_y \boldsymbol{Q}_{c,j,\tau}(x, y) = C_d\big((y+\tau)^{(1/d)-1} - y^{(1/d)-1}\big). \tag{8-9}$$

Hence the desired estimate (8-7) follows immediately from the mean value theorem. The bound (8-8) can be obtained similarly. □

Lemma 8.3. *Let $\boldsymbol{I}$ be a fixed interval of length 1, and let θ be a bump function supported on $\big[\frac{1}{100}, 2\big]$ (or $\big[-2, -\frac{1}{100}\big]$). Suppose that $\phi_{d,j,m}$ is a phase function defined by*

$$\phi_{d,j,m}(x, y) = C_{d,j,m}2^m\big(x - y^{1/d} + c_{j,m}\big)^d, \tag{8-10}$$

where $C_{d,j,m}$, $c_{j,m}$ are constants independent of x, y such that $2^{-200} \le |C_{d,j,m}| \le 2^{200}$. Let $\Lambda_{d,j,m,\boldsymbol{I}}$ be a bilinear form defined by

$$\Lambda_{d,j,m,\boldsymbol{I}}(f, g) = \iint e^{i\phi_{d,j,m}(x,t)} f\big(x - 2^{-(d-1)j}t\big)g(x)\mathbf{1}_{\boldsymbol{I}}(x)\theta(t)\,dx\,dt. \tag{8-11}$$

Then we have

$$\big|\Lambda_{d,j,m,\boldsymbol{I}}(f, g)\big| \le C_d 2^{-\mathfrak{D}(d-1)m/2}\|f\|_2\,\|g\|_\infty \tag{8-12}$$

for all $f \in L^2$ and $g \in L^\infty$, where C_d is a constant depending only on d.

Proof. The bilinear form $\Lambda_{d,j,m,\boldsymbol{I}}(f,g)$ equals $\langle \boldsymbol{T}_{d,j,m,\boldsymbol{I}}(g), f\rangle$, where $\boldsymbol{T}_{d,j,m,\boldsymbol{I}}$ is defined by

$$\boldsymbol{T}_{d,j,m,\boldsymbol{I}}\, g(x) = \int \exp\bigl(i\phi_{d,j,m}(x+2^{-(d-1)j}t, t)\bigr)(g\mathbf{1}_{\boldsymbol{I}})\bigl(x+2^{-(d-1)j}t\bigr)\theta(t)\,dt. \tag{8-13}$$

By a change of variables, $\|\boldsymbol{T}_{d,j,m,\boldsymbol{I}}\, g\|_2^2$ can be expressed as

$$\int \left(\iint e^{i\Phi_{d,j,m,\tau}(x,t)} G_\tau\bigl(x+2^{-(d-1)j}t\bigr)\Theta_\tau(t)\,dx\,dt\right) d\tau,$$

where

$$\begin{aligned}\Phi_{d,j,m,\tau}(x,t) &= \phi_{d,j,m}\bigl(x+2^{-(d-1)j}t, t\bigr) - \phi_{d,j,m}\bigl(x+2^{-(d-1)j}t+2^{-(d-1)j}\tau, t+\tau\bigr),\\ G_\tau(x) &= (\mathbf{1}_{\boldsymbol{I}}g)(x)\overline{(\mathbf{1}_{\boldsymbol{I}}g)\bigl(x+2^{-(d-1)j}\tau\bigr)},\\ \Theta_\tau(t) &= \theta(t)\theta(t+\tau).\end{aligned}$$

Changing coordinates $(x,t)\mapsto(u,v)$ by $u = x+2^{-(d-1)j}t$ and $v=t$, we write the inner double integral in the previous integral as

$$\iint \exp\bigl(iC_{d,j,m}2^m \boldsymbol{Q}_{c_{j,m},j,\tau}(u,v)\bigr)G_\tau(u)\Theta_\tau(v)\,du\,dv,$$

where $\boldsymbol{Q}_{c_{j,m},j,\tau}$ is defined as in (8-6). From (8-7), (8-8) and Lemma 8.1, we then estimate $\|\boldsymbol{T}_{d,j,m,\boldsymbol{I}}\, g\|_2^2$ by

$$C_d\int_{-10}^{10}\min\bigl\{1, 2^{-\mathfrak{D}(d-1)m}\tau^{-\mathfrak{D}(d-1)}\bigr\}\|G_\tau\|_2\,\|\Theta_\tau\|_2\,d\tau,$$

which clearly is bounded by

$$C_d 2^{-\mathfrak{D}(d-1)m}\|g\|_\infty^2.$$

Hence (8-12) follows and therefore we complete the proof. □

We now turn to the proof of Proposition 7.3. For simplicity, we assume ρ is supported on $\bigl[\frac{1}{8}, 2\bigr]$. For any function $q_1 = a\xi^{d/(d-1)} + b\xi \in \mathfrak{Q}_1$, we have

$$R_{\Phi_1}(\check{e^{iq_1}})(x) = \int \widehat{\Phi}_1(\xi)\exp(ia\xi^{d/(d-1)})\exp(i(x+b)\xi)\,d\xi, \tag{8-14}$$

where $|a|\sim 2^m$. The stationary phase method yields that the principal part of (8-14) is

$$\mathcal{P}(q_1)(x) = C_d|a|^{-1/2}\exp(ic_1a^{-(d-1)}(x+b)^d)\widehat{\Phi}_1\bigl(c_2a^{-(d-1)}(x+b)^{d-1}\bigr), \tag{8-15}$$

where C_d, c_1, c_2 are constants depending only on d. Thus to obtain Proposition 7.3, it suffices to prove that there exists a constant C such that

$$\bigl|\tilde\Lambda_{j,m}(e^{iq_1}, f_2, f_3)\bigr| \le C2^{-\frac{\mathfrak{D}(d-1)m}{2}}\|\check f_2\|_2\,\|\check f_3\|_\infty \tag{8-16}$$

holds for all $q_1\in\mathfrak{Q}_1$, $\check f_2\in L^2$, and $\check f_3\in L^\infty$, where $\tilde\Lambda_{j,m,n}(e^{iq_1}, f_2, f_3)$ is defined to be

$$2^{-(d-1)j/2}\iint \mathcal{P}(q_1)\bigl(2^{-(d-1)j}x - 2^m t\bigr)\check f_2\bigl(x-2^m t^d\bigr)\bigl(\mathbf{1}_{I_0}^*\check f_3\bigr)(x)\rho(t)\,dt\,dx.$$

Observe that $\widehat{\Phi}_1$ is supported essentially in a bounded interval away from 0. Thus we can restrict the variable x in a bounded interval $\boldsymbol{I}_{d,j,m}$ whose length is comparable to $2^{(d-1)j+m}$ and reduce the problem to showing that

$$\bigl|\Lambda_{j,m,\boldsymbol{I}_{d,j,m}}(f_2, f_3)\bigr| \le C 2^{-\frac{\mathfrak{D}(d-1)m}{2}} \|\check{f}_2\|_2 \|\check{f}_3\|_\infty \tag{8-17}$$

for an absolute constant C and all $\check{f}_2 \in L^2$, $\check{f}_3 \in L^\infty$, where $\Lambda_{j,m,n,\boldsymbol{I}_{d,j,m}}(f_2, f_3)$ is equal to

$$2^{-\frac{(d-1)j}{2}-\frac{m}{2}} \iint \mathcal{P}_{d,j,m}(2^{-(d-1)j}x - 2^m t) \check{f}_2(x - 2^m t^d)(\mathbf{1}_{\boldsymbol{I}_{d,j,m}} \check{f}_3)(x) \rho(t)\, dt\, dx. \tag{8-18}$$

Here

$$\mathcal{P}_{d,j,m}(x) = \exp(i c_1 a^{-(d-1)}(x+b)^d) \widehat{\Phi}_1\bigl(c_2 a^{-(d-1)}(x+b)^{d-1}\bigr). \tag{8-19}$$

Let $\boldsymbol{I}$ be an interval of length 1. A rescaling argument then reduces (8-17) to an estimate of a bilinear form $\Lambda_{j,m,n,\boldsymbol{I}}$ associated to $\boldsymbol{I}$, that is,

$$\bigl|\Lambda_{j,m,\boldsymbol{I}}(f, g)\bigr| \le C 2^{-\frac{\mathfrak{D}(d-1)m}{2}} \|f\|_2 \|g\|_\infty, \tag{8-20}$$

where $\Lambda_{j,m,\boldsymbol{I}}(f, g)$ is defined by

$$\iint \mathcal{P}_{d,j,m}\bigl(2^m x - 2^m t\bigr) f\bigl(x - 2^{-(d-1)j} t^d\bigr) g(x) \mathbf{1}_{\boldsymbol{I}}(x) \rho(t)\, dt\, dx.$$

Notice that

$$\mathcal{P}_{d,j,m}(2^m x - 2^m t) = \exp(i C_{d,j,m} 2^m (x - t + c_{j,m})^d) \widehat{\Phi}_1\bigl(C_d C_{d,m}(x - t + c_m)^{d-1}\bigr), \tag{8-21}$$

where $C_{d,j,m}$, $C_{d,m}$, $c_{j,m}$, c_m, C_d are constants such that $|C_{d,j,m}|, |C_{d,m}| \in [2^{-100}, 2^{100}]$. Clearly

$$\widehat{\Phi_1}\bigl(C_d C_{d,m}(x - t + c_m)^{d-1}\bigr)$$

can be dropped by utilizing Fourier series since $\widehat{\Phi}_1$ is a Schwartz function, because $x \in \boldsymbol{I}$, $t \in \operatorname{supp} \rho$ are restricted in bounded intervals. Then (8-20) can be reduced to Lemma 8.3 by a change of variable $t^d \mapsto t$. This proves Proposition 7.3.

9. Appendix

In this appendix, we consider a simple bilinear operator associated to a polynomial curve without singularity. A counterexample is given to indicate that the range of $(1/p, 1/q, 1/r)$ must depend on the degree of the polynomial when the linear term does not vanish. Let ρ be a Schwartz function supported in the union of two intervals $\bigl[-2, -\frac{1}{2}\bigr]$ and $\bigl[\frac{1}{2}, 2\bigr]$.

Lemma 9.1. *Let P be a real polynomial with degree $d \ge 2$. And let $2 \le n \le d$. Suppose that the n-th order derivative of P, $P^{(n)}$, does not vanish. Let $T(f, g)(x) = \int f(x-t) g(x - P(t)) \rho(t)\, dt$. Then T is bounded from $L^p \times L^q$ to L^r for $p, q > 1$, $r > (n-1)/n$ and $1/p + 1/q = 1/r$.*

Proof. We may without loss of generality restrict x, and hence likewise the supports of f, g, to fixed bounded intervals whose sizes depend on the coefficients of the polynomial P. This is possible because of the restriction $|t| \le 2$ in the integral. Let us restrict x in a bounded interval I_P. It is obvious that T is bounded uniformly from $L^\infty \times L^\infty$ to L^∞ and from $L^p \times L^{p'}$ to L^1 for $1 \le p \le \infty$ and $1/p + 1/p' = 1$. When $P'(t) \ne 1$ in $\frac{1}{2} \le |t| \le 2$, the boundedness from $L^1 \times L^1$ to L^1 can be obtained immediately by changing variable $u = x - t$ and $v = x - P(t)$, since the Jacobian $\partial(u, v)/\partial(x, t) = 1 - P'(t)$. Thus T is bounded from $L^1 \times L^1$ to $L^{1/2}$, since x is restricted to a bounded interval I_P, and then the lemma follows by interpolation. When there is a real solution in $\frac{1}{2} \le |t| \le 2$ to the equation $P'(t) = 1$, the trouble happens at a neighborhood of t_0, where $t_0 \in \{t : \frac{1}{2} \le |t| \le 2\}$ is the real solution to $P'(t) = 1$. There are at most $d - 1$ real solutions to the equation $P'(t) - 1 = 0$. Thus we only need to consider a small neighborhood containing only one real solution t_0 to $P'(t) = 1$. Let $I(t_0)$ be a small neighborhood of t_0 which contains only one real solution to $P'(t) - 1 = 0$. We should prove that

$$\int_{I_P} \left| \int_{I(t_0)} f(x-t)g(x-P(t))\rho(t)\,dt \right|^r dx \le C_P \|f\|_p^r \|g\|_q^r, \tag{9-1}$$

for $p > 1$, $q > 1$ and $r > (n-1)/n$ with $1/p + 1/q = 1/r$. Let ρ_0 be a suitable bump function supported in $\frac{1}{2} \le |t| \le 2$ such that $\sum_j \rho_0(2^j t) = 1$. To get (9-1), it suffices to prove that there is a positive ε such that

$$\int_{I_P} \left| \int_{I(t_0)} f(x-t)g(x-P(t))\rho(t)\rho_0(2^j(t-t_0))\,dt \right|^r dx \le C 2^{-\varepsilon j} \|f\|_p^r \|g\|_q^r, \tag{9-2}$$

for all large j, $p > 1$, $q > 1$ and $r > (n-1)/n$ with $1/p + 1/q = 1/r$, since (9-1) follows by summing for all possible $j \ge 1$. By a translation argument, we need to show that

$$\int_{I_P} \left| \int f(x-t)g(x-P_1(t))\rho_0(2^j t)\,dt \right|^r dx \le C 2^{-\varepsilon j} \|f\|_p^r \|g\|_q^r, \tag{9-3}$$

for all large j, $p > 1$, $q > 1$ and $r > (n-1)/n$ with $1/p + 1/q = 1/r$, where P_1 is a polynomial of degree d defined by $P_1(t) = P(t + t_0) - P(t_0)$. It is clear that $P_1'(0) = 1$ and $P_1^{(n)} \ne 0$. When $|t| \le 2^{-j+1}$, $|P_1(t)| \le C_P 2^{-j}$ for some constant $C_P \ge 1$ depending on the coefficients of P. Let $I_P = [a_P, b_P]$ and A_N be defined by

$$A_N = \left[a_P + N C_P 2^{-j}, a_P + (N+1) C_P 2^{-j}\right] \quad \text{for } N = -1, \dots, \frac{(b_P - a_P) \cdot 2^j}{C_P}.$$

Notice that for a fixed $x \in I_P$, $x - t$, $x - P_1(t)$ is in $A_{N-1} \cup A_N \cup A_{N+1}$ for some N. So we can restrict x in one of the A_N. Now let $T_N(f, g)(x) = 1_{A_N}(x) \int f(x-t)g(x-P_1(t))\rho_0(2^j t)\,dt$. Due to the restriction of x, we only need to show that

$$\left\| T_N(f, g) \right\|_r^r \le C 2^{-\varepsilon j} \|f_N\|_p^r \|g_N\|_q^r \tag{9-4}$$

for all large $j \ge 1$, $p > 1$, $q > 1$ and $r > (n-1)/n$ with $1/p + 1/q = 1/r$, where $f_N = f 1_{A_N}$, $g_N = g 1_{A_N}$ and C is independent of N.

By inserting absolute values throughout, we get that T_N maps $L^p \times L^q$ to L^r with a bound $C 2^{-j}$ uniform in N, whenever $(1/p, 1/q, 1/r)$ belongs to the closed convex hull of the points $(1, 0, 1)$, $(0, 1, 1)$ and $(0, 0, 0)$. Observe that $P_1'(t) = 1 + \sum_{k=2}^{d-1} \bigl(P_1^{(k)}(0)/(k-1)!\bigr) t^{k-1}$ since $P_1'(0) = 1$. By $P_1^{(n)}(0) \ne 0$

and applying the Cauchy–Schwarz inequality, we obtain, for all j large enough,

$$\int \big|T_N(f,g)(x)\big|^{1/2}dx \le C_P 2^{-j/2}\|T_N(f,g)\|_1^{1/2}$$
$$\le C_P 2^{-j/2}2^{(n-1)j/2}\|f\|_1^{1/2}\|g\|_1^{1/2} = C_P 2^{(n-2)j/2}\|f\|_1^{1/2}\,\|g\|_1^{1/2}.$$

Hence, an interpolation then yields a bound $C2^{-\varepsilon j}$ for all triples of reciprocal exponents within the convex hull of $\big(1, 1/(n-1), n/(n-1)\big)$, $\big(1/(n-1), 1, n/(n-1)\big)$, $(1,0,1)$, $(0,1,1)$ and $(0,0,0)$. This finishes the proof of (9-4). Therefore we complete the proof of Lemma 9.1 □

Notice that if P is a monomial t^d, then the lower bound for r in Lemma 9.1 can be improved to $\frac{1}{2}$. This is because $P_1(t)=P(t+t_0)-P(t_0)=(t+t_0)^d-t_0^d$ has nonvanishing $P_1^{(2)}(0)$ when $\frac{1}{2}\le|t_0|\le 1$. We now give a counterexample to indicate that the lower bound $(n-1)/n$ for r is sharp in Lemma 9.1.

Proposition 9.2. *Let d, n be integers such that $d\ge 2$ and $2\le n\le d$. There is a real polynomial Q of degree $d\ge 2$ whose n-th order derivative does not vanish such that T_Q is unbounded from $L^p\times L^q$ to L^r for all $p,q>1$ and $r<(n-1)/n$ with $1/p+1/q=1/r$, where T_Q is the bilinear operator defined by $T_Q(f,g)(x)=\int f(x-t)g(x-Q(t))\rho(t)\,dt$.*

Proof. Let A be a very large number. We define $Q(t)$ by

$$Q(t)=\frac{1}{Ad!}(t-1)^d+\frac{1}{An!}(t-1)^n+(t-1). \tag{9-5}$$

It is sufficient to prove that if T_Q is bounded from $L^p\times L^q$ to L^r for some $p,q>1$ and $1/r=1/p+1/q$, then $r\ge(n-1)/n$. Suppose there is a constant C such that $\big\|T_Q(f,g)\big\|_r\le C\|f\|_p\|g\|_q$ for all $f\in L^p$ and $g\in L^q$. Let δ be a small positive number, and let $f_\delta=1_{[0,2^n\delta]}$ and $g_\delta=1_{[1-\delta,1]}$. Let D_1 be the intersection point of the curves $x=Q(t)+1$ and $x=t+2^n\delta$ in the tx-plane with $t>1$, and let D_2 be the intersection point of the curves $x=Q(t)+1-\delta$ and $x=t$ in the tx-plane with $t>1$. Let $D_1=(t_1,x_1)$ and $D_2=(t_2,x_2)$. Then

$$1+2^{1-1/n}(An!)^{1/n}\delta^{1/n}\le t_1\le 1+2(An!)^{1/n}\delta^{1/n}\quad\text{and}$$
$$1+2^{-1/n}(An!)^{1/n}\delta^{1/n}\le t_2\le 1+(An!)^{1/n}\delta^{1/n}.$$

Thus we have

$$1+2^{1-1/n}(An!)^{1/n}\delta^{1/n}+2^n\delta\le x_1\le 1+2(An!)^{1/n}\delta^{1/n}+2^n\delta\quad\text{and}$$
$$1+2^{-1/n}(An!)^{1/n}\delta^{1/n}\le x_2\le 1+(An!)^{1/n}\delta^{1/n}.$$

When A is large and δ is small, any horizontal line between line $x=x_1$ and line $x=x_2$ has a line segment of length $\delta/2$ staying within the region bounded by curves $x=t$, $x=Q(x)+1-\delta$, $x=t+2^n\delta$ and $x=Q(t)+1$. Hence, we have

$$\big\|T_Q(f_\delta,g_\delta)\big\|_r^r\ge\frac{\left(\frac{\delta}{2}\right)^r(An!)^{1/n}\delta^{1/n}}{100}. \tag{9-6}$$

By the boundedness of T_Q, we have

$$\big\|T_Q(f_\delta,g_\delta)\big\|_r^r\le C^r(2^n\delta)^{r/p}\delta^{r/q}=C^r2^{nr/p}\delta.$$

By (9-6), we have

$$\delta^r \le \frac{1002^{r+nr/p} C^r}{(An!)^{1/n}} \delta^{\frac{n-1}{n}}. \tag{9-7}$$

Since A can be chosen to be a very large number and δ can be very small, (9-7) implies $r \ge (n-1)/n$, which completes the proof of Proposition 9.2. $\square$

References

[Carbery et al. 1999] A. Carbery, M. Christ, and J. Wright, "Multidimensional van der Corput and sublevel set estimates", *J. Amer. Math. Soc.* **12**:4 (1999), 981–1015. MR 2000h:42010 Zbl 0938.42008

[Christ 1985a] M. Christ, "Hilbert transforms along curves, I: Nilpotent groups", *Ann. of Math.* (2) **122**:3 (1985), 575–596. MR 87f:42039a Zbl 0593.43011

[Christ 1985b] M. Christ, "Hilbert transforms along curves, II: A flat case", *Duke Math. J.* **52**:4 (1985), 887–894. MR 87f:42039b Zbl 0677.42011

[Christ et al. 1999] M. Christ, A. Nagel, E. M. Stein, and S. Wainger, "Singular and maximal Radon transforms: analysis and geometry", *Ann. of Math.* (2) **150**:2 (1999), 489–577. MR 2000j:42023 Zbl 0960.44001

[Christ et al. 2005] M. Christ, X. Li, T. Tao, and C. Thiele, "On multilinear oscillatory integrals, nonsingular and singular", *Duke Math. J.* **130**:2 (2005), 321–351. MR 2006g:42012 Zbl 1119.42005

[Duoandikoetxea and Rubio de Francia 1986] J. Duoandikoetxea and J. L. Rubio de Francia, "Maximal and singular integral operators via Fourier transform estimates", *Invent. Math.* **84**:3 (1986), 541–561. MR 87f:42046 Zbl 0568.42012

[Fabes and Rivière 1966] E. B. Fabes and N. M. Rivière, "Singular integrals with mixed homogeneity", *Studia Math.* **27** (1966), 19–38. MR 35 #683 Zbl 0161.32403

[Furstenberg 1990] H. Furstenberg, "Nonconventional ergodic averages", pp. 43–56 in *The legacy of John von Neumann* (Hempstead, NY, 1988), edited by J. Glimm et al., Proc. Sympos. Pure Math. **50**, American Mathematical Society, Providence, RI, 1990. MR 91i:28011 Zbl 0711.28006

[Gowers 1998] W. T. Gowers, "A new proof of Szemerédi's theorem for arithmetic progressions of length four", *Geom. Funct. Anal.* **8**:3 (1998), 529–551. MR 2000d:11019 Zbl 0907.11005

[Hörmander 1973] L. Hörmander, "Oscillatory integrals and multipliers on FL^p", *Ark. Mat.* **11** (1973), 1–11. MR 49 #5674 Zbl 0254.42010

[Host and Kra 2005] B. Host and B. Kra, "Nonconventional ergodic averages and nilmanifolds", *Ann. of Math.* (2) **161**:1 (2005), 397–488. MR 2007b:37004 Zbl 1077.37002

[Lacey and Thiele 1997] M. Lacey and C. Thiele, "L^p estimates on the bilinear Hilbert transform for $2 < p < \infty$", *Ann. of Math.* (2) **146**:3 (1997), 693–724. MR 99b:42014 Zbl 0914.46034

[Lacey and Thiele 1999] M. Lacey and C. Thiele, "On Calderón's conjecture", *Ann. of Math.* (2) **149**:2 (1999), 475–496. MR 2000d:42003 Zbl 0934.42012

[Li 2008] X. Li, "Uniform estimates for some paraproducts", *New York J. Math.* **14** (2008), 145–192. MR 2009m:46027 Zbl 1213.42046

[Nagel et al. 1983] A. Nagel, J. Vance, S. Wainger, and D. Weinberg, "Hilbert transforms for convex curves", *Duke Math. J.* **50**:3 (1983), 735–744. MR 85a:42025 Zbl 0524.44001

[Phong and Stein 1994] D. H. Phong and E. M. Stein, "On a stopping process for oscillatory integrals", *J. Geom. Anal.* **4**:1 (1994), 105–120. MR 95k:42025 Zbl 0803.42003

[Sogge 1993] C. D. Sogge, *Fourier integrals in classical analysis*, Cambridge Tracts in Mathematics **105**, Cambridge University Press, Cambridge, 1993. MR 94c:35178 Zbl 0783.35001

[Stein 1993] E. M. Stein, *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals*, Princeton Mathematical Series **43**, Princeton University Press, Princeton, NJ, 1993. MR 95c:42002 Zbl 0821.42001

[Stein and Wainger 1978] E. M. Stein and S. Wainger, "Problems in harmonic analysis related to curvature", *Bull. Amer. Math. Soc.* **84**:6 (1978), 1239–1295. MR 80k:42023 Zbl 0393.42010

Received 13 Oct 2011. Accepted 18 Nov 2011.

