Part II
The simple Galton-Watson process: Genealogical approach

Chapter 4
The Ulam-Harris model and Galton-Watson trees

The purpose of this chapter is to lay the foundations for a study of Galton-Watson branching processes within an extended framework that beyond mere generation sizes also incorporates the genealogical structure of the considered population and therefore requires the introduction of labeled trees with a distinguished root as random elements on a suitable probability space.

4.1 Basic setup

In the following we will define labeled trees with a distinguished root in a canonical way as subsets of the infinite *Ulam-Harris tree* with vertex set

$$\mathbb{V} \;=\; \bigcup_{n\ge 0} \mathbb{N}^n,$$

where $\mathbb{N}^0 := \{\varnothing\}$ consists of the root [☞ also Section 1.2]. Each vertex $\mathsf{v} = \mathsf{v}_1 ... \mathsf{v}_n \in \mathbb{V}\setminus\{\varnothing\}$ is connected to the root via the unique shortest path

$$\varnothing \;\to\; \mathsf{v}_1 \;\to\; \mathsf{v}_1\mathsf{v}_2 \;\to\; ... \;\to \mathsf{v}_1 ... \mathsf{v}_n.$$

The the length of v is denoted by $|\mathsf{v}|$, thus $|\mathsf{v}_1 ... \mathsf{v}_n| = n$ and particularly $|\varnothing| = 0$. Further let $\mathsf{u}\mathsf{v} = \mathsf{u}_1 ... \mathsf{u}_m \mathsf{v}_1 ... \mathsf{v}_n$ denote the concatenation of two vectors $\mathsf{u} = \mathsf{u}_1 ... \mathsf{u}_m$ and $\mathsf{v} = \mathsf{v}_1 ... \mathsf{v}_n$.

Definition 4.1. A subset τ of $\mathbb{V}$ is called *(labeled) tree* if

(T1) $\varnothing \in \tau$.
(T2) $\mathsf{v}_1 ... \mathsf{v}_n \in \tau$ implies $\mathsf{v}_1 ... \mathsf{v}_k \in \tau$ for each $k = 1, ..., n-1$.
(T3) $\mathsf{v}_1 ... \mathsf{v}_n \in \tau$ implies $\mathsf{v}_1 ... \mathsf{v}_{n-1} j \in \tau$ for each $j \in \{1, ..., \mathsf{v}_n\}$.

If, furthermore,

(T4) $z_n(\tau) := |\tau \cap \mathbb{N}^n| < \infty$ for any $n \in \mathbb{N}_0$,

then τ is called *locally finite*. The elements of τ are called *nodes, vertices* or *individuals*, the individual $\varnothing$ is called the *root* or *ancestor*. Finally, the *height of* τ is defined as

$$H(\tau) \;=\; \sup\{n \geq 0 : z_n(\tau) > 0\} \;\in\; \overline{\mathbb{N}}_0.$$

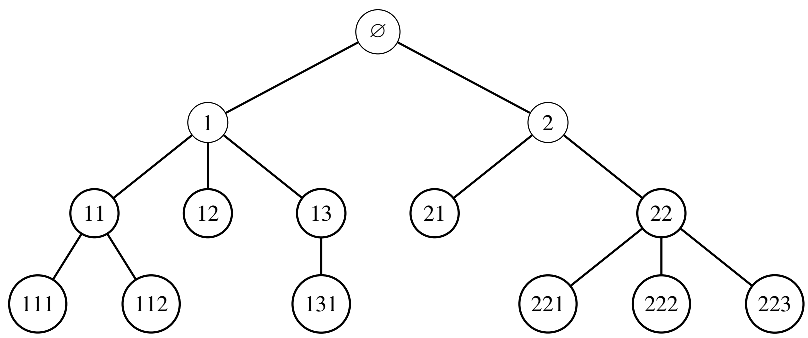

Fig. 4.1 A finite tree with Ulam-Harris labeling.

