

Ice core theory

Thea Quistgaard

Master Thesis 2020/2021

Contents

1	Ice Theory	1
1.1	Diffusion and densification	1
1.1.1	densification	1
	Herron Langway Empirical Model	2
1.1.2	Diffusion	5
	In Firn	5
	In Solid Phase	6
	Reconstruction of temperatures	7
1.2	Ice core analysis	9
1.2.1	Water Isotopes	9
	δ notation and water isotopes	9
	Electrical conductivity measurements	10
	Dielectric Profiling	11
1.2.2	Synthetic Data	13
	Step function and response	13
1.3	Signal Analysis	13
1.3.1	Power Spectral Densities	13
	Discrete and Fast Fourier Transform	15

	Discrete Cosine Transform	17
	Maximum Entropy Method	17
1.3.2	Noise and Signal Fitting	19
1.3.3	Creating a Restoration Filter	19
1.3.4	Deconvolution	19
1.4	Measuring water isotopes	19
1.4.1	Cavity Ring-Down Spectroscopy	19
	Calibration	19
	Discrete v. Continuous	19
Bibliography		21

Chapter 1

The theory of ice cores

1.1 Diffusion and densification

Throughout the firn column the important processes of diffusion and densification takes places. Both processes need to be well understood and examined when analyzing ice core data, as diffusion and densification play a large role in thinning of annual layers due to compression of snow to ice and in washing out the measured signals through diffusion in the firn.

1.1.1 Densification

Densification is the process of compression of snow to ice. It plays an important role in the annual layer thickness in the data as snow will be compressed to a smaller volume under pressure from the firn column above until it reaches a solid ice state with a, almost, constant density.

Commonly three stages of densification are described in the firn column. The first stage is between the initial precipitated snow density and the 'critical density' at $0.55 \frac{\text{Mg}}{\text{m}^3}$, the second stage is between critical density and the close-off density at $0.82 - 0.84 \frac{\text{Mg}}{\text{m}^3}$, and the third stage is from close-off and all the way through the ice.

At the first stage the densification is mostly due to grain settling and packing and the densification rate is very rapid. At the second stage, the snow is close to isolating air bubbles. At the third stage, the dominating densification taking place is by the compression of air bubbles.

For these three stages it is of interest to develop a depth-density profile, which is dependent on snow accumulation rate and temperature. The focus is on developing an empirical model for the first and second stages of densification,

as they are the most dramatic sections of the firn column considering densification and diffusion.

A number of different densification models have been developed (REFERENCES), and in this thesis will be presented the ones used for the analysis.

Herron Langway Empirical Model

Sorge's law (REFERENCES) assumes that the relation between snow density ρ and depth h is invariant with time, given a constant snow accumulation and temperature. Furthermore, annual layer thinning by plastic flow is ignored.

Densification of firn, which can be described as the proportional change in air space, is linearly related to change in stress due to the weight of the overlying snow (REFERENCES):

$$\frac{d\rho}{\rho_i - \rho} = \text{const.} \cdot \rho \, dh \quad (1.1)$$

By integration, this implies a linear relation between $\ln \left[\frac{\rho}{\rho_i - \rho} \right]$ and h .

When considering real data, analysis shows that $\ln \left[\frac{\rho}{\rho_i - \rho} \right]$ vs h plots have two linear segments (EXAMPLE), corresponding to the first and second stages of densification, with separation of segments at $\rho = 0.55$ and $\rho = 0.8$. These segments on the plots will yield two different slopes with slope constants:

$$C = \frac{d \ln \left[\frac{\rho}{\rho_i - \rho} \right]}{dh}, \rho < 0.55 \quad (1.2a)$$

$$C' = \frac{d \ln \left[\frac{\rho}{\rho_i - \rho} \right]}{dh}, 0.55 < \rho < 0.8 \quad (1.2b)$$

To find the densification rate, $\frac{d\rho}{dt}$, substitute $\frac{dh}{dt} = \frac{A}{\rho} \rightarrow dt = \frac{\rho}{A} dh$ and use the differentiation $\frac{\partial}{\partial t} \left[\ln \left[\frac{x(t)}{k - x(t)} \right] \right] = \frac{k \frac{dx}{dt}}{(k - x(t))x(t)}$

$$\begin{aligned} C &= \frac{\rho}{A} \frac{d \ln \left[\frac{\rho}{\rho_i - \rho} \right]}{dt} \\ &= \frac{\rho}{A} \frac{\rho_i}{\rho(\rho_i - \rho)} \frac{d\rho}{dt} \\ &= \frac{1}{A} \frac{\rho_i}{\rho_i - \rho} \frac{d\rho}{dt} \end{aligned}$$

leading to

$$\frac{d\rho}{dt} = \frac{CA}{\rho_i}(\rho_i - \rho) \quad (1.3a)$$

$$\frac{d\rho}{dt} = \frac{C'A}{\rho_i}(\rho_i - \rho) \quad (1.3b)$$

To continue from here two assumptions are made. The first is that the temperature and the accumulation rate dependencies may be separated, and that they thereby have no inter-correlation. The second is that the rate equations may be written as:

$$\frac{d\rho}{dt} = k_0 A^a (\rho_i - \rho), \rho < 0.55 \quad (1.4a)$$

$$\frac{d\rho}{dt} = k_1 A^b (\rho_i - \rho), 0.55 < \rho < 0.8 \quad (1.4b)$$

where k_0 and k_1 are Arrhenius type rate constants which are only temperature dependent, and a and b are constants determining the significance of the accumulation rate and are dependent on the densification mechanisms.