XIAOCHUN LI: xcli@math.uiuc.edu
Department of Mathematics, University of Illinois at Urbana-Champaign, Urbana, IL 61801, United States

ANALYSIS AND PDE
Vol. 6, No. 1, 2013
dx.doi.org/10.2140/apde.2013.6.221

A GLUING FORMULA FOR THE ANALYTIC TORSION ON SINGULAR SPACES

MATTHIAS LESCH

To my family

We prove a gluing formula for the analytic torsion on noncompact (i.e., singular) Riemannian manifolds. Let $M = U \cup_{\partial M_1} M_1$, where M_1 is a compact manifold with boundary and U represents a model of the singularity. For general elliptic operators we formulate a criterion, which can be checked solely on U, for the existence of a global heat expansion, in particular for the existence of the analytic torsion in the case of the Laplace operator. The main result then is the gluing formula for the analytic torsion. Here, decompositions $M = M_1 \cup_Y M_2$ along any compact closed hypersurface Y with M_1, M_2 both noncompact are allowed; however a product structure near Y is assumed. We work with the de Rham complex coupled to an arbitrary flat bundle F; the metric on F is not assumed to be flat. In an appendix the corresponding algebraic gluing formula is proved. As a consequence we obtain a framework for proving a Cheeger–Müller-type theorem for singular manifolds; the latter has been the main motivation for this work.

The main tool is Vishik's theory of moving boundary value problems for the de Rham complex which has also been successfully applied to Dirac-type operators and the eta invariant by J. Brüning and the author. The paper also serves as a new, self-contained, and brief approach to Vishik's important work.

1. Introduction

The Cheeger–Müller theorem [Cheeger 1979a; Müller 1978; 1993] on the equality of the analytic and combinatorial torsion is one of the cornerstones of modern global analysis. To extend the theorem to certain singular manifolds is an intriguing open challenge.

Supported by the Hausdorff Center for Mathematics, Bonn.
MSC2010: primary 58J52; secondary 58J05, 58J10, 58J35.
Keywords: analytic torsion, determinants.

In his seminal work, Cheeger [1979b; 1983] initiated the program of "extending the theory of the Laplace operator to certain Riemannian spaces with singularities". Since then a lot of work on this program has been done. It is impossible to give a proper account here, but let us mention [Brüning and Seeley 1988; 1987], Melrose [1993] and collaborators, and Schulze [1991] and collaborators. While the basic spectral theory (index theory, heat kernel analysis) for several types of singularities (cones [Lesch 1997], cylinders [Melrose 1993], cusps [Müller 1983], edges [Mazzeo 1991]) is fairly well understood, an analogue of the Cheeger–Müller theorem has not yet been established for any type of singular manifold, except compact manifolds with boundary.

We will not solve this problem in this paper. However, we will provide a framework for attacking the problem.

To describe this we must go back a little. Let M be a Riemannian manifold (boundaryless but not necessarily compact; also the *interior* of a manifold with boundary is allowed) and let P^0 be an elliptic differential operator acting on the sections $\Gamma^\infty(E)$ of the Hermitian vector bundle E. We consider P^0 as an unbounded operator in the Hilbert space $L^2(M,E)$ of L^2-sections of E. Moreover, we assume P^0 to be bounded below; for example, $P^0 = D^t D$ for an elliptic operator D. Fix a bounded below self-adjoint extension $P \geq -C > -\infty$.

We know that e^{-tP} is an integral operator with a smooth kernel $k_t(x,y)$ which on the diagonal has a pointwise asymptotic expansion

$$k_t(x,x) \sim_{t\searrow 0} \sum_{j=0}^{\infty} a_j(x)\, t^{\frac{j-\dim M}{\operatorname{ord} P}}. \tag{1-1}$$

This asymptotic expansion is *uniform on compact subsets of* M and hence if, e.g., M is compact, it may be integrated over the manifold to obtain an asymptotic expansion for the trace of e^{-tP}. For general noncompact M one cannot expect the operator e^{-tP} to be of trace class. Even if it is of trace class and even if the coefficients $a_j(x)$ in (1-1) are integrable, integration of (1-1) does not necessarily lead to an asymptotic expansion of $\operatorname{Tr}(e^{-tP})$. It is therefore a fundamental problem to give criteria which ensure that e^{-tP} is of trace class and such that there is an asymptotic expansion

$$\operatorname{Tr}(e^{-tP}) \sim_{t\searrow 0} \sum_{\substack{\Re\alpha\to\infty \\ 0\leq k\leq k(\alpha)}} a_{\alpha k}\, t^\alpha \log^k t. \tag{1-2}$$

It is not realistic to find such criteria for arbitrary open manifolds. Instead one looks at geometric differential operators on manifolds with singular exits which occur in geometry. A rather generic description of this situation can be given as follows: suppose that there is a compact manifold $M_1 \subset M$ and a "well understood" model manifold U such that

$$M = U \cup_{\partial M_1} M_1. \tag{1-3}$$

We list a couple of examples for U which are reasonably well understood and which are of geometrical significance:

1. *Smooth boundary.* $U = (0, \epsilon) \times Y$ is a cylinder with metric $dx^2 + g_Y$ over a smooth compact boundaryless manifold Y. Then M is just the interior of a compact manifold with boundary. To this situation the theory of elliptic boundary value problems applies. Heat trace expansions are established, for example, for all well-posed elliptic boundary value problems associated to Laplace-type operators [Grubb 1999].

2. *Isolated asymptotically conical singularities.* $U = (0, \epsilon) \times Y$ with metric $dx^2 + x^2 g_Y(x)$. Then M is a manifold with an isolated (asymptotically) conical singularity. This is the best understood case of a singular manifold; it is impossible here to do justice to all the scientists who contributed. So we just reiterate that its study was initiated by Cheeger [1979b; 1983].

3. *Simple edge singularities.* In the hierarchy of singularities of stratified spaces, which are in general of iterated cone type, this is the next simple class after isolated conical ones: simplifying a little, U is of the form $(0, \epsilon) \times F \times B$ with metric $dx^2 + x^2 g_F(x) + g_B(x)$. The heat trace expansion and the existence of the analytic torsion for this class of singularities has been established recently by Mazzeo and Vertman [2012].

4. *Complete cylindrical ends.* This case is at the heart of Melrose's celebrated b-calculus [1993]. An exact b-metric on $(0, \epsilon) \times Y$ is of the form $dx^2/x^2 + g_Y$. Making the change of variables $x = e^{-y}$ we obtain a metric cylinder $(-\log \epsilon, \infty) \times Y$ with metric $dy^2 + g_Y$. M is then a complete manifold. Therefore, the Laplacian, for example, is essentially self-adjoint. However, it is not a discrete operator and hence its heat operator is not of trace class.

5. *Cusps.* Cusps occur naturally as singularities of Riemann surfaces of constant negative curvature. A cusp is given by $U = (0, \infty) \times Y$ with metric $dx^2 + e^{-2x} g_Y$. Then M has finite volume. As in the previous case, however, the Laplacian is not a discrete operator. In this situation (and also in the previous one) one employs methods from scattering theory. There has been seminal work on this by Werner Müller [1992].

The results of this paper apply to situations where the operator P is discrete (has compact resolvent). This is the case in the examples 1–3 above, but *not* in 4 and 5. Nevertheless we are confident that our method can be extended to relative heat traces and relative determinants, for example, for surfaces of finite area.

To explain our results without becoming too technical, suppose that for $P_U = P \restriction U$ and $P_1 = P \restriction M_1$ (of course suitable extensions have to be chosen for P_U and P_1) we have proved expansions (1-2). Then in terms of a suitable cut-off function φ which is 1 in a neighborhood of M_1 one expects to hold:

Principle 1.1 (Duhamel's principle for heat asymptotics; informal version). If P_U and P_1 are discrete with trace-class heat kernels then so is P and

$$\operatorname{Tr}(e^{-tP}) = \operatorname{Tr}(\varphi e^{-tP_1}) + \operatorname{Tr}((1-\varphi)e^{-tP_U}) + O(t^N) \quad \text{as } t \to 0+, \quad \text{for all } N. \tag{1-4}$$

We reiterate that the heat operator is a *global operator*. On a *closed* manifold its short-time asymptotic expansion is local in the sense that the heat trace coefficients are integrals over local densities as described above. This kind of local behavior cannot be expected on noncompact manifolds. However, Principle 1.1 shows that the heat trace coefficients localize near the singularity; they may still be global *on* the singularity as is the case, for example, for Atiyah–Patodi–Singer boundary conditions [Atiyah et al. 1975].

Principle 1.1 is a folklore theorem which appears in various versions in the literature. In Section 3 below we will prove a fairly general rigorous version of it (Corollary 3.7).

Once the asymptotic expansion (1-2) is in place one obtains, via the Mellin transform, the meromorphic continuation of the ζ-function

$$\zeta(P;s) := \sum_{\lambda\in\operatorname{spec}(P)\setminus\{0\}} \lambda^{-s} = \frac{1}{\Gamma(s)}\int_0^\infty t^{s-1}\operatorname{Tr}((I-\Pi_{\ker P})e^{-tP})\,dt. \tag{1-5}$$

Let us specialize to the de Rham complex. Suppose that we have chosen an ideal boundary condition (essentially this means that we have chosen closed extensions for the exterior derivative) $(\mathcal{D}, D)$ for the de Rham complex such that the corresponding extensions $\Delta_j = D_j^* D_j + D_{j-1} D_{j-1}^*$ of the Laplace operators satisfy (1-2). Then we can form the analytic torsion of $(\mathcal{D}, D)$:

$$\log T(\mathcal{D}, D) := \frac{1}{2}\sum_{j\geq 0}(-1)^j\, j \,\frac{d}{ds}\Big|_{s=0}\zeta(\Delta_j; s). \tag{1-6}$$

For a closed manifold the celebrated Cheeger–Müller theorem [Cheeger 1979a; Müller 1978] relates the analytic torsion to the combinatorial torsion (Reidemeister torsion).

In terms of the decomposition (1-3) the problem of proving a CM-type theorem for the singular manifold M decomposes into the following steps.

(1) Prove that the analytic torsion exists for the model manifold U.

(2) Compare the analytic torsion with a suitable combinatorial torsion for U.

(3) Prove a gluing formula for the analytic and combinatorial torsion and apply the known Cheeger–Müller theorem for the manifold with boundary M_1.

A gluing formula for the combinatorial torsion is more or less an algebraic fact due to Milnor; see also the Appendix. The following theorem, which follows from our gluing formula, solves (3) under a product structure assumption:

Theorem 1.2. *Let M be a singular manifold expressed as* (1-3) *and assume that near ∂M_1 all structures are product. Then for establishing a Cheeger–Müller theorem for M it suffices to prove it for the model space U of the singularity.*

The theorem basically says that, under product assumptions, one gets step (3) for free. Otherwise the specific form of U is completely irrelevant. We conjecture that the product assumption in Theorem 1.2 can be dispensed with. This would follow once the anomaly formula of Brüning and Ma [2006] were established for the model U of the singularity; this would allow us to compare the analytic torsion for (U, g) to the torsion of (U, g_1), where g_1 is product near ∂M_1 and coincides with g outside a relatively compact collar.

The theorem is less obvious than it sounds since torsion invariants are global in nature. However, we will show here that under minimal technical assumptions the analytic torsion satisfies a gluing formula. That the combinatorial torsion satisfies a gluing formula is a purely algebraic fact (see Appendix). The blueprint for our proof is a technique of moving boundary conditions due to Vishik [1995] who applied it to prove the Cheeger–Müller theorem for compact manifolds with smooth boundary. Brüning and the author [Brüning and Lesch 1999] applied Vishik's moving boundary conditions to generalized Atiyah–Patodi–Singer nonlocal boundary conditions and to give an alternative proof of the gluing formula for the eta-invariant. We emphasize, however, that the technical part of the present paper is completely independent of (and in our slightly biased view simpler than) [Vishik 1995]. Also we work with the de Rham complex coupled to an arbitrary flat bundle F. Besides the product structure assumption we do not impose any restrictions on the metric h^F on F; in particular h^F is not assumed to be flat.

We note here that in the context of *closed* manifolds gluing formulas for the analytic torsion have been proved in [Vishik 1995; Burghelea et al. 1999], and recently [Brüning and Ma 2013]. In contrast our method applies to a wide class of singular manifolds.

Some more comments on conic singularities, the most basic singularities, are in order: let (N, g) be a compact closed Riemannian manifold and let $CN = (0, 1) \times N$ with metric $dx^2 + x^2 g$ be the cone over N. We emphasize that sadly near $\partial CN = \{1\} \times N$ we do not have product structure. Let g_1 be a metric on CN that is product near $\{1\} \times N$ and that coincides with g near the cone tip.

Vertman [2009] gave formulas for the torsion of the cone (CN, g) in terms of spectral data of the cone base. What is still not yet understood is how these formulas for the analytic torsion can be related to a combinatorial torsion of the cone, at least not in the interesting odd-dimensional case. For CN even-dimensional, Hartmann and Spreafico [2010] express the torsion of (CN, g) in terms of the intersection torsion introduced by A. Dar [1987] and the anomaly term of Brüning and Ma [2006]. If it were also possible to apply the latter to the singular manifold CN to compare the torsion of the metric cone (CN, g) to that of the cone (CN, g_1), where the metric near $\{1\} \times N$ is modified to a product metric, then one would obtain a (very sophisticated) new proof of Dar's theorem that for an even-dimensional manifold with conical singularities the analytic and the intersection torsions both vanish.[1] It would be more interesting, of course, to have this program worked out in the odd-dimensional case.

The paper is organized as follows. Section 2 serves to introduce some terminology and notation. In a purely functional analytic context we discuss self-adjoint operators with *discrete dimension spectrum*; this terminology is borrowed from Connes and Moscovici's celebrated paper [1995] on the local index theorem in noncommutative geometry. For *Hilbert complexes* [Brüning and Lesch 1992] whose Laplacians have discrete dimension spectrum one can introduce the analytic torsion. We state a formula for the torsion of a product complex (Proposition 2.3) and in Section 2B we collect some algebraic facts about determinants and the torsion of a finite-dimensional Hilbert complex. The main result of the section is Proposition 2.4 which, under appropriate assumptions, provides a variation formula for the analytic torsion of a one-parameter family of Hilbert complexes.

[1] For this to hold one needs to assume that the metric on the twisting bundle F is also flat.

In Section 3 we discuss the gluing of operators in a fairly general setting: we assume that we have two pairs (M_j, P_j^0), $j = 1, 2$, consisting of Riemannian manifolds M_j^m and elliptic operators P_j^0 such that each M_j is the interior of a manifold $\overline{M}_j$ with compact boundary Y ($\overline{M}_j$ is not necessarily compact). Let $W = Y \times (-c, c)$ be a common collar of Y in M_1 and in M_2 such that $\partial M_1 = Y \times \{1\}$ and $\partial M_2 = Y \times \{-1\}$ and such that P_1^0 coincides with P_2^0 over W. Then P_j^0 give rise naturally to a differential operator $P^0 = P_1^0 \cup P_2^0$ on $M := (M_1 \setminus (Y \times (0, c))) \cup_{Y \times \{0\}} (M_2 \setminus (Y \times (-c, 0)))$. Without becoming too technical here we will show in Proposition 3.5 that certain semibounded symmetric extensions P_j, $j = 1, 2$, of P_j^0 satisfying a noninteraction condition (3-18) give rise naturally to a semibounded self-adjoint extension of P^0. Furthermore, if the P_j have discrete dimension spectrum outside W (compare the paragraph before Corollary 3.7), then the operator P has discrete dimension spectrum and up to an error of order $O(t^\infty)$ the short-time heat trace expansion of P can be calculated easily from the corresponding expansions of P_j.

We also prove similar results for perturbed operators of the form $P_j + V_j$, where V_j is a certain non-pseudodifferential operator; such operators will occur naturally in Section 5, our main technical section.

In Section 4 we describe the details of the gluing situation, review Vishik's moving boundary conditions for the de Rham complex in this context, and introduce various one-parameter families of de Rham complexes. The main technical result of the paper is Theorem 4.1 which analyzes the variation of the torsions of these various families of de Rham complexes. The proof of Theorem 4.1 occupies the whole Section 5. The proof is completely independent of Vishik's original approach. The main feature of our proof is a gauge transformation à la Witten of the de Rham complex which transforms the de Rham operator, originally a family of operators with varying domains, onto a family of operators with constant domain; this family can then easily be differentiated by the parameter.

Theorem 6.1 in Section 6 then finally is the main result of the paper, whose proof, thanks to Theorem 4.1, is now more or less an exercise in diagram chasing.

The Appendix contains the analogues of our main results for finite-dimensional Hilbert complexes.

The paper has a somewhat lengthy history. The material of Sections 4 and 5, in the context of smooth manifolds only, was developed in the summer of 1999, while I was on a Heisenberg fellowship in Bonn. In light of the negative feedback received at conferences I felt that the subject was dying and abandoned it.

In recent years there has been revived interest in generalizing the Cheeger–Müller theorem to manifolds with singularities [Mazzeo and Vertman 2012; Vertman 2009; Müller and Vertman 2011; Hartmann and Spreafico 2010]. I noticed that my techniques (an adaption of Vishik's work [1995] plus simple observations based on Duhamel's principle) do not require the manifold to be closed. The bare minimal assumptions required for the analytic torsion to exist ("discrete dimension spectrum"; see Section 2) and a mild but obvious noninteraction restriction on the choice of the ideal boundary conditions (Definition 3.4) for the de Rham complex actually suffice to prove a gluing formula for the analytic torsion. Since a more concise and more accessible account of the long and important paper [Vishik 1995] is overdue anyway, I eventually made a final effort to write up this paper, in part because Werner Müller and Boris Vertman had been pushing me to do so.

2. Operators with meromorphic ζ-function

Let $\mathcal{H}$ be a separable complex Hilbert space, T a nonnegative self-adjoint operator in $\mathcal{H}$ with p-summable resolvent for some $1 \le p < \infty$. The summability condition implies that T is a *discrete* operator; that is, the spectrum of T consists of eigenvalues of finite multiplicity with $+\infty$ being the only accumulation point. Moreover,

$$\operatorname{Tr}(e^{-tT}) = \sum_{\lambda \in \operatorname{spec} T} e^{-t\lambda} = \dim\ker T + O(e^{-t\lambda_1}) \quad \text{as } t \to \infty \tag{2-1}$$

and

$$\operatorname{Tr}(e^{-tT}) = O(t^{-p}) \quad \text{as } t \to 0+. \tag{2-2}$$

Here $\lambda_1 := \min(\operatorname{spec} T \setminus \{0\})$ denotes the smallest nonzero eigenvalue of T.

As a consequence, the ζ-function

$$\zeta(T;s) := \sum_{\lambda \in \operatorname{spec}(T)\setminus\{0\}} \lambda^{-s} = \frac{1}{\Gamma(s)} \int_0^\infty t^{s-1} \operatorname{Tr}((I - \Pi_{\ker T})e^{-tT})\, dt \tag{2-3}$$

is a holomorphic function in the half plane $\Re s > p$; $\Pi_{\ker T}$ denotes the orthogonal projection onto $\ker T$.

Definition 2.1. Following [Connes and Moscovici 1995] we say that T has *discrete dimension spectrum* if $\zeta(T;s)$ extends meromorphically to the complex plane $\mathbb{C}$ such that on finite vertical strips $|\Gamma(s)\zeta(T;s)| = O(|s|^{-N})$, $|\Im s| \to \infty$, for each N. Denote by $\Sigma(T)$ the set of poles of the function $\Gamma(s)\zeta(T;s)$.

