— Pour voir que $(3) \iff (4)$, il suffit de remarquer que

$$\begin{aligned}
&\sup_{O \text{ ouvert}} \left(\mu(O) - \underline{\lim}_{n\to+\infty} \mu_n(O)\right) \\
= &\sup_{F \text{ fermé}} \left(\mu(F^c) - \underline{\lim}_{n\to+\infty} \mu_n(F^c)\right) \\
= &\sup_{F \text{ fermé}} \left(1-\mu(F) - \underline{\lim}_{n\to+\infty} (1-\mu_n(F))\right) \\
= &\sup_{F \text{ fermé}} \left(-\mu(F) - \underline{\lim}_{n\to+\infty} -\mu_n(F)\right) \\
= &\sup_{F \text{ fermé}} \left(-\mu(F) + \overline{\lim}_{n\to+\infty} \mu_n(F)\right).
\end{aligned}$$

Ainsi, si l'un des supremum est négatif, alors l'autre l'est aussi.

— Preuve de $(2) \Longrightarrow (3)$. Soit F un fermé de $\mathbb{R}^d$. Pour tout $x \in \mathbb{R}^d$, on pose $d_F(x) = d(x,F) = \inf(\|y-x\| : y \in F)$ et, pour $\varepsilon > 0$, G_ε est la fonction continue définie sur $\mathbb{R}$ par $G_\varepsilon(x) = \left(1 - \frac{|x|}{\varepsilon}\right)^+$. On a $G_\varepsilon \circ d_F \ge \mathbb{1}_F$, donc

$$\overline{\lim}_{n\to+\infty} \int G_\varepsilon \circ d_F \, d\mu_n \ge \overline{\lim}_{n\to+\infty} \int \mathbb{1}_F \, d\mu_n.$$

Comme $G_\varepsilon \circ d_F$ est uniformément continue (c'est la composée d'une application 1-lipschitzienne et d'une application $\frac{1}{\varepsilon}$-lipschitzienne), on a

$$\overline{\lim}_{n\to+\infty} \int G_\varepsilon \circ d_F \, d\mu_n = \int G_\varepsilon \circ d_F \, d\mu,$$

d'où

$$\forall \varepsilon > 0, \quad \int G_\varepsilon \circ d_F \, d\mu \ge \overline{\lim}_{n\to+\infty} \mu_n(F).$$

Or par définition de la mesure image, on a $\int G_\varepsilon \circ d_F \, d\mu = \int G_\varepsilon \, d\mu_{d_F}$. Lorsque ε tend vers 0, G_ε converge vers l'indicatrice de 0 et donc par convergence dominée, $\int G_\varepsilon \circ d_F \, d\mu = \int G_\varepsilon \, d\mu_{d_F}$ converge vers

$$\int \mathbb{1}_{\{0\}} \, d\mu_{d_F} = \mu_{d_F}(\{0\}) = \mu(d_F = 0) = \mu(F).$$

— Preuve de $(3,4) \Longrightarrow 5$. On a $\mathring{A} \subset A \subset \overline{A}$, d'où

$$\begin{aligned}
&\overline{\lim}_{n\to+\infty} \mu_n(A) \le \overline{\lim}_{n\to+\infty} \mu_n(\overline{A}) \le \mu(\overline{A}), \\
\text{et} \quad &\underline{\lim}_{n\to+\infty} \mu_n(A) \ge \underline{\lim}_{n\to+\infty} \mu_n(\mathring{A}) \ge \mu(\mathring{A}).
\end{aligned}$$

Par ces deux inégalités, on obtient

$$\mu(\mathring{A}) \leq \underline{\lim}_{n\to+\infty} \mu_n(A) \leq \overline{\lim}_{n\to+\infty} \mu_n(A) \leq \mu(\overline{A}).$$

Comme $\mu(\overline{A}) - \mu(\mathring{A}) = \mu(\partial A) = 0$, la suite $(\mu_n(A))_{n\geq 1}$ admet une limite supérieure qui coïncide avec sa limite inférieure. Elle converge donc vers $\mu(\overline{A}) = \mu(\mathring{A})$, c'est-à-dire vers $\mu(A)$, car $\mu(\mathring{A}) \leq \mu(A) \leq \mu(\overline{A})$.

— $(5) \Longrightarrow (6)$ est évident

— Preuve de $(6) \Longrightarrow (1)$. L'idée est d'approcher la fonction f par une somme d'indicatrices de pavés dont la frontière est de mesure nulle. On va commencer par exhiber un grand ensemble de mesure nulle.
Soit T une base de $\mathbb{R}$ vu comme un $\mathbb{Q}$-espace vectoriel telle que $1 \in T$. On peut toujours se ramener au cas où tous les éléments de la base sont dans $[0,1]$. Pour $t \in T$ et $k \in \{1,\dots,d\}$, notons

$$\mathcal{P}_t^k = \{x \in \mathbb{R}^d; x_k - t \in \mathbb{Q}\}.$$

$\mathcal{P}_t^k$ est une réunion d'hyperplans orthogonaux au k-ième vecteur de la base de $\mathbb{R}^d$. À k fixé, les ensembles $(\mathcal{P}_t^k)_{t\in[0,1[\cap T}$ sont disjoints. Comme μ est une probabilité, l'ensemble des $t \in [0,1[\cap T$ tels que $\mu(\mathcal{P}_t^k) > 0$ est au plus dénombrable. Comme T n'est pas dénombrable [3], il existe t_k tel que $\mu(\mathcal{P}_{t_k}^k) = 0$. Pour un entier p, on pose $\bar{B}_p = \{x \in \mathbb{R}^d; \|x-t\|_\infty \leq p\}$, où $t = (t_1,\dots,t_d)$. Soit $\varepsilon > 0$. On peut trouver p tel que $\mu(\bar{B}_p) \geq 1 - \frac{\varepsilon}{2}$. D'après ce qui précède, $\mu(\partial\bar{B}_p) = 0$. On peut donc trouver n_0 tel que $n \geq n_0$ entraîne $\mu_n(\bar{B}_p) \geq 1 - \varepsilon$. Or $\bar{B}_p$ est compact. On note ω_f le module de continuité de la restriction de f à $\bar{B}_p$, défini pour $\eta > 0$ par

$$\omega_f(\eta) = \sup(|f(x) - f(y)| : \ \|x-y\|_\infty \leq \eta, x, y \in \bar{B}_p).$$

Pour $N \geq 1$, on pose

$$f_N(x) = \mathbb{1}_{\bar{B}_p}(x)\ f\left(t_1 + \frac{\lceil N(x-t_1)\rceil}{N}, \dots, t_d + \frac{\lceil N(x-t_d)\rceil}{N}\right).$$

On a

$$|f_N(x) - f(x)| \leq \omega_f(1/N) + 2\|f\|_\infty\, \mathbb{1}_{\bar{B}_p^c}(x).$$

3. En effet, si on avait $T = \{t^i; i \in \mathbb{N}^*\}$, on aurait $\mathbb{R} = \bigcup_{n=1}^{+\infty} \{\sum_{i=1}^{n} \lambda_i t^i; (\lambda_1,\dots,\lambda_n) \in \mathbb{Q}^n\}$, ce qui contredirait le fait que $\mathbb{R}$ n'est pas dénombrable.