HAL
open science

Bases and linear transforms of TU-games and cooperation systems

Ulrich Faigle, Michel Grabisch

To cite this version:

Ulrich Faigle, Michel Grabisch. Bases and linear transforms of TU-games and cooperation systems. International Journal of Game Theory, 2016, 45 (4), pp.875-892. 10.1007/s00182-015-0490-x . hal-01404509

HAL Id: hal-01404509

https://hal.science/hal-01404509v1

Submitted on 28 Nov 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Bases and Linear Transforms of TU-Games and Cooperation Systems

Ulrich FAIGLE* Michel GRABISCH†

July 16, 2015

Abstract

We study linear properties of TU-games, revisiting well-known issues like interaction transforms, the inverse Shapley value problem and potentials. We embed TU-games into the model of cooperation systems and influence patterns, which allows us to introduce linear operators on games in a natural way. We focus on transforms, which are linear invertible maps, relate them to bases and investigate many examples (Möbius transform, interaction transform, Walsh transform and Fourier analysis *etc.*). In particular, we present a simple solution to the inverse problem in its general form: Given a linear value Φ and a game v, find all games v' such that $\Phi(v) = \Phi(v')$. Generalizing Hart and Mas-Colell's concept of a potential, we introduce general potentials and show that every linear value is induced by an appropriate potential.

Keywords: Cooperation system, cooperative game, basis, Fourier analysis, inverse problem, potential, transform
JEL Classification: C71

*Mathematisches Institut, Universität zu Köln, Weyertal 80, 50931 Köln, Germany. Email: faigle@zpr.uni-koeln.de

†Corresponding author. Paris School of Economics, University of Paris I, 106-112, Bd. de l'Hôpital, 75013 Paris, France. Tel. (33) 144-07-82-85, Fax (33)-144-07-83-01. Email: michel.grabisch@univ-paris1.fr

1 Introduction

It is well known that finite TU-games with n players, and more generally set functions on a set N of $n = |N|$ elements, form a 2^n-dimensional vector space. Usually, particular attention is paid to two special bases: the set of unanimity games ζ_S, $S \in 2^N$, and the set of identity (Dirac) games δ_S. The coefficients of a game in the basis of unanimity games are known as the *Harsanyi dividends* [14], or the *Möbius inverse* [20]. But also other transforms (namely, invertible linear operators) have been proposed and studied in the literature (*e.g.*, the interaction transform [8], the Walsh transform [23], which is also known as Fourier transform, *etc.*).

Although recognized to be important in discrete mathematics, these transforms are not very well known in the game theory community so far. In fact, it seems that the linear properties of TU-games–while often used indirectly–have not yet been fully exploited in game theoretic research in their own right. To the best of our knowledge, the obvious correspondence between bases and linear transforms has never been addressed explicitly, for example. As a consequence, the famous "inverse problem" for TU-games, which asks for a description of all games v' having the same Shapley value as a given game v, has been solved in a somewhat tedious way (see Kleinberg and Weiss [17] or Dragan [3, 6]). We solve the inverse problem for general linear values in Section 3.3.

In our present study of linear properties of game theoretic concepts it is convenient (and natural) to embed TU-games into the context of *cooperation systems* and *influence patterns* of coalitions. We introduce this model in Section 2. Linear transforms and their interplay with bases are investigated in Section 3 and illustrated with fundamental examples and applications and relate them to discrete Fourier analysis.

Hart and Mas-Collel [15] have given an interpretation of the Shapley value as a *potential value* which provides an interesting link to the (non-cooperative!) *potential games* of Monderer and Shapley [18]. It turns out that a comprehensive framework for linear potentials as generalizations of transforms exists in the context of cooperation systems. We show in Section 4 how this general framework allows us to exhibit every linear value as a potential value.

2 Cooperation systems

In order to set up a suitable context for the analysis of transforms and values, we extend the classical model of cooperative games to the model of cooperation systems. We consider finite sets N of $n = |N| \geq 1$ players, denoting by $\mathcal{N} = 2^N$ the collection of all subsets (or *coalitions*) $S \subseteq N$.

A *cooperation system* is a pair (N, F), where $F : \mathcal{N} \times \mathcal{N} \to \mathbb{R}$ is a map that reflects for every pair (S, T) of coalitions, the amount $F(S, T)$ of *influence* S exerts on T. We refer to F as the *influence pattern* of (N, F). A *valuation* on (N, F) is a map $v : \mathcal{N} \to \mathbb{R}$ that assigns to every coalition S a value $v(S)$. In the terminology of classical cooperative game theory, a valuation v with $v(\emptyset) = 0$ is the *characteristic function* of a *cooperative TU game* (N, v).

The valuations v form the 2^n-dimensional vector space $\mathbb{R}^{\mathcal{N}}$, while the possible influence patterns F define the vector space $\mathbb{R}^{\mathcal{N}\times\mathcal{N}}$. Mathematically, one may think of a valuation $v \in \mathbb{R}^{\mathcal{N}}$ as a parameter vector, indexed by the $S \in \mathcal{N}$. We denote the (euclidean) inner product of $\mathbb{R}^{\mathcal{N}}$ as

$$\langle v, w \rangle = \sum_{S\in\mathcal{N}} v_S w_S = \sum_{S\in\mathcal{N}} v(S)w(S).$$

$F \in \mathbb{R}^{\mathcal{N}\times\mathcal{N}}$ corresponds to a matrix $[f_{ST}]$ with rows and columns indexed by the members of $\mathcal{N}$ and coefficients $f_{ST} = F(S, T)$. We denote by f_S the row vector of F that corresponds to $S \subseteq N$ and interpret it as the associated *influence function* $f_S : \mathbb{R}^{\mathcal{N}} \to \mathbb{R}$ with values

$$f_S(T) = f_{ST} = F(S, T) \qquad (T \subseteq N).$$

2.1 Examples

Counting pattern. Let $C = [c_{ST}] \in \mathbb{R}^{\mathcal{N}\times\mathcal{N}}$ be the pattern defined by

$$c_{ST} = |S \cap T|.$$

Here the individual players act independently. The influence of a coalition S on another coalition T depends only on the number of players in S that are also members of T.