It then follows that for fixed real numbers $a < b$ there are only finitely many poles in the strip $a < \Re s < b$. Moreover, as explained, e.g., in [Brüning and Lesch 1999, Section 2], the discrete dimension spectrum condition is equivalent to the existence of an asymptotic expansion

$$\operatorname{Tr}(e^{-tT}) \sim_{t\to 0+} \sum_{\substack{\alpha \in -\Sigma \\ 0 \le k \le k(\alpha)}} a_{\alpha k}\, t^\alpha \log^k t. \tag{2-4}$$

Furthermore, there is the following simple relation between the coefficients of the asymptotic expansion and the principal parts of the Laurent expansion at the poles of $\Gamma(s)\zeta(T;s)$:

$$\Gamma(s)\zeta(T;s) \sim \sum_{\substack{\alpha \in -\Sigma \\ 0 \le k \le k(\alpha)}} \frac{a_{\alpha k}(-1)^k k!}{(s+\alpha)^{k+1}} - \frac{\dim\ker T}{s}. \tag{2-5}$$

2A. ***Hilbert complexes and the analytic torsion.*** We use the convenient language of Hilbert complexes as outlined in [Brüning and Lesch 1992]. Recall that a Hilbert complex $(\mathcal{D}, D)$ consists of a sequence of Hilbert spaces H_j, $0 \le j \le N$, together with closed operators D_j mapping a dense linear subspace $\mathcal{D}_j \subset H_j$ into H_{j+1}. The complex property means that actually $\operatorname{ran} D_j \subset \mathcal{D}_{j+1}$ and $D_{j+1} \circ D_j = 0$. We say that a Hilbert complex has discrete dimension spectrum if all its Laplace operators $\Delta_j = D_j^* D_j + D_{j-1} D_{j-1}^*$ do have discrete dimension spectrum in the sense of Definition 2.1. Note that since Δ_j has compact resolvent, $(\mathcal{D}, D)$ is automatically a Fredholm complex, by [loc. cit., Theorem 2.4]. For a Hilbert complex $(\mathcal{D}, D)$ which is Fredholm, the finite-dimensional cohomology group $H^j(\mathcal{D}, D) = \ker D_j / \operatorname{ran} D_{j-1}$ is the quotient space of the Hilbert space $\ker D_j$ by the closed subspace $\operatorname{ran} D_{j-1}$ and therefore is naturally equipped with a Hilbert space structure. From the Hodge decomposition [Brüning and Lesch 1992, Corollary 2.5]

$$H_j = \ker D_j \cap \ker D_{j-1}^* \oplus \operatorname{ran} D_{j-1} \oplus \operatorname{ran} D_j^* = \ker \Delta_j \oplus \operatorname{ran} D_{j-1} \oplus \operatorname{ran} D_j^*, \tag{2-6}$$

one then sees that the natural isomorphism $\hat{H}^j(\mathcal{D}, D) := \ker \Delta_j = \ker D_j \cap \ker D_{j-1}^* \to H^j(\mathcal{D}, D)$ is an isometric isomorphism. We will always tacitly assume that the cohomology groups are equipped with this natural Hilbert space structure.

Recall the *Euler characteristic*

$$\chi(\mathcal{D}, D) := \sum_{j \geq 0} (-1)^j \dim H^j(\mathcal{D}, D) = \sum_{j \geq 0} (-1)^j \dim \ker \Delta_j. \tag{2-7}$$

The discrete dimension spectrum assumption implies the validity of the *McKean–Singer formula*

$$\chi(\mathcal{D}, D) = \sum_{j \geq 0} (-1)^j \operatorname{Tr}(e^{-t\Delta_j}) \quad \text{for } t > 0. \tag{2-8}$$

Definition 2.2. Let $(\mathcal{D}, D)$ be a Hilbert complex with discrete dimension spectrum. The *analytic torsion* of $(\mathcal{D}, D)$ is defined by

$$\log T(\mathcal{D}, D) := \frac{1}{2} \sum_{j \geq 0} (-1)^j j \frac{d}{ds}\Big|_{s=0} \zeta(\Delta_j; s).$$

If $\zeta(\Delta_j; s)$ has a pole at $s = 0$ then by $\frac{d}{ds}|_{s=0}\zeta(\Delta_j; s)$ we understand the coefficient of s in the Laurent expansion at 0.

Obviously $\log T(\mathcal{D}, D)$ can be defined under the weaker assumption that the function

$$F(\mathcal{D}, D; s) := \frac{1}{2} \sum_{j \geq 0} (-1)^j j \zeta(\Delta_j; s) \tag{2-9}$$

extends meromorphically to $\mathbb{C}$.

The analytic torsion can also be expressed in terms of the *closed* and *coclosed* Laplacians: put

$$\Delta_{j,\mathrm{cl}} := \Delta_j \upharpoonright \operatorname{ran} D_{j-1} = D_{j-1} D_{j-1}^* \upharpoonright \operatorname{ran} D_{j-1}, \tag{2-10}$$

$$\Delta_{j,\mathrm{ccl}} := \Delta_j \upharpoonright \operatorname{ran} D_j^* = D_j^* D_j \upharpoonright \operatorname{ran} D_j^*. \tag{2-11}$$

Note that by definition $\Delta_{0,\mathrm{ccl}} = 0$ and $\Delta_{N,\mathrm{cl}} = 0$ act on the trivial Hilbert space $\{0\}$; recall that N is the length of the Hilbert complex. By the Hodge decomposition (2-6) the operators $\Delta_{j,\mathrm{cl}}$ and $\Delta_{j,\mathrm{ccl}}$ are invertible. Moreover,

$$\Delta_{j+1,\mathrm{cl}} D_j \upharpoonright \operatorname{ran} D_j^* = D_j \Delta_{j,\mathrm{ccl}}. \tag{2-12}$$

Hence the eigenvalues of $\Delta_{j,\mathrm{ccl}}$ and $\Delta_{j+1,\mathrm{cl}}$ coincide including multiplicities. Putting for the moment $A_j := \operatorname{Tr}(e^{-t\Delta_{j,\mathrm{cl}}}) = \operatorname{Tr}(e^{-t\Delta_{j-1,\mathrm{ccl}}})$ for $j \geq 1$ and $A_0 := 0$ we therefore have

$$\operatorname{Tr}(e^{-t\Delta_j}) - \dim H^j(\mathcal{D}, D) = \operatorname{Tr}(e^{-t\Delta_{j,\mathrm{cl}}}) + \operatorname{Tr}(e^{-t\Delta_{j,\mathrm{ccl}}}) = A_j + A_{j+1}, \tag{2-13}$$

and hence

$$\begin{aligned}
&\sum_{j\geq 0}(-1)^j j(\operatorname{Tr}(e^{-t\Delta_j})-\dim H^j(\mathcal{D},D))\\
&\qquad\qquad=\sum_{j\geq 0}(-1)^j j(A_j+A_{j+1})=\sum_{j\geq 0}(-1)^j jA_j-\sum_{j\geq 0}(-1)^j(j-1)A_j\\
&\qquad\qquad=\sum_{j\geq 0}(-1)^j\operatorname{Tr}(e^{-t\Delta_{j,\mathrm{cl}}})=-\sum_{j\geq 0}(-1)^j\operatorname{Tr}(e^{-t\Delta_{j,\mathrm{ccl}}}).
\end{aligned}\tag{2-14}$$

To avoid cumbersome distinction of cases we understand that $\operatorname{Tr}(e^{-t\Delta_{0,\mathrm{ccl}}})=0$.

Proposition 2.3. *Let* $(\mathcal{D}',D')$, $(\mathcal{D}'',D'')$ *be two Hilbert complexes with discrete dimension spectrum. Let* $(\mathcal{D},D):=(\mathcal{D}',D')\hat{\otimes}(\mathcal{D}'',D'')$ *be their tensor product. Denote by* Δ', Δ'', Δ *the Laplacians of* $(\mathcal{D}',D')$, $(\mathcal{D}'',D'')$, $(\mathcal{D},D)$, *respectively.*

Then the function $F(\mathcal{D},D;s):=\frac{1}{2}\sum_{j\geq 0}(-1)^j j\zeta(\Delta_j;s)$ *extends meromorphically to* $\mathbb{C}$. *More precisely, in terms of the corresponding function for the complexes* $(\mathcal{D}',D')$, $(\mathcal{D}'',D'')$ *we have the equations*

$$\chi(\mathcal{D},D)=\chi(\mathcal{D}',D')\cdot\chi(\mathcal{D}'',D''),\tag{2-15}$$

$$F(\mathcal{D},D;s)=\chi(\mathcal{D}',D')\cdot F(\mathcal{D}'',D'';s)+\chi(\mathcal{D}'',D'')\cdot F(\mathcal{D}',D';s);\tag{2-16}$$

in particular

$$\log T(\mathcal{D},D)=\chi(\mathcal{D}',D')\cdot\log T(\mathcal{D}'',D'')+\chi(\mathcal{D}'',D'')\cdot\log T(\mathcal{D}',D').\tag{2-17}$$

Proof. This is an elementary calculation; compare [Vishik 1995, Proposition 2.1] and [Ray and Singer 1971, Theorem 2.5]. Since

$$\Delta_k=\bigoplus_{i+j=k}\Delta'_i\otimes I+I\otimes\Delta''_j,$$

we have

$$\ker(\Delta_k-\lambda)=\bigoplus_{\lambda'+\lambda''=\lambda}\ \bigoplus_{i+j=k}\ker(\Delta'_i-\lambda')\otimes\ker(\Delta''_j-\lambda'').\tag{2-18}$$

This proves (2-15), which follows also from the Künneth theorem for Hilbert complexes [Brüning and Lesch 1992, Corollary 2.15]. Furthermore,

$$\begin{aligned}
&\sum_{k\geq 0}(-1)^k k\operatorname{Tr}e^{-t\Delta_k}\\
&=\sum_{k\geq 0}(-1)^k k\sum_{i+j=k}\ \sum_{\substack{\lambda'\in\operatorname{spec}\Delta'_i\\ \lambda''\in\operatorname{spec}\Delta''_j}}e^{-t\lambda'}e^{-t\lambda''}=\sum_{i,j\geq 0}(-1)^{i+j}(i+j)\sum_{\substack{\lambda'\in\operatorname{spec}\Delta'_i\\ \lambda''\in\operatorname{spec}\Delta''_j}}e^{-t\lambda'}e^{-t\lambda''}\\
&=\Big(\sum_{i\geq 0}(-1)^i\operatorname{Tr}e^{-t\Delta'_i}\Big)\Big(\sum_{j\geq 0}(-1)^j j\operatorname{Tr}e^{-t\Delta''_j}\Big)+\Big(\sum_{j\geq 0}(-1)^j\operatorname{Tr}e^{-t\Delta''_j}\Big)\Big(\sum_{i\geq 0}(-1)^i i\operatorname{Tr}e^{-t\Delta'_i}\Big).
\end{aligned}\tag{2-19}$$

The claim now follows from (2-3) and the McKean–Singer formula (2-8) applied to Δ'_i, Δ''_j. ☐

Next we state an abstract differentiability result (compare [Dai and Freed 1994, Appendix; Bohn 2009, Appendix D]).

Proposition 2.4. *Let $(\mathcal{D}^\theta, D^\theta)$, where $\theta \in J \subset \mathbb{R}$, be a one-parameter family of Hilbert complexes with discrete dimension spectrum. Let $\Delta_j^\theta = (D_j^\theta)^* D_j^\theta + D_{j-1}^\theta (D_{j-1}^\theta)^*$ be the corresponding Laplacians. Assume that*

(1) $H_T(\mathcal{D}^\theta, D^\theta)(t) = \sum_{j\geq 0} (-1)^j j \operatorname{Tr}(e^{-t\Delta_j^\theta})$ *is differentiable in* $(t,\theta) \in (0,\infty) \times J$ *and*
$$\frac{d}{d\theta} H_T(\mathcal{D}^\theta, D^\theta)(t) = t \frac{d}{dt} \operatorname{Tr}(P e^{-t\Delta^\theta}) \tag{2-20}$$
with some operator P *in* $H = \bigoplus_{j\geq 0} H_j$ *with* $P(I+\Delta^\theta)^{-N}$ *bounded for some* N;

(2) Δ^θ *is a graph smooth family of self-adjoint operators with constant domain and* $\dim\ker\Delta^\theta$ *independent of* θ;

(3) *there is an asymptotic expansion*
$$\operatorname{Tr}(P e^{-t\Delta^\theta}) \sim_{t\to 0+} \sum_{\substack{\alpha\in-\Sigma\\ 0\leq k\leq k(\alpha)}} a_{\alpha k}^\theta\, t^\alpha \log^k t, \tag{2-21}$$
which is locally uniformly in θ *and with* $a_{\alpha k}^\theta$ *depending smoothly on* θ;

(4) $a_{0k}^\theta = 0$ *for* $k > 0$; *that is, in the asymptotic expansion* (2-21) *there are no terms of the form* $t^0 \log^k t$ *for* $k > 0$.

Then $\theta \mapsto \log T(\mathcal{D}^\theta, D^\theta)$ *is differentiable and*
$$\frac{d}{d\theta} \log T(\mathcal{D}^\theta, D^\theta) = -\tfrac12 \operatorname*{LIM}_{t\to 0+} \operatorname{Tr}(P e^{-t\Delta^\theta}) + \tfrac12 \operatorname*{LIM}_{t\to\infty} \operatorname{Tr}(P e^{-t\Delta^\theta}) = -\tfrac12 a_{00}^\theta + \tfrac12 \operatorname{Tr}(P \restriction \ker\Delta^\theta).$$

Here $\mathrm{LIM}_{t\to a}$ stands, as usual, for the constant term in the asymptotic expansion as $t \to a$. In (1) we have used the abbreviation $\Delta^\theta := \bigoplus_{j\geq 0} \Delta_j^\theta$.

Proof. Assumptions (2) and (3) of Proposition 2.4 guarantee that in the following we may interchange differentiation by s and by θ:
$$\begin{aligned}
2\frac{d}{d\theta} \log T(\mathcal{D}^\theta, D^\theta) &= \frac{d}{d\theta} \frac{d}{ds}\Big|_{s=0} \frac{1}{\Gamma(s)} \int_0^\infty t^{s-1} \sum_{j\geq 0} (-1)^j j \operatorname{Tr}(e^{-t\Delta_j^\theta} - \Pi_{\ker\Delta_j^\theta}) dt \\
&= \frac{d}{ds}\Big|_{s=0} \frac{1}{\Gamma(s)} \int_0^\infty t^s \frac{d}{dt} \operatorname{Tr}(P e^{-t\Delta^\theta}) dt \\
&= -\frac{d}{ds}\Big|_{s=0} \frac{s}{\Gamma(s)} \int_0^\infty t^{s-1} \operatorname{Tr}(P e^{-t\Delta^\theta}) dt \\
&= -\frac{d}{ds}\Big|_{s=0} \frac{s}{\Gamma(s)} \Big[\Big(\frac{a_{00}^\theta}{s} + c_0^\theta + c_1^\theta s + \cdots \Big) - \frac{\operatorname{Tr}(P \restriction \ker\Delta^\theta)}{s} \Big] \\
&= -a_{00}^\theta + \operatorname{Tr}(P \restriction \ker\Delta^\theta).
\end{aligned} \tag{2-22}$$

Assumption (1) was used in the second equality and assumptions (3), (4) were used in the penultimate equality. Without assumption (4) the higher derivatives of the function $1/\Gamma(s)$ at $s = 0$ would cause additional terms. Assumption (2) guarantees in particular that $\operatorname{Tr}(\Pi_{\ker \Delta_j^\theta})$ is independent of θ. ☐

2B. *Torsion of a finite-dimensional Hilbert complex.* This subsection mainly serves the purpose of fixing some notation. Let H_1, H_2 be finite-dimensional Hilbert spaces. For a linear map $T : H_1 \to H_2$ we put

$$\operatorname{Det}(T) := \det(T^*T)^{1/2}. \tag{2-23}$$

If $T : H_1 \to H_2$, $S : H_2 \to H_3$ are linear maps then obviously $\operatorname{Det}(TS) = \operatorname{Det}(T)\operatorname{Det}(S)$. Furthermore, given orthogonal decompositions $H_j = H_j^{(1)} \oplus H_j^{(2)}$, $j = 1, 2$, such that with respect to these decompositions we have

$$T = \begin{pmatrix} T_1 & T_{12} \\ 0 & T_2 \end{pmatrix}, \tag{2-24}$$

then $\operatorname{Det}(T) = \operatorname{Det}(T_1)\operatorname{Det}(T_2)$.

Let $0 \to C^0 \xrightarrow{d_0} C^1 \xrightarrow{d_1} \cdots \xrightarrow{d_{n-1}} C^n \to 0$ be a finite-dimensional Hilbert complex. Then the torsion of this complex satisfies

$$\log T(C^*, d) = \sum_{p\geq 0} (-1)^p \log \operatorname{Det}(d_p : \ker d_p^\perp \to \operatorname{im} d_p) =: \log \tau(C^*, d). \tag{2-25}$$

Needless to say each finite-dimensional Hilbert complex is automatically a Hilbert complex with discrete dimension spectrum. In fact, since the zeta function is entire in this case, for the Laplacian of the complex the set $\Sigma(\Delta)$ defined in Definition 2.1 then equals the set of poles of the Γ-function, $\{0, -1, -2, \dots\}$.

The following two standard results about the torsion and the determinant will be needed at several places. The first one is elementary; the second one is due to Milnor [1966].

Lemma 2.5. *Let (C_k^*, d^k), $k = 1, 2$, be finite-dimensional Hilbert complexes and $\alpha : (C_1^*, d^1) \to (C_2^*, d^2)$ be a chain isomorphism. Then*

$$\begin{aligned} \log \tau(C_1^*, d^1) = \log \tau(C_2^*, d^2) &+ \sum_{j\geq 0} (-1)^j \log \operatorname{Det}(\alpha_j : C_1^j \to C_2^j) \\ &- \sum_{j\geq 0} (-1)^j \log \operatorname{Det}(\alpha_{j,*} : H^j(C_1^*, d^1) \to H^j(C_2^*, d^2)). \end{aligned} \tag{2-26}$$

Proof. For complexes of length 2 the formula follows directly from (2-24). Then one proceeds by induction on the length of the complexes C_1, C_2. We omit the elementary but a little tedious details. ☐

Proposition 2.6 [Milnor 1966, Theorem 3.1/3.2]. *Let $0 \to C_1 \xrightarrow{\alpha} C \xrightarrow{\beta} C_2 \to 0$ be an exact sequence of finite-dimensional Hilbert complexes and let*

$$\mathcal{H} : 0 \to H^0(C_1) \xrightarrow{\alpha_*} H^0(C) \xrightarrow{\beta_*} H^0(C_2) \xrightarrow{\delta} H^1(C_1) \to \cdots \tag{2-27}$$

be their long exact cohomology sequence. Then

$$\log\tau(C^*,d)=\log\tau(C_1^*,d^1)+\log\tau(C_2^*,d^2)+\log\tau(\mathscr{H})\\ -\sum_{j\geq 0}(-1)^j\log\tau\big(0\to C_1^j\xrightarrow{\alpha}C^j\xrightarrow{\beta}C_2^j\to 0\big). \quad (2\text{-}28)$$

In fact the proposition as stated is a combination of [Milnor 1966, Theorem 3.2] and Lemma 2.5. The last term in (2-28) does not appear in [Milnor 1966, Theorem 3.2] since there one is given *preferred* bases of C_1, C, C_2 which are *compatible*. In our Hilbert complex setting the preferred bases are the orthonormal ones. The last term in (2-28) makes up for the fact that in general it is not possible to choose orthonormal bases of C_1, C, C_2 which are compatible in the sense of [loc. cit.]. For a proof in the more general von Neumann setting see [Burghelea et al. 1999, Theorem 1.14].

For future reference we note that for the acyclic complex $(0\to C_1^j\xrightarrow{\alpha}C^j\xrightarrow{\beta}C_2^j\to 0)$ of length 2 on the right of (2-28) it follows from the definition (2-25) that

$$\log\tau\big(0\to C_1^j\xrightarrow{\alpha}C^j\xrightarrow{\beta}C_2^j\to 0\big)=\tfrac{1}{2}\log\operatorname{Det}\big(C_1^j\xrightarrow{\alpha^*\alpha}C_1^j\big)-\tfrac{1}{2}\log\operatorname{Det}\big(C_2^j\xrightarrow{\beta\beta^*}C_2^j\big). \quad (2\text{-}29)$$

Finally, we remind the reader of the (trivial) fact that if in Proposition 2.6 the complex C equals $C_1\oplus C_2$, α the inclusion and β the projection onto the second summand, then $\log\tau(\mathscr{H})=0$ and $\log\tau(C^*,d)=\log\tau(C_1^*,d^1)+\log\tau(C_2^*,d^2)$.

3. Elementary operator gluing and heat kernel estimates on noncompact manifolds

3A. ***Standing assumptions.*** Let M^m be a Riemannian manifold of dimension m; it is essential to note that M^m is *not* necessarily complete; see Figure 1. Furthermore, let $P_0:\Gamma_c^\infty(M,E)\to\Gamma_c^\infty(M,E)$ be a second-order formally self-adjoint elliptic differential operator acting on the compactly supported sections, $\Gamma_c^\infty(M,E)$, of the Hermitian vector bundle E. We assume that P_0 is bounded below and we fix

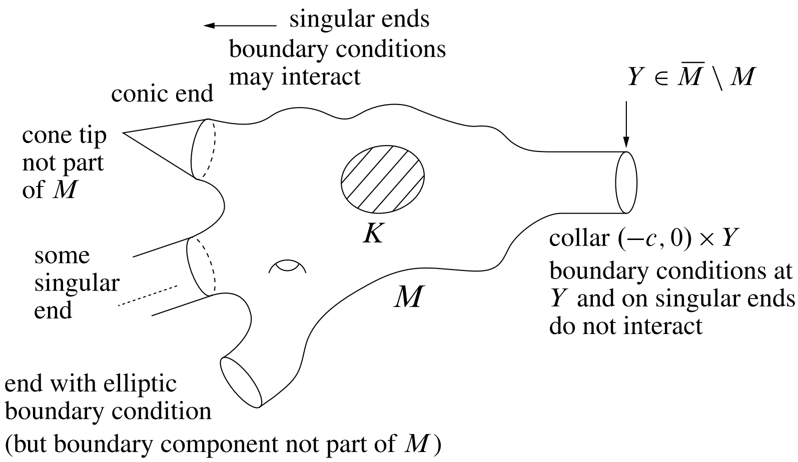

Figure 1. Example of a singular manifold.

once and for all a bounded below self-adjoint extension P of P_0 in the Hilbert space of square-integrable sections $L^2(M, E)$, for example, the Friedrichs extension.

