Note di Matematica ISSN 1123-2536, e-ISSN 1590-0932
Note Mat. **30** (2010) n. 1, 147–177. doi:10.1285/i15900932v30n1p147

Characterization of idempotent 2-copulas

William F. Darsow
Department of Mathematics, Illinois Institute of Technology, Chicago IL (retired)

Elwood T. Olsen
Department of Mathematics, Illinois Institute of Technology, Chicago IL (retired)
etocon@hotmail.com

Received: 15.7.2009; accepted: 1.2.2010.

Abstract. A 2-copula A induces a transition probability function p_A via

$$p_A(x,S) = \frac{d}{dx}\int_S \frac{\partial}{\partial t} A(x,t)\,dt.$$

where $S \in \mathcal{B}$, $\mathcal{B}$ denoting the Lebesgue measurable subsets of $[0,1]$. We say that a set S is invariant under A if $p_A(x,S) = \chi_S(x)$ for almost all $x \in [0,1]$, χ_S being the characteristic function of S. The sets S invariant under A form a sub-σ-algebra of the Lebesgue measurable sets, which we denote $\mathcal{B}_A$. A set $S \in \mathcal{B}_A$ is called an atom if it has positive measure and if for any $S' \in \mathcal{B}_A$, $\lambda(S' \cap S)$ is either $\lambda(S)$ or 0.

A 2-copula F is idempotent if $F * F = F$. Here $*$ denotes the product defined in [1]. Idempotent 2-copulas are classified and characterized as follows:

(i) An idempotent F is said to be nonatomic if $\mathcal{B}_F$ contains no atoms. If F is a nonatomic idempotent, then it is the product of a left invertible copula and its transpose. That is, there exists a copula B such that

$$B * B^T = F, \qquad \text{and}$$
$$B^T * B = M,$$

where $M(x,y) = \min(x,y)$.

(ii) An idempotent F is said to be totally atomic if there exist essentially disjoint atoms $S_n \in \mathcal{B}_F$ with

$$\sum_n \lambda(S_n) = 1.$$

If F is a totally atomic idempotent, then it is conjugate to an ordinal sum of copies of the product copula. That is, there exists a copula C satisfying $C * C^T = C^T * C = M$ and a partition $\mathcal{P}$ of $[0,1]$ such that

$$F = C * (\oplus_{\mathcal{P}} F_k) * C^T \tag{1}$$

where each component F_k in the ordinal sum is the product copula P.

(iii) An idempotent F is said to be atomic (but not totally atomic) if $\mathcal{B}_F$ contains atoms but the sum of the measures of a maximal collection of essentially disjoint atoms is strictly less than 1. In this mixed case, there exists a copula C invertible with respect to M and a partition $\mathcal{P}$ of $[0,1]$ for which (1) holds, with F_1 being a nonatomic idempotent copula and with $F_k = P$ for $k > 1$.

Some of the immediate consequences of this characterization are discussed.

Keywords: copula, idempotent, star product

MSC 2000 classification: primary 60G07, secondary 60J05, 60J25

1 Introduction

We address here idempotent copulas, meaning 2-copulas A for which $A * A = A$. The motivation for studying idempotents is a hope of developing a roadmap to the large and amorphous set of all 2-copulas. We present here what we believe to be a rather thorough analysis of idempotents. Whether this can lead to a general roadmap remains to be seen.

The $*$ product of two 2-copulas is defined as follows:

$$A * B(x, y) = \int_0^1 A_{,2}(x, t) B_{,1}(t, y)\, dt.$$

Here, and in general throughout this paper, $C_{,1}$ and $C_{,2}$ denote the partial derivatives of a 2-copula C with respect to its first and second arguments, respectively. The $*$ product of two copulas is always a copula. The $*$ product is an associative operation: $A*(B*C) = (A*B)*C$ for all copulas A, B and C. Furthermore, the $*$ product is continuous in each place with respect to uniform convergence: if $A_n \to A$ uniformly, then $A_n * B \to A * B$ and $B * A_n \to B * A$ uniformly. These and other properties of the $*$ product are proved in [1]. For convenience, we will generally write AB for the $*$ product of A and B, omitting the $*$, except for emphasis or where clarity may require it. Thus, also, we write A^2 for $A * A$ and the like.

The product copula P, given by $P(x, y) = xy$ and the min copula M given by $M(x, y) = \min(x, y)$ are idempotent. They are the most important idempotent copulas in an algebraic sense, since P is the (unique) universal annihilator – $PC = CP = P$ for all copulas C – and since M is the (unique) universal unit – $MC = CM = C$ for all copulas M.

What other copulas are idempotent? We know that there exist copulas which are invertible on the left but not on the right, with respect to M. A copula A has a left inverse if and only if for all y, $A_{,1}(x, y)$ is 1 or 0 for almost all x, and if this is the case, its left inverse is its transpose A^T, where A^T is defined by $A^T(x, y) = A(y, x)$. [1], Theorem 7.1, discussed further below. Similarly, a copula A possesses a right inverse if and only if for all x, $A_{,2}(x, y)$ is 1 or 0 for almost all y, and if this is the case, its right inverse is its transpose A^T. An example of a copula with a left inverse but no right inverse is the hat copula Λ given by

$$\Lambda(x, y) = \begin{cases} x, & 0 \le x \le 1/2, 2x \le y \le 1 \\ \frac{y}{2}, & 0 \le x \le 1/2, 0 \le y \le 2x \\ \frac{y}{2}, & 1/2 \le x \le 1, 0 \le y \le 2(1-x) \\ x + y - 1, & 1/2 \le x \le 1, 2(1-x) \le y \le 1. \end{cases}$$

We leave it to the reader to verify that Λ is a copula. Observe that since $\Lambda_{,1}$ is 0 or 1 almost everywhere, necessarily Λ possesses a left inverse which is equal to Λ^T. On the other hand $\Lambda_{,2}$ is not 0 or 1 almost everywhere; it has the value 1/2 in the triangular region bounded by $y = 2x$, $y = 2(1-x)$ and $y = 0$. According to the theorem, therefore Λ possesses no right inverse. By direct calculation, one obtains $\Lambda\Lambda^T = (M + W)/2$, where W is the copula given by $W(x, y) = \max(0, x + y - 1)$. One verifies readily that $(M + W)/2$ is idempotent. Indeed, it is always true, whenever a copula A possesses a left but not a right inverse, that, whereas $A^T A = M$, AA^T is an idempotent copula different from M. To see this, observe that $(AA^T)^2 = (AA^T)(AA^T) = A(A^T A)A^T = AMA^T = AA^T$, using the facts that the $*$ product is associative, that $A^T A = M$ and that M is a universal unit. Hence, AA^T is idempotent; $AA^T \ne M$, since by hypothesis A possesses no right inverse with respect to M. As will appear, it is easy to construct copulas with a left but not a right inverse, so this construction gives a large class of idempotent copulas.

Does it give them all? The answer is no. We cannot, for example, find a left but not right invertible copula A for which $P = AA^T$. For if we could, then it would have to be true that $P = A^T PA = A^T(AA^T)A = (A^T A)(A^T A) = M^2 = M$, contradiction, since P and M are not equal.

What is so different about P and, for example, $(M + W)/2$, which allows the latter but not the former to be decomposed in the form AA^T for a left invertible copula A? We show here that there is a significant difference in the invariant sets of the idempotents P and $(M+W)/2$. We turn now to this issue.

For any copula A, we define a transition probability function p_A via

$$p_A(x, S) = \frac{d}{dx}\int_S A_{,2}(x, t)\, dt.$$

where S is a measurable set. This is called a transition probability since, if we imagine a process with uniformly distributed random variables, for which A is the joint distribution of X_s and X_t, say, with $s < t$, then $p_A(x, S)$ gives the conditional probability $E(X_t \in S|X_s = x)$. We say that a set S is invariant under A if the equation $p_A(x, S) = \chi_S(x)$ is satisfied for almost all $x \in [0, 1]$. If S is invariant under A, and the process is in S at time s (i.e. $X_s \in S$), then the process will be in S at time t (i.e. $X_t \in S$), with probability 1.

The equation for the invariant sets of P is

$$\chi_S(x) = \frac{d}{dx}\int_S P_{,2}(x, t)\, dt = \lambda(S),$$

for a.a. x, where $\lambda(S)$ denotes the Lebesgue measure of S. This equation can be satisfied only if $\lambda(S) = 0$ or 1, and it is trivally satisfied for all S whose measure is 0 or 1. We conclude that the invariant sets of P are the measurable sets whose measure is 0 or 1.

On the other hand, the equation for the invariant sets of $(M + W)/2$ is

$$\begin{aligned}\chi_S(x) &= \frac{d}{dx}\int_S \left(\frac{1}{2}\chi_{[0,x]}(t) + \frac{1}{2}\chi_{[1-x,1]}(t)\right) dt\\ &= \frac{1}{2}\frac{d}{dx}\int_0^x \chi_S(t)\, dt + \frac{1}{2}\frac{d}{dx}\int_{1-x}^1 \chi_S(t)\, dt\\ &= \frac{1}{2}\chi_S(x) + \frac{1}{2}\chi_S(1-x) \text{ for a.a. } x.\end{aligned}$$

We conclude that S is an invariant set of $(M + W)/2$ if and only if $1 - x \in S$ for a.a. $x \in S$. These sets are just those which are essentially symmetric about the point $x = 1/2$. They constitute a σ-algebra, since the symmetry about $x = 1/2$ is preserved by the complement operation and by taking countable unions.

Note first that the invariant sets in both of these cases form a σ-algebra. This is true in general, as we will show later on. The difference between the two families of invariant sets is that $\mathcal{B}_P$, the collection of invariant sets of P, is an atomic σ-algebra, whereas $\mathcal{B}_{(M+W)/2}$, the collection of invariant sets of $(M + W)/2$, is a nonatomic σ-algebra. This difference turns out to be crucial for our analysis.

Terminology: Let $\mathcal{S}$ be a sub-σ-algebra of the Lebesgue measurable subsets of $[0, 1]$. $S \in \mathcal{S}$ is an atom if it has positive measure and if for all $S' \in \mathcal{S}$ either $\lambda(S \cap S') = \lambda(S)$ or $\lambda(S \cap S') = 0$. An atom is accordingly a set of positive measure, no non-trivial subset of which is a member of $\mathcal{S}$, sets which differ from from S or the empty set ϕ by a null set being considered trivial for this purpose. In the case of P, the invariant sets consist essentially of a single atom, the interval $[0, 1]$

itself, and the empty set ϕ. On the other hand, in the case of $(M+W)/2$, there are no atoms among the invariant sets. To see this, suppose S is a set of positive measure which is invariant under $(M+W)/2$. Then by the symmetry condition on S, $\lambda(S \cap [1/2,1]) = \lambda(S \cap [0,1/2])$ so that both $S \cap [1/2,1]$ and $S \cap [0,1/2]$ must have positive measure. Define

$$f(x) = \lambda([0,x] \cap S) = \int_0^x \chi_S(t)\,dt.$$

Then f is continuous and non-decreasing and maps $[0,1/2]$ onto $[0,\lambda(S)/2]$. Whenever $0 < a < \lambda(S)/2$ there is a number $x_0 \in (0,1/2)$ for which $f(x_0) = a$, by the intermediate value theorem. Set $Q = ([0,x_0] \cup [1-x_0,1]) \cap S$. Then Q satisfies the symmetry condition for an invariant set of $(M+W)/2$, hence is in $\mathcal{B}_{(M+W)/2}$, $Q = Q \cap S$, and Q has measure $2a > 0$. Since $2a < \lambda(S)$, by construction, S is not an atom. Since S was an arbitrary invariant set of positive measure, we conclude that $\mathcal{B}_{(M+W)/2}$ contains no atoms.

We will say that a copula C is atomic, if there are atoms among its invariant sets. We will say that C is nonatomic, if there are no atoms among its invariant sets. We will say that C is totally atomic, if among its invariant sets there is a collection of pairwise essentially disjoint atoms, the sum of whose measures is 1, so that the invariant sets consist essentially of atoms, their unions and the empty set. Using this terminology, P is totally atomic, since among its invariant sets $\mathcal{B}_P$ there is an atom whose measure is 1, namely $[0,1]$, and $(M+W)/2$ is nonatomic, since among its invariant sets $\mathcal{B}_{(M+W)/2}$ there are no atoms. We will see later that, in general, when an idempotent E has the form AA^T for some copula possessing a left inverse, it is a nonatomic idempotent. We will show that, conversely, if E is a nonatomic idempotent, it must have the form AA^T for some left invertible copula A.

Are there other atomic idempotents besides P? The answer is yes. Let $\mathcal{P} : \{(a_k, b_k)\}$ be any partition of $[0,1]$ (a partition is a collection of disjoint open intervals the sum of whose lengths is 1). Assign to each interval (a_k, b_k) a copula A_k. Define A via

$$A(x,y) = \begin{cases} a_k + (b_k - a_k) A_k\left(\frac{x-a_k}{b_k-a_k}, \frac{y-a_k}{b_k-a_k}\right), & a_k \le x, y \le b_k \\ M(x,y), & \text{otherwise.} \end{cases}$$

Then A is called an ordinal sum on the partition $\mathcal{P}$ with components A_k; it is a copula, [5, 7]. We will use the notation $A = \oplus_{\mathcal{P}} A_k$ for an ordinal sum. If $A = \oplus_{\mathcal{P}} A_k$ and $B = \oplus_{\mathcal{P}} B_k$ are ordinal sums on the same partition $\mathcal{P}$, then $A = B$ if and only if $A_k = B_k$ for all k, as is obvious, $A^T = \oplus_{\mathcal{P}} A_k^T$, as is also obvious, and $A * B = \oplus_{\mathcal{P}} A_k * B_k$, as is not so obvious, maybe, but follows from an elementary calculation. From these facts, it follows that an ordinal sum A is idempotent if and only if A_k is idempotent for all k. If $A = \oplus_{\mathcal{P}} A_k$ is an ordinal sum, the intervals (a_k, b_k) in the partition are invariant sets of A, and the only other invariant sets S are sets whose intersection with (a_k, b_k), appropriately scaled and translated, is an invariant set of A_k for all k. This is proved later on. From this, it follows that if $A = \oplus_{\mathcal{P}} A_k$ and each A_k is the product copula P, then A is idempotent, and the invariant sets of A are null sets, and sets which differ from the intervals (a_k, b_k), or unions of them, by null sets. Since the sum of the measures of the intervals (a_k, b_k) in $\mathcal{P}$ is 1, such an idempotent is totally atomic. This still does not exhaust the totally atomic copulas. But we will show that if a copula E is totally atomic, then it is conjugate to an ordinal sum of copies of the product copula P. That is, there exists a partition $\mathcal{P}$ and a copula C with a two-sided inverse with respect to M such that

$$E = C^T (\oplus_{\mathcal{P}} F_k) C,$$

where each component F_k is the product copula.

In the mixed case, where E is idempotent and atomic, but not totally atomic, we will show that E is conjugate to an ordinal sum $\oplus_{\mathcal{P}} F_k$ where one of the components F_k is a nonatomic idempotent and each of the other components F_k is the product copula P.

The plan of this paper is to develop the tools needed for the analysis in Section 2, then to address nonatomic, totally atomic, and atomic but not totally atomic copulas separately in Sections 3, 4 and 5. In Section 6, we will set forth some of the consequences of the characterization of idempotent copulas presented in the earlier sections.

