Chapter 3

The Continuous Wavelet Transform

The wavelet analysis has been implemented in the studies of the large-scale structure and CMB anisotropy because of its capability to decompose the cosmological perturbations into space and scale simultaneously. In non-Gaussianity detections, its implementation is motivated from the difficulty to detect the weak non-Gaussian signature in real and scale space due to the obscuring effect from the central limit theorem of random field [17] [18] [19]; on the contrary, the central limit theorem effect is reduced very much by the detections in wavelet space. Using wavelets, moreover, we can detect the non-Gaussian signature in each interval of scale which is useful to specify the sources of non-Gaussianity, because different scales of the anisotropy is affected from different cosmological processes.

In this thesis, we review the works of the non-Gaussianity investigation using the spherical continuous wavelet transform. In this chapter, we introduce the continuous wavelet transform, a version of wavelet transform, and its properties that make it an optimal method for detecting the non-Gaussianity. Then we introduce the 2-D continuous wavelet transform.

3.1 Space-Scale Analysis

The wavelet analysis is a space-scale (or time-frequency) analysis. While the Fourier analysis represents only the frequency of a signal, a time-frequency analysis provides both time and frequency localization. The Fourier analysis, which has been well-known in physics, gives an interpretation of a function as a superposition of all modes of sinusoidal waves and is very useful in physics. However, it is inappropriate in some situations that have the characteristic of localization in real space such as a transient, a spike and a point source, etc. In other words, the information of the position is diffused through Fourier transformation. While the Fourier spectrum can represent the amplitude of each Fourier mode, the information of localization in real space is diffused into the phases of all modes. For example, the Dirac delta function $\delta(x - x_0)$ which is the extreme case of the localization in real space around x_0 has the Fourier transform whose amplitude of all modes are the same and the phase factor of every mode k is expressed as e^{ikx_0}. In addition, the Fourier transform is very uneconomic. One needs nearly infinite number of modes to reconstruct a function whose little information lies on a small segment because the sum of large number of modes is required to recover the major region with zero value. A lack of any mode would change a reconstruction of the function very much. In other words, we can say that the Fourier analysis is unstable to a perturbation localized in real space.

To overcome these problems, several space-scale (or time-frequency) representations have been proposed. In physics, for example, the Wigner functions in quantum mechanics and the Gabor transform, the first version of the windowed Fourier transform, were early approaches. In the present, the wavelet transform is a widely accepted method in space-scale decomposition. There are two versions of the wavelet transforms: the continuous wavelet transform (CWT) and discrete wavelet transform (DWT). While the CWT is an integral transform analogous to the Fourier transform, the DWT is presented as the series expansion in the wavelet bases which is the analogue of the Fourier series. The DWT is very useful in several subjects such as signal processing, image processing, etc. Indeed, it has been also used for detecting the non-Gaussianity in the CMB, including the early works, but only in the 2-D plane corresponding to a small patch on the sky. In this work, however, we consider the wavelet analysis of all-sky CMB anisotropy using the spherical Mexican hat wavelets which is a CWT, hence we choose to describe only the properties of the continuous wavelet transform.