Later on we will need a class of operators which is slightly more general than (pseudo)differential operators. For our purposes it will suffice to consider an auxiliary operator

$$V : H^s_{\mathrm{loc}}(M, E) \to H^{s-1}_{\mathrm{comp}}(M, E), \tag{3-1}$$

which for each real s maps the space $H^s_{\mathrm{loc}}(M, E)$ of sections, which are locally of Sobolev class s, continuously into the space of compactly supported sections of Sobolev class $s-1$; see [Shubin 2001, Section I.7]. We assume that V is symmetric with respect to the L^2-scalar product on E; that is, $\langle Vf, g\rangle = \langle f, Vg\rangle$ for $f \in H^1_{\mathrm{loc}}(M, E)$, $g \in L^2_{\mathrm{loc}}(M, E)$.

Finally, we assume that V is confined to a compact subset $\mathcal{K} \subset M$ in the sense that

$$M_\varphi V = V M_\varphi = 0, \tag{3-2}$$

for any smooth function vanishing in a neighborhood of $\mathcal{K}$. Equation (3-2) implies that V commutes with M_φ for any smooth function which is *constant* in a neighborhood of $\mathcal{K}$. Our main example is the operator $\tilde{\Delta}^\theta$ defined after (5-8) below.

In view of (3-1) and the ellipticity of P_0, the operator V is P-bounded with arbitrarily small bound; thus $P + V$ is self-adjoint and bounded below as well.

With regard to the mapping property (3-1) of V we introduce the space $\mathrm{Op}^a_c(M, E)$ of linear operators A mapping H^s_{loc} continuously into H^{s-a}_{comp} and whose Schwartz kernel K_A is compactly supported. Obvious examples are pseudodifferential operators with compactly supported Schwartz kernel, but also certain Fourier integral operators. The point is that elements in Op_c are not necessarily pseudolocal. Note that V is in $\mathrm{Op}^1_c(M, E)$.

The set-up outlined in this Section 3A will be in effect during the remainder of Section 3.

3B. *Heat kernel estimates for $P + V$.*

Lemma 3.1. *For all $s \geq 0$ we have $\mathcal{D}(P + V)^s = \mathcal{D}(P^s)$. Furthermore, the operator $e^{-t(P+V)}$, $t > 0$, has a smooth integral kernel.*

Proof. By complex interpolation [Taylor 1996, Section 4.2] it suffices to prove the first claim for $s = k \in \mathbb{N}$ where it follows easily by induction exploiting the elliptic regularity for P and (3-1).

Consequently, $e^{-t(P+V)}$ is a self-adjoint operator which maps $L^2(M, E)$ into

$$\bigcap_{k\geq 0} \mathcal{D}((P + V)^k) = \bigcap_{k\geq 0} \mathcal{D}(P^k), \tag{3-3}$$

and the latter is contained in $\Gamma^\infty(M, E)$ by elliptic regularity. This implies smoothness of the kernel of $e^{-t(P+V)}$. □

Proposition 3.2. *Let $A \in \mathrm{Op}^a_c(M, E)$, $B \in \mathrm{Op}^b_c(M, E)$ with compactly supported Schwartz kernels K_A, K_B. Denote by $\pi_j : M \times M \to M$, $j = 1, 2$, the projections onto the first and second factor and suppose that $\pi_2(\mathrm{supp}\, K_A) \cap \pi_1(\mathrm{supp}\, K_B) = \varnothing$ and $\pi_2(\mathrm{supp}\, K_A) \cap \mathcal{K} = \varnothing$.* (For $\mathcal{K}$, see Section 3A.)

Then $Ae^{-t(P+V)}B$ is a trace class operator and

$$\|Ae^{-t(P+V)}B\|_{\mathrm{tr}} = O(t^\infty), \quad t \to 0+. \tag{3-4}$$

Here $O(t^\infty)$ is an abbreviation for $O(t^N)$ for any N; the O-constant may depend on N. Furthermore, $\|\cdot\|_{\mathrm{tr}}$ denotes the trace norm on the Schatten ideal of trace class operators.

Proof. (Compare [Lesch 1997, Section I.4].) Since the Schwartz kernels are compactly supported it suffices to prove that for all real α, β and all $N > 0$ we have

$$\|Ae^{-t(P+V)}B\|_{\alpha,\beta} = O(t^N), \quad t \to 0+. \tag{3-5}$$

Here, $\|\cdot\|_{\alpha,\beta}$ stands for the mapping norm between the Sobolev spaces

$$H^\alpha(\pi_2(\operatorname{supp} K_B), E) \quad \text{and} \quad H^\beta(\pi_1(\operatorname{supp} K_A), E).$$

The O-constant may depend on A, B, α, β, N.

Equation (3-5) follows from Duhamel's formula by a standard bootstrapping argument as follows: note first that the mapping properties of A, B and $P + V$ imply that, for real α,

$$\|Ae^{-t(P+V)}B\|_{\alpha,\alpha-a-b} = O(1), \quad t \to 0+. \tag{3-6}$$

Assume by induction that, for fixed l, N, for all A, B satisfying our assumptions and for all real α,

$$\|Ae^{-t(P+V)}B\|_{\alpha,\alpha-a-b+l} = O(t^N), \quad t \to 0+. \tag{3-7}$$

Fix plateau functions χ, φ, $\psi \in C_c^\infty(M)$ with the following properties:

(1) $\varphi \equiv 1$ in a neighborhood of $\pi_2(\operatorname{supp} K_A)$ and $\operatorname{supp}\varphi \cap \mathscr{K} = \varnothing$.

(2) $\psi \equiv 1$ in a neighborhood of $\pi_1(\operatorname{supp} K_B)$.

(3) $\chi \equiv 1$ in a neighborhood of $\operatorname{supp}\varphi$ and $\operatorname{supp}\chi \cap \mathscr{K} = \varnothing$.

(4) $\operatorname{supp}\chi \cap \operatorname{supp}\psi = \varnothing$.

Then

$$\begin{aligned}\|Ae^{-t(P+V)}B\|_{\alpha,\alpha-a-b+l+1/2} &= \|A\varphi e^{-t(P+V)}\psi B\|_{\alpha,\alpha-a-b+l+1/2} \\ &\le C_1\|\varphi e^{-t(P+V)}\psi\|_{\alpha-b,\alpha-b+l+1/2}.\end{aligned} \tag{3-8}$$

From

$$(\partial_t + P + V)\varphi e^{-t(P+V)}\psi = \chi[P_0,\varphi]e^{-t(P+V)}\psi, \tag{3-9}$$

where $[P_0,\varphi]$ denotes the commutator between the differential expression P_0 and multiplication by φ, we infer

$$\varphi e^{-t(P+V)}\psi = \int_0^t \chi e^{-(t-s)(P+V)}\chi[P_0,\varphi]e^{-s(P+V)}\psi\, ds; \tag{3-10}$$

here we have used the assumptions on the supports of χ, ψ, φ and (3-2). In the displayed formulas we wrote, to save some space, χ, ψ, φ for the multiplication operators M_χ, M_ψ, M_φ.

For $\tilde\alpha = \alpha - b$ we now find

$$\|\varphi e^{-t(P+V)}\psi\|_{\tilde\alpha,\tilde\alpha+l+\frac12} \le \int_0^t \|\chi e^{-(t-s)(P+V)}\chi\|_{\tilde\alpha-1+l,\tilde\alpha+l+\frac12}\, \|[P_0,\varphi]e^{-s(P+V)}\psi\|_{\tilde\alpha,\tilde\alpha-1+l}\, ds. \tag{3-11}$$

Since $[P_0,\varphi]$ is in Op_c^1 we find, using (3-7),

$$\|[P_0,\varphi]e^{-s(P+V)}\psi\|_{\tilde\alpha,\tilde\alpha-1+l} = O(s^N) \quad \text{as } s \to 0+. \tag{3-12}$$

Furthermore, denoting by C a constant such that $P \ge -C+1$,

$$\begin{aligned} &\|\chi e^{-u(P+V)}\chi\|_{\tilde\alpha-1+l,\tilde\alpha+l+\frac12} \\ &\le \|(P+V+C)^{\frac{\tilde\alpha+l-1}{2}}\chi\|_{\tilde\alpha-1+l,0}\, \|(P+V+C)^{\frac34}e^{-u(P+V)}\|_{0,0}\, \|\chi(P+V+C)^{-\frac{\tilde\alpha+l+1/2}{2}}\|_{0,\tilde\alpha+l+\frac12}. \end{aligned} \tag{3-13}$$

The first and the third factors on the right are bounded while for the second factor we have, by the spectral theorem,

$$\|(P+V+C)^{3/4}e^{-u(P+V)}\|_{0,0} = O(u^{-3/4}) \quad \text{as } u \to 0+. \tag{3-14}$$

Thus

$$\|\varphi e^{-t(P+V)}\psi\|_{\tilde\alpha,\tilde\alpha+l+1/2} \le C_1 \int_0^t (t-s)^{-3/4}s^N\, ds = O(t^{N+1/4}), \quad t \to 0+. \tag{3-15}$$

Thus we have improved the parameters l and N in (3-7) by $\frac12$ and $\frac14$, respectively, and therefore the result follows by induction. □

Proposition 3.3. *Under the standing assumptions of Section 3A, let $\varphi, \psi \in C^\infty(M)$ with* $\operatorname{supp}\varphi \cap \operatorname{supp}\psi$ *being compact (the individual supports of φ or ψ may be noncompact!) such that $d\varphi$, $d\psi$ are compactly supported and that* $\operatorname{supp} d\varphi \cap \mathcal{K} = \varnothing = \operatorname{supp} d\psi \cap \mathcal{K}$. *Furthermore, assume that multiplication by φ and by ψ preserves $\mathcal{D}(P+V) = \mathcal{D}(P)$.*

Then for $t > 0$ the operator $\varphi e^{-t(P+V)}\psi$ is trace class and

$$\|\varphi e^{-t(P+V)}\psi\|_{\mathrm{tr}} = O(t^{-m/2-0}), \quad t \to 0+. \tag{3-16}$$

If $\operatorname{supp}\varphi \cap \operatorname{supp}\psi = \varnothing$ *then the right-hand side can be improved to $O(t^\infty)$, $t \to 0+$.*

Here $O(t^{-m/2-0})$ is an abbreviation for $O(t^{-m/2-\epsilon})$ for any $\epsilon > 0$; the O-constant may depend on ϵ.

Proof. Assume first that additionally ψ is compactly supported. Again applying Duhamel we find

$$\varphi e^{-t(P+V)}\psi = \int_0^t e^{-(t-s)(P+V)}[P_0,\varphi]e^{-s(P+V)}\psi\, ds. \tag{3-17}$$

Now apply Lemma 3.1 and Proposition 3.2 to the operator $[P_0,\varphi]e^{-s(P+V)}\psi$. If $\operatorname{supp}\varphi \cap \operatorname{supp}\psi \ne \varnothing$ then the trace norm estimate is a simple consequence of Sobolev embedding and the established mapping

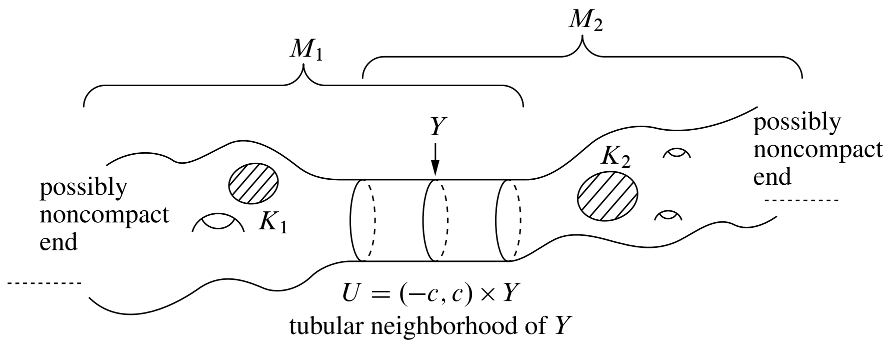

Figure 2. The gluing situation.

properties. If $\operatorname{supp} \varphi \cap \operatorname{supp} \psi = \varnothing$ then Proposition 3.2 implies $\|[P_0, \varphi]e^{-s(P+V)}\psi\|_{\mathrm{tr}} = O(t^\infty)$ and the claim follows in this case.

Since $e^{-t(P+V)}$ is self-adjoint the roles of φ, ψ may be interchanged by taking adjoints and hence the proposition is proved if φ *or* ψ is compactly supported. The general case now follows from (3-17) since the compactness of $\operatorname{supp} d\varphi$ implies the compactness of the support of the Schwartz kernel of $[P_0, \varphi]$. □

3C. *Operator gluing.* Now we assume that we have two triples (M_j, P_j^0, V_j), $j = 1, 2$, consisting of Riemannian manifolds M_j^m and operators P_j^0, V_j satisfying the standing assumptions of Section 3A.

Furthermore, we assume that each M_j is the interior of a manifold $\overline{M}_j$ with compact boundary Y (it is essential that $\overline{M}_j$ is not necessarily compact). Let $U = Y \times (-c, c)$ be a common collar of Y in M_1 and in M_2, such that $\partial M_1 = Y \times \{1\}$ and $\partial M_2 = Y \times \{-1\}$.

We assume that the sets $\mathcal{K}_j$ corresponding to V_j (cf. (3-2)) lie in $M_j \setminus U$ and that P_1^0 coincides with P_2^0 over U. Then the P_j^0 gives rise naturally to a differential operator $P^0 = P_1^0 \cup P_2^0$ on

$$M := \big(M_1 \setminus (Y \times (0, c))\big) \cup_{Y \times \{0\}} \big(M_2 \setminus (Y \times (-c, 0))\big),$$

and the V_j to an operator $V = V_1 + V_2 \in \mathrm{Op}_c^1(M, E)$, where E is the bundle obtained by gluing the bundles E_1 and E_2 in the obvious way (due to (3-2) the operators V_1, V_2 extend to M in a natural way).

Definition 3.4. By $C_U^\infty(M_j)$ we denote the space of those smooth functions $\varphi \in C^\infty(M_j)$ such that φ is constant in a neighborhood of $M_j \setminus U$ and $\varphi \equiv 0$ in a neighborhood of $\partial \overline{M}_j$; see Figure 3.

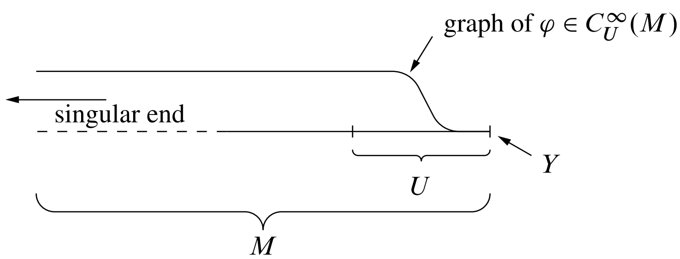

Figure 3. Schematic sketch of a function in $C_U^\infty(M)$. The line indicates the manifold M; to the left are the possible noncomplete ends. On the right there is the collar U.

A function $\varphi \in C_U^\infty(M_j)$ extends by 0 to a smooth function on M.

Proposition 3.5. *Let P_j, $j = 1, 2$, be closed symmetric extensions of P_j^0 which are bounded below and for which*

$$\varphi \mathcal{D}(P_j^*) \subset \mathcal{D}(P_j) \quad \textit{for all } \varphi \in C_U^\infty(M_j). \tag{3-18}$$

For a fixed pair of functions $\varphi_j \in C_U^\infty(M_j)$, $j = 1, 2$, put

$$\mathcal{D}(P) := \big\{ f \in \mathcal{D}(P_{\max}^0) \;\big|\; \varphi_j f \in \mathcal{D}(P_j),\ j = 1, 2 \big\} = H_{\mathrm{comp}}^2(U, E) + \varphi_1 \mathcal{D}(P_1) + \varphi_2 \mathcal{D}(P_2). \tag{3-19}$$

$\mathcal{D}(P)$ is indeed independent of the particular choice of φ_j and the operator P which is defined by restricting $P_{\max}^0 = (P^0)^*$ to $\mathcal{D}(P)$ is self-adjoint and bounded below. V is P-bounded with arbitrarily small bound and hence $P + V$ is self-adjoint and bounded below as well.

Furthermore, if for fixed $j \in \{1, 2\}$ we have φ, $\psi \in C_U^\infty(M_j)$ satisfying (3-18), *then the operator $\varphi e^{-t(P_j+V)}\psi - \varphi e^{-t(P+V)}\psi$ is trace class and its trace norm is $O(t^\infty)$ as $t \to 0+$.*

Remark 3.6. 1. It is not assumed that $\varphi e^{-t(P_j+V)}\psi$ or $\varphi e^{-t(P+V)}\psi$ is of trace class individually!

2. Equation (3-18) says that the "boundary conditions" at the exits of M_1 and M_2 are separated. We illustrate this with an example: let $M_1 = \big(-1, \tfrac{1}{2}\big)$, $M_2 = \big(-\tfrac{1}{2}, 1\big)$, $U = \big(-\tfrac{1}{2}, \tfrac{1}{2}\big)$, $M = (-1, 1)$, and let $P_1^0 = P_2^0 = -d^2/dx^2 = \Delta$ be the Laplacian on functions. Let P_1^{per} be the Laplacian Δ on M_1 with periodic boundary conditions. These boundary conditions are not separated and indeed for $\varphi \in C^\infty\big(-1, \tfrac{1}{2}\big)$ with $\varphi(x) = 1$ for $x \le -\tfrac{1}{4}$ and $\varphi(x) = 0$ for $x \ge \tfrac{1}{4}$ the space $\varphi \mathcal{D}(P_1^{\mathrm{per}})$ equals $\varphi H^2\big(\big[-1, \tfrac{1}{2}\big]\big)$ and this is not contained in $\mathcal{D}(P_1^{\mathrm{per}})$.

However, for any pair of self-adjoint extensions P_j of P_j^0, $j = 1, 2$, with separated boundary conditions at the ends of the intervals M_j one has $\varphi \mathcal{D}(P_j) \subset \mathcal{D}(P_j)$; that is, the condition (3-18) is satisfied and Proposition 3.5 applies to this pair.

Proof. Since $H_{\mathrm{comp}}^2(U, E) \subset \mathcal{D}(P_{j,\min}^0)$ the second equality in (3-19), the symmetry of P and the independence of $\mathcal{D}(P)$ of the particular choice of φ_j are easy consequences of (3-18).

To prove self-adjointness let $f \in \mathcal{D}(P^*)$. We claim that for $\varphi_1 \in C_U^\infty(M_1)$ we have $\varphi_1 f \in \mathcal{D}(P_1^*)$. Indeed for $g \in \mathcal{D}(P_1)$ we have

$$\langle \varphi_1 f, P_1 g \rangle = \langle f, \overline{\varphi}_1 P_1 g \rangle = \langle f, [\overline{\varphi}_1, P_1^0] g \rangle + \langle f, P \overline{\varphi}_1 g \rangle. \tag{3-20}$$

Since $\operatorname{supp} d\varphi_1 \subset U$ is compact and since $[\overline{\varphi}_1, P_1^0]$ is a compactly supported first-order differential operator on U we find

$$\cdots = \langle [P_1^0, \varphi_1] f + \varphi_1 P^* f, g \rangle, \tag{3-21}$$

proving $\varphi_1 f \in \mathcal{D}(P_1^*)$. In view of (3-18) we see, by choosing another plateau function $\psi \in C_U^\infty(M_1)$ with $\psi \varphi_1 = \varphi_1$, that $\varphi_1 f \in \mathcal{D}(P_1)$. In the same way we conclude $\varphi_2 f \in \mathcal{D}(P_2)$ for $\varphi_2 \in C_U^\infty(M_2)$ and thus $f \in \mathcal{D}(P)$.

To prove the trace class property and the trace estimate we choose another plateau function $\chi \in C_U^\infty(M_j)$ such that $\chi \equiv 1$ in a neighborhood of $\operatorname{supp} \psi$ with $\chi - \psi \in C_c^\infty(M_j)$; hence χ also satisfies (3-18).

Consider first $K_t := \chi e^{-t(P_j+V_j)}\psi - \chi e^{-t(P+V)}\psi$. $K_{t=0} = 0$ and

$$(\partial_t + P + V)K_t = [P_j^0, \chi]e^{-t(P_j+V_j)}\psi - [P^0, \chi]e^{-t(P+V)}\psi. \tag{3-22}$$

Here we have used that multiplication by χ commutes with V and V_j; see (3-2). Propositions 3.2 and 3.3 now imply that K_t is trace class for $t > 0$ and that $\|K_t\|_{\mathrm{tr}} = O(t^\infty)$ as $t \to 0+$. Consequently

$$\|\chi\varphi e^{-t(P_j+V)}\psi - \chi\varphi e^{-t(P+V)}\psi\|_{\mathrm{tr}} \le \|\varphi\|_\infty \, \|K_t\|_{\mathrm{tr}} = O(t^\infty).$$

To $(1-\chi)\varphi e^{-t(P_j+V)}\psi - (1-\chi)\varphi e^{-t(P+V)}\psi$ we can apply Proposition 3.3 since $(\operatorname{supp}\psi) \cap \operatorname{supp}(1-\chi) = \varnothing$ and the proof is complete. □

Finally, we discuss heat expansions. Under the assumptions of Proposition 3.5 assume that $P_j + V_j$ has *discrete dimension spectrum outside* U. By this we understand that for $\varphi \in C_U^\infty(M_j)$ the operator $\varphi e^{-t(P_j+V_j)}$ is trace class and that there is an asymptotic expansion of the form (2-4) with $a_{\alpha k} = a_{\alpha k}(\varphi)$. Then:

Corollary 3.7. *Under the additional assumption of discrete dimension spectrum for $P_j + V_j$ outside U the operator $P + V$ has discrete dimension spectrum and for any $\varphi \in C_U^\infty(M_1)$ we have*

$$\operatorname{Tr}(e^{-t(P+V)}) = \operatorname{Tr}(\varphi e^{-t(P_1+V_1)}) + \operatorname{Tr}((1-\varphi)e^{-t(P_2+V_2)}) + O(t^\infty) \quad \textit{as } t \to 0+. \tag{3-23}$$

Proof. This is immediate from Proposition 3.5 and the discrete dimension spectrum assumption. □

We add, however, a little more explanation since the term "discrete dimension spectrum outside U" might lead to some confusion: since $\mathscr{H} \cap U = \varnothing$ (cf. (3-2) and the second paragraph of this section, on page 236) for $f \in \Gamma_c^\infty(U, E)$ we have $(P+V)f = Pf$. The classical interior parametric elliptic calculus (see [Shubin 2001], for example) then implies that for $\varphi \in C_c^\infty(U)$ there is an asymptotic expansion

$$\operatorname{Tr}(\varphi e^{-t(P+V)}) \sim_{t\searrow 0} \sum_{j\ge 0} a_j(P,\varphi)\, t^{j-m/2}, \tag{3-24}$$

where $a_j(P,\varphi) = \int_M \tilde{a}_j(x,P)\varphi(x)\,dx$ and $\tilde{a}_j(x,P)$ are the local heat invariants of P. Thus over any compact subset in the *interior* of $M \setminus \mathscr{H}$ the discrete dimension spectrum assumption follows from standard elliptic theory and hence is a nonissue. Rather it is a condition on the behavior of P on noncompact "ends" and a condition on V over $\mathscr{H}$.

