PDMI PREPRINT — 20/2003

# THE SHARP CONSTANT IN THE JOHN–NIRENBERG INEQUALITY

**V. Vasyunin**

Russian Academy of Sciences
Steklov Mathematical Institute

191011, St.Petersburg, Fontanka, 27
October 3, 2003

ABSTRACT

The sharp constant in the integral form of the John–Nirenberg inequality is found by constructing of the exact Bellman function of the corresponding extremal problem. The explicit solution of the Bellman equation is found. Moreover, the extremal function is presented that proves sharpness of the constant.

This research was partially supported by grants RFBR 02–01–00264 and "Veduschie Nauchnye Schkoly" NSh–2266.2003.1

**Notation.** For any interval $I \subset \mathbb{R}$ and for a function $\varphi \in L^1(I)$ the symbol $\langle\varphi\rangle_I$ means its average over this interval: $\langle\varphi\rangle_I \stackrel{\text{def}}{=} \frac{1}{|I|}\int_I \varphi(t)dt$. By $\mathrm{BMO}_\delta(J)$ we denote the set of all functions $\varphi \in L^2(J)$ such that

$$\sup_{I\subset J}\{\langle\varphi^2\rangle_I - \langle\varphi\rangle_I^2\} \le \delta^2. \tag{1}$$

**Theorem 1.** *For any function $\varphi \in \mathrm{BMO}_\delta(J)$ with $\delta < 1$ the following inequality is true*

$$\langle e^\varphi\rangle_J \ \le\ C(\delta)e^{\langle\varphi\rangle_J}, \tag{2}$$

*where the constant*

$$C(\delta) = \frac{e^{-\delta}}{1-\delta} \tag{3}$$

*is sharp.*

Note that $\langle\varphi\rangle_I^2 \le \langle\varphi^2\rangle_I$, and to arbitrary function $\varphi \in \mathrm{BMO}_\delta(J)$ we can put in correspondence the point $x = (x_1, x_2) = (\langle\varphi\rangle_I, \langle\varphi^2\rangle_I)$ in the domain

$$\Omega_\delta \stackrel{\text{def}}{=} \{(x_1, x_2)\colon x_1 \in \mathbb{R},\, x_1^2 \le x_2 \le \delta^2 + x_1^2\}.$$

We define on $\Omega_\delta$ two following functions:

$$\mathbf{B}_\delta^+(x) \stackrel{\text{def}}{=} \sup_{\varphi\in\mathrm{BMO}_\delta(I)}\Big\{\langle e^\varphi\rangle_I\colon \langle\varphi\rangle_I = x_1,\ \langle\varphi^2\rangle_I = x_2\Big\}, \tag{4}$$

$$\mathbf{B}_\delta^-(x) \stackrel{\text{def}}{=} \inf_{\varphi\in\mathrm{BMO}_\delta(I)}\Big\{\langle e^\varphi\rangle_I\colon \langle\varphi\rangle_I = x_1,\ \langle\varphi^2\rangle_I = x_2\Big\}. \tag{5}$$

Note, that $\mathbf{B}_\delta^\pm$ do not depend on $I$. In spite of the fact that $\mathbf{B}_\delta^\pm$ essentially depend on $\delta$, we shall omit sometimes this subscript, when it is clear that $\delta$ is fixed.

Theorem 1 is an immediate consequence of the following more refinement estimate.

**Theorem 2.** *If $\delta \in [0,1)$, then*

$$\mathbf{B}_\delta^+(x) = \frac{1-\sqrt{\delta^2+x_1^2-x_2}}{1-\delta}\exp\Big(x_1 + \sqrt{\delta^2+x_1^2-x_2} - \delta\Big); \tag{6}$$

*if $\delta \ge 1$, then*

$$\mathbf{B}_\delta^+(x) = \begin{cases} e^{x_1} & \text{for} \quad x_2 = x_1^2, \\ +\infty & \text{for} \quad x_2 > x_1^2; \end{cases}$$

*and*

$$\mathbf{B}_\delta^-(x) = \frac{1+\sqrt{\delta^2+x_1^2-x_2}}{1+\delta}\exp\Big(x_1 - \sqrt{\delta^2+x_1^2-x_2} + \delta\Big). \tag{7}$$

*for all $\delta \ge 0$.*

Indeed, since the function $t \mapsto (1-t)e^t$ monotonously decreases for positive $t$, the function $\mathbf{B}_\delta^+(x)$ takes its maximal value when $x_2 = x_1^2 + \delta^2$, i.e.,

$$\mathbf{B}_\delta^+(x) \le \frac{e^{-\delta}}{1-\delta}e^{x_1},$$

what gives (2) with a sharp constant.