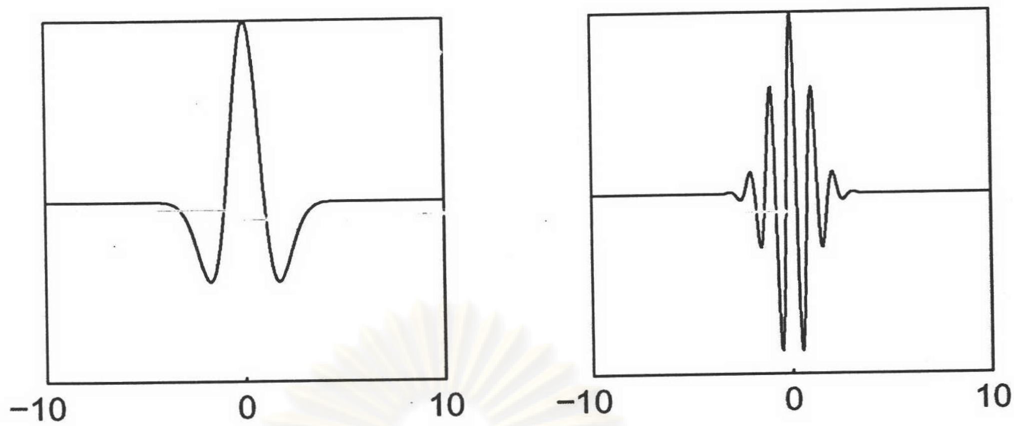

Figure 3.1: Two examples of wavelets: (left) The Mexican hat wavelet and (right) The Morlet wavelet [20].

3.2 The Continuous Wavelet Transform

The term "wavelet" reflects its characteristic which is like a small wave: a wavelet $\psi(x)$ is an oscillatory function implied by

$$\int \mathrm{d}x \, \psi(x) = 0, \tag{3.1}$$

and has the energy concentrated on a finite region (see Fig. 3.1). This characteristic causes the localization in both space and scale of a continuous wavelet transform (CWT), as can be seen later. From this property, we can have the information of a function covering the whole of position-scale plane by translations and dilations (or contractions) of a *mother* wavelet. Note that any wavelet acquired from a translation and dilation of the mother wavelet and the mother wavelet should have the same norm, usually the unit norm[1] in order to have the unique weight for all regions of space-scale plane.

A wavelet built from a translation b and dilation a, where $b \in \mathbf{R}, a \in \mathbf{R}^+$, of a mother wavelet $\psi(x)$ localized around $x = 0$ is

$$\psi_{a,b}(x) = \frac{1}{\sqrt{a}} \psi\left(\frac{x-b}{a}\right), \tag{3.2}$$

[1]Here, the unit norm of the wavelets is defined as $\int_{-\infty}^{\infty} \mathrm{d}x |\psi_{a,b}(x)|^2 = \int_{-\infty}^{\infty} \mathrm{d}x |\psi(x)|^2 = 1$, which is called in mathematics L^2 norm.

where $\frac{1}{\sqrt{a}}$ on the right hand side is the normalization factor. A continuous wavelet transform (CWT) of a function $f(x)$ on the real line, $x \in \mathbf{R}$, with respect to the continuous wavelets $\psi_{a,b}(x)$ is

$$w(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} \mathrm{d}x \, \overline{\psi\left(\frac{x-b}{a}\right)} f(x), \tag{3.3}$$

where the overbar denotes complex conjugation and $\psi(x)$ satisfies *the admissibility condition*:

$$C_\psi \equiv \int_{-\infty}^{\infty} \frac{\mathrm{d}\omega}{|\omega|} |\hat{\psi}(\omega)|^2 < \infty, \tag{3.4}$$

where

$$\hat{\psi}(\omega) = \int_{-\infty}^{\infty} \mathrm{d}x \, e^{-i\omega x} \psi(x) \tag{3.5}$$

is the Fourier transform of the wavelet $\psi(x)$, in order to have the invertibility of the wavelet transform. When the admissibility condition, $C_\psi < \infty$, is held, we have the inverse wavelet transform, or the reconstruction formula,

$$f(x) = \frac{1}{C_\psi} \int_{-\infty}^{\infty} \frac{\mathrm{d}a \, \mathrm{d}b}{a^2} w(a,b) \psi_{a,b}(x). \tag{3.6}$$

