(iii) $\implies$ (iv)

Fix a $\mathscr{C}_0^\infty$ function ϕ with $0 \le \phi \le 1$, supported in the ball $B(0,4)$, and equal to 1 on the ball $B(0,2)$. We consider the functions $\phi(\cdot/R)$ that tend to 1 as $R \to \infty$ and we show that $T(1)$ is the weak limit of the functions $T(\phi(\cdot/R))$. This means that for all $g \in \mathscr{D}_0$ (smooth functions with compact support and integral zero) one has

$$\big\langle T(\phi(\cdot/R)), g\big\rangle \to \big\langle T(1), g\big\rangle \tag{4.3.5}$$

as $R \to \infty$. To prove (4.3.5) we fix a $\mathscr{C}_0^\infty$ function η that is equal to one on a neighborhood of the support of g. Then we write

$$\begin{aligned}\big\langle T(\phi(\cdot/R)), g\big\rangle &= \big\langle T(\eta\phi(\cdot/R)), g\big\rangle + \big\langle T((1-\eta)\phi(\cdot/R)), g\big\rangle\\ &= \big\langle T(\eta\phi(\cdot/R)), g\big\rangle\\ &\qquad + \int_{\mathbf{R}^n}\int_{\mathbf{R}^n}\big(K(x,y)-K(x_0,y)\big)g(x)(1-\eta(y))\phi(y/R)\,dy\,dx\,,\end{aligned}$$

where x_0 is a point in the support of g. There exists an $R_0 > 0$ such that for $R \ge R_0$, $\phi(\cdot/R)$ is equal to 1 on the support of η, and moreover the expressions

$$\int_{\mathbf{R}^n}\int_{\mathbf{R}^n}\big(K(x,y)-K(x_0,y)\big)g(x)(1-\eta(y))\phi(y/R)\,dy\,dx$$

converge to

$$\int_{\mathbf{R}^n}\int_{\mathbf{R}^n}\big(K(x,y)-K(x_0,y)\big)g(x)(1-\eta(y))\,dy\,dx$$

as $R \to \infty$ by the Lebesgue dominated convergence theorem. Using Definition 4.1.16, we obtain the validity of (4.3.5).

Next we observe that the functions $\phi(\cdot/R)$ are in L^2. We show that

$$\big\|T(\phi(\cdot/R))\big\|_{BMO} \le C_{n,\delta}(A+B_3) \tag{4.3.6}$$

uniformly in $R > 0$. Once (4.3.6) is established, then the sequence $\{T(\phi(\cdot/j))\}_{j=1}^\infty$ lies in a multiple of the unit ball of $BMO = (H^1)^*$, and by the Banach–Alaoglou theorem, there is a subsequence of the positive integers R_j such that $T(\phi(\cdot/R_j))$ converges weakly* to an element b in BMO. This means that

$$\big\langle T(\phi(\cdot/R_j)), g\big\rangle \to \big\langle b, g\big\rangle \tag{4.3.7}$$

as $j \to \infty$ for all $g \in \mathscr{D}_0$. Using (4.3.5), we conclude that $T(1)$ can be identified with the BMO function b, and as a consequence of (4.3.6) it satisfies

$$\big\|T(1)\big\|_{BMO} \le C_{n,\delta}(A+B_3)\,.$$

In a similar fashion, we identify $T^t(1)$ with a BMO function with norm satisfying

$$\big\|T^t(1)\big\|_{BMO} \le C_{n,\delta}(A+B_3)\,.$$

We return to the proof of (4.3.6). We fix a ball $B = B(x_0, r)$ with radius $r > 0$ centered at $x_0 \in \mathbf{R}^n$. If for all $R > 0$ we had a constant $c_{B,R}$ such that

$$\frac{1}{|B|}\int_B |T(\phi(\cdot/R))(x) - c_{B,R}|\,dx \le c_{n,\delta}\,(A + B_3), \tag{4.3.8}$$

~~for all $R > 0$,~~ then property (3) in Proposition 3.1.2 (adapted to balls) would yield (4.3.6). Obviously, (4.3.8) is a consequence of the two estimates

$$\frac{1}{|B|}\int_B |T\big[\phi(\tfrac{\cdot - x_0}{r})\phi(\tfrac{\cdot}{R})\big](x)|\,dx \le c_n B_3\,, \tag{4.3.9}$$

$$\frac{1}{|B|}\int_B \Big|T\big[(1-\phi(\tfrac{\cdot - x_0}{r}))\phi(\tfrac{\cdot}{R})\big](x) - T\big[(1-\phi(\tfrac{\cdot - x_0}{r}))\phi(\tfrac{\cdot}{R})\big](x_0)\Big|\,dx \le \frac{c_n}{\delta}A\,. \tag{4.3.10}$$