Parity pattern. Let $\Pi = [\pi_{ST}] \in \mathbb{R}^{\mathcal{N}\times\mathcal{N}}$ be the pattern with coefficients

$$\pi_{ST} = (-1)^{|S\cap T|}.$$

Π is called the *parity pattern*. Two coalitions S and T exert a positive ("$+1$") or negative ("-1") influence on each other depending on whether they have an even or an odd number of players in common. Π plays a major role in the Fourier analysis of $\mathbb{R}^{\mathcal{N}}$ (see Section 3.2 below), which is based on the following fundamental observation.

Lemma 2.1 *For all $S, T \subseteq N$, one has the orthogonality property*

$$\langle \pi_S, \pi_T \rangle = \sum_{K\subseteq N} \pi_{SK}\pi_{TK} = \begin{cases} 2^n & \text{if } S = T, \\ 0 & \text{if } S \neq T. \end{cases}$$

Proof. In the case $S = T$, one has

$$\langle \pi_S, \pi_T \rangle = \sum_{K\in\mathcal{N}} (-1)^{|S\cap K|}(-1)^{|S\cap K|} = \sum_{K\subseteq N} 1 = 2^n.$$

So we may assume the existence of some $t \in T \setminus S$ without loss of generality. In this case, we observe

$$\sum_{K\subseteq N\setminus t} (-1)^{|S\cap K|}(-1)^{|T\cap K|} = (-1)\cdot \sum_{K\ni t} (-1)^{|S\cap K|}(-1)^{|T\cap K|}$$

and hence

$$\langle \pi_S, \pi_T \rangle = \sum_{K\subseteq N} (-1)^{|S\cap K|}(-1)^{|T\cap K|} = 0.$$

□

Containment pattern. Let $Z = [\zeta_{ST}] \in \mathbb{R}^{\mathcal{N}\times\mathcal{N}}$ be the pattern with the coefficients

$$\zeta_{ST} = \begin{cases} 1 & \text{if } S \subseteq T \\ 0 & \text{otherwise.} \end{cases}$$

Z is the containment pattern, where a coalition S is thought to be able to exert an influence on any coalition T containing it. For any $S \in \mathcal{N}$, the influence function ζ_S is $(0,1)$-valued. If $S \neq \emptyset$, (N, ζ_S) is commonly known as a *unanimity game*.

Influence in voting In Grabisch and Rusinowska [12], influence is studied in terms of voting dynamics as follows. Let S be the set of 'yes'-voters at time t and let m_{ST} be the probability for T to be the set of 'yes'-voters at time $t+1$. The transition $S \to T$ is the result of a round of discussion among voters. So m_{ST} is a measure for the influence the constellation S exerts on the constellation T. Here, the pattern matrix $M = [m_{ST}]$ is row-stochastic and defines a Markov process on $\mathcal{N}$.

2.2 Influence spaces, bases and additive games

We define the *influence space* $\mathcal{F}$ of F as the collection of all linear combinations of influence functions:

$$\mathcal{F} = \{v \in \mathbb{R}^{\mathcal{N}} \mid v = \sum_{S \in \mathcal{N}} \lambda_S f_S, \lambda_S \in \mathbb{R}\}.$$

$\mathcal{F}$ is a subspace of $\mathbb{R}^{\mathcal{N}}$ and corresponds to the row space of the matrix F. We think of parameters vectors λ as a row vectors and can therefore have in matrix notation

$$\mathcal{F} = \{\lambda F \mid \lambda \in \mathbb{R}^{\mathcal{N}}\} = \mathbb{R}^{\mathcal{N}} F.$$

We state a well-known fundamental linear algebraic fact.

Lemma 2.2 (Basis lemma) *Equality $\mathcal{F} = \mathbb{R}^{\mathcal{N}}$ holds if and only if the 2^n influence functions f_S are linearly independent and hence form a basis of $\mathbb{R}^{\mathcal{N}}$.*

It is easy to see (see Ex. 2.1 below) that the 2^n influence functions (unanimity games) ζ_S of the containment pattern Z are linearly independent and hence form a basis of $\mathbb{R}^{\mathcal{N}}$. The inverse pattern is given by the *Möbius matrix* $Z^{-1} = [\mu_{ST}] = M$ with the coefficients (*cf.* Rota [20])

$$\mu_{ST} = \begin{cases} (-1)^{|T \setminus S|} & \text{if } S \subseteq T, \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

On the other hand, the influence functions c_S of the counting pattern C are *not* linearly independent. Indeed, setting $\zeta_i = \zeta_{\{i\}}$, we have

$$\zeta_i = c_{\{i\}} \quad \text{and} \quad c_S = \sum_{i \in S} \zeta_i \quad \text{for all } i \in N \text{ and } S \in \mathcal{N} \setminus \{\emptyset\}. \tag{2}$$

It follows that the n influence functions ζ_i form a basis of the influence space $\mathcal{C}$ of the counting pattern C. In other words, $\mathcal{C}$ is the vector space of all characteristic functions v of the form

$$v = \sum_{i \in N} v(\{i\})\zeta_i \quad \text{or} \quad v(S) = \sum_{i \in S} v(\{i\}) \ \ \forall S \subseteq N.$$

So $\mathcal{C}$ is the n-dimensional space of all *additive* cooperative games on N.

Example 2.1 *Let $F = [f_{ST}]$ be such that for all $S, T \in \mathcal{N}$, one has $f_{SS} \neq 0$ and $f_{ST} = 0$ unless $S \subseteq T$. Label the rows and columns of F so that S precedes T whenever $S \subset T$. This exhibits F as (upper) triangular with non-zero diagonal. So the 2^n influence functions f_S are seen to be linearly independent (see Grabisch et al. [10] and Denneberg and Grabisch [2] for a detailed treatment of this type of matrix in the context of interaction).*

