Abstract. A quick review of BV functions from the point of view of complex Borel measures and the Riesz representation theorem.

1. BV Functions

We denote by an interval any set $I \subseteq \mathbb{R}$ with the property that for any $x, y \in I$ and $x < z < y$ implies $z \in I$. We always exclude the case of a single point $I = \{x\}$. For a function $F : I \to \mathbb{C}$ we define the (possibly infinite) quantity:

$$\|F\|_{BV(I)} = \sup_{x_0 < x_1 < \ldots < x_N} \sum_{k=1}^{N} |F(x_k) - F(x_{k-1})|, \quad \text{where } x_k \in I. $$

If $I = \bigcup_i I_i$ is a disjoint union (of possibly infinitely many terms), then its easy to see this quantity adds:

$$\|F\|_{BV(I)} = \sum_i \|F\|_{BV(I_i)}, \quad \text{where } I = \bigcup_i I_i \text{ is a disjoint union}. $$

Note that in the case where $I = \bigcup_{i=1}^N I_i$ this is immediate from the definition because any partition of $I$ produces partitions of the $I_i$ and vice versa. In the case of infinite unions a similar reason shows $\sum_{i=1}^N \|F\|_{BV(I_i)} \leq \|F\|_{BV(I)}$ for each $N$ which is the main direction to prove in this case.

We say $F \in BV(I)$ if $\|F\|_{BV(I)} < \infty$. Note that this becomes a seminorm on the subset of functions $F : I \to \mathbb{C}$ for which this finiteness condition holds. One nice thing about BV functions is that they can’t have many discontinuities, in fact:

Lemma 1.1. Let $F \in BV(I)$. Then the set of discontinuities is a countable subset of $I$. In fact if one defines:

$$\text{jump}(F) = \{x \in I \mid F(x^-) \neq F(x^+)\},$$

then $\text{jump}(F)$ is at most countably infinite and one has:

$$\sum_{x \in \text{jump}(F)} \left(\|F(x) - F(x^-)\| + \|F(x) - F(x^+)\|\right) \leq \|F\|_{BV(I)}. $$

Proof. The key observation here is that if $x \in I$ is not an an endpoint then for $[x-h,x) \subseteq I$ where $h > 0$ one has $\lim_{h \to 0^+} \|F\|_{BV([x-h,x))} = 0$. This comes by writing $[x - \frac{1}{k}, x) = \bigcup_{n \geq k} [x - \frac{1}{n}, x - \frac{1}{n+1})$ for some $k$ large enough that $[x - \frac{1}{k}, x) \subseteq I$ to begin with, and then using (1) which says:

$$\|F\|_{BV([x-h,x))} = \sum_{n=N}^{\infty} \|F\|_{BV([x - \frac{1}{n}, x - \frac{1}{n+1})}) = o_N(1),$$
	hanks{thanks to the fact that the entire sum starting at $n = k$ is finite (so it has small tail).}

The limit $\lim_{h \to 0^+} \|F\|_{BV([x-h,x))} = 0$ implies that $F(x_n)$ is Cauchy for any $x_n \nearrow x$, so $F(x_n)$ has some limit. This also implies that all such limits must be the same because if $y_n \nearrow x$ the interleaved sequence $\{x_1, y_1, x_2, y_2, \ldots\}$ is also Cauchy. A similar argument shows $F(x^-)$ exists as well.

Finally let $x_i \in \text{jump}(F)$ for $i = 1, \ldots, N$ be a finite collection and suppose $x_1 < \ldots < x_N$. Choosing $\epsilon \leq \frac{1}{2} \min_{i \neq j} \{|x_i - x_j|\}$ we have that $(x_i - \epsilon, x_i + \epsilon)$ are each disjoint (and contained in $I$ for small enough $\epsilon$). Then:

$$\sum_{i=1}^{N} \left(\|F(x) - F(x_i - \epsilon)\| + \|F(x) - F(x_i + \epsilon)\|\right) \leq \|F\|_{BV(I)}. $$


Taking the limit as $\epsilon \to 0$ gives (2) for all finite subsets of $jump(F)$. This shows $jump(F)$ must be countable, and that in fact the sum over all jumps must converge to a value $\leq \|F\|_{BV(I)}$.

