Definition. *A regularizing kernel (or* mollifier*) is a function* $\rho \in C_c^\infty(B, [0,\infty))$ *with* $\int_B \rho = 1$ *and* $\rho(-x) = \rho(x)$ *for every* $x \in \mathbb{R}^n$*.*

Definition. *Given* $\varepsilon \in (0,1)$*, and* $u \in L^1_{\text{loc}}(\mathbb{R}^n)$ *we define the* ε*-dilation of* ρ

$$\rho_\varepsilon(x) := \frac{1}{\varepsilon^n}\rho\Big(\frac{x}{\varepsilon}\Big)$$

and the ε*-regularization of* u *as the convolution between* u *and* ρ_ε*, that is*

$$u_\varepsilon(x) = (u * \rho_\varepsilon)(x) = \int_{\mathbb{R}^n} \rho_\varepsilon(x-y)\, u(y)\, dy.$$

Note that ρ_ε is a smooth approximation to a delta function, since $\rho_\varepsilon \in C_c^\infty(B_\varepsilon, [0,\infty))$ and $\int_{\mathbb{R}^n} \rho_\varepsilon(x)\, dx = 1$.

There are two important facts about regularization of functions.

Fact (1)**.** *The regularization* u_ε *is smooth. Since* ρ_ε *is smooth, we can compute*

$$\nabla u_\varepsilon(x) = \nabla(u * \rho_\varepsilon)(x) = \nabla_x \int_{\mathbb{R}^n} \rho_\varepsilon(x-y)\, u(y)\, dy = \int_{\mathbb{R}^n} \nabla\rho_\varepsilon(x-y)\, u(y)\, dy = (u * \nabla\rho_\varepsilon)(x).$$

Fact (2)**.** *If* $u \in C_c^0(\mathbb{R}^n)$ *then* $u_\varepsilon \to u$ *in* $C_c^0(\mathbb{R}^n)$*. This follows from the fact that* ρ_ε *has mass 1 and is supported in the ball* B_ε*.*

Now, we can extend this construction to vector-valued Radon measures.

We first define the convolutions $(\mu * \rho_\varepsilon) : \mathbb{R}^n \to \mathbb{R}^m$ as

$$(\mu * \rho_\varepsilon)(x) = \int_{\mathbb{R}^n} \rho_\varepsilon(x-y)\, d\mu(y).$$

We want to use this convolution to define the regularization μ_ε, but we cannot do this directly, because we essentially have a type error: $\mu * \rho_\varepsilon$ is a function but μ_ε is a measure. So, we do the expected thing: take μ_ε to be the measure with density function $\mu * \rho_\varepsilon$.

We can either do this by integrating over Borel sets E:

$$\mu_\varepsilon(E) := \int_E (\mu * \rho_\varepsilon)(x)\, dx,$$

or, equivalently, via the duality with C_c^0:

$$\langle \mu_\varepsilon, \varphi\rangle := \int_{\mathbb{R}^n} \varphi(x)\cdot(\mu * \rho_\varepsilon)(x)\, dx, \qquad \varphi \in C_c^0(\mathbb{R}^n; \mathbb{R}^m).$$

So, what about important facts (1) and (2)?

It is easily seen that we have an analogue of fact (1).

Fact. *μ_ε has a density function $(\mu * \rho_\varepsilon) \in C^\infty(\mathbb{R}^n, \mathbb{R}^m)$, and*

$$\nabla(\mu * \rho_\varepsilon)(x) = \big(\mu * (\nabla \rho_\varepsilon)\big)(x).$$

Fact (2) is covered by the following theorem.

Theorem (4.36)**.** *If μ is a $\mathbb{R}^m$-valued Radon measure on $\mathbb{R}^n$, then, as $\varepsilon \to 0^+$,*

$$\mu_\varepsilon \overset{*}{\rightharpoonup} \mu, \qquad |\mu_\varepsilon| \overset{*}{\rightharpoonup} |\mu|.$$

Moreover, if $I_\varepsilon(E) = \{x \in \mathbb{R}^n : \operatorname{dist}(x, E) < \varepsilon\}$, then for every Borel set $E \subset \mathbb{R}^n$

$$|\mu_\varepsilon|(E) \leq |\mu|(I_\varepsilon(E)).$$

Proof. The proof proceeds in three parts.

(1) $\mu_\varepsilon \overset{*}{\rightharpoonup} \mu$. By Fubini and symmetry of ρ_ε, we have

$$\begin{aligned}\int \varphi(x) \cdot d\mu_\varepsilon(x) &= \int \varphi(x) \cdot \left(\int \rho_\varepsilon(x-y)\, d\mu(y) \right) dx \\ &= \int \left(\int \rho_\varepsilon(y-x)\, \varphi(x)\, dx \right) d\mu(y) = \int \varphi_\varepsilon(y)\, d\mu(y).\end{aligned}$$

Then since $\varphi_\varepsilon \to \varphi$ in $C_c^0(\mathbb{R}^n)$, we have

$$\langle \varphi, \mu_\varepsilon \rangle = \langle \varphi_\varepsilon, \mu \rangle \to \langle \varphi, \mu \rangle$$

for all $\varphi \in C_c^0$, and so $\mu_\varepsilon \overset{*}{\rightharpoonup} \mu$.

