The geometry of relative plausibilities

Fabio Cuzzolin

The geometry of relative plausibilities

Fabio Cuzzolin
Computer Science Department, University of California at Los Angeles
cuzzolin@cs.ucla.edu

Abstract

The study of the interplay between belief and probability has recently been posed in a geometric setup, in which belief functions are represented as points of a simplex called belief space. Each belief function b is associated with three different geometric entities, which in turn determine three different Bayesian function related to b. In this paper, in particular, we study the geometry of *relative plausibilities*, and show how equidistribution can be possibly a *trait d'union* among all the Bayesian relatives of a belief function.

Keywords: Belief space, Bayesian relatives, relative plausibility, equidistribution, orthogonality.

1 Introduction

In the last decades a number of different uncertainty measures have been proposed, as alternatives or extensions of the classical probability theory. The theory of evidence is one the most popular formalism, extending quite naturally probabilities on finite spaces through the notion of *belief function*. Naturally enough, the connection between belief and probability plays a major role in the theory of evidence, and is the basement of a popular approach to evidential reasoning, Smets' *pignistic model* [12]. In fact, the problem of finding a correct probabilistic approximation of a belief function has been widely studied [9, 2], mainly in order to find efficient implementations of the rule of combination, as in Tessem [13], or Lowrance *et al.* [10]. Voorbraak [14], on his side, proposed to adopt the so-called *relative plausibility* function $\tilde{pl}_b$, i.e. the unique probability that, given a belief function b with plausibility pl_b, assigns to each singleton its normalized plausibility. He proved that $\tilde{pl}_b$ is a perfect representative of b when combined with other probabilities, $\tilde{pl}_b \oplus p = b \oplus p \ \forall p \in \mathcal{P}$. Cobb and Shenoy [4] described the properties of the relative plausibility of singletons and discussed its nature of probability function which is equivalent to the original belief function.

The study of the interplay between belief and probability has recently been posed in a geometric setup [8, 3], in which belief functions are represented as points of a simplex called *belief space* [7]. Each belief function is associated with three different geometric entities, namely the line (b, pl_b), the orthogonal complement $\mathcal{P}^\perp$ of the probabilistic subspace $\mathcal{P}$, and the simplex of consistent probabilities $P[b] = \{p \in \mathcal{P} : p(A) \geq b(A) \ \forall A \subset \Theta\}$. These in turn determine three different probabilities associated with b, i.e. the orthogonal projection $\pi[b]$ of b onto $\mathcal{P}$, the barycenter of $P[b]$, and the pignistic function $BetP[b]$. We recently showed that if the belief function b is 2-additive then pignistic function, orthogonal projection and intersection of (b, pl_b) with $\mathcal{P}$ coincide [5].

However, the analysis of the simplest binary case suggests that the relative plausibility function does not fit in this picture. In this paper we will study the geometry of relative plausibilities in the general case, pointing out its richness and complexity. We will see that their geometry can be described by three planes in the belief space, in turn related to an additional set of three Bayesian functions. As 2-additivity is not able to encompass both the above Bayesian functions and Voorbraak's relative plausibility function, we show that equidistribution can be a *trait d'union* for all the "Bayesian relatives" of a belief function.

2 Geometry of belief functions

In the *theory of evidence* [11] degrees of (subjective) belief are represented as *belief functions.* A *basic probability assignment* (b.p.a.) over a finite set Θ (frame of discernment) is a function $m : 2^\Theta \to [0,1]$ such that

$$m(\emptyset) = 0, \ \sum_{A\subset\Theta} m(A) = 1, \ \ m(A) \geq 0 \ \ \forall A \subset \Theta.$$

Subsets of Θ associated with non-zero values of m are called *focal elements* (f.e.). The *belief function* (b.f.) $b : 2^\Theta \to [0,1]$ associated with a basic probability assignment m is simply

$$b(A) = \sum_{B\subset A} m(B).$$

The basic probability assignment m of a belief function b can be uniquely recovered by means of the *Moebius inversion formula* $m(A) = \sum_{B\subset A}(-1)^{|A-B|}b(B)$. In particular, a probability function is a peculiar belief function called *Bayesian* b.f. assigning mass to singletons only: $m(A) = 0, |A| > 1$.

A dual representation of the evidence encoded by a belief function b is the *plausibility function*, whose value $pl_b(A)$ expresses the amount of evidence *not against* a proposition A: $pl_b(A) \doteq 1 - b(A^c) = \sum_{B\cap A\neq\emptyset} m(B) \geq b(A)$. From pl_b we can derive a Bayesian belief function called *relative plausibility of singletons* [14] as follows:

$$\begin{array}{rcl} \tilde{pl}_b : \ \Theta & \to & [0,1] \\ x & \mapsto & \tilde{pl}_b(x) = \dfrac{pl_b(x)}{\sum_{y\in\Theta} pl_b(y)}. \end{array} \quad (1)$$

