Geochemical Implications of Stirring and Mixing in the Earth's Mantle

John Frederick Rudge

Trinity College

September 2006

This dissertation is submitted for the degree of Doctor of Philosophy.

Declaration

This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration except where specifically indicated in the text. It does not exceed the page limit set out by the Degree Committee for Earth Sciences and Geography, and is not substantially the same as any work that has been, or is being, submitted to any other university for any degree, diploma, or any other qualification.

John Frederick Rudge
22nd September 2006

Acknowledgements

I wholeheartedly thank both my supervisors, Dan McKenzie and Peter Haynes, for coming together on this project and guiding me through it. As a mathematician by training, entering the world of geochemistry was at first quite daunting, but it has turned out to be a most enjoyable and rewarding experience. On this journey Dan has been a inexhaustible source of enthusiasm, encouragement, advice and anecdotes. I thank Dan for teaching me the power of observations, and Peter for trying at the same time to make a honest mathematician of me.

This interdisciplinary PhD has involved the important transfer of knowledge between Earth Sciences and Applied Mathematics, mostly in the form of papers in my rucksack as I have cycled back and forth between the Bullard Labs and DAMTP! At Bullard, I thank everyone for making it a very friendly place to work, with special mention to Gill, Barbara, Angie and Joyce. I thank my seismologist office mates Brian, Matt and Richard for their entertaining company, and treasure those rare moments when we were all in the office at the same time. I particularly thank Alistair and Brian for proof-reading some of this thesis, and Dave Lyness for his coffee time history lessons.

At DAMTP, I thank my atmospheric dynamicist office mates Steve and Sue for their pleasant company, Paul Metcalfe and Chris Heaton for their mathematical advice, and John Hinch for first introducing me to research in fluid dynamics. I particularly thank John Lister for his insights into two phase flow, and thank Stephen Cowley and Kalvis Jansons for my tea time education in politics, computing and lock picking.

I thank all the geologists, geophysicists, geochemists, and fluid dynamicists who have patiently tried to explain things to me over the last three years. I thank Karl Grönvold, Richard Walker, James Jackson, Nick McCave and Albert Galy for introducing me to field geology. I have benefited greatly from the geochemical advice of Andreas Stracke, Al Hofmann, Steve Galer, Janne Blichert-Toft, Nobu Shimizu, and Ken Sims. I particularly thank Francis Albarède for many useful discussions, and his patience with me! I also thank all the staff and fellows at the WHOI Geophysical Fluid Dynamics program 2005 for a most enjoyable summer break away from the PhD.

Summary

Measurements of radiogenic isotopes can in principle constrain the melting, melt migration, and solid state convection that occurs in the Earth's mantle, but to do so requires suitable quantitative models. A new statistical model is introduced to better understand the observed heterogeneity in isotopic ratios $^{143}\text{Nd}/^{144}\text{Nd}$, $^{87}\text{Sr}/^{86}\text{Sr}$, $^{176}\text{Hf}/^{177}\text{Hf}$, $^{208}\text{Pb}/^{204}\text{Pb}$, $^{206}\text{Pb}/^{204}\text{Pb}$ and $^{207}\text{Pb}/^{204}\text{Pb}$ measured on mid-ocean ridge basalt. The model is highly idealised, analytically tractable, and contains the essential physical processes involved: radioactive decay, the stirring and recycling of mantle convection, partial melting, and the mixing of melts. Comparison of the modelled heterogeneity with that observed constrains model parameters, which in turn constrains aspects of mantle convection and melting.

The model provides a new interpretation of the 2.0 Ga lead-lead pseudo-isochron age in terms of an age distribution of mantle material. Simple equations relate the pseudo-isochron age to the rate of melting and decay constants. These equations are different from, but related to and more general than, those found previously for standard geochemical box models. The results are in good agreement with numerical simulations of mantle convection. The 2.0 Ga pseudo-isochron age is shown to infer a 0.5 Ga average time scale for melting of mantle material.

Geochemical and geological evidence suggests that melt travels to the surface via a network of channels under the ridge. Motivated by this, the fluid dynamical problem of a open melt conduit surrounded by a deformable porous medium is studied. Previous work has shown that the conduit supports solitary waves of elevation, with a region of trapped melt travelling with the wave. The new analysis comes to a different conclusion, showing that the solitary wave is instead one of depression, without a region of trapped melt.

Publications arising from this thesis

Chapters 2 and 3 are based on work published in:

J. F. Rudge, D. McKenzie, and P. H. Haynes. A theoretical approach to understanding the isotopic heterogeneity of mid-ocean ridge basalt. *Geochim. Cosmochim. Acta*, 69:3873-3887, 2005. doi:10.1016/j.gca.2005.03.004.

Chapters 4 and 5 are based on work published in:

J. F. Rudge. Mantle pseudo-isochrons revisited. *Earth Planet. Sci. Lett.*, 249:494-513, 2006. doi:10.1016/j.epsl.2006.06.046.

Chapters 6 and 7 are unpublished work.

Contents

Chapter 1

Introduction

We can not observe the Earth's mantle directly: all inferences must come from measurements made at the surface. Geophysical observations, notably from seismology, tell us a great deal about the structure, and a little about the composition, of the Earth, but can tell us almost nothing of its history. Geochemical observations, particularly of radiogenic isotopes, fill this gap and can tell us a great deal about the time scales for processes occurring within the Earth. Extracting the geochemical information requires quantitative models, and the role of this thesis is to develop such models.

This chapter briefly reviews some of the basic geochemical arguments and observations underpinning the thesis. For a more detailed exposition of the topics raised here the reader is referred to the textbooks [Albarède, 2003, Dickin, 2005, Faure, 1986, White, 2005] and recent review articles [Hofmann, 2003, Kelemen et al., 1997, van Keken et al., 2002].

1.1 Trace elements

It is perhaps surprising that one of the best sources of geochemical information comes from the elements that are the least abundant. These elements are the trace elements, which are usually defined to be those elements making up less than $\sim 0.1\%$ weight of a rock. They are most useful as passive tracers: They have negligible effect on any process, but are affected by the process. A great advantage is their diversity: there are a wide variety of different trace elements with a wide variety of different behaviours.

A key process in this thesis is melting, and of particular interest is the partitioning of trace elements between liquid melt and solid residue. This is usually measured in terms of a solid/liquid partition coefficient $D_{s/l}$, which measures the ratio of the trace element concentration C_s in the solid phase to the concentration C_l in the liquid phase in chemical equilibrium, $D_{s/l} = C_s/C_l$. Different trace elements have different partitioning behaviour. Those elements with $D_{s/l} \ll 1$ readily enter the melt and are referred to as incompatible elements; those with $D_{s/l} \geq 1$ prefer to stay behind in the residue and are called compatible elements. All the trace elements associated with the radiogenic isotopes discussed in this thesis are incompatible to varying degrees, and this makes them ideal for studying melting processes.

There are several factors that influence the partition coefficients. Two important factors are the ionic charge and ionic radius of an element relative to those of a potential host crystal lattice site. The more similar in ionic charge and ionic radius the trace element is to the atom it replaces in the crystal lattice, the larger the partition coefficient and the more compatible the trace element. Partition coefficients are also affected by pressure and temperature, and by the different mineral phases present in the host rock. This is of

particular note, since melting in the Earth occurs at different depths and thus at different pressures, and also with different mineral assemblages as a result of phase transitions in the mantle. For example, garnet is a stable mantle mineral at depths $\gtrsim 80$ km, but not at shallower depths, and this has a large effect on trace element partitioning. Measurements of trace element concentrations can thus tell us a great deal about melting and crystallisation processes on Earth. For a more in-depth discussion of the factors controlling trace element partitioning see Blundy and Wood [2003].

The rare earth elements (or lanthanides: La to Lu on the periodic table) are a particularly useful set of trace elements. They have similar chemical properties and, with two exceptions (Ce and Eu), always have the same ionic charge of +3. The ionic radius of the rare earth elements decreases with increasing atomic number. As such, there is a systematic variation in the partition coefficients of the rare earths: In mantle melting the lighter rare earths (such as La) are much more incompatible than the heavier rare earths (such as Lu). As a result, the elemental ratio La/Lu will be much higher in the melt than in the source rock. La and Lu are said to be fractionated by the melting process, and such fractionations are crucial to understanding isotope geochemistry.

1.2   Isochron dating

The most well-known use of radiogenic isotopes is in dating. The age of the Earth was first determined accurately by lead isotopic dating. Patterson [1956] measured the lead isotopic ratios ^{206}Pb/^{204}Pb and ^{207}Pb/^{204}Pb for five different meteorites. ^{206}Pb is a stable isotope produced by the radioactive decay of ^{238}U, and ^{207}Pb is a stable isotope produced by the decay of ^{235}U. ^{204}Pb is a reference isotope: it neither decays nor is a decay product. Uranium and lead are both trace elements, and uranium is more incompatible than lead. Geochemists typically measure isotopic ratios rather than absolute concentrations of isotopes as the ratios can be measured more precisely and more accurately.

The presolar nebula from which the Solar System formed is thought to have been well mixed, and thus chemically homogeneous. However, the condensation process that led to the formation of the different meteorites and planets had a chemical fractionation effect, with different elements being concentrated in different bodies to different degrees. In particular, Pb is much more volatile than U and thus condensed later. As a result, the five meteorites formed with different U/Pb elemental ratios. Isotopes of the same element (such as ^{204}Pb, ^{206}Pb and ^{207}Pb) effectively behave identically in such chemical processes, and are not fractionated from each other. Thus the lead isotopic ratios ^{206}Pb/^{204}Pb and ^{207}Pb/^{204}Pb are assumed to be initially identical for all five meteorites. The five meteorites are then assumed to have remained unaltered (a closed system) since their formation at the birth of the Solar System.

The radioactive decay law for a given meteorite can be written as

$$\left(\frac{^{207}\text{Pb}}{^{204}\text{Pb}}\right)_t = \left(\frac{^{207}\text{Pb}}{^{204}\text{Pb}}\right)_0 + \left(e^{\lambda_{235}t} - 1\right)\left(\frac{^{235}\text{U}}{^{204}\text{Pb}}\right)_t, \tag{1.1}$$

$$\left(\frac{^{206}\text{Pb}}{^{204}\text{Pb}}\right)_t = \left(\frac{^{206}\text{Pb}}{^{204}\text{Pb}}\right)_0 + \left(e^{\lambda_{238}t} - 1\right)\left(\frac{^{238}\text{U}}{^{204}\text{Pb}}\right)_t, \tag{1.2}$$

where subscript 0 refers to the time of formation, and t the present time. $\lambda_{235} = 9.85 \times 10^{-10}$ a^{-1} and $\lambda_{238} = 1.55 \times 10^{-10}$ a^{-1} are the decay constants of ^{235}U and ^{238}U respectively (a = annum = year). The half-life $t_{1/2}$ of an isotope is related to its decay constant by
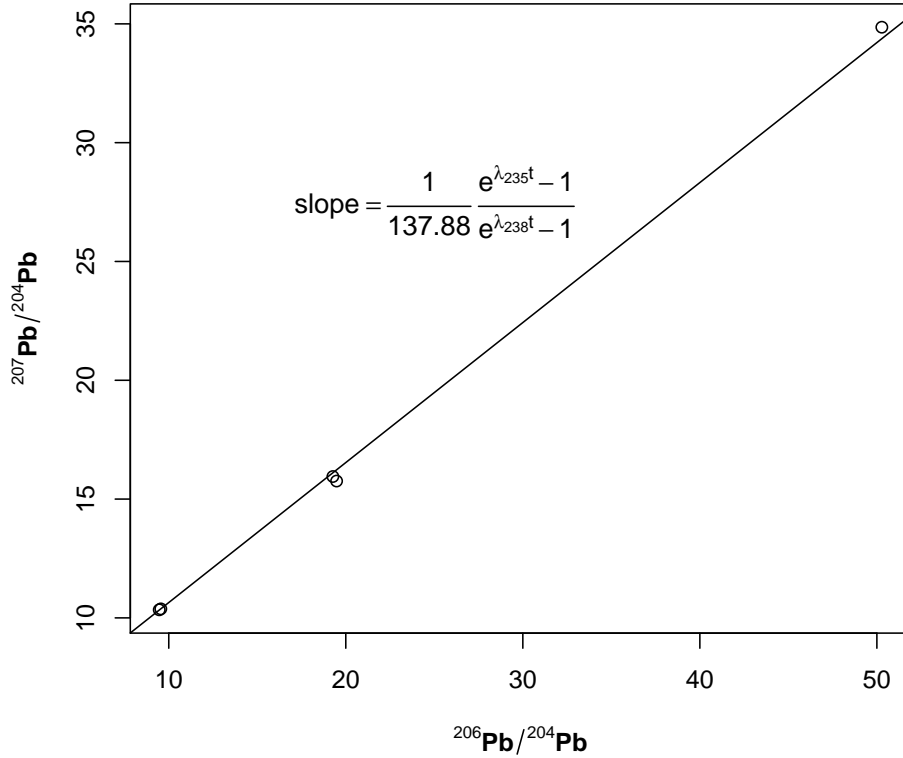
Figure 1.1: Patterson's original lead-lead meteorite isochron, yielding an age of 4.55 Ga.

$t_{1/2} = (\log 2)/\lambda$. By combining (1.1) and (1.2) it follows that

$$\frac{\left({}^{207}\mathrm{Pb}/{}^{204}\mathrm{Pb}\right)_t - \left({}^{207}\mathrm{Pb}/{}^{204}\mathrm{Pb}\right)_0}{\left({}^{206}\mathrm{Pb}/{}^{204}\mathrm{Pb}\right)_t - \left({}^{206}\mathrm{Pb}/{}^{204}\mathrm{Pb}\right)_0} = \frac{e^{\lambda_{235}t} - 1}{e^{\lambda_{238}t} - 1} \left(\frac{{}^{235}\mathrm{U}}{{}^{238}\mathrm{U}}\right)_t. \tag{1.3}$$

This equation forms the basis for lead-lead isochron dating. At the present day, ${}^{235}\mathrm{U}/{}^{238}\mathrm{U} = 1/137.88$ everywhere in the Solar System. Plotting the five meteorites on a diagram of ${}^{207}\mathrm{Pb}/{}^{204}\mathrm{Pb}$ against ${}^{206}\mathrm{Pb}/{}^{204}\mathrm{Pb}$ (Figure 1.1) yields a straight line called an isochron: a line of equal time. The slope of the isochron can then be used in (1.3) to calculate the age of formation of the meteorites, and thus the Solar System, as $t = 4.55 \pm 0.07$ Ga (1 Ga $= 10^9$ years). Note that Patterson's isochron does not date the Earth directly, but we know from other arguments (such as the method of extinct radioactivities) that it must have formed within 10-20 million years of these meteorites. Modern estimates for the age of the Earth differ very little from Patterson's.

1.3 The Sm-Nd system

${}^{147}\mathrm{Sm}$ decays slowly to the stable isotope ${}^{143}\mathrm{Nd}$ with a half-life of 106 Ga. Measurements of the isotopic ratio ${}^{143}\mathrm{Nd}/{}^{144}\mathrm{Nd}$ (where ${}^{144}\mathrm{Nd}$ is a reference isotope) can place important constraints on the evolution of the Earth. Sm and Nd are both rare earth elements. Rare earth elements are refractory (i.e. they have high condensation temperatures), and are generally similar in chemical behaviour. As a result, it is believed that no process in the early solar system fractionated the rare earths from one another (as opposed the fractionation of refractory U from volatile Pb in the solar nebula, discussed in the previous section). In particular, this means that the relative concentrations of the rare earth elements in the

Earth today are the same as those of chondrites (a class of meteorites that have remained essentially unchanged since condensation from the solar nebula). Thus the initial Sm/Nd elemental ratio, and the ^{147}Sm/^{144}Nd and ^{143}Nd/^{144}Nd isotopic ratios of the bulk Earth are the same as those of chondrites. The entire evolution of ^{143}Nd/^{144}Nd for the bulk Earth is thus the same as chondrites and is known (solid line, Figure 1.2).

The Earth has chemically segregated into crust, mantle and core. Core formation is thought to have happened very soon after the Earth's formation, within around 30-50 million years. The Earth excluding the core is referred to as the Bulk Silicate Earth (BSE). Since the rare earths do not partition into the core, their relative concentrations in the BSE are the same as those in chondrites. The continental crust (CC) is thought to have been produced by partial melting of the mantle over the Earth's history, followed by further less well understood differentiation processes. Incompatible elements are thus concentrated in the continental crust (and the crust is said to be enriched in these elements). The residue remaining after extraction of continental crust from the BSE is correspondingly known as the depleted mantle (DM). Since Nd is more incompatible than Sm, the Sm/Nd ratio of the continental crust is lower than chondritic, and the Sm/Nd ratio of the depleted mantle is higher than chondritic. Over time this leads to an elevated ^{143}Nd/^{144}Nd for the depleted mantle, and a lowered ^{143}Nd/^{144}Nd for the continental crust, relative to chondrites. Due to the long half-life of ^{147}Sm, and the small amounts of fractionation between Sm and Nd, the differences in ^{143}Nd/^{144}Nd ratios are rather small, but are readily measurable with modern mass spectrometry. ^{143}Nd/^{144}Nd ratios are often reported relative to the chondritic value as

$$\epsilon_{\text{Nd}} = 10^4 \times \left(\frac{\left(^{143}\text{Nd}/^{144}\text{Nd}\right)_{\text{sample}}}{\left(^{143}\text{Nd}/^{144}\text{Nd}\right)_{\text{CHUR}}} - 1 \right), \tag{1.4}$$

where CHUR is the chondritic value (CHondritic Uniform Reservoir), and the scaling by 10^4 is due to the small differences in measured ^{143}Nd/^{144}Nd.

A simple model of crustal formation is shown in Figure 1.2. Here it is supposed that CC and DM were produced by partial melting of the BSE at 3.5 Ga before the present. The radioactive decay law is given by

$$\left(\frac{^{143}\text{Nd}}{^{144}\text{Nd}} \right)_t = \left(\frac{^{143}\text{Nd}}{^{144}\text{Nd}} \right)_0 + \lambda_{147} t \left(\frac{^{147}\text{Sm}}{^{144}\text{Nd}} \right)_t, \tag{1.5}$$

where a linear approximation $e^{\lambda t} - 1 \approx \lambda t$ has been made due to the long half-life of ^{147}Sm. The partial melting at 3.5 Ga fractionates the Sm/Nd ratios, which over time leads to different present day ^{143}Nd/^{144}Nd ratios for CC and DM. The simple model demonstrates the fundamental processes involved in the Sm-Nd system, but the situation in the real Earth is rather more complicated. A key difference is that the continental crust has continually formed over the Earth's history and not just at one specific time. Moreover, there is good evidence that continental crust has also been recycled back into the mantle [Galer et al., 1989]. Thus more sophisticated models are needed that take into account the open system behaviour of the mantle.

One of the most powerful geochemical modelling tools for understanding the isotopic evolution of the Earth is reservoir (or box) modelling. Reservoir models divide the Earth into a number of physically distinct reservoirs. For example, the Bulk Silicate Earth could be divided into the three reservoirs: continental crust, depleted mantle and primitive mantle. Simple ordinary differential equations can then be written down expressing conservation of chemical species in terms of mass fluxes between the different reservoirs. Different mass transport histories can be investigated and then compared with the geochemical
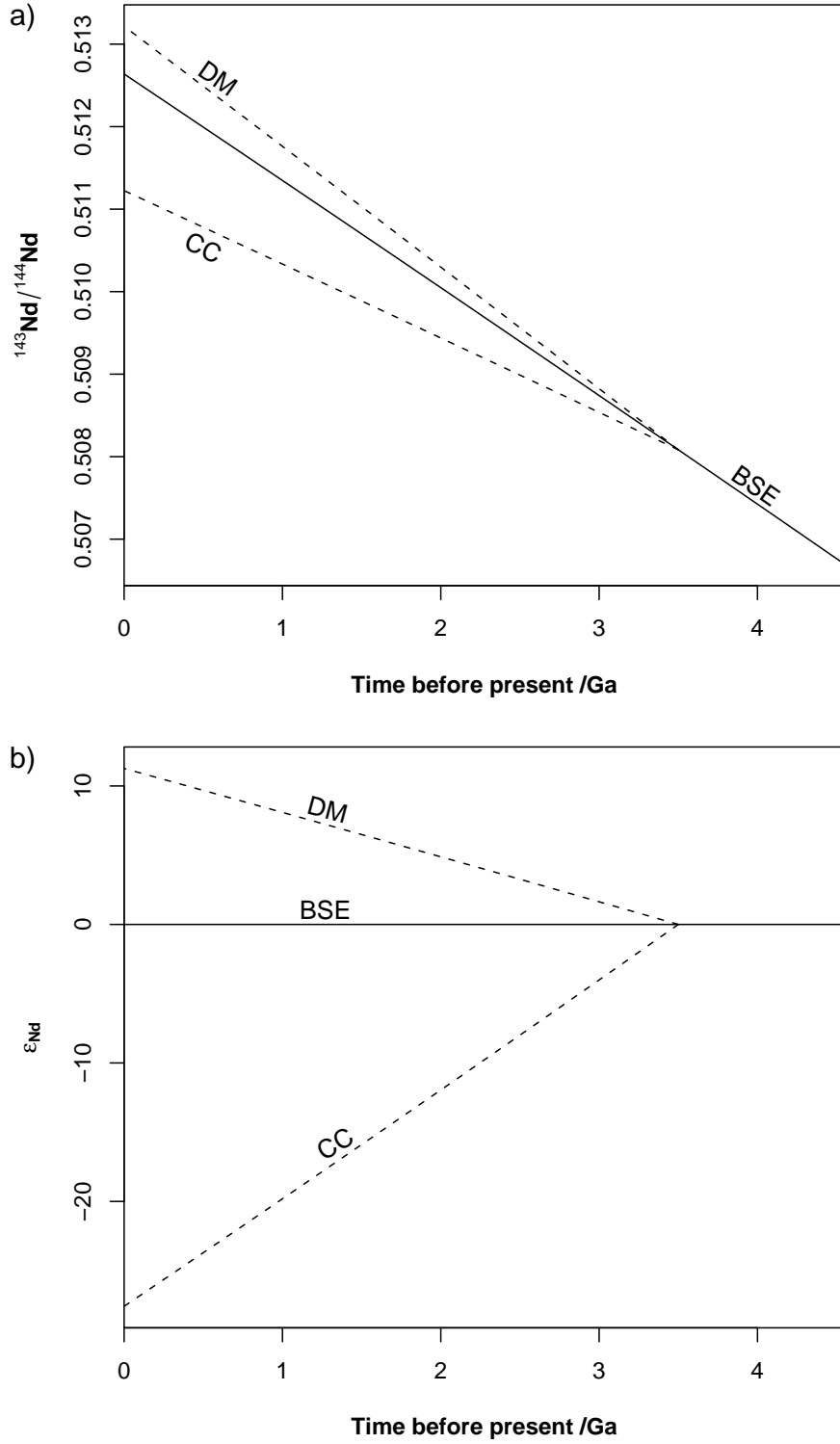
Figure 1.2: a) A simple model of ^{143}Nd/^{144}Nd evolution of mantle and crust. By convention the present is plotted on the left of the diagram, and so the arrow of time is from right to left. The Bulk Silicate Earth (BSE) evolution (solid line) is known, and is the same as that of chondrites. In the above figure it is supposed that continental crust (CC) and depleted mantle (DM) were produced 3.5 Ga before the present by melting from the BSE (dashed lines). Note that the evolution curves are nearly linear as a result of the long half-life of ^{147}Sm. The slopes of the lines are proportional to the Sm/Nd ratio. Present day BSE values are ^{143}Nd/^{144}Nd = 0.512638 and ^{147}Sm/^{144}Nd = 0.1967. b) Same diagram using ϵ_{Nd} notation of (1.4). (Based on Figure 8.9 of White [2005])

observations. Reservoir models are responsible for much of our current understanding of the isotopic evolution of mantle and crust. However, they are limited in that they only consider the mean compositions of each reservoir and thus neglect the differences in composition within each reservoir.

1.4 Mantle heterogeneity

We can only sample the mantle after it has partially melted and that melt has erupted at the Earth's surface. Such melting happens within three main environments on Earth: at mid-ocean ridges (such as the Mid-Atlantic ridge) where two plates separate and melt passively upwells due to decompression; at ocean islands (such as Hawaii), where there is an active upwelling of melt caused by a plume in the mantle circulation; and at island arcs (such as the Aleutians) near subduction zones, where the presence of water and other volatiles causes a region around the subducted plate to melt. The focus of this thesis is the isotopic measurements that have been made on mid-ocean ridge basalt (MORB), as this is the simplest of the three situations to model. There will also be some discussion of ocean island basalt (OIB). Mid-ocean ridges are where most melting of the mantle occurs. MORB has very similar (but not uniform) isotopic characteristics regardless of where in the world it has come from, although there are some small differences between ridges with different spreading rates. OIB varies much more from island to island. The overall degree of melting is also quite different between ridges and ocean islands. There is much more melting under mid-ocean ridges with around $\sim 10\%$ of the underlying mantle melting. The degree of melting at ocean islands is more variable but usually much smaller, around $\sim 1\%$.