2.3 Linear values

In game theoretic terminology, a *value* is a function $\Phi : \mathbb{R}^{\mathcal{N}} \to \mathbb{R}^N$ that evaluates the strength (or power or reward *etc.*) of player $i \in N$ relative to the valuation v as $\Phi_i(v)$. The value Φ can be considered as a mapping Φ from $\mathbb{R}^{\mathcal{N}}$ to $\mathcal{C}$, the influence space of the counting pattern, which assigns to any valuation v an additive game Φ^v defined by

$$\Phi^v(S) = \sum_{i \in S} \Phi_i(v) \quad \text{for any } S \in \mathcal{N} \setminus \{\emptyset\} \tag{3}$$

with the convention $\Phi^v(\emptyset) = 0$. Φ is *linear* if it is a linear map. A classical linear example is Shapley's [22] value Φ^{Sh}, defined by

$$\Phi_i^{\mathrm{Sh}}(v) = \sum_{S \in \mathcal{N} \setminus \{\emptyset\}} \frac{(n - |S|)!(|S| - 1)!}{n!} \big(v(S) - v(S \setminus i)\big) \qquad (i \in N). \tag{4}$$

In applications, one is often interested in values Φ that satisfy additional conditions. For example, Φ might be required to be *efficient* in the sense

$$\sum_{i \in N} \Phi_i(v) = v(N) \quad \text{for all valuations } v \in \mathbb{R}^{\mathcal{N}}.$$

3 Bases and linear transforms

Consider a cooperation system (N, F) with a full-dimensional influence space $\mathcal{F} = \mathbb{R}^{\mathcal{N}}$. Hence, by Lemma 2.2, F is invertible, and the rows f_S, $S \in \mathcal{N}$, of F form a basis of $\mathbb{R}^{\mathcal{N}}$, yielding for any $v \in \mathbb{R}^{\mathcal{N}}$ a unique representation

$$v = \sum_{S \in \mathcal{N}} w_S f_S,$$

or in matrix notation $v = wF$ with w the row-vector $[w_S]_{S \in \mathcal{N}}$, and consequently $w = vF^{-1}$. Following some tradition, one may view the mapping $v \mapsto w$ as a *transform*, namely a linear and invertible operator on $\mathbb{R}^{\mathcal{N}}$ with F^{-1} as its standard matrix representation (see below for such examples). Letting Ψ be any such transform $v \mapsto \Psi^v$, the above observations yield the well-known one-to-one correspondence between bases and transforms:

Lemma 3.1 (Equivalence between bases and transforms) *For every basis $F = [f_S]_{S \in \mathcal{N}}$ of $\mathbb{R}^{\mathcal{N}}$, there is a (unique) transform Ψ such that for any $v \in \mathbb{R}^{\mathcal{N}}$,*

$$v = \sum_{S \in \mathcal{N}} \Psi^v(S) f_S, \tag{5}$$

whose inverse Ψ^{-1} is given by $v \mapsto (\Psi^{-1})^v = \sum_{T \in \mathcal{N}} v(T) f_T = vF$.

Conversely, to any transform Ψ there corresponds a unique basis F of $\mathbb{R}^{\mathcal{N}}$ such that (5) holds, given by $f_S = (\Psi^{-1})^{\delta_S}$, where δ_S is the identity game with $\delta_S(T) = 1$ if $S = T$ and 0 otherwise.

Game theoretic investigations have traditionally been restricted to the use of mainly two bases: the basis of identity games and the basis of unanimity games. For the latter, it is well-known the coordinates of a game are its Harsanyi dividends (see the next section). Lemma 3.1 above, although straightforward from an abstract linear algebraic point of view, exhibits the general duality between bases and transforms, which, to the best of our knowledge, has never been noticed nor exploited game theoretically. As a first consequence of the lemma, the use of the various transforms already existing in the fields of game theory, operations research and computer science, puts at our disposal a variety of new bases for the analysis of games. It is well known that the choice of a "good" basis can be of crucial importance. Recall, for example, that the characterization of the Shapley value by linearity, symmetry, null player and efficiency can be established in a few lines with the basis of unanimity games. So the knowledge of new bases can be

of considerable help in the study of cooperative games. Sections 3.1 and 3.2 give many examples of transforms and their associated bases.

A second important consequence is that the solution of the well-known "inverse problem", which asks for finding all games having the same Shapley value (or other linear value), can be obtained in an elegant way (see Section 3.3 below).

3.1 Examples

Harsanyi dividends and Möbius transform. Harsanyi [14] has shown that a valuation v admits coefficients m_S^v (the so-called *Harsanyi dividends* in game-theoretic language) such that one has

$$v(T) = \sum_{S \subseteq T} m_S^v = \sum_{S \subseteq N} m_S^v \zeta_S(T) \quad \text{for all } T \subseteq N. \tag{6}$$

To see that such coefficients exist indeed, just observe that the second equality in (6) defines the the so-called *Möbius transform* $v \mapsto m^v$ relative to the containment pattern Z:

$$v = \sum_{S \subseteq N} m_S^v \zeta_S.$$

The inverse (Möbius) pattern $Z^{-1} = [\mu_{ST}]$, therefore, yields the representation

$$m^v = \sum_{S \subseteq N} v(S) \mu_S$$

and hence the explicit formula for the Harsanyi dividends:

$$m^v(S) = \sum_{T \in \mathcal{N}} v(T) \mu_T(S) = \sum_{T \subseteq S} (-1)^{|S \setminus T|} v(T). \tag{7}$$

The commonality transform. The *commonality coefficients* $\check{m}_S^v$ of a valuation v were studied by Shafer [21] (see also Grabisch *et al.*[10]) as parameters with the property

$$v(S) = \sum_{T \subseteq N \setminus S} (-1)^{|T|} \check{m}^v(T). \tag{8}$$

To demonstrate their existence, we set $\tilde{v}(S) = (-1)^{|S|} v(N \setminus S)$ and conclude from (7) that the numbers $\tilde{v}(S)$ must be the Harsanyi dividends of $\check{m}^v$:

$$\tilde{v}(S) = \sum_{T \subseteq S} (-1)^{|S \setminus T|} \check{m}^v(T).$$

In view of (6), we thus find

$$\check{m}^v(S) = \sum_{T \subseteq S} \tilde{v}(T) = \sum_{T \subseteq S} (-1)^{|T|} v(N \setminus T). \tag{9}$$

The *commonality transform* (or *co-Möbius function*) is the operator $v \mapsto \check{m}^v$. From (8), we infer immediately that the basis of the commonality transform consists of the valuations f_S with values

$$f_S(T) = \sum_{B \subseteq N \setminus S} (-1)^{|B|} \delta_S(B) = \begin{cases} (-1)^{|S|} & \text{if } S \cap T = \emptyset, \\ 0 & \text{otherwise.} \end{cases}$$