Consider now a frame of discernment Θ and introduce in the Euclidean space $\mathbb{R}^N$, $N = |2^\Theta| - 1$, an orthonormal reference frame $\{x_A\}_{A\subset\Theta, A\neq\emptyset}$ in which each coordinate x_A measures the belief value $b(A)$. The *belief space* associated with Θ is then the set of points $\mathcal{B}_\Theta$ of $\mathbb{R}^N$ which correspond to a belief function. We assume the domain Θ fixed, and denote the belief space with $\mathcal{B}$. It is not difficult to prove [7] that $\mathcal{B}$ *is convex.* More precisely, if we call

$$b_A \doteq b \in \mathcal{B} \ s.t. \ m(A) = 1, \ m(B) = 0 \ \forall B \neq A$$

the unique belief function assigning all the mass to a single subset A of Θ, the belief space $\mathcal{B}$ is the convex closure Cl of all the basis belief functions b_A:

$$\mathcal{B} = Cl(b_A, \ A \subset \Theta, \ A \neq \emptyset). \quad (2)$$

Furthermore, each belief function $b \in \mathcal{B}$ can be written as a convex sum as follows:

$$b = \sum_{A\subset\Theta, \ A\neq\emptyset} m(A) b_A. \quad (3)$$

Since a probability is a belief function assigning non zero masses to singletons only, the set $\mathcal{P}$ of all the Bayesian belief functions is the simplex determined by all the basis functions associated with singletons $\mathcal{P} = Cl(b_{\{x\}}, x \in \Theta)$. Analogously, we call *plausibility space* the region $\mathcal{PL}$ of $\mathbb{R}^N$ whose points correspond to admissible plausibility functions. It can be proved [6] that this is also a simplex $\mathcal{PL} = Cl(pl_A, A \subset \Theta)$, whose vertices are

$$pl_A = -\sum_{B\subset A} (-1)^{|B|} b_B. \quad (4)$$

The A-th vertex pl_A of the plausibility space is the plausibility vector associated with the basis belief function b_A, $pl_A = pl_{b_A}$. Each plausibility function pl_b can be uniquely expressed as a convex combination of basis plausibility functions (4) as follows

$$pl_b = \sum_{A\subset\Theta} m(A) pl_A. \quad (5)$$

Belief functions are then nothing but special points in an Euclidean space. On the other side, given an ordering $A_1, .., A_N$ for the subsets of Θ, each point $x = [x_1, ..., x_N]' \in \mathbb{R}^N$ (where x' denotes the transpose of a vector x), $N \doteq 2^{|\Theta|} - 1$, can be thought of as a function $\varsigma : 2^\Theta \setminus \emptyset \to \mathbb{R}$ s.t. $\varsigma(A) = x_{ord(A)}$, where $ord(A)$ is the position of A in the ordering. As the Moebius transformation is invertible, for each of these functions ς there always exists another function $m_\varsigma : 2^\Theta \setminus \emptyset \to \mathbb{R}$ such that $\varsigma(A) = \sum_{B \subset A} m_\varsigma(B)$. Each vector ς of $\mathbb{R}^N$ can be thought of as a *sum function* [1], even though m_ς does not in general satisfy the positivity constraint $m_\varsigma(A) \geq 0 \; \forall \; A \subset \Theta$. The section $x_{ord(\Theta)} = 1$ of $\mathbb{R}^N$ corresponds to the constraint $\varsigma(\Theta) = 1$, so that points of this section are sum functions meeting the normalization axiom, $\sum_{A \subset \Theta} m_\varsigma(A) = 1$. We call them *normalized sum functions* (n.s.f.). Analogously to the case of belief functions, we call *Bayesian* normalized sum function a n.s.f. ς such that $\sum_{x \in \Theta} m_\varsigma(\{x\}) = 1$. Note that this implies $\sum_{|A|>1} m_\varsigma(A) = 0$, but not necessarily $m_\varsigma(A) = 0 \; \forall |A| > 1$.

3 A geometric interplay of belief and probability

As $b(\Theta) = pl_b(\Theta) = 1$ for all b.f. b, we can neglect the coordinate x_Θ and think of $\mathcal{B}$ as a region of $\mathbb{R}^{N-1}$, in which $b_\Theta = \mathbf{0} = [0, ..., 0]'$. Figure 1 shows the geometry of belief and plausibility spaces for a binary frame $\Theta_2 = \{x, y\}$, where belief and plausibility vectors are points of a plane with coordinates

$$
\begin{gathered}
b = [b(x) = m(x), b(y) = m(y)] \\
pl_b = [pl_b(x) = 1 - m(y), pl_b(y) = 1 - m(x)]
\end{gathered}
$$

respectively. These two simplices are symmetric with respect to $\mathcal{P}$. Furthermore, each pair of functions (b, pl_b) determines a line which is orthogonal to $\mathcal{P}$, on which they lay on symmetric positions on the two sides of the Bayesian region. The orthogonal projection of b onto $\mathcal{P}$, the pignistic function $BetP[b]$ and the center of mass $\bar{P}[b]$ of the probabilities $P[b] \doteq \{p \in \mathcal{P} : p(A) \geq b(A) \forall A \subset \Theta\}$ consistent with b *coincide* $\pi[b] = BetP[b] = \bar{P}[b]$. b is then associated with three geometric entities: the line (b, pl_b), the set of consistent probabilities $P[b]$, and the orthogonal com-

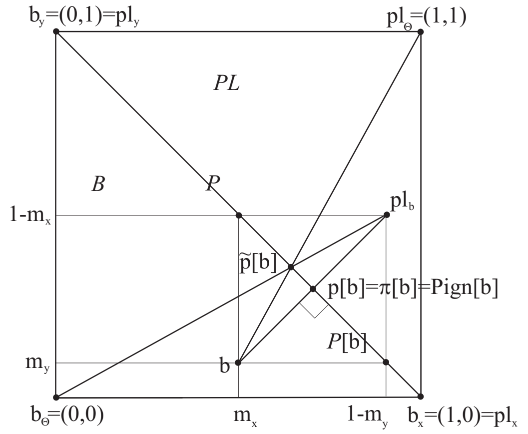

Figure 1: Geometry of the relative plausibility function in the bi-dimensional case $\Theta = \{x, y\}$.