The mantle is chemically and isotopically heterogeneous, and this is reflected in the chemical and isotopic heterogeneity of MORB. The observed heterogeneity of isotopic ratios such as ^{143}Nd/^{144}Nd reflects ancient parent/daughter fractionations, indicative of the recycling of old crust and residue from ancient melting events caused by plate tectonics and mantle convection. We wish to understand quantitatively what this heterogeneity can tell us about processes occurring in the mantle, and this requires suitable models of the physical processes.

1.5 Geochemical end-members

A popular approach to understanding the observed isotopic heterogeneity of oceanic basalts is by the identification of end-member components. The concept of geochemical end-members is somewhat loosely defined. The key idea is that the data is considered to result from mixing of a small number of uniform components in different proportions. These uniform components are usually taken to lie at the extremes of the dataset, and are thus called *end*-members.

The simplest example of end-member identification is shown in Figure 1.3, which plots ^{143}Nd/^{144}Nd against ^{87}Sr/^{86}Sr for MORB data. This particular diagram is commonly referred to as "the mantle array". There is a strong negative correlation between the two isotopic ratios; this may be interpreted in terms of mixing between two end-members, for example those indicated by A and B. Any mixture of A and B lies on a line between A and B (called a two-component mixing line). In general, two-component mixing lines on ratio-ratio plots like Figure 1.3 are not straight but curved, unless the denominators in both ratios are the same (such as ^{207}Pb/^{204}Pb against ^{206}Pb/^{204}Pb) or proportional. The linear appearance of the array is often taken to infer that the denominators of the end-members should be chosen to be approximately proportional.
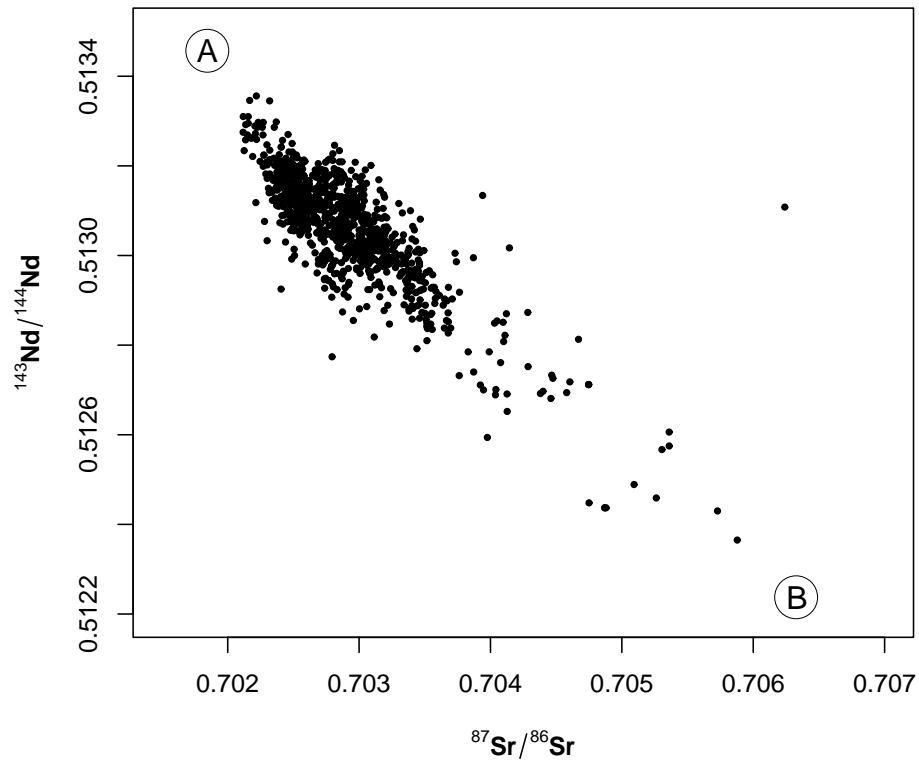
Figure 1.3: The mantle array: a plot of ^{143}Nd/^{144}Nd against ^{87}Sr/^{86}Sr for MORB data. Note the spread of different values, and the clear negative correlation. A and B are respectively possible depleted and enriched end-members. Mixing between these two end-members could account for the majority of the observed correlation.
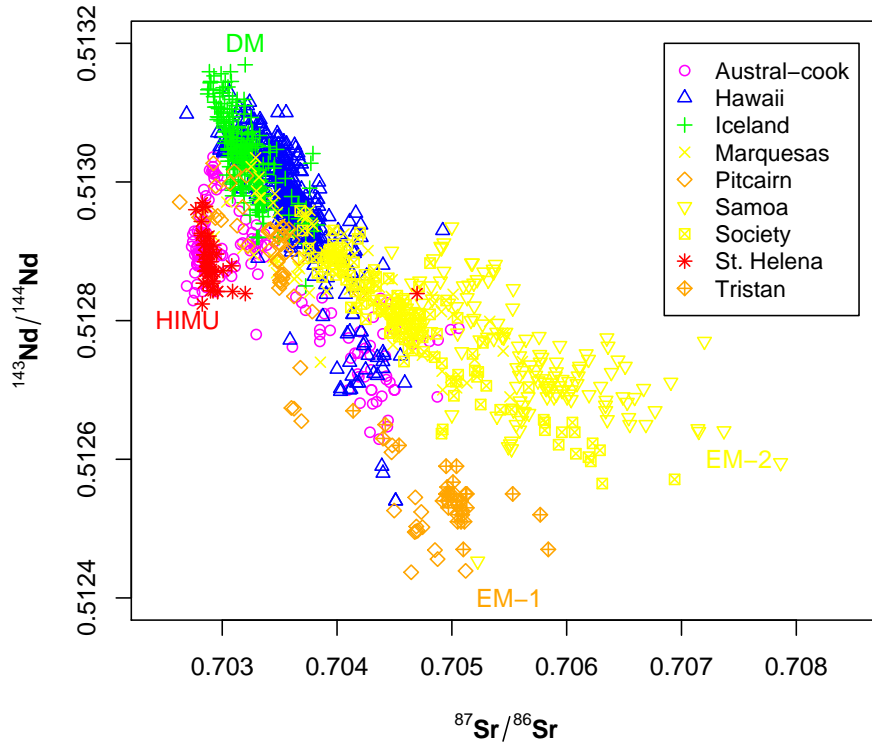
Figure 1.4: The mantle zoo: a plot of ^{143}Nd/^{144}Nd against ^{87}Sr/^{86}Sr for OIB data (based on Figure 15 of Hofmann [2003]). Islands thought to be of the same type are coloured the same, and the four standard end-members DM, HIMU, EM-I and EM-II have been plotted.

Based on data from several different isotopic systems, and from the wide range of ocean island basalts, geochemists have identified a whole host of mantle end-member components. An example, showing the four standard end-members for OIB data, is shown in Figure 1.4. These four end-members have the acronyms DM (depleted mantle), EM-I (enriched mantle 1), EM-II (enriched mantle 2) and HIMU (high $\mu = {}^{238}$U/^{204}Pb) [White, 1985, Zindler and Hart, 1986]. Sometimes an additional component is identified in the centre of the array, referred to by different authors as C (common), FOZO (focal zone) or PREMA (prevalent mantle). This taxonomy of oceanic basalts in terms of end-member components has, quite rightly, been called the "mantle zoo". The different geochemical end-members have been interpreted in terms of different processes occurring in the Earth, or in terms of different reservoirs (for example, the acronym DM refers to both an end-member and a mantle reservoir).

How are end-members chosen? The simplest approach is by eye: to examine the data for linear arrays, and assign end-members to either end. However, with this approach a single isotope ratio - isotope ratio plot can lead quickly to a large number of end-members, the choice of which is largely subjective. A more objective approach is to ask: what is the minimum number of end-members that is consistent with a data set? If we have a data set of n isotopic ratios, then this requires mixing between at most $n + 1$ end-members. Having more than $n + 1$ end-members means that the individual observations cannot be expressed uniquely in terms of the end-members. With statistical techniques, such as principal component analysis (PCA), we can go further and determine the number of significant end-members [Debaille et al., 2006]. PCA reduces the dimensionality of the

data to the m most statistically significant principal components, where $m \leq n$. The data can then be interpreted in terms of $m + 1$ significant end-members. However, it should again be noted that there is a potential problem here: PCA is a linear analysis technique, and the mixing lines between end-members are not necessarily linear.

Understanding isotopic heterogeneity with end-members has a number of problems. Firstly, given a dataset, there is no unique way to identify the end-members. For example, end-members can always lie an arbitrary distance further away from the arrays. The second issue is one of interpretation. Oceanic basalts are simply not made by different amounts of mixing between a small number of mantle components. In reality, the continual processes of fractionation occurring on Earth will produce an infinite range of different mantle components which may be mixed together. End-member analysis helps describe some of the main features of the complex isotopic systematics of oceanic basalt, but new modelling approaches are certainly needed.

1.6 Statistical box models

This thesis takes a different view of mantle heterogeneity from that of end-member components and reservoir models. Rather than considering mixing between a few uniform mantle components, a model is produced for the overall statistical distribution of isotopic ratios. The focus is on understanding the broad features of the observed heterogeneity, and most importantly on modelling in terms of the physical processes responsible. There has been some recent work in this direction, either through modified reservoir models [Kellogg et al., 2002] or through the advection of tracers in mantle convection simulations [Christensen and Hofmann, 1994, Ferrachat and Ricard, 2001, Xie and Tackley, 2004a,b]. Both of these approaches require involved numerical calculations, with the modified reservoir models being rather less involved than the convection simulations. The approach taken in this thesis is somewhat akin to the modified reservoir models of Kellogg et al. [2002], but is much more straightforward and has the great advantage of being analytically tractable. The model is highly idealised, but includes the essential physical processes involved: radioactive decay, fractionation produced by melting, recycling and stirring by mantle convection, and the mixing of melts that occurs before sampling.

The bulk of this thesis, chapters 2 to 6, describe this new approach to modelling mantle isotopic heterogeneity. To make the thesis more accessible, the detailed mathematical derivations have been separated out, and are described in chapters 3 and 5.

1.6.1 Long-lived isotopes

Chapters 2 and 3 describe the simplest form of the model, which tackles the isotopic ratios $^{143}\text{Nd}/^{144}\text{Nd}$, $^{176}\text{Hf}/^{177}\text{Hf}$, $^{87}\text{Sr}/^{86}\text{Sr}$, $^{208}\text{Pb}/^{204}\text{Pb}$ and $^{206}\text{Pb}/^{204}\text{Pb}$ which result from the decay of long-lived (with respect to the age of the Earth) parent isotopes. The model is a simple statistical box model of mantle processes. A single melt region produces a melt fraction F of melt, and the average time since a given parcel of mantle material last visited this region is given by a time scale τ_{melt}. The melt region fractionates the parent/daughter ratios. Over time this leads to variations in the mantle isotopic ratios as the parent decays to the daughter. Key assumptions are that the flow is strongly stirring, and that the mantle has reached a statistical steady state. These assumptions allow the neglect of the specifics of the underlying flow. Sampling from the model mantle is performed by averaging over a large number N of samples, representative of the mixing that occurs after melting [Meibom and Anderson, 2004].

The model predicts a probability density for isotopic ratios. Fitting the model to the MORB data gives estimates of the parameters F, τ_{melt} and N. Small melt fractions with F around 0.5% are essential for a good fit, whereas τ_{melt} and N are less well constrained. τ_{melt} is estimated from the variance of the Nd, Sr and Hf isotopic ratios at around 1.4 to 2.4 Ga, and N is of the order of hundreds. With these parameters, the model predicts a larger variability for the Pb isotopes than that observed. Fundamental differences are shown to exist between the dynamics of Pb isotopes and those of Nd, Sr and Hf isotopes.

1.6.2 Pseudo-isochrons

Chapters 4 and 5 generalise the statistical box model to tackle shorter-lived isotopes, notably the isotopic ratio $^{207}\text{Pb}/^{204}\text{Pb}$, and discuss the problem of mantle pseudo-isochrons. If MORB data is plotted on the same diagram of $^{207}\text{Pb}/^{204}\text{Pb}$ against $^{206}\text{Pb}/^{204}\text{Pb}$ used to date the meteorites (Figure 1.1), then an approximate linear relationship is found. If naively the isochron method is used on the slope of a regression line through the data to calculate an age, a figure of around 2.0 Ga is found. However, since the mantle is not a closed system (as the meteorites are) the isochron dating method is not valid, and thus this age of 2.0 Ga is often referred to as a pseudo-isochron age. An important question in mantle geochemistry is: how to interpret this pseudo-isochron age? It certainly does not date a single fractionation event as the mantle is constantly processed through melting. The statistical box model offers a solution to this problem and provides a way of relating the pseudo-isochron age to physical parameters.

Simple equations are presented which relate the pseudo-isochron age to the decay constants and distribution of heterogeneity ages in the model mantle. In turn, this age distribution is simply related to the history of melting. The equations are in good agreement with results from mantle convection simulations [Christensen and Hofmann, 1994, Xie and Tackley, 2004a]. The equations are different from, but related to and more general than, those found previously for reservoir models [Albarède, 2001, Allègre and Lewin, 1995a, Donnelly et al., 2004]. While the pseudo-isochron age does not signify a mean age in the usual sense, in the model presented it is related to a "generalised mean" over the distribution of heterogeneity ages. If a constant melt rate over the Earth's history is assumed, the 2.0 Ga lead-lead pseudo-isochron age is shown to infer a mean remelting time τ_{melt} of 0.5 Ga.

Concluding the statistical box model work, Chapter 6 discusses the shortcomings of the approach, and suggests some generalisations that may be useful in future.

1.7 Uranium series disequilibria

The remaining part of the thesis relates to an altogether different problem, and to a different use of radiogenic isotopes. The decay of ^{238}U to ^{206}Pb, ^{235}U to ^{207}Pb, and ^{232}Th to ^{208}Pb takes place through a series of short-lived intermediate decay products (Figure 1.5). These decay series can be used to place important constraints on the time scales of melt extraction from the mantle. Over long times in a closed system, a state of radioactive equilibrium arises, whereby each of the daughter isotopes in the decay chain are produced at the same rate as they decay. It is useful to consider the activity of an isotope, defined as the product of an isotope's concentration C and its decay constant λ, and is denoted by round brackets, e.g. $(X) = \lambda_X C_X$. In radioactive equilibrium, the activities of all the isotopes in the chain are identical.
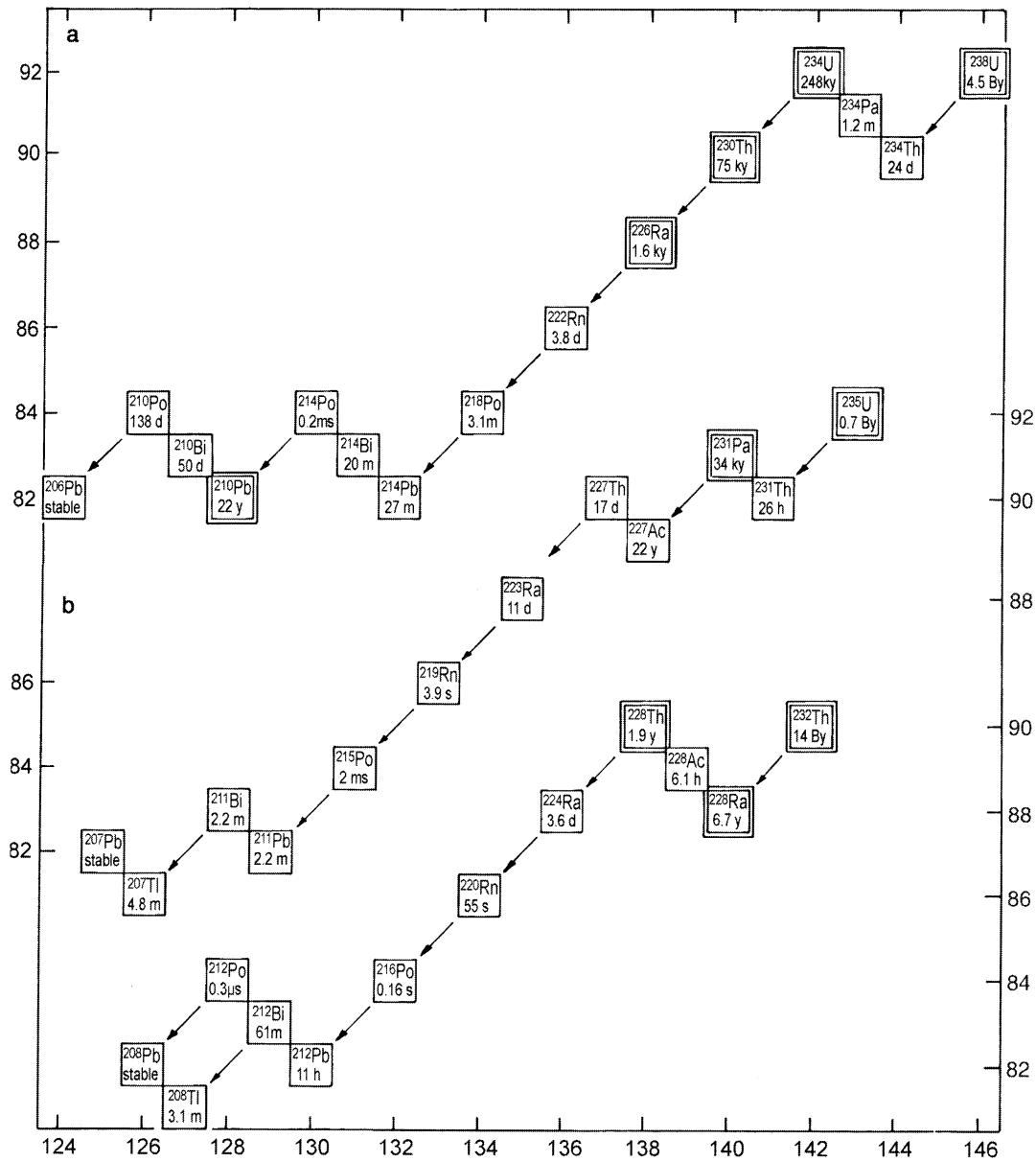
Figure 1.5: The decay series of ^{238}U, ^{235}U and ^{232}Th (reproduced from Dickin [2005]), showing the intermediate decay products and their half-lives. Some of the geochemically more useful isotopes are indicated by double boxes. Arrows indicate alpha decay; diagonally adjacent isotopes indicate beta decay.

An important feature of decay chains is the time taken to return to radioactive equilibrium after all of a particular isotope is removed. Perhaps counterintuitively, this return to equilibrium is governed by the decay constant of the isotope removed and not by the decay constant of its parent. Disequilibrium is usually measured by an activity ratio, e.g. $(^{230}\text{Th}/^{238}\text{U})$. In a state of radioactive equilibrium $(^{230}\text{Th}/^{238}\text{U}) = 1$, and a ratio different from 1 signifies disequilibrium. After removal of ^{230}Th, the activity ratio $(^{230}\text{Th}/^{238}\text{U})$ returns to 1 on a characteristic time scale $1/\lambda_{230\text{Th}} \approx 100,000$ years.

Since the half-life of ^{230}Th is very short compared with the time scales involved in mantle convection, mantle material just before melting must be in radioactive equilibrium with $(^{230}\text{Th}/^{238}\text{U}) = 1$. However, disequilibrium activity ratios with $(^{230}\text{Th}/^{238}\text{U}) \approx 1.15$ are measured on MORB. This disequilibria must arise from the melting process, and requires that Th is extracted more quickly than U during melting. Since Th is only more incompatible than U at depths greater ~ 80 km in the mantle (in the garnet stability field), this disequilibrium signal can only be generated near the onset of melting, which occurs at such depths. Migration of melt to the surface from where it is first produced must happen in less than 100,000 years, in order to preserve this disequilibrium signal.

Shorter-lived isotopes place further constraints on melt migration. Disequilibrium activity ratios have been measured for $(^{226}\text{Ra}/^{230}\text{Th})$ (half-life of $^{226}\text{Ra} = 1,600$ years) and more recently for $(^{210}\text{Pb}/^{226}\text{Ra})$ (half-life of $^{210}\text{Pb} = 22$ years) [Rubin et al., 2005]. However, equilibrium activity ratios have been measured for $(^{228}\text{Ra}/^{232}\text{Th})$ (half-life of $^{228}\text{Ra} = 6.7$ years) and more recently for $(^{227}\text{Ac}/^{231}\text{Pa})$ (half-life of $^{227}\text{Ac} = 22$ years) [K. Sims, personal communication 2006]. It is harder to constrain the exact depths at which the different disequilibria signals are generated, but it appears likely that the degassing of radon (Rn) near the surface is responsible for the disequilibrium $(^{210}\text{Pb}/^{226}\text{Ra})$. Thus, current observations suggest that the time scale of melt extraction is constrained by the half-lives of ^{227}Ac and ^{226}Ra at between 22 and 1,600 years.

An important question is now: what are the physical mechanisms whereby melt travels to the surface so quickly? One possibility is that the melt is organised in a series of connected melt conduits. Evidence for such a network of conduits is also provided by geological observations of ophiolites (rocks thought to be samples of oceanic crust and upper mantle material) [Kelemen et al., 1997, 2000]. Channel flow in the conduits allows melt to be extracted fast enough to satisfy the U-series constraints. Motivated by this, Richardson et al. [1996] studied the fluid dynamics of a isolated melt conduit surrounded by a partially molten region. The conduit was shown to support solitary waves, with a region of trapped melt travelling with the wave. Chapter 7 revisits their analysis, and comes to a different conclusion. It seems that the region of trapped melt may not exist, and this may have implications for the chemistry. Finally, chapter 8 concludes the thesis.

Chapter 2

The isotopic heterogeneity of mid-ocean ridge basalt

2.1 Introduction

Many measurements of isotopic ratios such as ^{143}Nd/^{144}Nd, ^{176}Hf/^{177}Hf and ^{87}Sr/^{86}Sr have been made on mid-ocean ridge basalt (MORB). These measurements in principle contain significant information on melting and convective processes in the Earth's mantle, but to extract that information requires a suitable quantitative model. In this chapter a simplified model is presented of what are believed to be the fundamental processes that give rise to the isotopic heterogeneity of MORB. This chapter describes the formulation of the model, summarises the results, and compares them with observations. The following chapter presents the mathematical derivations involved.

The model presented in this chapter describes well the isotopes ^{143}Nd/^{144}Nd, ^{176}Hf/^{177}Hf, and ^{87}Sr/^{86}Sr, but not the isotopes of Pb. Generalisations of the model to tackle Pb isotopes and noble gases (such as He) will be discussed in later chapters.

2.2 Key processes

The isotopic systems studied consist of three isotopes, whose molar concentrations are labelled as p, d and d'. p is the parent isotope, which decays with decay constant λ to the daughter isotope d. d' is a stable reference isotope of the same element as d, whose abundance does not change due to radioactive decay. It is measurements of isotopic ratios d/d' that are made on MORB.

The heterogeneity of d/d' in the model arises from partial melting processes in the mantle which fractionate the parent element from the daughter element. It is assumed that isotopes of the same element have the same behaviour under melting, and so the ratio d/d' is unchanged after melting. The fractionation leads to differences in the p/d' ratio of different parts of mantle material. Over time this leads to differences in d/d' as the parent isotope p decays to the daughter isotope d. While melting processes create heterogeneity, they also are responsible for destroying it. When the mantle melts, there is a mixing of different melts which smooths out the variations in d/d'.

These processes occur on the background of mantle convection and plate tectonics. Mantle convection stirs the mantle and reduces the length scale of heterogeneities. The distinction between stirring by the convection, and the mixing of the melts should be emphasised - stirring does not reduce the heterogeneity. The convection also plays a key role in recycling. Melted material comes to the surface to form new crust, which may

later become subducted back in to the mantle and subsequently undergo further melting events.

Finally, it is important to note that we cannot sample from the mantle directly: we can only do so after it has melted and come to the surface. Therefore we expect the variations observed in MORB to be less than those in the underlying mantle.

2.3 The model

The model consists of two stages. First the processes that create and destroy heterogeneity in the mantle are examined to derive the form of the distribution of d/d' in the interior. Secondly the processes involved in sampling after melting from this model mantle are examined to predict how this interior distribution relates to a distribution expected to be observed at mid-ocean ridges.

An episodic model is considered with a box representing the mantle. For an interval of time Δt parcels are advected around the box, and the appropriate radioactive decay is followed. At the end of this interval, a melting event occurs. In a melting event a single melt fraction F of melt is produced. The concentrations of p, d and d' are averaged over in the melting region of the box, and this is used to calculate the relevant concentrations in the melt and residue produced. These concentrations depend on the partition coefficients for each isotope, and it is assumed that isotopes of the same element, such as d and d', have the same behaviour under melting. A fraction F of the material in the melting region is then set with the uniform melt value composition, and a fraction $1-F$ with the uniform residue composition. In this way the melt region both creates and destroys heterogeneity.