**Outline of the proof of Theorem 2.** By $B_\delta^+(x)$ and $B_\delta^-(x)$ we denote the functions on the right hand side of (6) and (7), respectively. First we prove easier inequalities:

**Lemma 1.** *The inequalities*

$$\mathbf{B}_\delta^+(x) \geq B_\delta^+(x), \qquad \mathbf{B}_\delta^-(x) \leq B_\delta^-(x)$$

*are true for every $x \in \Omega_\delta$.*

These inequalities will be proved by direct writing down the extremal functions. The opposite inequalities we cannot check directly and we use an approximating procedure, namely, we prove the following lemma.

**Lemma 2.** *The inequalities*

$$\mathbf{B}_\delta^+(x) \leq B_\varepsilon^+(x), \qquad \mathbf{B}_\delta^-(x) \geq B_\varepsilon^-(x).$$

*are true for every $x \in \Omega_\delta$ and for every $\varepsilon > \delta$.*

But since $B_\delta$ are continuous from above in $\delta$, we can pass to the limit $\varepsilon \to \delta$ and obtain the second pair of the desired inequalities.

We start to prove this lemma by verifying that $B_\delta^+$ is a concave function and $B_\delta^-$ is a convex one.

**Lemma 3.** *If the whole interval with the endpoints $x^\pm$ is in $\Omega_\delta$, then the inequalities*

$$B_\delta^+(\alpha_- x^- + \alpha_+ x^+) \geq \alpha_- B_\delta^+(x^-) + \alpha_+ B_\delta^+(x^+), \tag{8}$$

$$B_\delta^-(\alpha_- x^- + \alpha_+ x^+) \leq \alpha_- B_\delta^-(x^-) + \alpha_+ B_\delta^-(x^+), \tag{9}$$

*are true for an arbitrary pair of nonnegative numbers $\alpha_\pm$ such that $\alpha_- + \alpha_+ = 1$.*

For a given function $\varphi \in \mathrm{BMO}_\delta(J)$ and a given interval $I \subset J$, if we take an arbitrary splitting $I = I_- \cup I_+$, then the interval with the end points $x^\pm = (\langle\varphi\rangle_{I_\pm}, \langle\varphi^2\rangle_{I_\pm})$ not obligatory is inside the domain $\Omega_\delta$. We avoid this difficulty in the following lemma.

**Lemma 4.** *Let the parameter $\delta$ is given. Then for an arbitrary $\varepsilon > \delta$ and for an arbitrary function $\varphi$, $\varphi \in \mathrm{BMO}_\delta(I)$, there exists a splitting $I = I_- \cup I_+$ such that the whole interval with the end points $x^\pm = (\langle\varphi\rangle_{I_\pm}, \langle\varphi^2\rangle_{I_\pm})$ is in $\Omega_\varepsilon$. Moreover, the splitting parameter $\alpha_+$ can be chosen uniformly (with respect to $\varphi$, and therefore with respect to $I$ as well) separated from $0$ and $1$.*

Now we start to realize the programm outlined above.

**Proof of Lemma 1.** Since $x_2 = x_1^2$ ocuurs if and only if $\varphi = x_1 = \text{const}$, it is clear that $\mathbf{B}_0^\pm = B_0^\pm = e^{x_1}$. So we need to consider $\delta > 0$.

Take $I = [0,1]$, $a \in (0,1]$, $b \in \mathbb{R}$, and $\gamma \in \mathbb{R} \setminus \{0\}$, and put

$$\varphi_{a,b,\gamma}(t) = \begin{cases} \gamma \log \frac{a}{t} + b & \text{for} \quad 0 \leq t \leq a, \\ b & \text{for} \quad a \leq t \leq 1. \end{cases}$$

By direct calculations we check that $\varphi_{a,b,\gamma} \in \mathrm{BMO}_{|\gamma|}(I)$, $\langle \varphi_{a,b,\gamma}\rangle_I = \gamma a + b$, $\langle \varphi^2_{a,b,\gamma}\rangle_I = 2\gamma^2 a + 2\gamma ab + b^2$, and

$$\langle e^{\varphi_{a,b,\gamma}}\rangle_I = \begin{cases} \frac{1-\gamma+a\gamma}{1-\gamma} e^b, & \text{for} \quad \gamma < 1; \\ \infty, & \text{for} \quad \gamma \geq 1. \end{cases}$$