The admissibility condition implies that $\hat{\psi}(\omega) = 0$ at $\omega = 0$. This means that the wavelet ψ has zero mean following Eq. (3.1). In other words, this condition implies the wave-like characteristic. Again, the wavelet is like a "small wave" whose amplitude decays outside a band of space in which its energy concentrate. Additionally, the wavelet is local in scale space within a scale band which is implied by the mathematical conditions $\hat{\psi}(\omega) = 0$ at $\omega = 0$ and $\hat{\psi}(\omega) = 0$ as $|\omega| \to \infty$. The localizations in both space and scale of the wavelet imply that the continuous wavelet transform has this property too. Obviously, the localization in space can be seen from Eq. (3.3) that the integration of $f(x)$ with a wavelet $\psi_{a,b}(x)$ cut up the function $f(x)$ in the region around the position b. The localization in scale of the CWT can be seen indirectly from Eq. (3.3). From the convolution theorem, which states that

$$\widehat{f * g}(\omega) = \hat{f}(\omega) \, \hat{g}(\omega), \tag{3.7}$$

where the star denotes the convolution operation: $f * g(y) = \int \mathrm{d}x \, f(y-x) g(x)$, we have the absolute of the Fourier transform of the CWT

$$|\hat{w}(a,\omega)| = |\sqrt{a} \, \hat{\psi}(a\omega)| |\hat{f}(\omega)|. \tag{3.8}$$

This equation tells us that the Fourier transform of the CWT is non-zero only at the scale that the Fourier transform of the wavelet is non-zero; hence, the CWT cut up the scale information of a function on a finite band. For a mother wavelet $\psi(x)$ localized at $x = 0$ and having the Fourier transform $\hat{\psi}(\omega)$ localized around the scale $\omega = \omega_0$, the CWT with respect to the wavelet $\psi_{a,b}(x)$ has the localization property at the position $x = 0$ and the scale $\omega = \omega_0/a$.

Another property of wavelets involves the uncertainty principle. Assume that the support, which is the quantity that represents the space interval in which the wavelet is well non-zero, of the mother wavelet $\psi(x)$ has the length L and that of $\hat{\psi}(\omega)$ has the size Ω, centered around the mean ω_0. The forms of the wavelet and its Fourier transform: $a^{-1/2}\psi(\frac{x-b}{a})$ and $a^{1/2}\hat{\psi}(a\omega)$, respectively, imply that the wavelet $\psi_{a,b}(x)$ has the support of length aL around $x = b$ and its Fourier transform, $a^{1/2}\hat{\psi}(a\omega)$, has the support Ω/a around $\omega = \omega_0/a$. Since the CWT with respect to $\psi_{a,b}(x)$ cut up a function into a piece of space with support's size aL around a position b and, similarly, cut up into an interval of scale space with size Ω/a around ω_0/a, the supports of wavelet can be interpreted as the uncertainty of "measuring" the position and scale by the CWT tool: $\Delta x = aL$ and $\Delta p = \Omega/a$. Obviously, the wavelet tool has the property of the uncertainty principle, namely,

$$\Delta x \Delta p = L\Omega = \text{constant}, \tag{3.9}$$

which implies the preservation of the area of uncertainty box $\Delta x \Delta p$ (see Fig. 3.2). In addition, the CWT probe a small scale (large ω) with a narrow window in position space (small Δx) while a large scale (small ω) is probed with a wide window, which we can see from the centered "frequency" $\omega_* \propto a^{-1}$ and the support of wavelets $\Delta x \propto a$ (see Fig. 3.2). This is reasonable because we should have the size of space window comparable to the scale we are interested. If we have a small window to probe the structure with large wavelength, the information that we obtain do not cover one wavelength which is not enough to specify the "frequency". In contrast, if the window is large relative to the wavelength of structure, the localization in position space is lost; the window covers large number of regions with different wavelengths. This causes the wavelet analysis to give a reasonable space-scale decomposition.

