1. First note that for any $\mathcal{L}^1$ function ϕ and for any x,

$$\int r^{-n}\phi((x-y)/r)\,dy = \int \phi.$$

(This is true for indicator functions of measurable sets, hence for simple functions, hence for nonnegative measurable functions, hence for $\mathcal{L}^1$ functions.)

Thus

$$\begin{aligned}
|(\phi_r * f)(x) - f(x)| &= \left| r^{-n}\int \phi((x-y)/r) f(y)\,dy - f(x)\right| \\
&= \left| r^{-n}\int \phi((x-y)/r) f(y)\,dy - f(x)\, r^{-n}\int \phi((y-x)/r)\,dy\right| \\
&= r^{-n}\left|\int \phi((x-y)/r)(f(y)-f(x))\,dy\right| \\
&\le r^{-n}\int |\phi((x-y)/r)|\cdot|f(y)-f(x)|\,dy \\
&= r^{-n}\int_{\mathbf{B}(x,r)} |\phi((x-y)/r)|\cdot|f(y)-f(x)|\,dy
\end{aligned}$$

(since $\phi(z)=0$ for $|z|>1$)

$$(*)\qquad \le Kr^{-n}\int_{\mathbf{B}(x,r)} |f(y)-f(x)|\,dy$$

where K is a bound for $|\phi|$. By definition, this goes to 0 as $r\to 0$ if x is a Lebesgue point of f. Since almost every point is a Lebesgue point, we are done.

2. One method: let a be a Lebesgue point of 1_A and b be a Lebesgue point of B. Translating A and/or B does not affect the hypothesis or the conclusion, so we may assume that $a=b=0$.

Let $\frac{3}{4}<\theta<1$. Then for all sufficiently small $r>0$,

$$\begin{aligned}
\lambda(A\cap[-r,r]) &\ge 2r\theta, \\
\lambda(B\cap[-r,r]) &\ge 2r\theta.
\end{aligned}$$

Let $A' = A\cap[-r,r]$ and $B' = B\cap[-r,r]$, and let $t\in[0,r]$. Then $(A'+t)\cup B' \subset [-r,2r]$, so

$$\begin{aligned}
3r &\ge \lambda((A'+t)\cup B') \\
&= \lambda(A'+t) + \lambda(B') - \lambda((A'+t)\cap B') \\
&\ge 4r\theta - \lambda((A'+t)\cap B'),
\end{aligned}$$

so

$$\lambda((A'+t)\cap B') \ge (4\theta-3)r > 0$$

since $\theta>\frac{3}{4}$. Thus $(A'+t)\cap B'$ is nonempty or, equivalently $t\in B'-A'$. Thus $[0,r]\subset B'-A'\subset B-A$.

Another solution: We can assume that A and B are bounded. (Otherwise replace A and B by $A\cap[-n,n]$ and $B\cap[-n,n]$ where n is large enough that these sets have positive measure.) Thus 1_A and 1_B are in L^p for every p. In particular, they are both in L^2. Thus the function $1_A * 1_B$ is continuous. Note that

$$\begin{aligned}\int (1_A * 1_B)(x)dx &= \int_x \int_y 1_A(x-y)1_B(y)\,dy\,dx \\ &= \int_y \int_x 1_A(x-y)1_B(y)\,dx\,dy \\ &= \int_y 1_B(y) \int_x 1_A(x-y)\,dx\,dy \\ &= \int_y 1_B(y)\lambda(A)\,dy \\ &= \lambda(A)\lambda(B) \\ &> 0.\end{aligned}$$

Now $1_A * 1_B$ is a continuous function with positive integral, so there is an interval I on which is it is positive. Now for $x \in I$,

$$\begin{aligned} 0 &< (1_A * 1_B)(x) \\ &= \int_y 1_A(x-y)1_B(y)\,dy\,dx \\ &= \lambda\{y : x-y \in A \text{ and } y \in B\}\,dx \\ &= \lambda\{y : y \in x-A \text{ and } y \in B\}\,dx \\ &= \lambda((x-A)\cap B)\,dx.\end{aligned}$$

In particular, $(x-A)\cap B$ is non-empty, or, equivalently, $x \in A+B$. We have shown: there is an interval I such that $I \subset A+B$. Of course the same is true for the sets $-A := \{-a : a \in A\}$ and B.