Since $\mathbf{B}^{\pm}_{\delta}(x_1, x_1^2) = B^{\pm}_{\delta}(x_1, x_1^2) = e^{x_1}$ for all $\delta$, we need to consider only the points $x \in \Omega_\delta$ with $x_2 > x_1^2$. Then we can take $a = 1 - \frac{1}{\gamma}\sqrt{\gamma^2 + x_1^2 - x_2} \in (0,1]$ and $b = x_1 - \gamma a$, and we get $\langle \varphi_{a,b,\gamma}\rangle_I = x_1$, $\langle \varphi^2_{a,b,\gamma}\rangle_I = x_2$. Now, if we put $\gamma = \delta \geq 1$, then we have $\mathbf{B}^+_\delta(x) = \infty$. For $\gamma = \delta \in (0,1)$ we get

$$\mathbf{B}^+_\delta(x_1, x_2) \geq \langle e^{\varphi_{a,b,\delta}}\rangle_I = \frac{1 - \sqrt{\delta^2 + x_1^2 - x_2}}{1-\delta} \exp\Big(x_1 + \sqrt{\delta^2 + x_1^2 - x_2} - \delta\Big) = B^+_\delta(x_1, x_2).$$

If we take $\gamma = -\delta \in (-\infty, 0)$, then we conclude

$$\mathbf{B}^-_\delta(x_1, x_2) \leq \langle e^{\varphi_{a,b,-\delta}}\rangle_I$$
$$= \frac{1 + \sqrt{\delta^2 + x_1^2 - x_2}}{1+\delta} \exp\Big(x_1 - \sqrt{\delta^2 + x_1^2 - x_2} + \delta\Big) = B^-_\delta(x_1, x_2). \quad \square$$

**Proof of Lemma 3.** To prove the Lemma we need to check that

$$\mp \frac{\partial^2 B^{\pm}}{\partial x_i \partial x_j} \tag{10}$$

is a nonnegative matrix. Direct calcuation yields

$$\begin{aligned}
\frac{\partial B^{\pm}}{\partial x_1} &= \frac{1 - x_1 \mp \sqrt{\delta^2 + x_1^2 - x_2}}{1 \mp \delta} \exp\left\{x_1 \pm \sqrt{\delta^2 + x_1^2 - x_2} \mp \delta\right\}, \\
\frac{\partial B^{\pm}}{\partial x_2} &= \frac{1}{2(1 \mp \delta)} \exp\left\{x_1 \pm \sqrt{\delta^2 + x_1^2 - x_2} \mp \delta\right\}, \\
\frac{\partial^2 B^{\pm}}{\partial x_1^2} &= \mp \frac{\Big(x_1 \pm \sqrt{\delta^2 + x_1^2 - x_2}\Big)^2}{\sqrt{\delta^2 + x_1^2 - x_2}(1 \mp \delta)} \exp\left\{x_1 \pm \sqrt{\delta^2 + x_1^2 - x_2} \mp \delta\right\}, \\
\frac{\partial^2 B^{\pm}}{\partial x_1 \partial x_2} &= \pm \frac{x_1 \pm \sqrt{\delta^2 + x_1^2 - x_2}}{2\sqrt{\delta^2 + x_1^2 - x_2}(1 \mp \delta)} \exp\left\{x_1 \pm \sqrt{\delta^2 + x_1^2 - x_2} \mp \delta\right\}, \\
\frac{\partial^2 B^{\pm}}{\partial x_2^2} &= \mp \frac{1}{4\sqrt{\delta^2 + x_1^2 - x_2}(1 \mp \delta)} \exp\left\{x_1 \pm \sqrt{\delta^2 + x_1^2 - x_2} \mp \delta\right\},
\end{aligned}$$