2 Tools

There are several results which we will make use of in more than one subsequent section of this paper. They are collected here.

Left invertible and right invertible copulas. Many of our results depend ultimately on the characterization of copulas which possess a left or right inverse with respect to M. We state the relevant theorem here for reference throughout the paper.

Theorem 2.1. *A 2-copula A has a left inverse with respect to M if and only if for each $y \in [0,1]$, $A_{,1}(x,y) = 0$ or 1 for a.a. $x \in [0,1]$, and in that case A^T is the unique left inverse of A. A 2-copula A has a right inverse with respect to M if and only if for each $x \in [0,1]$, $A_{,2}(x,y) = 0$ or 1 for a.a. $y \in [0,1]$, and in that case A^T is the unique right inverse of A.*

Proof. Omitted. The theorem stated here combines Theorems 7.1 and 7.3 of [1]. $\boxed{QED}$

Representation of copulas using pairs of measure preserving functions. Many of the results herein exploit the association of a copula A with a pair of measure preserving Borel functions. Let $\mathcal{F}$ denote the set of all Borel measure preserving transformations of the interval $[0,1]$. That is, $f \in \mathcal{F}$ if and only if it is a Borel function whose range is contained in $[0,1]$ and it has the property that for all Borel sets B,

$$\lambda(f^{-1}B) = \lambda(B).$$

Theorem 2.2. *For $f, g \in \mathcal{F}$, define a function $C_{f,g}$ via*

$$C_{f,g}(x,y) = \lambda(f^{-1}[0,x] \cap g^{-1}[0,y]).$$

Then $C_{f,g}$ is a 2-copula. Furthermore, for every 2-copula C there exist $f, g \in \mathcal{F}$ such that $C = C_{f,g}$.

Proof. Omitted. See, e.g. [4, 9]. We presented a constructive proof in [2]. $\boxed{QED}$

For convenience, we will normally write C_{fg} for $C_{f,g}$, omitting the comma, unless it is necessary for clarity. The following theorem gives some useful properties of C_{fg}.

Theorem 2.3. *Let f, g and h denote elements of $\mathcal{F}$. Let e denote the identity function $e(x) = x$ for all $x \in [0,1]$. Let C_{fg} denote the 2-copula determined by f, g, as in Theorem 2.2 above. Let M denote the min copula: $M(x,y) = \min\{x,y\}$. Then for all $f \in \mathcal{F}$, $f, g \in \mathcal{F}$ or, $f, g, h \in \mathcal{F}$, as the case may be,*

(1) For all $x \in [0,1]$,

$$C_{fe,2}(x,y) = \frac{\partial}{\partial y} C_{fe}(x,y) = \chi_{f^{-1}[0,x]}(y) = \chi_{[0,x]}(f(y)) = \chi_{[f(y),1]}(x)$$

for a.a. $y \in [0,1]$. For all $y \in [0,1]$,

$$C_{ef,1}(x,y) = \frac{\partial}{\partial x} C_{ef}(x,y) = \chi_{f^{-1}[0,y]}(x) = \chi_{[0,y]}(f(x)) = \chi_{[f(x),1]}(y),$$

for a.a $x \in [0,1]$.

(2) C_{fe} *possesses a right inverse, and* C_{ef}*, a left inverse, with respect to* M.

(3) $C_{fg} = C_{fe} * C_{eg}$.

(4) $C_{ff} = M$.

(5) $C_{fg} = M$ *if and only if* $f = g$ *a.s.*

(6) $C^T_{fg} = C_{gf}$.

(7) Cancellation law: $C_{f\circ h, g\circ h} = C_{fg}$.

(8) $C_{fe} = C_{ge}$ *if and only if* $f = g$ *a.s.*

(9) $C_{fe} * C_{ge} = C_{f\circ g, e}$*, and* $C_{eg} * C_{ef} = C_{e, f\circ g}$.

Proof. (1) We can write

$$C_{fe}(x,y) = \lambda(f^{-1}[0,x] \cap [0,y]) = \int_0^y \chi_{f^{-1}[0,x]}(t)\,dt,$$

from which the existence a.e. of the partial derivative with respect to y follows. Observe that $y \in f^{-1}[0,x]$ iff $f(y) \in [0,x]$ iff $x \in [f(y),1]$. It follows that $\chi_{f^{-1}[0,x]} = \chi_{[0,x]} \circ f$ and $\chi_{[0,x]}(f(y)) = \chi_{[f(y),1]}(x)$. The conclusions regarding C_{ef} are proved similarly.

(2) It follows from (1) that for all x, $C_{fe,2}(x,y) = 0$ or 1 for a.a. y. Thus, by Theorem 2.1 above, C_{fe} possesses a right inverse with respect to M. The conclusion that C_{ef} possesses a left inverse is proved similarly.

(3) By (1),

$$C_{fe} * C_{eg}(x,y) = \int_0^1 \chi_{f^{-1}[0,x]}(t)\chi_{g^{-1}[0,y]}(t)\,dt = C_{fg}(x,y).$$

(4) $C_{ff}(x,y) = \lambda(f^{-1}[0,x] \cap f^{-1}[0,y]) = \lambda(f^{-1}\{[0,x] \cap [0,y]\})$. Since f is measure preserving, it follows that $C_{ff}(x,y) = \lambda([0,x] \cap [0,y]) = \min(x,y)$.

(5) If $f = g$ a.s., then $f^{-1}[0,y]$ and $g^{-1}[0,y]$ differ by a null set for all y, whence

$$C_{fg}(x,y) = \lambda(f^{-1}[0,x] \cap g^{-1}[0,y]) = \lambda(f^{-1}[0,x] \cap f^{-1}[0,y]) = M(x,y),$$

using (4). On the other hand, if $C_{fg} = M$, then, since $C_{fg}(x,x) = x$ and $C_{fg}(x,1) = x$ for all x, we must have

$$\lambda(f^{-1}[0,x) \cap g^{-1}(x,1]) = \lambda\{f < x < g\} = 0$$

for all x. Let r_n be an enumeration of the rational numbers in $[0,1]$. Here and elsewhere, we use the terminology $\{f < x\}$ to denote the set $\{t \mid f(t) < x\}$, and similarly for other relations, so $\{f < x < g\} = \{t \mid f(t) < x < g(t)\}$. Since $f(t) < g(t)$ iff there is a rational number r_k such that $f(t) < r_k < g(t)$, we can write

$$\{f < g\} = \cup_{k=1}^{\infty}\{f < r_k < g\},$$

whence $\{f < g\}$, being the union of a countable collection of null sets, is itself a null set. Similarly, $\{g < f\}$ is a null set. Therefore, $\{f \neq g\} = \{f < g\} \cup \{g < f\}$ is a null set.

(6) $C^T_{fg}(x,y) = C_{fg}(y,x) = \lambda(f^{-1}[0,y] \cap g^{-1}[0,x]) = C_{gf}(x,y)$.

(7) We have

$$\begin{aligned}C_{f\circ h,g\circ h}(x,y) &= \lambda(h^{-1}f^{-1}[0,x]\cap h^{-1}g^{-1}[0,y])\\ &= \lambda(h^{-1}\{f^{-1}[0,x]\cap g^{-1}[0,y]\})\\ &= \lambda(f^{-1}[0,x]\cap g^{-1}[0,y])\\ &= C_{fg}(x,y),\end{aligned}$$

using the fact that h is measure preserving.

(8) If $C_{fe} = C_{ge}$ then also $C_{fe}^T = C_{ge}^T$, so by (6) $C_{ef} = C_{eg}$. Then

$$M = C_{ff} = C_{fe}C_{ef} = C_{fe}C_{eg} = C_{fg},$$

using (3) and (4). Thus, by (5), $f = g$ a.s. On the other hand, if $f = g$ a.s., then

$$M = C_{fg} = C_{fe}C_{eg},$$

by (3) and (5). Since the left inverse of a copula, when it exists, is the transpose of the copula, $C_{fe} = C_{eg}^T = C_{ge}$, by (6).

(9) We have

$$\begin{aligned}C_{ge} * C_{e,f\circ g} &= C_{g,f\circ g} \text{ by (3)}\\ &= C_{e\circ g,f\circ g} \text{ using } e\circ g = g\\ &= C_{ef} \text{ by (7).}\end{aligned}$$

It follows that $C_{fe} * C_{ge} * C_{e,f\circ g} = C_{fe} * C_{ef} = C_{ff} = M$, by (3) and (4). Thus, $C_{fe}C_{ge}$ is the left inverse of $C_{e,f\circ g}$ with respect to M. Since the left inverse of $C_{e,f\circ g}$ is $C_{e,f\circ g}^T = C_{f\circ g,e}$, by (3) and (4), and since left inverses, when they exist, are unique, Theorem 2.1 above, we must have $C_{f\circ g,e} = C_{fe}C_{ge}$. Taking transposes, it follows also that $C_{eg}C_{ef} = C_{e,f\circ g}$. $\boxed{QED}$

Some of the properties set forth in Theorem 2.3 have been established by others, e.g. [4]; we include them here for completeness and easy reference.

The following theorem is interesting in itself, and it leads to some useful results, which are stated as corollaries to the theorem.

Theorem 2.4. *Same terminology as in Theorem 2.3. If C is a copula which possesses a left inverse with respect to M, then the function f defined by*

$$f(x) = \inf\{y \mid C_{,1}(x,y) = 1\}$$

is a measure preserving Borel function, and $C = C_{ef}$.

Proof. Since by hypothesis C possesses a left inverse, necessarily for all y, $C_{,1}(x,y)$ is 0 or 1 for a.a. x, by Theorem 2.1. We claim first that if $C_{,1}(x,y_0) = 1$ and $y > y_0$, then necessarily $C_{,1}(x,y)$ exists and equals 1. To see this, observe that for $h > 0$ and $y_0 < y$ we must have

$$\begin{aligned}0 \le \frac{C(x+h,y_0) - C(x,y_0)}{h} &\le \frac{C(x+h,y) - C(x,y)}{h}\\ &\le \frac{C(x+h,1) - C(x,1)}{h} = 1,\end{aligned}$$

since the mass assigned by C to the rectangle $[x, x+h] \times [0, y_0]$ is bounded above by the mass assigned to the rectangle $[x, x+h] \times [0, y]$, which in turn is bounded above by the mass assigned to the entire verticle strip $[x, x+h] \times [0,1]$. Since this is true for all $h > 0$, and by hypothesis the limit as $h \downarrow 0$ of the term on the left exists and equals 1, the term in the middle is constrained to have the limit 1 as $h \downarrow 0$. Hence, C has a right first partial derivative equal to 1 at (x, y). By a similar argument on the adjoining vertical strip $[x-h, x] \times [0,1]$, we conclude that C also has a left first partial derivative equal to 1 at (x, y). Hence $C_{,1}(x, y)$ exists and equals 1, as claimed.

Next, set $S_a = \{x | C_{,1}(x, a) = 1\}$. We claim that S_a is a Borel set for all a. To see this, let $h_n \downarrow 0$, and observe that the function $x \to C_{,1}(x, a)$ is the pointwise limit of the continuous (hence Borel) functions

$$x \to \frac{C(x+h_n, a) - C(x, a)}{h_n},$$

where it exists. Furthermore, the set of points where the limit fails to exist is the set of points where the lim inf and lim sup of this collection, both Borel functions, and the lim inf and lim sup of the analogous collection

$$x \to \frac{C(x, a) - C(x-h_n, a)}{h_n},$$

also Borel functions, are not all the same, hence is also a Borel set. In fact, the points where the derivative fails to exist necessarily constitute a Borel null set, since C is a copula whose first partial derivative $C_{,1}(x, a)$ exists for almost all x. For definiteness, redefine $C_{,1}(x, a)$ to be zero where the derivative does not exist, in order to obtain a Borel function defined for all x. The sets S_a are level sets of these Borel functions, hence Borel sets. These sets are also nested: for $a < b$, $S_a \subset S_b$ by the claim proved above.

We are now in a position to use the classical method of constructing a measurable function from suitable level sets. Let r_n be an enumeration of the rational numbers in $[0,1]$ and define a function f on $[0,1]$ via

$$f(x) = \inf\{r_n | x \in S_{r_n}\}.$$

Then f is a Borel function, by the classical argument. E.g. [6], Chapter 11, Lemma 9.

To complete the proof, we show that f is measure preserving and that $C = C_{ef}$. Since $f(x) < a$ implies that there is a rational number r_n between $f(x)$ and a such that $C_{,1}(x, r_n) = 1$, necessarily $C_{,1}(x, a) = 1$, by the claim proved above. On the other hand, if $C_{,1}(x, a) = 1$, then by the claim proved above, $C_{,1}(x, r_n) = 1$ for all rational numbers $r_n > a$, whence $f(x) \leq a$. Thus, for all a, $\{f < a\} \subset \{x | C_{,1}(x, a) = 1\} \subset \{f \leq a\}$. Let $\epsilon > 0$. Then

$$\{x | C_{,1}(x, a) = 1\} \subset \{f \leq a\} \subset \{f < a + \epsilon\} \subset \{x | C_{,1}(x, a+\epsilon) = 1\}.$$

It follows that

$$\lambda(\{x | C_{,1}(x, a) = 1\}) \leq \lambda(\{f \leq a\}) \leq \lambda(\{x | C_{,1}(x, a+\epsilon) = 1\}).$$

Since $C_{,1}(x, a)$ is 0 where it is not 1, we have

$$\lambda(\{x | C_{,1}(x, a) = 1\}) = \int_0^1 C_{,1}(x, a)\, dx = C(1, a) - C(0, a) = a.$$

Similarly, $\lambda(\{x | C_{,1}(x, a+\epsilon) = 1\}) = a + \epsilon$. Thus, $a \leq \lambda(\{f \leq a\}) \leq a + \epsilon$ holds for all $\epsilon > 0$. It follows that $\lambda(\{f \leq a\}) = a$, that is, f is measure preserving. Then also $\lambda(\{f < a\}) = a$ and $\lambda(\{f = a\}) = 0$. It follows that for each a, $\chi_{f^{-1}[0,a]}(t) = C_{,1}(t, a)$, for a.a. t. Integrate from 0 to x to obtain $C_{ef} = C$. $\boxed{QED}$

By Theorem 2.3, C_{ef} is left invertible for all $f \in \mathcal{F}$ and by Theorem 2.4, a left invertible copula A can always be written in the form C_{ef}. Observe that if a function A is known to be a copula, we can test A to determine whether it possesses a left inverse by taking its first partial derivative. If that derivative is 0 or 1 a.e., A possesses a left inverse, if not, it does not, by Theorem 2.1. In addition, if the derivative is 0 or 1 almost everywhere, and we want to write A in the form C_{ef}, Theorem 2.4 says that we can read off from the derivative what the function f must be. For example, the hat copula Λ discussed in Section 1 has a first partial derivative which is 0 on the triangular region bounded by the lines $y = 2x$, $y = 2(1 - x)$ and $y = 0$ and is 1 elsewhere. By Theorem 2.3, therefore, the hat function f defined by

$$f(x) = \begin{cases} 2x, & 0 \le x \le 1/2 \\ 2(1-x) & 1/2 \le x \le 1 \end{cases}$$

is a measure preserving Borel function and $\Lambda = C_{ef}$.