We bound the double integral in (4.3.10) by

$$\frac{1}{|B|}\int_B \int_{|y-x_0|\ge 2r} |K(x,y) - K(x_0,y)|\,\phi(y/R)\,dy\,dx\,, \tag{4.3.11}$$

since $1-\phi((y-x_0)/r) = 0$ when $|y-x_0| \le 2r$. Since $|x-x_0| \le r \le \frac{1}{2}|y-x_0|$, condition (4.1.2) gives that (4.3.10) holds with $c_n = \omega_{n-1} = |\mathbf{S}^{n-1}|$.

It remains to prove (4.3.9). It is easy to verify that there is a constant $C_0 = C_0(n,\phi)$ such that for $0 < \varepsilon \le 1$ and for all $a \in \mathbf{R}^n$ the functions

$$C_0^{-1}\phi(\varepsilon(x+a))\phi(x), \qquad C_0^{-1}\phi(x)\phi(-a+\varepsilon x) \tag{4.3.12}$$

are normalized bumps. The important observation is that with $a = x_0/r$ we have

$$\phi(\tfrac{x}{R})\phi(\tfrac{x-x_0}{r}) = r^n\tau^{x_0}\Big[\Big(\phi\big(\tfrac{r}{R}(\cdot + a)\big)\phi(\cdot)\Big)_r\Big](x) \tag{4.3.13}$$

$$= R^n\Big(\phi(\cdot)\phi\big(-a + \tfrac{R}{r}(\cdot)\big)\Big)_R(x), \tag{4.3.14}$$

and thus in either case $r \le R$ or $R \le r$, one may express the product $\phi(\frac{x}{R})\phi(\frac{x-x_0}{r})$ as a multiple of a translation of an L^1 dilation of a normalized bump.

Let us suppose that $r \le R$. In view of (4.3.13) we write

$$T\big[\phi(\tfrac{\cdot - x_0}{r})\phi(\tfrac{\cdot}{R})\big](x) = C_0\,r^n T\big[\tau^{x_0}\varphi_r\big](x)$$

for some normalized bump φ. Using this fact and the Cauchy–Schwarz inequality, we estimate the expression on the left in (4.3.9) by

$$\frac{C_0\,r^{n/2}}{|B|^{\frac{1}{2}}}r^{n/2}\left(\int_B |T\big[\tau^{x_0}\varphi_r\big](x)|^2\,dx\right)^{\frac{1}{2}} \le \frac{C_0\,r^{n/2}}{|B|^{\frac{1}{2}}}B_3 = c_n B_3\,,$$

where the first inequality follows by applying hypothesis (iii).

We now consider the case $R \le r$. In view of (4.3.14) we write

$$T\big[\phi(\tfrac{\cdot - x_0}{r})\phi(\tfrac{\cdot}{R})\big](x) = C_0 R^n T\big(\varphi_R\big)(x)$$

for some other normalized bump φ. Using this fact and the Cauchy–Schwarz inequality, we estimate the expression on the left in (4.3.9) by

$$\frac{C_0 R^{n/2}}{|B|^{\frac{1}{2}}} R^{n/2} \left(\int_B |T(\varphi_R)(x)|^2 \, dx \right)^{\frac{1}{2}} \le \frac{C_0 R^{n/2}}{|B|^{\frac{1}{2}}} B_3 \le c_n B_3$$

applying hypothesis (iii) and recalling that $R \le r$. This proves (4.3.9).

To finish the proof of (iv), we need to prove that T satisfies the weak boundedness property. But this is elementary, since for all normalized bumps φ and ψ and all $x \in \mathbf{R}^n$ and $R > 0$ we have

$$\begin{aligned} \big|\big\langle T(\tau^x \psi_R), \tau^x \varphi_R \big\rangle\big| &\le \big\|T(\tau^x \psi_R)\big\|_{L^2} \big\|\tau^x \varphi_R\big\|_{L^2} \\ &\le B_3 R^{-\frac{n}{2}} \big\|\tau^x \varphi_R\big\|_{L^2} \\ &\le C_n B_3 R^{-n}. \end{aligned}$$

This gives $\|T\|_{WB} \le C_n B_3$, which implies the estimate $B_4 \le C_{n,\delta}(A + B_3)$ and concludes the proof of the fact that condition (iii) implies (iv).