In order to simplify the model two key assumptions are made. The first is that the radioactive decay is slow, so that the equation of radioactive decay can be linearised (this assumption will later be dropped in Chapter 4). The second is that the advection is strongly stirring. By this it is meant that the stirring is such that the statistics of the melting region at the end of the interval are the same as the statistics over the whole box. This situation gives rise to a statistical steady state [Armstrong, 1968], and a model mantle which is statistically of the same isotopic composition everywhere (although not a uniform composition). For calculational convenience, the limit is also taken as $\Delta t \to 0$, turning the episodic model into a continuous one. An important parameterisation of the melting process is then the timescale τ_{melt}, which determines how often on average it has been since an individual parcel last experienced a melting event. The assumption that the radioactive decay is slow is then formally that $\lambda \tau_{\mathrm{melt}} \ll 1$.

It is then possible to derive analytically an expression for the distribution of d/d' in the model mantle (section 3.1).

2.4 Sampling

As mentioned earlier we cannot sample from the mantle directly, but instead sample after melting. To model this the ideas of "sampling under melting and averaging" (SUMA) of Meibom and Anderson [2004] have been extended. To model the sampling a large number N of independent identically distributed samples from the model mantle are mixed together. N characterises the mixing - if N is large, there is a large amount of mixing between different components and thus a large reduction in the variability observed. N is a simple single parameter characterisation of the averaging process, which in practice is determined by a variety of different factors such as the length scale of heterogeneities, the

length scale of the mantle region undergoing partial melting, the diffusivity of chemical species in the melt, and the time spent in magma chambers.

In the case of averaging large numbers of independent samples, the central limit theorem is potentially relevant [Allègre and Lewin, 1995b, Meibom and Anderson, 2004]. It states that the average of N independent random samples from a distribution with mean μ_0 and standard deviation σ_0 approaches a normal distribution with mean μ_0 and standard deviation $\sigma_0/\sqrt{N}$ as N increases. However, the central limit theorem is not directly applicable to the distribution of the ratio d/d' as it is the individual concentrations d and d' that are averaged and then the ratio d/d' calculated and not the other way round. However, it can be shown that the model d/d' distribution does approach a normal distribution as N increases, but the standard deviation is not simply $\sigma_0/\sqrt{N}$. Section 3.2 explains the details of calculating the distribution after sampling, and section 3.3 considers the large N asymptotics.

2.5   Analytic results

For N large there are a number of asymptotic results that can be derived for the properties of the d/d' distribution after sampling. The most important analytic result is for the asymptotic behaviour of the standard deviation σ of the distribution,

$$\sigma \approx \lambda \tau_{\mathrm{melt}} |G_p - G_d| \frac{\bar{p}}{\bar{d}'} \sqrt{\frac{2}{NF(1-F)}}. \tag{2.1}$$

Here λ is the decay constant of the parent isotope; $\bar{p}/\bar{d}'$ is the ratio of the mean parent isotope concentration to mean reference isotope concentration over the whole box at the present day; G_p and G_d are functions of the melt fraction F and the relevant partition coefficients, where the lower case subscripts p and d refer to the parent and daughter isotopes under consideration. G is the molar fraction of a chemical species that enters the melt, and thus is a measure of the incompatibility of a given isotope under a particular melt fraction. If $G \approx 0$ then the isotope is compatible (stays in the residue), while if $G \approx 1$ then it is incompatible (readily enters the melt). For the Shaw [1970] melting model G_p is given by

$$G_p = 1 - \left(1 - \frac{P_p}{D_p}F\right)^{1/P_p}, \tag{2.2}$$

where capital P and D are different weighted averages of partition coefficients.

$$D_p = \sum_i x^i K_p^i, \tag{2.3}$$

$$P_p = \sum_i q^i K_p^i. \tag{2.4}$$

K_p^i is the mineral/melt partition coefficient for the parent in mineral i; x^i is the mass fraction of mineral i in the source rock; and q^i is the fractional contribution of mineral i to the melt. G_d is defined similarly.

Some features of the equation for σ can be easily interpreted. Variations in d/d' arise from differences in p/d'. Since linearised decay is assumed, σ depends linearly on λ and $\bar{p}/\bar{d}'$. Also, because the only other time scale in the problem is τ_{melt} and since σ is non-dimensional, then on dimensional grounds σ must have a linear dependence on τ_{melt} also. So the larger τ_{melt}, the longer it is between melting events and thus the greater the heterogeneity that can arise.

The role of the melting behaviour on σ is less obvious. The dependence on $|G_p - G_d|$ is indicative of the fact there has to be a fractionation between parent and daughter to produce the heterogeneity. As such, small melt fractions are essential. For example, large melt fraction MORB melting would have $G \approx 1$ for nearly all elements under consideration and as such can not produce significant heterogeneity. As mentioned earlier N is indicative of the amount of mixing after melting, so asymptotically the larger N the smaller σ.

Another analytic result of particular interest is for the skewness parameter γ_1 (section 3.3.2), which measures the asymmetry of the distribution about its mean value,

$$\gamma_1 \approx \frac{3 \left(1 - 2G_d\right) \operatorname{sgn}(G_p - G_d)}{\sqrt{2NF(1 - F)}}. \tag{2.5}$$

Here $\operatorname{sgn}(x)$ is the sign function: it equals 1 if x is positive, -1 if x is negative. Hence the sign of the skew is determined by the sign of $(G_p - G_d)\left(1 - 2G_d\right)$. For example, if $G_p < G_d$ (parent more compatible than daughter) and $G_d < 1/2$ (a sufficiently small melt fraction say) then the distribution will be asymptotically negatively skewed. Furthermore, note the dependence on $NF(1 - F)$ in the denominator. While N may be thought of as a large parameter, F is usually small, so the skew and non-normality in the distributions can be quite significant.

It is common in geochemistry to plot one isotopic ratio against another. Consider two isotopic systems 1 and 2 with different parent and daughter elements. Suppose $(d/d')_2$ is plotted against $(d/d')_1$. For N large, the model predicts that samples lie approximately on a straight line through the mean value, and gives a prediction for the gradient $\bar{\beta}$ of this line (section 3.3.4)

$$\bar{\beta} \approx \frac{\lambda_2 \left(G_{p2} - G_{d2}\right) \left(\bar{p}/\bar{d}'\right)_2}{\lambda_1 \left(G_{p1} - G_{d1}\right) \left(\bar{p}/\bar{d}'\right)_1}. \tag{2.6}$$

Thus, if the partition coefficients, half lives, and mean parent/daughter isotopic ratios are known for each system, the melt fraction can be inferred from the gradient in plots of $(d/d')_2$ against $(d/d')_1$.

This equation simplifies in a particular special case. If the daughter isotopes are of the same element for both systems (for example, both isotopes of lead) then all samples lie precisely on a straight line. If it is also the case that the parent elements are the same (or at least behave similarly under melting such that $G_{p1} \approx G_{p2}$) then the gradient β of this line is independent of the melt model and is simply

$$\beta = \frac{\lambda_2 \bar{p}_2}{\lambda_1 \bar{p}_1}. \tag{2.7}$$

Note that this equation will be valid for $^{208}\text{Pb}/^{204}\text{Pb}$ and $^{206}\text{Pb}/^{204}\text{Pb}$, but not $^{207}\text{Pb}/^{204}\text{Pb}$. The short half-life of ^{235}U means the $^{207}\text{Pb}/^{204}\text{Pb}$ system does not satisfy the slow decay approximation. As such, this version of the model cannot be used for $^{207}\text{Pb}/^{204}\text{Pb}$. It has been included in the data plots for completeness, and the model will be extended in chapter 4 to handle isotopes with shorter half-lives.

2.6   Comparison with observations

The model has the following input parameters: λ, $\bar{p}/\bar{d}'$, $\bar{d}/\bar{d}'$, P_p, D_p, P_d, D_d, τ_{melt}, F, and N. The key concern is with studying the variations in d/d' so $\bar{d}/\bar{d}'$ is not important. $\bar{d}/\bar{d}'$ is chosen appropriately to get a match with the mean of the observations. Other parameters are based on estimates by previous authors as shown in Table 2.1. The melting

Table 2.1: Given parameters.

		^{147}Sm	^{87}Rb	^{176}Lu	^{232}Th	^{238}U	^{235}U
parent isotope	p	^{147}Sm	^{87}Rb	^{176}Lu	^{232}Th	^{238}U	^{235}U
daughter isotope	d	^{143}Nd	^{87}Sr	^{176}Hf	^{208}Pb	^{206}Pb	^{207}Pb
reference isotope	d'	^{144}Nd	^{86}Sr	^{177}Hf	^{204}Pb	^{204}Pb	^{204}Pb
decay rate / Ga^{-1a}	λ	0.00654	0.0142	0.0193	0.0495	0.155	0.985
parent part. coeff.^{b}	D_p	0.03990	0.00039	0.29739	0.00021	0.00054	0.00054
parent part. coeff.^{b}	P_p	0.12392	0.00058	0.94719	0.00068	0.00178	0.00178
daughter part. coeff.^{b}	D_d	0.02787	0.02983	0.04508	0.02793	0.02793	0.02793
daughter part. coeff.^{b}	P_d	0.08653	0.08908	0.14333	0.07175	0.07175	0.07175
mean isotopic par./dau.^{c}	$\bar{p}/\bar{d}'$	0.253	0.0188	0.0324	20.0	8.04	0.0583

(a) Decay constants λ of the parent isotopes are accurately known [Faure, 1986]. Some λ have been rounded.

(b) Partition coefficients based on the estimates of McKenzie and O'Nions [1991], taking a linear combination of the values for garnet and spinel peridotite (Appendix A).

(c) $\bar{p}/\bar{d}'$ estimated from the depleted mantle model of McKenzie and O'Nions [1991] (Appendix A).

is assumed to take place somewhere in the garnet/ spinel transition region. As such the partition coefficients used are a linear combination of the coefficients for garnet and spinel peridotite. With the exception of the Lu-Hf system, there is little difference in the partition coefficients for these two rock types, so the Lu-Hf system was used to constrain this linear combination.

There are just three unknown parameters left in this model: the melting timescale τ_{melt}, the effective melt fraction F producing the heterogeneity, and the parameter N characterising the mixing before sampling. Note also that these three parameters are independent of the isotopic system under consideration.

2.6.1 Samarium - neodymium

The Sm-Nd system is the best constrained of the systems under examination. The Sm/Nd ratio of the bulk Earth is taken to be chondritic, and there is general agreement on the values for the partition coefficients of Sm and Nd [Green, 1994]. Hence the Sm/Nd ratio in depleted mantle models can be calculated accurately.

2.6.2 Rubidium - strontium

The Rb-Sr system is less well constrained. Owing to the volatility of Rb, the bulk Earth Rb/Sr ratio is not precisely known. Furthermore, the observed behaviour of Sr under melting is not well described by the partition coefficients that have been measured in the laboratory [Hanson, 1977]. Hence estimates of Rb/Sr in depleted mantle models vary considerably (for example, a 35% larger estimate than that in Table 2.1 is given by Salters and Stracke [2004]).

2.6.3 Lutetium - hafnium

The Lu-Hf system is similar in many ways to the Sm-Nd system, in that good estimates on the Lu/Hf ratio can be obtained from chondrites. However, this system is complicated by the large difference between partition coefficients for Lu in the spinel and garnet stability
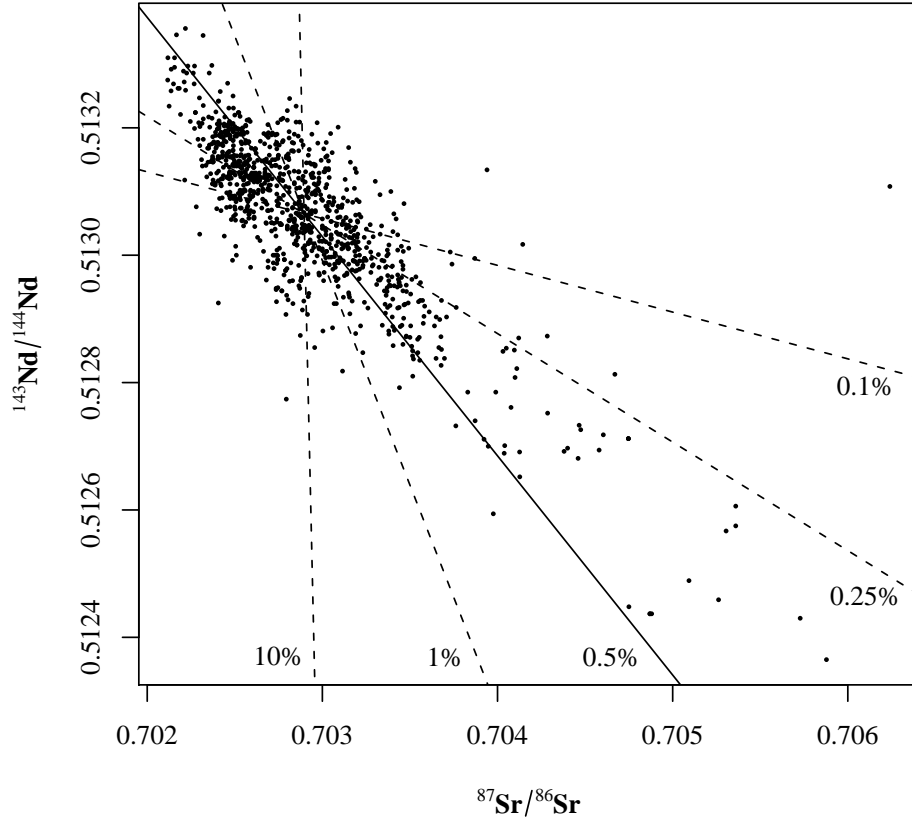
Figure 2.1: Plot of ^{143}Nd/^{144}Nd against ^{87}Sr/^{86}Sr. Measured MORB data points are plotted. Straight lines are model predictions using (2.6) with melt fraction F as indicated. A melt fraction of 0.5% provides a good fit.

fields. A 37% larger estimate for Lu/Hf is given in Salters and Stracke [2004]. Also, fewer measurements of hafnium isotopes have been made.

2.6.4   Uranium - thorium - lead

The U-Th-Pb systems are the most difficult to model. The ratio of U/Pb in the depleted mantle is not well known (discussed later), and the behaviour of lead under melting is not well understood [Tatsumoto, 1978, White, 1993].

2.7   Results

The three parameters F, N and τ_{melt} were varied in order to fit the model to the measurements of MORB isotopic ratios (sources given in Appendix A).

   F is the most straightforward parameter to constrain, provided the decay rates, partition coefficients and mean parent/ daughter isotopic ratios are known. (2.6) shows that the average gradient in plots of one isotopic ratio against another are determined by the melt fraction F, and is independent of N or τ_{melt} to leading order. Figure 2.1 gives an

example of such a plot, and shows how F can be varied to provide a good fit. The Sm-Nd and Rb-Sr systems constrain F to be around 0.5%.

N and τ_{melt} are harder to constrain, mainly due to the dependence on $\tau_{\mathrm{melt}}/\sqrt{N}$ of the standard deviation for N large. Given F, the standard deviations only constrain $\tau_{\mathrm{melt}}/\sqrt{N}$, and so a longer timescale could be offset by a larger amount of mixing before sampling. Increasing τ_{melt} stretches the density linearly. The behaviour of N is more complicated, as it also has an effect on the shape of the distributions (Figure 2.2). For N large the model distribution approaches a normal distribution, and yet there is a definite skew to the distributions that are observed. Hence all that can be said with any certainty is that τ_{melt} is around 1.2 - 2.4 Ga, and N is of the order of hundreds to get a reasonable fit to the data.

Motivated by these constraints, Figures 2.3 and 2.4 compare the measurements with model data using F=0.5%, τ_{melt}=1.7 Ga and N=500. Figure 2.3 shows a plot of kernel smoothed probability density estimates for model and measured data. Kernel smoothed density estimates (section 3.4) are a generalisation and improvement on histograms. They have the advantage of being smooth, and have no dependence on the end points of bins as histograms do. They do have a dependence on the bandwidth of the kernel, which determines how smoothed out distribution features become. However, there are algorithms for choosing the most appropriate bandwidth, and these have been used. Figure 2.4 shows a more traditional set of scatterplots for pairs of the isotopic systems. These plot an artificial sample from the model, as well as the measured data. Table 2.2 lists the values of a few important calculated parameters.

In this table $G_d < 1/2$ for all isotopic systems. Hence (2.5) implies that the sign of the skew is determined by the sign of $G_p - G_d$, i.e. by the relative compatibility of parent and daughter. Where the parent is more compatible than the daughter ($G_p - G_d < 0$ e.g. Sm-Nd, Lu-Hf) a negatively skewed distribution is expected, and where the parent is less compatible ($G_p - G_d > 0$ e.g. Rb-Sr, U-Th-Pb) a positively skewed distribution. This is what is observed. Note also that $NF(1-F) = 2.5$ so that the skews and the non-normality in these distributions are appreciable. Furthermore, note that the sign of the correlations in the plots of one isotope against another is given by $(G_{p2} - G_{d2})/(G_{p1} - G_{d1})$ and thus the observed correlations are also governed by the relative compatibilities of parent and daughter elements.

The special case that yields (2.7) applies to the plot of ^{206}Pb/^{204}Pb against ^{208}Pb/^{204}Pb. In this case the model predicts a single straight line whose gradient β depends only on the decay constants of ^{238}U and ^{232}Th, and κ, the average isotopic ratio ^{232}Th/^{238}U in the model mantle ($\beta = (\lambda_{232}/\lambda_{238})\kappa$). Since the decay constants are known, the slope of the plot of ^{206}Pb/^{204}Pb against ^{208}Pb/^{204}Pb can be used to estimate κ. The value of κ used was 2.5, and this provides a reasonable fit to the observed slope.

There is reasonably good agreement between the model and the measurements of Nd, Sr and Hf isotopic ratios. The MORB data shows more scatter than the model, but the general trends are the same. Such scatter could possibly be accounted for by the variability in melt fractions in the real mantle. The Pb isotopes, however, do not fit this simple model. The standard deviation of Pb calculated from the model is larger than that observed. An ad hoc solution to this problem is to reduce the average ratio ^{238}U/^{204}Pb, μ, while keeping κ unchanged. A value for mean depleted mantle μ of around 2.7 (as opposed to the value of 8.0 used) gives a good match with these parameters. There is still much debate on the value of μ in the depleted mantle, with some authors arguing it should be less than 4.7 [White, 1993], or as high as 14.4 [Salters and Stracke, 2004].

The isotopes of Nd, Sr and Hf correlate well with each other. The isotopes of Pb
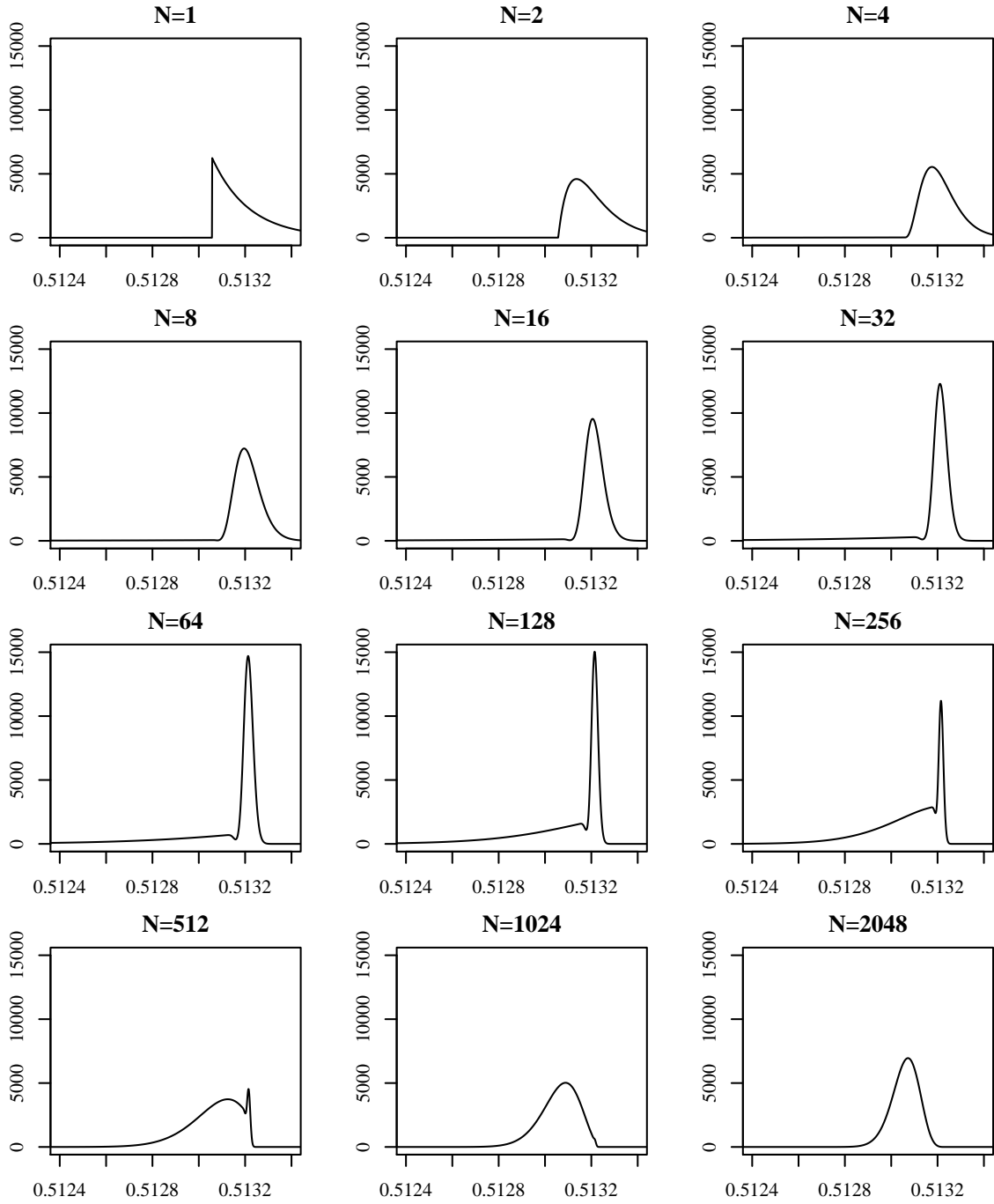
Figure 2.2: Plot of model probability densities for ^{143}Nd/^{144}Nd with F=0.5%, τ_{melt}=1.7 Ga, and N varied as shown. Surprisingly, for N between 2 and 30 the standard deviation actually increases as N is increased. For N large the distribution becomes more symmetric and gaussian in appearance, and standard deviation decreases as $1/\sqrt{N}$ (2.1). N=500 gives a similar distribution shape to that observed (Figure 2.3). Note that $\bar{d}/\bar{d}'$ has been chosen to match the observed mean with N=500.
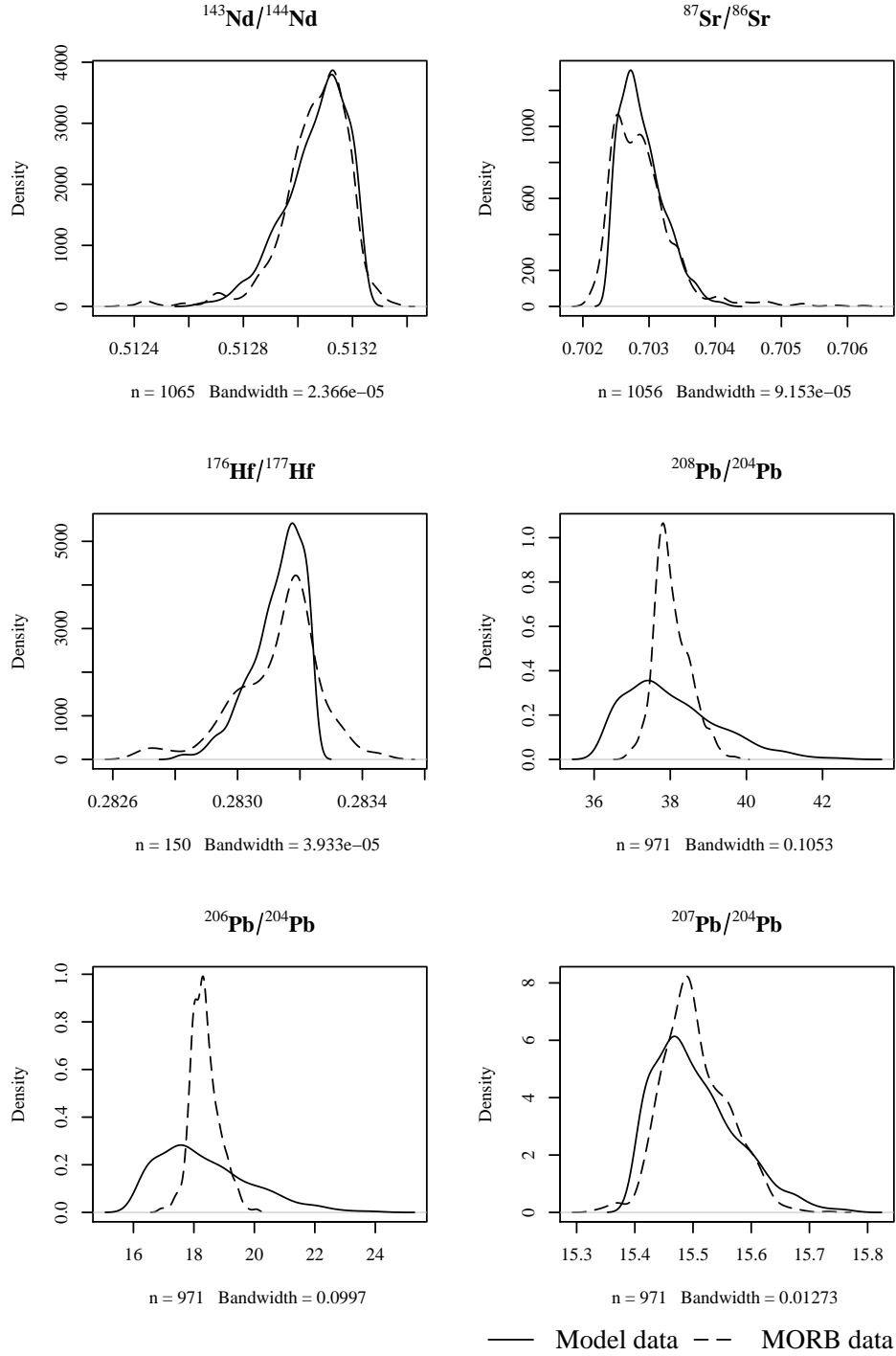
Figure 2.3: Plots of kernel smoothed probability density estimates for model and observed MORB data. n gives the number of measured MORB isotopic ratios. The model curves were calculated from a sample of size 1000 generated by numerical simulation. The bandwidth quoted is the standard deviation of the gaussian kernel smoother applied.
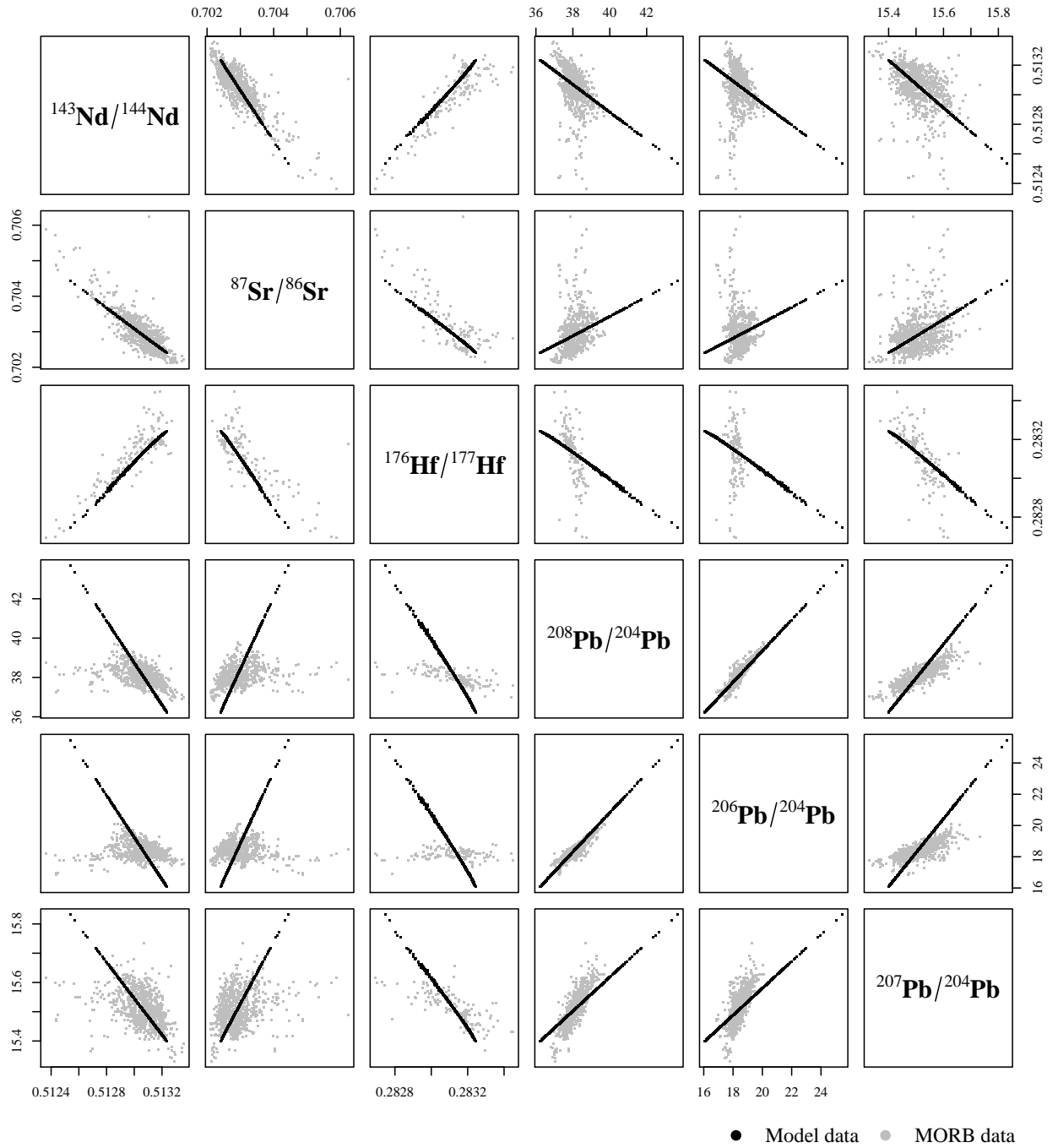
Figure 2.4: Scatter plots of model and observed MORB data.