a and b may be determined by comparing slopes for densification at different sites of nearly equivalent conditions as:

$$a = \frac{\ln\left(\frac{C_1}{C_2}\right)}{\ln\left(\frac{A_1}{A_2}\right)} + 1 \quad (1.5)$$

and equivalently for b , with C'_1 and C'_2 .

k_0 and k_1 can be estimated by observing values of k at different temperatures and plotting $\ln(k)$ versus temperature - a so-called Arrhenius plot (REFERENCES) - to find A and E_a in equations:

$$k = Ae^{-\frac{E_a}{k_B T}} = Ae^{-\frac{E_a}{RT}} \quad (1.6)$$

$$\ln(k) = \ln(A) - \frac{E_a}{R} \frac{1}{T}$$

leading to values of k_0 and k_1 of:

$$k_0 = 11e^{-\frac{10160}{RT}} \quad (1.7a)$$

$$k_1 = 575e^{-\frac{21400}{RT}} \quad (1.7b)$$

Depth-density and depth-age calculations

Assuming that temperature, annual accumulation rate and initial snow density are known, the following calculations can be made:

- Density at depth h , $\rho(h)$
- Depth at pore close-off, $\rho = 0.55$
- Depth-age relationship from surface to pore close-off (stage 1 and 2).

1. stage of densification: Depth-density profile:

$$\rho(h) = \frac{\rho_i Z_0}{1 + Z_0} \quad (1.8)$$

where $Z_0 = e^{\rho_i k_0 h + \ln \left[\frac{\rho_0}{\rho_i - \rho_0} \right]}$. In this segment, the depth-density is independent of accumulation rate. The critical density depth can be calculated as:

$$h_{0.55} = \frac{1}{\rho_i k_0} \left[\ln \left[\frac{0.55}{\rho_i - 0.55} \right] - \ln \left[\frac{\rho_0}{\rho_i - \rho_0} \right] \right] \quad (1.9)$$

and the age at close-off depth as:

$$t_{0.55} = \frac{1}{k_0 A} \ln \left[\frac{\rho_i - \rho_0}{\rho_i - 0.55} \right] \quad (1.10)$$

WHERE DOES THIS COME FROM? 2. stage of densificaion: The depth-density profile

$$\rho(h) = \frac{\rho_i Z_1}{1 + Z_1} \quad (1.11)$$

where $Z_1 = e^{\rho_i k_1 (h - h_{0.55}) \frac{1}{A^{0.5}} + \ln \left[\frac{0.55}{\rho_i - 0.55} \right]}$. The age of firm at a given density ρ :

$$t_\rho = \frac{1}{k_1 A^{0.5}} \ln \left[\frac{\rho_1 - 0.55}{\rho_1 - \rho} \right] \quad (1.12)$$

An estimate of the mean annual accumulation rate can be made from the slope C' and the mean annual temperature:

$$A = \left(\frac{\rho_i k_1}{C'} \right)^2 \quad (1.13)$$

1.1.2 Diffusion

In Firn

Diffusion describes the attenuation of a given signal, e.g. a water isotopic signal, due to vapor phase diffusion in the porous firn column. To develop accurate knowledge of paleo climate and temperatures it is of great importance to understand this process, as a reconstruction of the part of the signal lost will reveal finer details in the signal and thus a more detailed knowledge of past times. Diffusion can be described through Fick's 2nd law, which describes the change in concentration of a substance with time, due to diffusion:

$$\frac{\partial \phi}{\partial t} = D(t) \frac{\partial^2 \phi}{\partial z^2} - \dot{\epsilon}_z(t) z \frac{\partial \phi}{\partial z} \quad (1.14)$$

If we say the diffusion is focused on water isotopes, then we can approximate the water isotopic signal with the concentration, $\phi \approx \delta$, so:

$$\frac{\partial \delta}{\partial t} = D(t) \frac{\partial^2 \delta}{\partial z^2} - \dot{\epsilon}_z(t) z \frac{\partial \delta}{\partial z} \quad (1.15)$$

Through attenuation with depth and time due to diffusion there is a loss of information. But the diffusion constant and the vertical strain rate $\dot{\epsilon}_z(t)$ in Fick's 2nd law are dependent on temperature and accumulation on site, this information loss process can be used to infer temperature of firn and accumulation on site. The solution of Eq. 1.15 can be found by deconvolution. The attenuated, directly measured, isotopic signal, $\delta(z)$, can be described as the convolution between the initial isotopic signal, $\delta'(z)$, and a Gaussian filter, $\mathcal{G}(z)$, multiplied by the thinning function, $S(z)$, which describes the total thinning of a given layer at depth z due to the vertical strain from the above firn column.:

$$\delta(z) = S(z) [\delta'(z) * \mathcal{G}(z)] \quad (1.16)$$

where

$$S(z) = e^{\int_0^z \dot{\epsilon}_z(z') dz'} \quad (1.17)$$

and

$$\mathcal{G}(z) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{z^2}{2\sigma^2}} \quad (1.18)$$