The Shapley interaction transform. The *Shapley (interaction) transform* on $\mathbb{R}^{\mathcal{N}}$ is the function $v \mapsto I^v$ defined by

$$I^v(S) \quad = \quad \sum_{K \subseteq N} \frac{|N \setminus (S \cup K)|! |K \setminus S|!}{(n - |S| + 1)!} (-1)^{|S \setminus K|} v(K).$$

It extends the Shapley value in the sense

$$I^v(\{i\}) = \sum_{T \subseteq N \setminus i} \frac{(n - |T| - 1)! |T|!}{n!} \big(v(T \cup i) - v(T) \big) = \Phi_i^{\mathrm{Sh}}(v). \tag{10}$$

It was shown by Grabisch [8] that v can be recaptured from I^v:

$$v(S) = \sum_{K \subseteq N} \beta_{|S \cap K|}^{|K|} I^v(K), \tag{11}$$

where

$$\beta_k^l = \sum_{j=0}^{k} \binom{k}{j} B_{l-j} \qquad (k \leq l),$$

and $B_0, B_1, \ldots$ are the Bernoulli numbers. The first values of β_k^l are given in Table 1.

Using Lemma 3.1, we find that the corresponding basis consists of the 2^n valuations b_T^I with values

$$b_T^I(S) = \beta_{|T \cap S|}^{|T|} \quad \text{for all } S \in \mathcal{N}. \tag{12}$$

$k \setminus l$	0	1	2	3	4
0	1	$-\frac{1}{2}$	$\frac{1}{6}$	0	$-\frac{1}{30}$
1		$\frac{1}{2}$	$-\frac{1}{3}$	$\frac{1}{6}$	$-\frac{1}{30}$
2			$\frac{1}{6}$	$-\frac{1}{6}$	$\frac{2}{15}$
3				0	$-\frac{1}{30}$
4					$-\frac{1}{30}$

Table 1: The coefficients β_k^l

The Banzhaf interaction transform. The well-known *Banzhaf value* of a cooperative game (N, v) is the linear value defined as follows:

$$\Phi_i^{\mathrm{B}}(v) = \frac{1}{2^{n-1}} \sum_{T \subseteq N \setminus i} [v(T \cup i) - v(T)] \qquad (i \in N). \tag{13}$$

REMARK. The value Φ^{B} was introduced by Banzhaf [1] for voting games (*i.e.*, monotone cooperative games (N, v) with $v : \mathcal{N} \to \{0, 1\}$) and is commonly known as the *Banzhaf power index* in the voting context.

Grabisch *et al.* [10] have extended the Banzhaf value to arbitrary coalitions $S \subseteq N$ *via*

$$I_{\mathrm{B}}^v(S) = \left(\frac{1}{2}\right)^{n-|S|} \sum_{T \subseteq N} (-1)^{|S \setminus T|} v(T), \tag{14}$$

and thus the property $I_B^v(\{i\}) = B_i^v$ for all $i \in N$, and have shown that v can be reconstructed from I_B^v:

$$v(S) = \sum_{T \subseteq N} \frac{(-1)^{|T \setminus S|}}{2^{|T|}} I_B^v(T) \qquad (S \subseteq N). \tag{15}$$

It follows that the *Banzhaf (interaction) transform* $v \mapsto I_B^v$ is the linear transform associated with the 2^n basis functions $b_T^{I_B}$ with values

$$b_T^{I_B}(S) = \frac{(-1)^{|T \setminus S|}}{2^{|T|}}. \tag{16}$$

The inversion relation (14) $\leftrightarrow$ (15) can be verified by direct computation. We will show below that it follows easily within the setting of game-theoretic Fourier analysis (and the Walsh transform there, in particular).

REMARK. There is a general theory of *interaction values* of cooperative games that includes in particular Owen's [19] local interaction values ("co-values") and their extension to arbitrary subsets by Grabisch and Roubens [11]. As it turns out all these values are linear and hence fall into the scope of the present model. We do not go into further details here but refer the reader to Faigle and Voss [7]. The conversion formulas between the transforms presented above were originally established in [10] (see also a summary in [9]).

3.2 Fourier analysis

Roughly speaking, we mean with "Fourier analysis" the analysis of valuations in the context of the parity pattern $\Pi = [\pi_{ST}]$ and hence consider the *parity transform* $v \mapsto P^v$ with the property

$$v(S) = \sum_{T \in \mathcal{N}} P^v(T)\pi_T(S) = \sum_{T \in \mathcal{N}} (-1)^{|S \cap T|} P^v(T) \qquad (S \in \mathcal{N}). \tag{17}$$

Because of $\Pi^{-1} = 2^{-n}\Pi$ (*cf.* Lemma 2.1), we immediately find

$$P^v(S) = \frac{1}{2^n} \sum_{T \in \mathcal{N}} (-1)^{|S \cap T|} v(T) \qquad (S \in \mathcal{N}). \tag{18}$$

REMARK. The equivalent *Hadamard transform* $v \mapsto H^v$ with $H^v = 2^{n/2} P^v$, *i.e.*,

$$H^v(S) = \frac{1}{2^{n/2}} \sum_{T \subseteq N} (-1)^{|S \cap K|} v(T) \qquad (S \in \mathcal{N}) \tag{19}$$

is self-inverse (*i.e.*, $H^{-1} = H$). H is of particular importance in quantum computing, for example, and known as an instance of the *discrete (quantum) Fourier transform* (see, *e.g.*, Gruska [13]).