(iv) $\implies$ (L^2 boundedness of T)

We now assume condition (iv) and we present the most important step of the proof, establishing the fact that T has an extension that maps $L^2(\mathbf{R}^n)$ to itself. The assumption that the distributions $T(1)$ and $T^t(1)$ coincide with BMO functions leads to the construction of Carleson measures that provide the key tool in the boundedness of T.

We pick a smooth radial function Φ with compact support that is supported in the ball $B(0, \frac{1}{2})$ and that satisfies $\int_{\mathbf{R}^n} \Phi(x)\, dx = 1$. For $t > 0$ we define $\Phi_t(x) = t^{-n}\Phi(\frac{x}{t})$. Since Φ is a radial function, the operator

$$P_t(f) = f * \Phi_t \tag{4.3.15}$$

is self-transpose. The operator P_t is a continuous analogue of $S_j = \sum_{k \le j} \Delta_k$, where the Δ_j's are the Littlewood–Paley operators.

We now fix a Schwartz function f whose Fourier transform is supported away from a neighborhood of the origin. We discuss an integral representation for $T(f)$. We begin with the facts, which can be found in Exercises 4.3.1 and 4.3.2, that

$$\begin{aligned} T(f) &= \lim_{s \to 0} P_s^2 T P_s^2(f), \\ 0 &= \lim_{s \to \infty} P_s^2 T P_s^2(f), \end{aligned}$$

where the first limit is in the topology of $\mathscr{S}'(\mathbf{R}^n)$ and the second one is in the topology of $\mathscr{S}'(\mathbf{R}^n)/\mathscr{P}(\mathbf{R}^n)$. Thus, with the use of the fundamental theorem of calculus and the product rule, we are able to write

$$\begin{aligned} T(f) &= \lim_{s\to 0} P_s^2 T P_s^2(f) - \lim_{s\to\infty} P_s^2 T P_s^2(f) \\ &= -\lim_{\varepsilon\to 0}\int_\varepsilon^{\frac{1}{\varepsilon}} s\frac{d}{ds}\big(P_s^2 T P_s^2\big)(f)\,\frac{ds}{s} \\ &= -\lim_{\varepsilon\to 0}\int_\varepsilon^{\frac{1}{\varepsilon}} \left[s\left(\frac{d}{ds}P_s^2\right)TP_s^2(f) + P_s^2\left(Ts\frac{d}{ds}P_s^2\right)(f)\right]\frac{ds}{s}\,, \end{aligned} \tag{4.3.16}$$

where the limit is in the sense of $\mathscr{S}'(\mathbf{R}^n)/\mathscr{P}(\mathbf{R}^n)$. For a Schwartz function g we have

$$\begin{aligned} \left(s\frac{d}{ds}P_s^2(g)\right)^{\widehat{}}(\xi) &= \widehat{g}(\xi)\,s\frac{d}{ds}\widehat{\Phi}(s\xi)^2 \\ &= \widehat{g}(\xi)\,\widehat{\Phi}(s\xi)\,\big(2s\xi\cdot\nabla\widehat{\Phi}(s\xi)\big) \\ &= \widehat{g}(\xi)\sum_{k=1}^n \widehat{\Psi_k}(s\xi)\widehat{\Theta_k}(s\xi) \\ &= \sum_{k=1}^n\Big(\widetilde{Q}_{k,s}Q_{k,s}(g)\Big)^{\widehat{}}(\xi) = \sum_{k=1}^n\Big(Q_{k,s}\widetilde{Q}_{k,s}(g)\Big)^{\widehat{}}(\xi)\,, \end{aligned}$$

where for $1\le k\le n$, $\widehat{\Psi_k}(\xi) = 2\xi_k\widehat{\Phi}(\xi)$, $\widehat{\Theta_k}(\xi) = \partial_k\widehat{\Phi}(\xi)$, and $Q_{k,s}$, $\widetilde{Q}_{k,s}$ are operators defined by

$$Q_{k,s}(g) = g*(\Psi_k)_s\,, \qquad \widetilde{Q}_{k,s}(g) = g*(\Theta_k)_s\,;$$

here $(\Theta_k)_s(x) = s^{-n}\Theta_k(s^{-1}x)$ and $(\Psi_k)_s$ are defined similarly. Observe that Ψ_k and Θ_k are smooth odd bumps supported in $B(0,\frac{1}{2})$ and have integral zero. Since Ψ_k and Θ_k are odd, they are anti-self-transpose, meaning that $(Q_{k,s})^t = -Q_{k,s}$ and $(\widetilde{Q}_{k,s})^t = -\widetilde{Q}_{k,s}$. We now write the expression in (4.3.16) as