Table 2.2: Calculated parameters given $F=0.5\%$, $\tau_{\text{melt}}=1.7$ Ga, $N=500$.

		^{147}Sm	^{87}Rb	^{176}Lu	^{232}Th	^{238}U	^{235}U
parent isotope	p	^{147}Sm	^{87}Rb	^{176}Lu	^{232}Th	^{238}U	^{235}U
daughter isotope	d	^{143}Nd	^{87}Sr	^{176}Hf	^{208}Pb	^{206}Pb	^{207}Pb
reference isotope	d'	^{144}Nd	^{86}Sr	^{177}Hf	^{204}Pb	^{204}Pb	^{204}Pb
parent incompatibility	G_p	0.119	1.000	0.017	1.000	1.000	1.000
daughter incompatibility	G_d	0.165	0.155	0.106	0.165	0.165	0.165
relative compatibility	$G_p - G_d$	-0.047	0.845	-0.089	0.835	0.835	0.835
model asymptotic s.d.[a]	σ_{asym}	1.18×10^{-4}	3.43×10^{-4}	8.49×10^{-5}	1.260	1.587	0.073
MORB data s.d.[b]	σ_{MORB}	1.26×10^{-4}	4.93×10^{-4}	1.36×10^{-4}	0.464	0.469	0.056

[a] σ_{asym} is the asymptotic standard deviation of the model calculated using (2.1).
[b] σ_{MORB} is the standard deviation of the isotopic measurements on MORB.

also correlate well with each other, but there is little correlation between the Pb isotopes and the Nd, Sr and Hf isotopes. It is important to note that the model Pb isotopes are more variable than the measured. It implies it is not necessary to include additional mechanisms for producing heterogeneity, such as subducted sediments, to explain the observed variations. Note also that Sm, Nd, Lu and Hf are elements which are little depleted in the mantle by the formation of continental crust. Sr is also only little depleted, but its parent, Rb, is very depleted. Th, U and Pb are all much depleted in the mantle by the formation of continental crust. As has been stated by many previous authors, the dynamics of the Pb systems seem notably different to those of Nd, Sr and Hf.

2.8 Previous work

There has been some recent work done on this problem by other authors. Most recently Meibom and Anderson [2004] introduced the ideas of the statistical upper mantle assemblage, and of sampling under melting and averaging (which they refer to as SUMA). Their ideas form a key part of the model presented here. The only slight criticism of their work is in their assessment of the standard deviation within the framework of the central limit theorem (see section 3.5 for a fuller discussion). Their paper describes the ideas behind SUMA, argues strongly for statistical ways of thinking about these distributions, and discusses in more detail possible interpretations of the mixing parameter N (referred to as M in that paper).

Another detailed examination of this problem was recently given by Kellogg et al. [2002]. Many of the main ideas in their modelling approach are shared with the approach taken here. They too have partial melting as the primary cause of heterogeneity, and have a sampling process which averages over different components in their model mantle reducing heterogeneity. The main difference in the two approaches lies in their complexity. The Kellogg et al. [2002] model analyses the behaviour of multiple reservoirs with fluxes between them, whereas in the present model we look at essentially just one statistical depleted mantle reservoir. Their model has to track subreservoirs within these reservoirs to study the heterogeneity, and further sub-subreservoirs to handle the sampling. As such, in order to implement their model it is necessary to solve numerous differential equations numerically. Furthermore their approach leads to a proliferation of unknown free parameters, and even more so in extensions of their model to fit Pb data [Kellogg, 2004]. Their findings for the Sm-Nd and Rb-Sr systems are broadly similar to our own, and importantly they also found that small melt fractions (also around 0.5%) were essential to

getting a good fit to the data. The advantage of our model is that it is able to reproduce such features using a simpler and analytically tractable framework with just a few free parameters.

Modelling mixing by averaging over N components was used for trace element observations by Slater et al. [2001]. The variability in trace elements in the Slater et al. [2001] model arises from mixing between different fractional melts from a uniform MORB source. They found the standard deviation of their model fractional melts to be a factor of 4 greater than that observed in melt inclusions, which in turn was a factor of 4 greater than that observed in the host lavas. Hence their $N \approx 16^2 \approx 250$, which is of similar magnitude to our model. However, our model differs in using a heterogeneous MORB source, which is necessary when studying isotopic variability rather than elemental.

A number of the important ideas in this chapter can be found in the earlier work of Allègre and Lewin [1995a,b]. The first of these papers [Allègre and Lewin, 1995b] stresses the importance of looking at the distributions of geochemical measurements, and examines in general terms how various distributions can arise from the fundamental processes of mixing and fractionation. In the context of this, the model mantle considered here is "a well-stirred homogeneous mixture of components from heterogeneous sources". There is heterogeneity in the source components because there is melt and residue components of varying ages. These are strongly stirred in the model mantle, and then mixed in the melting before sampling. Importantly, Allègre and Lewin [1995b] stress the role of the central limit theorem and the approach to a normal distribution in such a case. The subsequent paper [Allègre and Lewin, 1995a] models the standard deviation of isotopic ratios in terms of these fundamental processes, and their model will be discussed in detail in Chapter 4. For now, note that this chapter has demonstrated that the behaviour of the standard deviation of ratios can be quite non intuitive and can not be modelled by a simple linear evolution equation as is done in Allègre and Lewin [1995a].

2.9   Conclusions

The minimal model presented in this chapter goes some way to explaining the isotopic heterogeneity seen in MORB. In this version of the model the heterogeneity in the model mantle is produced by a small melt fraction (0.5%) partial melting process, where on average a parcel of mantle material undergoes such melting around every 2 Ga. In particular, the heterogeneity observed in MORB cannot originate in the large melt fraction melting that produces MORB. One possible place for a small melt fraction partial melting process to occur to generate this heterogeneity would be underneath ocean island hot spots.

The simplified model does not fit well with the lead isotopes, but this is unsurprising since lead isotopes are generally inconsistent with simple evolutionary models of the depleted mantle. Any resolutions to the various "lead paradoxes" [Galer and O'Nions, 1985, Murphy et al., 2003] should also be able to explain the variability found in the isotopic measurements. The main message here is one of consistency. With the exception of lead, the heterogeneity seen is consistent with a simple model of mantle processes.

The model presented is minimal, and neglects many processes that have been proposed to account for the isotopic variability of the mantle. But it is hoped it contains the most essential physics. Subsequent chapters will discuss further refinements of this model. In particular, Chapter 4 will discuss the effect of not making the linear decay approximation, not assuming a steady state, and not having a constant melt rate over the Earth's history.

Chapter 3

Statistical box model derivations

This chapter presents the mathematical derivations behind the statistical box model discussed in the previous chapter.

3.1 The basic model

3.1.1 Radioactive decay

We track the molar concentrations of the three isotopes p, d and d'. Consider a fluid parcel that remains outside the melting region. Suppose that initially $p = p_{\text{start}}$, $d = d_{\text{start}}$ and $d' = d'_{\text{start}}$ for such a fluid parcel. Following this fluid parcel the radioactive decay law is given by

$$\frac{\mathrm{D}p}{\mathrm{D}t} = -\lambda p, \qquad \frac{\mathrm{D}d}{\mathrm{D}t} = \lambda p, \qquad \frac{\mathrm{D}d'}{\mathrm{D}t} = 0, \tag{3.1}$$

where the Lagrangian derivative D is to emphasise that we are following the fluid parcel. Thus, after a time t following this fluid parcel we have

$$p_t = p_{\text{start}} \mathrm{e}^{-\lambda t}, \tag{3.2}$$

$$d_t = d_{\text{start}} + p_{\text{start}} \left(1 - \mathrm{e}^{-\lambda t}\right), \tag{3.3}$$

$$d'_t = d'_{\text{start}}, \tag{3.4}$$

and in particular

$$\left(\frac{d}{d'}\right)_t = \left(\frac{d}{d'}\right)_{\text{start}} + \left(1 - \mathrm{e}^{-\lambda t}\right) \left(\frac{p}{d'}\right)_{\text{start}}. \tag{3.5}$$

It is often more convenient to work from the present backwards in time, especially since we are usually much more certain of the present state than the initial one. This perspective will be used often throughout this work. Suppose that at the present $p = p_{\text{finish}}$, $d = d_{\text{finish}}$ and $d' = d'_{\text{finish}}$ for the fluid parcel. Following the fluid parcel backward in time we have

$$\frac{\mathrm{D}p}{\mathrm{D}\tau} = \lambda p, \qquad \frac{\mathrm{D}d}{\mathrm{D}\tau} = -\lambda p, \qquad \frac{\mathrm{D}d'}{\mathrm{D}\tau} = 0. \tag{3.6}$$

Thus, at an age τ before the present we have for the fluid parcel

$$p_\tau = p_{\text{finish}} \mathrm{e}^{\lambda \tau}, \tag{3.7}$$

$$d_\tau = d_{\text{finish}} - p_{\text{finish}} \left(\mathrm{e}^{\lambda \tau} - 1\right), \tag{3.8}$$

$$d'_\tau = d'_{\text{finish}}, \tag{3.9}$$

and

$$\left(\frac{d}{d'}\right)_\tau = \left(\frac{d}{d'}\right)_{\text{finish}} - \left(e^{\lambda\tau} - 1\right)\left(\frac{p}{d'}\right)_{\text{finish}}. \tag{3.10}$$

The mean isotope concentrations in the box satisfy the same radioactive decay law as a individual fluid parcel that remains outside the melting region. We will denote by $\bar{p}$, $\bar{d}$ and $\bar{d}'$ the mean values at the present day, and $\bar{p}_\tau$, $\bar{d}_\tau$ and $\bar{d}'_\tau$ the corresponding mean values at an age τ before the present. As in (3.7-3.9) we have

$$\bar{p}_\tau = \bar{p}e^{\lambda\tau}, \tag{3.11}$$

$$\bar{d}_\tau = \bar{d} - \bar{p}\left(e^{\lambda\tau} - 1\right), \tag{3.12}$$

$$\bar{d}'_\tau = \bar{d}'. \tag{3.13}$$

3.1.2 Melting

At intervals of time Δt a melting event occurs in the melting region. In a melting event we find the mean values of p, d and d' over fluid parcels in the melting region and call these our p_{source}, d_{source} and d'_{source} values. Note that in the strong stirring approximation the mean of the melting region is the same as the mean over the whole box. That is to say at an age τ before the present $p_{\text{source},\tau} = \bar{p}_\tau$, $d_{\text{source},\tau} = \bar{d}_\tau$ and $d'_{\text{source},\tau} = \bar{d}'_\tau$.

The melting region produces a melt fraction F of melt. As such, a fraction F of material in the melting region is set with $p = p_{\text{melt},\tau}$, $d = d_{\text{melt},\tau}$ and $d' = d'_{\text{melt},\tau}$, where these values are calculated from the expressions below. Similarly, a fraction $(1 - F)$ is set to the appropriate residue values. The melting equations are

$$p_{\text{melt},\tau} = \frac{G_p}{F}p_{\text{source},\tau}, \qquad p_{\text{res},\tau} = \frac{1 - G_p}{1 - F}p_{\text{source},\tau}, \tag{3.14}$$

$$d_{\text{melt},\tau} = \frac{G_d}{F}d_{\text{source},\tau}, \qquad d_{\text{res},\tau} = \frac{1 - G_d}{1 - F}d_{\text{source},\tau}, \tag{3.15}$$

$$d'_{\text{melt},\tau} = \frac{G_{d'}}{F}d'_{\text{source},\tau}, \qquad d'_{\text{res},\tau} = \frac{1 - G_{d'}}{1 - F}d'_{\text{source},\tau}, \tag{3.16}$$

where G_p, G_d and $G_{d'}$ come from the melting model used, and are related to the partition coefficients for the relevant isotopes and the melt fraction F. G is the molar fraction of each chemical species that goes into the melt. A good overview of different melt models is given in Zou [1998]. Here we have used the Shaw [1970] model, where

$$G_p = 1 - \left(1 - \frac{P_p}{D_p}F\right)^{1/P_p}, \tag{3.17}$$

with similar definitions for G_d and $G_{d'}$. This is slightly confusing notation - upper case P and D are different weighted averages of partition coefficients for a given isotope, whereas lower case p and d refer to the parent and daughter isotopes respectively.

$$D_p = \sum_i x^i K_p^i, \tag{3.18}$$

$$P_p = \sum_i q^i K_p^i. \tag{3.19}$$

K_p^i is the mineral/ melt partition coefficient for the parent in mineral i, x^i is the mass fraction of mineral i in the source rock, and q^i is the fractional contribution of mineral i to the melt. Note that

$$G_p \to \frac{F}{D_p} \quad \text{as} \quad F \to 0. \tag{3.20}$$

We will assume that isotopes of the same element have the same behaviour under melting (and thus for a given melt fraction the same G), and in particular that $G_d = G_{d'}$.

Suppose a fluid parcel last entered the melt region at an age τ before the present, and at that time was made as melt. Denote its *present day* isotopic concentrations by $p_{\text{melt}}(\tau)$, $d_{\text{melt}}(\tau)$, and $d'_{\text{melt}}(\tau)$. The radioactive decay law (3.2-3.4) gives

$$p_{\text{melt}}(\tau) = p_{\text{melt},\tau}\text{e}^{-\lambda\tau}, \tag{3.21}$$

$$d_{\text{melt}}(\tau) = d_{\text{melt},\tau} + \left(1 - \text{e}^{-\lambda\tau}\right)p_{\text{melt},\tau}, \tag{3.22}$$

$$d'_{\text{melt}}(\tau) = d'_{\text{melt},\tau}. \tag{3.23}$$

Similarly, for a fluid parcel that was made instead as residue at an age τ

$$p_{\text{res}}(\tau) = p_{\text{res},\tau}\text{e}^{-\lambda\tau}, \tag{3.24}$$

$$d_{\text{res}}(\tau) = d_{\text{res},\tau} + \left(1 - \text{e}^{-\lambda\tau}\right)p_{\text{res},\tau}, \tag{3.25}$$

$$d'_{\text{res}}(\tau) = d'_{\text{res},\tau}. \tag{3.26}$$

Substituting the melting equations (3.14-3.16) yields

$$p_{\text{melt}}(\tau) = \frac{G_p}{F}\bar{p}, \tag{3.27}$$

$$d_{\text{melt}}(\tau) = \frac{G_d}{F}\bar{d} + \frac{G_p - G_d}{F}\left(\text{e}^{\lambda\tau} - 1\right)\bar{p}, \tag{3.28}$$

$$d'_{\text{melt}}(\tau) = \frac{G_d}{F}\bar{d'}. \tag{3.29}$$

$$p_{\text{res}}(\tau) = \frac{1 - G_p}{1 - F}\bar{p}, \tag{3.30}$$

$$d_{\text{res}}(\tau) = \frac{1 - G_d}{1 - F}\bar{d} - \frac{G_p - G_d}{1 - F}\left(\text{e}^{\lambda\tau} - 1\right)\bar{p}, \tag{3.31}$$

$$d'_{\text{res}}(\tau) = \frac{1 - G_d}{1 - F}\bar{d'}. \tag{3.32}$$

For slowly decaying isotopes $\lambda\tau \ll 1$, the radioactive decay law can be linearised ($\text{e}^{\lambda\tau} - 1 \approx \lambda\tau$). In this case (3.28) and (3.31) can be written as

$$d_{\text{melt}}(\tau) = \frac{G_d}{F}\bar{d} + \frac{G_p - G_d}{F}\lambda\tau\bar{p}, \tag{3.33}$$

$$d_{\text{res}}(\tau) = \frac{1 - G_d}{1 - F}\bar{d} - \frac{G_p - G_d}{1 - F}\lambda\tau\bar{p}. \tag{3.34}$$

For simplicity, this linear approximation will be used for the remainder of this chapter. In the following chapter the consequences of not making this approximation will be examined in detail.

3.1.3 Distribution of ages

Let the melting region have non-dimensional area A relative to the size of the box. Define the iterate age s to be the number of iterates since a given parcel last entered the melting region. Let $\hat{S}$ be a discrete random variable giving the distribution of iterate ages. If the advection is strongly stirring, then the probability of any given parcel entering the melting

region at that timestep is A. Considering a long time statistical steady state, $\hat{S}$ is given by a geometric distribution, where the probability of an iterate age of s is

$$\mathbb{P}(\hat{S} = s) = A(1 - A)^s, \qquad s = 0, 1, 2... \qquad (3.35)$$

Note that it is thus assumed that all parcels have passed through the melting region at some point. It is easier to work with a continuous distribution rather than a discrete distribution. The continuous analogue of the geometric distribution is the exponential distribution. The age τ is the time since a given parcel last entered the melting region and so $\tau = s\Delta t$, where Δt is the time step in the discrete model. Note that

$$A(1 - A)^s = A \exp\left(-\tau \frac{-\log(1 - A)}{\Delta t}\right), \qquad (3.36)$$

and so the discrete distribution has an associated melting timescale given by

$$\tau_{\text{melt}} = \frac{\Delta t}{-\log(1 - A)}. \qquad (3.37)$$

Taking the limit as $\Delta t \to 0$, $A \to 0$ with τ_{melt} fixed allows us to relate the discrete and continuous distributions. Let $\hat{T}$ be the continuous random variable giving the distribution of ages. Then $\hat{T}$ is distributed exponentially with mean τ_{melt}, and melting is thus being modelled as a Poisson process [Albarède, 2005, 2003]. Let $\hat{E}$ be an exponential random variable with mean 1. Then we can write $\hat{T} = \tau_{\text{melt}}\hat{E}$. $\hat{E}$ has probability density function given by

$$f_{\hat{E}}(x) = e^{-x}, \qquad x > 0. \qquad (3.38)$$

In this limit both the time interval Δt and the melting region area A become infinitesimal. As such the assumption that strong stirring occurs between time intervals cannot be made. However, the continuous limit should be a reasonable approximation provided the timescale for stirring is short compared to the melting timescale τ_{melt}, and for calculational convenience it is used from now on.

3.1.4 Distribution of d/d'

Hence, in terms of the random variable $\hat{T}$ giving the distribution of ages we have the following for the distribution of d/d' in the interior of the box (using (3.29), (3.32), (3.33) and (3.34))

$$\frac{\hat{d}}{\hat{d'}} - \frac{\bar{d}}{\bar{d'}} \sim \lambda\left(G_p - G_d\right)\frac{\bar{p}}{\bar{d'}} \begin{cases} \hat{T}/G_d & \text{(melt)} \\ -\hat{T}/(1 - G_d) & \text{(res)} \end{cases}, \qquad (3.39)$$

where we have a probability F of getting a melt value, and $(1 - F)$ of getting a residue value. Here hats on $\hat{d}$ and $\hat{d'}$ are used to denote that they are random variables. (3.39) can also be expressed in terms of the exponential random variable $\hat{E}$ as

$$\frac{\hat{d}}{\hat{d'}} - \frac{\bar{d}}{\bar{d'}} \sim l \begin{cases} \hat{E}/G_d & \text{(melt)} \\ -\hat{E}/(1 - G_d) & \text{(res)} \end{cases}, \qquad (3.40)$$

where

$$l = \lambda\tau_{\text{melt}}\left(G_p - G_d\right)\frac{\bar{p}}{\bar{d'}}. \qquad (3.41)$$

3.2   Sampling

When we sample the mantle we do not sample directly from the distribution that has just been derived. In fact we sample after numerous components from this distribution have melted and mixed together. This is the "sampling under melting and averaging" (SUMA) concept [Meibom and Anderson, 2004]. We will assume there is no bias in melting and mixing - the likelihood of a component entering the melt is proportional to its abundance in the source.

