

HAAR WAVELET AS A VARIATION ON FOURIER SERIES

JASON NGO

1. INTRODUCTION

Trigonometric polynomials, though most widely used in the computation and study of Fourier series, are not the only orthonormal basis for L^2 . It is known that they exhibit slow convergence and large oscillations near discontinuity points, also known as Gibbs Phenomenon (see Figure 1). Moreover, the Fourier coefficients, and hence the Fourier series expansion, depend on all values of the function; therefore, if a function f is changed on a small interval, it is possible that all the Fourier coefficients also change.¹

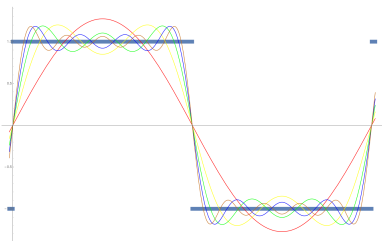


FIGURE 1. Fourier Series and Square Wave [9]

These failings of Fourier series motivate us to look for another expansion with the good properties of trigonometric polynomials (such as orthogonality and translation-invariant) but with better local properties: small change on one interval affects only a few of the series coefficients. This local property that we are seeking is extremely important for the analysis of signals with sudden transitions [5].

Luckily, bases that have such properties do exist, one of which is the Haar wavelet, first proposed in 1909 by Alfréd Haar and further developed in the 1980s (see Figure 2 for elements in the Haar system). Since then, we have seen rapid development in the theory of wavelets and its practical importance in various fields, namely signal denoising and image processing.

In this paper, we will prove that the Haar system is an orthonormal basis for L^2 and discuss its local property advantage over Fourier series expansion.

Date: 2019-05-03.

¹We will talk more about these shortcoming in Section 5.

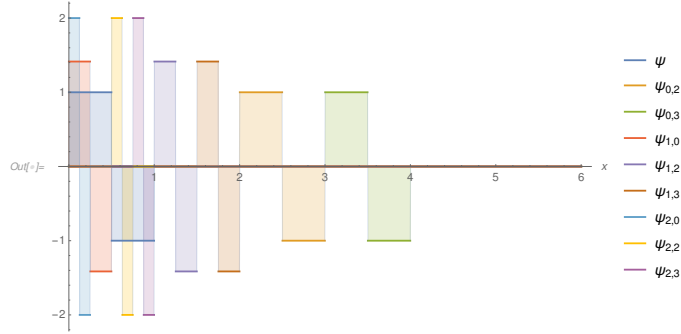


FIGURE 2. Elements of the Haar system

First, Section 2 will provide definitions of the Haar function and of an orthonormal wavelet. In Section 3, we will prove that the Haar system is orthonormal. Then, Section 4 will show that it spans $L^2(\mathbb{R})$. Finally, Section 5 presents a discussion of the Haar wavelet's local property and the absence of Gibbs phenomenon.

2. DEFINITION

To get started, recall that $L^2(\mathbb{R})$ is a vector space of square integrable functions, taken with the inner product

$$\langle f, g \rangle = \int_{\mathbb{R}} f(x)g(x)dx.$$

Definition 2.1 ([2, p. 515]). The *Haar function* is the function $\Psi = \chi_{[0,0.5)} - \chi_{[0.5,1)}$. The *Haar system* is the family

$$\{\Psi_{j,k}(x) = 2^{j/2}\Psi(2^j x - k) \mid j, k \in \mathbb{Z}\}.$$

Note that each term in the Haar system is constructed by translating and/or dilating the original Ψ (again, see Figure 2). By this construction, since $\int_{\mathbb{R}} \Psi(x)dx = 0$, it follows that $\int_{\mathbb{R}} \Psi_{j,k}(x)dx = 0$ for $j, k \in \mathbb{Z}$. Now, let us calculate the width and height for each $\Psi_{j,k}$. Because $\Psi(2^j - k) = 1$ when $0 \leq 2^j - k < \frac{1}{2}$, or equivalently, $\frac{k}{2^j} \leq x < \frac{k+\frac{1}{2}}{2^j}$, in this interval, we have $\Psi_{j,k} = 2^{j/2}$. Similar calculations yield:

$$(2.1) \quad \Psi_{j,k} = \begin{cases} 2^{j/2} & \text{if } \frac{k}{2^j} \leq x < \frac{k+\frac{1}{2}}{2^j}, \\ -2^{j/2} & \text{if } \frac{k+\frac{1}{2}}{2^j} \leq x < \frac{k+1}{2^j}, \\ 0 & \text{otherwise.} \end{cases}$$

Next, we will present and then prove in later sections that the Haar function Ψ satisfies Definition 2.2 of an *orthonormal wavelet*.