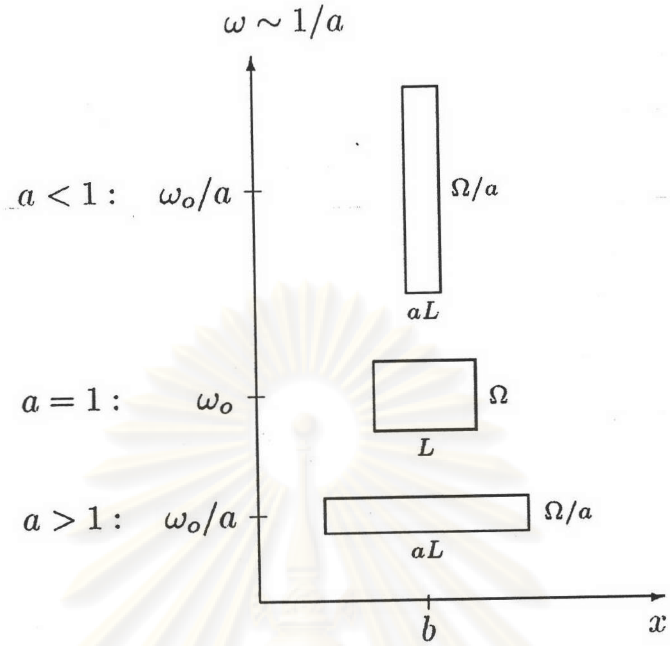

Figure 3.2: Space-frequency uncertainty window at different frequency derived by varying the dilation a: $a < 1$, $a = 1$ and $a > 1$, but at the same position b [20].

3.3 The CWT in Two dimensions

3.3.1 The Basic Properties

The CWT can be extended to two dimensions with the basic properties completely analogous to those in the one dimensional case, i.e. the admissibility condition, the analysis and reconstruction formula. However, the extension changes the number of operations for the construction of wavelets from the mother wavelet. While the wavelets in one dimension can be built from their mother wavelet with only two operations: the translation and dilation, those in two dimensions are in general constructed from the mother wavelet with four operations: two translations, the rotation around the wavelet itself and the global dilation. Certainly, one more translation is expected for translating the wavelet to cover the two dimensional space, but the rotation maybe a surprise at first sight. The rotation is considered when the wavelet which is a localized function embedded on the plane has the directional preference. In fact, this is common for any object in two dimensions.

In order to hold the reconstruction property, as in one dimension, the mother wavelet in two dimensions $\psi(\mathbf{x})$, where $\mathbf{x}$ is the position vector in two dimensions, satisfies the admissibility condition:

$$0 < C_\psi \equiv (2\pi)^2 \int \frac{d^2\mathbf{k}}{|\mathbf{k}|^2} |\hat{\psi}(\mathbf{k})|^2 < \infty, \tag{3.10}$$

where $\hat{\psi}(\mathbf{k})$ is the Fourier transform of $\psi(\mathbf{x})$, which implies that, as in one dimension, the wavelet has zero mean:

$$\hat{\psi}(\mathbf{0}) = 0 \Leftrightarrow \int d^2\mathbf{x}\, \psi(\mathbf{x}) = 0. \tag{3.11}$$

Using the translations $\mathbf{b}$ ($= (b_1, b_2)$ for cartesian coordinate), rotation $\mathscr{R}_{-\theta}$ and global dilation a, the wavelet constructed from the mother wavelet $\psi(\mathbf{x})$ and its Fourier transform are

$$\psi_{a,\theta,\mathbf{b}}(\mathbf{x}) = \frac{1}{a}\, \psi(\frac{1}{a}\mathscr{R}_{-\theta}(\mathbf{x} - \mathbf{b})), \tag{3.12}$$

$$\hat{\psi}_{a,\theta,\mathbf{b}}(\mathbf{k}) = a\, e^{-i\mathbf{b}\cdot\mathbf{k}} \hat{\psi}(a\mathscr{R}_{-\theta}(\mathbf{k})), \tag{3.13}$$

respectively, where $\mathbf{b} \in \mathbf{R}$, $a \in \mathbf{R}^+$, $\mathscr{R}_{-\theta}$, where $0 \le \theta < 2\pi$, denotes the 2×2 rotation matrix and the hat denotes a 2-D Fourier transform. Note that the rotation matrix $\mathscr{R}_\theta$ acting on the points $\mathbf{x} = (x_1, x_2)$ for Cartesian coordinate can be written as

$$\mathscr{R}_\theta \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}. \tag{3.14}$$

The factor a^{-1} on the right hand side of Eq. (3.12) is for preserving the unit norm.