(2) $|\mu_\varepsilon|(E) \leq |\mu|(I_\varepsilon(E))$. First, if A is an open set in $\mathbb{R}^n$, then

$$\begin{aligned}|\mu_\varepsilon|(A) &:= \sup_{|\varphi| \leq 1} \int_A \varphi(x) \cdot d\mu_\varepsilon(x) = \sup_{|\varphi| \leq 1} \int_A \varphi(x) \cdot \int \rho_\varepsilon(x-y)\, d\mu(y)\, dx \\ &= \int_A \sup_{|\varphi(x)| \leq 1} \varphi(x) \cdot \int \rho_\varepsilon(x-y)\, d\mu(y)\, dx = \int_A \left| \int \rho_\varepsilon(x-y)\, d\mu(y) \right| dx,\end{aligned}$$

since the supremum is almost achieved pointwise in x. (The convolution is smooth, the optimizer = its direction, direction function is smooth almost everywhere on its support.) If E is a Borel set in $\mathbb{R}^n$, then we obtain the same equality (replacing A with E) by infimizing over open sets $A \supset E$, noting that the integrand is C^∞.

Applying the polar decomposition $d\mu = g d|\mu|$, we then have

$$\begin{aligned}\left| \int \rho_\varepsilon(x-y)\, d\mu(y) \right| &= \left| \int \rho_\varepsilon(x-y)\, g(y)\, d|\mu|(y) \right| \leq \int \rho_\varepsilon(x-y)\, |g(y)|\, d|\mu|(y) \\ &= \int \rho_\varepsilon(x-y)\, d|\mu|(y).\end{aligned}$$

Then, we have

$$\begin{aligned}|\mu_\varepsilon|(E) &= \int_E \left| \int \rho_\varepsilon(x-y)\, d\mu(y) \right| dx \le \int_E \int \rho_\varepsilon(x-y)\, d|\mu|(y)\, dx \\ &= \int_E \int_{B(x,\varepsilon)} \rho_\varepsilon(x-y)\, d|\mu|(y)\, dx = \int_{I_\varepsilon(E)} \int_{B(y,\varepsilon)\cap E} \rho_\varepsilon(x-y)\, dx\, d|\mu|(y) \\ &\le \int_{I_\varepsilon(E)} \int_{\mathbb{R}^n} \rho_\varepsilon(x-y)\, dx\, d|\mu|(y) = |\mu|(I_\varepsilon(E)).\end{aligned}$$

The third step uses the fact that ρ_ε is supported on B_ε. The fourth step follows from Fubini, noting that $x \in E$ and $y \in B(x,\varepsilon)$ imply $\operatorname{dist}(y,E) < \varepsilon$, so $y \in I_\varepsilon(E)$. The final two steps enlarge the domain and use the fact that $\int \rho_\varepsilon = 1$. This gives the result.

(3) $|\mu_\varepsilon| \overset{*}{\rightharpoonup} |\mu|$. The proof is built on Exercise 4.31, which states that if $\mu_h \overset{*}{\rightharpoonup} \mu$ and there exists $r_k \to \infty$, $k \in \mathbb{N}$ such that

$$\lim_{h\to\infty} |\mu_h|(B_{r_k}) = |\mu|(B_{r_k}) \quad \forall k,$$

then $|\mu_h| \overset{*}{\rightharpoonup} |\mu|$. In other words, weak convergence of μ_h implies weak convergence of its total variation as long as the total variation is converging "correctly" on arbitrarily large sets.

As an example of what can happen without this condition, consider the high-frequency oscillations $\sin(hx)\, dx$: measure is converging to 0, but total variation is converging to $\frac{2}{\pi}\, dx$.

Thus, it suffices to find a sequence of radii $r_k \to \infty$ along which $|\mu_\varepsilon|(B_{r_k}) \to |\mu|(B_{r_k})$ as $\varepsilon \to 0$.

We pick our radii so that $|\mu|(\partial B_{r_k}) = 0$, which ensures that $\overline{B}_{r_k}$ and B_{r_k} have the same $|\mu|$-mass. This is possible due to Proposition 2.16, which ensures that at most countably many radii can satisfy $|\mu|(\partial B_r) > 0$, so such a sequence $r_k \to +\infty$ exists.

With this choice, we establish the chain:

$$|\mu|(B_{r_k}) \le \liminf_{\varepsilon\to 0} |\mu_\varepsilon|(B_{r_k}) \le \limsup_{\varepsilon\to 0} |\mu_\varepsilon|(B_{r_k}) \le \limsup_{\varepsilon\to 0} |\mu|(B_{r_k+\varepsilon}) = |\mu|(\overline{B}_{r_k}) = |\mu|(B_{r_k}).$$

The first inequality comes from Proposition 4.29. The second is trivial. The third applies the inequality from part (2). The fourth uses continuity of measure from above, and the fifth the special choice of radii.

Thus the limit $\lim_{\varepsilon\to 0} |\mu_\varepsilon|(B_{r_k})$ exists and is equal to $|\mu|(B_{r_k})$ for each k, and the claim follows. □