In the gaussian filter, the variance σ^2 is referred to commonly as the diffusion length: the distance a water molecule is displaced along the z-axis. This quantity is directly related to both $D(t)$ and $\dot{\epsilon}_z(t)$ (the strain rate being approximately proportional to the densification rate in the column). Thus an

accurate estimate of the diffusion length is crucial for describing the diffusion process. The change of diffusion length over time is given as

$$\frac{d\sigma^2}{dt} - 2\dot{\epsilon}_z(t)\sigma^2 = 2D(t) \quad (1.19)$$

given by JOHNSEN1977, which also states that for the strain rate, the following approximation can be made:

$$\dot{\epsilon}_z(t) \approx -\frac{d\rho}{dt} \frac{1}{\rho} \quad (1.20)$$

where ρ is the density and $\frac{d\rho}{dt}$ is the densification rate. With this approximation, the solution to the equation for evolution of the diffusion length in the firn column can be found, defined only through density and densification rate, as(DESCRIBE HOW TO SOLVE FOR SIGMA):

$$\sigma^2(\rho) = \frac{1}{\rho^2} \int_{\rho_0}^{\rho} 2\rho'^2 \left(\frac{d\rho'}{dt} \right)^{-1} D(\rho') d\rho' \quad (1.21)$$

Certain densities and corresponding depths are of special interest as they indicate a specific stage of the firn and ice column. At top and bottom, we find the two extremum densities of settled snow, $\rho_{\text{snow}} = 330 \frac{\text{kg}}{\text{m}^3}$, and ice, $\rho_{\text{ice}} = 917 \frac{\text{kg}}{\text{m}^3}$. In between these two there are two more densities of importance: the critical density, $\rho_{\text{Cr}} = 550 \frac{\text{kg}}{\text{m}^3}$, describing the transition between the two firn stages (see Section 1.1.1), and the pore close off density, $\rho_{\text{CO}} = 330 \frac{\text{kg}}{\text{m}^3}$, describing the density at which air pockets in firn will seal off from each other to form single bubbles. From the close off density, further densification will be due to compression of these closed off air bubbles until the density reaches ρ_{ice} . If we assume that the diffusion constant, $D(\rho)$, and the densification rate, $\frac{d\rho}{dt}$ are known, then it is possible to give an estimate of the diffusion length profile by integrating from top, at density ρ_0 , to pore close-off depth, ρ_{co} .

In Solid Phase

When firn reaches solid state, below close-off depth, the isotope diffusion is driven not as much by densification any more, but by isotopic gradients within the ice crystal lattice structure. This diffusion process is much slower than the diffusion in vapor phase taking place in firn, and thus does not contribute as much to the information loss and attenuation of the signal. For solid ice, at $\rho \leq \rho_{\text{ice}}$, the diffusion constant is only dependent on temperature, and can

be described through an Arrhenius type equation as(ref: RAMSEIER1967, JOHNSENetal2000):

$$D_{ice} = 9.2 \cdot 10^{-4} e^{-\frac{7186}{T}} \left[\frac{\text{m}^2}{\text{s}} \right] \quad (1.22)$$

The diffusion length in ice ice given from the diffusion constant in ice and the thinning function as:

$$\sigma_{ice}^2(t) = S(t)^2 \int_0^t 2D_{ice}(t') S(t')^{-2} dt' \quad (1.23)$$

(DESCRIBE HOW TO SOLVE FOR SIGMA and a discussion of the ice diffusion constant.)

Reconstruction of temperatures

Reconstruction of paleotemperatures can be attempted through a number of various techniques (REFERENCES). For precise and accurate results, the single isotopologue diffusion methods have proven useful(REFERENCES).

As is known, convolution in time domain is equal to multiplication in the frequency domain. According to equation (1.16), the transfer function to the frequency domain, will be the Fourier transform of the Gaussian filter:

$$\mathcal{F}[\mathcal{G}(z)] = \hat{\mathcal{G}} = e^{-\frac{k^2 \sigma^2}{2}}, \quad k = 2\pi f = \frac{2\pi}{\Delta} \quad (1.24)$$

where Δ is the discrete sampling size. This filter keeps larger wavelength frequencies (≥ 50 cm) unaltered but attenuates short wavelengths (≤ 20 cm) heavily, which is exactly the effect of diffusion on the isotopic signal. An estimate of the diffusion length σ^2 can be made from the power spectral density(PSD) of an isotopic time series. In the frequency domain a PSD composed of an initial signal, a filter function and a noise term is given by:

$$P_s = P_0(k) e^{-k^2 \sigma^2} + |\hat{\eta}(k)|^2, \quad f \in [0, f_{Nq}] \quad (1.25)$$

where the diffused and noise-affected signal, P_s , is equal to the original signal, $P_0(k)$, times a filter, $e^{-k^2 \sigma^2}$ (our previously inspected Gaussian filter), plus a noise term, $|\hat{\eta}(k)|^2$, over a frequency space ranging from zero to the Nyquist frequency, f_{Nq} . The Nyquist frequency is dependent on the sampling resolution by $f_{Nq} = \frac{1}{2\Delta}$. The noise term, often categorized as white noise, but red noise is also seen in isotopic signals(REFERENCES), is given as

$$|\hat{\eta}(k)|^2 = \frac{\sigma_n^2 \Delta}{|1 - a_1 e^{ik\Delta}|^2} \quad (1.26)$$