The Walsh transform. Closely related to the parity transform is the *Walsh transform* $v \mapsto W^v$ relative to the modified parity basis functions $w_S = (-1)^{|S|}\pi_S$ and hence the property

$$v(S) = \sum_{T \in \mathcal{N}} W^v(T) w_T(S) = \sum_{T \in \mathcal{N}} (-1)^{|T|} W^v(T) \pi_S. \tag{20}$$

The functions w_S were introduced by Walsh [23] and have the values

$$w_S(T) = (-1)^{|S|}(-1)^{|S \cap T|} = (-1)^{|S \setminus T|}.$$

In view of the uniqueness of the transformation coefficients, we have equality $P^v(T) = (-1)^{|T|}W^v(T)$ and hence deduce from (18):

$$W^v(S) = (-1)^{|S|}P^v(S) = \frac{1}{2^n}\sum_{T\in\mathcal{N}}(-1)^{|S\setminus T|}v(T) \qquad (S \in \mathcal{N}). \tag{21}$$

Recalling the Banzhaf transform (14), we thus observe its intimate connection with the Walsh transform:

$$I_B^v(S) = 2^{|S|}W^v(S) \quad \text{for all } S \in \mathcal{N}. \tag{22}$$

Moreover, (20) can be re-written in the form

$$v(S) = \sum_{T\in\mathcal{N}} I_B^v(T)2^{-|T|}w_T(S) = \sum_{T\in\mathcal{N}} I_B^v(T)b_T^{I_B}(S)$$

with $b_T^{I_B}(S) = 2^{-|T|}(-1)^{|T\setminus S|}$, as claimed in (16).

REMARK. The Fourier/Walsh approach is particularly appropriate in the context of social choice theory when one thinks of a function $f : \mathcal{N} \to \{0, 1\}$ as a "social choice function". For example, Kalai [16] has demonstrated that Arrow's theorem admits a short proof in this setting.

3.3 The inverse problem

In game theory, the following "inverse problem" is well-known: for a given linear value Φ and game v, find all games v' such that

$$\Phi(v) = \Phi(v') \quad \text{or, equivalently,} \quad \Phi(v - v') = 0.$$

This problem was solved[1] by Kleinberg and Weiss [17] for the Shapley value by exhibiting a basis for the associated *null space* or *kernel*:

$$\ker(\Phi) = \{v \in \mathbb{R}^{\mathcal{N}} \mid \Phi_i(v) = 0 \ \forall i \in N\}.$$

Our linear analysis provides adequate tools for solving the problem easily in its full generality. We present two approaches, the first one being very simple

[1] See also Yokote et al. [24] for recent work on this topic, and Dragan [5], who solved this problem for the Shapley value [3] and later for all semivalues [4] in a simpler way than Kleinberg and Weiss.

but requiring some extra condition to hold, while the second is applicable in any situation.

The first construction requires a transform Ψ to be known such that the linear value Φ in question corresponds to the transform restricted to singletons, *i.e.*, $\Phi_i(v) = \Psi^v(\{i\})$ for every $i \in N$ and game v. Using the corresponding basis $\{b_T^\Psi\}_{T\in\mathcal{N}}$, we can then write:

$$v = \sum_{S\in\mathcal{N}} \Psi^v(S) b_S^\Psi = \sum_{i\in N} \Phi_i(v) b_{\{i\}}^\Psi + \sum_{|S|\neq 1} \Psi^v(S) b_S^\Psi,$$

which implies

$$v \in \ker(\Phi) \quad \Longleftrightarrow \quad v = \sum_{|S|\neq 1} \Psi^v(S) b_S^\Psi,$$

i.e.,

$$\ker(\Phi) = \Big\{ \sum_{|S|\neq 1} \lambda_S b_S^\Psi \mid \lambda_S \in \mathbb{R} \Big\}. \tag{23}$$

This method can be readily applied to the Shapley and Banzhaf values since the Shapley and Banzhaf interaction transforms extend these values in the above sense:

$$\Phi_i^{\mathrm{Sh}}(v) = I^v(\{i\}), \quad \Phi_i^{\mathrm{B}}(v) = I_{\mathrm{B}}^v(\{i\}) \qquad (i \in N).$$

Example 3.1 *Applying the first construction to the Shapley value by means of the Shapley interaction transform, the representation (23) for valuations* $v \in \ker(\Phi^{\mathrm{Sh}})$ *yields*

$$v(S) = \sum_{T\subseteq N, |T|\neq 1} \lambda_T \beta_{|S\cap T|}^{|T|} \quad \textit{for all } S \subseteq N.$$

In the case $N = \{1, 2, 3\}$, one thus obtains

$$\begin{aligned}
v(\emptyset) &= \lambda_\emptyset + \frac{1}{6}(\lambda_{12} + \lambda_{13} + \lambda_{23}) \\
v(1) &= \lambda_\emptyset - \frac{1}{3}\lambda_{12} - \frac{1}{3}\lambda_{13} + \frac{1}{6}\lambda_{23} + \frac{1}{6}\lambda_{123} \\
v(2) &= \lambda_\emptyset - \frac{1}{3}\lambda_{12} + \frac{1}{6}\lambda_{13} - \frac{1}{3}\lambda_{23} + \frac{1}{6}\lambda_{123} \\
v(3) &= \lambda_\emptyset + \frac{1}{6}\lambda_{12} - \frac{1}{3}\lambda_{13} - \frac{1}{3}\lambda_{23} + \frac{1}{6}\lambda_{123} \\
v(12) &= \lambda_\emptyset + \frac{1}{6}\lambda_{12} - \frac{1}{3}\lambda_{13} - \frac{1}{3}\lambda_{23} - \frac{1}{6}\lambda_{123} \\
v(13) &= \lambda_\emptyset - \frac{1}{3}\lambda_{12} + \frac{1}{6}\lambda_{13} - \frac{1}{3}\lambda_{23} - \frac{1}{6}\lambda_{123} \\
v(23) &= \lambda_\emptyset - \frac{1}{3}\lambda_{12} - \frac{1}{3}\lambda_{13} + \frac{1}{6}\lambda_{23} - \frac{1}{6}\lambda_{123} \\
v(123) &= \lambda_\emptyset + \frac{1}{6}(\lambda_{12} + \lambda_{13} + \lambda_{23}).
\end{aligned}$$

The second approach allows us to construct a basis for the null space of an arbitrary linear value $\Phi : \mathbb{R}^{\mathcal{N}} \to \mathbb{R}^N$. Moreover, while the constructions in the literature often use the condition $\dim \Phi(\mathbb{R}^{\mathcal{N}}) = n$, our method is general and needs no *a priori* assumption on $\dim \Phi(\mathbb{R}^{\mathcal{N}})$.