Using Eqs. (3.12) and (3.13), the CWT of a signal $f(\mathbf{x})$, as a linear transform from 2-D Euclidean space to the space of these four parameters, is:

$$W(a, \theta, \mathbf{b}) = \frac{1}{a} \int d^2\mathbf{x}\, \overline{\psi(\frac{1}{a}\mathscr{R}_{-\theta}(\mathbf{x} - \mathbf{b}))} f(\mathbf{x}) \tag{3.15}$$

$$= a \int d^2\mathbf{k}\, e^{i\mathbf{b}\cdot\mathbf{k}} \overline{\hat{\psi}(a\mathscr{R}_{-\theta}(\mathbf{k}))} \hat{f}(\mathbf{k}), \tag{3.16}$$

where the overbar, again, denotes complex conjugation. Note that Eq. (3.16) has been acquired from Eq. (3.15) by using the Parseval identity. We can see that the CWT on the plane is able to detect not only the information of the locations on real space and the scales of $f(\mathbf{x})$ as the CWT in one dimension, but also the directional feature. This will be discussed in the next subsection.

Finally, as mentioned above, when the wavelets satisfy the admissibility condition, Eq. (3.10), we can reconstruct the function $f(\mathbf{x})$ from its wavelet coefficients $W(a, \theta, \mathbf{b})$. Analogous to the 1-D CWT, the synthesis of a signal $f(\mathbf{x})$ from its wavelet coefficients is:

$$f(\mathbf{x}) = \frac{1}{C_\psi} \int_o^{2\pi} \mathrm{d}\theta \int_0^\infty \frac{\mathrm{d}a}{a^3} \int_{\mathbf{R}^2} \mathrm{d}^2\mathbf{b}\, \psi_{a,\theta,\mathbf{b}}(\mathbf{x}) W(a, \theta, \mathbf{b}). \tag{3.17}$$

3.3.2 Choice of the Analyzing Wavelets

An advantage of wavelet analysis is that we can choose or construct the wavelets which match with the aim of our works. For wavelet analysis on the plane, one can classify the analyzing wavelets into two types where each of which is suitable to the different purposes, namely the isotropic wavelet and the directional wavelet:

Isotropic wavelets

The isotropic wavelet is rotational invariant. Hence, its distribution has the feature of a smeared point. Using it, we perform only the space-scale decomposition as in one dimension, the directional features are not involved. In addition, it is appropriate for the detection of the pointwise feature on the image because the CWT detects well the features that match with its diffusion on the space. A typical example of isotropic wavelets is the isotropic 2-D Mexican hat wavelet or Marr wavelet:

$$\psi_H(\mathbf{x}) = (2 - |\mathbf{x}|^2)e^{-|\mathbf{x}|^2/2}, \tag{3.18}$$

which is simply the Laplacian of a Gaussian function, i.e. $\psi_H(\mathbf{x}) = -\nabla^2 e^{-|\mathbf{x}|^2/2}$. We can see from its mathematical formula that it is a rotation invariant wavelet. It has many applications in astrophysics mostly for pointwise feature detection. In particular, it is applied for the detection of the point sources in the CMB map due to astrophysical objects, as a foreground, not of cosmological origin. As a result, they can be removed in order to not disturb the cosmological considerations.