Therefore the quadratic form of the matrix (10) is

$$\begin{aligned}
&\mp \sum_{i,j=1}^{2} \frac{\partial^2 B^{\pm}}{\partial x_i \partial x_j} \Delta_i \Delta_j \\
&= \frac{\Big(\big(x_1 \pm \sqrt{\delta^2 + x_1^2 - x_2}\big)\Delta_1 - \frac{1}{2}\Delta_2\Big)^2}{\sqrt{\delta^2 + x_1^2 - x_2}(1 \mp \delta)} \exp\left\{x_1 \pm \sqrt{\delta^2 + x_1^2 - x_2} \mp \delta\right\} \geq 0 \qquad \square
\end{aligned} \tag{11}$$

**Proof of Lemma 4.** For any interval $I$ we define a procedure how to split it in two parts $I_-$ and $I_+$, $\alpha_\pm = |I_\pm|/|I|$. Here, as before, we put, $x_1^\pm = \langle\varphi\rangle_{I_\pm}$, $x_2^\pm = \langle\varphi^2\rangle_{I_\pm}$ and $x_1^0 = \langle\varphi\rangle_I$, $x_2^0 = \langle\varphi^2\rangle_I$.

First we take $\alpha_- = \alpha_+ = \frac{1}{2}$. If the whole segment of the straight line connecting the points $x^\pm$ is in the domain $\Omega_\varepsilon$, then we fix this splitting. And in the rest of the proof we assume that it is not the case, i.e., there exists a point on this interval with $x_2 - x_1^2 > \varepsilon^2$. Since any straight line intersects the parabola $x_2 = x_1^2 + \varepsilon^2$ at most at two points, we can choose $\xi$ to be the end point of our segment such that the points $x$ with $x_2 - x_1^2 > \varepsilon^2$ are between $x^0$ and $\xi$. Define a function $\rho(\alpha_+)$ as follows: for a given $\alpha_+$ we put $\rho(\alpha_+)$ to be the maximum value of $x_2 - x_1^2$, when $x$ runs over the segment between $x^0$ and $\xi$. By our assumption $\rho(\frac{1}{2}) > \varepsilon^2$. We begin to change $\alpha_+$ choosing the direction, where $\xi$ goes to $x^0$, i.e., we shall increase $\alpha_+$, if $\xi = x^+$, and decrease it, if $\xi = x^-$. We stop at the moment, when $\rho(\alpha_+) = \varepsilon^2$ at the first time. This will be our procedure how to choose the splitting of a given interval $I$. Now we have to check that such an intersection occurs at all, and moreover, the corresponding value of $\alpha_+$ is uniformly (with respect to $\varphi$ and $I$) separated from 0 and 1.

Since $x_2^0 - (x_1^0)^2 \le \delta^2$, we have $x_2 - x_1^2 < \varepsilon^2$ on the whole segment between $x^0$ and $\xi$, if $\xi$ is close enough to $x^0$, i.e., $\rho(\alpha_+) < \varepsilon^2$ for the corresponding $\alpha_+$. Since the function $\rho$ is continuous, there exists an $\alpha_+$ such that $\rho(\alpha_+) = \varepsilon^2$.

Thus we have proved that the described stop point exists and we have to check that the corresponding $\alpha_+$ is not too close to 0 and 1. If $\xi = x_+$, then $\alpha_+ > \frac{1}{2}$ and $\xi_1 - x_1^0 = x_1^+ - x_1^0 = \alpha_-(x_1^+ - x_1^-)$. Symmetricaly, if $\xi = x_-$, then $\alpha_- > \frac{1}{2}$ and $\xi_1 - x_1^0 = x_1^- - x_1^0 = \alpha_+(x_1^- - x_1^+)$. Thus $|\xi_1 - x_1^0| = \min\{\alpha_\pm\}|x_1^- - x_1^+|$.

For the stop value of $\alpha_+$ our straight line passing through the points $x^\pm$ and $x^0$ is tangent to the parabola $x_2 = x_1^2 + \varepsilon^2$, touching it at some point, say $y$, therefore the equation of this line is $x_2 = 2x_1y_1 - y_1^2 + \varepsilon^2$. This line intersects the graph $x_2 = x_1^2 + \delta^2$ at the points

$$x(\delta)^\pm = (y_1 \pm \sqrt{\varepsilon^2 - \delta^2}, y_2 \pm 2y_1\sqrt{\varepsilon^2 - \delta^2}).$$

and the graph $x_2 = x_1^2$ at the points

$$x(0)^\pm = (y_1 \pm \varepsilon, y_2 \pm 2y_1\varepsilon).$$

Then we have the following embedded segments of our line: $[x(\delta)^-, x(\delta)^+] \subset [x^0, \xi] \subset [x^-, x^+] \subset [x(0)^-, x(0)^+]$, and therefore