plement $\mathcal{P}^\perp$ of $\mathcal{P}$. In [5] we proved that the line (b, pl_b) is always *orthogonal* to $\mathcal{P}$, even in the general case. However, (b, pl_b) *does not* intersect the probabilistic subspace in general, but there exists a unique Bayesian normalized sum function $\varsigma[b]$ such that

$$
\varsigma[b] = \beta[b] \cdot pl_b + (1 - \beta[b]) \cdot b \qquad (6)
$$

where

$$
\beta[b] = \frac{\sum_{|A|>1} m(A)}{\sum_{|A|>1} m(A)|A|} \qquad (7)
$$

is a scalar value between 0 and 1 which depends on b. $\varsigma[b]$'s values on singletons are

$$
m_{\varsigma[b]}(x) = m(x) + \beta[b] \cdot \sum_{A \supset x, A \neq x} m(A). \qquad (8)
$$

Hence, $\varsigma[b]$ is naturally associated with a Bayesian *belief function*, assigning the same mass (8) to each singleton, namely

$$
\sigma[b] \doteq \sum_{x \in \Theta} m_{\varsigma[b]}(x) \cdot b_{\{x\}}. \qquad (9)
$$

In [5] we also derived the convex shape of $P[b]$, and proved how the center of mass of this simplex $\bar{P}[b]$ is in fact the pignistic function. CITARE REFERENZA
The three geometric loci which arise from the analysis of the binary case are in conclusion each associated with a different Bayesian b.f. $(b, pl_b) \leftrightarrow \sigma[b]$, $\mathcal{P}^\perp \leftrightarrow \pi[b]$, $P[b] \leftrightarrow BetP[b] = \bar{P}[b]$. The *trait d'union* between

those Bayesian functions is again suggested by the binary case [5].

Proposition 1. *If the belief function b is 2-additive ($m(A) = 0$, $|A| > 2$) then pignistic function, orthogonal projection onto $\mathcal{P}$ and intersection of the line (b, pl_b) with $\mathcal{P}$ coincide, $BetP[b] = \bar{P}[b] = \pi[b] = \varsigma[b] = \sigma[b]$.*

However, it is clear from Figure 1 that Voorbraak's relative plausibility function does not fit in this picture, as 2-additivity is then not general enough to encompass both $\tilde{pl}_b$ and the above Bayesian functions. In the following sections we will study its geometry: for sake of simplicity, we will drop the dependence on b from the notations for σ, ς, β.

4 Geometry of relative plausibility

From the definition of relative plausibilities (1) it follows a straightforward geometric interpretation of $\tilde{pl}_b$. Let us call *plausibility of singletons* the quantity

$$\bar{pl}_b \doteq \sum_{x \in \Theta} pl_b(x) \cdot b_{\{x\}}. \tag{10}$$

$\bar{pl}_b$ can be written as $\bar{pl}_b = \sum_x pl_b(x) \cdot b_{\{x\}} + (1 - k_1) \cdot b_\Theta$ with $k_1 \doteq \sum_x pl_b(x)$ since $b_\Theta = \mathbf{0}$, in order for it to meet the normalization constraint. As $1 - k_1 \leq 0$ $\bar{pl}_b$ is a normalized sum function.

Now, by definition (1) $\tilde{pl}_b = \bar{pl}_b / k_1$. Since b_Θ is the origin of the reference frame, $\tilde{pl}_b$ lies on the segment $(\bar{pl}_b, b_\Theta)$. This in turn implies that $\tilde{pl}_b = (\bar{pl}_b, b_\Theta) \cap \mathcal{P}$ i.e. *$\tilde{pl}_b$ is the intersection of the line joining the vacuous belief function b_Θ and the plausibility of singletons $\bar{pl}_b$ with the probabilistic subspace.*

From the definition of σ (9) it follows that $\sigma = \sum_x m(x) b_{\{x\}} + \beta \sum_x (pl_b(x) - m(x)) b_{\{x\}} = (1-\beta) \sum_x m(x) b_{\{x\}} + \beta \sum_x pl_b(x) b_{\{x\}}$ so that, after defining

$$\bar{b} = \sum_{x \in \Theta} m(x) b_{\{x\}} = \sum_{x \in \Theta} m(x) b_{\{x\}} + (1 - k_2) b_\Theta, \tag{11}$$

with $k_2 \doteq \sum_{|A|>1} m(A)$, Eq. (10) implies

$$\sigma = (1 - \beta) \cdot \bar{b} + \beta \cdot \bar{pl}_b \tag{12}$$

where $\bar{b} \in Cl(b_\Theta, \mathcal{P}) \doteq \mathcal{D}$ and $\bar{pl}_b \in span(\mathcal{D})$, the affine (vector) space generated by the simplex $\mathcal{D}$. $\mathcal{D}$ is the simplex of *discounted* probabilities [11], i.e. belief functions assigning mass to singletons or Θ only. $\bar{b}$ is in fact the unique belief function assigning to Θ all the mass $(1 - k_2)$ which b gives to non-singletons. In the general case, then, the functions on the line (b, pl_b) are associated with other quantities laying on the line $(\bar{b}, \bar{pl}_b)$ in the affine subspace $span(\mathcal{D})$ of discounted probabilities $b \leftrightarrow \bar{b}$, $pl_b \leftrightarrow \bar{pl}_b$, $\varsigma \leftrightarrow \sigma$ which are in the same relative positions on the segment (just compare (6) and (12)). Again, 2-additivity plays a role, since [5]

Theorem 1. *$\varsigma[b] = \sigma[b]$ iff b is 2-additive, in which case $\sigma[b]$ is the orthogonal projection $\pi[b]$ of b onto $\mathcal{P}$.*