Suppose we mix together N independent identically distributed (i.i.d.) random samples from the whole box distribution. Note in particular that the d/d' ratio of the mixture is *not* the mean of the ratios of the independent samples. Instead the individual concentrations d and d' are averaged and the ratio is then calculated i.e. the quantity of interest is

$$\hat{Z}_d = \sum_{i=1}^{N} \hat{d}_i / \sum_{i=1}^{N} \hat{d}'_i. \tag{3.42}$$

Calculations are made easier if we introduce new starred variables

$$d^\star = \frac{1}{\bar{d}'} \left(d - \frac{\bar{d}}{\bar{d}'} d' \right), \tag{3.43}$$

$$d'^\star = \frac{1}{\bar{d}'} \left(d' - \bar{d}' \right), \tag{3.44}$$

and note the important identity

$$\hat{Z}_d - \frac{\bar{d}}{\bar{d}'} = \frac{\sum_i \hat{d}_i^\star}{\sum_i (1 + \hat{d}_i'^\star)}. \tag{3.45}$$

In terms of the starred variables, the components as a function of age τ are

$$d_{\mathrm{melt}}^\star(\tau) = \frac{\bar{p}}{\bar{d}'} \frac{G_p - G_d}{F} \lambda \tau, \tag{3.46}$$

$$d_{\mathrm{melt}}'^\star(\tau) = \frac{G_d}{F} - 1, \tag{3.47}$$

$$d_{\mathrm{res}}^\star(\tau) = -\frac{\bar{p}}{\bar{d}'} \frac{G_p - G_d}{1 - F} \lambda \tau, \tag{3.48}$$

$$d_{\mathrm{res}}'^\star(\tau) = \frac{1 - G_d}{1 - F} - 1. \tag{3.49}$$

The $\hat{d}_i^\star$'s are thus i.i.d. random variables with distribution

$$\hat{d}_i^\star = l \begin{cases} \hat{E}_i / F & \text{(melt)} \\ -\hat{E}_i / (1 - F) & \text{(res)} \end{cases}, \tag{3.50}$$

where the $\hat{E}_i$'s are i.i.d. exponential random variables with mean 1. Again, there is a probability F of getting a melt value, and $1 - F$ of getting a residue value. $\sum_i (1 + \hat{d}_i'^\star)$ depends only on how many melt and how many residue components are mixed together. Suppose we have r melt components and $N - r$ residue components. Then $\sum_i (1 + \hat{d}_i'^\star) = M(r)$, where

$$M(r) = r \frac{G_d}{F} + (N - r) \frac{1 - G_d}{1 - F}. \tag{3.51}$$

Assuming the sampling is unbiased, the probability q_r of getting r melt components and $N - r$ residue components in the mix is given by the binomial distribution

$$q_r = \binom{N}{r} F^r (1 - F)^{N-r}. \tag{3.52}$$

Hence $\hat{Z}_d$ has distribution given by

$$\hat{Z}_d - \frac{\bar{d}}{\bar{d'}} = \hat{Y}_r \qquad \text{with probability } q_r, \tag{3.53}$$

where $\hat{Y}_r$ is the result of combining r melt and $N - r$ residue components, with distribution given in terms of the $\hat{E}_i$'s as

$$\hat{Y}_r = \frac{l}{M(r)} \left(\sum_{i=1}^r \frac{\hat{E}_i}{F} - \sum_{i=r+1}^N \frac{\hat{E}_i}{1 - F} \right). \tag{3.54}$$

The above expression forms the basis of one method of estimating the probability density function for d/d' after sampling. Exponential random variables can be simulated easily, and so an artificial sample from this distribution can be constructed. From this it is possible to numerically estimate the density (e.g. by kernel smoothing, section 3.4).

3.2.1 Characteristic functions

In order to get an analytic expression for the probability density function of $\hat{Z}_d$ it is useful to use characteristic functions. Suppose a random variable $\hat{X}$ has probability density function $f(x)$. Then its characteristic function $\phi(k)$ is defined by its Fourier transform

$$\phi(k) = \int_{-\infty}^{\infty} e^{-ikx} f(x) \, dx. \tag{3.55}$$

The moments of the distribution can be read off from the Taylor expansion about $k = 0$ of the characteristic function. If the nth moment is

$$\alpha_n = \mathbb{E}\hat{X}^n = \int_{-\infty}^{\infty} x^n f(x) \, dx, \tag{3.56}$$

where $\mathbb{E}$ denotes expectation, we have

$$\phi(k) = \sum_{n=0}^{\infty} \alpha_n \frac{(-ik)^n}{n!}. \tag{3.57}$$

Similarly, the cumulants of the distribution can be read off from the Taylor expansion about $k = 0$ of the log of the characteristic function. If the nth cumulant is κ_n we have

$$\log \phi(k) = \sum_{n=0}^{\infty} \kappa_n \frac{(-ik)^n}{n!}. \tag{3.58}$$

There are two important properties that we wish to exploit. If $\hat{X}$ is a random variable with characteristic function $\phi(k)$ then $a\hat{X} + b$ has characteristic function $e^{-ikb}\phi(ak)$. Also, if $\hat{X}_1, \hat{X}_2, ... \hat{X}_n$ are independently distributed with characteristic functions $\phi_1(k), \phi_2(k), ... \phi_n(k)$, then $\sum_{i=1}^n \hat{X}_i$ has characteristic function $\prod_{i=1}^n \phi_i(k)$.

The exponential random variable $\hat{E}$ has characteristic function

$$\phi_{\hat{E}}(k) = (1 + ik)^{-1}. \tag{3.59}$$

Hence, using the two properties stated above, $\hat{Y}_r$ has characteristic function

$$\phi_{\hat{Y}_r}(k) = \left(1 + \frac{ikl}{M(r)F}\right)^{-r}\left(1 - \frac{ikl}{M(r)(1 - F)}\right)^{-(N-r)}. \tag{3.60}$$

Hence $\hat{Z}_d$ has characteristic function

$$\phi_{\hat{Z}_d}(k) = \mathrm{e}^{-ik\bar{d}/\bar{d}'}\sum_{r=0}^{N}\binom{N}{r}F^r(1 - F)^{N-r}\phi_{\hat{Y}_r}(k). \tag{3.61}$$

By Taylor expanding this expression about $k = 0$ the moments and cumulants of the distribution can be found. Furthermore, by using the Fourier inversion theorem we may recover the probability density function of $\hat{Z}_d$,

$$f_{\hat{Z}_d}(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\mathrm{e}^{ikx}\phi_{\hat{Z}_d}(k)\,\mathrm{d}k. \tag{3.62}$$

This is another way of constructing the density of d/d' after sampling, and gives the precise density function. Doing this Fourier inversion analytically leads to complicated sums of a large number of terms. These are very time consuming to calculate, and difficult to interpret. However, this integral can be done efficiently numerically using a Fast Fourier Transform (FFT), and this method was used to produce Figure 2.2. Moreover, simpler asymptotic results for large N can be obtained that are easier to interpret and much quicker to evaluate.

3.3 Asymptotic results for large N

3.3.1 The asymptotics of averaging ratio quantities

This section describes some important general results on the asymptotics of averaging ratio quantities. Consider i.i.d. pairs of random variables $\{\hat{x}_i, \hat{y}_i\}$, $i = 1, 2, ...N$. Suppose $\hat{y}_i > 0$. We are interested in the asymptotic behaviour of the ratio of sums

$$\hat{Z} = \sum_i \hat{x}_i / \sum_i \hat{y}_i \tag{3.63}$$

for N large. Let

$$\hat{x}^\star = \frac{1}{\bar{y}}\left(\hat{x} - \frac{\bar{x}}{\bar{y}}\hat{y}\right), \tag{3.64}$$

$$\hat{y}^\star = \frac{1}{\bar{y}}\left(\hat{y} - \bar{y}\right), \tag{3.65}$$

where $\bar{x} = \mathbb{E}(\hat{x})$, $\bar{y} = \mathbb{E}(\hat{y})$. Note that $\mathbb{E}(\hat{x}^\star) = \mathbb{E}(\hat{y}^\star) = 0$. Note also that

$$\hat{Z} = \frac{\sum \hat{x}_i}{\sum \hat{y}_i} = \frac{\bar{x}}{\bar{y}} + \frac{\sum \hat{x}_i^\star/N}{1 + \sum \hat{y}_i^\star/N}. \tag{3.66}$$

It can be shown [Novak, 1996] that $\hat{Z}$ is asymptotically normal for large N under appropriate assumptions (assumptions which give the central limit theorem for $\sum \hat{x}_i^\star$ and give

the law of large numbers for $\sum \hat{y}_i$). The condition $\hat{y}_i > 0$ ensures that moments of $\hat{Z}$ are well defined, and Cauchy distribution problems do not arise. By Taylor expanding (3.66) and taking expectations the following expressions for the asymptotic central moments can be derived:

$$\mu = \bar{Z} = \mathbb{E}(\hat{Z}) = \frac{\bar{x}}{\bar{y}} - \frac{1}{N}\mathbb{E}\left(\hat{x}^\star \hat{y}^\star\right) + O\left(\frac{1}{N^2}\right), \tag{3.67}$$

$$\mu_2 = \sigma^2 = \mathbb{E}\left(\hat{Z} - \bar{Z}\right)^2 = \frac{1}{N}\mathbb{E}\left(\hat{x}^{\star 2}\right) + O\left(\frac{1}{N^2}\right), \tag{3.68}$$

$$\mu_3 = \mathbb{E}\left(\hat{Z} - \bar{Z}\right)^3 = \frac{1}{N^2}\left(\mathbb{E}\left(\hat{x}^{\star 3}\right) - 6\mathbb{E}\left(\hat{x}^\star \hat{y}^\star\right)\mathbb{E}\left(\hat{x}^{\star 2}\right)\right) + O\left(\frac{1}{N^3}\right). \tag{3.69}$$

Hence the skew parameter γ_1 is

$$\gamma_1 = \frac{\mu_3}{(\mu_2)^{3/2}} = \frac{\mathbb{E}\left(\hat{x}^{\star 3}\right) - 6\mathbb{E}\left(\hat{x}^\star \hat{y}^\star\right)\mathbb{E}\left(\hat{x}^{\star 2}\right)}{N^{1/2}\left(\mathbb{E}\left(x^{\star 2}\right)\right)^{3/2}} + O\left(\frac{1}{N^{3/2}}\right). \tag{3.70}$$

The kurtosis and higher order moments can be derived similarly by expanding to higher orders.

We are often concerned with plots of one ratio against another. So consider two sets of i.i.d. pairs of random variables $\{\hat{x}_i, \hat{y}_i\}_1$ and $\{\hat{x}_i, \hat{y}_i\}_2$, $i = 1, 2, ...N$. Then the covariance of $\hat{Z}_1$ and $\hat{Z}_2$ is given by

$$\text{cov}\left(\hat{Z}_1, \hat{Z}_2\right) = \mathbb{E}\left(\left(\hat{Z}_1 - \bar{Z}_1\right)\left(\hat{Z}_2 - \bar{Z}_2\right)\right) = \frac{1}{N}\mathbb{E}\left(\hat{x}_1^\star \hat{x}_2^\star\right) + O\left(\frac{1}{N^2}\right)$$

$$= \frac{1}{N}\text{cov}\left(\hat{x}_1^\star, \hat{x}_2^\star\right) + O\left(\frac{1}{N^2}\right), \tag{3.71}$$

with corresponding correlation

$$r_{12} = \text{cor}\left(\hat{Z}_1, \hat{Z}_2\right) = \frac{\text{cov}\left(\hat{Z}_1, \hat{Z}_2\right)}{\sigma_1 \sigma_2} = \frac{\mathbb{E}\left(\hat{x}_1^\star \hat{x}_2^\star\right)}{\sqrt{\mathbb{E}\left(\hat{x}_1^{\star 2}\right)\mathbb{E}\left(\hat{x}_2^{\star 2}\right)}} + O\left(\frac{1}{N}\right)$$

$$= \text{cor}\left(\hat{x}_1^\star, \hat{x}_2^\star\right) + O\left(\frac{1}{N}\right). \tag{3.72}$$

We are particularly interested in calculating the slope of a regression line of one system against another. There are many different methods for fitting regression lines to a cloud of data points which make various assumptions about the underlying data. Three commonly used estimates are

$$\beta_{12} = r_{12}\frac{\sigma_2}{\sigma_1} = \frac{\mathbb{E}(\hat{x}_1^\star \hat{x}_2^\star)}{\mathbb{E}(\hat{x}_1^{\star 2})} + O\left(\frac{1}{N}\right), \tag{3.73}$$

$$\bar{\beta} = \frac{\sigma_2}{\sigma_1} = \sqrt{\frac{\mathbb{E}(\hat{x}_2^{\star 2})}{\mathbb{E}(\hat{x}_1^{\star 2})}} + O\left(\frac{1}{N}\right), \tag{3.74}$$

$$\beta_{21} = \frac{1}{r_{12}}\frac{\sigma_2}{\sigma_1} = \frac{\mathbb{E}(\hat{x}_2^{\star 2})}{\mathbb{E}(\hat{x}_1^\star \hat{x}_2^\star)} + O\left(\frac{1}{N}\right), \tag{3.75}$$

where β_{12} is the slope of the linear least squares regression line of system 2 on system 1, and β_{21} is the same line but for system 1 on system 2. $\bar{\beta}$ is the slope of the geometric mean regression line, and its sign is chosen to be the same as the correlation r_{12}. Note that $|\beta_{12}| \leq |\bar{\beta}| \leq |\beta_{21}|$. If the correlation is good all three estimates will be similar.

3.3.2 Asymptotic moments

Using the general expressions of the previous section the asymptotic moments of $\hat{Z}_d$ can be calculated. For example, the leading order variance can be calculated using (3.50) and (3.68) as

$$
\begin{aligned}
\mathrm{var}\hat{Z}_d = \mathbb{E}\left(\hat{Z}_d - \bar{Z}_d\right)^2 &\sim \frac{1}{N}\mathbb{E}\left(\hat{d}^{\star 2}\right) \\
&= \frac{1}{N}\left(F\mathbb{E}\left(\frac{l\hat{E}}{F}\right)^2 + (1-F)\mathbb{E}\left(-\frac{l\hat{E}}{1-F}\right)^2\right) \\
&= \frac{l^2\mathbb{E}\left(\hat{E}^2\right)}{NF(1-F)} \\
&= \frac{2l^2}{NF(1-F)}.
\end{aligned}
\tag{3.76}
$$

Performing a similar calculation for the other moments yields

$$
\bar{Z}_d = \mathbb{E}\hat{Z}_d = \frac{\bar{d}}{\bar{d'}} + \frac{l(F-G_d)}{NF(1-F)} + O\left(\frac{1}{N^2}\right),
\tag{3.77}
$$

$$
\mathrm{var}\hat{Z}_d = \mathbb{E}\left(\hat{Z}_d - \bar{Z}_d\right)^2 = \frac{2l^2}{NF(1-F)} + O\left(\frac{1}{N^2}\right),
\tag{3.78}
$$

$$
\mathbb{E}\left(\hat{Z}_d - \bar{Z}_d\right)^3 = \frac{6l^3\left(1-2G_d\right)}{(NF(1-F))^2} + O\left(\frac{1}{N^3}\right).
\tag{3.79}
$$

In particular, note that the skew parameter γ_1 is (using (3.70))

$$
\gamma_1 = \mathrm{sgn}(l)\frac{3\left(1-2G_d\right)}{(2NF(1-F))^{1/2}} + O\left(\frac{1}{N^{3/2}}\right).
\tag{3.80}
$$

Most importantly from all this we see that the asymptotic standard deviation $\sigma = \sqrt{\mathrm{var}\hat{Z}_d}$ of this distribution is

$$
\sigma \sim \lambda\tau_{\mathrm{melt}}|G_p - G_d|\frac{\bar{p}}{\bar{d'}}\sqrt{\frac{2}{NF(1-F)}}.
\tag{3.81}
$$

Less important is the small $O(1/N)$ correction to the mean which is not significant compared with the $O(1/N^{1/2})$ standard deviation.

3.3.3 Asymptotic density: Edgeworth expansion

An asymptotic expansion of the density for large N can obtained by using an Edgeworth expansion (see Blinnikov and Moessner [1998] for a fuller discussion of these expansions). Let $\hat{Q} = (\hat{Z} - \bar{Z})/\sigma$. Then the first two terms in the Edgeworth expansion for the density of $\hat{Q}$ are

$$
f_{\hat{Q}}(x) = \frac{\mathrm{e}^{-x^2/2}}{\sqrt{2\pi}}\left\{1 + \sigma\frac{S_3}{6}\mathrm{He}_3(x) + \sigma^2\left(\frac{S_4}{24}\mathrm{He}_4(x) + \frac{S_3^2}{72}\mathrm{He}_6(x)\right) + O\left(\sigma^3\right)\right\},
\tag{3.82}
$$

where $S_n = \kappa_n/\sigma^{2n-2}$ and κ_n is the nth cumulant (defined by (3.58)). $\text{He}_n(x)$ is the nth Chebyshev-Hermite polynomial, the relevant ones being

$$\text{He}_3(x) = x^3 - 3x, \tag{3.83}$$

$$\text{He}_4(x) = x^4 - 6x^2 + 3, \tag{3.84}$$

$$\text{He}_6(x) = x^6 - 15x^4 + 45x^2 - 15. \tag{3.85}$$

We have

$$\sigma = |l|\sqrt{\frac{2}{NF(1-F)}} + O\left(\frac{1}{N^{3/2}}\right). \tag{3.86}$$

Using the results of the previous section we can calculate the cumulants of the distribution in powers of σ. Thus

$$S_3 = \frac{\kappa_3}{\sigma^4} = \frac{3}{2l}(1 - 2G_d) + O\left(\sigma^2\right), \tag{3.87}$$

$$S_4 = \frac{\kappa_4}{\sigma^6} = \frac{3}{2l^2}\left(14G_d^2 + F^2 - 8FG_d - 10G_d + 3F + 2\right) + O\left(\sigma^2\right). \tag{3.88}$$

Substituting these two expressions back in to (3.82) gives an expansion with errors of $O\left(\sigma^3\right)$. Substituting back the leading order expression for σ (3.86) gives an expansion in powers of N with errors of $O\left(1/N^{3/2}\right)$. As an example, working to first order in σ (as the next term is more cumbersome) we get

$$f_{\hat{Q}}(x) = \frac{e^{-x^2/2}}{\sqrt{2\pi}}\left\{1 + \frac{(1-2G_d)\,\text{sgn}(l)}{2\sqrt{2NF(1-F)}}\text{He}_3(x) + O\left(\frac{1}{N}\right)\right\}. \tag{3.89}$$

The above Edgeworth expansion is a third way of estimating the distribution of d/d' after sampling. Its main strength is that it is the quickest to evaluate. It is an approximation to the true density and can contain features that the true density would not. For example, it is possible for an Edgeworth expansion of a density to become negative, although it still always integrates to 1.

3.3.4   Isotope ratio - isotope ratio plots

Consider two isotopic systems 1 and 2 with potentially different parent and daughter elements. Suppose we plot $(d/d')_2$ against $(d/d')_1$. For large N the correlation r_{12} between the two ratios is given by (3.72)

$$r_{12} = \frac{\mathbb{E}\left(\hat{d}_1^\star \hat{d}_2^\star\right)}{\sqrt{\mathbb{E}\left(\hat{d}_1^{\star 2}\right)\mathbb{E}\left(\hat{d}_2^{\star 2}\right)}} + O\left(\frac{1}{N}\right)$$

$$= \text{sgn}\left(l_1 l_2\right) + O\left(\frac{1}{N}\right)$$

$$= \text{sgn}\left((G_{p1} - G_{d1})(G_{p2} - G_{d2})\right) + O\left(\frac{1}{N}\right). \tag{3.90}$$

Hence as $N \to \infty$ the correlation is perfect ($|r_{12}| \sim 1$) with sign depending on the relative compatibilities of parents and daughters. The slope $\bar{\beta}$ of the geometric mean regression

line is given by the ratio of the standard deviations, where the sign is chosen appropriately to match r_{12}. Using (3.74) we get

$$
\begin{aligned}
\bar{\beta} &= \sqrt{\frac{\mathbb{E}d_2^{\star 2}}{\mathbb{E}d_1^{\star 2}}} + O\left(\frac{1}{N}\right) \\
&= \frac{l_2}{l_1} + O\left(\frac{1}{N}\right) \\
&= \frac{\lambda_2\left(G_{p2} - G_{d2}\right)\left(\bar{p}/\bar{d}'\right)_2}{\lambda_1\left(G_{p1} - G_{d1}\right)\left(\bar{p}/\bar{d}'\right)_1} + O\left(\frac{1}{N}\right).
\end{aligned}
\tag{3.91}
$$

As $N \to \infty$ all samples will plot on a single line with this slope.

Note a particular special case - if for both systems the daughters are isotopes of the same element (for example, both isotopes of lead), then $M_1(r) = M_2(r)$ and $\hat{Y}_{r2}/\hat{Y}_{r1} = l_2/l_1$, and thus all samples lie on a single straight line of gradient exactly $\beta = l_2/l_1$, regardless of N. In this case it is usual for the reference isotopes to be the same $d_1' = d_2'$. Furthermore, if it also the case that the parent elements are the same (or at least behave similarly under melting $G_{p1} \approx G_{p2}$) then this gradient is independent of the melt model and is simply

$$
\beta = \frac{\lambda_2 \bar{p}_2}{\lambda_1 \bar{p}_1}.
\tag{3.92}
$$

3.4 Comparison of methods for density estimation

We have outlined in the previous sections (3.2, 3.2.1, and 3.3.3) three methods of estimating the density of d/d' in the model. The first method, numerical simulation, generates a sample from the model distribution by using computer generated samples from an exponential distribution. The probability density is then estimated using a kernel smoothed estimate for this sample. Given n independent observations $x_1, x_2, ..., x_n$ from the random variable $\hat{X}$, the kernel density estimator $\tilde{f}_h(x)$ of the density value $f(x)$ is defined as

$$
\tilde{f}_h(x) = \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{x_i - x}{h}\right),
\tag{3.93}
$$

where h is the bandwidth and $K(u)$ the kernel function. For a gaussian kernel

$$
K(u) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}u^2\right).
\tag{3.94}
$$

The function `density` in R (R Development Core Team [2006]) was used to calculate this efficiently. One advantage of using kernel smoothed estimates, over plotting a histogram for example, is that it is easy to plot more than one density on the same plot. In this work, the default gaussian kernel was used, with optimal bandwidth as chosen by R. Kernel smoothed estimates are sensitive to choices of kernel and bandwidth, see Sheather and Jones [1991], Silverman [1986], and the R documentation for a fuller discussion.

The second method involves calculating the full analytic distribution by numerically doing the Fourier inversion of (3.62). The advantage of this method is that you get the density precisely.

The third method is to use the asymptotic approximation provided by the Edgeworth expansion of 3.89. This is an approximation to the true density, but is the quickest of the three methods to calculate, and becomes more accurate as N is increased.
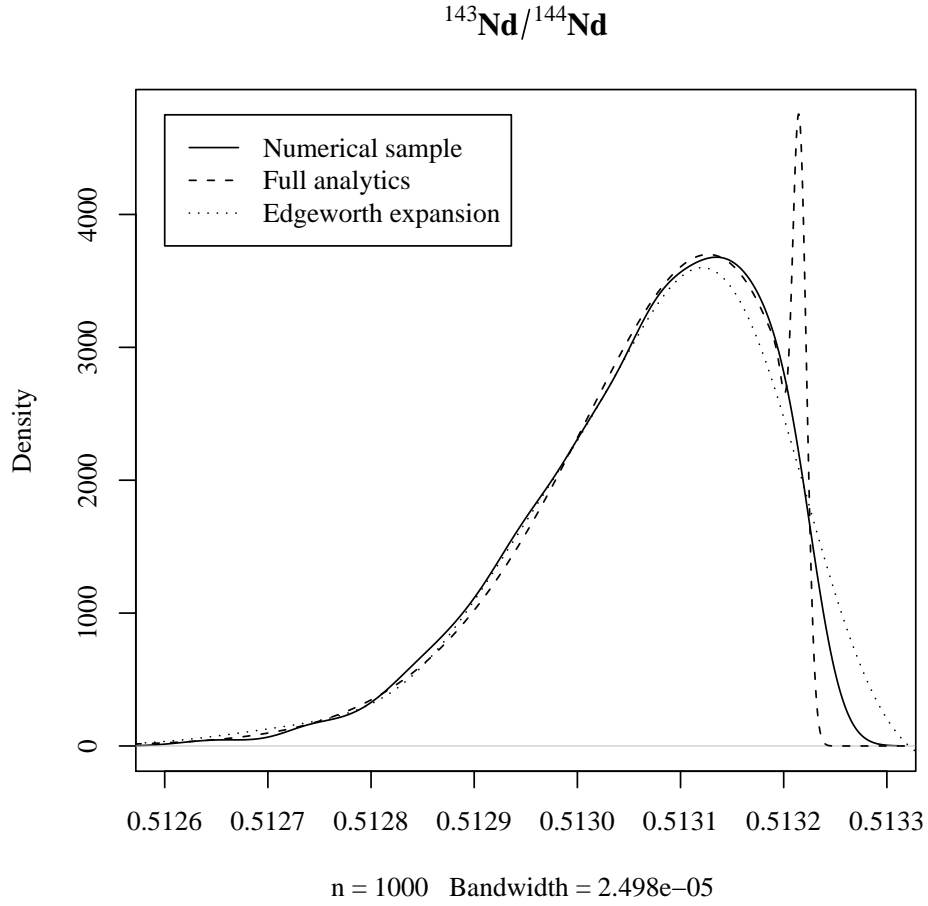
Figure 3.1: Plot comparing the three methods of density estimation in the model for ^{143}Nd/^{144}Nd, with F=0.5%, τ_{melt}=1.7 Ga, and N=500. For the numerical sample, the sample size n is shown along with the bandwidth of the kernel smoother. The Edgeworth expansion is to second order (up to and including terms of order σ^2).