Since any tree considered hereafter has the described Ulam-Harris labeling and the distinguished root $\varnothing$ we may omit, unlike other texts, the attributes 'labeled' and 'rooted'. We are further interested only in locally finite trees and therefore define

$$\mathbb{T} \;=\; \{\tau \subset \mathbb{V} : \tau \text{ is a locally finite tree}\}.$$

The subset of finite trees is denoted by $\mathbb{T}^e$, i.e.

$$\mathbb{T}^e \;=\; \{\tau \in \mathbb{T} : |\tau| < \infty\} \;=\; \{\tau \in \mathbb{T} : H(\tau) < \infty\}.$$

In order to study random quantities taking values in $\mathbb{T}$, or functionals thereof, we must first endow this set with a suitable σ-field $\mathscr{T}$ so as to render measurability. This will be accomplished by defining a metric d on $\mathbb{T}$ and then choosing $\mathscr{T}$ as the associated Borel σ-field generated by the topology induced by d.

For $\tau \in \mathbb{T}$ and $n \in \mathbb{N}_0$, we define

$$\tau_n \;=\; \tau\cap\mathbb{N}^n \;=\; \{\mathsf{v}\in\tau : |\mathsf{v}|=n\},$$
$$\tau_{|n} \;=\; \bigcup_{k=0}^{n}\tau_k \;=\; \{\mathsf{v}\in\tau : |\mathsf{v}|\le n\},$$
$$[\tau]_n \;=\; \{\tau'\in\mathbb{T} : \tau'_{|n}=\tau_n\}.$$

Lemma 4.2. *Defining* $d:\mathbb{T}\times\mathbb{T}\to[0,1]$ *by*

$$d(\tau,\tau') \;=\; \exp\Big(-\sup\{n\ge 0 : \tau_{|n}=\tau'_{|n}\}\Big)$$

with $e^{-\infty}:=0$*, the pair* $(\mathbb{T},d)$ *forms a separable metric space with countable dense subset* $\mathbb{T}^e$*.*

Proof. To see that d is a metric, we must only verify the triangular inequality. But for $\tau,\tau',\tau''\in\mathbb{T}$, we have

$$\sup\{n\ge 0 : \tau_{|n}=\tau''_{|n}\} \;\ge\; \sup\{n\ge 0 : \tau_{|n}=\tau'_{|n}\}\wedge\sup\{n\ge 0 : \tau'_{|n}=\tau''_{|n}\}$$

and so

$$d(\tau,\tau'') \;\le\; d(\tau,\tau')\vee d(\tau',\tau'') \;\le\; d(\tau,\tau')+d(\tau',\tau'').$$

Any metric satisfying the first stronger form of the triangular inequality is called *ultrametric*.

Setting $\mathbb{T}_n=\{\tau\in\mathbb{T} : H(\tau)=n\}$ for $n\in\mathbb{N}_0$, we have

$$\mathbb{T}^e \;=\; \bigcup_{n\ge 0}\mathbb{T}_n,$$

and since any $\mathbb{T}_n$ is obviously countable, the same holds true for $\mathbb{T}^e$. Furthermore, for any $\tau\in\mathbb{T}$ and $n\in\mathbb{N}_0$, it follows that

$$d(\tau,\tau_{|n}) \;\le\; e^{-n}$$

which together with $\tau_{|n}\in\mathbb{T}^e$ for any n shows that $\mathbb{T}^e$ is a dense subset of $\mathbb{T}$. □

Remark 4.3. Although not needed for our purposes, we mention that $(\mathbb{T},d)$ is also complete and leave the proof as an exercise [☞ Problem 4.7].

Remark 4.4. For $\tau\in\mathbb{T}$ and $\varepsilon>0$, let

$$\mathbb{B}(\tau,\varepsilon) \;=\; \{\tau'\in\mathbb{T} : d(\tau,\tau')<\varepsilon\}$$

be the open ε-ball about τ with respect to d. Since d takes values only in the countable set $\{e^{-n} : n\in\overline{\mathbb{N}}_0\}$, we infer that $\mathbb{B}(\tau,\varepsilon)=\mathbb{T}$ if $\varepsilon>1$, and

$$\{\mathbb{B}(\tau,\varepsilon) : \tau\in\mathbb{T}, 0<\varepsilon\le 1\} \;=\; \{\mathbb{B}(\tau,e^{-n}) : \tau\in\mathbb{T}, n\ge 0\} \;=\; \{[\tau]_n : \tau\in\mathbb{T}, n\ge 1\},$$

where the second equality follows from

$$\begin{aligned} \tau' \in \mathbb{B}(\tau, e^{-n}) \quad &\Leftrightarrow \quad \sup\{k : \tau_{|k} = \tau'_{|k}\} > n \\ &\Leftrightarrow \quad \tau_{|n+1} = \tau'_{|n+1} \\ &\Leftrightarrow \quad \tau' \in [\tau]_{n+1} \end{aligned}$$