3. Let $H_n(x) = \sum_{i>n} F_i(x)$. Note that $H_n(x) \to 0$ for all x. Since F, the F_n's and the H_n's are increasing, they are differentiable almost everywhere and have nonnegative derivatives. Since

$$F(x) = \sum_{i=1}^{n} F_i(x) + H_n(x),$$

we have

$$F'(x) = \sum_{i=1}^{n} F_i'(x) + H_n'(x) \tag{*}$$

for almost all x (indeed, for all x where these derivatives exist.) Thus it suffices to show that $H_n'(x) \to 0$ almost everywhere as $n \to \infty$.

By (*), $H_n'(x)$ is decreasing as a function of n, so $\lim_n H_n'$ exists. Thus

$$\begin{aligned}\int \lim_n H_n' &\le \liminf_n \int H_n' \qquad \text{(by Fatou)} \\ &\le \liminf(H_n(b) - H_n(a)) \\ &= 0.\end{aligned}$$

Thus $\lim_n H_n' = 0$ almost everywhere, so letting $n \to \infty$ in (*) gives the desired result.

4. Consider such a collection $\overline{\mathbf{B}}_i = \overline{\mathbf{B}}(x_i, r_i)$, $i = 1, \dots, n$. By translating and scaling, we can assume that $\mathbf{B}(x_1, r_1) = \mathbf{B}(0,1)$. We can also assume that $r_i = \max\{1/2, |x_i| - 1\}$ since if we replace each r_i by $\max\{1/2, |x_i| - 1\}$, the new balls still satisfy the hypotheses. We can assume that the balls are labelled so that

$$|x_1| \le |x_2| \le \dots |x_{k-1}| \le \frac{3}{2} < |x_k| \le \dots |x_n|.$$

Let $\lambda = \frac{3}{2|x_k|}$. Consider the balls

$$\overline{\mathbf{B}}_i(\lambda x_i, \lambda r_i), \quad i = k, \dots, n. \tag{*}$$

None of these balls contains the center of another, since they are obtained from the $\overline{\mathbf{B}}_i$, $i \geq k$, by scaling. Now consider the balls

$$\overline{\mathbf{B}}_i^* = \overline{\mathbf{B}}_i(\lambda_i x_i, |\lambda_i x_i| - 1), \quad i = k, \ldots, n.$$

None of these balls contains the center of another, since they are obtained from the balls (*) by decreasing the radii and keeping the centers the same. Also, if $j < k$, then $\overline{\mathbf{B}}_i^*$ cannot contain x_j since $\overline{\mathbf{B}}_i^* \subset \overline{\mathbf{B}}_i$. Since radius$(\overline{\mathbf{B}}_j) \leq$ radius$(\overline{\mathbf{B}}_i^*)$, it follows that $\overline{\mathbf{B}}_i$ cannot contain the center of $\overline{\mathbf{B}}_i^*$.

Thus the new family $\overline{\mathbf{B}}_1, \ldots, \overline{\mathbf{B}}_k, \overline{\mathbf{B}}_{k+1}^*, \ldots, \overline{\mathbf{B}}_n^*$ satisfies the hypotheses, and $(k+1)$ of the centers lie in $\overline{\mathbf{B}}(0, 3/2)$.

Repeating the process, we get a collection of n balls $\overline{\mathbf{B}}(y_i, 1/2)$ satisfying the hypotheses with all n balls having radius 1/2 and lying in $\overline{\mathbf{B}}(0, 2)$.

The balls $\mathbf{B}(y_i, 1/4)$ are disjoint balls in $\mathbf{B}(0, 2)$, so

$$(\dagger) \qquad \sum_{i=1}^{n} \lambda \mathbf{B}(y_i, 1/4) \leq \lambda \mathbf{B}(0, 2)$$

Now $\lambda \mathbf{B}(x, r) = \lambda \mathbf{B}(0, 1) r^d$, so by (†),

$$\frac{n}{4^d} \leq 2^d,$$

so $n \leq 8^d$.

(b). Let $\overline{\mathbf{B}}(x_1, r_1)$ be the largest ball in $\mathcal{F}$. (If there are several largest, choose one of them to be $\overline{\mathbf{B}}(x_1, r_1)$.) Now suppose we have chosen $\overline{\mathbf{B}}(x_i, r_i)$ for $i < n$. Among all of the balls in $\mathcal{F}$ with center not in $\cup_{i<n} \overline{\mathbf{B}}(x_i, r_i)$, let $\overline{\mathbf{B}}(x_n, r_n)$ be one of largest possible radius.