A function $f \in \mathcal{F}$ is said to possess an essential inverse g if $g \in \mathcal{F}$ and $g \circ f = e$ a.s. and $f \circ g = e$ a.s.

Corollary 2.4.1. *Same terminology as in Theorem 2.3. A copula C possesses a two-sided inverse with respect to M if and only if $C = C_{fe}$ for some measure preserving Borel function $f \in \mathcal{F}$ which possesses an essential inverse $g \in \mathcal{F}$.*

Proof. If $C = C_{fe}$ and g is an essential inverse of f, then $C_{ge}C_{fe} = C_{g\circ f,e} = M$, by parts (5) and (9) of Theorem 2.3, and similarly $C_{fe}C_{ge} = C_{f\circ g,e} = M$. Thus, C_{fe} possesses a two-sided inverse with respect to M. On the other hand, if $CC^T = C^TC = M$ then C^T has a left inverse, so by Theorem 2.4 there exists $f \in \mathcal{F}$ such that $C^T = C_{ef}$. Then $C = C_{fe}$. Since also C has a left inverse, Theorem 2.4 says that there exists $g \in \mathcal{F}$ such that $C_{fe} = C_{eg}$. Using properties from Theorem 2.3, we obtain

$$M = C_{gg} = C_{ge}C_{eg} = C_{ge}C_{fe} = C_{g\circ f,e}.$$

It then follows from Theorem 2.3, part (5), that $g \circ f = e$ a.s. Similarly,

$$M = C_{ff} = C_{fe}C_{ef} = C_{fe}C_{eg}^T = C_{fe}C_{ge} = C_{f\circ g,e},$$

whence $f \circ g = e$ a.s. $\boxed{QED}$

Remark: Let $\mathcal{G}$ be the set of $f \in \mathcal{F}$ possessing an essential inverse with respect to e, and let $\hat{\mathcal{G}}$ be the set of equivalence classes in $\mathcal{G}$ under the relation $f \sim g$ if $\lambda(\{f \neq g\}) = 0$. Then $\hat{\mathcal{G}}$ is a group under the binary operation on equivalence classes inherited from composition of functions, and it is in one-to-one correspondence with G_M, the group of 2-copulas invertible with respect to M, via the map $E(g) \to C_{ge}$, where $E(g) \in \hat{\mathcal{G}}$ is the equivalence class of $g \in \mathcal{G}$. The proof is based on the foregoing result and part (8) of Theorem 2.3. Details are left to the reader. In addition, this map is a group homomorphism; this follows from part (9) of Theorem 2.3.

Corollary 2.4.2. *Same terminology as in the statement of Theorem 2.3. Idempotent copulas $C_{ef}C_{fe}$ and $C_{eh}C_{he}$ are equal if and only if there is a Borel function $g \in \mathcal{F}$ possessing an essential inverse, for which $g \circ h = f$ a.s.*

Proof. Suppose $f = g \circ h$ a.s. and g possesses an essential inverse. Then

$$C_{ef}C_{fe} = C_{e,g\circ h}C_{g\circ h,e} = C_{eh}C_{eg}C_{ge}C_{he},$$

using part (9) of Theorem 2.3. Since g possesses an essential inverse, C_{ge} possesses a two-sided inverse with respect to M, by Corollary 2.4.1 above, and since that inverse must be $C_{ge}^T = C_{eg}$, we have $C_{eg}C_{ge} = M$. Insert this in the expression above to obtain $C_{ef}C_{fe} = C_{eh}MC_{he} = C_{eh}C_{he}$.

If, on the other hand, $C_{ef}C_{fe} = C_{eh}C_{he}$, pre- and post-multiply by C_{fe} and C_{ef} to obtain $M = C_{fh}C_{hf}$. If we pre- and post-multiply the original expression instead by C_{he} and C_{eh}, we obtain $C_{hf}C_{fh} = M$. We conclude that C_{fh} has a two-sided inverse with respect to M. Thus, by Corollary 2.4.1 above, there exists an essentially invertible Borel function $g \in \mathcal{F}$ for which $C_{fh} = C_{ge}$. Then

$$M = C_{ge}C_{hf} = C_{ge}C_{he}C_{ef} = C_{g\circ h,f}.$$

It follows from part (5) of Theorem 2.3 that $g \circ h = f$ a.s. $\boxed{QED}$

Essential equivalence of sub-σ-algebras of measurable sets. We say that two sub-σ-algebras $\mathcal{S}$ and $\mathcal{T}$ of the measurable subsets of $[0,1]$ (meaning the Lebesgue measurable subsets of $[0,1]$) are essentially equivalent if for all $S \in \mathcal{S}$ there is a set $T \in \mathcal{T}$ such that $\lambda(S\Delta T) = 0$ and for all $T \in \mathcal{T}$ there is an $S \in \mathcal{S}$ such that $\lambda(S\Delta T) = 0$. The symbol Δ denotes the symmetric difference operator. In general, the sets $S\Delta T$ are not members of either $\mathcal{S}$ or $\mathcal{T}$; the relationship accordingly demands some overarching family of subsets which includes both $\mathcal{S}$ and $\mathcal{T}$. This overarching family is the Lebesgue measurable subsets of $[0,1]$, denoted here $\mathcal{B}_L$ (and elsewhere just $\mathcal{B}$, since, for most purposes, we make no distinction between Lebesgue measurable and Borel measurable sets). The essential equivalence relation is in fact an equivalence relation among sub-σ-algebras of $\mathcal{B}_L$; we leave it to the reader to verify this.

We present here some minor theorems, whose intent is to show that, for our purposes, the essential equivalence relation among sub-σ-algebras is such a close relationship, that we can treat essentially equivalent sub-σ-algebras as equal (which is what we will do, in the remainder of this paper).

Theorem 2.5. *If ψ and θ are measurable functions, and $\lambda(\{\psi \neq \theta\}) = 0$, then $\psi^{-1}(\mathcal{B})$ is essentially equivalent to $\theta^{-1}(\mathcal{B})$. Conversely, if ψ is measurable with respect to $\mathcal{S}$ and $\mathcal{T}$ is essentially equivalent to $\mathcal{S}$, there is a function θ measurable with respect to $\mathcal{T}$ such that $\lambda(\{\psi \neq \theta\}) = 0$.*

Proof. If $S \in \psi^{-1}(\mathcal{B})$ there is a Borel set B such that $S = \psi^{-1}(B)$. Set $T = \theta^{-1}(B)$ and observe that $S\Delta T \subset \{\psi \neq \theta\}$, whence $S\Delta T$ is a null set. The argument in the other direction, starting with $T \in \theta^{-1}(\mathcal{B})$ is identical.

As for the converse, ψ is measurable with respect to $\mathcal{S}$ if $\psi^{-1}(B) \in \mathcal{S}$ for every Borel set B, or $\psi^{-1}(\mathcal{B}) \subset \mathcal{S}$. It is not necessarily the case that $\psi^{-1}(\mathcal{B}) = \mathcal{S}$. We outline the proof when ψ is a bounded function. The proof can be extended to unbounded functions by taking limits of bounded functions. Suppose ψ is such that $\psi^{-1}(B) \in \mathcal{S}$ for all Borel sets B and ψ is bounded, that is, there is a number M such that $|\psi(x)| < M$ for all x. Let

$$\psi_n = \sum_{k=1}^{m(n)} a_{nk}\chi_{S_{nk}}$$

be a sequence of $\mathcal{S}$-simple functions converging pointwise almost everywhere to ψ. Arrange it so that for all n $S_{nk} \cap S_{n\ell} = \phi$ when $k \neq \ell$, and arrange it so that $|a_{nk}| < M$ for all n, k. The set N_0 where the sequence ψ_n fails to converge is the set where $\liminf \psi_n \neq \limsup \psi_n$, which is in $\mathcal{S}$, since the lim inf and lim sup are $\mathcal{S}$-measurable functions. For each S_{nk} let $T_{nk} \in \mathcal{T}$ be such that $\lambda(S_{nk}\Delta T_{nk}) = 0$, set

$$N_n = \cup_{k=1}^{m(n)-1} \cup_{\ell=k+1}^{m(n)} (T_{nk} \cap T_{n\ell})$$

and observe that N_n is a null set in $\mathcal{T}$ and that the sets $T_{nk} \setminus N_n, k = 1, \ldots, m(n)$ are pairwise disjoint. Define $\theta_n = \sum_{k=1}^{m(n)} a_{nk}\chi_{T_{nk}\setminus N_n}$. Then by construction the functions θ_n are uniformly bounded by M and, since $\theta_n = \psi_n$ a.e., they converge pointwise almost everywhere. The set N where they fail to converge is in $\mathcal{T}$, by an argument parallel to that used above. Define

$$\theta(x) = \begin{cases} \lim_n \theta_n(x), & \text{where the limit exists} \\ 0, & \text{where the limit does not exist.} \end{cases}$$

Then θ is a $\mathcal{T}$-measurable function, and

$$\{\psi \neq \theta\} \subset N_0 \cup N \cup (\cup_n \{\psi_n \neq \theta_n\}).$$

Since all of the sets on the right are null sets, $\{\psi \neq \theta\}$ is a null set. $\boxed{QED}$

Note that Theorem 2.5 implies that if f and h are equivalent measure preserving Borel functions, then $f^{-1}(\mathcal{B})$ and $h^{-1}(\mathcal{B})$ are essentially equivalent.

The essential equivalence relation among sub-σ-algebras preserves the properties of foremost interest here:

Theorem 2.6. *Suppose $\mathcal{S}$ and $\mathcal{T}$ are essentially equivalent sub-σ-algebras of $\mathcal{B}_L$. If $S \in \mathcal{S}$ and $T \in \mathcal{T}$ satisfy $\lambda(S\Delta T) = 0$, then S is an atom if and only if T is an atom. $\mathcal{S}$ is nonatomic if and only if $\mathcal{T}$ is nonatomic. If $S_k \in \mathcal{S}$ and $T_k \in \mathcal{T}$ and $\lambda(S_k \Delta T_k) = 0$ for all k, then $\{S_k\}$ is a maximal collection of essentially disjoint atoms in $\mathcal{S}$ if and only if $\{T_k\}$ is a maximal collection of essentially disjoint atoms in $\mathcal{T}$.*

Proof. Suppose $S \in \mathcal{S}$ and $T \in \mathcal{T}$ satisfy $\lambda(S\Delta T) = 0$, and suppose $A \in \mathcal{S}$ and $A \subset S$. By essential equivalence of $\mathcal{S}$ and $\mathcal{T}$, there exists $B \in \mathcal{T}$ such that $\lambda(A\Delta B) = 0$. Write $N_1 = S\Delta T$, $N_2 = A\Delta B$ and $N = N_1 \cup N_2$. All of these sets are Lebesgue null sets. Then

$$\begin{aligned} B &= A\Delta N_2 \subset A \cup N_2 \\ A &= B\Delta N_2 \subset B \cup N_2 \\ S &= T\Delta N_1 \subset T \cup N_1 \text{ and} \\ T &= S\Delta N_1 \subset S \cup N_1, \end{aligned}$$

as is readily verified. It follows that

$$A = A \cap S \subset (B \cap T) \cup N$$

and that

$$B \cap T \subset (A \cap S) \cup N = A \cup N.$$

We have exhibited a subset $B \cap T \subset T$ which differs from A by a null set and which accordingly satisfies $\lambda(B \cap T) = \lambda(A)$. It follows that if S is not an atom, T cannot be an atom (take A such that $0 < \lambda(A) < \lambda(S)$) and accordingly that if T is an atom, S must be an atom. By parallel argument, starting with a set $B \in \mathcal{T}$ which is a subset of T, we conclude that if T is not an atom, S cannot be an atom, and accordingly that if S is an atom, T must likewise be an atom. This completes the proof of the first assertion of the theorem. The remaining assertions follow from the first by standard arguments and are left to the reader. $\boxed{QED}$

Denote the Borel sets of measure 0 by $\mathcal{N}_B$, and define

$$\mathcal{N}_L = \{N \mid \text{ there exists } S \in \mathcal{N}_B \text{ such that } N \subset S.\}$$

For any sub-σ-algebra $\mathcal{S}$ of the Lebesgue measurable sets, we define the completion of $\mathcal{S}$, denoted $\mathcal{S}_{\max}$ here, to be the collection

$$\mathcal{S}_{\max} = \{S \Delta N \mid S \in \mathcal{S}, N \in \mathcal{N}_L\}.$$

It is easy to show that $\mathcal{S}_{\max}$ is a σ-algebra; we leave this to the reader.

Theorem 2.7. *Let $\mathcal{S}$ be a sub-σ-algebra of $\mathcal{B}_L$ and let $\mathcal{S}_{\max}$ be its completion. Then any sub-σ-algebra $\mathcal{T}$ which is essentially equivalent to $\mathcal{S}$ is essentially equivalent to $\mathcal{S}_{\max}$. Furthermore, $\mathcal{T} \subset \mathcal{S}_{\max}$. $\mathcal{S}_{\max}$ is accordingly the largest sub-σ-algebra in the equivalence class of $\mathcal{S}$ under the essential equivalence relation.*

Proof. Assume $T \in \mathcal{T}$ and $\mathcal{T}$ essentially equivalent to $\mathcal{S}$. We want to show that $T \in \mathcal{S}_{\max}$, the completion of $\mathcal{S}$. There is a set S in $\mathcal{S}$ such that $\lambda(T\Delta S) = 0$. Set $N = T\Delta S$. Then

$$N\Delta S = T\Delta S\Delta S = T,$$

so $T \in \mathcal{S}_{\max}$. In addition, if $S \in \mathcal{S}_{\max}$, we want to show there is a set $T \in \mathcal{T}$ such that $\lambda(S\Delta T) = 0$. Since $S \in \mathcal{S}_{\max}$ and $\mathcal{S}_{\max}$ is the completion of $\mathcal{S}$, there is a set $\tilde{S} \in \mathcal{S}$ such that $S\Delta\tilde{S} \in \mathcal{N}_L$, and since $\mathcal{T}$ is essentially equivalent to $\mathcal{S}$, there is a set $T \in \mathcal{T}$ such that $\lambda(\tilde{S}\Delta T) = 0$. Set $N = \tilde{S}\Delta T$. Then N is a null set and $T\Delta N = T\Delta\tilde{S}\Delta T = \tilde{S}$, whence $S\Delta\tilde{S} = S\Delta T\Delta N$ and $S\Delta\tilde{S}\Delta N = S\Delta T\Delta N\Delta N = S\Delta T$, using properties of the symmetric difference. It follows that $S\Delta T \subset (S\Delta\tilde{S}) \cup N$, whence

$$\lambda(S\Delta T) \le \lambda(S\Delta\tilde{S}) + \lambda(N) = 0.$$

We conclude that for every $S \in \mathcal{S}_{\max}$ there exists a $T \in \mathcal{T}$ such that $\lambda(S\Delta T) = 0$. Since also $\mathcal{T} \subset \mathcal{S}_{\max}$, it follows that $\mathcal{T}$ is essentially equivalent to $\mathcal{S}_{\max}$. $\boxed{QED}$

The foregoing theorem justifies the use of the term $\mathcal{S}_{\max}$ for the completion of $\mathcal{S}$: $\mathcal{S}_{\max}$ is the maximal element in the equivalence class of $\mathcal{S}$; all other sub-σ-algebras in the equivalence class are subsets of $\mathcal{S}_{\max}$.