4.1 A three plane geometry

In the binary case $\mathcal{D} = \mathcal{B}$, $\bar{b} = b$ and $\bar{pl}_b = pl_b$, so that the plausibility of singletons is a plausibility function (Figure 1). In the general case, though, $\mathcal{D} = Cl(b_\Theta, \mathcal{P})$ is in correspondence with a "dual" simplex $\hat{\mathcal{D}} \subset \mathcal{PL}$, namely the set of plausibilities associated with a discounted probability. As pl_Θ is the dual of b_Θ, and each probability $b_{\{x\}} = pl_{\{x\}}$ is the dual of itself, $\hat{\mathcal{D}} = Cl(pl_\Theta, \mathcal{P})$. The functions $\bar{b}, \bar{pl}_b \in \mathcal{D}$ then correspond to two dual quantities in $\hat{\mathcal{D}}$ (recalling Equation (5))

$$\begin{aligned} \hat{b} &= \sum_{x \in \Theta} m(x) pl_{\{x\}} + (1 - k_2) pl_\Theta = \\ &= \bar{b} + (1 - k_2) pl_\Theta \\ \hat{pl}_b &= \sum_{x \in \Theta} pl_b(x) pl_{\{x\}} + (1 - k_1) pl_\Theta = \\ &= \bar{pl}_b + (1 - k_1) pl_\Theta. \end{aligned} \tag{13}$$

Theorem 2. *The line passing through the duals of plausibility of singletons (10) and belief of singletons (11) crosses $\sigma[b]$ too, and*

$$\beta(\hat{pl}_b - \hat{b}) + \hat{b} = \sigma[b] = \beta(\bar{pl}_b - \bar{b}) + \bar{b}. \tag{14}$$

The geometry of the relative plausibility function can then be described in terms of the *three planes* $(\bar{pl}_b, \sigma, \hat{pl}_b)$, $(b_\Theta, \tilde{pl}_b, pl_\Theta)$ and $(b_\Theta, \tilde{b}, pl_\Theta)$ (see Figure 2), where

$$\tilde{b} \doteq \bar{b} / k_2 \tag{15}$$

is another probability which is natural to call *relative belief of singletons*. More precisely, we have just seen how σ is the intersection of both $(\bar{b}, \bar{pl}_b)$ and $(\hat{b}, \hat{pl}_b)$ and lays in the same relative position on the two lines. They then determine a plane $(\bar{b}, \sigma, \hat{b}) = (\bar{pl}_b, \sigma, \hat{pl}_b)$. Furthermore, by definition $\tilde{pl}_b - b_\Theta = (\bar{pl}_b - b_\Theta)/k_1$ while (13) implies $\tilde{pl}_b = \bar{pl}_b/k_1 = (\hat{pl}_b - (1-k_1)pl_\Theta)/k_1$ so that $\tilde{pl}_b - pl_\Theta = (\hat{pl}_b - pl_\Theta)/k_1$. The relative plausibility function then lies in the same relative position on the two lines $(b_\Theta, \bar{pl}_b)$ and $(pl_\Theta, \hat{pl}_b)$, which intersect exactly in $\tilde{pl}_b$. $b_\Theta, pl_\Theta, \tilde{pl}_b, \bar{pl}_b$ and $\hat{pl}_b$ then determine another unique plane which we can denote with $(b_\Theta, \tilde{pl}_b, pl_\Theta)$. Analogously, by definition $\tilde{b} - b_\Theta = (\bar{b} - b_\Theta)/k_2$ while (13) yields $\tilde{b} - pl_\Theta = (\hat{b} - pl_\Theta)/k_2$, so that the relative belief of singletons lies in the same relative position on the two lines $(b_\Theta, \bar{b})$ and $(pl_\Theta, \hat{b})$, which intersect exactly in $\tilde{b}$. $b_\Theta, pl_\Theta, \tilde{b}, \bar{b}$ and $\hat{b}$ then determine a single plane denoted by $(b_\Theta, \tilde{b}, pl_\Theta)$.

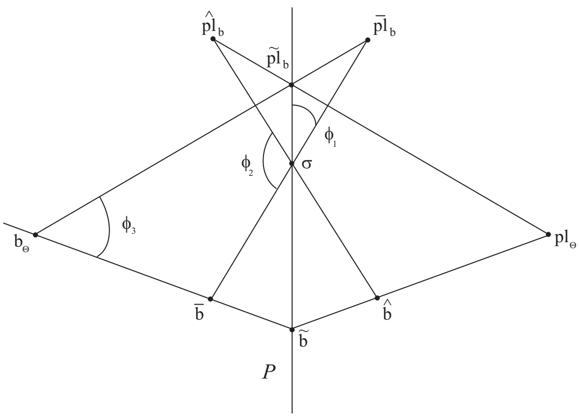

Figure 2: Planes and angles describing the geometry of the relative plausibility function.

4.2 A geometry of three angles

In the binary case, $b = \bar{b} = \hat{pl}_b$, $pl_b = \bar{pl}_b = \hat{b}$ and all these quantities are coplanar. This suggests a description of the geometry of $\tilde{pl}_b$ in terms of the three angles

$$\phi_1[b] \doteq \widehat{\tilde{pl}_b \sigma \bar{pl}_b}, \;\; \phi_2[b] \doteq \widehat{\bar{b}\sigma\tilde{pl}_b}, \;\; \phi_3[b] \doteq \widehat{\bar{b} b_\Theta \tilde{pl}_b}. \tag{16}$$

In fact, even though the line (b, pl_b) is always orthogonal to $\mathcal{P}$, $(\bar{b}, \bar{pl}_b)$ is *not* in general orthogonal to the probabilistic subspace:

$$\langle \bar{pl}_b - \bar{b}, b_{\{y\}} - b_{\{x\}} \rangle \propto \cos(\phi_1) \neq 0, \quad y \neq x \tag{17}$$

where $\langle \cdot \rangle$ denotes the scalar product, and $\{b_{\{y\}} - b_{\{x\}} \forall y \neq x\}$ are the basis vectors of $\mathcal{P}$. Nevertheless, ϕ_1 has an interesting interpretation in terms of belief.