for any $n \in \mathbb{N}_0$. Next observe that, for any $\tau, \tau' \in \mathbb{T}$ and $n \geq k \geq 1$,

$$[\tau]_n \cap [\tau']_k \;=\; \{\chi \in \mathbb{T} : \chi_{|n} = \tau_{|n}, \chi_{|k} = \tau'_{|k}\} \;=\; \begin{cases} [\tau]_n, & \text{if } \tau_{|k} = \tau'_{|k}, \\ \emptyset, & \text{otherwise} \end{cases}$$

holds true. Consequently,

$$\mathscr{E} \;:=\; \{\emptyset, \mathbb{T}\} \cup \{[\tau]_n : \tau \in \mathbb{T}, n \geq 1\} \tag{4.1}$$

forms a $\cap$-stable system of open neighborhoods of $\mathbb{T}$.

We now define $\mathscr{T}$ as the σ-field generated by $\mathscr{E}$, i.e.

$$\mathscr{T} \;=\; \sigma(\mathscr{E}) \;=\; \sigma\left(\{[\tau]_n : \tau \in \mathbb{T}, n \geq 1\}\right). \tag{4.2}$$

Based on the previous considerations, the following lemma is easily proved.

Lemma 4.5. *The σ-field $\mathscr{T}$ defined in* (4.2) *equals the Borel σ-field $\mathscr{B}(\mathbb{T})$ induced by d, that is, the σ-field generated by the open subsets of $\mathbb{T}$ with respect to d.*

Proof. Clearly, $\mathscr{T} \subset \mathscr{B}(\mathbb{T})$. In view of Remark 4.4 it therefore suffices to note that any nonempty open subset of a separable metric space can be obtained as a countable union of ε-balls. □

We close this section with the definition of an ordering on $\mathbb{V}$ that reflects the kinship of its vertices when interpreted as individuals of a genealogical tree.

Definition 4.6. Let $\mathsf{v} = \mathsf{v}_1 \ldots \mathsf{v}_m$ and $\mathsf{w} = \mathsf{w}_1 \ldots \mathsf{w}_n$ be elements of $\mathbb{V}$.

(a) If $\mathsf{w} = \mathsf{vu}$ for some $\mathsf{u} \in \mathbb{V}$, then v is called an *ancestor* or *progenitor* of w and, conversely, w a *descendant* of v. In this case we write $\mathsf{v} \preceq \mathsf{w}$ or $\mathsf{w} \succeq \mathsf{w}$.
(b) If $\mathsf{u} \in \mathbb{N}$ in (a), then v is also called *mother* of w and, conversely, w *child* or *offspring* of v.

(c) Setting $\phi(\mathsf{v},\mathsf{w}) := \inf\{k \geq 1 : \mathsf{v}_k \neq \mathsf{w}_k\}$, the *most recent common ancestor (MRCA)* of v and w is defined as

$$\mathsf{v} \wedge \mathsf{w} := \mathsf{v}_1 \ldots \mathsf{v}_{\phi(\mathsf{v},\mathsf{w})-1}.$$

(d) In the case $\mathsf{v} \neq \mathsf{w}$, we further define

$$\mathsf{v} < \mathsf{w} \quad \overset{\text{def}}{\Longleftrightarrow} \quad \begin{cases} |\mathsf{v}| < |\mathsf{w}|, & \text{if } |\mathsf{v}| \neq |\mathsf{w}|, \\ \mathsf{v}_{\phi(\mathsf{v},\mathsf{w})} < \mathsf{w}_{\phi(\mathsf{v},\mathsf{w})}, & \text{if } |\mathsf{v}| = |\mathsf{w}| \end{cases}$$

and then generally $\mathsf{v} \leq \mathsf{w}$ iff $\mathsf{v} = \mathsf{w}$ or $\mathsf{v} < \mathsf{w}$.

With these definitions $\varnothing \preceq \mathsf{v}$ as well as $\varnothing \leq \mathsf{v}$ for any $\mathsf{v} \in \mathbb{V}$ holds true.