Figure 3.1 uses these three methods to produce model densities for ^{143}Nd/^{144}Nd. All three methods produce very similar shapes, clearly showing similar standard deviations and skews. The analytic density contains a small scale spike at 0.51322. The area under this spike is small compared with 1. The spike is not seen in the kernel smoothed estimate because it is a feature of smaller scale than the bandwidth. Similarly it is not seen in the Edgeworth expansion because there are not enough terms in the expansion to express such a feature. Our interest in this work is in broad features of the distributions, and such small features are liable to be disturbed by other factors (for example, measurement error) and so this spike is not an important feature. Note also that the Edgeworth expansion becomes negative on the far right of the plot, which is a disadvantage in using such an expansion far from the mean.

When demonstrating the behaviour of the model distribution with varying N in Figure 2.2 numerical Fourier inversion was used for greatest accuracy. However, when comparing the model with measured data in Figure 2.3 we have opted to use kernel smoothed densities to provide a more like-for-like comparison between the model and measured data. When generating a model sample, a sample size of 1000 was used to provide a close comparison with the measured data used which is of a similar size.

3.5   Remarks on the central limit theorem

As has been discussed in section 3.2, when we sample from our model mantle we do so after averaging over a number of components. For a set of independent identically distributed random variables $\hat{X}_1, \hat{X}_2, ..., \hat{X}_N$ with mean μ_0 and standard deviation σ_0, the central limit theorem states that the distribution of $\sum \hat{X}_i/N$ will approach a normal distribution with mean μ_0 and standard deviation $\sigma_0/\sqrt{N}$. However, this cannot directly be applied in mixing isotopic ratios since it is the individual concentrations of d and d' that are averaged and then the ratio taken, i.e. $\sum \hat{d}_i / \sum \hat{d}'_i \neq \sum (\hat{d}_i/\hat{d}'_i)/N$. As always in geochemistry, great care must be taken in dealing with ratios of quantities. However, it does turn out that in our model we do approach a normal distribution, but the standard deviation is quite different from a naive use of the central limit theorem. As an example, using Meibom and Anderson's notation, the standard deviation σ_{um} of the d/d' distribution before averaging is

$$\sigma_{\text{um}} = |l| \frac{\sqrt{(F - G_d)^2 + 2F(1 - F)}}{G_d(1 - G_d)} \tag{3.95}$$

whereas the standard deviation σ_{meas} we would measure after the averaging is

$$\sigma_{\text{meas}} = |l| \sqrt{\frac{2}{NF(1 - F)}} + O\left(\frac{1}{N^{3/2}}\right) \tag{3.96}$$

and in particular it is clear that

$$\sigma_{\text{meas}} \neq \frac{\sigma_{\text{um}}}{\sqrt{N}} \tag{3.97}$$

In fact, for the model values for Nd, there is a factor of 15 difference between the left and right hand sides of the above expression when N is large. It is also important to note that the approach to a normal distribution is governed by $NF(1 - F)$ and not N, as indicated by the expression for the skew of the distribution (3.80) and the first correction to the normal distribution in the Edgeworth expansion (3.82). Since F is small, the distributions we observe are significantly skewed despite N being large.

Chapter 4

Mantle pseudo-isochrons

4.1 Introduction

When ^{207}Pb/^{204}Pb is plotted against ^{206}Pb/^{204}Pb for data from mid-ocean ridge basalt (MORB) or ocean island basalt (OIB) an approximate linear relationship is found (Figure 4.1). The slope of a regression line through these data points can be used to an infer an age by treating the regression line as if it were an isochron [Brooks et al., 1976, Tatsumoto, 1978]. Formally an isochron age dates a single fractionation event, which is not the case for MORB and OIB; the isotopic systematics of these basalts result from multiple fractionations due to repeated melting and recycling over the course of the Earth's history. As such the ages calculated by the isochron method are often referred to as pseudo-isochron ages, and the aim of this chapter is to relate the pseudo-isochron ages to real physical parameters.

Recall that the isotopic systems we are studying consist of a parent isotope p which decays to a daughter isotope d with decay constant λ. There is a reference isotope d' with respect to which these isotopes are measured. The reference isotope is of the same element as the daughter d, but neither decays nor is a decay product. There are two particular isochrons we will focus on, and we will refer to these as the parent-daughter isochron and the daughter-daughter isochron.

The parent-daughter isochron involves plotting d/d' against p/d'. The parent-daughter isochron age τ_{pdi} is related to the slope β of the regression line by

$$\mathrm{e}^{\lambda \tau_{\mathrm{pdi}}} - 1 = \beta. \tag{4.1}$$

An example of this is the ^{147}Sm/^{144}Nd-^{143}Nd/^{144}Nd diagram. In practice parent-daughter isochrons are not very useful for analysing basalt as the melting that occurs shortly before sampling fractionates parent from daughter, destroying the underlying p/d' signal. As such our main focus will be on the daughter-daughter isochron which is not affected in this way by the melting that occurs before sampling. This is where $(d/d')_2$ is plotted against $(d/d')_1$ for two different isotopic systems which have the same parent and daughter elements. In practice there is just one case in which this applies: the ^{206}Pb/^{204}Pb-^{207}Pb/^{204}Pb diagram. The lead-lead (daughter-daughter) isochron age τ_{ddi} is related to the slope β of the regression line by

$$\frac{^{235}\mathrm{U}}{^{238}\mathrm{U}} \cdot \frac{\mathrm{e}^{\lambda_{235} \tau_{\mathrm{ddi}}} - 1}{\mathrm{e}^{\lambda_{238} \tau_{\mathrm{ddi}}} - 1} = \beta, \tag{4.2}$$

where λ_{235} and λ_{238} are the decay constants of ^{235}U and ^{238}U respectively, and the ratio ^{235}U/^{238}U has constant value of 1/137.88 at the present day throughout the solar system.
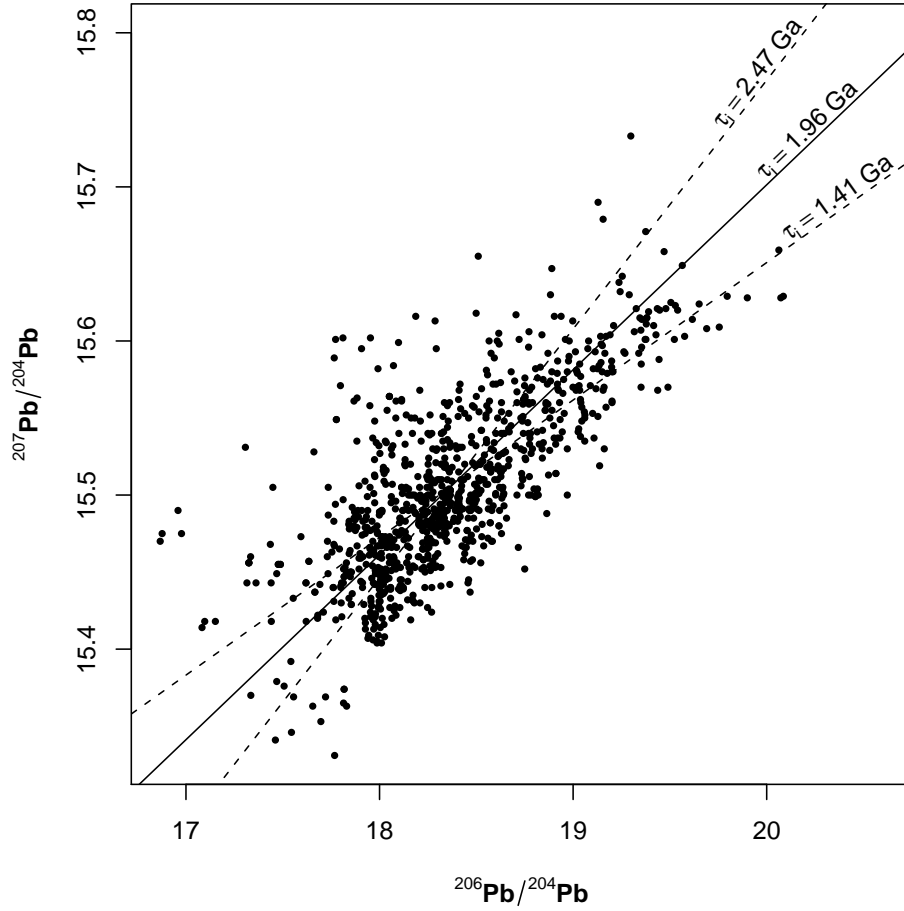
Figure 4.1: Scatterplot of measured MORB data (see Appendix A for a list of data sources). Three regression lines are plotted, and the corresponding pseudo-isochron ages shown. Linear regression of ^{207}Pb/^{204}Pb against ^{206}Pb/^{204}Pb gives a pseudo-isochron age of 1.41 Ga. Linear regression of ^{206}Pb/^{204}Pb against ^{207}Pb/^{204}Pb gives a pseudo-isochron age of 2.47 Ga. The geometric mean regression line has a pseudo-isochron age of 1.96 Ga.

Measured MORB data leads to pseudo-isochron ages around 2.0 Ga (Figure 4.1) and it is this age that we are most keen on understanding.

There have been a number of attempts to model the MORB and OIB pseudo-isochrons, and they have all been based on the same physical principles of radioactive decay, stirring, melting and mixing. There have been three main approaches: mean box models [Albarède, 2001, Allègre and Lewin, 1995a, Donnelly et al., 2004], statistical box models [Kellogg, 2004], and numerical simulations of mantle convection [Christensen and Hofmann, 1994, Xie and Tackley, 2004a]. This work aims to unite the different approaches and show that most of the modelling work that has been done so far can be best explained by some simple equations derived from a statistical box model of mantle processes. For easier reading, this chapter discusses the pseudo-isochron equations resulting from the model, and compares them to previous work. The following chapter presents the model derivations.

4.2   The model

The model used is a generalisation of the simple statistical box model described in the previous two chapters. Recall that the model treats the mantle as a box containing a

melting region which is visited by parcels of mantle material on an average time scale τ_{melt}. When parcels enter the melting region a melting event is performed which fractionates parent from daughter, producing a fraction F of melt and a fraction $1 - F$ of residue. G is the molar fraction of a chemical species that enters the melt, and depends on partition coefficients and the melt fraction F. For a given concentration C of species entering the melting region the concentration of the melt produced is CG/F and of the residue is $C(1 - G)/(1 - F)$ (Section 3.1.2). The melt and residue produced are then recycled back into the box. Importantly, we assume that the stirring in the box is strong, and this allows us to treat melting as a Poisson process [Albarède, 2005, 2003] with the statistics of material entering the melting region being the same as those over the whole box. Sampling is modelled by drawing N samples from the box and averaging [Meibom and Anderson, 2004], representative of the mixing that occurs after melting.

There are three crucial differences in the generalised model of this chapter from that described in the previous chapters. Firstly, the decay constant λ is no longer assumed small which allows $^{207}\mathrm{Pb}/^{204}\mathrm{Pb}$ to be modelled. Secondly, a steady state is not assumed, but instead the box is assumed to start with uniform concentrations of isotopes at an age τ_{s} before the present. Finally, the melt rate is no longer assumed constant over time but instead given by some function of age $\gamma_{\mathrm{melt}}(\tau)$. For constant melt rate $\gamma_{\mathrm{melt}}(\tau) = 1/\tau_{\mathrm{melt}}$, and we now redefine τ_{melt} as just the melting time scale at the present day.

A number of key analytic results can be derived from the model in the asymptotic limit where $N \to \infty$ (heavy averaging), and all the results in the next section are based on this limit (Section 5.2). Numerical simulation suggests that the dependence of the pseudo-isochron age on N is fairly weak (Figure 4.2), and so using the $N \to \infty$ asymptotics seems well justified. In this limit the distribution of isotopic ratios tends to a multivariate normal distribution, and expressions for the corresponding covariance matrix can be derived. In particular these expressions allow us to estimate the slopes of regression lines in plots of one isotopic ratio against another, and thus to derive expressions for the model pseudo-isochron ages.

There are many different ways of fitting a line to a cloud of data points. Following Allègre and Lewin [1995a] we have focused on the geometric mean regression line (also known as the reduced major axis regression line) whose slope is given by the ratio of the standard deviations of the two isotopic ratios in question (solid line in Figure 4.1). To give an estimate of the uncertainty in fitting a line we have also included results from using the two linear regression lines (dotted lines in Figure 4.1) in some of the figures. If the correlation is good all three lines should be similar.

4.3 The pseudo-isochron equations

The key to determining the pseudo-isochron ages in the model is the random variable $\hat{T}_{\mathrm{m}}$ which gives the distribution of parcel ages for those parcels that have passed through the melting region. The age of a parcel is defined as the time since last visit to the melting region. The parcels that have not visited the melting region are referred to as primordial parcels, and are not assigned an age. Let $\mathbb{E}$ be an expectation over a random variable, so that

$$\mathbb{E}f(\hat{T}_{\mathrm{m}}) = \int_0^{\tau_{\mathrm{s}}} f(\tau) q_{\mathrm{m}}(\tau) \, \mathrm{d}\tau, \qquad (4.3)$$

where $q_{\mathrm{m}}(\tau)$ is the probability density function of parcel ages, and $f(\tau)$ is an arbitrary given function. The subscript 'm' is to emphasise that we consider only those parcels that have passed through the melting region. In this model the primordial parcels make no
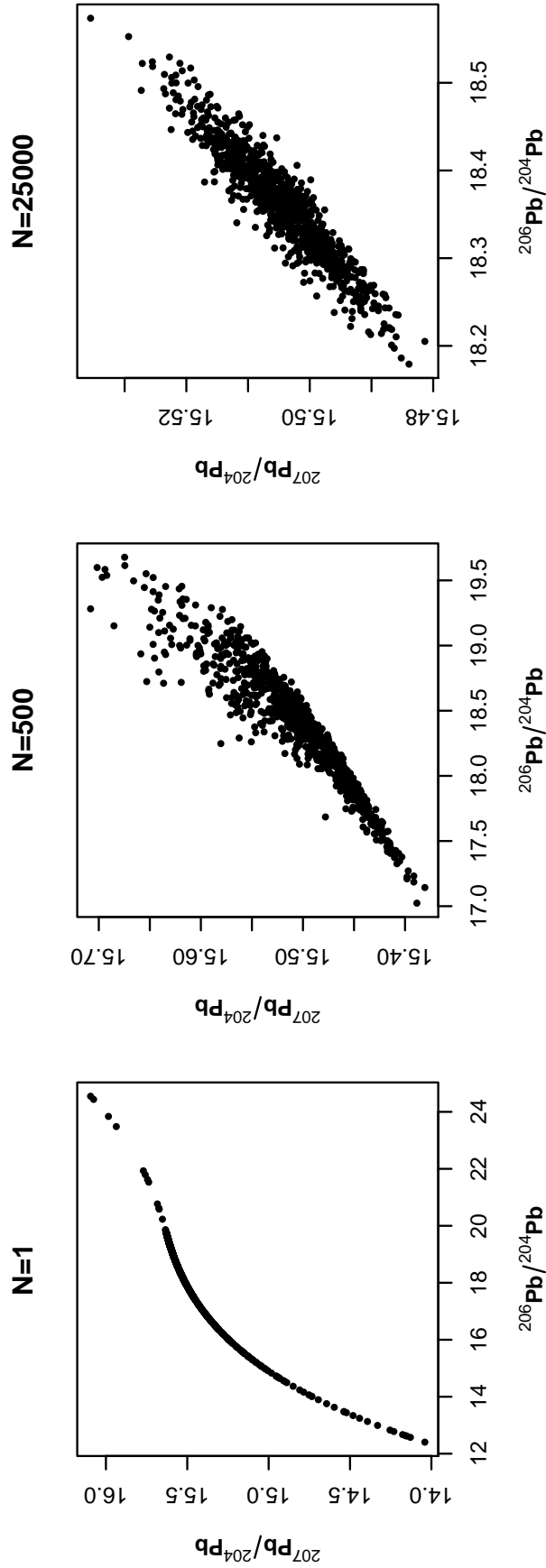
Figure 4.2: Constant melt rate model numerical simulations with $N = 1$, 500, and 25000, $\tau_{\mathrm{melt}} = 0.75$ Ga, $\tau_{\mathrm{s}} = 3.0$ Ga, $F = 1.5\%$, and a sample size of 1000. Partition coefficients are those used in the previous chapters, giving $G_{\mathrm{Pb}} = 0.42$ and $G_{\mathrm{U}} = 1.00$ using (2.2). The middle picture has similar variance and slope as the measured MORB data (Figure 4.1). The left picture shows unaveraged compositions, the right heavily averaged compositions. Note that the data forms an ellipse under heavy averaging, indicative of a bivariate normal distribution as expected by the central limit theorem [Novak, 1996]. The geometric mean regression line for the three cases leads to pseudo-isochron ages of 2.30 Ga, 1.98 Ga, and 2.06 Ga respectively. The pseudo-isochron equation (4.8) predicts an age of 2.03 Ga as $N \to \infty$.

contribution to the pseudo-isochron ages, since primordial parcels have uniform isotopic concentrations equal to the mean over the whole box. In terms of $\hat{T}_\mathrm{m}$ the model pseudo-isochron equations are simply (Sections 5.3 and 5.4)

$$(e^{\lambda \tau_\mathrm{pdi}} - 1)^2 = \mathbb{E}(e^{\lambda \hat{T}_\mathrm{m}} - 1)^2, \tag{4.4}$$

$$\frac{(e^{\lambda_{235} \tau_\mathrm{ddi}} - 1)^2}{(e^{\lambda_{238} \tau_\mathrm{ddi}} - 1)^2} = \frac{\mathbb{E}(e^{\lambda_{235} \hat{T}_\mathrm{m}} - 1)^2}{\mathbb{E}(e^{\lambda_{238} \hat{T}_\mathrm{m}} - 1)^2}, \tag{4.5}$$

where τ_pdi and τ_ddi are the parent-daughter and lead-lead pseudo-isochron ages respectively. Note that these expressions depend only on the decay constants and the distribution of parcel ages. The expressions clearly show that the pseudo-isochron ages are just particular weighted averages over the parcel ages. In fact the pseudo-isochron ages are examples of "generalised means" [Hardy et al., 1934, Qi, 2000] (Section 5.5). It also follows from these expressions that the pseudo-isochron ages are always greater than or equal to $\bar{\tau}_\mathrm{m} = \mathbb{E}\hat{T}_\mathrm{m}$ the mean age of parcels that have passed through the melting region. This is an important result as it means that models with $\bar{\tau}_\mathrm{m}$ greater than the observed lead-lead pseudo-isochron age of around 2.0 Ga cannot be compatible with the isotopic observations (such as the Davies [2002] model). If the parcel ages $\ll 1/\lambda$ then (4.4) reduces to $\tau_\mathrm{pdi} = \sqrt{\mathbb{E}\hat{T}_\mathrm{m}^2}$, a result which is independent of λ. Hence parent-daughter pseudo-isochron ages will be the same for all slowly decaying isotopes, namely for the $^{147}\mathrm{Sm}/^{144}\mathrm{Nd}$-$^{143}\mathrm{Nd}/^{144}\mathrm{Nd}$, $^{87}\mathrm{Rb}/^{86}\mathrm{Sr}$-$^{87}\mathrm{Sr}/^{86}\mathrm{Sr}$, $^{176}\mathrm{Lu}/^{177}\mathrm{Hf}$-$^{176}\mathrm{Hf}/^{177}\mathrm{Hf}$, and $^{232}\mathrm{Th}/^{204}\mathrm{Pb}$-$^{208}\mathrm{Pb}/^{204}\mathrm{Pb}$ diagrams.

The history of the rate of melting can be directly related to the distribution of parcel ages in the model. If $\gamma_\mathrm{melt}(\tau)$ is the melt rate as a function of age then $q_\mathrm{m}(\tau) = q(\tau)/\int_0^{\tau_\mathrm{s}} q(\tau)\,\mathrm{d}\tau$ where (Section 5.6)

$$q(\tau) = \gamma_\mathrm{melt}(\tau) \exp\left(-\int_0^\tau \gamma_\mathrm{melt}(\tau)\,\mathrm{d}\tau\right). \tag{4.6}$$

An important special case is where melt rate is constant $\gamma_\mathrm{melt}(\tau) = 1/\tau_\mathrm{melt}$, where τ_melt is a constant melting time scale. In this case $q_\mathrm{m}(\tau) = e^{-\tau/\tau_\mathrm{melt}}/(\tau_\mathrm{melt}(1 - e^{-\tau_\mathrm{s}/\tau_\mathrm{melt}}))$ and the model parent-daughter pseudo-isochron equation is (Section 5.7)

$$(e^{\lambda \tau_\mathrm{pdi}} - 1)^2 = \frac{\int_0^{\tau_\mathrm{s}} (e^{\lambda \tau} - 1)^2 e^{-\tau/\tau_\mathrm{melt}}\,\mathrm{d}\tau}{\tau_\mathrm{melt}\left(1 - e^{-\tau_\mathrm{s}/\tau_\mathrm{melt}}\right)}. \tag{4.7}$$

The most important feature of this equation is that it depends only on the three time scale parameters in the problem: the melting time scale τ_melt, the start age τ_s, and the decay constant λ; and not on any of the other parameters. Figure 4.3 plots solutions to (4.7) for $^{143}\mathrm{Nd}/^{144}\mathrm{Nd}$-$^{147}\mathrm{Sm}/^{144}\mathrm{Nd}$ ($\lambda = 0.00654\ \mathrm{Ga}^{-1}$) in two different ways. Figure 4.4 shows similar graphs for $^{235}\mathrm{U}/^{204}\mathrm{Pb}$-$^{207}\mathrm{Pb}/^{204}\mathrm{Pb}$ ($\lambda = 0.985\ \mathrm{Ga}^{-1}$), a case for which the decay is not linearisable. Note that there is a reasonable uncertainty in the parent-daughter pseudo-isochron ages indicated by the wide grey region in Figures 4.3b and 4.4b. This expected (and indeed observed) lack of very good correlation is another reason why parent-daughter pseudo-isochron ages are not particularly useful.