Theorem 3. *$(\bar{b}, \bar{pl}_b) \perp \mathcal{P}$ ($\phi_1[b] = \pi/2$) iff $\forall x$ $\sum_{A \supset x, A \neq \{x\}} m(A) = pl_b(x) - m(x) = const$.*

In this case, recalling the definition of σ (9), $\sigma(x) = m(x) + \beta(pl_b(x) - m(x)) = m(x) + \frac{1-k_2}{\sum_y (pl_b(y) - m(y))}(pl_b(x) - m(x)) = m(x) + \frac{1-k_2}{n}$ and σ assigns the mass originally given by b to non-singletons *to all singletons on equal basis*. This can be expressed in terms of another probability function $R[b]$

$$R[b] \doteq \frac{\bar{pl}_b - \bar{b}}{k_1 - k_2} = \sum_x \frac{pl_b(x) - m(x)}{k_1 - k_2} b_{\{x\}}. \tag{18}$$

Corollary 1. *The dual line $(\bar{b}, \bar{pl}_b)$ is orthogonal to $\mathcal{P}$ iff the Bayesian b.f. $R[b]$ is the uninformative probability $\bar{\mathcal{P}} \doteq \sum_x b_{\{x\}}/n$.*

In fact $\sum_{x \in \Theta} \sum_{A \supset x, A \neq x} m(A) = \sum_{x \in \Theta}(pl_b(x) - m(x)) = k_1 - k_2$ so that the condition of Theorem 3 can be written as $pl_b(x) - m(x) = \sum_{A \supset x, A \neq x} m(A) = \frac{k_1 - k_2}{n}$ for all x. Replacing this in (18) yields $R[b] = \sum_{x \in \Theta} \frac{1}{n} b_{\{x\}}$. The value of ϕ_2 also depends on this mysterious probability (18).

Lemma 1. *ϕ_2 is nil if and only if*

$$\langle \mathbf{1}, R[b] \rangle^2 = \|R[b]\|^2 \langle \mathbf{1}, \mathbf{1} \rangle. \tag{19}$$

where $pl_\Theta = \mathbf{1}$ is the N-dim vector $[1, .., 1]'$.

Theorem 4. *The angle $\phi_2[b]$ is zero if and only if $R[b]$ is parallel to $pl_\Theta = \mathbf{1}$.*

Condition (19) has the form $\langle A, B \rangle^2 = \|A\|^2 \|B\|^2 \cos^2(\widehat{AB}) = \|A\|^2\|B\|^2$ i.e. $\cos^2(\widehat{AB}) = 1$, with $A = pl_\Theta$, $B = R[b]$. This yields $\cos(\widehat{R[b]pl_\Theta}) = 1$ or $\cos(\widehat{R[b]pl_\Theta}) = -1$, but as both $R[b]$ and pl_Θ have all positive components it must be $\cos(\widehat{R[b]pl_\Theta}) = 1$ so that $\widehat{R[b]pl_\Theta} = 0$. In this case the two lines $(\bar{b}, \bar{pl}_b)$ and $(\hat{b}, \hat{pl}_b)$ coincide. However, $R[b] \| pl_\Theta$ means $R[b] = \alpha \cdot pl_\Theta$ for some scalar

α, i.e. $R[b] = -\alpha \cdot \sum_{A \subset \Theta}(-1)^{|A|} b_A$ (since $pl_\Theta = -\sum_{A \subset \Theta}(-1)^{|A|} b_A$ by Equation (4)). On the other side, $R[b] = \sum_{x \in \Theta} R[b](x) b_{\{x\}}$ is a probability and the two conditions are never met together, unless $|\Theta| = 2$. In conclusion

Corollary 2. *$\phi_2[b] \neq 0$ and the lines $(\bar{b}, \bar{pl}_b)$, $(\hat{b}, \hat{pl}_b)$ never coincide $\forall b \in \mathcal{B}_\Theta$ when $|\Theta| > 2$; instead $\phi_2[b] = 0 \ \forall \ b \in \mathcal{B}_\Theta$ when $|\Theta| \leq 2$.*

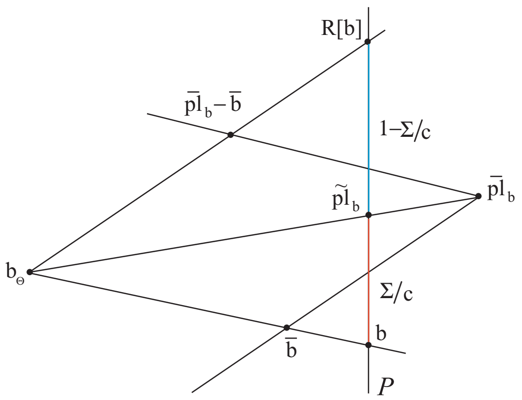

Figure 3: Location of $R[b]$ in $\mathcal{P}$.