$$\begin{aligned} 2\sqrt{\varepsilon^2 - \delta^2} &= |x(\delta)_1^+ - x(\delta)_1^-| \le |x_1^0 - \xi_1| = \min\{\alpha_-, \alpha_+\}|x_1^+ - x_1^-| \\ &\le \min\{\alpha_-, \alpha_+\}|x(0)_1^+ - x(0)_1^-| = \min\{\alpha_-, \alpha_+\}2\varepsilon, \end{aligned}$$

and

$$\sqrt{1 - \left(\frac{\delta}{\varepsilon}\right)^2} \le \alpha_+ \le 1 - \sqrt{1 - \left(\frac{\delta}{\varepsilon}\right)^2}.$$

We see that this estimate depends on $\delta$ and $\varepsilon$ only, but not on the choice of $\varphi$ and $I$. $\square$

Now, we are ready to prove the main lemma.

**Proof of Lemma 2.** Fix an arbitrary $\varepsilon > \delta$. Starting with an arbitrary interval $I = I^{0,0}$ we construct a chain of intervals $I^{n,m}$ according to the rule from Lemma 4. By $I^{n,m}$ we denote the intervals that appear in the $n$-th generation as follows: $I^{n,2k} = I_-^{n-1,k}$ and $I^{n,2k+1} = I_+^{n-1,k}$, so the second index runs from 0 to $2^n - 1$. The corresponding mean values will be indexed by the same pair of indices, and $\alpha_{n,m} = |I^{n,m}|/|I|$. Note that due to the fact that in Lemma 4 we have the value $\alpha_+$ separated from 0 and 1, the maximal length of the intervals of $n$-th generation tends to 0, when $n \to \infty$. So, we have

$$
\begin{aligned}
B_\varepsilon^+(x^{0,0}) &\geq \alpha_{1,0} B_\varepsilon^+(x^{1,0}) + \alpha_{1,1} B_\varepsilon^+(x^{1,1}) \geq \sum_{m=0}^{2^n-1} \alpha_{n,m} B_\varepsilon^+(x^{n,m}) \\
&\geq \sum_{m=0}^{2^n-1} \alpha_{n,m} e^{x_1^{n,m}} = \frac{1}{|I|} \int_I e^{\varphi_n(s)} ds,
\end{aligned} \tag{12}
$$

where $\varphi_n$ is the step function $\varphi_n(s) = x_1^{m,n}$ for $s \in I^{m,n}$. Since $\varphi_n$ converges almost everywhere to $\varphi$, Fatou theorem yields:

$$
\begin{aligned}
B_\varepsilon^+(x^{0,0}) &\geq \frac{1}{|I|} \limsup \int_I e^{\varphi_n(s)} ds \geq \frac{1}{|I|} \int_I \limsup e^{\varphi_n(s)} ds \\
&= \frac{1}{|I|} \int_I e^{\varphi(s)} ds = \langle e^\varphi \rangle_I.
\end{aligned} \tag{13}
$$

Taking supremum over all $\varphi$ with $\langle \varphi \rangle_I = x_1^{0,0} = x_1$ and $\langle \varphi^2 \rangle_I = x_2^{0,0} = x_2$ we get one of the required estimate $B_\varepsilon^+(x_1, x_2) \geq \mathbf{B}_\delta^+(x_1, x_2)$. To obtain the inequality $B_\varepsilon^-(x_1, x_2) \leq \mathbf{B}_\delta^-(x_1, x_2)$ we follow exactly the consideration above replacing in (12) and (13) the sign of the inequality by the opposite one and taking in (13) $\liminf$ rather than $\limsup$. $\square$

**Appendix 1: how to find the Bellman functions** (6)–(7).

We look for a concave function $\mathbf{B}^+$ and a convex function $\mathbf{B}^-$ on some domain $\Omega_\delta$. First we note that the function $\mathbf{B}^\pm$ have to be of the following form

$$
\mathbf{B}_\delta^\pm(x_1, x_2) = \exp\{x_1 + w_\delta^\pm(x_2 - x_1^2)\}, \tag{14}
$$

where $w_\delta^\pm$ are positive functions on $[0, \delta^2]$ with $w_\delta^\pm(0) = 0$.