$$-\lim_{\varepsilon\to 0}\sum_{k=1}^n\left[\int_\varepsilon^{\frac{1}{\varepsilon}}\widetilde{Q}_{k,s}Q_{k,s}TP_sP_s(f)\frac{ds}{s} + \int_\varepsilon^{\frac{1}{\varepsilon}}P_sP_sTQ_{k,s}\widetilde{Q}_{k,s}(f)\frac{ds}{s}\right], \tag{4.3.17}$$

where the limit is in the sense of $\mathscr{S}'(\mathbf{R}^n)/\mathscr{P}(\mathbf{R}^n)$. We set

$$T_{k,s} = Q_{k,s}TP_s\,,$$

and we observe that the operator $P_sTQ_{k,s}$ is equal to $-((T^t)_{k,s})^t$.

Recall the notation $\tau^x h(z) = h(z-x)$. For a given $\varphi\in\mathscr{S}(\mathbf{R}^n)$ we have

$$\begin{aligned} Q_{k,s}TP_s(\varphi)(x) &= -\Big\langle TP_s(\varphi)\,,\,\tau^x(\Psi_k)_s\Big\rangle \\ &= -\Big\langle T\big(\Phi_s*\varphi\big)\,,\,\tau^x(\Psi_k)_s\Big\rangle \end{aligned}$$

$$= -\Big\langle T\Big(\int_{\mathbf{R}^n} \varphi(y)\,(\tau^y \Phi_s)\,dy\Big)\,,\, \tau^x(\Psi_k)_s\Big\rangle$$

$$= -\int_{\mathbf{R}^n} \big\langle T\big(\tau^y \Phi_s\big)\,,\, \tau^x(\Psi_k)_s\big\rangle \varphi(y)\,dy\,. \tag{4.3.18}$$

The last equality is justified by the convergence of the Riemann sums R_N of the integral $I = \int_{\mathbf{R}^n} \varphi(y)\,(\tau^y \Phi_s)(\cdot)\,dy$ to itself in the topology of $\mathscr{S}$ (this is contained in the proof of Theorem 2.3.20 in [156]); by the continuity of T, $T(R_N)$ converges to $T(I)$ in $\mathscr{S}'$ and thus $\langle T(R_N), \tau^x(\Psi_k)_s\rangle$ converges to $\langle T(I), \tau^x(\Psi_k)_s\rangle$. But $\langle T(R_N), \tau^x(\Psi_k)_s\rangle$ is also a Riemann sum for the rapidly convergent integral in (4.3.18); hence it converges to it as well.

We deduce that the operator $T_{k,s} = Q_{k,s} T P_s$ has kernel

$$K_{k,s}(x,y) = -\big\langle T(\tau^y \Phi_s), \tau^x(\Psi_k)_s\big\rangle = -\big\langle T^t(\tau^x(\Psi_k)_s), \tau^y \Phi_s\big\rangle\,. \tag{4.3.19}$$

Hence, the operator $P_s T Q_{k,s} = -((T^t)_{k,s})^t$ has kernel

$$\big\langle T^t(\tau^x \Phi_s), \tau^y(\Psi_k)_s\big\rangle = \big\langle T(\tau^y(\Psi_k)_s), \tau^x \Phi_s\big\rangle\,.$$

For $1 \le k \le n$ we need the following facts regarding these kernels:

$$\big|\big\langle T(\tau^y(\Psi_k)_s), \tau^x \Phi_s\big\rangle\big| \;\le\; C_{n,\delta}\big(\|T\|_{WB} + A\big)\, p_s(x-y)\,, \tag{4.3.20}$$

$$\big|\big\langle T^t(\tau^x(\Psi_k)_s), \tau^y \Phi_s\big\rangle\big| \;\le\; C_{n,\delta}\big(\|T\|_{WB} + A\big)\, p_s(x-y)\,, \tag{4.3.21}$$

where

$$p_t(u) = \frac{1}{t^n}\frac{1}{(1+|\frac{u}{t}|)^{n+\delta}}$$

is the L^1 dilation of the function $p(u) = (1+|u|)^{-n-\delta}$.

To prove (4.3.21), we consider the following two cases: If $|x-y| \le 5s$, then the weak boundedness property gives

$$\big|\big\langle T(\tau^y \Phi_s), \tau^x(\Psi_k)_s\big\rangle\big| = \big|\big\langle T(\tau^x((\tau^{\frac{y-x}{s}} \Phi)_s)), \tau^x(\Psi_k)_s\big\rangle\big| \le \frac{C_{n,\Phi}\|T\|_{WB}}{s^n}\,,$$

since both Ψ_k and $\tau^{\frac{y-x}{s}}\Phi$ are multiples of normalized bumps. Notice here that both of these functions are supported in $B(0,10)$, since $\frac{1}{s}|x-y| \le 5$. This estimate proves (4.3.21) when $|x-y| \le 5s$.