The binary case is then really anomalous, a situation signaled by $R[b]$ being parallel to the vacuous plausibility pl_Θ. $R[b]$ has in fact a simple geometric interpretation, as Equation (18) implies $R[b](k_1 - k_2) = \bar{pl}_b - \bar{b} = \tilde{pl}_b \cdot k_1 - \tilde{b} \cdot k_2 = \tilde{pl}_b \cdot k_1 - \tilde{b} \cdot k_2 + k_1 \cdot \tilde{b} - k_1 \cdot \tilde{b} = k_1 \cdot (\tilde{pl}_b - \tilde{b}) + \tilde{b}(k_1 - k_2)$ from which

$$R[b] = \tilde{b} + \frac{k_1}{k_1 - k_2}(\tilde{pl}_b - \tilde{b}). \qquad (20)$$

$R[b]$ lies on the line joining $\tilde{b}$ and $\tilde{pl}_b$, outside the segment. Figure 3 illustrates how its location depends on the ratio between the total *belief* of singletons k_2 and the total *plausibility* of singletons k_1.

Finally, the angle $\phi_3[b]$ between $\tilde{b}$ and $\tilde{pl}_b$ nullifies iff $\tilde{b} = \tilde{pl}_b$, which is equivalent to $m(x)/k_2 = pl_b(x)/k_1 \ \forall x \in \Theta$. In other words, *relative plausibility and belief coincide iff mass and plausibility are distributed in the same way among the singletons.*

Again, a necessary and sufficient condition for $\phi_3[b] = 0$ can expressed in terms of $R[b]$, as $R[b](x) = (pl_b(x) - m(x))/(k_1 - k_2) = \frac{1}{k_1 - k_2}(\frac{k_1}{k_2} m(x) - m(x)) = m(x)/k_2 \ \forall x$, i.e. $R[b] = \tilde{b}$, with $R[b]$ "squashing" $\tilde{pl}_b$ onto $\tilde{b}$ from the outside. In this case the quantities $\bar{pl}_b, \hat{pl}_b, \tilde{pl}_b, \sigma[b], \bar{b}, \hat{b}, \tilde{b}$ all lie in the same plane.

5 Equidistribution, orthogonality and Bayesianity

This geometric analysis brings to full light the existence of several Bayesian functions linked to the original belief function by means of the balance between the overall plausibility k_1 of the elements of the frame, and the total mass k_2 assigned to them. This flag of the "non-Bayesianity" of b is symbolized by the Bayesian function $R[b]$. Those functions must be added to the three other probabilities discussed in Section 3. However, 2-additivity falls short of unifying this extended set of Bayesian relatives of b. The binary case again provides us with an intuition of how *equidistribution* can provide such a common ground. From Figure 1, we can appreciate how belief functions with $m(x) = m(y)$ lay on the bisector of the first quadrant which is orthogonal to $\mathcal{P}$. Their relative plausibility is also equal to their orthogonal projection $\pi[b]$.

In fact, Theorem 3 can be interpreted in terms of equal distribution of mass among focal elements. Consider the quantity

$$M_k(x) \doteq \sum_{A \supset x, |A| = k} m(A) \qquad (21)$$

representing the contribution of size k subsets to the plausibility of $x \in \Theta$. Suppose that $M_k(x) = const = M_k \ \forall x \in \Theta$ for all $k = 2, ..., n-1$ (trivially true for $k = n$). Then $\sum_{A \supset x, A \neq x} m(A) = \sum_{k=2}^{n} \sum_{|A|=k, A \supset x} m(A) = m(\Theta) + \sum_{k=2}^{n-1} M_k$ which is constant $\forall x \in \Theta$. In other words,

Corollary 3. *If $M_k(x) = const$ for all $x \in \Theta$, $\forall k = 2, ..., n-1$ then the dual line $(\bar{b}, \bar{pl}_b)$ is orthogonal to $\mathcal{P}$.*

It is worth to notice that [5]

Proposition 2. *If the belief function b is such that for all $x \in \Theta, k = 3, .., n-1$ $M_k(x) = const = M_k$ then pignistic function $\bar{P}[b]$ and orthogonal projection coincide, $\bar{P}[b] = \pi[b]$.*

The quantity $M_k(x)$ is then somehow connected to the geometric condition of orthogonality in the belief space framework. In [5] we have shown that a belief function b is orthogonal to the probabilistic subspace $\mathcal{P}$ iff

$$\sum_{B\supset y, B\not\supset x} m(B)2^{1-|B|} = \sum_{B\supset x, B\not\supset y} m(B)2^{1-|B|}. \tag{22}$$

We can use (22) to prove that

Theorem 5. *If a belief function b is such that $M_k(x) = const = M_k$ for all $k = 1, ..., n-1$ then b is orthogonal to the probabilistic subspace, $b \perp \mathcal{P}$.*

In fact, condition (22) is equivalent to $\sum_{B\supset y} m(B)2^{1-|B|} = \sum_{B\supset x} m(B)2^{1-|B|} \equiv \sum_{k=1}^{n-1} \frac{1}{2^k} \sum_{|B|=k, B\supset y} m(B) = \sum_{k=1}^{n-1} \frac{1}{2^k} \sum_{|B|=k, B\supset x} m(B) \equiv \sum_{k=1}^{n-1} \frac{1}{2^k} M_k(y) = \sum_{k=1}^{n-1} \frac{1}{2^k} M_k(x)$ for all $y \neq x$, and if $M_k(x) = M_k(y)$ $\forall y \neq x$ the condition is met.