Indeed, if a function $\varphi$ runs over the whole $\mathrm{BMO}_\delta(J)$ then the function $\tilde\varphi = \varphi + c$ does, where $c$ is an arbitrary constant. Then we have $\langle \tilde\varphi \rangle = \langle \varphi \rangle + c$, $\langle \tilde\varphi^2 \rangle = \langle \varphi^2 \rangle + 2c\langle \varphi \rangle + c^2$, and $\langle e^{\tilde\varphi} \rangle = e^c \langle e^\varphi \rangle$. Taking supremum (or infimum) in the latter identity we get

$$
\mathbf{B}(x_1 + c, x_2 + 2cx_1 + c^2) = e^c \mathbf{B}(x_1, x_2).
$$

Putting $c = -x_1$ we obtain the following identity

$$
\mathbf{B}(0, x_2 - x_1^2) = e^{-x_1} \mathbf{B}(x_1, x_2),
$$

By the Jensen inequality ($\langle e^\varphi \rangle \geq e^{\langle \varphi \rangle}$) the values of the latter function are not less then one. Hense there exist positive functions $w = w_\delta^\pm$ defined on the interval $[0, \delta^2]$ such that (14) is fulfilled. Moreover, since the equality $x_2 = x_1 = 0$ occurs iff $\varphi = 0$, we have $\mathbf{B}(0,0) = 1$, i.e., $w(0) = 0$.

Now, we calculate the matrix (10) using representation (14):

$$\begin{aligned}
\frac{\partial \mathbf{B}}{\partial x_1} =&(1-2x_1 w')\mathbf{B}, \qquad \frac{\partial \mathbf{B}}{\partial x_2} = w'\mathbf{B},\\
\frac{\partial^2 \mathbf{B}}{\partial x_1^2} =&\big((1-2x_1 w')^2 - 2w' + 4x_1^2 w''\big)\mathbf{B},\\
\frac{\partial^2 \mathbf{B}}{\partial x_1 \partial x_2} =&\big(w'(1-2x_1 w') - 2x_1 w''\big)\mathbf{B},\\
\frac{\partial^2 \mathbf{B}}{\partial x_2^2} =&\big((w')^2 + w''\big)\mathbf{B}.
\end{aligned}$$

Thus, the matrix (10) turns into

$$\begin{pmatrix} \frac{\partial^2 \mathbf{B}}{\partial x_1^2} & \frac{\partial^2 \mathbf{B}}{\partial x_1 \partial x_2} \\ \frac{\partial^2 \mathbf{B}}{\partial x_1 \partial x_2} & \frac{\partial^2 \mathbf{B}}{\partial x_2^2} \end{pmatrix} = \mathbf{B} \begin{pmatrix} 1 & -2x_1 \\ 0 & 1 \end{pmatrix} R \begin{pmatrix} 1 & 0 \\ -2x_1 & 1 \end{pmatrix}, \tag{15}$$

where

$$R = \begin{pmatrix} 1-2w' & w' \\ w' & (w')^2 + w'' \end{pmatrix}.$$

Since on an extremal function (if any) the equality has to be on each step in (12), the matrix (10) has to be degenerated. Due to representation (15) we can rewrite these conditions in the form

$$(1-2w')\big((w')^2 + w''\big) = (w')^2, \tag{16}$$

$$\pm\,(2(w^{\pm})' - 1) \geq 0. \tag{17}$$

Let us solve equation (16).

$$\begin{gathered}
(1-2w')w'' = 2(w')^3\\
\left(\frac{1}{2(w')^3} - \frac{1}{(w')^2}\right) w'' = 1\\
\left(\frac{1}{w'} - \frac{1}{4(w')^2}\right)' = 1\\
\frac{1}{w'} - \frac{1}{4(w')^2} = t + \mathrm{const}\\
-\left(1 - \frac{1}{2w'}\right)^2 = t + \mathrm{const}
\end{gathered}$$

For the latter relation to determine a real valued function $w'$ at least at the point $t = 0$, the constant has to be nonpositive. So, we parametrize the family of possible solutions by a positive parameter $\delta$ by putting $\mathrm{const} = -\delta^2$, then

$$\left(1 - \frac{1}{2w'}\right)^2 = \delta^2 - t$$

and

$$1 - \frac{1}{2(w^{\pm})'} = \pm\sqrt{\delta^2 - t}. \tag{18}$$

Condition (17) for $w^+$ requires for the square root to be strictly less than 1. Therefore, we can get a required solution $w_\delta^+$ only for $\delta < 1$ and this solution will be defined on the interwal $t \in [0, \delta^2]$. The second solution $w_\delta^-$ is defined on the same interval and satisfies (17) for all $\delta > 0$.