Definition 2.2 (Orthonormal Wavelet, [6, p. 303]). If $\Psi \in L^2(\mathbb{R})$, $\Psi_{j,k}(x) = 2^{j/2}\Psi(2^jx - k)$, and the set $\{\Psi_{j,k} : j, k \in \mathbb{Z}\}$ is an orthonormal basis for $L^2(\mathbb{R})$, then Ψ is called an *orthonormal wavelet*.

To prove the Haar system is an *orthonormal basis*, we have to show:

- (1) The Haar system $\{\Psi_{j,k}\}$ is an orthonormal set, i.e. $\|\Psi_{j,k}\| = 1$ for all $j, k \in \mathbb{Z}$ and $\langle \Psi_{j,k}, \Psi_{j',k'} \rangle = 0$ for all $(j, k) \neq (j', k')$.
- (2) The span of Haar system, denoted $\text{span}\{\Psi_{j,k}\}$, is dense in $L^2(\mathbb{R})$.

3. ORTHONORMALITY

First, we will start out by showing that the Haar system is orthonormal. The proof below is built upon the terse proof in [2], with computations and explanation added.

Lemma 3.1 ([2, p. 409]). *The Haar system is an orthonormal set in $L^2(\mathbb{R})$.*

Proof. First, it is easy to see that $\|\Psi\| = 1$. Let us compute:

$$\begin{aligned} \|\Psi_{j,k}\|^2 &= \int_{\mathbb{R}} \left| 2^{j/2}\Psi(2^jx - k) \right|^2 dx = \int_{\mathbb{R}} 2^j |\Psi(2^jx - k)|^2 dx \\ &= \int_{\mathbb{R}} |\Psi(y)|^2 dy = \|\Psi\|^2, \end{aligned}$$

by change of variable $y = 2^jx - k$ in the integral. Therefore, all the elements in the Haar system has norm 1.

Now, we want to show the orthogonality. Consider $\Psi_{j,k}$ and $\Psi_{j',k'}$, for $k \neq k'$. Since these two elements have disjoint support², their inner product is 0. If $j < j'$, then either $\Psi_{j,k}$ and $\Psi_{j',k'}$ have disjoint supports (when $k \neq k'$), or $\text{supp } \Psi_{j',k'}$ is contained in an interval on which $\Psi_{j,k}$ is constant to $2^{j/2}$ (when $k = k'$). For the former case, their inner product is obviously 0; for the latter case, we have:

$$\langle \Psi_{j,k}, \Psi_{j',k'} \rangle = \int_{\mathbb{R}} 2^{j/2} \cdot \Psi_{j',k'} dx = 2^{j/2} \int_{\mathbb{R}} \Psi_{j',k'} dx = 0.$$

Hence, the set is orthonormal. \square

Remark 3.2. The Haar function is a dyadic function, meaning that the dilations are taken to be powers of 2. Here, the factor $2^{j/2}$ is called the normalization factor, which is there so that the dilated and translated Haar function has norm 1. Although powers of 2 are a common choice, they are certainly not the most general one³.

²The *support* of a function, denoted supp , is the subset of the domain containing those elements which are not mapped to zero.

³See examples of non-dyadic wavelets and a discussion why they are more appropriate for statistical data analysis in [7].

4. HAAR WAVELET BASIS FOR $L^2(\mathbb{R})$

Now that we have shown the Haar system is orthonormal, this section will finish showing Ψ satisfies Definition 2.2 by proving the following theorem:

Theorem 4.1 ([2, p. 411]). *The Haar system $\{\Psi_{j,k}\}$ spans all of $L^2(\mathbb{R})$.*

To prove this theorem, we will start out with $C_c(\mathbb{R})$, the set of continuous functions with compact support; since $C_c(\mathbb{R})$ is dense in $L^2(\mathbb{R})$ ⁴, if we can show that any function $f \in C_c(\mathbb{R})$ is in the closure of the span of $\{\Psi_{j,k}\}$, then it follows that $\{\Psi_{j,k}\}$ spans $L^2(\mathbb{R})$.