The relation ”$\leq$” introduced in (d) defines a total ordering on $\mathbb{V}$ which, when restricted to $\mathbb{N}$, coincides with the usual one. Thus, for any two elements v, w of $\mathbb{V}$, either $\mathsf{v} \leq \mathsf{w}$ or $\mathsf{w} \leq \mathsf{v}$ holds true, and each finite subset of $\mathbb{V}$ possesses a minimum and a maximum. If this subset consists of two elements v, w, then the minimum equals their MRCA $\mathsf{v} \wedge \mathsf{w}$ defined in (c) which explains the chosen notation ”$\wedge$”. Finally, we note that ”$\mathsf{v} < \mathsf{w}$” may be interpreted as ”v is older than w” when introducing a suitable ordering of ages of the individuals of $\mathbb{V}$. Details are left to the reader [☞ Problem 4.9].

Problems

Problem 4.7. Prove that the metric space $(\mathbb{T}, d)$ is complete.

Problem 4.8. Prove that

$$d_2(\tau, \tau') := \frac{1}{1 + \sup\{n \geq 0 : \tau_{|n} = \tau'_{|n}\}}$$

defines another metric on $\mathbb{T}$ which generates the same topology and thus the same Borel σ-field on $\mathbb{T}$ as the metric d.

Problem 4.9. Let $\mathsf{v}, \mathsf{w} \in \mathbb{V}$ and $\tau \in \mathbb{T}^e$ be a finite tree.

(a) Let $\mathsf{v} \vee \mathsf{w}$ be the maximum of v, w with respect to the ordering given in Def. 4.6(d). Give an intuitive characterization of this element.

(b) Give an intuitive characterization of the minimal and maximal element of τ.

Problem 4.10. Show that $\mathbb{V}$ is a multiplicative semigroup with neutral element $\varnothing$ when multiplication is defined as concatenation. [This provides a justification for the notation $\mathsf{v}_1 \ldots \mathsf{v}_n$ for $(\mathsf{v}_1, \ldots, \mathsf{v}_n)$.]

4.2 The Galton-Watson tree: formal definition and properties

Given an offspring distribution $(p_n)_{n\ge 0}$, we are now able to provide the formal definition of an associated *Galton-Watson tree (GWT)* $\boldsymbol{GW}$ as a random element in $(\mathbb{T},\mathscr{B}(\mathbb{T}))$, which is most conveniently accomplished in the following *standard model* similar to the one described in Section 1.1 for the GWP.

Let $\{X_{\mathsf{v}}:\mathsf{v}\in\mathbb{V}\}$ be a family of iid random variables with common distribution $(p_n)_{n\ge 0}$ and defined on a probability space $(\Omega,\mathfrak{A},\mathbb{P})$. Define $\boldsymbol{GW}_0=\{\varnothing\}$,

$$\boldsymbol{GW}_n \;=\; \{\mathsf{v}_1...\mathsf{v}_n\in\mathbb{N}^n:\mathsf{v}_1...\mathsf{v}_{n-1}\in\boldsymbol{GW}_{n-1}\text{ and }1\le\mathsf{v}_n\le X_{\mathsf{v}_1...\mathsf{v}_{n-1}}\}$$

for $n\ge 1$ (with the usual convention $\mathsf{v}_1...\mathsf{v}_{n-1}:=\varnothing$ if $n=1$) and finally

$$\boldsymbol{GW} \;=\; \bigcup_{n\ge 0}\boldsymbol{GW}_n.$$

Obviously, $\boldsymbol{GW}$ then is a $\mathbb{T}$-valued map when stipulating that edges are put between $\mathsf{v},\mathsf{w}\in\boldsymbol{GW}$ whenever w is a child of v. We further define (with $z_n:\mathbb{T}\to\mathbb{N}_0$ as in (T4) of Definition 4.1)

$$Z_n \;=\; |\boldsymbol{GW}_n| \;=\; z_n\circ\boldsymbol{GW}$$

for $n\in\mathbb{N}_0$.

Lemma 4.11. *The following assertions hold true for the previously defined mappings:*

(a) $\boldsymbol{GW}:\Omega\to\mathbb{T}$ *is* $\mathfrak{A}$-$\mathscr{B}(\mathbb{T})$*-measurable and thus a* $\mathbb{T}$*-valued random element defined on* $(\Omega,\mathfrak{A},\mathbb{P})$.
(b) *For any* $n\in\mathbb{N}_0$*, the mapping* $z_n:\mathbb{T}\to\mathbb{N}_0$ *is* $\mathscr{B}(\mathbb{T})$*-measurable and thus* $Z_n=z_n\circ\boldsymbol{GW}$ *an integer-valued random variable defined on* $(\Omega,\mathfrak{A},\mathbb{P})$.