One thing we'll see in a moment is that $jump(F)$ contains essential information about $F$, while the precise values of $F$ can be irrelevant aside from giving a spuriously large value for $\|F\|_{BV(I)}$. This is illustrated by the example of letting $F(x) \equiv c$ for some constant, except for $x$ at finitely many points $x_1, \ldots, x_N$. In this case $jump(F) = \emptyset$ even though $\|F\|_{BV} = \sum_{i=1}^{N} |F(x_i)|$ can be quite large. But for all intents and purposes a function like this should be thought of as a constant (again we’ll make the idea here more precise in a bit).

2. CDF of Radon Measures

The main way to get $BV(\mathbb{R})$ functions is through the following construction: Let $\mu \in \mathcal{M}(\mathbb{R})$ be a complex Borel measure (in particular if $\mu$ is nonnegative it is still finite). Also set $BV_0(\mathbb{R})$ to be all $F \in BV(\mathbb{R})$ with the property that $\lim_{x \to -\infty} F(x) = 0$ (note that infinite limits of $BV(\mathbb{R})$ functions always exist by an argument that is similar to the one in the last section for left and right limits). We define the quantity:

$$F(x) = \mu(((-\infty, x])$$

which is sometimes called the right continuous cumulative distribution of $\mu$. The name comes from the fact that if $x_n \searrow x$ then since $(-\infty, x_n]$ are nested decreasing with intersection $(-\infty, x]$ one gets $\lim_n F(x_n) = F(x)$ thanks to continuity of measures for intersections. Thus $F(x) = F(x^+)$ for all $x$. On the other thanks to continuity of measures for unions one gets $F(x^-) = \mu((-\infty, x]) = \mu((-\infty, x]) - \mu(\{x\}) = F(x^+)-\mu(\{x\})$. This leads to:

**Proposition 2.1.** Let $\mu \in \mathcal{M}(\mathbb{R})$ and $F(x)$ its right continuous CDF. Then one has $F \in BV_0(\mathbb{R})$ and also:

$$\mu = \text{atoms}(\mu)$$

where atoms($\mu$) = $\{x \in \mathbb{R} \mid \mu(\{x\}) \neq 0\}$. Moreover one has the identity:

$$\|F\|_{BV(\mathbb{R})} = \|\mu\|_{\mathcal{M}(\mathbb{R})}.$$  

*Proof.* The statement about the atoms comes from the identity $F(x^+)-F(x^-) = \mu(\{x\})$ which we just proved. To get the statement about $BV(\mathbb{R})$ note that for for any partition $x_0 < x_1 < \ldots < x_N$ we have:

$$\sum_{i=1}^{N} |F(x_i) - F(x_{i-1})| \leq \sum_{i=1}^{N} |\mu|(x_{i-1}, x_i] \leq |\mu|(x_0, x_N] \leq |\mu|(\mathbb{R}) = \|\mu\|_{\mathcal{M}(\mathbb{R})}.$$  

To get the other direction of (3) is just a little more work because we need to use the regularity properties of Radon measures. Recall that the definition of the total variation of $\mu$ is the positive measure defined by:

$$|\mu|(E) = \sup \left\{ \sum_{i=1}^{N} |\mu(E_i)| \mid \cup_i E_i \subseteq E \text{ is a disjoint union} \right\}.$$  

For each finite disjoint collection of sets $E_i$ we can find compact and open sets $K_i \subseteq E_i$ with the property that $|\mu|(E_i) < |\mu(K_i)| + \epsilon 2^{-i}$, and for each of these $K_i$ we can find *finitely* many disjoint open intervals $I_{i,j} = (a_{i,j}, b_{i,j})$ with $K_i \subseteq \cup_j I_{i,j}$ and $\sum_j |\mu|(I_{i,j}) < |\mu|(K_i) + \epsilon 2^{-i}$. In fact we can do a little bit better because each $K_i$ is compact, so it must be bounded away from the right endpoint $b_{i,j}$ of each of these intervals. Using this and monotonicity we can find a disjoint collection of intervals of the form $J_{i,j} = (a_{i,j}, c_{i,j})$ with $K_i \subseteq \cup_j J_{i,j}$ and $\sum_j |\mu|(J_{i,j}) < |\mu|(K_i) + \epsilon 2^{-i}$. By the containments of everything in this construction we have:

$$\mu(E_i) = \mu(K_i) + \mu(E_i \setminus K_i) = \mu(\cup_j J_{i,j}) - \mu(\cup_j J_{i,j} \setminus K_i) + \mu(E_i \setminus K_i).$$  