3D. ***Ideal boundary conditions with discrete dimension spectrum.*** The remarks of Section 3C extend to ideal boundary conditions of elliptic complexes in a straightforward fashion. Let X be a Riemannian manifold which is the interior of a Riemannian manifold $\overline{X}$ with compact boundary Y, and let $U = (-c, 0) \times Y$ be a collar of the boundary. Since $\overline{X}$ is allowed to be noncompact it is not excluded that away from U there are "ends" of $\overline{X}$ which can be completed by adding another boundary component; see Figure 1.

As an example which illustrates what can happen consider a compact manifold Z with boundary, where $\partial Z = Y_1 \cup Y_2 \cup Y_3$ consists of the disjoint union of three compact closed manifolds Y_j, $j = 1, 2, 3$. Attach a cone $C(Y_3) = Y_3 \times (0,1)$ with metric $dr^2 + r^2 g_{Y_3}$ to Y_3 (and smooth it out near $Y_3 \times \{1\}$).

Then put $X := (Z \setminus (Y_1 \cup Y_2)) \cup_{Y_3} C(Y_3)$ and $\overline{X} := (Z \setminus Y_2) \cup_{Y_3} C(Y_3)$. Then Y_1 plays the role of Y above, but $\overline{X}$ is not compact. Compare Figure 1 on page 232.

When introducing closed extensions (that is to say, boundary conditions) for elliptic operators on X it is important that the boundary conditions at Y_1 and Y_2 on the one hand, and on the cone on the other, do *not* interact in order to ensure (3-18) holds.

Leaving this example behind, let $(\Gamma_c^\infty(E), d)$ be an elliptic complex and let $(\mathcal{D}, D)$ be an ideal boundary condition for $(\Gamma_c^\infty(E), d)$ — that is, a Hilbert complex such that D_j, $j = 1, 2$, is a closed extension of d_j.

We say that the ideal boundary condition $(\mathcal{D}, D)$ has discrete dimension spectrum outside U if the Laplacians $\Delta_j = D_j^* D_j + D_{j-1} D_{j-1}^*$ have discrete dimension spectrum outside U; cf. the paragraph before Corollary 3.7. Then Proposition 3.5 and Corollary 3.7 hold for the Laplacians.

More concretely, let X, Y be as before and let (F, ∇) be a flat bundle over $\overline{X}$. Assume that we are given an ideal boundary condition $(\mathcal{D}, D)$ of the de Rham complex $(\Omega^\bullet(X; F), d)$ with values in the flat bundle F with discrete dimension spectrum over the open set $X \setminus U$, $U = (-c, 0) \times Y$. Fix a smooth function $\varphi \in C^\infty(-c, 0)$ which is 1 near $-c$ and 0 near 0 and extend it to a smooth function on X in the obvious way.

We then define the absolute and relative boundary conditions at Y as follows:

$$\mathcal{D}^j(X; F) := \varphi\mathcal{D}(D_j) + (1-\varphi)\mathcal{D}(d_{j,\max}), \quad \mathcal{D}^j(X, Y; F) := \varphi\mathcal{D}(D_j) + (1-\varphi)\mathcal{D}(d_{j,\min}). \quad (3\text{-}25)$$

The Laplacians of the maximal and minimal ideal boundary conditions are, near Y, realizations of local elliptic boundary conditions (see, e.g., [Gilkey 1995, Section 2.7]). This, together with Proposition 3.5 and Corollary 3.7 applied to $M_1 = X$, $M_2 = Y \times (-c, 0)$, $U = Y \times (-c, -c/2)$, implies that the Hilbert complexes $(\mathcal{D}(X; F), d)$ and $(\mathcal{D}(X, Y; F), d)$ are Hilbert complexes with discrete dimension spectrum.

4. Vishik's moving boundary conditions

4A. ***Standing assumptions.*** We discuss here Vishik's [1995] moving boundary conditions for the de Rham complex in our slightly more general setting. Let X be a (not necessarily compact or complete!) Riemannian manifold; see Figure 1. Furthermore, let (F, ∇) be a flat bundle with a (not necessarily flat) Hermitian metric h^F. We assume furthermore that X contains a compact separating hypersurface $Y \subset X$ such that in a collar neighborhood $W = (-c, c) \times Y$ all structures are product. In particular we assume that ∇^F is in temporal gauge on W; that is, $\nabla^F \restriction W = \pi^* \tilde{\nabla}^F$ for a flat connection $\tilde{\nabla}^F$ on $F \restriction Y$, where π denotes the natural projection map $W \to Y$. In other words X is obtained by gluing two manifolds with boundary $X^\pm$ along their common boundary Y where all structures are product near Y; compare Figure 2.

We make the fundamental assumption that

we are given ideal boundary conditions $(\mathcal{D}^\pm, D^\pm)$ of the twisted de Rham complexes $(\Omega^\bullet(X^{\circ,\pm}; F), d)$ which have discrete dimension spectrum over $U^\pm := X^\pm \setminus W$. We put $X^{\mathrm{cut}} := X^- \amalg X^+$. (4-1)

4B. ***Some exact sequences and the main deformation result.*** As explained in Section 3D we therefore have the following Hilbert complexes with discrete dimension spectrum: $\mathcal{D}^\bullet(X^\pm;F)$ (absolute boundary condition at Y), $\mathcal{D}^\bullet(X^\pm,Y;F)$ (relative boundary condition at Y), $\mathcal{D}^\bullet(X;F)$ (continuous transmission condition at Y). By construction we have the exact sequences of Hilbert complexes

$$0 \longrightarrow \mathcal{D}^\bullet(X^-,Y;F) \xhookrightarrow{\alpha_-} \mathcal{D}^\bullet(X;F) \xrightarrow{\beta} \mathcal{D}^\bullet(X^+;F) \longrightarrow 0, \tag{4-2}$$

$$0 \longrightarrow \mathcal{D}^\bullet(X^\pm,Y;F) \xhookrightarrow{\gamma_\pm} \mathcal{D}^\bullet(X^\pm;F) \xrightarrow{i_\pm^*} \mathcal{D}^\bullet(Y;F) \longrightarrow 0, \tag{4-3}$$

$$0 \longrightarrow \mathcal{D}^\bullet(X^-,Y;F)\oplus\mathcal{D}^\bullet(X^+,Y;F) \xhookrightarrow{\alpha_++\alpha_-} \mathcal{D}^\bullet(X;F) \xrightarrow{r} \mathcal{D}^\bullet(Y;F) \longrightarrow 0. \tag{4-4}$$

Here $\alpha_\pm$ are extensions by 0, β is the pullback (i.e., restriction) to X^+, $\gamma_\pm$ is the natural inclusion of the complex $\mathcal{D}^\bullet(X^\pm,Y;F)$ with relative boundary condition at Y into the complex $\mathcal{D}^\bullet(X^\pm;F)$ with absolute boundary condition, and $i_\pm : Y \hookrightarrow X^\pm$ is the inclusion map. Finally $r\omega = \frac{\sqrt{2}}{2}(i_+^*\omega + i_-^*\omega) = \sqrt{2}i_\pm^*\omega$ for $\omega \in \mathcal{D}^\bullet(X;F)$.

It is a consequence of standard trace theorems for Sobolev spaces that $i_\pm^* : \mathcal{D}^\bullet(X^\pm;F) \to \mathcal{D}^\bullet(Y;F)$ is well-defined; see, for example, [Paquet 1982; Lions and Magenes 1972; Brüning and Lesch 2001, Section 1]. To save some space we have omitted the operator D from the notation in the complexes in (4-2), (4-3), and (4-4). Clearly, the complex differential is always the exterior derivative on the indicated domains.

Each of the complexes (4-2)–(4-4) induces a long exact sequence in cohomology. We abbreviate these long exact cohomology sequences by

$$\mathcal{H}((X^-,Y),X,X^+;F),\quad \mathcal{H}((X^\pm,Y),X^\pm,Y;F),\quad \mathcal{H}((X^-,Y)\cup(X^+,Y),X,Y;F),$$

respectively. The long exact cohomology sequences of the complexes (4-2), (4-3), (4-4) are exact sequences of finite-dimensional *Hilbert spaces* and therefore their torsion $\tau(\mathcal{H}(\cdots))$ is defined; cf. (2-25). The Euler characteristics — see (2-7) — of the complexes in (4-2)–(4-4) are denoted by $\chi(X^\pm,Y;F)$, $\chi(X^\pm;F)$, $\chi(X;F)$, $\chi(Y;F)$, etc.

Next we introduce parametrized versions of the exact sequences (4-2) and (4-4). The idea is due to Vishik [1995] who applied it to give a new proof of the Ray–Singer conjecture for compact smooth manifolds with boundary. Namely, for $\theta \in \mathbb{R}$ consider the following ideal boundary condition of the twisted de Rham complex on the disjoint union $X^{\text{cut}} = X^- \amalg X^+$:

$$\mathcal{D}_\theta^j(X;F) := \{(\omega_1,\omega_2) \in \mathcal{D}^j(X^-;F)\oplus\mathcal{D}^j(X^+;F) \mid \cos\theta \cdot i_-^*\omega_1 = \sin\theta \cdot i_+^*\omega_2\}. \tag{4-5}$$

We will see that for each real θ the complex $(\mathcal{D}_\theta^\bullet(X;F),d)$ is indeed a Hilbert complex with discrete dimension spectrum. In fact near Y it is a realization of a local elliptic boundary value problem for the de Rham complex on the manifold X^{cut}, and away from Y we may apply Corollary 3.7 and our assumption (4-1) that the Hilbert complexes $(\mathcal{D}^\pm, D^\pm)$ have discrete dimension spectrum over $X^\pm \setminus W$.

For $\theta = 0$ we have $\mathcal{D}_\theta^\bullet(X;F) = \mathcal{D}^\bullet(X^-,Y;F)\oplus\mathcal{D}^\bullet(X^+;F)$, and for $\theta = \pi/4$ we see that the total Gauss–Bonnet operators $d+d^*$ of the complexes $\mathcal{D}_{\pi/4}(X;F)$ and $\mathcal{D}(X;F)$ coincide (see [Vishik 1995, Proposition 1.1, p. 16]). Hence the family of complexes $(\mathcal{D}^\bullet_\theta(X;F), d^\theta)$ interpolates in a sense between the direct sum $\mathcal{D}^\bullet(X^-,Y;F)\oplus\mathcal{D}^\bullet(X^+;F)$ and the complex $\mathcal{D}^\bullet(X;F)$ on the manifold X.

The parametrized versions of (4-2), (4-4) are then

$$0 \longrightarrow \mathcal{D}^\bullet(X^-,Y;F) \overset{\alpha_\theta}{\hookrightarrow} \mathcal{D}^\bullet_\theta(X;F) \xrightarrow{\beta_\theta} \mathcal{D}^\bullet(X^+;F) \longrightarrow 0, \tag{4-6}$$

$$0 \longrightarrow \mathcal{D}^\bullet(X^-,Y;F)\oplus\mathcal{D}^\bullet(X^+,Y;F) \overset{\gamma_+ + \gamma_-}{\hookrightarrow} \mathcal{D}^\bullet_\theta(X;F) \xrightarrow{r_\theta} \mathcal{D}^\bullet(Y;F) \longrightarrow 0, \tag{4-7}$$

where $\alpha_\theta\omega=(\omega,0)$ is extension by 0, $\beta_\theta(\omega_1,\omega_2)=\omega_2$ is restriction to X^+, $\gamma_+\oplus\gamma_-(\omega_1,\omega_2)=(\omega_1,\omega_2)$ is inclusion and $r_\theta(\omega_1,\omega_2)=\sin\theta\cdot i_-^*\omega_1+\cos\theta\cdot i_+^*\omega_2$. We denote by $\mathcal{H}_\theta((X^-,Y),X,X^+;F)$ and $\mathcal{H}_\theta((X^-,Y)\cup(X^+,Y),X,Y;F)$ be the corresponding long exact cohomology sequences.

We denote the cohomology groups of the complex $\mathcal{D}^\bullet_\theta(X;F)$ by $H^j_\theta(X;F)$; the corresponding space of harmonic forms will be denoted by $\hat{H}^j_\theta(X;F)$. For the next result we need some more notation. Let $\mathcal{H}$ be a Hilbert space and let $T:\mathcal{H}\to\mathcal{H}$ be a bounded linear operator. For a finite-dimensional subspace $V\subset\mathcal{H}$ we write $\operatorname{Tr}(T\restriction V)$ for $\operatorname{Tr}(P_VTP_V)$ where P_V is the orthogonal projection onto V. If e_j, $j,\dots,n$, is an orthonormal basis of V then

$$\operatorname{Tr}(T\restriction V)=\sum_{j=1}^n\langle Te_j,e_j\rangle. \tag{4-8}$$

We will apply this to β_θ on the space $H^j_\theta(X;F)$. If e_j, $j,\dots,n$, is an orthonormal basis of $\hat{H}^j_\theta(X;F)$ then

$$\operatorname{Tr}(\beta_\theta\restriction H^j_\theta(X;F))=\sum_{j=1}^n\|e_j\restriction X^+\|^2_{X^+}=\sum_{j=1}^n\int_{X^+}e_j\wedge *e_j. \tag{4-9}$$

After these preparations we are able to state our main technical result. It is inspired by Lemma 2.2 and Section 2.6 in [Vishik 1995].

Theorem 4.1. *The functions*

$$\theta\mapsto\log T(\mathcal{D}^\bullet_\theta(X;F)),\quad \log\tau(\mathcal{H}_\theta((X^-,Y),X,X^+;F)),\quad \log\tau(\mathcal{H}_\theta((X^-,Y)\cup(X^+,Y),X,Y;F))$$

are differentiable for $0<\theta<\pi/2$. *Moreover, for* $0<\theta<\pi/2$,

$$\frac{d}{d\theta}\log T(\mathcal{D}^\bullet_\theta(X;F))=\frac{2}{\sin 2\theta}\Big[-\sum_{j\ge 0}(-1)^j\operatorname{Tr}(\beta_\theta\restriction H^j_\theta(X;F))+\chi(X^+;F)\Big]-\tan\theta\cdot\chi(Y;F), \tag{4-10}$$

$$\frac{d}{d\theta}\log\tau\big(\mathcal{H}_\theta((X^-,Y),X,X^+;F)\big)=\frac{2}{\sin 2\theta}\Big[-\sum_{j\ge 0}(-1)^j\operatorname{Tr}(\beta_\theta\restriction H^j_\theta(X;F))+\chi(X^+;F)\Big], \tag{4-11}$$

$$\frac{d}{d\theta}\log\tau\big(\mathcal{H}_\theta((X^-,Y)\cup(X^+,Y),X,Y;F)\big)=\frac{d}{d\theta}\log T(\mathcal{D}^\bullet_\theta(X;F)). \tag{4-12}$$

Furthermore, with $\mathcal{H}_\theta$ *standing for either* $\mathcal{H}_\theta((X^-,Y),X,X^+;F)$ *or* $\mathcal{H}_\theta((X^-,Y)\cup(X^+,Y),X,Y;F)$, *the map*

$$\theta \mapsto \log T(\mathcal{D}^\bullet_\theta(X;F)) - \log \tau(\mathcal{H}_\theta) \tag{4-13}$$

is differentiable for $0 \le \theta < \pi/2$.

The proof of Theorem 4.1 will occupy Section 5.

5. Gauge transforming the parametrized de Rham complex à la Witten

Consider the manifold X as described in Section 4. Recall that in the collar $W := (-c,c)\times Y$ of Y all structures are assumed to be product. We introduce $W^{\mathrm{cut}} := (-c,0]\times Y \amalg [0,c)\times Y$. Furthermore, let $S : W^{\mathrm{cut}} \to W^{\mathrm{cut}}$, $(t,p)\mapsto(-t,p)$ be the reflection map at Y. Finally, we introduce the map

$$T : \Omega^\bullet(W^{\mathrm{cut}};F) \to \Omega^\bullet(W^{\mathrm{cut}};F), \quad T(\omega_1,\omega_2) := (S^*\omega_2, -S^*\omega_1). \tag{5-1}$$

T is a skew-adjoint operator in $L^2(W^{\mathrm{cut}}, \Lambda^\bullet T^*W^{\mathrm{cut}} \otimes F)$ with $T^2 = -I$. Note furthermore that T commutes with the exterior derivative d. We denote by D^θ (on X^{cut} or W^{cut}) the closed extension of the exterior derivative with boundary conditions as in (4-5) along Y. More precisely, D^θ acts on the domain

$$\mathcal{D}^j_\theta(W;F) := \big\{(\omega_1,\omega_2)\in\mathcal{D}(d_{j,\max}) \;\big|\; \cos\theta\cdot i_-^*\omega_1 = \sin\theta\cdot i_+^*\omega_2\big\}. \tag{5-2}$$

The operator family has varying domain. In order to obtain variation formulas for functions of D^θ we will apply the method of gauge-transforming D^θ onto a family with constant domain; compare, for instance, [Douglas and Wojciechowski 1991] and [Lesch and Wojciechowski 1996].

We choose a cut-off function $\varphi \in C_c^\infty((-c,c)\times Y)$ with $\varphi \equiv 1$ in a neighborhood of $\{0\}\times Y$ and which satisfies $\varphi(-t,p) = \varphi(t,p)$, $(t,p)\in(-c,c)\times Y$. Then we introduce the gauge transformation

$$\Phi_\theta := e^{\theta\varphi T} = \cos(\theta\varphi)I + \sin(\theta\varphi)\,T : \Omega^\bullet(W^{\mathrm{cut}};F) \to \Omega^\bullet(W^{\mathrm{cut}};F). \tag{5-3}$$

Since $e^{\theta\varphi(t,p)T} = 1$ for $|t|$ sufficiently close to c, Φ_θ extends in an obvious way to a unitary transformation of $L^2(\Lambda^*T^*X^{\mathrm{cut}};F)$ which maps smooth forms to smooth forms.

Lemma 5.1. *For* θ, $\theta' \in \mathbb{R}$ *the operator* Φ_θ *maps* $\mathcal{D}^j_{\theta'}(X;F)$ *onto* $\mathcal{D}^j_{\theta+\theta'}(X;F)$, *and accordingly* $\mathcal{D}^j_{\theta'}(W^{\mathrm{cut}};F)$ *onto* $\mathcal{D}^j_{\theta+\theta'}(W^{\mathrm{cut}};F)$. *Furthermore,*

$$\Phi_\theta^* D^{\theta+\theta'}\Phi_\theta = D^{\theta'} + \theta\,\mathrm{ext}(d\varphi)T. \tag{5-4}$$

Proof. It obviously suffices to prove the lemma for W^{cut}. Consider $(\omega_1,\omega_2)\in\mathcal{D}^j_{\theta'}(W^{\mathrm{cut}};F)$. Then

$$i_-^*\Phi_\theta(\omega_1,\omega_2) = \cos\theta\cdot i_-^*\omega_1 + \sin\theta\cdot i_+^*\omega_2, \tag{5-5}$$

$$i_+^*\Phi_\theta(\omega_1,\omega_2) = \cos\theta\cdot i_-^*\omega_2 - \sin\theta\cdot i_+^*\omega_1. \tag{5-6}$$

A direct calculation now shows

$$\cos(\theta+\theta')\, i_-^*\Phi_\theta(\omega_1,\omega_2) = \sin(\theta+\theta')\, i_+^*\Phi_\theta(\omega_1,\omega_2), \tag{5-7}$$

proving the first claim. The formula (5-4) follows since T commutes with exterior differentiation. ☐

Note that $D^{\pi/4} + \theta\ \mathrm{ext}(d\varphi)T$ is a deformed de Rham operator acting on smooth differential forms on the smooth manifold X (respectively, W). T is not a differential operator. However, the reflection map S allows us to identify $(-c, 0) \times Y$ with $(0, c) \times Y$ and hence sections in a vector bundle E over $(-c, 0)\times Y \amalg (0, c)\times Y$ may be viewed as sections in the vector bundle $E \oplus S^*E$ over $(0, c)\times Y$. Therefore, since $\mathrm{supp}(d\varphi)$ is compact in $(-c, 0) \times Y \amalg (0, c) \times Y$, T may be viewed as a bundle endomorphism acting on the bundle $(\Lambda^* T^*(0, c) \times Y) \otimes (F \oplus F)$. In particular employing the classical interior parametric elliptic calculus, as, for example, in [Shubin 2001], we infer that the Laplacian corresponding to $D^{\pi/4} + \theta\,\mathrm{ext}(d\varphi)T$ has discrete dimension spectrum over any such compact neighborhood of $\mathrm{supp}(d\varphi)$ which does have positive distance from $\pm c \times Y$.