Proof. (a) We must only show that $\boldsymbol{GW}^{-1}(\mathscr{E})\subset\mathfrak{A}$. But for $A=[\tau]_n$, $\tau\in\mathbb{T}$ and $n\in\mathbb{N}$, we obtain

$$\begin{aligned}\boldsymbol{GW}^{-1}(A) \;&=\; \{\omega\in\Omega:\boldsymbol{GW}_{|n}(\omega)=\tau_{|n}\}\\ &=\; \bigcap_{k=1}^{n}\{\boldsymbol{GW}_k(\omega)=\tau_k\}\\ &\in\; \sigma(\{X_{\mathsf{v}}:|\mathsf{v}|\le n-1\})\;\subset\;\mathfrak{A}.\end{aligned}$$

(b) Here it suffices to note that for each $n\in\mathbb{N}$

$$z_n^{-1}(\{k\}) \;=\; \{\tau\in\mathbb{T}:z_n(\tau)=k\} \;=\; \bigcup_{\tau\in\mathbb{T}_n:\tau_n=k}[\tau]_n,$$

which is an element of $\mathscr{B}(\mathbb{T})$ because the last union is countable. □

After these preparations we are ready to give the following definition of the *Galton-Watson measure*:

Definition 4.12. We put $\mathsf{GW} := \mathbb{P}(\boldsymbol{GW} \in \cdot)$ and call this distribution of $\boldsymbol{GW}$ the *Galton-Watson measure (GWM)* on $(\mathbb{T}, \mathscr{B}(\mathbb{T}))$ associated with $(p_n)_{n\geq 0}$.

It now follows for each $n \in \mathbb{N}_0$ that

$$\mathbb{P}(Z_n \in \cdot) \;=\; \mathbb{P}(z_n \circ \boldsymbol{GW} \in \cdot) \;=\; \mathsf{GW}(z_n \in \cdot)$$

as well as, more generally,

$$\mathbb{P}((Z_n)_{n\geq 0} \in \cdot) \;=\; \mathbb{P}((z_n \circ \boldsymbol{GW})_{n\geq 0} \in \cdot) \;=\; \mathsf{GW}((z_n)_{n\geq 0} \in \cdot).$$

Since $(Z_n)_{n\geq 0}$ is clearly a GWP with one ancestor and offspring distribution $(p_n)_{n\geq 0}$, the last relation shows that such a process may also be realized as a stochastic sequence defined on the probability space $(\mathbb{T}, \mathscr{B}(\mathbb{T}), \boldsymbol{GW})$ which in fact means nothing but realizing the generation sizes of a given population as functionals of the associated GWT, namely as its numbers of nodes at each level.

We proceed with a result that, in view of the stochastically independent and identical reproductive behavior of individuals in a GWT, should not take by surprise, namely that the subtrees of descendants generated by individuals of the same generation independent with common distribution $\boldsymbol{GW}$. To make this precise, some further definition are needed.

Given $\tau \in \mathbb{T}$ and $\mathsf{u} \in \tau$, we call

$$\tau^{\mathsf{u}} \;:=\; \{\mathsf{v} \in \tau : \mathsf{v} \succeq \mathsf{u}\}$$

the *subtree of* τ *rooted in* u. It consists of all individuals which are descendants of u including u itself. In order to identify τ^{u} in a unique way with an isomorphic element of $\mathbb{T}$, let Θ_{u} be the u-*shift*, defined on $\mathsf{u}\mathbb{V} := \{\mathsf{uv} : \mathsf{v} \in \mathbb{V}\}$ by

$$\Theta_{\mathsf{u}}(\mathsf{uv}) \;=\; \mathsf{v},$$

in particular $\Theta_{\mathsf{u}}(\mathsf{u}) = \varnothing$. Evidently, $\Theta_{\mathsf{u}}(\tau^{\mathsf{u}})$ then just equals the unique element of $\mathbb{T}$ that coincides with τ^{u} apart from a relabeling of its nodes such that the root becomes $\varnothing$. By a similar argument as in the proof of Lemma 4.11(a) one can easily verify the $\mathfrak{A}$-$\mathscr{B}(\mathbb{T})$-measurability of $\Theta_{\mathsf{u}} \circ \boldsymbol{GW}^{\mathsf{u}}$ for any $\mathsf{u} \in \boldsymbol{GW}$ [☞ Problem 4.14 for a precise formulation of this statement].