Now we also have $\mu(\cup_j J_{i,j}) = \sum_j (F(c_{i,j}) - F(a_{i,j}))$, so that:

$$|\mu(E_i)| \leq \sum_j |F(c_{i,j}) - F(a_{i,j})| + \epsilon 2^{-i}.$$  

Now we are in business because for this collection of $E_i$ and disjoint half closed intervals $J_{i,j}$ we compute:

$$\sum_i |\mu(E_i)| \leq \sum_{i,j} |F(c_{i,j}) - F(a_{i,j})| + \epsilon \leq \|F\|_{BV(\mathbb{R})} + \epsilon.$$  

2
Since the $E_i$ were an arbitrary set of finite disjoint intervals in $\mathbb{R}$ we get $\|\mu\|_{\mathcal{M}(\mathbb{R})} \leq \|F\|_{\text{BV}(\mathbb{R})}$ as was to be shown.

3. DISTRIBUTIONAL DERIVATIVES OF BV FUNCTIONS

Next we give a Reisz representation version of the equivalence between $\mathcal{M}(\mathbb{R})$ and $\text{BV}_0(\mathbb{R})$. Doing this will also explain why we don’t care about discontinuities outside jump$(F)$. Let $F \in \text{BV}(\mathbb{R})$ (not necessarily vanishing at $-\infty$ or right continuous, just in $\text{BV}$). From this we can define a linear functional on the vector space $C^1_c(\mathbb{R}) = \{f \in C_c(\mathbb{R}) \mid f' \in C_c(\mathbb{R})\}$:

$$L(\varphi) = -\int_{-\infty}^{\infty} F(x)\varphi'(x)dx .$$

Note that this is completely classical as the integral can be defined in the sense of Riemann. What’s interesting is that $L$ can be extended continuously to the much larger space $\mathcal{C}_0(\mathbb{R})$ because of the condition $F \in \text{BV}(\mathbb{R})$. To see this let $\Delta_h \varphi(x) = h^{-1}(\varphi(x+h) - \varphi(x))$ be the difference quotient. Then $\varphi' = \lim_{h \to 0} \Delta_h \varphi$, and one has the “integration by parts” formula:

$$-\int_{-\infty}^{\infty} F\Delta_h \varphi dx = \int_{-\infty}^{\infty} \Delta_{-h} F \varphi dx , \quad h \neq 0 .$$

Now $\lim_{h \to 0} \Delta_h F$ can in general be quite singular so we don’t expect to get useful information by looking that pointwise. But what about its averages against $\mathcal{C}_0(X)$? First assuming $\varphi$ is compactly supported we can discretize the integral and for a fixed $h > 0$:

$$|\int_{-\infty}^{\infty} \Delta_{-h} F \varphi dx| \leq \sum_{n \in \mathbb{Z}} \frac{1}{h} \int_{hn}^{h(n+1)} |F(x+h) - F(x)| |\varphi(x)| dx \leq \|\varphi\|_{L^\infty(\mathbb{R})} \sum_{n \in \mathbb{Z}} \sup_{x \in [hn,h(n+1)]} |F(x+h) - F(x)| .$$

On the other hand $\sup_{x \in [hn,h(n+1)]} |F(x+h) - F(x)| \leq \|F\|_{\text{BV}([hn,h(n+2)])}$. Therefore using (1) the second term can be bounded by:

$$\sum_{n \in \mathbb{Z}} \sup_{x \in [hn,h(n+1)]} |F(x+h) - F(x)| \leq 2 \|F\|_{\text{BV}(\mathbb{R})} .$$