The more useful pseudo-isochron equation is the lead-lead pseudo-isochron equation given for constant melt rate by (Section 5.8)

$$\frac{(e^{\lambda_{235} \tau_\mathrm{ddi}} - 1)^2}{(e^{\lambda_{238} \tau_\mathrm{ddi}} - 1)^2} = \frac{\int_0^{\tau_\mathrm{s}} (e^{\lambda_{235} \tau} - 1)^2 e^{-\tau/\tau_\mathrm{melt}}\,\mathrm{d}\tau}{\int_0^{\tau_\mathrm{s}} (e^{\lambda_{238} \tau} - 1)^2 e^{-\tau/\tau_\mathrm{melt}}\,\mathrm{d}\tau}. \tag{4.8}$$
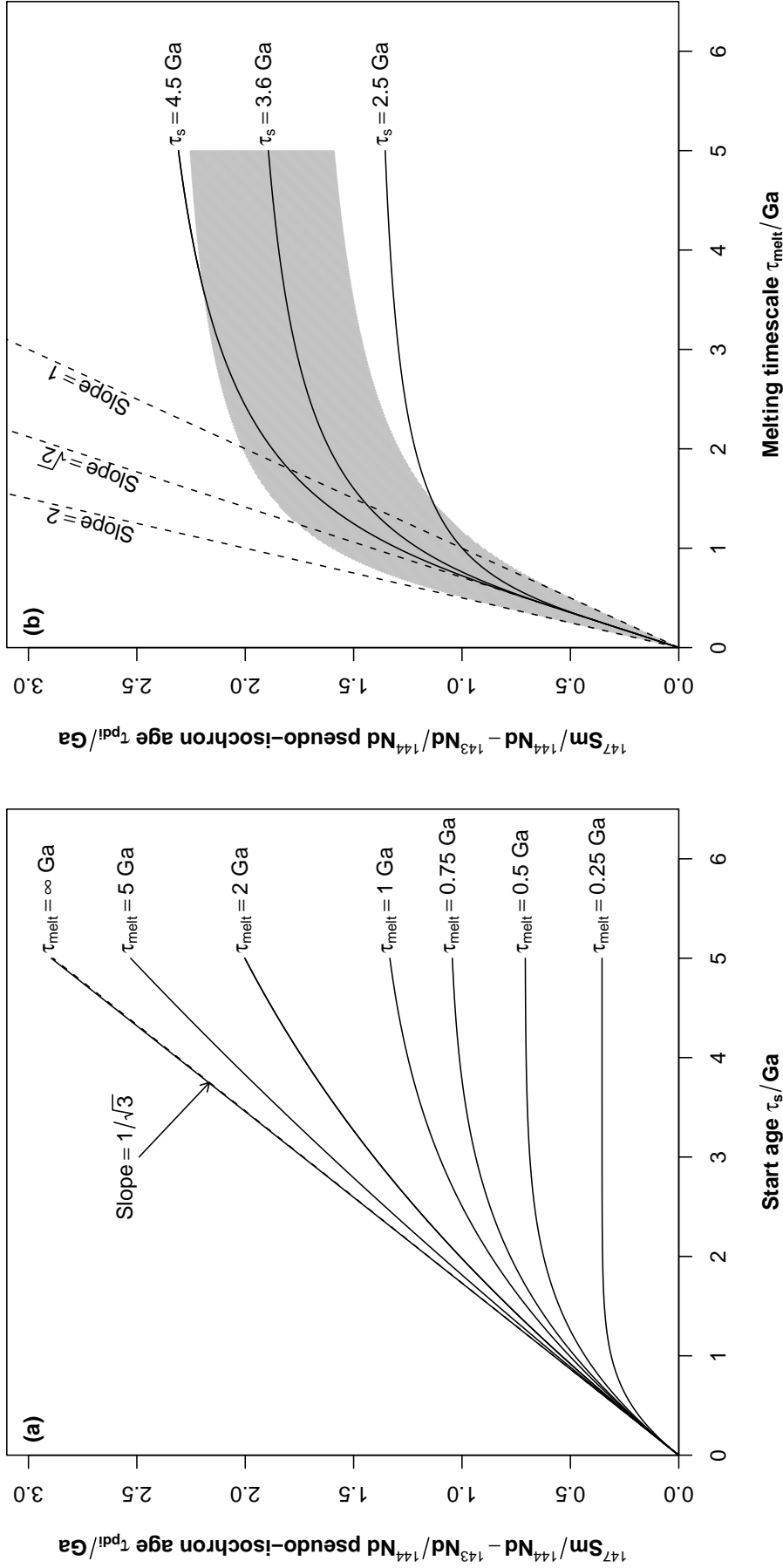
Figure 4.3: Plot showing the variation of $^{147}Sm/^{144}Nd$-$^{143}Nd/^{144}Nd$ ($\lambda = 0.00654$ Ga^{-1}) model pseudo-isochron age τ_{pdi} (4.7). Near identical curves will be produced for other isotopic systems where the decay is linearisable (τ_{melt} or $\tau_s \ll 1/\lambda$), such as $^{87}Rb/^{86}Sr$-$^{87}Sr/^{86}Sr$, $^{176}Lu/^{177}Hf$-$^{176}Hf/^{177}Hf$, and $^{232}Th/^{204}Pb$-$^{208}Pb/^{204}Pb$. (a) plots τ_{pdi} against start age τ_s for different values of the melting time scale τ_{melt}. For small values of τ_s the curves have slope $= 1/\sqrt{3}$. The curves shown asymptote to $\sqrt{2}\tau_{melt}$ for large values of τ_s with the exception of the $\tau_{melt} = \infty$ Ga curve. (b) plots τ_{pdi} against τ_{melt} for different values of τ_s. Values of $\tau_s = 2.5$ Ga, 3.6 Ga and 4.5 Ga have been chosen for comparison with numerical simulations of mantle convection [Christensen and Hofmann, 1994, Xie and Tackley, 2004a]. For small values of τ_{melt} the curves have a slope of $\sqrt{2}$. For large values of τ_{melt} the curves asymptote to pseudo-isochron ages of 1.45 Ga, 2.08 Ga and 2.60 Ga respectively ($\sim \tau_s/\sqrt{3}$). Shown in grey is the uncertainty in fitting a line for $\tau_s = 3.6$ Ga: the geometric mean regression age (solid line) lies between the two linear regression ages which bound the grey region. The corresponding regions for $\tau_s = 2.5$ Ga and 4.5 Ga have not been plotted as they substantially overlap the 3.6 Ga region.
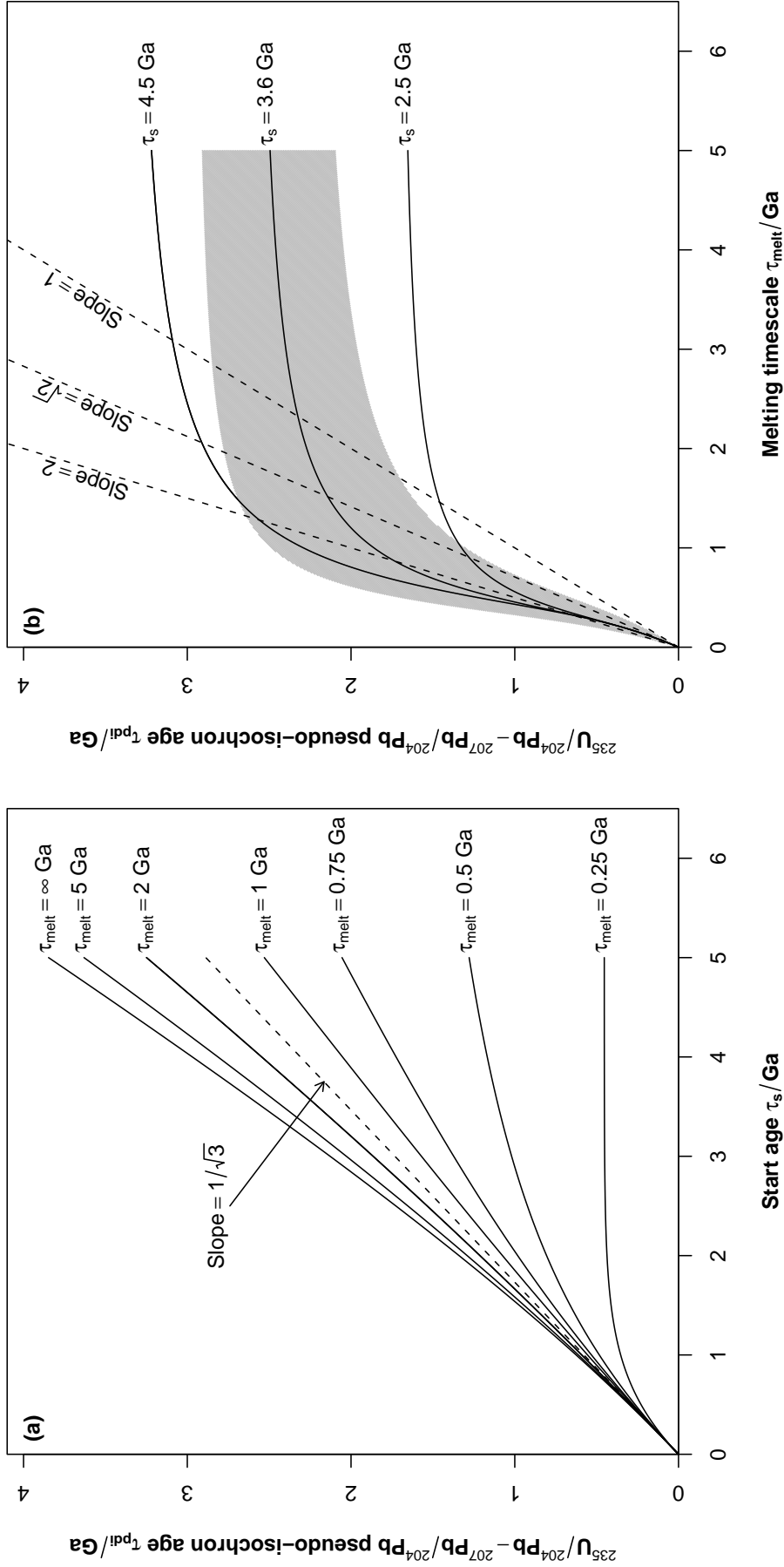
Figure 4.4: As Figure 4.3 but for ^{235}U/^{204}Pb-^{207}Pb/^{204}Pb ($\lambda = 0.985$ Ga^{-1}), where the decay is not linearisable. (a) plots τ_{pdi} against τ_{s} for different values of τ_{melt}. For $\tau_{\mathrm{melt}} < 1/2\lambda = 0.51$ Ga the curves flatten out for large τ_{s}. For $\tau_{\mathrm{melt}} > 1/2\lambda$ the curves grow linearly with τ_{s} for large τ_{s}. The $\tau_{\mathrm{melt}} = \infty$ Ga curve approaches a slope of 1 for large τ_{s}. (b) plots τ_{pdi} against τ_{melt} for different values of τ_{s}. For large values of τ_{melt} the curves asymptote to pseudo-isochron ages of 1.73 Ga, 2.63 Ga and 3.40 Ga respectively.
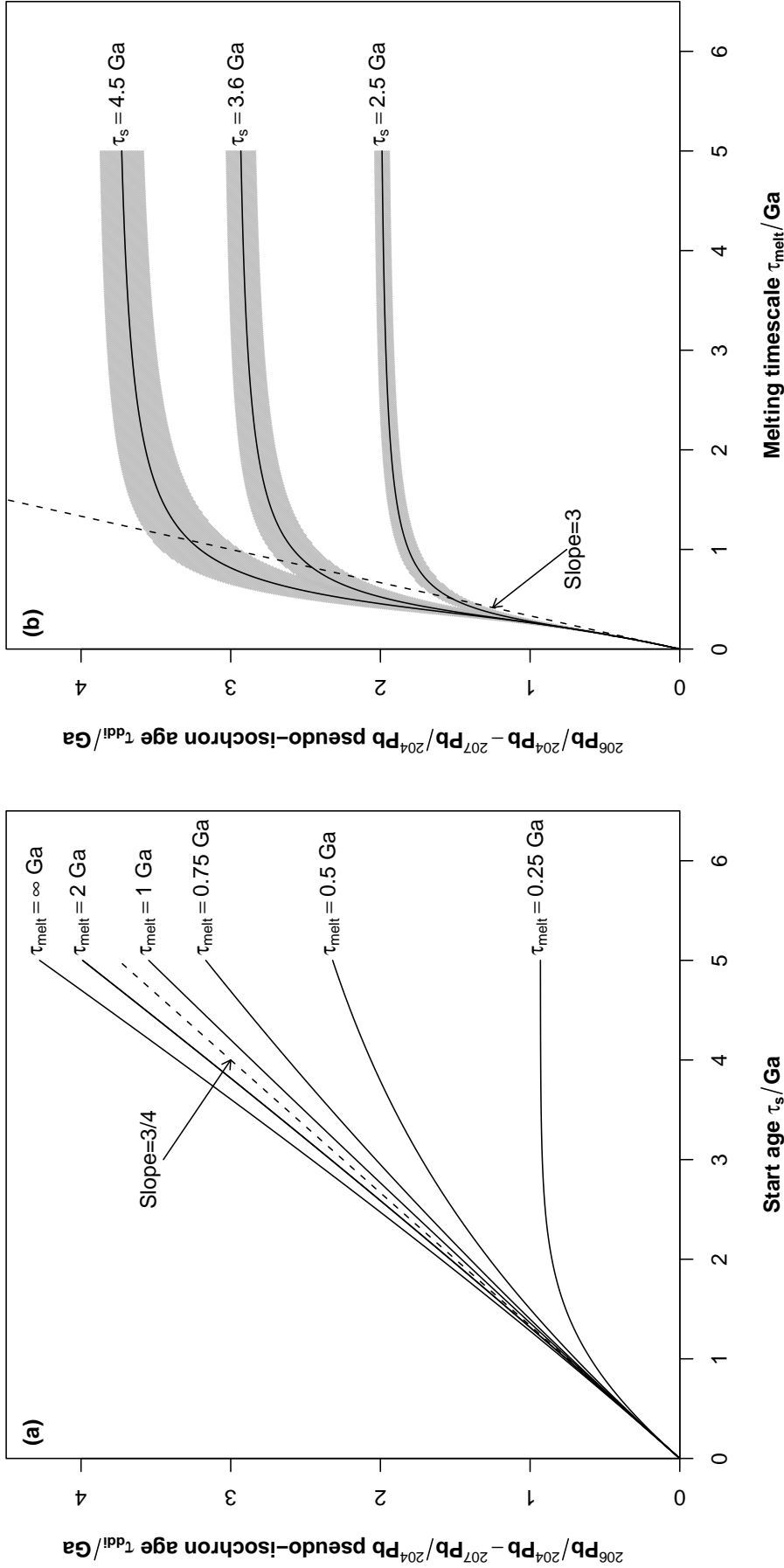
Figure 4.5: As Figure 4.3 but for the ^{206}Pb/^{204}Pb-^{207}Pb/^{204}Pb ($\lambda_{238} = 0.155$ Ga^{-1}, $\lambda_{235} = 0.985$ Ga^{-1}) model pseudo-isochron age τ_{ddi} (4.8). (a) plots τ_{ddi} against τ_s for different values of τ_{melt}. For small values of τ_s the curves have slope = 0.75. For values of τ_{melt} around 1-2 Ga the approximate relationship $\tau_{ddi} \approx 0.75\tau_s$ holds for the range of τ_s plotted. For $\tau_{melt} < 1/2\lambda_{235} = 0.51$ Ga the curves flatten out for large τ_s. For $\tau_{melt} > 1/2\lambda_{235}$ the curves grow linearly with τ_s for large τ_s. The $\tau_{melt} = \infty$ Ga curve approaches a slope of 1 for large τ_s. (b) plots τ_{ddi} against τ_{melt} for different values of τ_s. For small values of τ_{melt} the curves have a slope of 3. For large values of τ_{melt} the curves asymptote to pseudo-isochron ages of 2.02 Ga, 2.99 Ga and 3.81 Ga respectively. Shown in grey is the uncertainty in fitting a line for each of the τ_s values.

Figure 4.5 plots solutions to (4.8). Note that the lead-lead pseudo-isochron ages are fairly well constrained as the model correlation is good (as indicated by the narrow grey regions of Figure 4.5b).

The pseudo-isochron equations only encode age information, and do not involve the parameters G and F. However, if we are concerned with the variance of isotopic ratios, or the slopes in plots of one isotopic ratio against another which do not have common parent and daughter elements, then these parameters are involved. This was the main focus of the previous chapters, and the corresponding generalisation of (2.1) for the standard deviation σ of d/d' ratios after sampling is (Section 5.2)

$$\sigma = \frac{\bar{p}}{\bar{d}'} \cdot \frac{|G_p - G_d|}{\sqrt{NF(1-F)}} \cdot \sqrt{\int_0^{\tau_s} \left(e^{\lambda\tau} - 1\right)^2 q(\tau)\,\mathrm{d}\tau}, \tag{4.9}$$

where $\bar{p}/\bar{d}'$ is the ratio of mean parent isotope concentration to mean reference isotope concentration over the whole box at the present day.

4.4 Linear pseudo-isochron equations

There are three models based on linear evolution equations that have recently been proposed for the pseudo-isochrons [Albarède, 2001, Allègre and Lewin, 1995a, Donnelly et al., 2004]. In fact, all three models produce identical pseudo-isochron equations and are closely related. Albarède [2001] and Donnelly et al. [2004] both consider a two reservoir model, and derive the pseudo-isochron equations in precisely the same way. Allègre and Lewin [1995a] consider a different set of linear evolution equations for isotopic dispersion, which with some rearranging are almost equivalent to the two interacting reservoir equations.

The governing equations for two interacting reservoirs derived by Albarède [2001] and Donnelly et al. [2004] are

$$\frac{\mathrm{d}n_1}{\mathrm{d}t} = -\frac{n_1}{\tau_1} + \frac{n_2}{\tau_2} - \lambda n_1, \tag{4.10}$$

$$\frac{\mathrm{d}n_2}{\mathrm{d}t} = \frac{n_1}{\tau_1} - \frac{n_2}{\tau_2} - \lambda n_2, \tag{4.11}$$

$$\frac{\mathrm{d}m_1}{\mathrm{d}t} = -\frac{m_1}{\theta_1} + \frac{m_2}{\theta_2} + \lambda n_1, \tag{4.12}$$

$$\frac{\mathrm{d}m_2}{\mathrm{d}t} = \frac{m_1}{\theta_1} - \frac{m_2}{\theta_2} + \lambda n_2, \tag{4.13}$$

$$\frac{\mathrm{d}s_1}{\mathrm{d}t} = -\frac{s_1}{\theta_1} + \frac{s_2}{\theta_2}, \tag{4.14}$$

$$\frac{\mathrm{d}s_2}{\mathrm{d}t} = \frac{s_1}{\theta_1} - \frac{s_2}{\theta_2}, \tag{4.15}$$

where the notation of Albarède [2001] has been followed. n_i, m_i and s_i are the *total number of moles* of parent, daughter and reference isotopes respectively in reservoir i (these should not be confused with p, d and d' used throughout to represent *concentrations*). τ_i is the residence time of the parent element in reservoir i, θ_i the corresponding residence time for the daughter element. Let Σ_n, Σ_m and Σ_s be the total number of moles of parent, daughter and reference isotopes in both reservoirs: $\Sigma_n = n_1 + n_2$, $\Sigma_m = m_1 + m_2$, and $\Sigma_s = s_1 + s_2$. Then

$$\frac{\mathrm{d}\Sigma_n}{\mathrm{d}t} = -\lambda\Sigma_n, \quad \frac{\mathrm{d}\Sigma_m}{\mathrm{d}t} = \lambda\Sigma_n, \quad \frac{\mathrm{d}\Sigma_s}{\mathrm{d}t} = 0, \tag{4.16}$$

and the governing equations for reservoir 1 can be rewritten as

$$\frac{dn_1}{dt} = -\frac{n_1}{\tau} + \frac{\Sigma_n}{\tau_2} - \lambda n_1, \tag{4.17}$$

$$\frac{dm_1}{dt} = -\frac{m_1}{\theta} + \frac{\Sigma_m}{\theta_2} + \lambda n_1, \tag{4.18}$$

$$\frac{ds_1}{dt} = -\frac{s_1}{\theta} + \frac{\Sigma_s}{\theta_2}, \tag{4.19}$$

where τ and θ are the relaxation times of the two elements: the harmonic means of the residence times in the two reservoirs

$$\frac{1}{\tau} = \frac{1}{\tau_1} + \frac{1}{\tau_2}, \quad \frac{1}{\theta} = \frac{1}{\theta_1} + \frac{1}{\theta_2}. \tag{4.20}$$

We now rewrite these equations for closer comparison with Allègre and Lewin [1995a]. Introduce new variables $n_1^\star$ and $m_1^\star$ defined by

$$n_1^\star = \frac{1}{\Sigma_s}\left(n_1 - s_1 \left(\frac{n}{s}\right)_{1+2}\right), \tag{4.21}$$

$$m_1^\star = \frac{1}{\Sigma_s}\left(m_1 - s_1 \left(\frac{m}{s}\right)_{1+2}\right), \tag{4.22}$$

where the subscript 1+2 refers to the total system: $(n/s)_{1+2} = \Sigma_n/\Sigma_s$ and $(m/s)_{1+2} = \Sigma_m/\Sigma_s$. Note that $n_2^\star = -n_1^\star$ and $m_2^\star = -m_1^\star$. The governing equations in the starred variables can then be written as

$$\frac{dn_1^\star}{dt} = \left(\frac{1}{\tau_2} - \frac{1}{\theta_2}\right)\left(\frac{n}{s}\right)_{1+2} - \lambda n_1^\star - \frac{1}{\Sigma_s}\left(\frac{n_1}{\tau} - \frac{s_1}{\theta}\left(\frac{n}{s}\right)_{1+2}\right), \tag{4.23}$$

$$\frac{dm_1^\star}{dt} = \lambda n_1^\star - \frac{m_1^\star}{\theta}. \tag{4.24}$$

These governing equations take on a particularly simple form if the relaxation times for parent and daughter elements are the same, $\tau = \theta = 1/\gamma$ say. For later convenience, define $g_n = \tau/\tau_1$ and $g_m = \theta/\theta_1$, and note that $0 \le g_n, g_m \le 1$. Then

$$\frac{dn_1^\star}{dt} = -(g_n - g_m)\gamma\left(\frac{n}{s}\right)_{1+2} - (\lambda + \gamma)n_1^\star, \tag{4.25}$$

$$\frac{dm_1^\star}{dt} = \lambda n_1^\star - \gamma m_1^\star. \tag{4.26}$$

These can be compared to the governing equations (1) and (2) of Allègre and Lewin [1995a]

$$\frac{d\langle\mu\rangle(t)}{dt} = A(t) - (\lambda + M(t))\langle\mu\rangle(t), \tag{4.27}$$

$$\frac{d\langle\alpha\rangle(t)}{dt} = \lambda\langle\mu\rangle(t) + B(t) - M(t)\langle\alpha\rangle(t). \tag{4.28}$$

These governing equations are identical to (4.25) and (4.26) if the chemical dispersion $\langle\mu\rangle(t) = n_1^\star$, the isotopic dispersion $\langle\alpha\rangle(t) = m_1^\star$, the rate of injection of chemical heterogeneity $A(t) = -(g_n - g_m)\gamma(n/s)_{1+2}$, the rate of injection of isotopic heterogeneity $B(t) = 0$, and the stirring parameter $M(t) = \gamma$. The appendix of Allègre and Lewin

[1995a] discusses solutions to these equations in the simplified form ((A1-1) and (A1-2) of Allègre and Lewin [1995a])

$$\frac{\mathrm{d}\langle\mu\rangle(t)}{\mathrm{d}t} = \frac{\Delta\langle\mu\rangle\mathrm{e}^{-\lambda t}}{R} - \left(\lambda + \tau_{\mathrm{stir}}^{-1}\right)\langle\mu\rangle(t), \tag{4.29}$$

$$\frac{\mathrm{d}\langle\alpha\rangle(t)}{\mathrm{d}t} = \lambda\langle\mu\rangle(t) + \frac{\Delta\langle\alpha\rangle}{R} - \frac{\langle\alpha\rangle(t)}{\tau_{\mathrm{stir}}}, \tag{4.30}$$

where by comparison $\Delta\langle\mu\rangle/R = -\left(g_n - g_m\right)\gamma(n/s)_{1+2}(0)$, $\Delta\langle\alpha\rangle = 0$, and $\tau_{\mathrm{stir}} = \gamma^{-1}$. These are solved subject to initial conditions $\langle\mu\rangle(0) = 0$, $\langle\alpha\rangle(0) = 0$ which corresponds to the two reservoirs having initially identical isotopic ratios $(n/s)_1(0) = (n/s)_2(0) = (n/s)_{1+2}(0)$ and $(m/s)_1(0) = (m/s)_2(0) = (m/s)_{1+2}(0)$. Allègre and Lewin define the slope of their parent-daughter pseudo-isochron by the ratio $\langle\alpha\rangle(t)/\langle\mu\rangle(t)$. However,

$$\frac{\langle\alpha\rangle(t)}{\langle\mu\rangle(t)} = \frac{m_1^\star}{n_1^\star} = \frac{m_1 - s_1\,(m/s)_{1+2}}{n_1 - s_1\,(n/s)_{1+2}} = \frac{(m/s)_1 - (m/s)_{1+2}}{(n/s)_1 - (n/s)_{1+2}}. \tag{4.31}$$

Hence the definition of the slope by the ratio $\langle\alpha\rangle(t)/\langle\mu\rangle(t)$ is the same as the slope of the line through the reservoirs 1, 2 and 1+2 on the parent-daughter isochron diagram. Thus provided the relaxation times are equal in the two reservoir model, and there is no excess isotopic heterogeneity $\Delta\langle\alpha\rangle$ in the Allègre and Lewin model, the pseudo-isochron equations are exactly the same. Importantly, the stirring time τ_{stir} of Allègre and Lewin can be reinterpreted as the common relaxation time in the two reservoir model.

In one sense the Allègre and Lewin model is more general than the two reservoir model because of the excess isotopic heterogeneity term $\Delta\langle\alpha\rangle$, but in practice this was always set to zero for their pseudo-isochron calculations. On the other hand, the two reservoir model can be thought of as more general since it allows the parent and daughter elements to have different relaxation times. If parent and daughter are fractionated by the same melting process we might expect relaxation times to be the same. Also, note that in secular equilibrium only the relaxation time θ of the daughter element determines the pseudo-isochron [Albarède, 2001].

Standard deviations are not conservative and should not be modelled by linear evolution equations such as (4.27) and (4.28). The justification of these equations by Allègre and Lewin is rather ad hoc. The two reservoir model and the Allègre and Lewin model are both linear models and thus essentially concerned with mean values, whereas standard deviations are fundamentally nonlinear. In the statistical box model it possible to discuss standard deviations, variances and covariances, as well as means, because the underlying probability distributions are being modelled. Note that the starred variables which relate the Allègre and Lewin model to the two reservoir model arise naturally from the linearisation

$$\frac{x}{y} - \frac{\bar{x}}{\bar{y}} \approx \frac{1}{\bar{y}}\left(x - y\frac{\bar{x}}{\bar{y}}\right), \tag{4.32}$$

valid for $|x - \bar{x}| \ll \bar{x}$ and $|y - \bar{y}| \ll \bar{y}$. The starred variables turn out to be particularly useful when considering asymptotics for large N in the statistical box model (Section 3.3.1).