Proposition 4.13. *Let* $\boldsymbol{GW}$ *be a GWT with GWM* GW *and associated GWP* $(Z_n)_{n\geq 0}$. *Then the following assertion holds true for any* $n \in \mathbb{N}$: *If* $k \in \mathbb{N}$ *is such that* $\mathbb{P}(Z_n = k) > 0$, *then, given* $\boldsymbol{GW}_{|n} = \tau$ *for some* $\tau \in \mathbb{T}$ *with* $z_n(\tau) = k$, *the k shifted subtrees* $\Theta_{\mathsf{u}}\boldsymbol{GW}^{\mathsf{u}}$, $\mathsf{u}\in\tau_n$, *generated by the individuals of the* n^{th} *generation are conditionally iid with common distribution* GW.

Proof. Suppose that $\boldsymbol{GW}_{|n}=\tau$ for some $\tau\in\mathbb{T}$ with $z_n(\tau)=k$ and pick any w.l.o.g. nonempty $A_{\mathsf{u}}\in\mathscr{E}$ for $\mathsf{u}\in\tau_n$. With a suitable family $\{B_{\mathsf{v}}^{(\mathsf{u})}:\mathsf{v}\in\mathbb{V},\,\mathsf{u}\in\tau_n\}$ of subsets of $\mathbb{N}_0$, we then obtain

$$\begin{aligned}
\prod_{\mathsf{u}\in\tau_n}\mathsf{GW}(A_{\mathsf{u}}) &= \prod_{\mathsf{u}\in\tau_n}\prod_{\mathsf{v}\in\mathbb{V}}\mathbb{P}(X_{\mathsf{v}}\in B_{\mathsf{v}}^{(\mathsf{u})})\\
&= \frac{1}{\mathbb{P}(\boldsymbol{GW}_{|n}=\tau,Z_n=k)}\prod_{\mathsf{u}\in\tau_n}\prod_{\mathsf{v}\in\mathbb{V}}\mathbb{P}(Z_n=k)\,\mathbb{P}(X_{\mathsf{uv}}\in B_{\mathsf{v}}^{(\mathsf{u})})\\
&= \frac{1}{\mathbb{P}(\boldsymbol{GW}_{|n}=\tau,Z_n=k)}\mathbb{P}\left(\{Z_n=k\}\cap\bigcap_{\mathsf{u}\in\tau_n}\bigcap_{\mathsf{v}\in\mathbb{V}}\left\{X_{\mathsf{uv}}\in B_{\mathsf{v}}^{(\mathsf{u})}\right\}\right)\\
&= \frac{\mathbb{P}\left(\boldsymbol{GW}_{|n}=\tau,Z_n=k,\Theta_{\mathsf{u}}\boldsymbol{GW}^{\mathsf{u}}\in A_{\mathsf{u}}\text{ for }\mathsf{u}\in\tau_n\right)}{\mathbb{P}(\boldsymbol{GW}_{|n}=\tau,Z_n=k)}\\
&= \mathbb{P}\left(\Theta_{\mathsf{u}}\boldsymbol{GW}^{\mathsf{u}}\in A_{\mathsf{u}}\text{ for }\mathsf{u}\in\tau_n\,\middle|\,\boldsymbol{GW}_{|n}=\tau,Z_n=k\right)
\end{aligned}$$

having utilized that the X_{v} are iid. This proves the assertion because $\mathscr{E}$ is a $\cap$-stable generator of $\mathscr{B}(\mathbb{T})$ containing $\mathbb{T}$ [☞ e.g. [5, Thm. 22.2]]. □

Problems

Problem 4.14. Let $\boldsymbol{GW}$ be any GWT and $\mathsf{u}\in\mathbb{V}$. Prove that the map

$$\omega\;\mapsto\;\begin{cases}\Theta_{\mathsf{u}}\circ\boldsymbol{GW}^{\mathsf{u}}(\omega), & \text{if }\mathsf{u}\in\boldsymbol{GW}(\omega),\\ \{\varnothing\}, & \text{otherwise.}\end{cases}$$

from Ω to $\mathbb{T}$ is $\mathfrak{A}$-$\mathscr{B}(\mathbb{T})$-measurable.