Corollary 2.7.1. *Let $\mathcal{S}$ be a sub-σ-algebra of $\mathcal{B}_L$. There exists a sub-σ-algebra $\mathcal{T}$ consisting solely of Borel sets which is essentially equivalent to $\mathcal{S}$.*

Proof. Given $\mathcal{S}$, form $\mathcal{S}_{\max}$, the completion of $\mathcal{S}$. By Theorem 2.7, $\mathcal{S}_{\max}$ is essentially equivalent to $\mathcal{S}$. Then set $\mathcal{T} = \mathcal{S}_{\max} \cap \mathcal{B}$, where $\mathcal{B}$ denotes the Borel sets. Since the intersection of σ-algebras of subsets of $[0,1]$ is necessarily a σ-algebra, $\mathcal{T}$ is a σ-algebra, and it consists solely of Borel sets. We will show that $\mathcal{T}$ is essentially equivalent to $\mathcal{S}_{\max}$, hence by Theorem 2.7 also to $\mathcal{S}$. To that end, consider $S \in \mathcal{S}_{\max}$. We want to show that there is a set $T \in \mathcal{T}$ such that $\lambda(S\Delta T) = 0$. If $\lambda(S) = 0$, take $T = \phi$ and observe that $\lambda(S\Delta T) = \lambda(S) = 0$. If $\lambda(S) > 0$, observe that since Lebesgue measure is a regular Borel measure, we can find a Borel set B such that $S \subset B$ and $\lambda(B) = \lambda(S)$, so that $N = B \setminus S$ is a Lebesgue null set. Since, by construction, every Lebesgue null set is an element of $\mathcal{S}_{\max}$, and since $\mathcal{S}_{\max}$ is closed under unions, necessarily the set $B = S \cup N$ is in $\mathcal{S}_{\max}$, hence is in $\mathcal{T}$, the intersection of $\mathcal{B}$ and $\mathcal{S}_{\max}$. Hence, we may take $T = B$, and we have $\lambda(S\Delta T) = \lambda(N) = 0$. $\boxed{QED}$

We propose not to address the issues addressed in Theorems 2.5, 2.6 and 2.7 at any later point in this paper. We accordingly adopt two conventions. First, whenever we talk about a sub-σ-algebra $\mathcal{S}$ or $h^{-1}(\mathcal{B})$, we really mean the equivalence class of all essentially equivalent sub-σ-algebras. Second, when we say a function ψ is measurable with respect to a σ-algebra $\mathcal{S}$, we really mean either that ψ is measurable with respect to some σ-algebra essentially equivalent to $\mathcal{S}$, or, equivalently (in view of Theorem 2.5), that some function equal to ψ

almost everywhere is measurable with respect to $\mathcal{S}$. We will also when addressing the order relations among equivalence classes, as in Theorems 6.3 and 6.4, and in particular in Corollary 6.4.2, take as the representative of an equivalence class the complete sub-σ-algebra in the class, since the order relation among equivalence classes holds if and only if the subset relation holds for complete sub-σ-algebras in the classes.

Invariant sets and fixed points of Markov operators. In Section 1, we defined invariant sets of a copula A as the sets S for which $p_A(x, S) = \chi_S(x)$ for a.a. $x \in S$, where

$$p_A(x, S) = \frac{d}{dx} \int_S A_{,2}(x, t)\, dt$$

is the family of transition probabilities obtained from A. It will be convenient here to work with invariant sets of Markov operators. A set S is an invariant set of a Markov operator T on $L^1([0, 1])$ if S is measurable and $T\chi_S = \chi_S$ a.s., that is, if χ_S is a fixed point of T. We turn to a discussion of fixed points and invariant sets of Markov operators.

A linear operator on L^1 is called a Markov operator if

(1) for all $\psi \geq 0$, $T\psi \geq 0$,

(2) for all ψ, $\int_0^1 T\psi\, d\lambda = \int_0^1 \psi\, d\lambda$, and

(3) the constant function $\psi = 1$ is a fixed point of T: $T1 = 1$.

It follows from these conditions that a Markov operator T is necessarily a bounded linear operator on $L^1([0, 1])$. The proof is trivial and is omitted.

Given a copula A, we define an operator T_A via

$$[T_A\psi](x) = \frac{d}{dx} \int_0^1 A_{,2}(x, t)\psi(t)\, dt.$$

T_A is a Markov operator, [3]. Observe that, by inspection of the definitions of p_A and T_A, a measurable set S satisfies the equation $p_A(x, S) = \chi_S(x)$ a.s. if and only if it satisfies the equation $T_A\chi_S = \chi_S$ a.s.

Theorem 2.8. *Define a function Φ from the set of 2-copulas to the set of Markov operators on $L^1([0, 1])$ via $\Phi(A) = T_A$, where T_A is the Markov operator associated with A per the definition above. Then Φ is a one-to-one and onto map. Furthermore,*

(1) $\Phi(A * B) = \Phi(A) \circ \Phi(B)$

(2) $\Phi(aA + (1 - a)B) = a\Phi(A) + (1 - a)\Phi(B)$ *for all* $a \in (0, 1)$ *and*

(3) $\Phi(A^T) = \Phi(A)^\dagger$,

for all copulas A, B. Finally, if A_n is a sequence of copulas and $T_n = \Phi(A_n)$ is the corresponding sequence of Markov operators, then $A_n \to A$ uniformly if and only if for all $\psi \in L^\infty([0, 1])$ and $\theta \in L^1([0, 1])$,

$$\int_0^1 \psi(x)[T_n\theta](x)\, dx \to \int_0^1 \psi(x)[T_A\theta](x)\, dx,$$

where $T_A = \Phi(A)$.

Proof. Omitted. This combines Theorems 2.1 and 3.1 of [3]. QED

Theorem 2.9. *The invariant sets of a Markov operator T constitute a sub-σ-algebra of the measurable subsets of* $[0, 1]$.

Proof. Let S be invariant. Then S^c is invariant. To see this, observe that since the constant function 1 is invariant, and $\chi_S + \chi_{S^c} = 1$, χ_{S^c} is the difference of fixed points, hence itself a fixed point. Next, suppose S_1 and S_2 are invariant sets of T, and consider $S = S_1 \cap S_2$. We have $\chi_{S_k} - \chi_S \geq 0$ for $k = 1$, 2, whence

$$0 \leq T(\chi_{S_k} - \chi_S) = \chi_{S_k} - T\chi_S,\ k = 1,\ 2$$

using one of the properties of Markov operators and the fact that S_1 and S_2 are invariant sets. It follows that for almost all x

$$[T\chi_S](x) \leq \min(\chi_{S_1}(x), \chi_{S_2}(x)) = \chi_S(x),$$

whence $\chi_S - T\chi_S \geq 0$ a.s. By another property of Markov operators,

$$\int_0^1 (\chi_S - T\chi_S)\, d\lambda = 0.$$

Since the integrand is nonnegative a.s., it must vanish a.s. Therefore $T\psi_S = \psi_S$ a.s., and S is an invariant set. If S_1 and S_2 are invariant sets and $S = S_1 \cup S_2$, we can write $S = S_1 \cup (S_2 \setminus S_1)$. Since this is a disjoint union, $\chi_S = \chi_{S_1} + \chi_{S_2 \setminus S_1}$. Thus, χ_S is a sum of fixed points, hence a fixed point, using the fact that $S_2 \setminus S_1 = S_2 \cap S_1^c$ is an invariant set by the two foregoing results. By induction, finite unions of invariant sets are invariant sets. That countable unions of invariant sets are invariant follows from the fact that T is a bounded linear operator. We omit the argument.

Since the invariant sets of T are closed under complementation, intersection and countable unions, they constitute a σ algebra. $\boxed{QED}$

We will call the collection of invariant sets of a Markov operator T and a copula A $\mathcal{B}_T$ and $\mathcal{B}_A$, respectively.

The following theorem is the principal motivation for moving the discussion of invariant sets to the Markov operator context.

Theorem 2.10. *Let T be a Markov operator on $L^1[0,1]$. Let $\mathcal{B}_T$ denote the invariant sets of T. A function ψ in L^1 is a fixed point of T if and only if ψ is $\mathcal{B}_T$-measurable, that is, if and only if $\psi^{-1}(B) \in \mathcal{B}_T$ for all $B \in \mathcal{B}$.*

Proof. If ψ is $\mathcal{B}_T$ measurable, we can find a sequence of $\mathcal{B}_T$ simple functions converging to ψ pointwise almost everywhere. Since each such simple function is a linear combination of characteristic functions χ_S for $S \in \mathcal{B}_T$, and each such characteristic function is a fixed point of T, it follows that each such simple function is a fixed point of T and hence that their pointwise a.e. limit ψ is a fixed point.

Conversely, suppose ψ is a fixed point of T. We show first that if ψ is a fixed point, then necessarily so also are ψ^+ and ψ^- given by

$$\psi^+(x) = \max\ \{\psi(x), 0\} \text{ and}$$
$$\psi^-(x) = \max\ \{-\psi(x), 0\}.$$

To see this, observe that $T\psi^+$ and $T\psi^-$ are both nonnegative, since Markov operators map nonnegative functions into nonnegative functions, and that also $T\psi^+ - T\psi^- = T\psi = \psi$ a.s. It follows that for x for which $\psi(x) \geq 0$ we have

$$[T\psi^+](x) = \psi(x) + [T\psi^-](x) \geq \psi(x) = \psi^+(x).$$

For x for which $\psi(x) < 0$,

$$[T\psi^+](x) \geq 0 = \psi^+(x).$$

For all x, therefore, $[T\psi^+](x) \geq \psi^+(x)$. Now

$$\int_0^1 (T\psi^+ - \psi^+)\, d\lambda = 0,$$

and since the integrand is a.s. nonnegative, it must in fact vanish almost everywhere. Thus, ψ^+ is a fixed point. Then also $\psi^- = \psi^+ - \psi$ is a fixed point, since linear combinations of fixed points are fixed points.

Now consider the set $S = \{\psi > a\}$. The function $\theta = [\psi - a]^+$ is a fixed point of T, since it is the positive part of the difference of two fixed points, ψ and the constant function a. The function θ vanishes outside of S, and $S = \{\theta > 0\}$. Define $\theta_n(x) = n\theta(x) - [n\theta - 1]^+(x)$, multiplying θ by a factor of n, then replacing values greater than 1 by 1. The function θ_n is a fixed point of T, since $n\theta$, $n\theta - 1$ and, by the result above, $[n\theta - 1]^+$ are all fixed points. Furthermore, θ_n vanishes outside S, because θ does, and $\theta_n(x) \uparrow 1$ for all $x \in S$. It follows via the monotone convergence theorem that $\|\chi_S - \theta_n\|_1 \to 0$. Since T is a bounded operator,

$$\|T\chi_S - \chi_S\| \leq \|T\chi_S - T\theta_n\| + \|\theta_n - \chi_S\| \leq (\|T\| + 1)\|\chi_S - \theta_n\| \to 0.$$

Thus, $T\chi_S = \chi_S$ a.s., and $S \in \mathcal{B}_T$. Since intervals of the form (a, ∞) generate the Borel subsets, $\psi^{-1}(B) \in \mathcal{B}_T$ for all B. Thus, ψ is $\mathcal{B}_T$ measurable. $\boxed{QED}$

In the following theorem, and from time to time thereafter, we work with Markov operators derived from copulas of the form C_{ef} and C_{fe}. The Markov operators associated with such copulas will be written T_{ef} and T_{fe} instead of $T_{C_{ef}}$ and $T_{C_{fe}}$, in order to avoid double subscripting.

Theorem 2.11. *Let $f \in \mathcal{F}$ be a measure preserving Borel function. For all Borel measurable $\psi \in L^1([0,1])$,*

$$[T_{ef}\psi](x) = \psi \circ f(x),$$

for almost all $x \in [0,1]$. In addition, if for $\psi \in L^1([0,1])$ there exists an integrable function θ such that $\psi = \theta \circ f$ a.s., then

$$[T_{fe}\psi](x) = \theta \text{ a.s.}$$

Proof. If $\psi \in C^\infty([0,1])$, then

$$\begin{aligned}
[T_{ef}\psi](x) &= \frac{d}{dx}\int_0^1 C_{ef,2}(x,t)\psi(t)\, dt \\
&= \frac{d}{dx}(C_{ef}(x,1)\psi(1) - \int_0^1 C_{ef}(x,t)\psi'(t)\, dt) \\
&= \psi(1) - \int_0^1 \chi_{[f(x),1]}(t)\psi'(t)\, dt \\
&= \psi(1) - (\psi(1) - \psi(f(x))) \\
&= \psi(f(x)).
\end{aligned}$$

This uses a characterization of $C_{ef,1}$ from Theorem 2.3, part (1). Since C^∞ is dense in L^1 and f is measure preserving, we have the first result of the theorem. (The measure preserving property is used here, since if ψ is integrable and $\|\psi_n - \psi\|_1 \to 0$, then $\|\psi_n \circ f - \psi \circ f\|_1 \to 0$, when f is measure preserving but not necessarily otherwise.) For the second part of the theorem, observe that if $\psi \in L^1([0,1])$ has the form $\psi = \theta \circ f$ for some integrable Borel function θ, then

$$\begin{aligned}
[T_{fe}\psi](x) &= \frac{d}{dx}\int_0^1 C_{fe,2}(x,t)\psi(t)\,dt \\
&= \frac{d}{dx}\int_0^1 \chi_{[0,x]}(f(t))\theta(f(t))\,dt \\
&= \frac{d}{dx}\int_0^1 \chi_{[0,x]}(s)\theta(s)\,ds \\
&= \theta(x), \text{ for a.a. } x.
\end{aligned}$$

This likewise uses a characterization of $C_{fe,2}$ from Theorem 2.3. It also uses fact that f is measure preserving. $\boxed{QED}$

Corollary 2.11.1. *If $f \in \mathcal{F}$ is a measure preserving Borel function, then every set $S \in f^{-1}(\mathcal{B})$ is an invariant set of $C_{ef}C_{fe}$, and every function measurable with respect to $f^{-1}(\mathcal{B})$ is a fixed point of $T_{ef}T_{fe}$. In particular, if θ is any Borel function, $\psi = \theta \circ f$ is a fixed point of $T_{ef}T_{fe}$.*

Proof. If $S \in f^{-1}(\mathcal{B})$, let $B \in \mathcal{B}$ be such that $S = f^{-1}(B)$. Then observe that $\chi_S = \chi_{f^{-1}(B)} = \chi_B \circ f$, so that by Theorem 2.11,

$$T_{ef}T_{fe}\chi_S = T_{ef}T_{fe}\chi_B \circ f = T_{ef}\chi_B = \chi_B \circ f = \chi_S \text{ a.s.}$$

Since

$$T_{ef}T_{fe} = T_{C_{ef}C_{fe}},$$

and since an invariant set of $C_{ef}C_{fe}$ is an invariant set of the corresponding Markov operator, S is an invariant set of $C_{ef}C_{fe}$. Furthermore, this argument shows that the σ algebra $f^{-1}(\mathcal{B})$ is a sub-σ-algebra of $\mathcal{B}_T$, the invariant sets of $T = T_{ef}T_{fe}$. Since By Theorem 2.9, an integrable function ψ is a fixed point of T if and only if $\psi^{-1}(B) \in \mathcal{B}_T$ for every Borel set B, those functions ψ measurable with respect to $f^{-1}(\mathcal{B})$, that is, those functions ψ for which $\psi^{-1}(B) \in f^{-1}(\mathcal{B})$ for every Borel set B, are among the fixed points of T. Finally, if $\psi = \theta \circ f$, then for any Borel set B,

$$\psi^{-1}(B) = f^{-1}(\theta^{-1}(B)) \in f^{-1}(\mathcal{B}),$$

since $\theta^{-1}(B)$ is a Borel set. It follows that ψ is a fixed point of T. $\boxed{QED}$

We remark that we do not know (yet) that $\mathcal{B}_T = f^{-1}(\mathcal{B})$ essentially, only that $f^{-1}(\mathcal{B}) \subset \mathcal{B}_T$.