Now, rewriting equation (18) as

$$(w^{\pm})' = \frac{1}{2(1 \mp \sqrt{\delta^2 - t})}$$

and taking into account the boundary condition $w(0) = 0$ we obtain

$$w^{\pm}(t) = \frac{1}{2}\int_0^t \frac{dx}{1 \mp \sqrt{\delta^2 - x}} = \log\frac{1 \mp \sqrt{\delta^2 - t}}{1 \mp \delta} \pm \sqrt{\delta^2 - t} \mp \delta,$$

what yields (6) for $\mathbf{B}_\delta^+(x)$ and (7) for $\mathbf{B}_\delta^+(x)$.

**Appendix 2: how to find an extremal function.** In the proof of Lemma 1 the required extremal function appeared without any explanations, now we explain, how to find it.

It was already mentioned that for an extremal function we must have the equality in (13) (in the limit $\varepsilon \to \delta$) and on each splitting step in the chain of inequalities (12). This means that in the splitting process we "go" only along the vectorial field defined by the kernel vectors of the matrix (10). From (11) we get the equation for the trajectories, on which $\mathbf{B}^{\pm}$ are linear functions

$$\left(x_1 \pm \sqrt{\delta^2 + x_1^2 - x_2}\right) dx_1 = \frac{1}{2} dx_2. \tag{19}$$

If to introduce $t = \pm\sqrt{\delta^2 + x_1^2 - x_2}$ or $t^2 = \delta^2 + x_1^2 - x_2$, then we get $2t\,dt = 2x_1 dx_1 - dx_2$. Replacing $\frac{1}{2}dx_2$ in (19) by $x_1 dx_1 - t\,dt$ we get $t\,dx_1 = -t\,dt$, i.e., $t = c - x_1$ and

$$x_2 = \delta^2 + x_1^2 - t^2 = 2cx_1 + \delta^2 - c^2.$$

The corresponding trajectories are the family of the straight lines tangential to the upper bound $x_2 = x_1^2 + \delta^2$ of our domain $\Omega_\delta$ at the point $x = (c, c^2 + \delta^2)$. In fact this is two families of straight segments

$$\omega_\delta^+(c) = \{x = (x_1, 2cx_1 + \delta^2 - c^2) \colon c - \delta \le x_1 \le c\},$$
$$\omega_\delta^-(c) = \{x = (x_1, 2cx_1 + \delta^2 - c^2) \colon c \le x_1 \le c + \delta\}.$$

Each of them covers the whole domain $\Omega_\delta$:

$$\Omega_\delta = \bigcup_{c\in\mathbb{R}} \omega_\delta^+(c) = \bigcup_{c\in\mathbb{R}} \omega_\delta^-(c),$$

and they correspond to the functions $\mathbf{B}_\delta^+$ and $\mathbf{B}_\delta^-$ respectively: $\mathbf{B}_\delta^+$ is a linear function on each segment $\omega_\delta^+(c)$ and $\mathbf{B}_\delta^-$ is a linear function on each segment $\omega_\delta^-(c)$. Indeed, since $\sqrt{\delta^2+x_1^2-x_2}=|x_1-c|$ on the line $x_2=2cx_1+\delta^2-c^2$, we have

$$
\begin{aligned}
\mathbf{B}_\delta^+(x_1,2cx_1+\delta^2-c^2)&=\frac{1+x_1-c}{1-\delta}e^{c-\delta} \quad \text{for} \quad c-\delta\le x_1\le c,\\
\mathbf{B}_\delta^-(x_1,2cx_1+\delta^2-c^2)&=\frac{1+x_1-c}{1+\delta}e^{c+\delta} \quad \text{for} \quad c\le x_1\le c+\delta.
\end{aligned}
$$

Therefore, if the points $x^\pm$ are on a segment $\omega_\delta^+(c)$, then we have equality in (8), and if they are on a segment $\omega_\delta^-(c)$, then we have equality in (9).