Let $k = \dim \Phi(\mathbb{R}^{\mathcal{N}}) \leq n$ be the dimension of Φ and recall the well-known dimension formula for linear maps:

$$\dim \ker \Phi = \dim \mathbb{R}^{\mathcal{N}} - \dim \Phi(\mathbb{R}^{\mathcal{N}}) = 2^n - k. \tag{24}$$

In the case $k = 0$, one has $\ker \Phi = \mathbb{R}^{\mathcal{N}}$. So any basis of $\mathbb{R}^{\mathcal{N}}$ solves the inverse problem for Φ. Let us therefore investigate the non-trivial situation $k \geq 1$.

Select a basis $E = \{e_1, \ldots, e_k\}$ for the range $\Phi(\mathbb{R}^{\mathcal{N}})$ of Φ as well as k arbitrary valuations $b_1, \ldots, b_k \in \mathbb{R}^{\mathcal{N}}$ such that

$$\Phi(b_i) = e_i \quad (i = 1, \ldots, k).$$

Lemma 3.2 *The set $\{b_1, \ldots, b_k\}$ of valuations is linearly independent.*

Proof. Suppose the statement is false and there are non-trivial scalars λ_i such that

$$\sum_{i=1}^{k} \lambda_i b_i = 0 \in \mathbb{R}^{\mathcal{N}} \quad \text{and hence} \quad \Phi(\sum_{i=1}^{k} \lambda_i b_i) = 0 \in \mathbb{R}^N.$$

The linearity of Φ then implies $0 = \sum_{i=1}^{k} \lambda_i \Phi(b_i) = \sum_{i=1}^{k} \lambda_i e_i$, which contradicts the independence of the set $E \subseteq \mathbb{R}^N$, however. $\square$

By Lemma 3.2 and basic facts from linear algebra, $\{b_1, \ldots, b_k\}$ may be completed to a basis

$$B = \{b_1, \ldots, b_k, b_{k+1}, \ldots, b_{2^n}\}$$

for the domain $\mathbb{R}^{\mathcal{N}}$. Moreover, the basis E for the range $\Phi(\mathbb{R}^{\mathcal{N}})$ guarantees, for each $b_j \in B$, $j = k+1, \ldots, 2^n$, unique scalars $\epsilon_1^{(j)}, \ldots, \epsilon_k^{(j)}$ such that

$$\Phi(b_j) = \sum_{i=1}^{k} \epsilon_i^{(j)} e_i = \sum_{i=1}^{k} \epsilon_i^{(j)} \Phi(b_i).$$

Because Φ is linear, the valuations $b_j^{\Phi} = b_j - \sum_{i=1}^{k} \epsilon_i^{(j)} b_i$, $j = k+1, \ldots, 2^n$, are in $\ker \Phi$:

$$\Phi(b_j^{\Phi}) = \Phi(b_j) - \Phi(\sum_{i=1}^{k} \epsilon_i^{(j)} b_i) = 0.$$

We have thus arrived at a solution of the inverse problem.

Theorem 3.1 *Let $B^{\Phi} = \{b_1, \ldots, b_k, b_{k+1}^{\Phi}, \ldots, b_{2^n}^{\Phi}\}$. Then*

(i) B^{Φ} is a basis for $\mathbb{R}^{\mathcal{N}}$.

(ii) $B_0^{\Phi} = \{b_{k+1}^{\Phi}, \ldots, b_{2^n}^{\Phi}\}$ is a basis for $\ker \Phi$.

Proof. Every b_j is a linear combination of vectors in B^{Φ}:

$$b_j = b_j^{\Phi} + \sum_{i=1}^{k} \epsilon_i^{(j)} b_i \qquad (j = k+1, \ldots, 2^n).$$

Because B generates $\mathbb{R}^{\mathcal{N}}$, also B^{Φ} generates $\mathbb{R}^{\mathcal{N}}$. Because of $|B^{\Phi}| = 2^n$, B^{Φ} is linearly independent and, therefore a basis, which proves (i).

We have seen that $B_0^{\Phi} \subseteq \ker \Phi$ holds. Since $B_0^{\Phi} \subseteq B^{\Phi}$ is linearly independent and $|B_0^{\Phi}| = 2^n - k = \dim \ker \Phi$, B_0^{Φ} must be a basis of $\ker \Phi$. $\square$

We summarize the procedure to find a basis of the kernel:

(i) Select a basis $E = \{e_1, \ldots, e_k\}$ of the range $\Phi(\mathbb{R}^{\mathcal{N}})$.

(ii) Find valuations $b_1, \ldots, b_k \in \mathbb{R}^{\mathcal{N}}$ such that $\Phi(b_i) = e_i$, $i = 1, \ldots, k$.

(iii) Complete the independent set $\{b_1, \ldots, b_k\}$ to form a basis $B = \{b_1, \ldots, b_{2^n}\}$ of $\mathbb{R}^{\mathcal{N}}$.

(iv) Compute the coordinates $\epsilon_1^{(j)}, \ldots, \epsilon_k^{(j)}$ of $\Phi(b_j)$ in E for $j = k+1, \ldots, 2^n$.

(v) Compute $b_j^{\Phi} = b_j - \sum_{i=1}^{k} \epsilon_i^{(j)} b_i$ for $j = k+1, \ldots, 2^n$, which are the vectors of the basis of the kernel.

Assume, for example, that Φ satisfies the null player axiom and that $\Phi(v)$ is the null vector only if every player is null. Then $\Phi_i(\zeta_{\{i\}}) = \alpha_i$ for some $\alpha_i \neq 0$, and 0 for every other player. It follows that the basis in step (i) can be chosen as the set of all unit vectors e_i and the vectors b_i in step (ii) can be chosen as $\zeta_{\{i\}}$. Consequently, it suffices to take the collection of all unanimity games ζ_S, $|S| > 1$ to complete the basis in step (iii). Clearly, this works for a large class of linear values. We illustrate the method below with the Shapley value.