3 Nonatomic idempotent copulas

If we are given a nonatomic idempotent, and we want to show that it can be written in the form $C_{ef}C_{fe}$, as we do want to do, we must somehow be able to obtain a measure preserving function f for which this representation will hold. The problem is similar to that addressed in Theorem 2.4, where, given a copula known to have a left inverse, we had to demonstrate how to obtain a measure preserving function f for which the representation C_{ef} would hold. In this section, we first show that all sub-σ-algebras $\mathcal{S}$ obtained from measure preserving Borel functions f via $\mathcal{S} = f^{-1}(\mathcal{B})$ are nonatomic. Then we turn to the more difficult problem of pulling an appropriate measure preserving Borel function out of the hat, when all we know is that $\mathcal{S}$ is a nonatomic sub-σ-algebra.

Theorem 3.1. *If $h \in \mathcal{F}$ is a measure preserving Borel function, then $h^{-1}(\mathcal{B})$ is nonatomic.*

Proof. Let $S \in h^{-1}(\mathcal{B})$ have positive measure. There is a Borel set $B \in \mathcal{B}$ such that $S = h^{-1}(B)$, and since h is measure preserving, $\lambda(S) = \lambda(B)$, hence B has positive measure. The collection $\mathcal{B}$ of all Borel sets is known to be nonatomic. One can prove this by elementary means, by considering, for any set B of positive Borel measure, the sets $[0, x] \cap B$ for all x in $[0, 1]$. The sets are all Borel sets, and their measures vary continuously from 0 to $\lambda(B)$ as x varies from 0 to 1, by the basic properties of Lebesgue measure. Thus, there must exist a Borel subset $A \subset B$ satisfying $0 < \lambda(A) < \lambda(B)$. Then $h^{-1}(A) \subset S$, and since h is measure preserving, $0 < \lambda(h^{-1}A) = \lambda(A) < \lambda(S)$, so S is not an atom. Since S was an arbitrary set of positive measure, $h^{-1}(\mathcal{B})$ is nonatomic. $\boxed{QED}$

The following theorem is due to Carathéodory; the statement and proof of a slightly more general result can be found in [5], Theorem 4, Chapter 15.

Theorem 3.2. *Let $\mathcal{S}$ be a sub-σ-algebra of the Borel subsets $\mathcal{B}$ of $[0, 1]$, with null sets $\mathcal{N}_S$. If $\mathcal{S}$ is nonatomic, there exists a one-to-one and onto measure preserving set function $\Phi : \mathcal{S}/\mathcal{N}_S \to \mathcal{B}/\mathcal{N}_B$ which preserves order, complementation and the lattice operation on equivalence classes corresponding to countable unions of monotonic sequences of sets.*

Here, $\mathcal{N}_S = \{S \in \mathcal{S} \,|\, \lambda(S) = 0\}$ and $\mathcal{N}_B = \{S \in \mathcal{B} \,|\, \lambda(S) = 0\}$. The quotient space $\mathcal{S}/\mathcal{N}_S$ is the set of equivalence classes under the equivalence relation $S_1 \sim S_2$ if $\lambda(S_1 \Delta S_2) = 0$, that is, if $S_1 \Delta S_2 \in \mathcal{N}_S$. The order relation on equivalence classes is that inherited from set inclusion. If $E(S_k)$ denotes the equivalence class of S_k, then $E(S_1) \le E(S_2)$ if $\lambda(S_1 \setminus S_2) = 0$, which is true if and only if, modulo a null set, $S_1 \subset S_2$.

Theorem 3.3. *Let $\mathcal{S}$ be a sub-σ-algebra of the Borel subsets $\mathcal{B}$ of $[0, 1]$, with null sets $\mathcal{N}_S$. If $\mathcal{S}$ is nonatomic, there exists a measure preserving Borel function $h : [0, 1] \to [0, 1]$ such that $h^{-1}(\mathcal{B}) \subset \mathcal{S}$. Furthermore, h has the property that for each $S \in \mathcal{S}$ there exists a set $S_0 \in h^{-1}(\mathcal{B})$ for which $\lambda(S \Delta S_0) = 0$, so that $\mathcal{S}$ is essentially equivalent to $h^{-1}(\mathcal{B})$.*

Proof. We give an outline of the proof. Let r_n be an enumeration of the rational numbers in $[0, 1]$, set $I_n = [0, r_n]$, and write $E(I_n)$ for the equivalence class of I_n. Let $\Phi : \mathcal{S}/\mathcal{N}_S \to \mathcal{B}/\mathcal{N}_B$ be a one-to-one and onto measure preserving set function which preserves order, complementation and the lattice operation corresponding to countable unions of nested sets, per Theorem 3.2 above. For each n, choose $S_n \in \Phi^{-1}(E(I_n))$ ($\Phi^{-1}(E(I_n))$ is an equivalence class), and arrange it so that $r_k < r_j$ implies $S_k \subset S_j$. This can be done, for example, by an inductive process in which, if S_1 through S_{n-1} satisfy the inclusion condition, S_n is adjusted by adjoining to S_n the sets $S_k \setminus S_n$, necessarily Borel null sets, for $r_k < r_n$, and deleting from S_n the sets $S_n \setminus S_k$, for $r_k > r_n$. When $r_k = 1$, arrange it so that $S_k = [0, 1]$, so that every x is in some S_k. Define $h(x) = \inf\{r_k | x \in S_k\}$. For all $x \in S_k$, $h(x) \le r_k$, and for all $x \notin S_k$, $h(x) \ge r_k$. It follows that for any real $a \in [0, 1]$, $\{h < a\} = \cup_{r_k < a} S_k$, hence that h is a Borel function. It also follows that $S_\ell \subset \{h < r_k\} \subset S_k$ whenever $r_\ell < r_k$, hence, by an argument similar to one used in the proof of Theorem 2.4, that $\lambda(\{h < r_k\}) = \lambda(S_k)$. The function h thus inherits the measure preserving property from Φ: since $S_k \in \Phi^{-1}(E(I_k))$, $\lambda(S_k) = \lambda([0, r_k]) = r_k$ for all k, and so $\lambda(\{h < r_k\}) = r_k$. It remains to show that for any $S \in \mathcal{S}$ we can find an $S_0 \in h^{-1}(\mathcal{B})$ such that $\lambda(S \Delta S_0) = 0$, that is, for which $S_0 \in E(S)$, where $E(S)$ denotes the equivalence class of S. For $S \in \mathcal{S}$, choose $B \in \Phi(E(S))$ and set $S_0 = h^{-1}(B)$. Observe that since Φ is one-to-one, we can write $E(S) = \Phi^{-1}(E(B))$, and what we want to establish is that $h^{-1}(B) \in \Phi^{-1}(E(B))$. Consider the class $\mathcal{M}$ of all Borel sets C for which $h^{-1}(C) \in \Phi^{-1}(E(C))$. If $C \in \mathcal{M}$, then we claim that its complement C^c is in $\mathcal{M}$. This follows from the fact that $E(C^c) = E(C)^c$, which says that two sets are equivalent if and only if their complements are equivalent, which is trivial to verify, and the fact that Φ preserves complementation (or rather the analog of complementation on equivalence classes, which we are also denoting by superscript c). The formal argument is

$$\begin{aligned} E(h^{-1}(C^c)) &= E((h^{-1}C)^c) = E(h^{-1}C)^c \\ &= \Phi^{-1}(E(C))^c = \Phi^{-1}(E(C)^c) = \Phi^{-1}(E(C^c)), \end{aligned}$$

which proves the claim. Similarly, if C_n in $\mathcal{M}$, $C_n \subset C_{n+1}$ and $\cup_n C_n = C$, then we claim $C \in \mathcal{M}$. In this case, since the lattice operations on sets of equivalence classes are preserved for countable unions of sets of nested sets, we have

$$\sup_n (E(C_n)) = E(\cup_n C_n)$$

which says that if C_n and $\tilde{C}_n$ are equivalent for all n, then $\cup_n C_n$ is equivalent to $\cup_n \tilde{C}_n$. This again is trivial to verify. It follows that

$$\begin{aligned} E(h^{-1}C) &= E(h^{-1}(\cup_n C_n)) = E(\cup_n(h^{-1}C_n)) = \sup_n E(h^{-1}C_n) \\ &= \sup_n \Phi^{-1}(E(C_n)) = \Phi^{-1}(\sup_n E(C_n)) = \Phi^{-1}(E(\cup_n C_n)) = \Phi^{-1}(E(C)). \end{aligned}$$

Thus, $C \in \mathcal{M}$, as claimed. Because of these facts, $\mathcal{M}$ is a monotone class, and since it contains $[0, r_k]$ for all k, it contains all Borel sets, by a standard argument. Thus, $h^{-1}(B) \in \Phi^{-1}(E(B))$. This says that $S_0 \in E(S)$, which was to be proved. $\boxed{QED}$

Theorem 3.4. *If F is a nonatomic idempotent copula with invariant sets $\mathcal{B}_F$, and f is a measure preserving Borel function for which $\mathcal{B}_F = f^{-1}(\mathcal{B})$ essentially, then $F = C_{ef}C_{fe}$.*

Proof. By Theorem 2.10, the fixed points of the Markov operator T_F corresponding to F are those functions which are measurable with respect to $\mathcal{B}_F = f^{-1}(\mathcal{B})$ essentially, among which, by an argument set forth in the proof of Corollary 2.11.1, are functions of the form $\psi = \theta \circ f$, where θ is an integrable Borel function. By Theorem 2.11, for any integrable Borel function θ, $T_{ef}\theta = \theta \circ f$ a.s. and $T_{fe}\theta \circ f = \theta$ a.s. Thus, for all integrable Borel functions θ,

$$T_{fe}T_F T_{ef}\theta = \theta \text{ a.s.}$$

Since the Borel functions are dense in L^1 (all continuous functions are Borel functions, for example, and the continuous functions are dense), it follows that $T_{fe}T_F T_{ef}$ is the identity operator on $L^1([0,1])$, hence, by the canonical isomorphism of copulas and their corresponding Markov operators, Theorem 2.8, that $C_{fe}FC_{ef} = M$. Now one-sided inverses with respect to M, when they exist, are unique, Theorem 2.1, whence we must have $C_{fe}F = C_{fe}$, since both $C_{fe}F$ and C_{fe} are left inverses of C_{ef}. Pre-multiply this relation by C_{ef} to obtain

$$C_{ef}C_{fe}F = C_{ef}C_{fe}.$$

On the other hand, since $T_F\psi$ is a fixed point of T_F (using the fact that $T_F^2 = T_F$), it must by Theorem 2.10 be measurable with respect to $\mathcal{B}_F = f^{-1}(\mathcal{B})$ essentially, so that, by Corollary 2.11.1, it is a fixed point of $T_{ef}T_{fe}$. Thus

$$T_{ef}T_{fe}T_F\psi = T_F\psi \text{ a.s.}$$

Since this holds for all integrable Borel functions ψ, we must have $T_{ef}T_{fe}T_F = T_F$. It follows from the isomorphism of copulas and their corresponding Markov operators that $C_{ef}C_{fe}F = F$. Since we showed above that $C_{ef}C_{fe}F = C_{ef}C_{fe}$, it must be true that $F = C_{ef}C_{fe}$. $\boxed{QED}$

As noted above, Theorem 2.11 does not assert that elements of $f^{-1}(\mathcal{B})$ are the only invariant sets of $C_{ef}C_{fe}$, only that such such sets are among the invariant sets of $C_{ef}C_{fe}$. But the fact that these are the only invariant sets does now follow from Theorem 3.4.

Corollary 3.4.1. *Let $\mathcal{S}$ be a nonatomic sub-σ-algebra of the Borel subsets of* $[0,1]$. *If f is a measure preserving Borel function for which $\mathcal{S} = f^{-1}(\mathcal{B})$ essentially, the invariant sets of the idempotent copula $C_{ef}C_{fe}$ are $\mathcal{S}$ essentially.*

Proof. Given a nonatomic sub-σ-algebra $\mathcal{S}$, there exists a unique idempotent Markov operator whose invariant sets are essentially $\mathcal{S}$, namely the conditional expectation operator $T : \psi \to E(\psi \,|\, \mathcal{S})$, [8]. Let F be the idempotent copula associated to T under the correspondence of Theorem 2.8. Then $\mathcal{B}_F = \mathcal{S}$ essentially, since F and T have the same invariant sets. By Theorem 3.4, $F = C_{ef}C_{fe}$. The invariant sets of $C_{ef}C_{fe}$ are thus essentially the given nonatomic sub-σ-algebra $\mathcal{S}$. $\boxed{QED}$

In fact, in [8], Sempi demonstrates that conditional expectation operators on $L^1[0,1]$ are in one-to-one correspondence with idempotent Markov operators on $L^1[0,1]$, hence also with idempotent 2-copulas. In effect, we address here a question left open at the end of [8] – how to obtain an explicit expression for the idempotent copula whose invariant sets are the given sub-σ-algebra $\mathcal{S}$. Corollary 3.4.1 gives the answer to this question for nonatomic $\mathcal{S}$.

Observe that the measure preserving function f in Corollary 3.4.1 is not unique; it is only the idempotent copula $C_{ef}C_{fe}$ which is uniquely determined by $\mathcal{S}$. Cf. Corollary 2.4.2 above.

4 Totally atomic idempotent copulas

A totally atomic idempotent copula is conjugate to an ordinal sum consisting of copies of the product copula P. In this section, we define terminology and prove this result.

It was noted Section 1 that the product copula P is a totally atomic copula. Indeed, its invariant sets are the sets whose measure is 0 or 1. We first show that it is the only such idempotent copula.

Theorem 4.1. *Let E be an idempotent copula whose invariant sets are sets whose measure is 0 or 1. Then $E = P$.*

Proof. Let T_E be the Markov operator associated with E. For all $\psi \in L^1$, $T_E\psi$ is a fixed point of T_E, since $T_E^2 = T_E$, hence by Theorem 2.10 $T_E\psi$ is $\mathcal{B}_E$ measurable. Since $\mathcal{B}_E = \{\phi, [0,1]\}$ essentially, $T_E\psi$ is an essentially constant function, regardless of ψ. Set $\psi = \chi_{[0,y]}$ and let $\kappa(y)$ denote the a.s. constant value of $T_E\chi_{[0,y]}$. Then

$$\kappa(y) = T_E\chi_{[0,y]}(x) = \frac{d}{dx}\int_0^y E_{,2}(x,t)\,dt = E_{,1}(x,y).$$

Integrate this expression from 0 to x to obtain $E(x,y) = x\kappa(y)$. Since $E(1,y) = y$, $\kappa(y) = y$ and $E(x,y) = xy$. $\boxed{QED}$

We show next that if an idempotent copula E is totally atomic, and its atoms are intervals (a_k, a_{k+1}), the sum of whose lengths is 1, then E is an ordinal sum of copies of the product copula. We will then address the general case.