From now on let

$$\tilde{D}^\theta := D^{\pi/4} + \theta\,\mathrm{ext}(d\varphi)T, \tag{5-8}$$

with domain $\mathcal{D}^\bullet_{\theta=\pi/4}(X; F)$ and $\tilde{\Delta}^\theta = (\tilde{D}^\theta)^* \tilde{D}^\theta + \tilde{D}^\theta(\tilde{D}^\theta)^*$ the corresponding Laplacian. On the collar $(-c, c) \times Y$ the operator $\tilde{\Delta}^\theta$ is of the form $P + V$ as discussed in Section 3A, where P is the form Laplacian and $V = \tilde{\Delta}^\theta - \Delta$ is induced by $\theta\,\mathrm{ext}(d\varphi)T$. The subset $\mathcal{K}$ of (3-2) is the support of $d\varphi$. The operator $\tilde{\Delta}^\theta$ is now obtained as in (3-19) by gluing the domains of the form Laplacians of the given de Rham complexes on $X^\pm$. Proposition 3.5 and Corollary 3.7 now give:

Theorem 5.2. *The Hilbert complexes $\mathcal{D}^\bullet_\theta(X; F)$ defined in* (4-5) *are Hilbert complexes with discrete dimension spectrum.*

Theorem 5.3. *For $0 < \theta < \pi/2$ the Hilbert complexes $\mathcal{D}^\bullet_\theta(X; F)$ satisfy* (1)–(4) *of Proposition 2.4. More precisely,*

$$\frac{d}{d\theta} H_T(\mathcal{D}^\bullet_\theta(X; F)) = -t\frac{d}{dt}\,\frac{4}{\sin 2\theta} \sum_{j\geq 0} (-1)^j\ \mathrm{Tr}(\beta_\theta e^{-t\Delta^\theta_j}) \tag{5-9}$$

and

$$\sum_{j\geq 0} (-1)^j\ \mathrm{Tr}(\beta_\theta e^{-t\Delta^\theta_j}) = \chi(X^+; F) - \sin^2\theta \cdot \chi(Y; F) + O(t^\infty), \tag{5-10}$$

as $t \to 0+$.

5A. ***Proof of Theorem 5.3.*** Note that

$$\frac{d}{d\theta} D^{\pi/4} = \mathrm{ext}(d\varphi)T = [d, \varphi T]. \tag{5-11}$$

Let us reiterate that although $[d, \varphi T]$ is strictly speaking not a 0th-order differential operator it may be viewed as one over $(\Lambda^* T^*(0, c) \times Y) \otimes (F \oplus F)$, which implies that it lies in $\mathrm{Op}^0_c(W^{\mathrm{cut}}) \subset \mathrm{Op}^0_c(X^{\mathrm{cut}})$.

We remind the reader of the definition of the closed and coclosed Laplacians in (2-10), (2-11). We find

$$\begin{aligned}\frac{d}{d\theta}\,\mathrm{Tr}(e^{-t\Delta^\theta_{p,\mathrm{ccl}}}) = \frac{d}{d\theta}\,\mathrm{Tr}(e^{-t\tilde{\Delta}^\theta_{p,\mathrm{ccl}}}) &= -t\,\mathrm{Tr}\big(((\tilde{D}^\theta_p)^t\ \mathrm{ext}(d\varphi)T + (\mathrm{ext}(d\varphi)T)^t\,\tilde{D}^\theta_p)e^{-t\tilde{\Delta}^\theta_{p,\mathrm{ccl}}}\big)\\ &= -t\,\mathrm{Tr}\big(((D^\theta_p)^t\ \mathrm{ext}(d\varphi)T + (\mathrm{ext}(d\varphi)T)^t\,D^\theta_p)e^{-t\Delta^\theta_{p,\mathrm{ccl}}}\big),\end{aligned} \tag{5-12}$$

where in the last line we have used that Φ_θ commutes with $\mathrm{ext}(d\varphi)T$.

Next let $(e_n)_{n\in\mathbb{N}}$ be an orthonormal basis of $\operatorname{ran}(D_p^\theta)^*$ consisting of eigenvectors of $\Delta^\theta_{p,\mathrm{ccl}}$ to eigenvalues $\lambda_n > 0$. Then $(\tilde e_n = \lambda_n^{-1/2} e_n)_n$ is an orthonormal basis of $\operatorname{ran} D_p^\theta$ consisting of eigenvectors of $\Delta^\theta_{p+1,\mathrm{cl}}$ (see (2-10), (2-11) and thereafter). Equation (5-12) gives

$$\begin{aligned}\frac{d}{d\theta}\operatorname{Tr}(e^{-t\Delta^\theta_{p,\mathrm{ccl}}}) &= -t\sum_n \langle (d_p^\theta)^t \operatorname{ext}(d\varphi) T e^{-t\Delta^\theta_{p,\mathrm{ccl}}} e_n, e_n\rangle - t\sum_n \langle e^{-t\Delta^\theta_{p,\mathrm{ccl}}} e_n, (d_p^\theta)^t \operatorname{ext}(d\varphi) T e_n\rangle \\ &= -2t\Re\Bigl(\sum_n \langle (d_p^\theta)^t \operatorname{ext}(d\varphi) T e^{-t\Delta^\theta_{p,\mathrm{ccl}}} e_n, e_n\rangle\Bigr). \qquad (5\text{-}13)\end{aligned}$$

Stokes' theorem and the boundary conditions will allow us to rewrite the individual summands of the last sum. To this end let $\omega, \eta \in \mathcal{D}(\Delta_p^\theta)$. Then since $d\varphi$ is compactly supported in the interior of W^{cut} we have

$$\begin{aligned}\langle (d_p^\theta)^t \operatorname{ext}(d\varphi) T\omega, \eta\rangle &= \langle \operatorname{ext}(d\varphi)T\omega, d\eta\rangle = \langle d\varphi\wedge T\omega, d\eta\rangle = \langle d(\varphi T\omega), d\eta\rangle - \langle \varphi T d\omega, d\eta\rangle \\ &= \int_{\partial X^{\mathrm{cut}}} T\omega \wedge \tilde{*} d\eta + \langle \varphi T\omega, d^t d\eta\rangle - \langle \varphi T d\omega, d\eta\rangle. \qquad (5\text{-}14)\end{aligned}$$

Here, $\tilde{*}$ denotes the natural isometry $\bigwedge^p T^*M\otimes F \to \bigwedge^{m-p} T^*M \otimes F^\dagger$. In the last equality we have applied Stokes' theorem on the manifold with boundary X^{cut}. Note that $\varphi T\omega$ is a compactly supported (locally of Sobolev class at least 2) form on X^{cut}.

The boundary of X^{cut} consists of two copies of Y with opposite orientations. To calculate the integral in the last equation we orient Y as the boundary of X^+. Then using that ω and η satisfy the boundary conditions (4-5) at Y we find

$$\begin{aligned}\int_{\partial X^{\mathrm{cut}}} T\omega\wedge\tilde{*}d\eta &= \int_Y i_+^*(T\omega\wedge\tilde{*}d\eta) - i_-^*(T\omega\wedge\tilde{*}d\eta) = -\int_Y i_-^*\omega\wedge i_+^*\tilde{*}d\eta + i_+^*\omega\wedge i_-^*\tilde{*}d\eta \\ &= -(\tan\theta+\cot\theta)\int_Y i_+^*(\omega\wedge\tilde{*}d\eta) = -\frac{2}{\sin 2\theta}\Bigl(\int_{X^+} d\omega\wedge\tilde{*}d\eta + (-1)^{|\omega|}\omega\wedge d\tilde{*}d\eta\Bigr) \\ &= -\frac{2}{\sin 2\theta}(\langle d\omega, d\eta\rangle_{X^+} - \langle \omega, d^t d\eta\rangle_{X^+}). \qquad (5\text{-}15)\end{aligned}$$

Here $\langle\cdot,\cdot\rangle_{X^+}$ denotes the L^2-scalar product of forms over X^+.

Plugging into (5-14) gives

$$\langle (d_p^\theta)^t\operatorname{ext}(d\varphi)T\omega,\eta\rangle = \langle\varphi T\omega, d^t d\eta\rangle - \langle \varphi T d\omega, d\eta\rangle - \frac{2}{\sin 2\theta}(\langle d\omega, d\eta\rangle_{X^+} - \langle\omega, d^t d\eta\rangle_{X^+}). \qquad (5\text{-}16)$$

Similarly,

$$\langle (\operatorname{ext}(d\varphi)T)^t D_p^\theta\omega,\eta\rangle = \langle d^t d\omega, \varphi T\eta\rangle - \langle d\omega, \varphi T d\eta\rangle - \frac{2}{\sin 2\theta}(\langle d\omega, d\eta\rangle_{X^+} - \langle\omega, d^t d\eta\rangle_{X^+}). \qquad (5\text{-}17)$$

We now apply (5-16) to the summands on the right of (5-13) and find using (2-12)

$$\begin{aligned}\langle (d_p^\theta)^t\operatorname{ext}(d\varphi)Te^{-t\Delta^\theta_{p,\mathrm{ccl}}}e_n, e_n\rangle &= \langle \varphi T e^{-t\Delta^\theta_{p,\mathrm{ccl}}}\Delta^\theta_{p,\mathrm{ccl}} e_n, e_n\rangle - \langle\varphi T e^{-t\Delta^\theta_{p+1,\mathrm{cl}}}\Delta^\theta_{p+1,\mathrm{cl}}\tilde e_n, \tilde e_n\rangle \\ &\quad - \frac{2}{\sin 2\theta}\bigl(\langle\beta_\theta e^{-t\Delta^\theta_{p+1,\mathrm{cl}}}\Delta^\theta_{p+1,\mathrm{cl}}\tilde e_n,\tilde e_n\rangle - \langle\beta_\theta e^{-t\Delta^\theta_{p,\mathrm{ccl}}}\Delta^\theta_{p,\mathrm{ccl}} e_n, e_n\rangle\bigr), \qquad (5\text{-}18)\end{aligned}$$

and summing over n gives

$$\begin{aligned}\frac{d}{d\theta}\operatorname{Tr}(e^{-t\Delta^{\theta}_{p,\mathrm{ccl}}}) &= -2t\Re\big(\operatorname{Tr}(\varphi T e^{-t\Delta^{\theta}_{p,\mathrm{ccl}}}\Delta^{\theta}_{p,\mathrm{ccl}}) - \operatorname{Tr}(\varphi T e^{-t\Delta^{\theta}_{p+1,\mathrm{cl}}}\Delta^{\theta}_{p+1,\mathrm{cl}})\big)\\ &\qquad + \frac{4t}{\sin 2\theta}\Re\big(\operatorname{Tr}(\beta_\theta\Delta^{\theta}_{p+1,\mathrm{cl}}e^{-t\Delta^{\theta}_{p+1,\mathrm{cl}}}) - \operatorname{Tr}(\beta_\theta\Delta^{\theta}_{p,\mathrm{ccl}}e^{-t\Delta^{\theta}_{p,\mathrm{ccl}}})\big)\\ &= 2t\frac{d}{dt}\,\frac{2}{\sin 2\theta}\big(\operatorname{Tr}(\beta_\theta e^{-t\Delta^{\theta}_{p,\mathrm{ccl}}}) - \operatorname{Tr}(\beta_\theta e^{-t\Delta^{\theta}_{p+1,\mathrm{cl}}})\big). \end{aligned} \tag{5-19}$$

Here we have used that, since φT is skew-adjoint, $\operatorname{Tr}(\varphi T A)$ is purely imaginary for every self-adjoint trace class operator A and similarly that since β_θ is self-adjoint that $\operatorname{Tr}(\beta_\theta A)$ is real. Consequently, using (2-14),

$$\frac{d}{d\theta}H_T(\mathcal{D}^{\bullet}_{\theta}(X;F)) = \frac{d}{d\theta}\sum_{j\ge 0}(-1)^{j+1}\operatorname{Tr}(e^{-t\Delta^{\theta}_{j,\mathrm{ccl}}}) = -2t\frac{d}{dt}\,\frac{2}{\sin 2\theta}\sum_{j\ge 0}(-1)^j\operatorname{Tr}(\beta_\theta e^{-t\Delta^{\theta}_{j}}). \tag{5-20}$$

Finally, for calculating the asymptotic expansion (5-10) as $t\to 0+$ we may again invoke Corollary 3.7. The asymptotic expansion (5-10) on X^{cut} differs from the corresponding expansion for the double $-X^+ \amalg X^+$ by an error term $O(t^\infty)$; here $-X^+$ stands for X^+ with the opposite orientation. However, on the double $-X^+ \amalg X^+$ we may write down the heat kernel for Δ^{θ}_{p} explicitly in terms of the heat kernels for Δ_p with relative and absolute boundary conditions at Y [Vishik 1995, (2.118), p. 60]. Namely, let Δ^r_p, Δ^a_p be the Laplacians of the relative and absolute de Rham complexes on X^+ as in (3-25) and denote by $E^{p,r/a}_t$ their corresponding heat kernels. Let S be the reflection map which interchanges the two copies of X^+ in $-X^+ \amalg X^+$. Its restriction to W is the reflection map S defined before (5-1) and hence denoting it by the same letter is justified.

Finally, let $E^p_t(x,y)$ be the heat kernel of $\Delta^{\pi/4}_p$ on $-X^+ \amalg X^+$, that is, the Laplacian with continuous transmission boundary conditions at Y. The absolute/relative heat kernels are given in terms of E^p_t by

$$E^{p,a}_t = (E^p_t + S^*\circ E^p_t)\restriction X^+, \quad E^{p,r}_t = (E^p_t - S^*\circ E^p_t)\restriction X^+. \tag{5-21}$$

More generally, we put for $x, y \in -X^+ \amalg X^+$:

$$E^{p,\theta}_t(x,y) := \begin{cases} E^p_t(x,y) + \cos(2\theta)(S^*\circ E^p_t)(x,y) & \text{if } x,y\in X^+,\\ \sin(2\theta)E^p_t(x,y) & \text{if } x\in(-X^+),\ y\in X^+. \end{cases} \tag{5-22}$$

One immediately checks that $E^{p,\theta}_t$ is the heat kernel of Δ^{θ}_{p} on $-X^+ \amalg X^+$. Consequently

$$\operatorname{Tr}(\beta_\theta e^{-t\Delta^{\theta}_{p}}) = \operatorname{Tr}(\beta_\theta E^p_t) + \cos(2\theta)\operatorname{Tr}(S^*\circ E^p_t) = \cos^2(\theta)\operatorname{Tr}(E^{p,a}_t) + \sin^2(\theta)\operatorname{Tr}(E^{p,r}_t). \tag{5-23}$$

In view of our standing assumptions (Section 4A) the complexes $\mathcal{D}^{\bullet}(X^+,Y;F)$ and $\mathcal{D}^{\bullet}(X^+;F)$ are Fredholm complexes, so the McKean–Singer formula (2-8) holds and hence taking alternating sums yields

$$\sum_{j\ge 0}(-1)^j\operatorname{Tr}(\beta_\theta e^{-t\Delta^{\theta}_{j}}) = \cos^2\theta\cdot\chi(X^+;F) + \sin^2\theta\cdot\chi(X^+,Y;F) = \chi(X^+;F) - \sin^2\theta\cdot\chi(Y;F), \tag{5-24}$$

and the proof of (5-10) is complete. In the last equality we used that $\chi(X^+;F) = \chi(X^+,Y;F) + \chi(Y;F)$; this formula follows from the exact sequence (4-4). $\square$

5B. ***Proof of Theorem 4.1.***

Proof of (4-10). Combining Proposition 2.4 and Theorem 5.3 we find

$$\begin{aligned}\frac{d}{d\theta}&\log T(\mathcal{D}^\bullet_\theta(X;F))\\ &=-\frac{1}{2}\,\frac{-4}{\sin 2\theta}\,(\chi(X^+;F)-\sin^2\theta\,\chi(Y;F))+\frac{1}{2}\,\frac{-4}{\sin 2\theta}\operatorname{Tr}\Bigl(\sum_{j\geq 0}(-1)^j\beta_\theta\restriction H^j_\theta(X;F)\Bigr)\\ &=\frac{2}{\sin 2\theta}\Bigl(-\sum_{j\geq 0}(-1)^j\operatorname{Tr}(\beta_\theta\restriction H^j_\theta(X;F))+\chi(X^+;F)\Bigr)-\tan\theta\cdot\chi(Y;F),\end{aligned}\tag{5-25}$$

which is the right-hand side of (4-10). ☐

Proof of (4-11) *and* (4-12). Let $0<\theta,\theta'<\pi/2$ and consider the following commutative diagram (cf. (4-6)):

$$\begin{tikzcd} 0 \arrow[r] & \mathcal{D}^\bullet(X^-,Y;F) \arrow[r, "\alpha_\theta"] \arrow[d, "\mathrm{id}"] & \mathcal{D}^\bullet_\theta(X;F) \arrow[r, "\beta_\theta"] \arrow[d, "\phi_{\theta,\theta'}"] & \mathcal{D}^\bullet(X^+;F) \arrow[r] \arrow[d, "\phi^+_{\theta,\theta'}"] & 0\\ 0 \arrow[r] & \mathcal{D}^\bullet(X^-,Y;F) \arrow[r, "\alpha_{\theta'}"] & \mathcal{D}^\bullet_{\theta'}(X;F) \arrow[r, "\beta_{\theta'}"] & \mathcal{D}^\bullet(X^+;F) \arrow[r] & 0, \end{tikzcd}\tag{5-26}$$

where $\phi_{\theta,\theta'}$ and $\phi^+_{\theta,\theta'}$ are Hilbert complex isomorphisms, defined by

$$\phi_{\theta,\theta'}(\omega_1,\omega_2)=\Bigl(\omega_1,\frac{\tan\theta}{\tan\theta'}\omega_2\Bigr),\quad \phi^+_{\theta,\theta'}(\omega_2)=\frac{\tan\theta}{\tan\theta'}\omega_2.$$

Hence we obtain a cochain isomorphism between the long exact cohomology sequences of the upper and lower horizontal exact sequences (F omitted to save horizontal space):

$$\begin{tikzcd} \cdots H^k(X^-,Y) \arrow[r, "\alpha_{\theta,*}"] \arrow[d, "\mathrm{id}"] & H^k_\theta(X) \arrow[r, "\beta_{\theta,*}"] \arrow[d, "\phi_{\theta,\theta',*}"] & H^k(X^-) \arrow[r, "\delta_\theta"] \arrow[d, "\phi^+_{\theta,\theta',*}"] & H^{k+1}(X^-,Y)\cdots \arrow[d, "\mathrm{id}"]\\ \cdots H^k(X^-,Y) \arrow[r, "\alpha_{\theta',*}"] & H^k_{\theta'}(X) \arrow[r, "\beta_{\theta',*}"] & H^k(X^-) \arrow[r, "\delta_{\theta'}"] & H^{k+1}(X^-,Y)\cdots \end{tikzcd}\tag{5-27}$$

Let $e_1,\dots,e_r$ be an orthonormal basis of $H^k_\theta(X;F)$. Then

$$\operatorname{Det}(\phi^k_{\theta,\theta',*})^2=\operatorname{Det}(\langle\phi_{\theta,\theta',*}e_i,\phi_{\theta,\theta',*}e_j\rangle^r_{i,j=1});\tag{5-28}$$

hence

$$\begin{aligned}\frac{d}{d\theta'}\Big|_{\theta'=\theta}\log\operatorname{Det}(\phi^k_{\theta,\theta',*})^2&=\operatorname{Tr}\biggl(\Bigl(\Bigl\langle\frac{d}{d\theta'}\Big|_{\theta'=\theta}\phi_{\theta,\theta',*}e_i,e_j\Bigr\rangle+\Bigl\langle e_i,\frac{d}{d\theta'}\Big|_{\theta'=\theta}\phi_{\theta,\theta',*}e_j\Bigr\rangle\Bigr)^r_{i,j=1}\biggr)\\ &=-2\frac{2}{\sin 2\theta}\sum_{j=0}^r\langle\beta_\theta e_j,e_j\rangle=-2\frac{2}{\sin 2\theta}\operatorname{Tr}(\beta_\theta\restriction H^k_\theta(X;F));\end{aligned}\tag{5-29}$$

see (4-8). Furthermore, since $\phi^+_{\theta,\theta'}$ is multiplication by $\tan\theta/\tan\theta'$ we have

$$\operatorname{Det}(\phi^+_{\theta,\theta',*}) = \left(\frac{\tan\theta}{\tan\theta'}\right)^{\chi(X^+;F)}, \tag{5-30}$$

and hence

$$\frac{d}{d\theta'}\bigg|_{\theta'=\theta} \log\operatorname{Det}(\phi^+_{\theta,\theta',*})^2 = -2\frac{2}{\sin 2\theta}\,\chi(X^+;F). \tag{5-31}$$

By Lemma 2.5 we have

$$\begin{aligned}\log\tau(\mathcal{H}_{\theta'}((X^-,Y),X,X^+;F)) - \log\tau(\mathcal{H}_{\theta}((X^-,Y),X,X^+;F))&\\ = \tfrac{1}{2}\log\operatorname{Det}(\phi_{\theta,\theta',*})^2 - \tfrac{1}{2}\log\operatorname{Det}(\phi^+_{\theta,\theta',*})^2;&\end{aligned} \tag{5-32}$$

combined with (5-29) and (5-31) we therefore find (4-11).