We now turn to the case $|x-y| \ge 5s$. Then the functions $\tau^y \Phi_s$ and $\tau^x(\Psi_k)_s$ have disjoint supports and so we have the integral representation

$$\big\langle T^t(\tau^x(\Psi_k)_s), \tau^y \Phi_s\big\rangle = \int_{\mathbf{R}^n}\int_{\mathbf{R}^n} \Phi_s(v-y) K(u,v) (\Psi_k)_s(u-x)\,du\,dv\,.$$

Using that Ψ_k has mean value zero, we can write the previous expression as

$$\int_{\mathbf{R}^n}\int_{\mathbf{R}^n} \Phi_s(v-y)\big(K(u,v) - K(x,v)\big)(\Psi_k)_s(u-x)\,du\,dv\,.$$

We observe that $|u-x| \le s$ and $|v-y| \le s$ in the preceding double integral. Since $|x-y| \ge 5s$, this makes $|u-v| \ge |x-y| - 2s \ge 3s$, which implies that $|u-x| \le \frac{1}{2}|u-v|$. Using (4.1.2), we obtain

$$|K(u,v) - K(x,v)| \le \frac{A|x-u|^\delta}{(|u-v|+|x-v|)^{n+\delta}} \le C_{n,\delta} A \frac{s^\delta}{|x-y|^{n+\delta}}\,,$$

where we used the fact that $|u-v| \approx |x-y|$. Inserting this estimate in the double integral, we obtain (4.3.21). Estimate (4.3.20) is proved similarly.

At this point we drop the dependence of $Q_{k,s}$ and $\widetilde{Q}_{k,s}$ on the index k, since we can concentrate on one term of the sum in (4.3.17). We have managed to express $-T(f)$ as a finite sum of operators of the form

$$\int_0^\infty \widetilde{Q}_s T_s P_s(f)\,\frac{ds}{s} \tag{4.3.22}$$

and of the form

$$\int_0^\infty P_s T_s \widetilde{Q}_s(f)\,\frac{ds}{s}\,, \tag{4.3.23}$$

where the preceding integrals converge in $\mathscr{S}'(\mathbf{R}^n)/\mathscr{P}(\mathbf{R}^n)$ and the T_s's have kernels $K_s(x,y)$, which are pointwise dominated by a constant multiple of $(A+B_4)p_s(x-y)$.

It suffices to obtain L^2 bounds for an operator of the form (4.3.22) with constant at most a multiple of $A+B_4$. Then by duality the same estimate also holds for the operators of the form (4.3.23). We make one more observation. Using (4.3.19) (recall that we have dropped the indices k), we obtain

$$T_s(1)(x) = \int_{\mathbf{R}^n} K_s(x,y)\,dy = \cancel{\left\langle T^t(\tau^x \Psi_s), 1\right\rangle =} (\Psi_s * T(1))(x)\,, \tag{4.3.24}$$

where all integrals converge absolutely.

We can therefore concentrate on the L^2 boundedness of the operator in (4.3.22). We pair this operator with a Schwartz function g in $\mathscr{S}_0(\mathbf{R}^n)$ and we use the convergence of the integral in $\mathscr{S}'/\mathscr{P}(\mathbf{R}^n)$ and the property $(\widetilde{Q}_s)^t = -\widetilde{Q}_s$ to obtain

$$\left\langle \int_0^\infty \widetilde{Q}_s T_s P_s(f)\,\frac{ds}{s}, g\right\rangle = \int_0^\infty \left\langle \widetilde{Q}_s T_s P_s(f), g\right\rangle \frac{ds}{s} = -\int_0^\infty \left\langle T_s P_s(f), \widetilde{Q}_s(g)\right\rangle \frac{ds}{s}\,.$$

The intuition here is as follows: T_s is an averaging operator at scale s and $P_s(f)$ is essentially constant on that scale. Therefore, the expression $T_sP_s(f)$ must look like $T_s(1)P_s(f)$. To be precise, we introduce this term and try to estimate the error that occurs. We have

$$T_s P_s(f) = T_s(1)P_s(f) + \left[T_s P_s(f) - T_s(1)P_s(f)\right]. \tag{4.3.25}$$