Equation 1.26 describes an autoregressive process of the first order, with a_1 being an AR-1 coefficient. **WHAT IS THIS? DESCRIBE..**

The spectral estimate of the time series, \mathbb{P}_s , can be computed via a number of different numerical schemes, here Burg's method will be used, REFERENCES. To determine the diffusion length a fit to these estimated spectral data, P_s , is found through for example a least square optimization, from which the parameters P_0 , σ , a_1 , σ_η^2 can be estimated.

The diffusion length σ^2 can be calculated by least-square minimization of the misfit between \mathbb{P}_s and P_s .

This estimated diffusion length needs to be corrected: the obtained $\hat{\sigma}^2$ is affected by two further diffusion processes, taking place respectively in the ice and in the experimental sampling:

- **Sampling diffusion:** This diffusion is due to the sampling method. Sampling at a certain discrete resolution - be it discrete sections or resolution in CFA system due to step or impulse response - gives an additional diffusion length of

$$\sigma_{dis} = \frac{2\Delta^2}{\pi^2} \ln\left(\frac{\pi}{2}\right) \quad (1.27)$$

- **Ice diffusion** When below the close-off depth, a correction for the ice diffusion must also be made.

So to obtain the actual diffusion length from the raw data, both the sampling and the ice diffusion need to be subtracted from σ^2 , and a scaling factor due to thinning from the strain must be introduced:

$$\sigma_{firn}^2 = \frac{1}{S(z)^2} \hat{\sigma}_{firn}^2 = \frac{\hat{\sigma}^2 - \sigma_{dis}^2 - \sigma_{ice}^2}{S(z)^2} \quad (1.28)$$

Now, from the obtained estimate of the firn diffusion length, a temperature estimate can be made by numerically finding the root of:

$$\left(\frac{\rho_{co}}{\rho_i}\right)^2 \sigma^2(\rho = \rho_{co}, T(z), A(z)) - \sigma_{firn}^2 = 0 \quad (1.29)$$

NOTE: Annual spectral signals appearing as peaks in the PSD, can influence the estimate of diffusion lengths. This can be taken into account by introducing a weight function omitting the annual signal from the PSD:

$$w(f) = \begin{cases} 0, & f_\lambda - df_\lambda \leq f \leq f_\lambda + df_\lambda \\ 1, & f < f_\lambda - df_\lambda, f > f_\lambda + df_\lambda \end{cases} \quad (1.30)$$

1.2 Ice core analysis

Electrical conductivity measurements(ECM), dielectric profiling(DEP) and isotopic composition analysis are three distinct ways of analyzing an ice core to examine past temperatures, climate and atmospheric composition. Some of these methods are sensitive to violent volcanic eruptions, which makes it possible to use known eruptions visible in the ice cores as volcanic horizons, and thus making dating of the ice core more precise and absolute.

1.2.1 Water Isotopes

A corner stone in ice core analysis, which helps lay the basis for paleo climate research, is through measurements of the isotopic composition of the water - or that of the encapsulated air in bubbles - which makes up the ice cores. Water isotopes are sensitive to temperature changes and can thus be used as a proxy for paleo temperature along with being used as dating parameters, since the annual cycles often are detectable in water isotope data.

δ notation and water isotopes

Water isotopic ratios, i.e. the ratio of the minority isotope, H_2^{18}O or H_2^{17}O ($^2\text{H}_2\text{O}$), compared to the majority isotope, H_2^{16}O ($^1\text{H}_2\text{O}$), are used to report the quantities of isotopes in a sample relative to the ratio of a given reference water sample. This is commonly expressed in the δ -notation as:

$$\delta^i = \frac{{}^iR_{\text{sample}}}{{}^iR_{\text{reference}}} - 1 \quad (1.31)$$

where $^{18}R = \frac{n_{^{18}\text{O}}}{n_{^{16}\text{O}}}$, $^2R = \frac{n_{^2\text{H}}}{n_{^1\text{H}}}$ and $^{17}R = \frac{n_{^{17}\text{O}}}{n_{^{16}\text{O}}}$. Here n is the abundance of the given isotope.