That the left-hand side of (4-12) equals the right-hand side of (4-10) is proved analogously. One just has to replace the commutative diagram (5-26) by

$$\begin{tikzcd} 0 \arrow[r] & \mathcal{D}^\bullet(X^-,Y;F)\oplus\mathcal{D}^\bullet(X^+,Y;F) \arrow[r, "\gamma_++\gamma_-"] \arrow[d, "\mathrm{id}\oplus\phi^+_{\theta,\theta'}"] & \mathcal{D}^\bullet_\theta(X;F) \arrow[r, "r_\theta"] \arrow[d, "\phi_{\theta,\theta'}"] & \mathcal{D}^\bullet(Y;F) \arrow[r] \arrow[d, "\psi_{\theta,\theta'}"] & 0 \\ 0 \arrow[r] & \mathcal{D}^\bullet(X^-,Y;F)\oplus\mathcal{D}^\bullet(X^+,Y;F) \arrow[r, "\gamma_++\gamma_-"] & \mathcal{D}^\bullet_{\theta'}(X;F) \arrow[r, "r_{\theta'}"] & \mathcal{D}^\bullet(Y;F) \arrow[r] & 0, \end{tikzcd} \tag{5-33}$$

where $\psi_{\theta,\theta'}(\omega) = \dfrac{\sin\theta}{\sin\theta'}\omega$. See also (A-20) and thereafter. □

Proof of the differentiability of (4-13) *at* 0. The problem is that the dimensions of the cohomology groups $H^j_\theta(X;F)$ may jump at 0; note that the isomorphism $\phi_{\theta,\theta'}$ defined after (5-27) between $\mathcal{D}^\bullet_\theta(X;F)$ and $\mathcal{D}^\bullet_{\theta'}(X;F)$ is defined only for $0<\theta,\theta'<\pi/2$. By our standing assumptions of Section 4A (see also Section 3D), $\mathcal{D}^\bullet(X^-,Y;F)$ and $\mathcal{D}^\bullet(X^+;F)$ are Hilbert complexes with discrete dimension spectrum. Hence we may choose $a>0$ such that a is smaller than the smallest nonzero eigenvalues of the Laplacians of $\mathcal{D}^\bullet(X^-,Y;F)$ and $\mathcal{D}^\bullet(X^+,Y;F)$. Furthermore, we denote by Π^p_θ the orthogonal projection onto

$$H^p_{\theta,a}(X;F) := \bigoplus_{0\le\lambda<a}\ker(\Delta^\theta_p - \lambda). \tag{5-34}$$

Since for $\theta=0$ the complex $\mathcal{D}^\bullet_\theta(X;F)$ is canonically isomorphic to the direct sum $\mathcal{D}^\bullet(X^-,Y;F)\oplus\mathcal{D}^\bullet(X^+;F)$ and since the gauge-transformed Laplacian $\tilde\Delta^\theta$ of $\mathcal{D}^\bullet_\theta(X;F)$ in view of (5-8) certainly depends smoothly on θ there exists a $\theta_0>0$ such that the projection Π^p_θ depends smoothly on θ for $0\le\theta<\theta_0$. In particular,

$$\operatorname{rank}\Pi^p_\theta = \dim H^p_{\theta=0}(X;F)$$

is constant for $0\le\theta<\theta_0$.

$(H^\bullet_{\theta,a}(X;F),d)$ is a finite-dimensional Hilbert complex and the orthogonal projections Π^p_θ give rise to a natural orthogonal decomposition of Hilbert complexes

$$\mathcal{D}^\bullet_\theta(X;F) =: (H^\bullet_{\theta,a}(X;F),d)\oplus\mathcal{D}^\bullet_{\theta,a}(X;F). \tag{5-35}$$

By construction of Π_a^p we have

$$\log T(\mathcal{D}_\theta^\bullet(X;F)) = \log\tau(H_{\theta,a}^\bullet(X;F),d) + \log T(\mathcal{D}_{\theta,a}^\bullet(X;F)), \tag{5-36}$$

and $\theta\mapsto \log T(\mathcal{D}_{\theta,a}^\bullet(X;F))$ is differentiable for $0\le\theta<\theta_0$.

Since surjectivity is an open condition we conclude that the sequence

$$0 \longrightarrow H^*(X^-,Y;F) \xhookrightarrow{\alpha_\theta} H_{\theta,a}^*(X;F) \xrightarrow{\beta_\theta} H^*(X^+;F) \longrightarrow 0 \tag{5-37}$$

is exact for $0\le\theta<\theta_1\le\theta_0$. Here, α_θ is defined in the obvious way while

$$\beta_\theta := \text{ orthogonal projection onto } H^*(X^+;F) \text{ of } \omega\restriction X^+. \tag{5-38}$$

Note that the differentials of the left and right complexes vanish and hence so do their torsions. The space of harmonics of the middle complex equals the space of harmonics of the complex $\mathcal{D}_\theta^\bullet(X;F)$ and hence the cohomology of the middle complex is (isometrically) isomorphic to the cohomology of $\mathcal{D}_\theta^\bullet(X;F)$. One immediately checks that the long exact cohomology sequence of (5-37) is exactly the exact cohomology sequence $\mathcal{H}((X^-,Y),X,X^+;F)$. Hence Proposition 2.6 yields

$$\begin{aligned}\log\tau(H_{\theta,a}^*(X;F)) &= \log\tau(\mathcal{H}((X^-,Y),X,X^+;F))\\ &\qquad -\sum_{p\ge0}\log\tau\big(0\to H^p(X^-,Y;F)\overset{\alpha_\theta}{\to} H_{\theta,a}^p(X;F)\overset{\beta_\theta}{\to} H^p(X^+;F)\to0\big).\end{aligned} \tag{5-39}$$

This shows the differentiability of the difference $\log\tau(H_{\theta,a}^*(X;F)) - \log\tau(\mathcal{H}((X^-,Y),X,X^+;F))$ at $\theta=0$. In view of (5-36) the claim is proved. □

6. The gluing formula

We can now state and prove the main result of this paper. The standing assumptions (Section 4A) are still in effect. Furthermore, we will use freely the notation introduced in Section 4B.

Theorem 6.1. *For the analytic torsions of the Hilbert complexes $\mathcal{D}^\bullet(X^\pm,Y;F)$, $\mathcal{D}^\bullet(X^\pm;F)$, $\mathcal{D}^\bullet(X;F)$ we have the following formulas*:

$$\begin{aligned}\log T(\mathcal{D}^\bullet(X;F)) &= \log T(\mathcal{D}^\bullet(X^-,Y;F)) + \log T(\mathcal{D}^\bullet(X^+;F))\\ &\qquad + \log\tau\big(\mathcal{H}((X^-,Y),X,X^+;F)\big) - \tfrac12\log2\cdot\chi(Y;F),\end{aligned} \tag{6-1}$$

$$\log T(\mathcal{D}^\bullet(X^-;F)) = \log T(\mathcal{D}^\bullet(X^-,Y;F)) + \log T(\mathcal{D}^\bullet(Y;F)) + \log\tau\big(\mathcal{H}((X^-,Y),X^-,Y;F)\big), \tag{6-2}$$

$$\begin{aligned}\log T(\mathcal{D}^\bullet(X;F)) &= \log T(\mathcal{D}^\bullet(X^-,Y;F)) + \log T(\mathcal{D}^\bullet(X^+,Y;F))\\ &\qquad + \log\tau\big(\mathcal{H}((X^-,Y)\cup(X^+,Y),X,Y;F)\big) + \log T(\mathcal{D}^\bullet(Y;F)).\end{aligned} \tag{6-3}$$

6A. ***Proof of Theorem 6.1.*** In the course of the proof we will make heavy use of Theorem 4.1.

Proof of (6-1). As noted after (4-5) we have for $\theta = 0$ that $\mathcal{D}^\bullet_{\theta=0}(X; F) = \mathcal{D}^\bullet(X^-, Y; F) \oplus \mathcal{D}^\bullet(X^+; F)$ and that for $\theta = \pi/4$ the complexes $\mathcal{D}^\bullet_{\theta=\pi/4}(X; F)$ and $\mathcal{D}^\bullet(X; F)$ are isometric. Hence we have

$$\begin{aligned}
&\log T(\mathcal{D}^\bullet(X; F)) - \log T(\mathcal{D}^\bullet(X^-, Y; F)) - \log T(\mathcal{D}^\bullet(X^+; F))\\
&\quad = \log T(\mathcal{D}^\bullet_{\pi/4}(X; F)) - \log T(\mathcal{D}^\bullet_{\theta=0}(X; F))\\
&\quad = \log T(\mathcal{D}^\bullet_{\pi/4}(X; F)) - \log \tau(\mathcal{H}_{\pi/4}((X^-, Y), X, X^+; F)) - \log T(\mathcal{D}^\bullet_{\theta=0}(X; F))\\
&\qquad\qquad + \log \tau\big(\mathcal{H}_{\theta=0}((X^-, Y), X, X^+; F)\big) + \log \tau\big(\mathcal{H}_{\pi/4}((X^-, Y), X, X^+; F)\big). \qquad (6\text{-}4)
\end{aligned}$$

Recall that for $\theta = 0$ the complex $\mathcal{D}^\bullet_{\theta=0}(X; F)$ is just the direct sum complex $\mathcal{D}^\bullet(X^-, Y; F) \oplus \mathcal{D}^\bullet(X^+; F)$ and hence $\log \tau(\mathcal{H}_{\theta=0}((X^-, Y), X, X^+; F)) = 0$ (see also the sentence after (2-29)). Furthermore, $\log \tau(\mathcal{H}_{\pi/4}((X^-, Y), X, X^+; F)) = \log \tau(\mathcal{H}((X^-, Y), X, X^+; F))$; hence by Theorem 4.1

$$\begin{aligned}
\cdots &= \int_0^{\pi/4} - \tan\theta d\theta \ \chi(Y; F) + \log \tau\big(\mathcal{H}((X^-, Y), X, X^+; F)\big)\\
&= -\tfrac{1}{2} \log 2 \ \chi(Y; F) + \log \tau\big(\mathcal{H}((X^-, Y), X, X^+; F)\big), \qquad (6\text{-}5)
\end{aligned}$$

and we arrive at (6-1). □

Proof of (6-2). Consider $\epsilon > 0$ and apply the proved equation (6-1) to the manifold $X^-_\epsilon := X^- \cup_Y [0, \epsilon] \times Y$. Then

$$\begin{aligned}
\log T(\mathcal{D}^\bullet(X^-_\epsilon; F)) &= \log T(\mathcal{D}^\bullet(X^-, Y; F)) + \log T(\mathcal{D}^\bullet([0, \epsilon] \times Y; F))\\
&\qquad - \tfrac{1}{2} \log 2 \ \chi(Y; F) + \log \tau\big(\mathcal{H}((X^-_\epsilon, Y), X^-_\epsilon, [0, \epsilon] \times Y; F)\big). \qquad (6\text{-}6)
\end{aligned}$$

For the cylinder $[0, \epsilon] \times Y$ it is well-known (it also follows easily from Proposition 2.3) that

$$\chi([0, \epsilon] \times Y; F) = \chi(Y; F) = \chi(Y) \operatorname{rank} F, \qquad (6\text{-}7)$$

$$\begin{aligned}
\log T(\mathcal{D}^\bullet([0, \epsilon] \times Y; F)) &= \log T(\mathcal{D}^\bullet(Y; F)) \ \chi([0, \epsilon]) + \chi(Y; F) \log T(\mathcal{D}^\bullet([0, \epsilon])\\
&= \log T(\mathcal{D}^\bullet(Y; F)) + \tfrac{1}{2} \log(2\epsilon) \ \chi(Y; F). \qquad (6\text{-}8)
\end{aligned}$$

Hence

$$\begin{aligned}
\log T(\mathcal{D}^\bullet(X^-_\epsilon; F)) &= \log T(\mathcal{D}^\bullet(X^-, Y; F)) + \log T(\mathcal{D}^\bullet(Y; F))\\
&\qquad + \tfrac{1}{2} \log \epsilon \ \chi(Y; F) + \log \tau(\mathcal{H}((X^-_\epsilon, Y), X^-_\epsilon, [0, \epsilon] \times Y; F)). \qquad (6\text{-}9)
\end{aligned}$$

In the sequel we will, to save some space, omit the bundle F from the notation in commutative diagrams. Our first commutative diagram is

$$\begin{tikzcd}
\cdots \arrow[r] & H^k(X^-, Y) \arrow[r, "\alpha_{-,*}"] \arrow[d, "\psi^*_\epsilon"] & H^k(X^-_\epsilon) \arrow[r, "\beta_*"] \arrow[d, "\psi^*_\epsilon"] & H^k([0, \epsilon] \times Y) \arrow[r] \arrow[d, "\chi^*_\epsilon"] & \cdots \\
\cdots \arrow[r] & H^k(X^-, Y) \arrow[r] & H^k(X^-) \arrow[r] & H^k(Y) \arrow[r] & \cdots
\end{tikzcd} \qquad (6\text{-}10)$$

The first row is the long exact cohomology sequence of (4-2) for $X^-_\epsilon = X^- \cup_Y [0, \epsilon] \times Y$ instead of X;

the second row is the long exact cohomology sequence of (4-2) for $X = X^- \cup_Y X^+$. The map ψ_ϵ is a diffeomorphism $X^- \to X^-_\epsilon$ obtained as follows: choose a diffeomorphism $f : [-c, 0] \to [-c, \epsilon]$ such that $f(x) = x$ for x near $-c$ and $f(x) = x + \epsilon$ for x near 0. Then ψ_ϵ is obtained by patching the identity on $X^- \setminus [-c, 0] \times Y$ and $f \times \mathrm{id}_Y$. Furthermore $\chi_\epsilon : Y \to [0, \epsilon] \times Y$, $p \mapsto (\epsilon, p)$.

For a harmonic form $\omega \in \mathcal{D}(d_{k,\max}) \cap \mathcal{D}((d_{k-1,\max})^*) \subset \Omega^k([0, \epsilon] \times Y; F)$ one has $\omega = \pi^* \chi^*_\epsilon(\omega)$ ($\pi : [0, \epsilon] \times Y \to Y$ the projection) and thus

$$\int_{[0,\epsilon]\times Y} \omega \wedge \tilde{*}\omega = \epsilon \int_Y \chi^*_\epsilon \omega \wedge \tilde{*} \chi^*_\epsilon \omega. \tag{6-11}$$

Therefore the determinant (in the sense of (2-23)) of χ^*_ϵ on the cohomology is given by $\epsilon^{-\frac{1}{2}\chi(Y;F)}$. Consequently by Lemma 2.5

$$\begin{aligned} \log \tau\big(\mathcal{H}((X^-_\epsilon, Y), X^-_\epsilon, [0, \epsilon] \times Y; F)\big) = {} & \log \tau\big(\mathcal{H}((X^-, Y), X^-, Y; F)\big) - \tfrac{1}{2} \log \epsilon\, \chi(Y; F) \\ & + \log \mathrm{Det}\big(\psi^*_\epsilon : H^*(X^-, Y; F) \to H^*(X^-, Y; F)\big) \\ & - \log \mathrm{Det}\big(\psi^*_\epsilon : H^*(X^-_\epsilon; F) \to H^*(X^-; F)\big). \end{aligned} \tag{6-12}$$

Summing up (6-9), (6-12)

$$\begin{aligned} \log T(\mathcal{D}^\bullet(X^-_\epsilon; F)) = \log T(\mathcal{D}^\bullet(X^-, Y; F)) + \log T(\mathcal{D}^\bullet(Y; F)) & + \log \tau\big(\mathcal{H}((X^-, Y), X^-, Y; F)\big) \\ & + \log \mathrm{Det}\big(\psi^*_\epsilon : H^*(X^-, Y; F) \to H^*(X^-, Y; F)\big) \\ & - \log \mathrm{Det}\big(\psi^*_\epsilon : H^*(X^-_\epsilon; F) \to H^*(X^-; F)\big). \end{aligned} \tag{6-13}$$

As $\epsilon \to 0$ the determinants of

$$\psi^*_\epsilon : H^*(X^-, Y; F) \to H^*(X^-, Y; F)), \quad \text{resp.,} \quad \psi^*_\epsilon : H^*(X^-_\epsilon; F) \to H^*(X^-; F))$$

tend to 1 and we obtain (6-2). ☐

Proof of (6-3). We note that $\tau\big(\mathcal{H}_\theta((X^-, Y) \cup (X^+, Y), X, Y; F)\big)\big|_{\theta=0} = \tau\big(\mathcal{H}((X^+, Y), X^+, Y; F)\big)$; hence by (4-12) and (4-13)

$$\begin{aligned} \log T(\mathcal{D}^\bullet_{\theta=\pi/4}&(X; F)) - \log \tau\big(\mathcal{H}_\theta((X^-, Y) \cup (X^+, Y), X, Y; F)\big)\big|_{\theta=\pi/4} \\ &= \log T(\mathcal{D}^\bullet_{\theta=0}(X; F)) - \log \tau\big(\mathcal{H}_\theta((X^-, Y) \cup (X^+, Y), X, Y; F)\big)\big|_{\theta=0} \\ &= \log T(\mathcal{D}^\bullet(X^-, Y; F)) + \log T(\mathcal{D}^\bullet(X^+; F)) - \log \tau\big(\mathcal{H}((X^+, Y), X^+, Y; F)\big) \\ &= \log T(\mathcal{D}^\bullet(X^-, Y; F)) + \log T(\mathcal{D}^\bullet(X^+, Y; F)) + \log T(\mathcal{D}^\bullet(Y; F)), \end{aligned} \tag{6-14}$$

where in the last equality we have used the proved identity (6-2). ☐

Appendix: The homological algebra gluing formula

We present here the analogues of Theorems 6.1 and 4.1 for finite-dimensional Hilbert complexes. This applies, for example, to the cochain complexes of a triangulation twisted by a unitary representation of the fundamental group; see, e.g., [Müller 1993, Section 1].

Let (C_j^*, d^j), $j = 1, 2$, be finite-dimensional Hilbert complexes. Let (B^*, d) be another such Hilbert complex and assume that we are given *surjective* homomorphisms of cochain complexes

$$r_j : (C_j, d^j) \to (B, d), \quad j = 1, 2. \tag{A-1}$$

We denote by $C_{j,r} \subset C_j$ the kernel of r_j, by $\alpha : C_1 \to C_1 \oplus C_2$ the inclusion and by $\beta : C_1 \oplus C_2 \to C_2$ the projection onto the second factor.

For $\theta \in \mathbb{R}$ we define the following homological algebra analogue of the complex $\mathcal{D}_\theta^\bullet(X; F)$ of (4-5) by putting

$$(C_1 \oplus_\theta C_2)^j := \bigl\{(\xi_1, \xi_2) \in C_1^j \oplus C_2^j \bigm| \cos\theta \cdot r_1\xi_1 = \sin\theta \cdot r_2\xi_2\bigr\}. \tag{A-2}$$

$(C_1 \oplus_\theta C_2, d = d^1 \oplus d^2)$ is a subcomplex of $(C_1 \oplus C_2, d^1 \oplus d^2)$. For $\theta = 0$ we have $C_1 \oplus_\theta C_2 = C_{1,r} \oplus C_2$ and for $\theta = \pi/4$ we have a homological algebra analogue of the complex $\mathcal{D}_\theta^\bullet(X; F)$.

Furthermore, we have the following analogues of the exact sequences (4-3), (4-6), (4-7) (note that the exact sequences (4-2), (4-4) are special cases of the exact sequences (4-6), (4-7)):

$$0 \to C_{j,r} \xrightarrow{\gamma_j} C_j \xrightarrow{r_j} B \to 0, \tag{A-3}$$

$$0 \to C_{1,r} \xrightarrow{\alpha_\theta} C_1 \oplus_\theta C_2 \xrightarrow{\beta_\theta} C_2 \to 0, \tag{A-4}$$

$$0 \to C_{1,r} \oplus C_{2,r} \xrightarrow{\gamma_1+\gamma_2} C_1 \oplus_\theta C_2 \xrightarrow{r_\theta} B \to 0. \tag{A-5}$$

Here, γ_j is the natural inclusion, $\beta_\theta = \beta \restriction C_1 \oplus_\theta C_2$, $\alpha_\theta(\xi) = (\xi, 0)$, and $r_\theta(\xi_1, \xi_2) = \sin\theta \cdot r_1\,\xi_1 + \cos\theta \cdot r_2\xi_2$. Denote by $\mathcal{H}(C_{j,r}, C_j, B)$, $\mathcal{H}(C_{1,r}, C_1 \oplus_\theta C_2, C_2)$, $\mathcal{H}(C_{1,r} \oplus C_{2,r}, C_1 \oplus_\theta C_2, B)$ the long exact cohomology sequences of (A-3), (A-4), (A-5), respectively.

Since all complexes are finite-dimensional we have Lemma 2.5 and Proposition 2.6 at our disposal. The latter applied to (A-3) immediately gives the analogue of (6-2):

$$\log\tau(C_1) = \log\tau(C_{1,r}) + \log\tau(B) + \log\tau(\mathcal{H}(C_{1,r}, C_1, B)). \tag{A-6}$$

The other claims of Theorems 6.1 and 4.1 have exact counterparts in this context as summarized in the following:

Theorem A.2. (1) *The functions*

$$\theta \mapsto \log\tau(C_1 \oplus_\theta C_2), \quad \log\tau(\mathcal{H}(C_{1,r}, C_1 \oplus_\theta C_2, C_2)), \quad \log\tau(\mathcal{H}(C_{1,r} \oplus C_{2,r}, C_1 \oplus_\theta C_2, B))$$

are differentiable for $0 < \theta < \pi/2$. *Moreover, for* $0 < \theta < \pi/2$,

$$\frac{d}{d\theta}\log\tau(C_1 \oplus_\theta C_2) = \frac{2}{\sin 2\theta}\Bigl(-\sum_{j\ge 0}(-1)^j \operatorname{Tr}(\beta_\theta \restriction H^j(C_1 \oplus_\theta C_2)) + \sum_{j\ge 0}(-1)^j \operatorname{Tr}(\beta_\theta \restriction (C_1 \oplus_\theta C_2)^j)\Bigr), \tag{A-7}$$

$$\frac{d}{d\theta}\log\tau(\mathcal{H}(C_{1,r}, C_1 \oplus_\theta C_2, C_2)) = \frac{2}{\sin 2\theta}\Bigl(-\sum_{j\ge 0}(-1)^j \operatorname{Tr}(\beta_\theta \restriction H^j(C_1 \oplus_\theta C_2)) + \chi(C_2)\Bigr), \tag{A-8}$$

and

$$\frac{d}{d\theta}\log\tau(\mathscr{H}(C_{1,r}\oplus C_{2,r},C_1\oplus_\theta C_2,B))$$
$$=\frac{2}{\sin 2\theta}\Big(-\sum_{j\geq 0}(-1)^j\operatorname{Tr}\big(\beta_\theta\restriction H^j(C_1\oplus_\theta C_2)\big)+\chi(C_2)\Big)-\tan\theta\,\chi(B). \quad \text{(A-9)}$$

Furthermore,

$$\theta\mapsto\log T(C_1\oplus_\theta C_2)-\log\tau(\mathscr{H}_\theta) \quad \text{(A-10)}$$

is differentiable for $0\leq\theta<\pi/2$. *Here,* $\mathscr{H}_\theta$ *stands for either*

$$\mathscr{H}(C_{1,r},C_1\oplus_\theta C_2,C_2) \quad \textit{or} \quad \mathscr{H}(C_{1,r}\oplus C_{2,r},C_1\oplus_\theta C_2,B).$$

(2) *Under the additional assumption that the* r_j *are* partial isometries *we have*

$$\frac{d}{d\theta}\log\tau(C_1\oplus_\theta C_2)=\frac{d}{d\theta}\log\tau(\mathscr{H}(C_{1,r}\oplus C_{2,r},C_1\oplus_\theta C_2,B)) \quad \text{(A-11)}$$

and

$$\log\tau(C_1\oplus_\theta C_2)=\log\tau(C_{1,r})+\log\tau(C_{2,r})+\log\tau(\mathscr{H}(C_{1,r}\oplus C_{2,r},C_1\oplus_\theta C_2,B)) \quad \text{(A-12)}$$
$$=\log\tau(C_{1,r})+\log\tau(C_2)+\log\tau(\mathscr{H}(C_{1,r},C_1\oplus_\theta C_2,C_2))+\log\cos\theta\,\chi(B). \quad \text{(A-13)}$$

When comparing the last formula with Theorem 6.1 one should note that for $\theta=\pi/4$ we have

$$\log\cos\theta=\log\frac{1}{\sqrt{2}}=-\tfrac{1}{2}\log 2.$$

Proof. For $0<\theta,\theta'<\pi/2$ we have the cochain isomorphism (cf. (5-26))

$$\phi_{\theta,\theta'}:C_1\oplus_\theta C_2\to C_1\oplus_{\theta'}C_2, \quad (\xi_1,\xi_2)\mapsto\Big(\xi_1,\frac{\tan\theta}{\tan\theta'}\xi_2\Big); \quad \text{(A-14)}$$

hence by Lemma 2.5

$$\log\tau(C_1\oplus_\theta C_2)=\log\tau(C_1\oplus_{\theta'}C_2)-\sum_{j\geq 0}(-1)^j\log\operatorname{Det}(\phi_{\theta,\theta'}\restriction H^j(C_1\oplus_\theta C_2))$$
$$+\sum_{j\geq 0}(-1)^j\log\operatorname{Det}(\phi_{\theta,\theta'}\restriction(C_1\oplus_\theta C_2)^j). \quad \text{(A-15)}$$

Taking $d/d\theta'$ at $\theta'=\theta$ yields (A-7).