It is easy to verify that a copula C has an ordinal sum decomposition on a partition $\mathcal{P} : \{(a_k, b_k)\}$ if and only if $C(a_k, a_k) = a_k$ and $C(b_k, b_k) = b_k$ for all k. If this condition is satisfied, the copulas A_k assigned to intervals (a_k, b_k) are uniquely determined by the copula C

which possesses the ordinal sum decomposition. We leave it to the reader to verify the validity of the condition.

Two properties of ordinal sums are used here. First, if $A = \oplus_{\mathcal{P}} A_k$ and $B = \oplus_{\mathcal{P}} B_k$ are ordinal sums on the same partition $\mathcal{P}$, then $A * B = \oplus_{\mathcal{P}} A_k * B_k$. That is, $A * B$ is also an ordinal sum on the partition and its components are $A_k * B_k$. Thus, in particular, an ordinal sum $E = \oplus_{\mathcal{P}} E_k$ is idempotent if and only if each of the components E_k is idempotent. The verification is again left to the reader. The second property we state as a lemma:

Lemma 4.1. *Suppose that A has an ordinal sum decomposition $A = \oplus_{\mathcal{P}} A_k$. Then every interval (a_k, b_k) in $\mathcal{P}$ is invariant under A. Furthermore, a set $Q \subset (a_k, b_k)$ is invariant under A if and only if*

$$S = \{\frac{x - a_k}{b_k - a_k} | x \in Q)\}$$

is invariant under A_k.

Proof. If $x \in (a_k, b_k)$,

$$\begin{aligned}
[T_A \chi_{(a_k,b_k)}](x) &= \frac{d}{dx}\int_0^1 A_{,2}(x,t)\chi_{(a_k,b_k)}(t)\,dt \\
&= \frac{d}{dx}\int_{a_k}^{b_k} A_{k,2}(\frac{x-a_k}{b_k-a_k}, \frac{t-a_k}{b_k-a_k})\,dt \\
&= (b_k - a_k)\frac{d}{dx}\int_0^1 A_{k,2}(\frac{x-a_k}{b_k-a_k}, s)\,ds \\
&= (b_k - a_k)\frac{d}{dx}(\frac{x-a_k}{b_k-a_k}) \\
&= 1,
\end{aligned}$$

where the substitution $s = (t - a_k)/(b_k - a_k)$ was made to obtain the third expression on the right.

On the other hand, if $x \notin [a_k, b_k]$, then

$$\begin{aligned}
[T_A \chi_{(a_k,b_k)}](x) &= \frac{d}{dx}\int_0^1 A_{,2}(x,t)\chi_{(a_k,b_k)}(t)\,dt \\
&= \frac{d}{dx}\int_{a_k}^{b_k} M_{,2}(x,t)\,dt \\
&= \frac{d}{dx}(M(x,b_k) - M(x,a_k)) \\
&= 0,
\end{aligned}$$

for if $x \notin [a_k, b_k]$, either $x < a_k$, whence $M(x, b_k) - M(x, a_k) = x - x = 0$, or $b_k < x$, whence $M(x, b_k) - M(x, a_k) = b_k - a_k$, so that in each case the derivative vanishes. It follows that (a_k, b_k) is invariant under A.

Now let $Q \subset (a_k, b_k)$ and write

$$S = \frac{1}{b_k - a_k}(Q - a_k).$$

Observe that $\xi \in S$ iff $x = a_k + (b_k - a_k)\xi \in Q$, or equivalently $x \in Q$ iff $\xi = (x - a_k)/(b_k - a_k) \in$

S. For all x we have

$$\begin{aligned}
[T_A\chi_Q](x) &= \frac{d}{dx}\int_0^1 A_{,2}(x,t)\chi_Q(t)\,dt \\
&= \frac{d}{dx}\int_{a_k}^{b_k} A_{k,2}(\frac{x-a_k}{b_k-a_k},\frac{t-a_k}{b_k-a_k})\chi_Q(t)\,dt \text{ (since } Q\subset(a_k,b_k)) \\
&= (b_k-a_k)\frac{d}{dx}\int_0^1 A_{k,2}(\frac{x-a_k}{b_k-a_k},s)\chi_Q(a_k+(b_k-a_k)s)\,ds \\
&= (b_k-a_k)\frac{d}{dx}\int_0^1 A_{k,2}(\frac{x-a_k}{b_k-a_k},s)\chi_S(s)\,ds \\
&= \frac{d}{d\xi}\int_0^1 A_{k,2}(\xi,s)\chi_S(s)\,ds \text{ (substituting } \xi=(x-a_k)/(b_k-a_k)) \\
&= [T_{A_k}\chi_S](\xi).
\end{aligned}$$

¿From this it follows readily that Q is invariant under A if and only if S is invariant under A_k. $\boxed{QED}$

Note that it follows from Lemma 4.1 that a set S is invariant under Q if and only if $Q_k = S\cap(a_k,b_k)$ is invariant under A_k for all k.

We will say a partition $\mathcal{P} : (a_k,b_k)$ of $[0,1]$ is a special partition if $a_{k+1} = b_k$ for all $k>1$.

Theorem 4.2. *Let $\mathcal{P} : \{(a_k,a_{k+1})\}$ be a special partition of the interval $[0,1]$. If $E = \oplus_{\mathcal{P}}E_k$ and each $E_k = P$, then E is idempotent and the invariant sets of E are essentially intervals (a_k,a_{k+1}) in the partition, their unions and the empty set. Conversely, if E is an idempotent copula with the property that $\mathcal{B}_E$ consists essentially of intervals (a_k,a_{k+1}) in the partition, unions of them and the empty set, then $E=\oplus_{\mathcal{P}}E_k$ where each component E_k is the product copula P.*

Proof. If $E=\oplus_{\mathcal{P}}E_k$ and each $E_k = P$, the only sets invariant under E_k are sets whose measure is 0 or 1, by Theorem 4.1, hence by Lemma 4.1 the invariant sets of E are essentially intervals (a_k,b_k) in the partition, their unions, and the empty set. This proves the first assertion. For the converse, suppose E is an idempotent copula and its invariant sets are essentially intervals (a_k,a_{k+1}) in the partition $\mathcal{P}$, their unions, and the empty set. The plan is to show that E is then necessarily an ordinal sum on $\mathcal{P}$. For once that is established, it will necessarily follow that each component E_k is idempotent, and, by Lemma 4.1, that the only invariant sets of E_k are sets of measure 0 and 1. Then, by Theorem 4.1, each E_k must be the product copula P. Thus, the second assertion of the theorem will be proved if we can show that E has an ordinal sum decomposition on the special partition $\mathcal{P}$.

By a remark above, it will be sufficient to show that $E(a_k,a_k)=a_k$ for all k. To that end, observe that since (a_k,a_{k+1}) is an invariant set of E, we must have for all k

$$[T_E\chi_{(a_k,a_{k+1})}](x) = \chi_{(a_k,a_{k+1})}(x)$$

for a.a. x. This says that for all k

$$\frac{d}{dx}\int_0^1 E_{,2}(x,t)\chi_{(a_k,a_{k+1})}(t)\,dt = E_{,1}(x,a_{k+1}) - E_{,1}(x,a_k) = \chi_{(a_k,a_{k+1})}(x).$$

Integrate this expression from 0 to a_{k+1} to obtain, for all k,

$$E(a_{k+1},a_{k+1}) - E(a_{k+1},a_k) = a_{k+1}-a_k.$$

Now for $k = 1$, we have $a_1 = 0$, and $E(a_1, a_1) = 0 = a_1$, by the boundary conditions of a copula. Suppose that $E(a_n, a_n) = a_n$. Since also $E(1, a_n) = a_n$ by one of the copula boundary conditions, and since $x \to E(x, y)$ is nondecreasing, it necessarily follows that also $E(a_{n+1}, a_n) = a_n$. Then, by the equation above, $E(a_{n+1}, a_{n+1}) = a_{n+1}$. Accordingly, by induction, $E(a_k, a_k) = a_k$ for all k, so E is necessarily an ordinal sum on the special partition. $\boxed{QED}$

We address next rearrangements – Borel functions which rearrange the mass in the unit interval in an essentially one-to-one and onto manner. Rearrangements are the tool we use here to pull an appropriate measure preserving function out of the hat, so as to obtain the desired characterization of totally atomic idempotent copulas. Rearrangements are well known, but we could not find a rigorous proof of the existence of rearrangements with exactly the properties we need, so we offer a proof here.

Our main result depends on Lemma 4.2 below, whose proof is rather fussy. For a Borel set S, define a function f_S via

$$f_S(x) = \lambda([0, x] \cap S).$$

Then f_S maps $[0, 1]$ onto $[0, \lambda(S)]$ in a continuous and nondecreasing manner. The idea is to show that f_S can be modified in an appropriate manner to obtain a Borel function $h : S \to [0, \lambda(S)]$ which is measure preserving and which has an essential inverse g which is also a measure preserving Borel function. The construction is complicated by the fact that we have to take into account the possibility that S is a Cantor-like set. For a simple example, let r_n be an enumeration of the rational numbers in $[0, 1]$. Given $0 < \epsilon < 1$, set

$$S = (0, 1) \cap \{\cup_n (r_n - \frac{\epsilon}{2^{n+1}}, r_n + \frac{\epsilon}{2^{n+1}})\}.$$

Then $\lambda(S) < \epsilon$, as is easy to verify, and f_S is a strictly increasing function which maps $[0, 1]$ onto $[0, \lambda(S)]$, as is also easy to verify. Basically, if $x_1 < x_2$, there is a rational number r_n between them, and so

$$f_S(x_2) = f_S(x_1) + \lambda((x_1, x_2] \cap S) > f_S(x_1).$$

When we restrict f_S to S in this case, we delete from the domain a set of positive measure, and we simultaneously delete some nonempty set from the range. We must somehow account for this set in defining an essential inverse; to be specific, we must show that what remains of the range of f_S after deleting this set contains a suitable measurable Borel set of measure $\lambda(S)$ on which to define an essential inverse function. In the example this is trivial – f_S has a continuous, hence Borel measurable, inverse g, and $g^{-1}(S^C)$ is necessarily a Borel set. The essential inverse to the restriction of f_S to S can accordingly be defined on the Borel set $[0, \lambda(S)] \setminus g^{-1}(S^C)$. We found it difficult turn this into a general argument, and we use a somewhat more roundabout approach in the proof of the lemma.

Lemma 4.2. *Let S be a Borel subset of $[0, 1]$ with $\lambda(S) > 0$. There exist a measure preserving Borel function $h : S \to [0, \lambda(S)]$ and a measure preserving Borel function $g : [0, \lambda(S)] \to S$ such that $h \circ g = e$ essentially and $g \circ h = e$ essentially, where e is the identity function $e(x) = x$.*

Proof. Set $f_S(x) = \lambda([0, x] \cap S), x \in [0, 1]$. We claim that for all $t \in [0, \lambda(S)]$, $\lambda(S \cap f_S^{-1}[0, t)) = t$. To see this, let $t \in (0, \lambda(S))$ and write

$$a^- = \inf\{x | f(x) = t\}$$
$$a^+ = \sup\{x | f(x) = t\}.$$

Then $f_S(a^-) = f_S(a^+) = t$, by continuity of f_S. (Possibly $a^- = a^+$.) Also $f_S(x) < t$ iff $x < a^-$, so

$$\lambda(\{f_S < t\} \cap S) = \lambda([0, a^-) \cap S) = f_S(a^-) = t.$$

This establishes the claim. The program of the remainder of the proof is to modify f_S in such a way as to limit its domain to S, without destroying the property just established, so as to obtain a measure preserving function, and then to define an inverse function on the range of the modified version of f_S in an appropriate manner.

Since f_S is continuous and nondecreasing, it has a right quasi-inverse f_S^Q. That is, there exists $f_S^Q : [0, \lambda(S)] \to [0, 1]$ with the property that $f \circ f^Q = e$. Let D be the set of discontinuities of f_S^Q. For each $a \in D$, let $[a^-, a^+]$ be the jump interval of f_S^Q at a. Let

$$N_1 = \cup_{a\in D}\{[a^-, a^+] \cap S\}.$$

N_1 is a Borel set of measure 0, since $\lambda([a^-, a^+] \cap S) = f_S(a^+) - f_S(a^-) = 0$ for all a, and the number of discontinuities of f^Q is at most countable. Let h_0 be the restriction of f_S to $S \setminus N_1$. Then $h_0 : S \setminus N_1 \to [0, \lambda(S)]$ is one-to-one, since if $f_S(x_1) = f_S(x_2) = a$ for some $x_1 < x_2$, then $a \in D$, so neither x_1 nor x_2 is in the domain of h_0. The function h_0 is measure-preserving by the claim proved above.

It proved difficult to show directly that the range of h_0 is a measurable set (cf. remarks preceding the statement of the lemma). To complete the proof, therefore, we proceed in the following manner: Since Lebesgue measure is a regular Borel measure,

$$\lambda(S \setminus N_1) = \sup\{\lambda(K) \mid K \subset S \setminus N_1, K \text{ compact } \}.$$

We can find, therefore, a sequence of compact sets K_n with the properties that $K_n \subset S \setminus N_1$ for all n, $K_n \subset K_{n+1}$ for all n and $\lambda(K_n) \uparrow \lambda(S \setminus N_1)$. Set

$$N_2 = (S \setminus N_1) \setminus \{\cup_n K_n\}$$

and observe that N_2 is a Borel set of measure 0. We claim that h_0 maps closed sets into closed sets. To see this, let K be a closed subset of $S \setminus N_1$, and let $y_n \in h_0(K)$ be a convergent sequence with limit y. For each n there is a number $x_n \in K$ for which $h_0(x_n) = y_n$. Since K is compact, x_n possesses a convergent subsequence x_{n_k}, with limit, say, x. Then $f_S(x) = y$, by the continuity of f_S, and since $x \in K \subset S \setminus N_1$, x is in the domain of h_0, and we have also $h_0(x) = y$. Thus, y is in the set $h_0(K)$, as claimed.