In this case, confirming the intuition given by the binary case (Figure 1), all the Bayesian relatives of b *converge to the same probability*:

Theorem 6. *If $M_k(x) = const = M_k$ for all $k = 1, ..., n-1$ then $\tilde{pl}_b = \pi[b] = \bar{P}[b] = \bar{\mathcal{P}}$.*

To see this let us write the values of the pignistic function $\bar{P}[b](x)$ in terms of $M_k(x)$ as $\sum_{A\supset x} \frac{m(A)}{|A|} = \sum_{k=1}^{n} \sum_{A\supset x, |A|=k} \frac{m(A)}{k} = \sum_{k=1}^{n} \frac{M_k(x)}{k}$ which is constant under the hypothesis, yielding $\bar{P}[b] = \bar{\mathcal{P}}$. But now, as $pl_b(x) = \sum_{A\supset x} m(A) = \sum_{k=1}^{n} \sum_{A\supset x, |A|=k} m(A) = \sum_k M_k(x)$ we get $\tilde{pl}_b(x) = \frac{pl_b(x)}{k_1} = \frac{\sum_{k=1}^{n} M_k(x)}{\sum_x \sum_{k=1}^{n} M_k(x)}$ which is equal to $1/n$ if $M_k(x) = M_k$ $\forall k, x$.

6 Discussion

From the geometrical point of view, each belief function is associated with two different families of Bayesian functions (the Bayesian relatives of b). A first group of probabilities $\sigma[b], \pi[b], \bar{P}[b] = BetP[b]$ coincide under the assumption of 2-additivity. From the geometric study of the relative plausibility of singletons a second family of probabilities $\tilde{b}, \tilde{pl}_b, R[b]$ arises which does not meet this condition. The geometry of $\tilde{pl}_b$ can be described in terms of three planes and angles, which depend on the balance between total belief and total plausibility of singletons.

Equidistribution can provide a common ground for the two families. If subsets of the same size *contributes equally* to the plausibility of each singleton ($M_k(x) = const$),

$$\begin{array}{lll} k = 3, ..., n & \mapsto & \bar{P}[b] = \pi[b] \\ k = 2, ..., n & \mapsto & (\bar{b}, \bar{pl}_b) \perp \mathcal{P} \\ k = 1, ..., n & \mapsto & b \perp \mathcal{P}, \tilde{pl}_b = \bar{P}[b] = \pi[b]. \end{array}$$

Equidistribution is then inherently linked to the geometric notion of orthogonality, and provides a unifying viewpoint for all the geometric entities we encountered in the study of the Bayesian relatives of b: the orthogonal complement of $\mathcal{P}$, the probabilities $P[b]$ consistent with b, and the dual lines (b, pl_b) and $(\bar{b}, \bar{pl}_b)$. The next natural step will be the study of the behavior of all these function under the rule of combination.

Appendix

Proof. (Theorem 2) From Equation (13) we have that $\hat{b} - \bar{b} = (1 - k_2) pl_\Theta$ and $\hat{pl}_b - \bar{pl}_b = (1 - k_1) pl_\Theta$. Hence, $\beta(\hat{pl}_b - \hat{b}) + \hat{b} = \beta[\bar{pl}_b + (1-k_1)pl_\Theta - \bar{b} - (1-k_2)pl_\Theta] + \bar{b} + (1-k_2)pl_\Theta = \beta[\bar{pl}_b - \bar{b} + (k_2 - k_1)pl_\Theta] + \bar{b} + (1-k_2)pl_\Theta = \bar{b} + \beta(\bar{pl}_b - \bar{b}) + pl_\Theta[\beta(k_2 - k_1) + 1 - k_2]$ but $\beta(k_2 - k_1) + 1 - k_2 = \frac{1-k_2}{k_2-k_1}(k_2 - k_1) + 1 - k_2 = 0$ by definition of $\beta[b]$ (7), and (14) is met. $\square$

Proof. (Theorem 3) (17) can be written as $\langle \bar{pl}_b - \bar{b}, b_{\{y\}} - b_{\{x\}} \rangle = \langle \sum_{z\in\Theta} (pl_b(z) - m(z)) b_{\{z\}}, b_{\{y\}} - b_{\{x\}} \rangle = \sum_{z\in\Theta} (pl_b(z) - m(z)) \cdot [\langle b_{\{z\}}, b_{\{y\}} \rangle - \langle b_{\{z\}}, b_{\{x\}} \rangle] = \sum_{z\in\Theta} (pl_b(z) - m(z)) \cdot [\langle b_{\{z\}\cup\{y\}}, b_{\{z\}\cup\{y\}} \rangle - \langle b_{\{z\}\cup\{x\}}, b_{\{z\}\cup\{x\}} \rangle]$ since it is easy to see by the definition of b_A that $\langle b_A, b_B \rangle = \langle b_{A\cup B}, b_{A\cup B} \rangle$.

We distinguish three cases: if $z \neq x, y$ then $|z \cup x| = |z \cup y| = 2$ and the difference $\|b_{z\cup x}\|^2 - \|b_{z\cup y}\|^2$ goes to zero. If $z = x$ then $\|b_{z\cup x}\|^2 - \|b_{z\cup y}\|^2 = \|b_x\|^2 - \|b_{x\cup y}\|^2 = 2^{n-2} - 1 - (2^{n-1} - 1) = -2^{n-2}$, while if $z = y$ then $\|b_{z\cup x}\|^2 - \|b_{z\cup y}\|^2 = \|b_{x\cup y}\|^2 - \|b_y\|^2 = 2^{n-2}$. Hence (17) is equal to $2^{n-2} \cdot (pl_b(y) - m(y)) - 2^{n-2}(pl_b(x) - m(x))$

and the thesis follows as $\sum_{A\supset x, A\neq x} m(A) = pl_b(x) - m(x)$. $\square$