Next we look at the analogues of (5-26) and (5-27):

$$\begin{tikzcd} 0 \arrow[r] & C_{1,r} \arrow[r, "\alpha_\theta"] \arrow[d, "\mathrm{id}"] & C_1\oplus_\theta C_2 \arrow[r, "\beta_\theta"] \arrow[d, "\phi_{\theta,\theta'}"] & C_2 \arrow[r] \arrow[d, "\tilde\phi_{\theta,\theta'}"] & 0 \\ 0 \arrow[r] & C_{1,r} \arrow[r, "\alpha_{\theta'}"] & C_1\oplus_{\theta'}C_2 \arrow[r, "\beta_{\theta'}"] & C_2 \arrow[r] & 0, \end{tikzcd} \quad \text{(A-16)}$$

where $\tilde\phi_{\theta,\theta'}(\xi)=\dfrac{\tan\theta}{\tan\theta'}\xi$, and at the corresponding isomorphism between the long exact cohomology sequences

$$\begin{array}{ccccccc}
\cdots H^k(C_{1,r}) & \xrightarrow{\alpha_{\theta,*}} & H^k(C_1 \oplus_\theta C_2) & \xrightarrow{\beta_{\theta,*}} & H^k(C_2) & \xrightarrow{\delta_\theta} & H^{k+1}(C_{1,r})\cdots \\
\downarrow \mathrm{id} & & \downarrow \phi_{\theta,\theta',*} & & \downarrow \tilde{\phi}_{\theta,\theta',*} & & \downarrow \mathrm{id} \\
\cdots H^k(C_{1,r}) & \xrightarrow{\alpha_{\theta',*}} & H^k(C_1 \oplus_{\theta'} C_2) & \xrightarrow{\beta_{\theta',*}} & H^k(C_2) & \xrightarrow{\delta_{\theta'}} & H^{k+1}(C_{1,r})\cdots
\end{array} \tag{A-17}$$

Following the argument after (5-27) we find that

$$\frac{d}{d\theta'}\bigg|_{\theta'=\theta} \log \operatorname{Det}(\phi^j_{\theta,\theta',*})^2 = -2\frac{2}{\sin 2\theta}\operatorname{Tr}(\beta_\theta \restriction H^j(C_1 \oplus_\theta C_2)), \tag{A-18}$$

$$\frac{d}{d\theta'}\bigg|_{\theta'=\theta} \log \operatorname{Det}(\tilde{\phi}^j_{\theta,\theta',*}) = -2\frac{2}{\sin 2\theta}\dim C_2^j, \tag{A-19}$$

and hence with Lemma 2.5 applied to (A-17) we arrive at (A-8).

The analogue of (5-33) is

$$\begin{array}{ccccccccc}
0 & \longrightarrow & C_{1,r} \oplus C_{2,r} & \xrightarrow{\gamma_1 \oplus \gamma_2} & C_1 \oplus_\theta C_2 & \xrightarrow{r_\theta} & B & \longrightarrow & 0 \\
 & & \downarrow \mathrm{id} \oplus \tilde{\phi}_{\theta,\theta'} & & \downarrow \phi_{\theta,\theta'} & & \downarrow \psi_{\theta,\theta'} = \frac{\tan\theta}{\tan\theta'}\,\mathrm{id} & & \\
0 & \longrightarrow & C_{1,r} \oplus C_{2,r} & \xrightarrow{\alpha_{\theta'}} & C_1 \oplus_{\theta'} C_2 & \xrightarrow{r_{\theta'}} & B & \longrightarrow & 0.
\end{array} \tag{A-20}$$

We apply Lemma 2.5 to the induced isomorphism of the long exact cohomology sequences and find

$$\begin{aligned}
\log \tau(\mathcal{H}(C_{1,r} \oplus C_{2,r}, C_1 \oplus_\theta C_2, B)) &- \log \tau(\mathcal{H}(C_{1,r} \oplus C_{2,r}, C_1 \oplus_{\theta'} C_2, B)) \\
= -\sum_{j\ge 0}(-1)^j \log \operatorname{Det}(\phi_{\theta,\theta',*} &: H^j \to H^j) + \sum_{j\ge 0}(-1)^j \log \operatorname{Det}(\tilde{\phi}_{\theta,\theta',*} : H^j \to H^j) \\
&+ \sum_{j\ge 0}(-1)^j \log \operatorname{Det}(\Psi_{\theta,\theta',*} : H^j \to H^j),
\end{aligned} \tag{A-21}$$

where H^j is shorthand for the respective cohomology groups. Since $\tilde{\phi}_{\theta,\theta'}$ and $\phi_{\theta,\theta'}$ are multiplication operators we have

$$\sum_{j\ge 0}(-1)^j \log \operatorname{Det}(\tilde{\phi}_{\theta,\theta',*} : H^j \to H^j) = \chi(C_{2,r})\,\log\frac{\tan\theta}{\tan\theta'}, \tag{A-22}$$

$$\sum_{j\ge 0}(-1)^j \log \operatorname{Det}(\Psi_{\theta,\theta',*} : H^j \to H^j) = \chi(B)\,\log\frac{\sin\theta}{\sin\theta'}, \tag{A-23}$$

and together with (A-18) we obtain

$$\frac{d}{d\theta}\log\tau\big(\mathscr{H}(C_{1,r}\oplus C_{2,r},C_1\oplus_\theta C_2,B)\big)$$
$$=\frac{2}{\sin 2\theta}\Big(-\sum_{j\geq 0}(-1)^j\operatorname{Tr}\big(\beta_\theta\restriction H^j(C_1\oplus_\theta C_2)\big)+\chi(C_{2,r})\Big)-\frac{\cos\theta}{\sin\theta}\chi(B). \quad \text{(A-24)}$$

Taking into account that $\chi(C_{2,r})=\chi(C_2)-\chi(B)$ (see (A-3)) and that

$$\frac{\cos\theta}{\sin\theta}-\frac{2}{\sin 2\theta}=-\tan\theta,$$

we find (A-9).

Next we apply Proposition 2.6 to the exact sequence (A-4) and get

$$\log\tau(C_1\oplus_\theta C_2)=\log\tau(C_{1,r})+\log\tau(C_2)+\log\tau(\mathscr{H}(C_{1,r},C_1\oplus_\theta C_2,C_2))$$
$$+\frac{1}{2}\sum_{j\geq 0}(-1)^j\log\operatorname{Det}(\beta\beta^*:C_2^j\to C_2^j). \quad \text{(A-25)}$$

Here we have used (2-29) and that α is a partial isometry and thus $\alpha^*\alpha=\mathrm{id}$. Analogously, we infer from (A-5) that

$$\log\tau(C_1\oplus_\theta C_2)=\log\tau(C_{1,r})+\log\tau(C_{2,r})+\log\tau(\mathscr{H}(C_{1,r}\oplus C_{2,r},C_1\oplus_\theta C_2,B))$$
$$+\frac{1}{2}\sum_{j\geq 0}(-1)^j\log\operatorname{Det}(r_\theta r_\theta^*:B^j\to B^j). \quad \text{(A-26)}$$

From (A-25) and (A-26) one deduces the differentiability statement (A-10).

Finally we discuss the case that the maps r_j, $j=1,2$ are partial isometries. Then for $(\xi_1,\xi_2)\in C_1\oplus_\theta C_2$, $\eta\in B$ we calculate

$$\langle r_\theta(\xi_1,\xi_2),b\rangle=\sin\theta\cdot\langle r_1\xi_1,b\rangle+\cos\theta\cdot\langle r_2\xi_2,b\rangle=\langle(\xi_1,\xi_2),(\sin\theta\cdot r_1^*b,\cos\theta\cdot r_2^*b)\rangle. \quad \text{(A-27)}$$

If r_1 and r_2 are partial isometries then $(\sin\theta\cdot r_1^*b,\cos\theta\cdot r_2^*b)\in C_1\oplus_\theta C_2$ and hence it equals $r_\theta^*(b)$. Consequently $r_\theta r_\theta^*b=(\sin^2\theta+\cos^2\theta)b=b$ and thus $\operatorname{Det}(r_\theta r_\theta^*:B^j\to B^j)=1$. Therefore (A-26) reduces to (A-12).

Similarly, one calculates

$$\operatorname{Det}(\beta\beta^*:C_2^j\to C_2^j)=(1+\tan^2)^{-\dim B^j}; \quad \text{(A-28)}$$

then (A-13) follows from (A-25). □

Acknowledgements

I would like to thank Werner Müller and Boris Vertman for pushing and encouraging me to complete this project, the two referees for their careful reading and their valuable suggestions for improvement, and finally Luiz Hartmann for pointing out an embarrassing error in (6-3).

I also owe a lot of gratitude to my family. I dedicate this paper to my dearly beloved late aunt Annels Roth (1923–2012).

References

[Atiyah et al. 1975] M. F. Atiyah, V. K. Patodi, and I. M. Singer, "Spectral asymmetry and Riemannian geometry, I", *Math. Proc. Cambridge Philos. Soc.* **77** (1975), 43–69. MR 53 #1655a Zbl 0297.58008

[Bohn 2009] M. Bohn, *On rho invariants of fiber bundles*, Ph.D. thesis, Universität Bonn, 2009, http://hss.ulb.uni-bonn.de/2009/1830/1830.pdf.

[Brüning and Lesch 1992] J. Brüning and M. Lesch, "Hilbert complexes", *J. Funct. Anal.* **108**:1 (1992), 88–132. MR 93k:58208 Zbl 0826.46065

[Brüning and Lesch 1999] J. Brüning and M. Lesch, "On the η-invariant of certain nonlocal boundary value problems", *Duke Math. J.* **96**:2 (1999), 425–468. MR 99m:58180 Zbl 0956.58014

[Brüning and Lesch 2001] J. Brüning and M. Lesch, "On boundary value problems for Dirac type operators, I: Regularity and self-adjointness", *J. Funct. Anal.* **185**:1 (2001), 1–62. MR 2002g:58034 Zbl 1023.58013

[Brüning and Ma 2006] J. Brüning and X. Ma, "An anomaly formula for Ray–Singer metrics on manifolds with boundary", *Geom. Funct. Anal.* **16**:4 (2006), 767–837. MR 2007i:58042 Zbl 1111.58024

[Brüning and Ma 2013] J. Brüning and X. Ma, "On the gluing formula for the analytic torsion", *Math. Z.* **273**:3-4 (2013), 1085–1117. Zbl 06149055

[Brüning and Seeley 1987] J. Brüning and R. Seeley, "The resolvent expansion for second order regular singular operators", *J. Funct. Anal.* **73**:2 (1987), 369–429. MR 88g:35151 Zbl 0625.47040

[Brüning and Seeley 1988] J. Brüning and R. Seeley, "An index theorem for first order regular singular operators", *Amer. J. Math.* **110**:4 (1988), 659–714. MR 89k:58271 Zbl 0664.58035

[Burghelea et al. 1999] D. Burghelea, L. Friedlander, and T. Kappeler, "Torsions for manifolds with boundary and glueing formulas", *Math. Nachr.* **208** (1999), 31–91. MR 2000m:58053 Zbl 0959.58041

[Cheeger 1979a] J. Cheeger, "Analytic torsion and the heat equation", *Ann. of Math.* (2) **109**:2 (1979), 259–322. MR 80j:58065a Zbl 0412.58026

[Cheeger 1979b] J. Cheeger, "On the spectral geometry of spaces with cone-like singularities", *Proc. Nat. Acad. Sci. U.S.A.* **76**:5 (1979), 2103–2106. MR 80k:58098 Zbl 0411.58003

[Cheeger 1983] J. Cheeger, "Spectral geometry of singular Riemannian spaces", *J. Differential Geom.* **18**:4 (1983), 575–657. MR 85d:58083 Zbl 0529.58034

[Connes and Moscovici 1995] A. Connes and H. Moscovici, "The local index formula in noncommutative geometry", *Geom. Funct. Anal.* **5**:2 (1995), 174–243. MR 96e:58149 Zbl 0960.46048

[Dai and Freed 1994] X. Dai and D. S. Freed, "η-invariants and determinant lines", *J. Math. Phys.* **35**:10 (1994), 5155–5194. MR 96a:58204 Zbl 0822.58048

[Dar 1987] A. Dar, "Intersection R-torsion and analytic torsion for pseudomanifolds", *Math. Z.* **194**:2 (1987), 193–216. MR 88b:58132 Zbl 0605.57012

[Douglas and Wojciechowski 1991] R. G. Douglas and K. P. Wojciechowski, "Adiabatic limits of the η-invariants: The odd-dimensional Atiyah–Patodi–Singer problem", *Comm. Math. Phys.* **142**:1 (1991), 139–168. MR 92j:58110 Zbl 0746.58074

[Gilkey 1995] P. B. Gilkey, *Invariance theory, the heat equation, and the Atiyah–Singer index theorem*, 2nd ed., CRC Press, Boca Raton, FL, 1995. MR 98b:58156 Zbl 0856.58001

[Grubb 1999] G. Grubb, "Trace expansions for pseudodifferential boundary problems for Dirac-type operators and more general systems", *Ark. Mat.* **37**:1 (1999), 45–86. MR 2000c:35265 Zbl 1021.35046

[Hartmann and Spreafico 2010] L. Hartmann and M. Spreafico, "An extension of the Cheeger–Müller theorem for a cone", preprint, 2010. arXiv 1008.2987

[Lesch 1997] M. Lesch, *Operators of Fuchs type, conical singularities, and asymptotic methods*, Teubner-Texte zur Mathematik **136**, B. G. Teubner, Stuttgart, 1997. MR 98d:58174 Zbl 1156.58302

[Lesch and Wojciechowski 1996] M. Lesch and K. P. Wojciechowski, "On the η-invariant of generalized Atiyah–Patodi–Singer boundary value problems", *Illinois J. Math.* **40**:1 (1996), 30–46. MR 97d:58194 Zbl 0866.58054

[Lions and Magenes 1972] J.-L. Lions and E. Magenes, *Non-homogeneous boundary value problems and applications*, vol. 2, Grundlehren der Math. Wissenschaften **182**, Springer, New York, 1972. MR 50 #2671 Zbl 0227.35001

[Mazzeo 1991] R. Mazzeo, "Elliptic theory of differential edge operators, I", *Comm. Partial Differential Equations* **16**:10 (1991), 1615–1664. MR 93d:58152 Zbl 0745.58045

[Mazzeo and Vertman 2012] R. Mazzeo and B. Vertman, "Analytic torsion on manifolds with edges", *Adv. Math.* **231**:2 (2012), 1000–1040. MR 2955200 Zbl 06084090

[Melrose 1993] R. B. Melrose, *The Atiyah–Patodi–Singer index theorem*, Research Notes in Mathematics **4**, A K Peters, Wellesley, MA, 1993. MR 96g:58180 Zbl 0796.58050

[Milnor 1966] J. Milnor, "Whitehead torsion", *Bull. Amer. Math. Soc.* **72** (1966), 358–426. MR 33 #4922 Zbl 0147.23104

[Müller 1978] W. Müller, "Analytic torsion and R-torsion of Riemannian manifolds", *Adv. in Math.* **28**:3 (1978), 233–305. MR 80j:58065b Zbl 0395.57011

[Müller 1983] W. Müller, "Spectral theory for Riemannian manifolds with cusps and a related trace formula", *Math. Nachr.* **111** (1983), 197–288. MR 85i:58121 Zbl 0529.58035

[Müller 1992] W. Müller, "Spectral geometry and scattering theory for certain complete surfaces of finite volume", *Invent. Math.* **109**:2 (1992), 265–305. MR 93g:58151 Zbl 0772.58063

[Müller 1993] W. Müller, "Analytic torsion and R-torsion for unimodular representations", *J. Amer. Math. Soc.* **6**:3 (1993), 721–753. MR 93m:58119 Zbl 0789.58071

[Müller and Vertman 2011] W. Müller and B. Vertman, "The metric anomaly of analytic torsion on manifolds with conical singularities", preprint, 2011. arXiv 1004.2067v3

[Paquet 1982] L. Paquet, "Problèmes mixtes pour le système de Maxwell", *Ann. Fac. Sci. Toulouse Math.* (5) **4**:2 (1982), 103–141. MR 84e:58075 Zbl 0529.58038

[Ray and Singer 1971] D. B. Ray and I. M. Singer, "R-torsion and the Laplacian on Riemannian manifolds", *Advances in Math.* **7** (1971), 145–210. MR 45 #4447 Zbl 0239.58014

[Schulze 1991] B.-W. Schulze, *Pseudo-differential operators on manifolds with singularities*, Studies in Mathematics and its Applications **24**, North-Holland Publishing, Amsterdam, 1991. MR 93b:47109 Zbl 0747.58003

[Shubin 2001] M. A. Shubin, *Pseudodifferential operators and spectral theory*, 2nd ed., Springer, Berlin, 2001. MR 2002d:47073 Zbl 0980.35180

[Taylor 1996] M. E. Taylor, *Partial differential equations, I: Basic theory*, Applied Mathematical Sciences **115**, Springer, New York, 1996. MR 98b:35002b Zbl 0869.35002

[Vertman 2009] B. Vertman, "Analytic torsion of a bounded generalized cone", *Comm. Math. Phys.* **290**:3 (2009), 813–860. MR 2010d:58032 Zbl 1210.58026

[Vishik 1995] S. M. Vishik, "Generalized Ray–Singer conjecture, I: A manifold with a smooth boundary", *Comm. Math. Phys.* **167**:1 (1995), 1–102. MR 96f:58184 Zbl 0818.58049

Received 8 Jun 2012. Revised 10 Sep 2012. Accepted 18 Oct 2012.

MATTHIAS LESCH: *Mathematisches Institut, Universität Bonn, Endenicher Allee 60, D-53115 Bonn, Germany*
www.matthiaslesch.de

Guidelines for Authors

Authors may submit manuscripts in PDF format on-line at the Submission page at msp.berkeley.edu/apde.

Originality. Submission of a manuscript acknowledges that the manuscript is original and and is not, in whole or in part, published or under consideration for publication elsewhere. It is understood also that the manuscript will not be submitted elsewhere while under consideration for publication in this journal.

Language. Articles in APDE are usually in English, but articles written in other languages are welcome.

Required items. A brief abstract of about 150 words or less must be included. It should be self-contained and not make any reference to the bibliography. If the article is not in English, two versions of the abstract must be included, one in the language of the article and one in English. Also required are keywords and subject classifications for the article, and, for each author, postal address, affiliation (if appropriate), and email address.

Format. Authors are encouraged to use LaTeX but submissions in other varieties of TeX, and exceptionally in other formats, are acceptable. Initial uploads should be in PDF format; after the refereeing process we will ask you to submit all source material.

References. Bibliographical references should be complete, including article titles and page ranges. All references in the bibliography should be cited in the text. The use of BibTeX is preferred but not required. Tags will be converted to the house format, however, for submission you may use the format of your choice. Links will be provided to all literature with known web locations and authors are encouraged to provide their own links in addition to those supplied in the editorial process.

Figures. Figures must be of publication quality. After acceptance, you will need to submit the original source files in vector graphics format for all diagrams in your manuscript: vector EPS or vector PDF files are the most useful.

Most drawing and graphing packages (Mathematica, Adobe Illustrator, Corel Draw, MATLAB, etc.) allow the user to save files in one of these formats. Make sure that what you are saving is vector graphics and not a bitmap. If you need help, please write to graphics@msp.org with details about how your graphics were generated.

White space. Forced line breaks or page breaks should not be inserted in the document. There is no point in your trying to optimize line and page breaks in the original manuscript. The manuscript will be reformatted to use the journal's preferred fonts and layout.

Proofs. Page proofs will be made available to authors (or to the designated corresponding author) at a Web site in PDF format. Failure to acknowledge the receipt of proofs or to return corrections within the requested deadline may cause publication to be postponed.

ANALYSIS & PDE

Volume 6 No. 1 2013

2157-5045(2013)6:1;1-H