In other words we have the uniform bound:

$$|L_h(\varphi)| \leq 2 \|F\|_{\text{BV}(\mathbb{R})} \|\varphi\|_{\mathcal{C}_0(\mathbb{R})} , \quad \text{where} \quad L_h(\varphi) = \int_{-\infty}^{\infty} \Delta_{-h} F \varphi dx .$$

This shows that not only can we extend to all of $\mathcal{C}_0(\mathbb{R})$, but also that there holds a uniform bound $\|L_h\|_{\mathcal{C}_0(\mathbb{R})} \leq 2 \|F\|_{\text{BV}(\mathbb{R})}$. By weak-* compactness this also means that there exists a subsequence $L_{h_n} \to L_0$ for some $L_0 \in \mathcal{C}_0^*(\mathbb{R})$. But we already know that $L_{h_n}(\varphi) \to L(\varphi)$ where $L$ is defined by (4) above as long as $\varphi \in C^1_c(\mathbb{R})$. In fact we can say a little bit more because if $h_n \to 0$ is a any sequence ($h_n \neq 0$), then there is a further weak-* convergent subsequence $L_{h_{n_k}} \to L$. This shows that in fact we have $\lim_{h \to 0} L_h \to L$ for any sequence of $h \to 0$. Thus $L \in \mathcal{C}_0^*(\mathbb{R})$ (after extension) so there must exist some $\mu \in \mathcal{M}(\mathbb{R})$ with:

$$\int_{\mathbb{R}} \varphi d\mu = -\int_{-\infty}^{\infty} F(x)\varphi'(x)dx , \quad \text{for all} \quad \varphi \in C^1_c(\mathbb{R}) .$$

Because of this we write $\mu = \lim_{h \to 0} \Delta_h F$ “weakly in the sense of measures”. The key point here is that the measure $\mu$ captures all the information from the singular limit of $\Delta_h F$ that could be lost by considering this limit pointwise.

It’s also important to notice something at this point: If we change $F$ on a finite (or even countable) set of points in such a way that its still in $\text{BV}(\mathbb{R})$, then the RHS of formula (5) does not change. Therefore neither does the LHS. Thus, going from $\text{BV}(\mathbb{R}) \Rightarrow \mathcal{M}(\mathbb{R})$ looses some information, but as we’ll show next this can only be spurious discontinuities like in the example of the first section.

**Theorem 3.1** (Fundamental Theorem of Calculus for BV Functions). Let $G \in \text{BV}(\mathbb{R})$, and let $G' = \mu$ be its weak derivative. Let $F$ be the right continuous CDF of $\mu$. Then one has that:

$$\text{jump}(F) = \text{jump}(G) .$$
In addition there exists a constant $C$ such that the pointwise identity holds:

\begin{equation}
G(x) = F(x) + C , \quad \text{for all } x \notin \text{jump}(G) \cup \text{disc}(G).
\end{equation}

Moreover the set of discontinuities $D = \text{disc}(G)$ of $G(x)$ must be countable and we have a disjoint decomposition $D = \text{jump}(G) \cup D_{\text{spur}}$ where $G(x^-) = G(x^+)$ for all $x \in D_{\text{spur}}$. In particular after redefining $G$ at countably many points (including jumps) we can have (6) at every point. In other words we can have:

\begin{equation}
G(x) = \int_{(-\infty,x]} d\mu + C , \quad \text{where } G' = \mu \text{ in the sense of measures},
\end{equation}

that is where the relationship between $G$ and $\mu$ is given by (5).