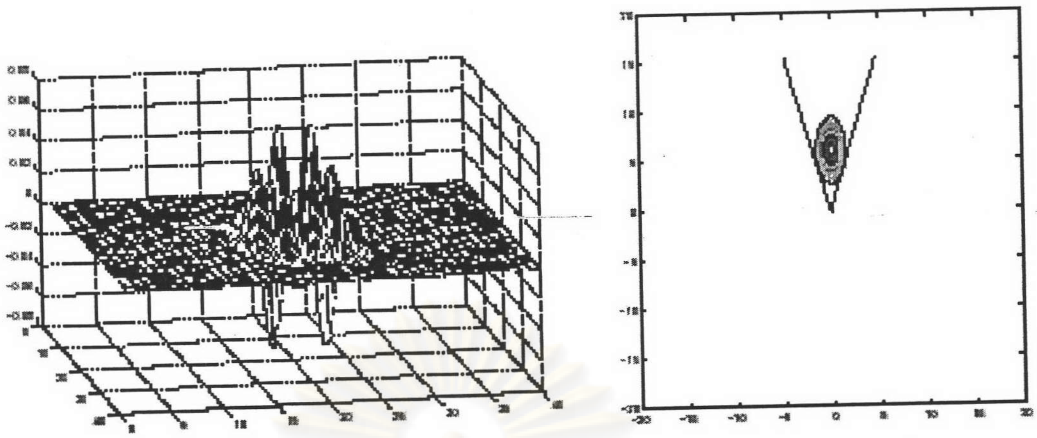

Figure 3.3: The 2-D Morlet wavelet, for $\epsilon = 2$, $\mathbf{k}_0 = (0, 6)$: (left) In position space (real part), (right) In spatial frequency space [20].

Directional wavelets

The directional wavelets are chosen when one wants to detect oriented feature, such as edges, linear discontinuities, vector field, etc. A wavelet ψ is directional when the effective support of its Fourier transform $\hat{\psi}$ is contained in an anisotropic area, which is directional selective, in scale space $\mathbf{k}$ [5]. We can see from Eq. (3.16) that the wavelet transform filters the signal in $\mathbf{k}$-space, i.e. the information of scales of the signal survives only in the region that contains the effective support of the Fourier transform of the wavelet $\hat{\psi}$. Suppose the signal $f(\mathbf{x})$ is strongly oriented like a long segment along the x-axis which fluctuates slowly in the wide region in the x-direction and fluctuates rapidly in the narrow region in the y-direction. According to the uncertainty principle, its Fourier transform $\hat{f}(\mathbf{k})$ is a long segment along the k_y-axis. More precisely, the wide (narrow) region of slow (rapid) fluctuations of $f(\mathbf{x})$ in the x-direction (y-direction) causes the narrow (wide) region of fluctuations of $\hat{f}(\mathbf{k})$ in the k_x-direction (k_y-direction). Hence, we need a wavelet ψ with the support in $\mathbf{k}$-space has the feature matched such a signal, i.e. a long segment along k_y-space, in order to detect the directional feature.

For example, the 2-D Morlet wavelet is a directional wavelet:

$$\psi_M(\mathbf{x}) = \exp(i\mathbf{k}_0 \cdot \mathbf{x}) \exp(-\frac{1}{2}|A\mathbf{x}|^2), \tag{3.19}$$

$$\hat{\psi}_M(\mathbf{k}) = \sqrt{\epsilon}\exp(-\frac{1}{2}|A^{-1}(\mathbf{k}-\mathbf{k}_0)|^2). \tag{3.20}$$

The parameter $\mathbf{k}_0$ is the wave vector, and $A = \text{diag}[\epsilon^{-1/2}, 1]$, $\epsilon \geq 1$, is a 2×2 anisotropy matrix. The wavelet ψ_M is elongated in the x direction if $\epsilon > 1$, but in the y direction if $\epsilon < 1$, and its phase is constant along the direction orthogonal to $\mathbf{k}_0$. Hence, the wavelet ψ_M can detect the sharp transitions in the direction perpendicular to $\mathbf{k}_0$. In Fourier space, the effective support of $\hat{\psi}_M$ contained in a convex cone is an ellipse centered at $\mathbf{k}_0$ and elongated in the k_y direction, that becomes narrower as ϵ increases. An example of the function $\hat{\psi}_M$ is shown in Fig. 3.3.