Besides the isotopic quantities $\delta^{17}\text{O}$, $\delta^{18}\text{O}$ and $\delta^2\text{H} = \delta\text{D}$, both deuterium excess and $\Delta^{17}\text{O}$, known as ^{17}O excess, can be of interest. Deuterium excess is usually used as a measure of the kinetic fractionation processes, taking place in the water vapor formation of polar precipitation, giving an indicator of the conditions during precipitation formation, and thus giving a pointer to the source of the water vapor. Like deuterium excess ^{17}O is sensitive to kinetic fractionation, but much less sensitive to equilibrium fractionation than both δD and $\delta^{18}\text{O}$. Along with being nearly insensitive to temperature(REFERENCES), these robustness factors leads to ^{17}O being usable as an independent parameter to be used to reveal the ways of the complicated mixing effects of fractionation due to evaporation, transportation, formation

and deposition. **Fractionation**

Electrical conductivity measurements

The conductivity of ice arises from the current emerging due to the build-up of space charges in the ice structure. This conductivity can be analyzed by measuring the electrical current(DC) - induced by the electric potential and the acid balance - between two electrodes which are moved along the ice cores length. This current will be connected to the acid impurity concentration (pH), in the form of H_3O^+ concentration, of the ice core. Higher levels of acid impurity concentration are due to volcanic eruptions. Large amounts of volcanic gases, i.e. SO_2 , in the atmosphere oxidizes and combines with water to form acid, i.e. sulphuric acid, which is the washed out of the air due to precipitation. Thus it is made possible to recognize volcanic horizons in ice cores, and - if the location of the eruption is known - from the amount of acid, the magnitude of the eruption can also be estimated.

High acidity of layers containing volcanic fall-out influence the dielectric constant of ice, so that these layers may be a possible explanation to the internal reflection horizons found in radio-echo sounding.

The measured current can then be transformed into acidity by a calibration curve relating the current, in μA , to the acidity, in $\mu\text{equivalents } \text{H}_3\text{O}^+$ per kilogram. To find the calibration parameters, the current and the acidity must be measured - the current through the above mentioned method, and the acidity through pH measurements of melted ice core samples. The pH measurements must further be corrected for any CO_2 induced H^+ ions (REFERENCES). The relation between acidity $[\text{H}^+]$ (corrected for CO_2 induced H^+) and current I can be expressed in two ways:

- $[\text{H}^+] = (0.017 I^2 + 1.2) \mu\text{equiv. } \text{H}^+/\text{kg}$
without a 50% correction for CO_2 surplus.
- $[\text{H}^+] = (0.045 I^{1.73}) \mu\text{equiv. } \text{H}^+/\text{kg}$
with a 50% correction for CO_2 surplus.

The salt concentration in the ice can be estimated from measurements of the specific conductivity σ of the melted samples. The salt contribution hereto can be expressed as:

$$\sigma_s = \sigma - \sigma(\text{H}^+) - \sigma(\text{X}^-) - \sigma(\text{HCO}_3^-) \quad (1.32)$$

where the three later terms correspond to the contributions from H^+ (through pH measurements) and its anions¹, HCO_3^- and any other anions X^- . The anion concentration will be equal to the cation concentration, which in this case is only H^+ concentration. Disregarding low acidity samples, the concentration of HCO_3^- is negligible and thus $concentration(X^-) \approx concentration(H^+)$. The current is thus heavily influenced on/determined by the H^+ concentration, and thus it is approximated that the salt concentration has no influence on the current readings, which is fortunate, since the ECM method only responds to acidity, and not to salt and ammonia concentrations. This is one of the methods limitations, which the later dielectric profiling method took into account.

Dielectric Profiling

A method was later developed to demonstrate how both acids and salts play a decisive role in the determination of the electrical behavior of ice. The dielectric response of an ice core can be used to determine the total ionic concentration of the core. For ECM the measurements are sensitive to the fluctuating distance between ice core and electrodes, and after each measurement a fresh piece of ice needs to be prepared to repeat a measurement.

A new dielectric profiling technique (DEP) was developed (REFERENCES) with the advantages over the ECM that no direct contact is needed between the electrodes and the ice, so that the ice can stay in a protective polythene sleeve and the experiment easily can be repeated on the same piece of ice. Together the ice core and the polythene sleeve creates a complete system, where the plastic acts as an electrical blocking layer.

The dielectric response is measured by a sweeping of the AF-LF frequency range for the entire ice-polythene system. At LF the conductivity of the composite system is within a few percentages of the intrinsic behavior of the ice itself. At HF-VHF frequencies it also approximates well enough (REFERENCES).

The measured dielectric parameters are the conductivity of ice at HF-VHF range, denoted σ_∞ where ∞ signifies a frequency much higher than the relaxation frequency, f_r , of the dominant dispersion in the system. Both of these parameters display clear chemical response signals which can be used either alone or in combination with other ice core analysis measurements like ECM and isotope analysis.