Application: The inverse Shapley value problem revisited. Our general construction includes Dragan's [3, 6] solution of the inverse problem for the (weighted) Shapley value Φ^{Sh} as a straightforward special case. To see this, consider the basis $B = \{\zeta_S \mid S \in \mathcal{N}\}$ of $\mathbb{R}^{\mathcal{N}}$. For each $i \in N$, we have

$$\Phi^{\mathrm{Sh}}(\zeta_i) = e_i = \ i\text{th unit vector in } \mathbb{R}^N,$$

which implies $\dim \ker \Phi^{\mathrm{Sh}} = 2^n - n$. For each coalition $S \neq \emptyset$, we have

$$\Phi^{\mathrm{Sh}}(\zeta_S) = \frac{1}{|S|} \sum_{i \in S} e_i = \frac{1}{|S|} \sum_{i \in S} \Phi^{\mathrm{Sh}}(\zeta_i) \quad \text{and thus} \quad \zeta_S^{\Phi^{\mathrm{Sh}}} = \zeta_S - \frac{1}{|S|} \sum_{i \in S} \zeta_i,$$

which yields the following set $B_0^{\Phi^{\mathrm{Sh}}}$ as a basis for the null space $\ker \Phi^{\mathrm{Sh}}$:

$$B_0^{\Psi^{\mathrm{Sh}}} = \{\zeta_\emptyset\} \cup \{\zeta_S^{\Phi^{\mathrm{Sh}}} \mid S \in \mathcal{N}, |S| \geq 2\}.$$

4 Potential functions and values

Given the cooperation system (N, F) with influence functions f_S, we associate with every $v \in \mathbb{R}^{\mathcal{N}}$ its F-*potential*

$$v^F = \sum_{S \in \mathcal{N}} v(S) f_S \quad (= vF \ \text{ in matrix notation}). \tag{25}$$

So the influence space $\mathcal{F}$ of F contains precisely the F-potentials:

$$\mathcal{F} = \{\sum_{S \in \mathcal{F}} \lambda_S f_S \mid \lambda_S \in \mathbb{R}\} = \{v^F \mid v \in \mathbb{R}^{\mathcal{N}}\}.$$

Note that $v \mapsto v^F$ is a linear operator on $\mathbb{R}^{\mathcal{N}}$ and is invertible (*i.e.*, every $v \in \mathbb{R}^{\mathcal{N}}$ is uniquely determined by its F-potential) if and only if the influence functions form a basis of $\mathbb{R}^{\mathcal{N}}$. In such a case, the potential corresponds to an inverse transform (and hence to a transform in its own right) in the sense of Section 3.

Potentials are closely related to linear values (see Section 2.3). Indeed, a pattern F gives rise to the (linear) *potential value* ∂^F, i.e., a linear map into $\mathbb{R}^N$ in the sense of Section 2.3, where

$$\partial_i^F(v) = v^F(N) - v^F(N \setminus i) \quad (i \in N).$$

Lemma 4.1 *Every linear value Φ arises as the potential value ∂^F relative to a suitable influence pattern F.*

Proof. Since $v \mapsto \Phi^v$ is a linear map, there is an influence pattern F such that $\Phi^v = vF = v^F$ holds for all $v \in \mathbb{R}^{\mathcal{N}}$. Hence we find for all $i \in N$,

$$\Phi_i(v) = \Phi^v(N) - \Phi^v(N \setminus i) = v^F(N) - v^F(N \setminus i) = \partial_i^F. \qquad \square$$

Although also Lemma 4.1 is a straightforward consequence of linear algebra, similarly to Lemma 3.1, it permits us a general view and a better understanding of the theory of potentials (as initiated by Hart and Mas-Colell), relating them to linear values. As a first consequence, we can derive Hart and Mas-Colell's well-known relation between the Shapley value and the potential, in a very simple way (see Theorem 4.1 below). A second consequence is the insight that any linear value can be obtained as the Shapley value of some F-potential (Theorem 4.2).

We consider the Shapley value and rewrite (4) in a more general form, where the game is restricted to the coalitions T contained in $S \subseteq N$:

$$\Phi_i^{\mathrm{Sh}}(v, S) = \sum_{T \subseteq S} \frac{(s-t)!(t-1)!}{s!}(v(T) - v(T \setminus i))$$

with $s = |S|$ and $t = |T|$. It is well-known and easy to see that for the influence functions ζ_U one has

$$\Phi_i^{\mathrm{Sh}}(\zeta_U, S) = \begin{cases} 1/u & \text{if } i \in U \subseteq S, \\ 0 & \text{otherwise.} \end{cases} \tag{26}$$

If $i \in U$, then $\zeta_U(T \setminus i) = 0$. So the second sum term in the expression for $\Phi^{\mathrm{Sh}}(\zeta_U, S)$ vanishes and we find for any coalition $U \neq \emptyset$,

$$\sum_{T \subseteq S} \frac{(t-1)!(s-t)!}{s!} \zeta_U(T) = \begin{cases} 0 & \text{if } U \not\subseteq S, \\ 1/u & \text{if } U \subseteq S. \end{cases}$$

Since $v \mapsto \Phi_i^{\mathrm{Sh}}(v, S)$ is a linear map, the Shapley value (for player i) can be equivalently defined as the linear functional with property (26) for all coalitions U. Setting

$$P^v(S) = \sum_{T \subseteq S} \frac{(t-1)!(s-t)!}{s!} v(T), \tag{27}$$

we see:

Theorem 4.1 (Hart and Mas-Colell [15]) *For every cooperative game (N, v) and player $i \in N$ one has*

$$\Phi_i^{\mathrm{Sh}}(v, N) = P^v(N) - P^v(N \setminus i).$$

Proof. By the linearity of $v \mapsto \Phi_i^{\mathrm{Sh}}(v, N)$, it suffices to verify the Theorem for potentials of the form $v = \zeta_U$. If $i \in U$, we have $P^{\zeta_U}(N \setminus i) = 0$ and therefore

$$\Phi_i^{\mathrm{Sh}}(\zeta_U, N) = 1/u = P^{\zeta_U}(N) - P^{\zeta_U}(N \setminus i).$$

If $i \notin U$, we have $P^{\zeta_U}(N \setminus i) = 1/u$ and thus

$$\Phi_i^{\mathrm{Sh}}(\zeta_U, N) = 0 = P^{\zeta_U}(N) - P^{\zeta_U}(N \setminus i).$$

□

Let $P = [p_{ST}]$ be the pattern with coefficients

$$p_{ST} = \begin{cases} 1 & \text{if } S = T = \emptyset, \\ (t-1)!(s-t)!/s! & \text{if } T \subseteq S \neq \emptyset, \\ 0 & \text{otherwise.} \end{cases} \tag{28}$$