**Proof.** It is enough to show the two identities:

\begin{equation}
F(x^+) - F(x^-) = G(x^+) - G(x^-) , \quad F(x^+) + F(x^-) = G(x^+) + G(x^-) - 2G(-\infty) ,
\end{equation}

at every point $x \in \mathbb{R}$. The first shows that $\text{jump}(F) = \text{jump}(G)$, while the second shows that $F(x) = G(x) + G(-\infty)$ at every point where $x \notin \text{jump}(F)$ and simultaneously $G(x) = G(x) = G(x^+)$. Recall that:

\begin{equation}
\int_{\mathbb{R}} \varphi d\mu = -\int_{-\infty}^{\infty} G(x)\varphi'(x)dx , \quad \text{for all } \varphi \in C^1_c(\mathbb{R}) ,
\end{equation}

and $F = \mu((-\infty,x])$. Now fix a point $x \in \mathbb{R}$ and first consider this identity with a sequence of test functions $\varphi_n(y-x)$ with the property:

\[
\varphi_n'(x) = \begin{cases} 
n\psi(nx + 1) , & x \leq 0 ; \\
n\psi(nx - 1) , & x > 0 , \end{cases}
\]

where $\psi(x)$ is a smooth bump function with $\text{supp}(\psi) \subseteq [-1,1]$ and $\int \psi dx = 1$. Then $\phi_n(0) = 1$ for all $n$ and $\text{supp}(\varphi_n) \subseteq [-1/n,1/n]$ for all $n$. In particular $\varphi_n(y-x) \to 1_{\{x\}}(y)$ for every $y$ so by DCT:

\[
\lim_n \int_{\mathbb{R}} \varphi_n(y-x) d\mu(y) = \int_{\mathbb{R}} 1_{\{x\}} d\mu = \mu(\{x\}) = F(x^+) - F(x^-) .
\]

On the other hand for each $n$ we have:

\[
\int_{-\infty}^{\infty} G(y)\varphi'_n(y-x)dy = \int_{-\infty}^{\infty} G(y+x)\varphi'_n(y)dy ,
\]

\[
= n \int_{-\infty}^{\infty} G(y+x)\psi(ny + 1))dy - n \int_{-\infty}^{\infty} G(y+x)\psi(ny - 1))dy ,
\]

\[
= \int_{-\infty}^{\infty} G\left(\frac{1}{n}y + x - \frac{1}{n}\right)\psi(y)dy - \int_{-\infty}^{\infty} G\left(\frac{1}{n}y + x + \frac{1}{n}\right)\psi(y)dy .
\]

Now $\text{supp}(\psi) \subseteq [-1,1]$, so we have:

\[
G\left(\frac{1}{n}y + x \pm \frac{1}{n}\right)\psi(y) = G(x^\pm)\psi(y) + o(1)\psi(y) ,
\]

where $o(1) \to 0$ as $n \to \infty$. This combined with the fact $\int \psi dx = 1$ shows that the limit of (9) is $G(x^-) - G(x^+)$. Equating this (with a minus sign) to (3) gives the first identity on line (7).

To prove the second identity on line (7) we use similar calculations with a slightly different sequence of test functions. Here we set:

\[
\varphi'_n(x) = \begin{cases} 
2n\psi(nx + n^2) , & x \leq -\frac{2}{n} ; \\
n\psi(nx + 1) , & -\frac{2}{n} \leq x \leq 0 ; \\
n\psi(nx - 1) , & x > 0 . \end{cases}
\]

The key difference now is twofold. First:

\[
\varphi_n(y-x) \to \begin{cases} 
2 , & x < 0 ; \\
1 , & x = 0 ; \\
0 , & x > 0 . \end{cases}
\]
so by DCT we pick up on the LHS of (8) the quantity:

$$\lim_n \int_{\mathbb{R}} \varphi_n(y-x) d\mu(y) = 2\mu((-\infty, x)) + \mu\{x\} = \mu((-\infty, x)) + \mu((-\infty, x]) = F(x^-) + F(x^+)\,.$$ 

On the other hand the RHS of (8) produces:

$$\int_{-\infty}^{\infty} G(x)\varphi'(x)dx = \int_{-\infty}^{\infty} G(\frac{1}{n}y + x - \frac{1}{n})\psi(y)dy + \int_{-\infty}^{\infty} G(\frac{1}{n}y + x + \frac{1}{n})\psi(y)dy - 2\int_{-\infty}^{\infty} G(\frac{1}{n}y + x - n)\psi(y)dy\,,$$

which limits to $G(x^-) + G(x^+) - 2G(-\infty)$. \qed