If the core under analysis is chemically analyzed for Na^+ , Mg^{2+} , Cl^- , SO_4^{2-}

²Anions are molecules losing a number of electrons to become negatively charged. Cations are molecules that gain a number of electrons to become positively charged.

and NO_3^- , a number of important parameters, which can be used to evaluate the response of the dielectric parameters, can be calculated (REFERENCES):

- The salt parameter, which represents the total marine cation concentration calculated with the assumed marine ratios as:

$$[\text{salt}] = 1.05([\text{Na}^+] + [\text{Mg}^{2+}]) \quad (1.33)$$

- XSO_4 , the excess sulphate, which represents the amount the sulphate concentration is above the expected if the salt and sulphate ions were in normal sea salt ratios. Excess sulphate is essentially sulphuric acid, which is the main acidic component of the ice.
- The strong acid content of the ice has been calculated as (assuming no other ions present in significant quantities):

$$[\text{acid}] = [\text{Cl}^-] + [\text{SO}_4^{2-}] + [\text{NO}_3^-] - 1.05([\text{Na}^+] + [\text{Mg}^{2+}]) \quad (1.34)$$

From data, it can be seen that acid and salt concentration peaks clearly affect σ_∞ and f_r (EXAMPLES, REFERENCES). The relationship between salt and acid, and the two dielectric parameters have been derived through non-linear regression analysis. In PAPER (REFERENCES) the linear responses for the DEP at -22°C were:

$$\sigma_\infty = (0.39 \pm 0.01)[\text{salt}] + (1.43 \pm 0.05)[\text{acid}] + (12.7 \pm 0.3) \quad (1.35)$$

with 76.6 % variance

$$f_r = (440 \pm 11)[\text{salt}] + (612 \pm 65)[\text{acid}] + (8200 \pm 400) \quad (1.36)$$

with 68.4 % variance. σ_∞ is measured in $\mu\text{S}/\text{m}$, f_r in Hz and $[\text{acid}]$ and $[\text{salt}]$ in $\mu\text{Eq}/\text{l}$. The total ionic concentration of the ice core is strongly linked to the dielectric parameters, and a regression between the total anion concentration and the dielectric parameters gives:

$$[\text{anions}] = [\text{salt}] + [\text{acid}] = 0.022\sigma_\infty^{1.89} + 10^{-6}f_r^{1.61} - 0.2 \quad (1.37)$$

with 86.7 % variance.

The DEP complements the ECM technique by not only reacting to acids alone, as ECM does, but responds to both neutral salts and acids. The acid term is here associated with the DC conductivity, the same way it is also detected by ECM. The dielectric dependence on salts is consistent with the

Bjerrum L defect² affecting every one or two salt ions in the ice, indicating that a large fraction of the neutral salt is incorporated into the ice lattice.³ The sensitivity to salt concentrations allows for identifications of periods with major storms and open seas which are also important identifiers for paleo climate research, along with the volcanic eruption detection made possible through the ECM.

1.2.2 Synthetic Data

Step function and response

1.3 Signal Analysis

The data obtained through various experimental measurements are easily compared with a time series, as they typically show some quantity measured all along the depth of an ice core. This depth is often, at short intervals, treated as a regular linear time series thus making it possible to use some of the known signal analysis methods. Of course, when considering the entirety of an ice core, the linearity disappears as thinning and compression makes the depth series non linear. But when considering short lengths of core it is possible to estimate a linearity, assuming conformity in this specific layer.

1.3.1 Power Spectral Densities

A very useful tool for analyzing signals exhibiting oscillatory effects is analysis of the signals power spectrum. Instead of considering the signal in time, it is transformed to the spectral domain, where it is possible to obtain an estimate of both the signal and the underlying noise. This is crucial for enhancing the signal and filtering away noise. But to be able to examine these effects, first the data must be transformed. A range of different methods may be used to compute the frequency transform of the depth series, here I present the three I have been working with. Since the data are discrete and experimental, I will be presenting the discrete and applicable mathematical models.

When considering a signal, it may be of interest to investigate how the energy of said signal is distributed with frequency. The total power is defined as:

$$\text{Total Power} = \int_{-\infty}^{\infty} |X(\tau)|^2 d\tau. \quad (1.38)$$

³A Bjerrum defect is a crystallographic defect specific to ice, partly responsible for the electrical properties of ice. Usually a hydrogen bond will normally have one proton, but with a Bjerrum defect it will have either two protons (D defect) or no proton (L defect). (REFERENCES)

Using Parseval's theorem (REFERENCE) (assuming that the signal has a finite total energy), the power of the signal can alternatively be written as

$$\int_{-\infty}^{\infty} |X(\tau)|^2 d\tau = \int_{-\infty}^{\infty} |\tilde{X}(f)|^2 df \quad (1.39)$$

where $\tilde{X}(f)$ is the spectral (Fourier) transform of the signal, from time to frequency domain, defined as:

$$\tilde{X}(f) = \int_{-\infty}^{\infty} X(\tau) e^{2\pi i f \tau} d\tau \quad (1.40)$$

and the inverse spectral (Fourier) transform, from frequency to time domain, defined as:

$$X(t) = \int_{-\infty}^{\infty} \tilde{X}(f) e^{-2\pi i f t} df. \quad (1.41)$$

Both $X(t)$ and $\tilde{X}(f)$ represent the same function, just in different variable domains. Often, the angular frequency ω is used instead, with the relation between ω and f being $\omega \equiv 2\pi f$, giving the Fourier and inverse Fourier transforms as:

$$\begin{aligned} \tilde{X}(\omega) &= \int_{-\infty}^{\infty} X(t) e^{i\omega t} dt \\ X(\tau) &= \int_{-\infty}^{\infty} \tilde{X}(\omega) e^{-i\omega \tau} d\omega \end{aligned} \quad (1.42)$$

From Equation 1.39 we can interpret the integrand on the right hand side $|\tilde{X}(f)|^2$ as a density function, describing the energy per unit frequency. This is a property which is able to reveal much information about the considered signal, and it is useful to define this as the (one-sided) Power Spectral Density:

$$P_X(f) \equiv |\tilde{X}(f)|^2 + |\tilde{X}(-f)|^2 \quad 0 \leq f < \infty \quad (1.43)$$

This entity ensures that the total power is found just by integrating over $P_X(f)$ from 0 to ∞ . When the function is purely real, the PSD reduces to $P_X(f) = 2|\tilde{X}(f)|^2$.