Now set

$$B = \cup_n h_0(K_n).$$

Then B is a union of closed sets, hence a Borel set. Define g_0 on B via $g_0(y) = x$ where x is the unique element of $S \setminus N_1$ for which $h_0(x) = y$. Consider the sets $S_n(x) = [0, x] \cap K_n$. Since $g_0^{-1}(S_n(x))$ is closed for all n, it follows that $g_0^{-1}([0, x] \cap S) = g_0^{-1}([0, x] \cap S \setminus (N_1 \cup N_2))$ is a union of closed sets, hence a Borel set, for all x. Thus, g_0 is a Borel function. Similarly, since h_0 and g_0 are one-to-one, and h_0 is measure preserving,

$$\lambda(g_0^{-1}(S_n(x))) = \lambda(h_0^{-1} g_0^{-1}(S_n(x))) = \lambda(S_n(x)),$$

for all n and x. Taking the limit as $n \to \infty$, it follows that

$$\lambda(g_0^{-1}([0, x] \cap S \setminus (N_1 \cup N_2))) = \lambda([0, x] \cap S \setminus (N_1 \cup N_2)).$$

Now $g_0^{-1}([0, x] \cap S) = g_0^{-1}([0, x] \cap S \setminus (N_1 \cup N_2))$, since the range of g_0 is contained in $S \setminus (N_1 \cup N_2)$, and $\lambda([0, x] \cap S) = \lambda([0, x] \cap S \setminus (N_1 \cup N_2))$, since both N_1 and N_2 are null sets. It follows that

$$\lambda(g_0^{-1}([0, x] \cap S)) = \lambda([0, x] \cap S)$$

for all x. Since sets of the form $[0,x] \cap S$ generate the Borel subsets of S, this implies that g_0 is measure preserving. Observe that when $x = 1$, we obtain $\lambda(B) = \lambda(g_0^{-1}(S)) = \lambda(S)$. Since $B \subset [0, \lambda(S)]$, we have that the points in $[0, \lambda(S)]$ where g_0 is not defined constitute a Borel set of measure 0. To obtain functions g and h satisfying the conclusions of the lemma, we choose a point $s_0 \in S$ and define

$$h(x) = \begin{cases} h_0(x), & x \in S \setminus (N_1 \cup N_2) \\ 0, & x \in S \cap (N_1 \cup N_2) \end{cases}$$

$$g(y) = \begin{cases} g_0(y), & y \in B \\ s_0, & y \in [0, \lambda(S)] \setminus B. \end{cases}$$

The domains and ranges of g and h differ from the domains and ranges of g_0 and h_0 by Borel sets of measure 0, which guarantees that g and h inherit the measure preserving property. Also, $g \circ h(x) = x$ for $x \in S \setminus (N_1 \cup N_2 \cup \{s_0\})$, and $h \circ g(y) = y$ for $y \in B$. It follows that h and g are essential inverses of one another and are measure preserving Borel functions with the desired domains and ranges. $\boxed{QED}$

Theorem 4.3 (Rearrangement Theorem). *Let $\{S_k\}$ be an essentially pairwise disjoint family of Borel sets of positive measure, the sum of whose measures is 1. Then there is an essentially invertible measure preserving Borel function $g : [0,1] \to [0,1]$, with essential inverse h, and a special partition $\mathcal{P} : \{(a_k, a_{k+1})\}$ of $[0,1]$ such that $h^{-1}([a_k, a_{k+1})) = S_k$ essentially and $g^{-1}(S_k) = [a_k, a_{k+1})$ essentially for all k.*

Proof. Outline. First, replace the sets S_k by an equivalent disjoint collection. Define

$$N_{kn} = S_k \cap S_n$$

for $n > k$, and set

$$N = \cup_{k=1}^{\infty} \cup_{n=k+1}^{\infty} N_{kn}.$$

Then N is a Borel null set. Define $\tilde{S}_k = S_k \setminus N$ for all k. The sets $\tilde{S}_k$ are pairwise disjoint, $\tilde{S}_k \subset S_k$ for all k, and $\lambda(S_k \setminus \tilde{S}_k) = 0$ for all k.

Next, set $a_1 = 0$ and define

$$a_k = \sum_{j=1}^{k-1} \lambda(S_j), k > 1.$$

Then $\mathcal{P} : \{(a_k, a_{k+1})\}$ is a special partition of $[0,1]$, since

$$\sum_{k=1}^{\infty} (a_{k+1} - a_k) = \sum_{k=1}^{\infty} \lambda(S_k) = 1.$$

Next define g and h, using Lemma 4.2. Lemma 4.2 guarantees the existence of an essentially invertible measure preserving Borel function $h_k : \tilde{S}_k \to [0, \lambda(S_k)]$ for all k. Let $g_k : [0, \lambda(S_k)] \to \tilde{S}_k$ be a measure preserving essential inverse of h_k. Let $V_k \subset \tilde{S}_k \setminus (h_k^{-1}(0) \cup h_k^{-1}(\lambda(S_k)))$ and $B_k \subset (0, \lambda(S_k))$ be Borel sets of measure $\lambda(S_k)$ on which $g_k \circ h_k = e$ and $h_k \circ g_k = e$, respectively. Define

$$g(y) = \begin{cases} g_k(y - a_k), & a_k \le y < a_{k+1} \\ 0, & y = 1. \end{cases}$$

$$h(x) = \begin{cases} a_k + h_k(x), & x \in \tilde{S}_k \\ 0, & x \notin \cup_n \tilde{S}_k. \end{cases}$$

Then g and h are Borel functions which inherit the measure-preserving property from the functions g_k and h_k, as is easy to verify, and $g \circ h = e$ on $\cup_k V_k$, a Borel set of measure 1, and $h \circ g = e$ on $\cup_k \{a_k + B_k\}$, a Borel set of measure 1. $\boxed{QED}$

We will sometimes refer to the functions g and h of Theorem 4.3 as "rearrangements."

Theorem 4.4. *Suppose E is a totally atomic idempotent copula. Then there exists a copula C possessing a two-sided inverse with respect to M and a special partition $\mathcal{P} : \{(a_k, a_{k+1})\}$ of $[0,1]$ for which*

$$C^T E C = \oplus_{\mathcal{P}} F_k,$$

with $F_k = P$ for all k.

Proof. Let $\mathcal{B}_E$ denote the invariant sets of of E. A maximal collection of essentially disjoint atoms in $\mathcal{B}_E$ is necessarily at most a countable collection, since at most a finite number of disjoint atoms can have measure exceeding any positive real number a. Thus, we write $\{S_k\}$ for the collection, indexing the atoms by the positive integers or a suitable subset thereof (we assume here that the collection is countable and leave to the reader the proof in the case that the collection is finite). Since E is by hypothesis totally atomic and since the collection is maximal, we must have

$$\sum_k \lambda(S_k) = 1.$$

Set $a_1 = 0$ and define

$$a_k = \sum_{j=1}^{k-1} \lambda(S_j), \text{ for } k > 1.$$

By Theorem 4.3, there exist measure preserving functions $h : [0,1] \to [0,1]$ and $g : [0,1] \to [0,1]$ which map $S_k \to [a_k, a_{k+1})$ and $[a_k, a_{k+1}) \to S_k$ essentially, respectively, and which are essential inverses of one another. Since g and h are measure preserving Borel functions, we may construct from them copulas C_{ge} and C_{he}, and by Theorem 2.3 part (9) and Corollary 2.4.1, these copulas are inverses of one another with respect to M. In fact $C_{he} = C_{eg}$, since both are inverses of the copula C_{ge}, and inverses are unique. Similarly, $C_{eh} = C_{ge}$. By Theorem 2.11, for any measure preserving Borel function f and any set S,

$$T_{ef}\chi_S = \chi_S \circ f = \chi_{f^{-1}(S)} \text{ a.s., and}$$
$$T_{fe}\chi_{f^{-1}(S)} = T_{fe}\chi_S \circ f = \chi_S \text{ a.s.}$$

Set $F = C_{eg}EC_{ge}$, so that $E = C_{eh}FC_{he}$. If a set S is invariant under E, that is, if $T_E\chi_S = \chi_S$ a.s., then we calculate, using Theorem 2.11,

$$T_F\chi_{g^{-1}(S)} = T_{eg}T_E T_{ge}\chi_S \circ g = T_{eg}T_E\chi_S = T_{eg}\chi_S = \chi_S \circ g = \chi_{g^{-1}(S)} \text{ a.s.}$$

Thus, $g^{-1}(S)$ is is invariant under F. By parallel reasoning, if S is invariant under F, then $h^{-1}(S)$ is invariant under E. Since each S_k is invariant under E, it follows that $g^{-1}(S_k) = [a_k, a_{k+1})$ essentially is invariant under F. Thus, by an argument set forth in the proof of Theorem 4.2, F has an ordinal sum decomposition on the special partition $\mathcal{P} : \{(a_k, a_{k+1})\}$. Furthermore, if Q is invariant under F, then $Q_k = Q \cap [a_k, a_{k+1})$, being the intersection of invariant sets, is invariant under F, whence $h^{-1}(Q_k)$ is invariant under E. But observe that $h^{-1}(Q_k) \subset S_k$ essentially and that S_k is an atom. Therefore $\lambda(h^{-1}(Q_k)) = 0$ or $\lambda(S_k)$. Since h is measure preserving, it follows that $\lambda(Q_k) = 0$ or $a_{k+1} - a_k$. Therefore, the sets invariant under F are essentially intervals (a_k, a_{k+1}), their unions, and the empty set. It follows from Theorem 4.2 that

$$C_{eg}EC_{ge} = F = \oplus_{\mathcal{P}} F_k,$$

with each component $F_k = P$. $\boxed{QED}$

As in the case of nonatomic sub-σ-algebras, we can start with the σ-algebra, and construct the idempotent copula:

Corollary 4.4.1. *Let $\mathcal{S}$ be a totally atomic sub-σ-algebra of measurable subsets of* $[0,1]$. *There exist a measure preserving essentially invertible Borel function $h : [0,1] \to [0,1]$ and a special partition $\mathcal{P} : (a_k, a_{k+1})$ of $[0,1]$ such that the copula*

$$E = C_{eh}\big(\oplus_{\mathcal{P}} F_k\big)C_{he},$$

where $F_k = P$ for all k, has invariant sets essentially equivalent to $\mathcal{S}$.

Proof. Given a totally atomic sub-σ-algebra $\mathcal{S}$, there is an idempotent Markov operator whose invariant sets are essentially $\mathcal{S}$, namely the Raffaele Vitolo, Dipartimento di Matematica 'E. De Giorgi' Universita' del Salento, via per Arnesano 73100 Lecce ITALY tel.: +39 0832 297425 (office) fax.: +39 0832 297594 home page: http://poincare.unisalento.it/vitolo

conditional expectation operator $T : \phi \to E(\phi \,|\, \mathcal{S})$, [8]. Let E be the copula associated to T under the correspondence of Theorem 2.8, and observe that T and E have essentially the same invariant sets. Apply Theorem 4.4 to obtain a partition $\mathcal{P}$ and invertible copulas C_{eh} and C_{he} such that

$$E = C_{eh}\big(\oplus_{\mathcal{P}} F_k\big)C_{he},$$

where $F_k = P$ for all k. $\boxed{QED}$

Note again that while the idempotent copula constructed in Corollary 4.4.1 is unique, neither the special partition $\mathcal{P}$ nor the rearrangement h is uniquely determined. The atoms S_k of $\mathcal{S}$ can be labelled in any desired order, for example, which changes $\mathcal{P}$, and the function h can be modified in any desired fashion to rearrange the mass within an atom $S_k \in \mathcal{S}$, before mapping S_k to the interval (a_k, a_{k+1}).

5 The mixed case: atomic idempotent copulas which are not totally atomic

If E is an idempotent copula which is atomic but not totally atomic, then a maximal collection $\{S_k\}$ of essentially disjoint atoms of $\mathcal{B}_E$ has total measure less than 1. We label the sets S_k starting with $k = 2$ in this case and set

$$S_1 = \{\cup_{k\geq 2} S_k\}^c.$$

Then S_1 is a member of $\mathcal{B}_E$, no subset of which is an atom. S_1 has positive measure, for otherwise E would be totally atomic. We define $a_1 = 0$ and

$$a_k = \sum_{j=1}^{k-1} \lambda(S_j) \text{ for } k > 1,$$

as in the proof of Theorem 4.4, and we construct rearrangements h and g which map $S_k \to [a_k, a_{k+1})$ and $[a_k, a_{k+1}) \to S_k$ essentially, and are essential inverses of one another, exactly as before. By the reasoning in the proof of Theorem 4.4, $[a_k, a_{k+1})$ is necessarily invariant under $F = C_{eg}EC_{ge}$ for all k, whence, by an argument set forth in the proof of Theorem 4.2, F is necessarily an ordinal sum on the special partition $\mathcal{P} : \{(a_k, a_{k+1})\}$. Accordingly, we can write

$$F = \oplus_{\mathcal{P}} F_k,$$

and since trivially F is idempotent, each F_k must be idempotent. By the same logic as used in the proof of Theorem 4.4, for $k \geq 2$, the only subsets of $[a_k, a_{k+1})$ invariant under F are null sets and sets of measure $a_{k+1} - a_k$, since for $k \geq 2$, each S_k is an atom. Thus, by Lemma 4.1, the invariant sets of F_k for $k \geq 2$ are essentially $[0,1]$ and the empty set. Each component F_k for $k \geq 2$ must therefore be P, Theorem 4.1. But F_1 is a nonatomic idempotent, since otherwise some invariant subset of $[a_1, a_2)$ would be an atom in $\mathcal{B}_F$, whence its inverse image under h would be an atom in $\mathcal{B}_E$, contradicting the maximality of the collection of essentially disjoint atoms we started with. Accordingly F_1 is nonatomic, and by the characterization of nonatomic idempotents in Theorem 3.4, it must have the form $F_1 = C_{ef}C_{fe}$ for some measure preserving Borel function f. This is an outline of the proof of the following theorem.

Theorem 5.1. *Suppose E is an idempotent copula which is atomic but not totally atomic. Then there exist a copula C possessing a two-sided inverse with respect to M, a copula B such that $BB^T = M$ and a special partition $\mathcal{P} : \{(a_k, a_{k+1})\}$ of $[0,1]$ for which*

$$E = C(\oplus_{\mathcal{P}} F_k)C^T,$$

with $F_k = P$ for all $k \geq 2$ and $F_1 = B^T B$.

If we start with a sub-σ-algebra $\mathcal{S}$ of $\mathcal{B}$ instead of an atomic idempotent copula E, we obtain a similar result:

Corollary 5.1.1. *Let $\mathcal{S}$ be an atomic but not totally atomic sub-σ-algebra of measurable subsets of $[0,1]$. There exist a measure preserving essentially invertible Borel function $h : [0,1] \to [0,1]$, a special partition $\mathcal{P} : (a_k, a_{k+1})$ of $[0,1]$ and a nonatomic idempotent F_1 such that the copula*

$$E = C_{eh}\big(\oplus_{\mathcal{P}} F_k\big)C_{he},$$

where $F_k = P$ for all $k \geq 2$, has invariant sets essentially equivalent to $\mathcal{S}$.

Proof. The proof follows the proofs of Corollaries 3.4.1 and 4.4.1 and is omitted. $\boxed{QED}$

Note again that while the idempotent copula constructed in Corollary 5.1.1 is unique, neither the special partition $\mathcal{P}$ nor the rearrangement h is uniquely determined.