Then Theorem 4.1 says that the Shapley value is also a potential value in our sense:

$$\Phi_i^{\mathrm{Sh}}(v, N) = P^v(N) - P^v(N \setminus i) = v^P(N) - v^P(N \setminus i) = \partial^P(v). \tag{29}$$

In fact, the Shapley value is the "typical" linear value:

Theorem 4.2 *Let Φ be an arbitrary linear value on $\mathbb{R}^{\mathcal{N}}$. Then there exists a pattern G such that Φ arises as the Shapley value relative to G:*

$$\Phi_i(v, N) = \Phi_i^{\mathrm{Sh}}(v^G, N) \quad \textit{for all } v \in \mathbb{R}^{\mathcal{N}},\ i \in N.$$

Proof. Notice that the pattern matrix P, given as in (28), admits an inverse P^{-1}. Indeed, arranging the rows and columns of P so that S always precedes T if $S \subset T$ holds, turns P into a triangular matrix with non-zero diagonal elements $p_{SS} \neq 0$.

By Lemma 4.1, Φ arises as the potential value relative to some pattern F. Letting $G = FP^{-1}$, we thus obtain for all $i \in N$,

$$v^F(N) - v^F(N \setminus i) = (v^G)^P(N) - (v^G)^P(N \setminus i) = \Phi_i^{\mathrm{Sh}}(v^G, N). \qquad \square$$

Monderer and Shapley [18] introduced (non-cooperative) *potential games* and embedded the Shapley value into the value theory of this class, thus establishing an important link between cooperative and non-cooperative game theory.

5 Concluding remarks

We have shown that basic models from linear algebra permit to revisit, extend and put into perspective many results and concepts of cooperative game theory. In this respect, the major achievements and "take-home messages" are:

- Bases of games and linear transforms acting on games (*e.g.*, Möbius transform, Shapley and Banzhaf interaction transforms, Fourier transform) are two faces of the same coin (Lemma 3.1). Consequently, a plentitude of new bases for games becomes available, each of them giving a specific representation of games which can be useful in practice (*e.g.*, representation of a game through interaction indices, Fourier coefficients, *etc.* in Sections 3.1 and 3.2).

- The inverse problem can be solved in an easy way for *any* linear value. The solution is particularly simple in the case of Shapley and Banzhaf values (Section 3.3).

- The Hart and Mas-Colell potential can be generalized and related to any linear value (Section 4).

Acknowledgment. The authors are grateful to the reviewers for their careful reading of the manuscript and their constructive remarks. The second author thanks the Agence Nationale de la Recherche for financial support under contract ANR-13-BSHS1-0010 (DynaMITE).

References

[1] J.F. Banzhaf. Weighted voting does not work: A mathematical analysis. *Rutgers Law Review*, 19:317–343, 1965.

[2] D. Denneberg and M. Grabisch. Interaction transform of set functions over a finite set. *Information Sciences*, 121:149–170, 1999.

[3] I. Dragan. The potential basis and the weighted Shapley value. *Libertas Mathematica*, 11:139–150, 1991.

[4] I. Dragan. On the inverse problem for semivalues of cooperative TU games. *International Journal of Pure and Applied Mathematics*, 22(4):539–555, 2005.

[5] I. Dragan. The least square values and the Shapley value for cooperative TU games. *TOP*, 14:61–73, 2006.

[6] I. Dragan. On the inverse problem for multiweighted Shapley values of cooperative TU games. *International Journal of Pure and Applied Mathematics*, 75:279–287, 2012.

[7] U. Faigle and J. Voss. A system-theoretic model for cooperation, interaction and allocation. *Discrete Applied Mathematics*, 159:1736–1750, 2011.

[8] M. Grabisch. k-order additive discrete fuzzy measures and their representation. *Fuzzy Sets and Systems*, 92:167–189, 1997.

[9] M. Grabisch. Set function over finite sets: transformations and integrals. In E. Pap, editor, *Handbook of Measure Theory*, pages 1381–1401. Elsevier Science Publ., 2002.

[10] M. Grabisch, J.-L. Marichal, and M. Roubens. Equivalent representations of set functions. *Mathematics of Operations Research*, 25(2):157–178, 2000.

[11] M. Grabisch and M. Roubens. An axiomatic approach to the concept of interaction among players in cooperative games. *Int. Journal of Game Theory*, 28:547–565, 1999.

[12] M. Grabisch and A. Rusinowska. A model of influence based on aggregation functions. *Mathematical Social Sciences*, 66:316-330, 2013.

[13] J. Gruska. *Quantum Computing.* McGraw-Hill, 1999.

[14] J. C. Harsanyi. A simplified bargaining model for the n-person cooperative game. *International Economic Review*, 4:194–220, 1963.

[15] S. Hart and A. Mas-Colell. Potential, value and consistency. *Econometrica*, 57:589–614, 1989.

[16] G. Kalai. A Fourier-theoretic perspective on the Condorcet paradox and Arrow's theorem. *Advances in Applied Mathematics*, 29:412–426, 2002.

[17] N. L. Kleinberg and J. H. Weiss. Equivalent n-person games and the null space of the Shapley value. *Mathematics of Operations Research*, 10(2):233–243, 1985.

[18] D. Monderer and L.S. Shapley. Potential games. *Games and Economic Behavior*, 14:124–143, 1996.

[19] G. Owen. Multilinear extensions of games. *Management Sci.*, 18:64–79, 1972.

[20] G. C. Rota. On the foundations of combinatorial theory I. Theory of Möbius functions. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 2:340–368, 1964.

[21] G. Shafer. *A Mathematical Theory of Evidence*. Princeton Univ. Press, 1976.

[22] L. S. Shapley. A value for n-person games. In H. W. Kuhn and A. W. Tucker, editors, *Contributions to the Theory of Games, Vol. II*, number 28 in Annals of Mathematics Studies, pages 307–317. Princeton University Press, 1953.

[23] J. Walsh. A closed set of normal orthogonal functions. *American Journal of Mathematics*, 45:5–24, 1923.

[24] K. Yokote, Y. Funaki, and Y. Kamijo. Linear basis to the Shapley value. Technical report, Waseda Economic Working Paper Series, 2013.