In the above the transform used to define the PSD was presented as the Fourier transform. When working with discrete data, as is very common when analyzing real world data, there are a number of different ways of estimating the PSD. In the following three different methods will be presented, all used in this thesis.

RETHINK THIS PART. DO NOT USE TIME ON ALL THE CALCULATIONS. WRITE THE GENERAL IDEAS OF THE METHODS AND STATE HOW TO CALCULATE/COMPUTE. SMALL CODE SNIP TO GIVE GENERAL IDEA.

Discrete and Fast Fourier Transform

The definition of the continuous Fourier transform and its inverse was presented in the above. The Fourier transform is as seen a way of representing the function under consideration as an infinite sum of periodic components. When the function is discrete, so will the Fourier transform be, and the integral is replaced with a sum. This gives us the Discrete Fourier Transform (DFT) which transforms the signal into a sum of separate components contributing at different frequencies. The DFT is dependent on the sampling interval, Δ , and we can describe our discrete signal X as a function of N discrete time steps $t_k = k \cdot \Delta$, where $k = 0, 1, \dots, N - 1$:

$$X_k \equiv X(t_k) \quad (1.44)$$

This sample size is supposed to be representative for the entire discrete function, if the function continues beyond the N sampled points. When sampling discretely at interval Δ , there will be a special frequency, the Nyquist critical frequency, defined through the sampling size as:

$$f_{NQ} \equiv \frac{1}{2\Delta}. \quad (1.45)$$

This frequency is of great importance in transformation of discrete signals. If the continuous signal is sampled at an interval Δ is bandwidth limited to frequencies smaller in magnitude than f_{NQ} , $\tilde{X}(f) = 0$ for $|f| \geq f_{NQ}$ - i.e. the transformed function has only non-zero values inside the Nyquist interval, $\tilde{X}(-f_{NQ}), \dots, \tilde{X}(f), \dots, \tilde{X}(f_{NQ})$. This means that the function is completely determined since we have all information about the signal contained in our available frequency space.

On the other hand, which is much more likely, if the continuous signal consists of frequencies both inside and outside the Nyquist interval, then all spectral information outside of this range will be falsely interpreted as being inside this range. Thus a wave inside the interval with a frequency of f_n will have a number of wave siblings outside of the interval, with frequencies of $k \cdot \frac{1}{\Delta} f_n$, k being integers, which will be aliased into the Nyquist interval and give rise to an increased power at the frequency f_n .

When analyzing an already measured discrete signal, this might give rise to some headache. What can be done is to assume that the signal has been sampled competently and then assume that the Fourier transform is zero outside of the Nyquist interval. After the analysis it will then be possible to determine if the signal was indeed competently sampled, as the Fourier series will go to zero at f_{NQ} given a correct assumption, and go to a fixed value, if the sampling was not done competently.

Now with the basics of understanding the limits of frequency transform of a discretely sampled signal, it is possible to estimate the DFT of the signal $X_k \equiv X(t_k)$. Since the Fourier transform is a symmetric transformation it is easiest to assume that N is even.

Since the input information is of size N we should expect only to sample the frequency transform $\tilde{X}(f)$ at only discrete values of f in the range between the upper and lower critical Nyquist frequencies, $-f_{NQ}$ to f_{NQ} :

$$f_n \equiv \frac{n}{N\Delta}, \quad n = -\frac{N}{2}, \dots, \frac{N}{2} \quad (1.46)$$

This will indeed actually give rise to $N + 1$ values, since 0 will be in the interval as well, but the limit frequencies are actually not independent, but all frequencies between are, which reduces it to N samples.

Now the integral from Equation 1.40 needs to be estimated as a sum:

$$\tilde{X}(f_n) = \int_{-\infty}^{\infty} X(\tau) e^{2\pi i f_n \tau} d\tau \approx \sum_{k=0}^{N-1} X_k e^{2\pi i f_n t_k} \Delta = \Delta \sum_{k=0}^{N-1} X_k e^{2\pi i k \frac{n}{N}} \quad (1.47)$$

The Discrete Fourier Transform is thus defined as:

$$\tilde{X}_n \equiv \sum_{k=0}^{N-1} X_k e^{2\pi i k \frac{n}{N}} \quad (1.48)$$

This gives the approximate relation between the DFT estimate and the continuous Fourier transform $\tilde{X}(f)$ when sampling at size Δ as:

$$\tilde{X}(f_n) \approx \Delta \tilde{X}_n \quad (1.49)$$

The inverse DFT is given as:

$$X_n \equiv \frac{1}{N} \sum_{n=0}^{N-1} X \tilde{X}_n e^{-2\pi i k \frac{n}{N}} \quad (1.50)$$