Proof. (Lemma 1) After noticing that $\beta[b] = \frac{1-k_2}{k_1-k_2}$ we can write $\hat{pl}_b - \sigma = (\hat{pl}_b - \bar{pl}_b) + (\bar{pl}_b - \sigma) = (1-k_1)pl_\Theta + (1-\beta)(\bar{pl}_b - \bar{b}) = \frac{k_1-1}{k_1-k_2} \cdot (\bar{pl}_b - \bar{b}) - (k_1-1)pl_\Theta$, so that $\langle \hat{pl}_b - \sigma, \bar{pl}_b - \sigma\rangle = \langle \bar{pl}_b - \sigma, \bar{pl}_b - \sigma\rangle + (1-k_1)\langle pl_\Theta, \bar{pl}_b - \sigma\rangle$. But now $\langle \bar{pl}_b - \sigma, \bar{pl}_b - \sigma\rangle = (1-\beta)^2\langle \bar{pl}_b - \bar{b}, \bar{pl}_b - \bar{b}\rangle$, being $\bar{pl}_b - \sigma = (1-\beta)(\bar{pl}_b - \bar{b})$, while $\langle pl_\Theta, \bar{pl}_b - \sigma\rangle = (1-\beta)\langle pl_\Theta, \bar{pl}_b - \bar{b}\rangle$ where $1-\beta = \frac{k_1-1}{k_1-k_2}$, so that by definition of $R[b]$ (18) we have $\langle \bar{pl}_b - \sigma, \bar{pl}_b - \sigma\rangle = (k_1-1)^2\langle R[b], R[b]\rangle$ and $(1-k_1)\langle pl_\Theta, \bar{pl}_b - \sigma\rangle = -(k_1-1)^2\langle \mathbf{1}, R[b]\rangle$. The desired scalar product is then $\langle \hat{pl}_b - \sigma, \bar{pl}_b - \sigma\rangle = (k_1-1)^2(\langle R[b], R[b]\rangle - \langle \mathbf{1}, R[b]\rangle)$. Now,

$$\cos(\pi - \phi_2) = \frac{\langle \hat{pl}_b - \sigma, \bar{pl}_b - \sigma\rangle}{\|\hat{pl}_b - \sigma\| \|\bar{pl}_b - \sigma\|}$$

where $\|\hat{pl}_b - \sigma\| = [\langle \hat{pl}_b - \sigma, \hat{pl}_b - \sigma\rangle]^{1/2} = (k_1-1)[\langle R[b] - pl_\Theta, R[b] - pl_\Theta\rangle]^{1/2}$ which is equal to $(k_1-1)[\langle R[b], R[b]\rangle + \langle pl_\Theta, pl_\Theta\rangle - 2\langle R[b], pl_\Theta\rangle]^{1/2}$ and $\|\bar{pl}_b - \sigma\| = \|(1-\beta)(\bar{pl}_b - \bar{b})\| = (k_1-1)\|R[b]\|$. So, as $\phi_2[b] = 0$ iff $\cos(\pi - \phi_2[b]) = -1$ we can write the condition as

$$-1 = \frac{(k_1-1)^2(\|R\|^2 - \langle \mathbf{1}, R\rangle)}{(k_1-1)^2\|R\|\sqrt{\langle R, R\rangle + \langle \mathbf{1}, \mathbf{1}\rangle - 2\langle R, \mathbf{1}\rangle}}$$

that is equivalent to (after elevating to the square both numerator and denominator) $\|R[b]\|^2(\|R[b]\|^2 + \langle \mathbf{1}, \mathbf{1}\rangle - 2\langle R[b], \mathbf{1}\rangle) = \|R[b]\|^4 + \langle \mathbf{1}, R[b]\rangle^2 - 2\langle \mathbf{1}, R[b]\rangle\|R[b]\|^2$ and by erasing the common terms we have as desired. $\square$

References

[1] Martin Aigner, *Combinatorial theory*, Classics in Mathematics, Springer, New York, 1979.

[2] Mathias Bauer, *Approximation algorithms and decision making in the Dempster-Shafer theory of evidence–an empirical study*, IJAR **17** (1997), 217–237.

[3] P. Black, *Geometric structure of lower probabilities*, Random Sets: Theory and Applications, Springer, 1997, pp. 361–383.

[4] B.R. Cobb and P.P. Shenoy, *On transforming belief function models to probability models*, Tech. report, University of Kansas, School of Business, Working Paper No. 293, February 2003.

[5] F. Cuzzolin, *Geometric interplays of belief and probability*, submitted to the IEEE Transactions on Systems, Man, and Cybernetics - Part B (2005).

[6] Fabio Cuzzolin, *Geometry of upper probabilities*, Proceedings of ISIPTA'03, July 2003.

[7] Fabio Cuzzolin and Ruggero Frezza, *Geometric analysis of belief space and conditional subspaces*, Proceedings of ISIPTA2001, 26-29 June 2001.

[8] V. Ha and P. Haddawy, *Theoretical foundations for abstraction-based probabilistic planning*, Proc. of the 12^{th} Conference on Uncertainty in Artificial Intelligence, August 1996, pp. 291–298.

[9] Ivan Kramosil, *Approximations of believeability functions under incomplete identification of sets of compatible states*, Kybernetika **31** (1995), 425–450.

[10] John D. Lowrance, T. D. Garvey, and Thomas M. Strat, *A framework for evidential-reasoning systems*, Proceedings of the National Conference on Artificial Intelligence, 1986, pp. 896–903.

[11] Glenn Shafer, *A mathematical theory of evidence*, Princeton Univ. Press, 1976.

[12] Philippe Smets, *Belief functions versus probability functions*, Uncertainty and Intelligent Systems, Springer Verlag, Berlin, 1988, pp. 17–24.

[13] Bjornar Tessem, *Approximations for efficient computation in the theory of evidence*, Artificial Intelligence **61:2** (1993), 315–329.

[14] F. Voorbraak, *A computationally efficient approximation of Dempster-Shafer theory*, International Journal on Man-Machine Studies **30** (1989), 525–536.