Motivated by [1, p. 3] and [2, p. 516], we will set up the integral transform $P_n f$ in Section 4.1, and then in Section 4.2, we proceed to define the Haar inner product expansion and prove that it is dense in $C_c(\mathbb{R})$.

4.1. Integral Transform. Set $\phi(x) = \chi_{[0,1]}$. For $n \in \mathbb{Z}$, we define the integral kernel

$$K_n(x, y) = 2^n \sum_{k \in \mathbb{Z}} \phi(2^n x - k) \phi(2^n y - k).$$

Note that K_n is either equal to 2^n (when there is some $k \in \mathbb{Z}$ s.t. $2^n x - k$ and $2^n y - k$ lie in $[0, 1)$) or equal to 0 (otherwise). In other words, if $K_n = 2^n$, then there is some k s.t.

$$(4.1) \quad x, y \in \left[\frac{k}{2^n}, \frac{k+1}{2^n} \right) = I_{n,k}.$$

Here, note that each $I_{n,k}$ has length 2^{-n} and that the sets $\{I_{n,k}\}$ are disjoint.

We define the integral transform

$$P_n f(x) = \int_{\mathbb{R}} K_n(x, y) f(y) dy.$$

If $x \in \mathbb{R}$ then there is a unique $k_x \in \mathbb{Z}$ with $x \in I_{n,k_x}$ and

$$(4.2) \quad P_n f(x) = 2^n \int_{I_{n,k_x}} f(y) dy.$$

Since I_{n,k_x} has length 2^{-n} , we can say that $P_n f$ is the average value of a function f on the dyadic interval containing x .

Notice that the integral kernel K_n is dependent on ϕ . In order to write any function in terms of the Haar system $\{\Psi_{j,k}\}$, we must eliminate ϕ somehow from our calculations. The following Lemma 4.2 will write the integral kernels in terms of $\{\Psi_{j,k}\}$.

⁴The proof that $C_c(\mathbb{R})$ is dense in $L^2(\mathbb{R})$ can be found in [3, p. 326].

Lemma 4.2 ([6, p. 293]). *If $n \in \mathbb{Z}$, then*

$$K_{n+1} - K_n = \sum_{k \in \mathbb{Z}} \Psi_{n,k}(x) \Psi_{n,k}(y), \quad x, y \in \mathbb{R}.$$

Proof. First, let us calculate the right-hand side. By Equation (2.1), $\Psi_{n,k}(x) = 2^{n/2}$ when $x \in I_{n+1,2k} = [\frac{2k}{2^{n+1}}, \frac{2k+1}{2^{n+1}})$, and similarly, $\Psi_{n,k}(x) = -2^{n/2}$ when $x \in I_{n+1,2k+1}$. Therefore, it follows that:

$$\Psi_{n,k}(x) \Psi_{n,k}(y) = \begin{cases} 2^n & \text{if } x, y \in I_{n+1,2k} \text{ or } x, y \in I_{n+1,2k+1}, \\ -2^n & \text{if one is in } I_{n+1,2k} \text{ and the other is in } I_{n+1,2k+1}, \\ 0 & \text{otherwise.} \end{cases}$$

Now, before calculating the left-hand side $K_{n+1} - K_n$, note that:

$$I_{n+1,2k} \cup I_{n+1,2k+1} = \left[\frac{2k}{2^{n+1}}, \frac{2k+2}{2^{n+1}} \right) = I_{n,k}.$$

Therefore, if both x, y are in $I_{n+1,2k}$ or $I_{n+1,2k+1}$, then we have (1) $K_{n+1} = 2^{n+1}$ (by Equation (4.1)) and (2) $K_n = 2^n$ (because $x, y \in I_{n,k}$). Hence,

$$K_{n+1}(x, y) - K_n(x, y) = 2^{n+1} - 2^n = 2^n.$$

By similar argument, if either x or y is in $I_{n+1,2k}$ and the other is in $I_{n+1,2k+1}$, then $K_{n+1}(x, y) - K_n(x, y) = 0 - 2^n = -2^n$.