Note that we have one more "acceptable trajectory", which is the the envelope of the segments $\omega_\delta^+(c)$ (or $\omega_\delta^-(c)$): $x_2=x_1^2+\delta^2$.

First we consider the case when the point is on this critical line: $x_1^0$ is arbitrary and $x_2^0=(x_1^0)^2+\delta^2$. Let $I=[0,1]$. Split it at a point $a$: $I_-=[0,a]$, $I_+=[a,1]$. The points $x^\pm$ has to be on our "acceptable trajectory", i.e., for $\mathbf{B}_\delta^+$ ($\mathbf{B}_\delta^-$, respectively) on the line $x_2=x_1^2+\delta^2$ to the right (to the left, respectively) from $x^0$ and on the line $x_2=2x_1^0x_1+\delta^2-(x_1^0)^2$ to the left (to te right, respectively). Let us consider the case $\mathbf{B}_\delta^+$, consideration of $\mathbf{B}_\delta^-$ differs in the sign only. Due to the symmetry we can choose, which point of $x^\pm$ is to the left from $x^0$. Let it will be $x^+$. Then for all $a\in[0,1]$ we have

$$
\delta^2=x_2^- -(x_1^-)^2=\frac{1}{a}\int_0^a\varphi(t)^2dt-\left(\frac{1}{a}\int_0^a\varphi(t)dt\right)^2. \tag{20}
$$

If we introduce $\psi(a)=\frac{1}{\delta}\int_0^a\varphi(t)dt$, then $\varphi=\delta\psi'$ and after multiplication by $a$ (20) turns into

$$
\begin{gathered}
\int_0^a\psi'(t)^2dt=a+\frac{1}{a}\psi(a)^2\\
\psi'(a)^2=\frac{d}{da}\Big(a+\frac{1}{a}\psi(a)^2\Big)=1-\frac{\psi(a)^2}{a^2}+\frac{2\psi(a)\psi'(a)}{a}\\
\Big(\psi'(a)-\frac{\psi(a)}{a}\Big)^2=1
\end{gathered}
$$

Once we consider $\mathbf{B}^+$, we have to take here the minus sign for the square root (plus has to be taken for $\mathbf{B}^-$):

$$
\begin{gathered}
\psi'(a)-\frac{\psi(a)}{a}=-1\\
\Big(\frac{\psi(a)}{a}\Big)'=-\frac{1}{a}\\
\psi(a)=a\big(\log\frac{1}{a}+\text{const}\big)
\end{gathered}
$$

Thus, finaly we have

$$
\varphi(t)=\delta\Big(\log\frac{1}{t}-1\Big)+x_1^0, \tag{21}
$$

where const is chosen in such a way to have $x_1^0$ as the mean value of $\varphi$.

Now let $x^0$ be an arbitrary point in the strip $\Omega_\delta$. Then we make the following splitting: the points $x^\pm$ are the end point of the seguement $\omega_\delta^+(x_1^-)$ containing the point $x^0$, i.e.,

$$x_2 = 2x_1^- x_1 + \delta^2 - (x_1^-)^2,$$

with $x_1^- = x_1^0 + \sqrt{\delta^2 + (x_1^0)^2 - x_2^0}$ (for $\mathbf{B}^-$ we have to choose the minus sign for the square root). For the second end point we have $x_2^+ = (x_1^+)^2 = 2x_1^- x_1^+ + \delta^2 - (x_1^-)^2$ and hence $x_1^+ = x_1^- - \delta$ ($x_1^+ = x_1^- + \delta$ for $\mathbf{B}^-$). Then for the splitting point $a = \alpha_-$ we have an equation

$$x_1^0 = \alpha_- x_1^- + \alpha_+ x_1^+ = a x_1^- + (1-a)x_1^+ = x_1^+ + a\delta,$$

whence

$$a = \frac{x_1^0 - x_1^+}{\delta} = 1 - \frac{x_1^- - x_1^0}{\delta} = 1 - \frac{1}{\delta}\sqrt{\delta^2 + (x_1^0)^2 - x_2^0}.$$

Since $x_2^+ = (x_1^+)^2$, the function $\varphi$ is a constant on $I_+$, i.e., $\varphi|_{I_+} = x_1^+$, and for the interval $I_-$ we have to take the extremal function (21) renormalizing to this interval: $\varphi|_{I_-} = \delta(\log \frac{a}{t} - 1) + x_1^-$. This yield the extremal function we used to prove Lemma 1.