(The sequence stops at the nth step if $X \subset \cup_{i \leq n} \overline{\mathbf{B}}(x_i, r_i)$.)

By construction, $r_1 \geq r_2 \geq \ldots$. Also, for $i < j$, $x_j \notin \overline{\mathbf{B}}(x_i, r_i)$ (by choice of $\overline{\mathbf{B}}(x_j, r_j)$) and therefore $x_i \notin \overline{\mathbf{B}}(x_j, r_j)$ since $r_i \geq r_j$.

It remains to divide the $\overline{\mathbf{B}}_n$ into families $\mathcal{F}_i$, $i = 1, 2, \ldots, c(d)$. We do this inductively as follows. We assign $\overline{\mathbf{B}}_1$ to the family $\mathcal{F}_1$. Having assigned $\overline{\mathbf{B}}_1, \overline{\mathbf{B}}_2, \ldots, \overline{\mathbf{B}}_{n-1}$, note by part (a) that at most $c(d) - 1$ of the the balls $\overline{\mathbf{B}}_i$, $i < n$, can intersect $\overline{\mathbf{B}}_n$. In other words, if we let

$$I(n) = \{i : \text{for some } k < n,\ \overline{\mathbf{B}}_k \in \mathcal{F}_i \text{ and } \overline{\mathbf{B}}_k \cap \overline{\mathbf{B}}_n \neq \emptyset\},$$

then $I(n)$ has at most $c(d) - 1$ elements. Hence there is an $i \in \{1, 2, \ldots, c(d)\}$ not in $I(n)$. We choose such an i and assign $\overline{\mathbf{B}}_n$ to the family $\mathcal{F}_i$.

5. Let $x \in S$. Let $T = \{2x - y : y \in S\} = \{y : 2x - y \in S\}$. Note that if $y \neq x$, then the midpoint of y and $2x - y$ cannot both be in S since their midpoint is x. Thus $S \cap T = \{x\}$. Hence

$$\begin{aligned} 2r &\geq \lambda(S \cap T \cap [x - r, x + r]) \\ &= \lambda(S \cap [x - r, x + r]) + \lambda(T \cap [x - r, x + r]) - \lambda\{x\} \\ &= 2\lambda(S \cap [x - r, x + r]) \end{aligned}$$

Thus

$$\limsup_{r \to 0} \frac{1}{2r} \lambda(S \cap [x - r, x + r]) \leq \frac{1}{2}.$$

This holds for every $x \in S$. On the other hand,

$$\lim_{r \to 0} \frac{1}{2r} \lambda(S \cap [x - r, x + r]) = 1$$

for almost every $x \in S$. This gives a contradiction unless $\lambda(S) = 0$.

6. First consider the case when F is a simple function. Then we can write

$$F = \sum_{i=1}^{n} a_i 1_{A_i}$$

where $a_i > 0$ and the A_i are disjoint.

Then

$$\begin{aligned} \mu\{x : F(x) > y\} &= \sum_{\{i : a_i > y\}} \mu(A_i) \\ &= \sum_i 1_{\{a_i > y\}} \mu(A_i), \end{aligned}$$

so

$$\begin{aligned} \int_{y=0}^{\infty} \mu\{x : F(x) > y\}\, d\lambda(y) &= \sum_i \left(\int_{y=0}^{\infty} 1_{\{a_i > y\}}\, d\lambda(y) \right) \mu(A_i) \\ &= \sum_i a_i \mu(A_i) \\ &= \int F\, d\mu. \end{aligned}$$

Thus we have proved it when F is simple.

Now suppose $F : X \to [0, \infty)$ is $\mathcal{A}$-measurable. Recall that there are nonnegative simple $\mathcal{A}$-measurable functions F_n such that $F_1 \le F_2 \le \dots$ and such that $\lim_n F_n(x) = F(x)$ for all x.

Then

$$\{x : F_1(x) > y\} \subset \{x : F_2(x) > y\} \subset \dots$$

and

$$\cup_n \{x : F_n(x) > y\} = \{x : F(x) > y\},$$

so

$$\mu\{x : F_n(x) > y\} \uparrow \mu\{x : F(x) > y\}. \tag{1}$$

Since the result is true for simple functions,

$$\int \mu\{x : F_n(x) > y\}\, d\lambda(y) = \int F_n\, d\mu \tag{2}$$

for each n. By (1), (2), and the Monotone Convergence Theorem,

$$\int \mu\{x : F(x) > y\}\, d\lambda(y) = \int F\, d\mu.$$