Otherwise, both K_{n+1} and K_n are 0. \square

Remark 4.3. From this Lemma, the difference between two integral transforms can be written as the *inner product expansion* of the function f :

$$\begin{aligned} (P_{n+1} - P_n)f(x) &= \int_{\mathbb{R}} K_{n+1}(x, y) f(y) dy - \int_{\mathbb{R}} K_n(x, y) f(y) dy \\ &= \int_{\mathbb{R}} \sum_{k \in \mathbb{Z}} \Psi_{n,k}(x) \Psi_{n,k}(y) f(y) dy \\ &= \sum_{k \in \mathbb{Z}} \Psi_{n,k}(x) \int_{\mathbb{R}} \Psi_{n,k}(y) f(y) dy \\ (4.3) \quad &= \sum_{k \in \mathbb{Z}} \langle f, \Psi_{n,k} \rangle \Psi_{n,k}(x). \end{aligned}$$

The terms $\langle f, \Psi_{j,k} \rangle$ are called the Haar coefficients. One can easily notice that this inner product expansion is only one-sided, meaning we keep the dilation factor $j = n$ fixed and vary the translation factor k . In the next section, we will abstract this expansion to a more general setting, varying both the translation and dilation factors.

4.2. Two-sided Haar Representation. From Equation (4.3), we can write

$$(4.4) \quad P_{n+1}f(x) - P_{-m}f(x) = \sum_{j=-m}^n \sum_{k \in \mathbb{Z}} \langle f, \Psi_{j,k} \rangle \Psi_{j,k}(x).$$

To prove that any function $f \in C_c(\mathbb{R})$ is in the closure of the span of $\{\Psi_{j,k}\}$, we want to show that the Haar inner product expansion on the right-hand side converges to f . By Equation (4.4) above, it suffices to show $P_{-m}f \rightarrow 0$ and $P_{n+1}f \rightarrow f$ with respect to the L^2 norm as $m, n \rightarrow \infty$. The two claims are presented as Lemmas below.

First, Lemma 4.4 tells us that the larger the intervals on which we average a continuous function, the smaller the supremum of the averaged function. The proof below is an extended version of the proof found in [6], with step-by-step calculations added.

Lemma 4.4 (Average Function over Dyadic Intervals, [6, p. 295]). *If $f \in C_c(\mathbb{R})$, then $\|P_{-m}f\|_2 \rightarrow 0$ when $m \rightarrow \infty$.*

Proof. Suppose $f \in C_c(\mathbb{R})$ and $\text{supp}(f) \subseteq [-K, K]$. There exists $M > 0$ s.t. $\text{supp}(f) \subseteq [-2^M, 2^M]$. Now, if $m > M$, then $\text{supp}(f) \subseteq I_{-m,-1} \cup I_{-m,0}$; Equation (4.2) therefore holds. We compute:

$$\begin{aligned} \|P_{-m}f\|_2^2 &= \int_{\mathbb{R}} |P_{-m}f(x)|^2 dx = \sum_{k \in \mathbb{Z}} \int_{I_{-m,k}} |P_{-m}f(x)|^2 dx \\ &= \sum_{k \in \mathbb{Z}} \int_{I_{-m,k}} \left| 2^{-m} \int_{I_{-m,k}} f(y) dy \right|^2 dx = 2^m 2^{-2m} \sum_{k \in \mathbb{Z}} \left| \int_{I_{-m,k}} f(y) dy \right|^2. \end{aligned}$$

Since f is only supported on a subset of $I_{-m,-1}$ and $I_{-m,0}$, we have:

$$\begin{aligned} \|P_{-m}f\|_2^2 &= 2^{-m} \left(\left| \int_{I_{-m,-1}} f(y) dy \right|^2 + \left| \int_{I_{-m,0}} f(y) dy \right|^2 \right) \\ &= 2^{-m} \left(\left| \int_{-K}^0 f(y) dy \right|^2 + \left| \int_0^K f(y) dy \right|^2 \right). \end{aligned}$$

Note that $\int_{-K}^0 f(y) dy = \langle f(y), \chi_{[-K,0)} \rangle$; by Cauchy-Schwarz inequality,

$$\|P_{-m}f\|_2^2 \leq 2^{-m} (K\|f\|_2^2 + K\|f\|_2^2) = 2^{-m} \cdot 2K\|f\|_2^2.$$

As $m \rightarrow \infty$, the right-hand side converges to 0. Therefore, by Squeeze Theorem, it follows that $\|P_{-m}f\|_2 \rightarrow 0$. \square

Now, before proving our next Lemma, recall that $P_n f(x)$ is the average value of a function f on a small dyadic interval of length 2^{-n} . If we increase n , then the intervals will get smaller, and the average value $P_n f$ will get closer to

f (see Figure 3 for the sequence $\{P_n f\}$ approaching $f(x) = 0.5 + e^{-x} \sin 2\pi x$ on the unit interval). The proof for Lemma 4.5 follows from this intuition.