6 Consequences of the characterization of idempotent copulas

Theorem 6.1. *Every idempotent copula is symmetric, that is, if E is idempotent, necessarily $E(x, y) = E(y, x)$ for all x and y.*

Proof. Let E be idempotent. If E is nonatomic, then $E = C_{ef}C_{fe}$. E is symmetric, since $(C_{ef}C_{fe})^T = C_{ef}C_{fe}$. If E is atomic, then the characterization of one or the other of Theorems 4.4 or 5.1 applies. An ordinal sum of copulas is symmetric if and only if every component is symmetric, so ordinal sums whose components are P or a nonatomic idempotent are symmetric. For any symmetric copula A, $C^T AC$ is symmetric. Hence the characterization theorems for nonatomic and atomic copulas imply $E^T = E$. $\boxed{QED}$

Theorem 6.2. *The class of idempotent copulas is a lattice under the partial ordering $E \leq F$ if $EF = FE = E$.*

Proof. We have to show that for any two idempotent copulas E and F, there is a greatest idempotent G such that $G \leq E$ and $G \leq F$ and there is a least idempotent H such that $E \leq H$ and $F \leq H$.

Given E and F, EF is a copula, and by a classical argument, the function G defined by

$$G = \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} (EF)^k$$

is an idempotent copula. ($G = EF = FE$, if E and F commute, else EF is a copula but not an idempotent copula since it is not symmetric, and necessarily $G \neq EF$.) Furthermore, G annihilates EF in the sense that

$$G(EF) = (EF)G = G.$$

In addition, G is the greatest idempotent with this property, since if H is any other copula which annihilates EF, then H annihilates all convex combinations of powers of EF, so that

$$H(\frac{1}{n} \sum_{k=1}^{n} (EF)^k) = (\frac{1}{n} \sum_{k=1}^{n} (EF)^k)H = H$$

for all n. Take the limit, using the one-sided continuity of the $*$ product, to obtain $HG = GH = H$, i.e. $H \leq G$. We call G the greatest annihilator of EF.

It remains to show that G is the greatest lower bound of E and F in the partially ordered set of idempotent copulas. To that end, first take the transpose of $G(EF) = (EF)G = G$ using the symmetry of E, F and G, to obtain $(FE)G = G(FE) = G$. Conclude: G is an annihilator of FE, and if H is the greatest annihilator of FE, necessarily $G \leq H$. The same argument, starting with FE rather than EF shows $H \leq G$, whence necessarily $G = H$. Now consider the equation

$$\frac{1}{n} \sum_{k=1}^{n} (FE)^k = F(\frac{1}{n}(M - (EF)^n + \sum_{k=1}^{n} (EF)^k))E,$$

which is derived using the fact that $(FE)^k = F(EF)^{k-1}E$ for all $k > 1$. Take the limit as $n \to \infty$, and use the one-sided continuity of the $*$ product, to obtain on the left G and on the right FGE (since M/n and $(FE)^n/n$ both converge to 0). Thus, $G = FGE$ and $FG = F^2GE = FGE = G$ and $GE = FGE^2 = FGE = G$. Taking transposes, we have also $GF = G$ and $EG = G$. Thus, $G \leq E$ and $G \leq F$. If K is any other idempotent which annihilates both E and F, then K annihilates EF, and since G is the greatest annihilator of EF, necessarily $K \leq G$. Thus, G is the greatest common annihilator of E and F, and the greatest lower bound of E and F in the partially ordered set of idempotent copulas.

As to a least upper bound: Observe first that the min copula M is a common unit for any pair of idempotent copulas E and F. Observe next that if K and L are two common units for E and F, then their greatest lower bound H is a common unit for E and F. This follows from the fact that for all k $(KL)^kE = E(KL)^k = E$ and similarly $(KL)^kF = F(KL)^k = F$, whence, by the one-sided continuity of the $*$ product, the Cesaro limit of the powers, which is H, satisfies both $HE = EH = E$ and $HF = FH = F$. A Zorn's lemma argument using maximal ordered chains of common units for E and F now yields the desired result. The least elements in any pair of such maximal ordered chains must be equal, else their greatest lower bound would be a common unit for E and F and could be added to each of the chains, contradicting maximality. The least element in any such chain, therefore, is the smallest common unit for E and F. $\boxed{QED}$

In Theorems 6.3 and 6.4, and in Corollary 6.4.1, the sub-σ-algebras referred to are all taken to be the complete representatives of their equivalence classes, cf. Theorem 2.7 and related discussion for terminology.

Theorem 6.3. *Let E and F be idempotent copulas, and let $\mathcal{B}_E$ and $\mathcal{B}_F$ denote (completions of) their sub-σ-algebras of invariant sets. $E \le F$ iff $\mathcal{B}_E \subset \mathcal{B}_F$.*

Proof. Write T_E and T_F for the Markov operators associated with E and F. If $E \le F$, then $T_E T_F = T_F T_E = T_E$. Let $S \in \mathcal{B}_E$. Then

$$T_F \chi_S = T_F(T_E \chi_S) = [T_F T_E]\chi_S = T_E \chi_S = \chi_S \text{ a.s.}$$

so that $B \in \mathcal{B}_F$. For the converse, suppose $\mathcal{B}_E \subset \mathcal{B}_F$ and let $\psi \in L^1$. For every Borel set B, $[T_E\psi]^{-1}(B) \in \mathcal{B}_E \subset \mathcal{B}_F$ so, by Theorem 2.10, $T_E\psi$ is a fixed point of T_F. This says that $T_F T_E \psi = T_E \psi$ a.s. for all $\psi \in L^1$, hence that $T_F T_E = T_E$. It follows, by the isomorphism of copulas and Markov operators, that $FE = E$. Take transposes, using the fact that idempotent copulas are symmetric, and obtain also $EF = E$. We have shown that $E \le F$. $\boxed{QED}$

Theorem 6.4. *For any sub-σ-algebra $\mathcal{S} \subset \mathcal{B}$, there exists a unique idempotent copula E for which $\mathcal{B}_E = \mathcal{S}$ essentially.*

Proof. This result restates Corollary 3.4.1, Corollary 4.4.1 and Corollary 5.1.1 in convenient form for use here. The basic result is due to Sempi, [8], and is not really a consequence of the characterization given here. $\boxed{QED}$

Corollary 6.4.1. *The lattice of idempotent copulas, partially ordered by the $\le$ relation, is lattice isomorphic to the lattice of complete sub-σ-algebras of $\mathcal{B}$, partially ordered by the $\subset$ relation, under the mapping $E \to \mathcal{B}_E$.*

Proof. Theorem 6.3 says that the map $E \to \mathcal{B}_E$ is a lattice homomorphism and guarantees that the map is one-to-one. Theorem 6.4 says that the map is onto. $\boxed{QED}$

Theorem 6.5. *For any copula A there exists an idempotent copula E_A which annihilates A, i.e., $E_A A = AE_A = E_A$, and is such that $E_A \ge E$ for any other idempotent annihilator E of A. Furthermore, for any copula A there exists an idempotent copula F_A which is a unit for A, i.e., $F_A A = AF_A = A$. and is such that $F_A \le F$ for any other idempotent unit F for A. In all cases $E_A \le F_A$.*

Proof. Set

$$E_A = \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} A^k.$$

The argument showing that E_A is idempotent and annihilates A is well known, and is omitted. That E_A is the greatest annihilator of A is proved by an argument similar to that used to show G is the greatest annihilator of EF in the proof of Theorem 6.2. The proof of the existence and properties of F_A is similar to the argument used to show that idempotents E and F possess a least common unit in the proof of Theorem 6.2. Both are omitted. To see that necessarily $E_A \le F_A$, observe that

$$F_A E_A = F_A(AE_A) = (F_A A)E_A = AE_A = E_A.$$

Taking the transpose of $F_A E_A = E_A$ and using the symmetry of idempotents yields also $E_A F_A = E_A$. Thus, $E_A \le F_A$. $\boxed{QED}$

Theorem 6.6. *An idempotent E is an annihilator of a copula A if and only if E is an annihilator of A^T. An idempotent F is a unit for a copula A if and only if F is a unit for A^T. The greatest annihilators of A and A^T are equal. The least units of A and A^T are equal.*

Proof. This is a direct consequence of the symmetry of idempotents, Theorem 6.1. Details are left to the reader. $\boxed{QED}$

Theorem 6.7. *Let F be a nonatomic idempotent copula. Suppose that F is a unit for A and that A has a left inverse with respect to F, that is, there is a copula B such that $BA = F$. Then A possesses a unique left inverse, call it C, with respect to F among the copulas for which F is a unit. Also, F is the least unit of A and of C. Finally,*

$$C = A^T.$$

Identical conclusions hold if F is a unit for A and A possesses a right inverse with respect to F, with the word "left" replaced by "right" where needed.

Proof. If $BA = F$, define $C = FBF$. Then

$$\begin{aligned} CA &= FBFA \\ &= FBA \text{ (since } F \text{ is a unit for } A) \\ &= FF \quad \text{(since } B \text{ is a left inverse of } A \text{ with respect to } F) \\ &= F \qquad \text{(since } F \text{ is idempotent).} \end{aligned}$$

Thus, C is a left inverse of A. F is trivially a unit for C. So far, we have not used the fact that F is nonatomic. We do so now. Since F is nonatomic, there is a measure preserving Borel function f such that $F = C_{ef}C_{fe}$, Theorem 3.4. Since F is a unit for A, $CFA = CA = F$. Substituting $C_{ef}C_{fe}$ in this expression, and pre- and post-multiplying by C_{fe} and C_{ef} yields

$$(C_{fe}CC_{ef})(C_{fe}AC_{ef}) = C_{fe}C_{ef}C_{fe}C_{ef} = C_{ff}^2 = M,$$

using parts of Theorem 2.3. It follows that $C_{fe}AC_{ef}$ has a left inverse with respect to M, and since a left inverse with respect to M, when it exists, must be the transpose, Theorem 2.1, we have necessarily

$$C_{fe}CC_{ef} = (C_{fe}AC_{ef})^T = C_{ef}^T A^T C_{ef}^T = C_{fe}A^T C_{ef},$$

using part (6) of Theorem 2.3. Now pre- and post-multiply by C_{ef} and C_{fe}, and use $F = C_{ef}C_{fe}$, to obtain $FCF = FA^TF$. Since F is a unit for both C and A^T, we have $C = A^T$. This implies the uniqueness of C. To complete the proof, we have to show that F is the least unit of A and of A^T. To that end, let G be any unit for A. Then since $A^TA = F$, we have

$$FG = (A^TA)G = A^T(AG) = A^TA = F,$$

whence $FG = F$. Taking the transpose of this equation, using the symmetry of idempotents, gives $GF = F$. F is accordingly the least unit for A and, by Theorem 6.6, also for $C = A^T$. $\boxed{QED}$

Remark: We conjecture that also when F is atomic it is true that, when F is a unit for A and A has a left or right inverse with respect to A, then A possesses a unique left or right inverse among copulas for which F is a unit, and the left or right inverse is necessarily A^T.

We will use the notation $\mathcal{C}_F$ for the family of copulas for which F is a unit. Define also $\mathcal{G}_F$ to be the subset of $\mathcal{C}_F$ consisting of copulas which have two-sided inverses with respect to F. Then $\mathcal{G}_F$ is a group, and when F is nonatomic, it is isomorphic to $\mathcal{G}_M$, the set of copulas possessing two-sided inverses with respect to M:

Theorem 6.8. *Suppose that F is a nonatomic idempotent copula, and write $F = C_{ef}C_{fe}$, where $f : [0,1] \to [0,1]$ is a measure preserving Borel function. The mapping $\Phi : \mathcal{C}_F \to \mathcal{C}$, of $\mathcal{C}_F$ into the the set $\mathcal{C}$ of all 2-copulas, given by*

$$\Phi(A) = C_{fe}AC_{ef},$$

is one-to-one and onto and preserves both the $*$ product and convex combinations. Φ maps $\mathcal{G}_F$ to $\mathcal{G}_M$ in one-to-one and onto fashion. Thus, the restriction of Φ to $\mathcal{G}_F$ is a group isomorphism. The mapping Φ and its inverse are both continuous with respect to uniform convergence.

Proof. The map Φ is onto, since for any $C \in \mathcal{C}$, $\Phi(C_{ef}CC_{fe}) = C$. It is trivial to verify that $F = C_{ef}C_{fe}$ is a unit for $C_{ef}CC_{fe}$, hence that $C_{ef}CC_{fe}$ is in the domain of Φ. Φ is one-to-one since if F is a unit for both A and B and $C_{fe}AC_{ef} = C_{fe}BC_{ef}$, then pre- and post-multiplying by C_{ef} and C_{fe} yields $FAF = FBF$, whence $A = B$. That Φ preserves convex combinations is trivial. To see that Φ preserves $*$ products of copulas, observe first that if F is a unit for A and B, then F is a unit for AB, since $F(AB) = (FA)B = AB$ and $(AB)F = A(BF) = AB$, whence the product of two copulas in the domain of F is also in the domain. We can always insert $F = C_{ef}C_{fe}$ between A and B in the product AB, since F is a unit for each of A and B. Thus, we have

$$\Phi(AB) = C_{fe}ABC_{ef} = C_{fe}AC_{ef}C_{fe}BC_{fe} = \Phi(A)\Phi(B).$$

A has a two-sided inverse with respect to F if and only if $\Phi(A) = C_{fe}AC_{ef}$ has a two-sided inverse with respect to M, by an argument similar to that used in the proof of Theorem 6.7. The group isomorphism assertion in the theorem follows directly from this and the fact that Φ preserves the $*$ product. As to continuity, if $A_n \to A$ uniformly, then $C_{fe}A_nC_{ef} \to C_{fe}AC_{ef}$ uniformly, by the one-sided continuity property of the $*$ product ($A_n \to A$ imples $C_{fe}A_n \to C_{fe}A$ implies $C_{fe}A_nC_{ef} \to C_{fe}AC_{ef}$). Thus Φ is continuous. Similarly, if $B_n \to B$ then $C_{ef}B_nC_{fe} \to C_{ef}BC_{fe}$, so Φ^{-1} is continuous. $\boxed{QED}$

References

[1] W.F. Darsow, B. Nguyen, E.T. Olsen: *Copulas and Markov Processes*, Ill. J. Math., **36**, 600-642 (1992).

[2] W.F. Darsow, E.T. Olsen: *A Representation for Doubly Stochastic Measures*, IIT Research Report 1994.

[3] E.T. Olsen, W.F. Darsow, B. Nguyen: *Copulas and Markov Operators*, in Distributions with Fixed Marginals and Related Topics, L. Rueschendorf, B. Schweizer and M.D. Taylor, eds., pp. 244-259, Inst. of Mathematical Statistics, Hayward, California 1996.

[4] A. Kolesárová, R. Mesiar, C. Sempi: *Measure-Preserving Transformations, Copulae and Compatibility*, Mediterr. J. Math., **5**, 325-339 (2008).

[5] R.B. Nelsen: An Introduction to Copulas, Springer, New York 1999.

[6] H.L. Royden: Real Analysis, 3d ed., Macmillan, New York 1988.

[7] B. Schweitzer, A. Sklar: Probabilistic Metric Spaces, Dover, Mineola NY 1983.

[8] C. Sempi: *Conditional Expectations and Idempotent Copulae*, in Distributions with Given Marginals and Statistical Modelling, C.M. Cuadras et al., eds., pp. 223-228, Kluwer, Netherlands 2002.

[9] R.A. Vitale: *Parametrizing Doubly Stochastic Measures*, in Distributions with Fixed Marginals and Related Topics, L. Rueschendorf, B. Schweizer and M.D. Taylor, eds., pp. 358-364, Inst. of Mathematical Statistics, Hayward, California 1996.