Computation of the DFT can be very slow and tiresome, since it involves complex multiplication between a number of vectors and matrices. If we write

Equation 1.48 as $\tilde{X}_n = \sum_{k=0}^{N-1} W^{nk} X_k$, where W is a complex number $W \equiv e^{2\pi i/N}$. This shows that the vector X_k must be multiplied with a complex matrix which (n,k)th component consists of the constant W to the power of nk . This matrix multiplication evidently leads to a process of $O(N^2)$. Fortunately, a number of different algorithms(REFERENCES) have been developed for fast and efficient computation of the discrete Fourier transform. One of these is called the Fast Fourier Transform (FFT), which can reduce the computations to just $O(N \log_2 N)$! In this thesis the FFT used is the one implemented in the numpy.fft Python package(REFERENCES) which is based on the works of (REFERENCES). See said article for implementation details. One important thing about this specific algorithm is that for the algorithm to function most efficiently, the number of points computed in the frequency space must be of a power of 2, following the use of base \log_2

Discrete Cosine Transform

The DFT is generally defined for complex inputs and outputs, where the sine components of the transform describe the complex part and the cosine describes the real part.

Since the data analyzed in this thesis is purely real, it makes sense, for computational speed, to only work with the cosine parts of the transform. This leads to the Discrete Cosine Transform (DCT) which transforms real inputs to real outputs.

Maximum Entropy Method

The following is based on the work of Tad J. Ulrych, [1].

An estimate of the PSd of a stationary and uniformly sampled process can be given by the spectrum computed through the Maximum Entropy Method, from here on: MEM. This method is build on one of the most fundamental concepts of physics, namely that of the second law of thermodynamics: *The entropy of an isolated system will never decrease* (REFERENCE). The MEM searches to find the spectrum which maximizes the given process' entropy, which is physically intuitively satisfying. To be able to develop a method to compute the Maximum Entropy PSD of a process, it is necessary to visit information and entropy theory.

Considering a set of $M = m_1, m_2, \dots, m_M$ events possible, each with a probability p_i , the total information available in the system about a single event is

related to that events probability:

$$I = k \cdot \ln\left(\frac{1}{p_i}\right) \quad (1.51)$$

with k , a constant, equal to 1 when using base \log_2 . When observing the system over a long period of time T , we expect $p_1 \cdot T$ of m_1 events happening, $p_2 \cdot T$ of m_2 events happening and so on, adding up to the total information of:

$$I_{tot} = k \cdot \sum_{i=1}^M p_i T \ln\left(\frac{1}{p_i}\right) \quad (1.52)$$

From the total information, the entropy can be defined as the average information per time:

$$H = \frac{I_{tot}}{T} = k \cdot \sum_{i=1}^M \left(p_i \ln\left(\frac{1}{p_i}\right)\right) = -k \cdot \sum_{i=1}^M (p_i \ln(p_i)) \quad (1.53)$$

A more intuitive way of understanding entropy is to view it as a measure of uncertainty or variance in the ensemble described by the individual probabilities. For example, if all p_i except for one is zero, the system is completely deterministic as it is only one event that is allowed to happen - thus the variance, and the entropy, will be zero. In all other cases, the entropy will be positive, $H > 0$.

Now, to go from information and entropy theory to estimation of PSDs consider Jaynes' Principle of maximum entropy:

The prior probability assignment that describes the available information but is maximally noncommittal with regard to the unavailable information is the one with maximum entropy.

*Out of all possible (hypotheses—PDFs—...) that agree with constraints, choose the one that is **maximally non-committal** with respect to missing information*

This principle describes the relationship between entropy and uncertainty or ignorance - that is, the knowledge that is outside of the given parameter space defined by the signal. From this it is possible to develop an approach which is consistent with the prior knowledge(the observed signal), and which estimates the prior probability assignment describing the prior information, without assuming anything outside of the sampled space. This means that a maximum entropy method will not bind any constraints to the unknown

knowledge (other methods assume for example cyclical or zero valued data outside the parameter space, which can give rise to errors in the spectral estimate, especially around boundaries).

The maximum entropy prior probability distribution estimation can be made through the following:

Consider a process X_t which can take on the different values X_1, X_2, \dots, X_n . The information available concerning this process is assumed in the form of average values of a number of different functions, $\langle f_1(X_t) \rangle, \langle f_2(X_t) \rangle, \dots, \langle f_m(X_t) \rangle$. The number of available functions is smaller than the number of representations of the process, $m < n$.

1.3.2 Noise and Signal Fitting

1.3.3 Creating a Restoration Filter

1.3.4 Deconvolution

1.4 Measuring water isotopes

1.4.1 Cavity Ring-Down Spectroscopy

Calibration

Discrete v. Continuous

Bibliography

- [1] Tad J Ulrych and Thomas N Bishop. Maximum entropy spectral analysis and autoregressive decomposition. *Reviews of Geophysics*, 13(1):183–200, 1975.