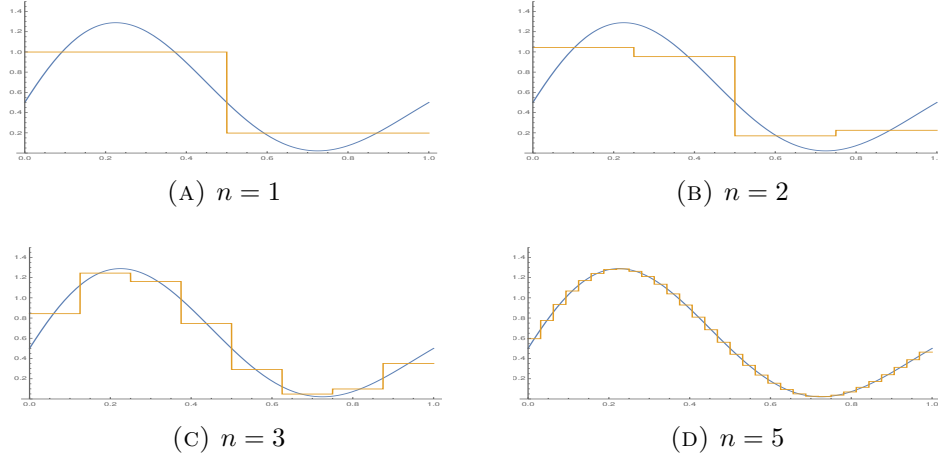


FIGURE 3. Approximation of $f(x) = 0.5 + e^{-x} \sin 2\pi x$ by $P_n f(x)$

Lemma 4.5 (Convergence of Integral Transform, [2, p. 517]). *If $f \in C_c(\mathbb{R})$, then $P_n f \rightarrow f$ in the uniform norm as well as in the L^2 norm.*

Proof. Suppose $f \in C_c(\mathbb{R})$ and $\text{supp}(f) \subset [-2^M, 2^M]$ for $M \geq 0$. Since f is continuous on a compact set, it is uniformly continuous.

Given $\epsilon > 0$. By uniform continuity, there exists $\delta > 0$ such that if $|x - y| < \delta$, then $|f(x) - f(y)| < \epsilon$.

Now, choose N such that $2^{-N} < \delta$; if $n > N$, then we have:

$$\begin{aligned}
 |P_n f(x) - f(x)| &= \left| 2^n \int_{I_{n,k_x}} f(y) dy - f(x) \right| \\
 &= \left| 2^n \int_{I_{n,k_x}} f(y) dy - 2^n \int_{I_{n,k_x}} f(x) dy \right| \\
 &\leq 2^n \int_{I_{n,k_x}} |f(y) - f(x)| dy < 2^n \int_{I_{n,k_x}} \epsilon dy = \epsilon.
 \end{aligned}$$

Therefore, $P_n f$ converges to f uniformly. Since uniform convergence implies L^2 convergence, it follows that $P_n f \rightarrow f$ with respect to the L^2 norm. \square

Now, we are ready to prove that the Haar system spans all of $L^2(\mathbb{R})$:

Proof of Theorem 4.1. By Lemma 4.4 and Lemma 4.5, it follows that for any $f \in C_c(\mathbb{R})$,

$$f = \sum_{j \in \mathbb{Z}} \sum_{k \in \mathbb{Z}} \langle f, \Psi_{j,k} \rangle \Psi_{j,k}.$$

Since $C_c(\mathbb{R})$ is a dense subspace of $L^2(\mathbb{R})$, we conclude that the Haar system spans $L^2(\mathbb{R})$. The Haar function Ψ is therefore an orthonormal wavelet; we will now refer to it as the *Haar wavelet*. \square

5. HAAR WAVELET'S LOCAL PROPERTY

So far we have seen the Haar wavelet has the good properties of trig functions: $\{\Psi_{j,k}\}$ is orthonormal and spans $L^2(\mathbb{R})$. In this final section, we will discuss in detail how this Haar wavelet basis helps resolve the issues for which Fourier series work badly.

5.1. Preservation of Series Coefficients. As opposed to the Fourier coefficients, each Haar coefficient $\langle f, \Psi_{j,k} \rangle$ depends only on the values of the function on $\text{supp } \Psi_{j,k}$. Therefore, when we change a function on an interval E , only the Haar coefficients $\langle f, \Psi_{j,k} \rangle$ with $\text{supp } \Psi_{j,k}$ overlapping with E change; the rest of the Haar coefficients remain unchanged. This preservation of series coefficients is desired when we want to modify the pitch of a signal at a point in time, yet unfortunately it does not apply to Fourier series.

Example 5.1. Consider $f(x) = 0.5 + e^{-x} \sin 2\pi x$ and $g(x) = f(x) + 0.1\chi_{[0.6, 0.61]}$.

Using **Mathematica**, we notice that all of the first 100 Fourier coefficients A_n of $f(x)$ and $g(x)$ change (see Figure 4). Although these changes are small, they pose unnecessary recalculations to the computation of Fourier series.

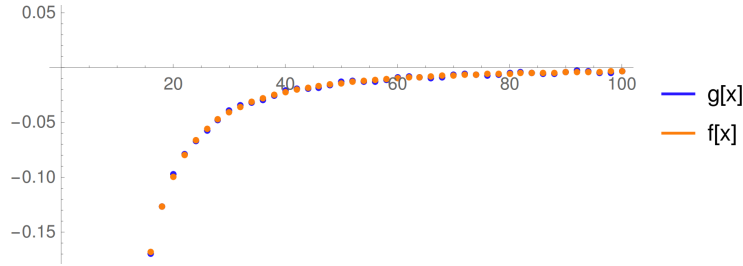


FIGURE 4. Fourier coefficients A_n of $f(x)$ and $g(x)$

5.2. Gibbs Phenomenon. Since the square wave presented in Figure 1 is a piecewise constant function, we can be certain that there is no Gibbs phenomenon when we approximate using the Haar wavelet basis, for we can always translate/dilate Haar functions to obtain the wave. For a more general case, it turns out that the Haar series expansions does not exhibit Gibbs phenomenon when approximating any $f \in L^2(\mathbb{R})$ either⁵. This property is highly desirable in signal processing since it does not cause clipping from the large oscillations near discontinuity points, as opposed to signal processing with Fourier series.

⁵The proof for this claim can be found in [8, Chap. 5] and [4].

REFERENCES

- [1] Jordan Bell. “Haar wavelets and multiresolution analysis”. 2014.
- [2] Kenneth R. Davidson and Allan P. Donsig. *Real Analysis with Real Applications*. Prentice Hall, 2002. ISBN: 0-13-041647-9.
- [3] RH Farrell. “Dense Algebras of Functions in L^p ”. In: *Proceedings of the American Mathematical Society* 13.2 (1962), pp. 324–328.
- [4] Susan E Kelly. “Gibbs phenomenon for wavelets”. In: *Applied and Computational Harmonic Analysis* 3.1 (1996), pp. 72–81.
- [5] BY Lee. “Application of the discrete wavelet transform to the monitoring of tool failure in end milling using the spindle motor current”. In: *The International Journal of Advanced Manufacturing Technology* 15.4 (1999), pp. 238–243.
- [6] Mark A Pinsky. *Introduction to Fourier analysis and wavelets*. Vol. 102. American Mathematical Soc., 2008.
- [7] Stephen Pollock and Iolanda Lo Cascio. “Non-dyadic wavelet analysis”. In: *Optimisation, Econometric and Financial Analysis*. Springer, 2007, pp. 167–203.
- [8] Kouros R. Raeen. “A study of the Gibbs phenomenon in Fourier series and wavelets”. PhD thesis. University of New Mexico, 2008.
- [9] Eric W. Weisstein. *Fourier Series–Square Wave*. From MathWorld–A Wolfram Web Resource. Visited on 30/04/19. URL: <http://mathworld.wolfram.com/FourierSeriesSquareWave.html>.