We can make an important connection between the statistical box model and the two reservoir model by dividing the parcels in the box into two groupings. Labelling the residue parcels and a fraction $1-F$ of the primordial parcels as one reservoir, and the melt parcels and a fraction F of the primordial parcels as another reservoir, gives a two reservoir system with a common relaxation time. The connection is made if $\gamma = \gamma_{\mathrm{melt}}$, $g_n = G_p$ and $g_m = G_d$ (Section 5.9). Furthermore, this suggests an alternative way of writing the linear

pseudo-isochron equations of Allègre and Lewin in terms of an age distribution. These are the same as (4.4) and (4.5) except with the squareds removed

$$e^{\lambda \tau_{\mathrm{pdil}}} - 1 = \mathbb{E}(e^{\lambda \hat{T}_{\mathrm{m}}} - 1), \tag{4.33}$$

$$\frac{e^{\lambda_{235} \tau_{\mathrm{ddil}}} - 1}{e^{\lambda_{238} \tau_{\mathrm{ddil}}} - 1} = \frac{\mathbb{E}(e^{\lambda_{235} \hat{T}_{\mathrm{m}}} - 1)}{\mathbb{E}(e^{\lambda_{238} \hat{T}_{\mathrm{m}}} - 1)}, \tag{4.34}$$

where τ_{pdil} and τ_{ddil} are the linear parent-daughter and lead-lead pseudo-isochron ages respectively. It should be noted that the linear pseudo-isochron ages will always be less than the corresponding ages obtained from (4.4) and (4.5). The pseudo-isochron equations (4.4) and (4.5) for our problem are different because they involve the variance of a mixture of melt, residue and primordial parcels, whereas the above equations result from mean values of reservoirs. The squareds reflect the difference between looking at a mean value and looking at a variance. To emphasise the similarities and differences between the linear pseudo-isochron equations and our pseudo-isochron equations, Figures 4.3, 4.4, and 4.5 mimic Figures 3, 4 and 9 of Allègre and Lewin [1995a].

A key result used in Allègre and Lewin [1995a] to estimate a stirring time for the mantle is the lead-lead pseudo-isochron age relationship $\tau_{\mathrm{ddil}} \sim 2\tau_{\mathrm{stir}}$ for vigorous stirring. In the context of the quadratic pseudo-isochron equations the corresponding asymptotic result is $\tau_{\mathrm{ddi}} \sim 3\tau_{\mathrm{melt}}$ for rapid remelting. In practice this result is only accurate for very rapid remelting, as it requires that $\tau_{\mathrm{melt}} \ll \tau_{\mathrm{s}}$ and $\tau_{\mathrm{melt}} \ll 1/\lambda_{235} = 1.0$ Ga. Hence comparison between the linear pseudo-isochron equations and our pseudo-isochron equations is best done using the full equations rather than any rapid remelting asymptotics. Using (A2-9) of Allègre and Lewin [1995a], if $\tau_{\mathrm{s}} = 4.5$ Ga then $\tau_{\mathrm{stir}} = 0.82$ Ga is needed to produce a lead-lead pseudo-isochron age of 2.0 Ga. Using our (4.8), if $\tau_{\mathrm{s}} = 4.5$ Ga then $\tau_{\mathrm{melt}} = 0.45$ Ga is needed, which is almost a factor of 2 less. For parent-daughter isochrons the vigorous stirring relationship $\tau_{\mathrm{pdil}} \sim \tau_{\mathrm{stir}}$ in Allègre and Lewin [1995a] becomes $\tau_{\mathrm{pdi}} \sim \sqrt{2}\tau_{\mathrm{melt}}$ for rapid remelting ($\tau_{\mathrm{melt}} \ll \tau_{\mathrm{s}}$ and $1/\lambda$).

The statistical model of Kellogg et al. [2002] also has a τ_{stir} parameter, but note that this takes on a different meaning to the τ_{stir} of the Allègre and Lewin [1995a] model. The Kellogg et al. τ_{stir} reduces the length scale of heterogeneities before sampling: essentially it relates to the parameter N in our model. The τ_{stir} in the Allègre and Lewin model describes the destruction of heterogeneity by repeated melting, and thus is τ_{melt} in our model.

4.5   Numerical simulations of mantle convection

Christensen and Hofmann [1994] put isotopic tracers into a numerical simulation of mantle convection. 252,000 tracers were used, and sampling was performed by dividing the domain into 40×20 sampling cells, and averaging over those cells. Thus N for their model is around 300, although note that their model is slightly different to that presented here as different cells may have different numbers of tracers. They have a constant rate of melting in their model, so (4.8) applies. Their standard model has $\tau_{\mathrm{s}} = 3.6$ Ga and $\tau_{\mathrm{melt}} = -1.36/\log(1-0.9) = 0.59$ Ga (inferred from the statement "in 1.36 Ga, statistically 90% of the total basalt content has been cycled through a melting zone in the model"). They found a lead-lead pseudo-isochron age of 2.10 Ga which compares very favourably to the figure of 2.15 Ga that is predicted by (4.8) (and not as favourably with the figure of 1.39 Ga predicted by the linear pseudo-isochron equation (A2-9) of Allègre and Lewin [1995a]).

Xie and Tackley [2004a] used a similar approach to Christensen and Hofmann [1994] for a different numerical simulation. Their model differs in having a melt rate which changes over time; melting being more vigorous in the past. 400,000 tracers were used, and the domain divided into 256×64 sampling cells, and thus N for their model is around 25. This smaller amount of averaging probably explains why their arrays show less correlation than the Christensen and Hofmann arrays. Their model has less frequent remelting than the Christensen and Hofmann model, which is why larger isotopic ages are observed. A rough rule of thumb for less frequent melting ($\tau_{\mathrm{melt}} = $ 1-2 Ga) is $\tau_{\mathrm{ddi}} \approx 0.75\tau_{\mathrm{s}}$ (Figure 4.5a). The rule works reasonably well at estimating the pseudo-isochron ages they found, but the simple constant melt rate formula of (4.8) does not actually apply in this case. To do a more accurate comparison we need to examine carefully the distribution of ages the Xie and Tackley model produces.

In Figures 5c and 5e of Xie and Tackley [2004a] the integrated crustal production and crustal production rate are plotted against time for a $\tau_{\mathrm{s}} = 4.5$ Ga run. We can use this information in our statistical box model to predict the distribution of ages (Section 5.6). Figure 4.6 shows there is good agreement between these calculated distributions and the age distribution observed (Figure 10a of Xie and Tackley [2004a]). Using (4.4) and (4.5) we can use these age distributions to calculate pseudo-isochron ages. The integrated crustal production gives a lead-lead pseudo-isochron age of 3.58 Ga, the crustal production rate an age of 3.76 Ga, and the observed age distribution an age of 3.67 Ga. These estimates are all slightly greater than the observed pseudo-isochron age of 3.39 Ga (Figure 8a of Xie and Tackley [2004a]). The small differences may be due to various factors, including differences in line fitting, the effects of binning and windowing, and the small value of N used in their simulations.

Xie and Tackley discussed the effect of greater rates of melting in the past on pseudo-isochron age. To examine this they introduced a box model and calculated a mean parcel age for three different scenarios. Using our statistical box model we can go further and calculate not only the mean age but also the age distributions and pseudo-isochron ages. Their three different scenarios were (1) melt rate constant, (2) melt rate proportional to the square of mantle heating rate (referred to as H^2), and (3) as H^2 but also proportional to depth of melting (referred to as H^2d). Mantle heating rate was assumed to decay with a half-life of 2.247 Ga. This corresponds to an exponentially growing history of melt rate $\gamma_{\mathrm{melt}}(\tau) = \mathrm{e}^{\nu\tau}/\tau_{\mathrm{melt}}$ where $\nu = (2\log 2)/2.247 = 0.62$ Ga^{-1} and τ_{melt} is the melting time scale at the present day. We have approximated the H^2d case by a similar exponential growth with $\nu = 0.74$ Ga^{-1}; only a small discrepancy in the mean ages results from this approximation. The age distributions are plotted in Figure 4.7 and the mean and pseudo-isochron ages are given in Table 4.1.

An important distinction should be drawn between the two mean ages given in Table 4.1. Xie and Tackley [2004a] calculated the mean mantle age $\bar{\tau}_{\mathrm{total}}$ which includes all parcels, and in particular it includes the primordial parcels which are assigned an age of τ_{s}. This is a misleading mean age to relate to the pseudo-isochron ages, since for the pseudo-isochron ages it is only the parcels that have passed through the melting region which matter. Hence of greater interest is $\bar{\tau}_{\mathrm{m}} = \mathbb{E}\hat{T}_{\mathrm{m}}$ which is a mean age that only includes the parcels that have passed through the melting region. It can be formally shown that the pseudo-isochron ages must be greater than or equal to $\bar{\tau}_{\mathrm{m}}$, but there is no such constraint on $\bar{\tau}_{\mathrm{total}}$. Indeed Table 4.1 shows pseudo-isochron ages both greater than and less than $\bar{\tau}_{\mathrm{total}}$.

Note that the smallest Pb-Pb pseudo-isochron age in Table 4.1 is 2.97 Ga, so none of the scenarios are compatible with the measured MORB data. Furthermore note that
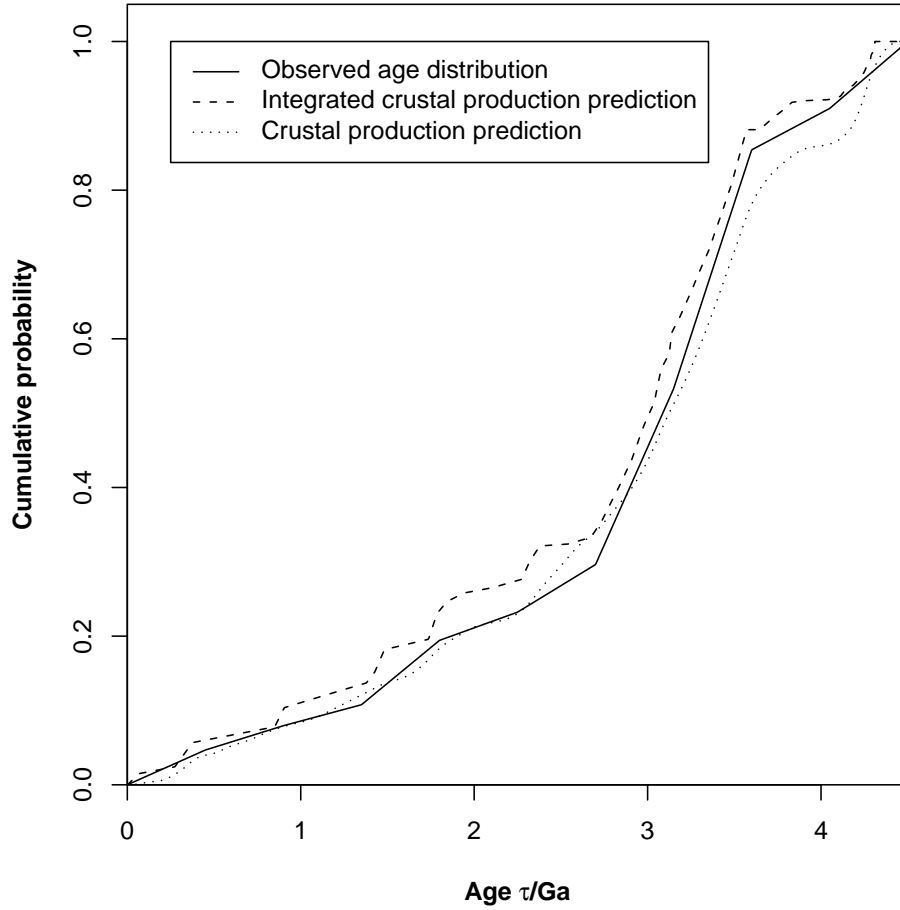
Figure 4.6: Plot of cumulative distribution function $Q_m(\tau) = \int_0^\tau q_m(\tau)\,\mathrm{d}\tau$ for a Xie and Tackley [2004a] numerical simulation. The solid line shows the observed age distribution (their Figure 10a), where residue and basalt tracers have been lumped together. The dashed line is a calculated age distribution (Section 5.6) based on their observed integrated crustal production (solid line of their Figure 5c). The dotted line is a calculated age distribution based on their observed crustal production rate (solid line of their Figure 5e), although note that this rate is actually an average over three slightly different runs. The three age distributions are very similar. Using (4.4) and (4.5), the age distributions correspond to lead-lead pseudo-isochron ages τ_{ddi} of 3.67 Ga, 3.58 Ga, and 3.76 Ga, Sm-Nd pseudo-isochron ages τ_{pdi} of 3.00 Ga, 2.88 Ga, and 3.06 Ga, and mean ages $\bar{\tau}_{\mathrm{m}}$ of 2.81 Ga, 2.67 Ga, and 2.86 Ga respectively.
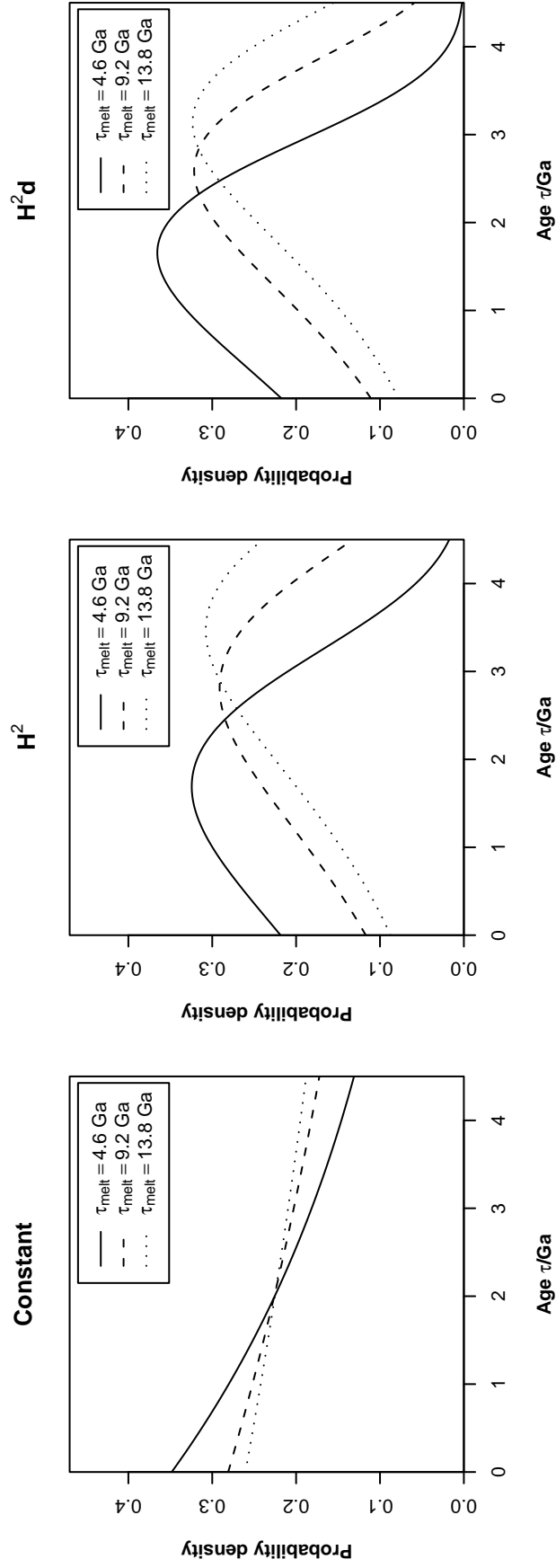
Figure 4.7: Plot of probability density function $q_{\mathrm{m}}(\tau)$. The melt rate is given by $\gamma_{\mathrm{melt}}(\tau) = e^{\nu\tau}/\tau_{\mathrm{melt}}$, and the three cases are constant ($\nu = 0$ Ga^{-1}), H^2 ($\nu = 0.62$ Ga^{-1}), and H^2d ($\nu = 0.74$ Ga^{-1}).

Table 4.1: Age calculations for varying melt rate $\gamma_{\text{melt}}(\tau) = e^{\nu\tau}/\tau_{\text{melt}}$ (compare with Table 4 of Xie and Tackley [2004a]). The three cases are constant ($\nu = 0$ Ga^{-1}), H^2 ($\nu = 0.62$ Ga^{-1}), and $H^2 d$ ($\nu = 0.74$ Ga^{-1}). The present day crustal production rate $b/\tau_{\text{melt}} = 1/46$ Ga^{-1} in all cases. $b = 0.1, 0.2,$ and 0.3, and hence $\tau_{\text{melt}} = 4.6$ Ga, 9.2 Ga and 13.8 Ga. $\tau_{\text{s}} = 4.5$ Ga and all ages are quoted in Ga.

|	Constant			H^2			$H^2 d$		
τ_{melt}	4.6	9.2	13.8	4.6	9.2	13.8	4.6	9.2	13.8
$\bar{\tau}_{\text{total}}$	2.87	3.56	3.84	1.82	2.54	2.96	1.67	2.31	2.72
$\bar{\tau}_{\text{m}}$	1.89	2.07	2.13	1.81	2.39	2.64	1.67	2.27	2.58
$\tau_{\text{pdi}}[\text{Sm-Nd}]$	2.28	2.44	2.50	2.09	2.66	2.89	1.92	2.53	2.82
$\tau_{\text{ddi}}[\text{Pb-Pb}]$	3.72	3.77	3.78	3.29	3.69	3.81	2.97	3.54	3.73

more frequent remelting in the past can have the effect of both increasing and decreasing the pseudo-isochron age (e.g. $\tau_{\text{melt}} = 13.8$ Ga in Table 4.1). This is due to two competing effects: if melting is more rapid in the past it can mean a lot of the early heterogeneity is destroyed, but it can also mean that a lot of heterogeneity is created early on. For $\nu = 0.62$ Ga^{-1} and $\nu = 0.74$ Ga^{-1} to get lead-lead pseudo-isochron ages of 2.0 Ga the present melting time scale τ_{melt} must be 1.45 Ga and 1.83 Ga respectively.

4.6 Conclusions

The key results of this work are the simple pseudo-isochron equations (4.4) and (4.5) which relate in a straightforward way melting history and pseudo-isochron age. The equations are similar to those examined by previous authors but differ in certain important details. In the case of constant melt rate these relationships can expressed in terms of just two unknown parameters: the age τ_{s} at which you begin the model, and τ_{melt} the melting time scale. The natural choice for τ_{s} is the age of the Earth, but the unmodelled process of continent formation motivates younger choices for τ_{s} such as 3.6 Ga [Christensen and Hofmann, 1994]. To produce the observed lead-lead pseudo-isochron ages of 2.0 Ga values of $\tau_{\text{melt}} = 0.45$ Ga (for $\tau_{\text{s}} = 4.5$ Ga) or $\tau_{\text{melt}} = 0.52$ Ga (for $\tau_{\text{s}} = 3.6$ Ga) are needed. If the melt rate is not constant but is instead greater in the past, then slower present day melting time scales may be needed, for example $\tau_{\text{melt}} = 1.45$ Ga found earlier when the rate of melting was proportional to the square of mantle heating rate. Figure 4.8 plots the corresponding probability density functions for a constant melt rate case and a variable melt rate case.

There is quantitative agreement between the predictions of this simple statistical box model and the much more complicated mantle convection calculations that have been done by previous authors. This helps justify some of the key assumptions that are made in the statistical box model, and in particular the neglect of the particular details of the underlying flow. Statistical box models are still in their infancy, but they seem to provide a powerful way of approaching the problem of isotopic heterogeneity in the mantle. A great advantage is that general analytical results can derived which can lead to better understanding of the problem without the need to run lots of numerical simulations. However, mantle convection simulations still have an important role to play; of most interest for the pseudo-isochron problem is knowing the kinds of melting history that they can produce. Future simulations should investigate further the validity of these
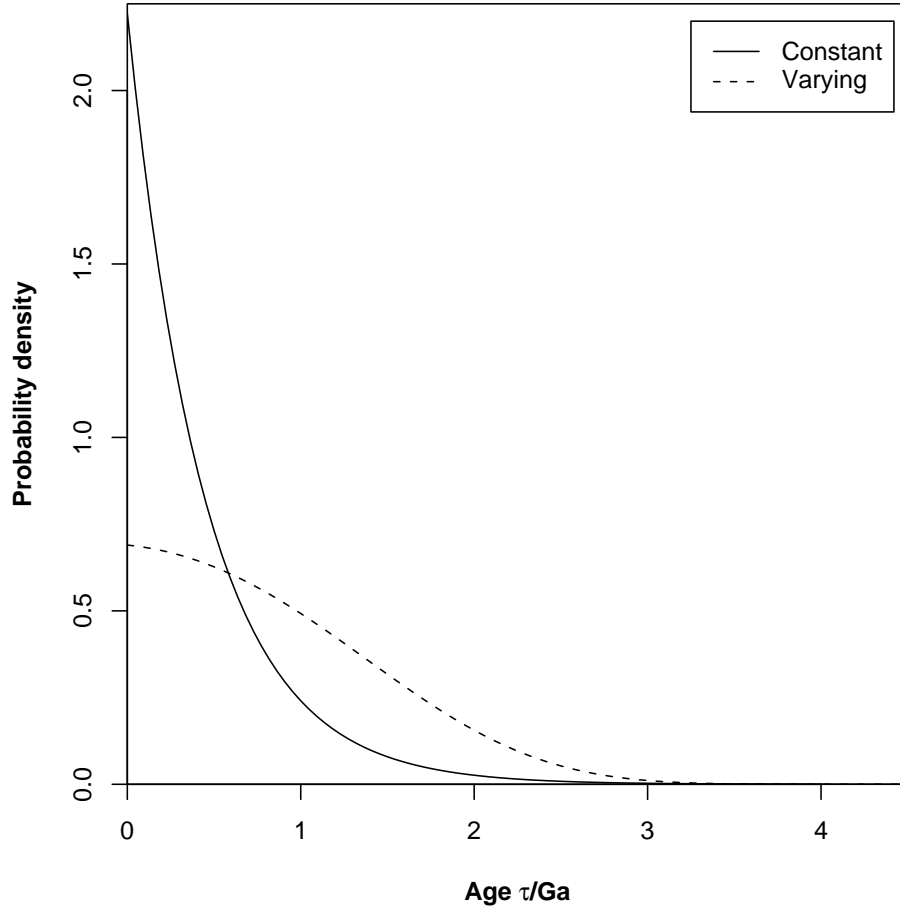
Figure 4.8: Plot of probability density function $q_{\mathrm{m}}(\tau)$. Two cases are shown which both yield a lead-lead pseudo-isochron age of 2.0 Ga. The solid line has constant melt rate $\gamma_{\mathrm{melt}}(\tau) = 1/\tau_{\mathrm{melt}}$ where $\tau_{\mathrm{melt}} = 0.45$ Ga. The dashed line has varying melt rate $\gamma_{\mathrm{melt}}(\tau) = \mathrm{e}^{\nu\tau}/\tau_{\mathrm{melt}}$ where $\nu = 0.62$ Ga^{-1} and $\tau_{\mathrm{melt}} = 1.45$ Ga. While the lead-lead ages are identical for the two cases, the parent-daughter and mean ages are not. Sm-Nd $\tau_{\mathrm{pdi}} = 0.64$ Ga and 1.10 Ga, and $\bar{\tau}_{\mathrm{m}} = 0.45$ Ga and 0.90 Ga respectively. Note that in both cases the very old (> 3.0 Ga) heterogeneity has been essentially eliminated.

simple analytic relationships between melt rate and age distribution, and between age distribution and pseudo-isochron age.

   An important caveat to bear in mind when using this model, and indeed in the convection simulations as well, is that no attempt is made to model continent formation. This is an important process to consider as early heterogeneity may become locked up in the continents and not be recycled into the mantle. It is this idea which prompts the use of younger values of τ_s. The multi-reservoir statistical box models of Kellogg [2004] have made some attempt at addressing this issue, although as is remarked it is somewhat difficult to see through the myriad of parameters used in that model. A key problem in modelling continent formation is that it requires modelling properly the mean lead isotopic evolution of mantle, which in itself is still not clearly understood. There is still yet to be any consensus on the resolution of the various lead paradoxes.

   For isotopic ratios d/d' for which the decay is linearisable the outcome of the model we present here is little changed from that described in the previous chapters. In particular the slopes in plots of one isotopic system against another are unchanged (2.6). There thus remains the fundamental inconsistency between the Pb isotopes and the Nd, Sr and Hf isotopes. Furthermore, the value of $\tau_{\mathrm{melt}} \sim 0.5$ Ga inferred here from the lead-lead pseudo-isochron is a factor of 3 different from the value of $\tau_{\mathrm{melt}} \sim 1.7$ Ga inferred from the standard deviations of the Nd, Sr and Hf isotopic ratios in chapter 2. In fact, $\tau_{\mathrm{melt}} \sim 0.5$ Ga provides a much better match to the observed standard deviations of the Pb isotopic ratios than $\tau_{\mathrm{melt}} \sim 1.7$ Ga (in chapter 2, model lead isotopic standard deviations were around 3 times larger than observed). Note that the pseudo-isochron is a much more robust way of estimating τ_{melt} as it depends only on the decay constants and start age, whereas the standard deviation of isotopic ratios depends on all the parameters in the model. It is still not possible in the model as it stands to match both the Pb isotopes and the Nd, Sr and Hf isotopes with a single set of parameters.

Chapter 5

Pseudo-isochron derivations

This chapter presents the mathematical derivations that lead to the pseudo-isochron equations discussed in the previous chapter.

5.1 Statistical box model

The simple statistical box model described in chapter 3 now has three restrictions lifted: the decay is no longer linearised, a steady state is not assumed, and the melt rate may vary with time. There are now three types of parcel in our model mantle: melt parcels (parcels last formed as melt in the melting region), residue parcels (parcels last formed as residue in the melting region), and primordial parcels. Melt and residue parcels have an associated age τ of when they last visited the melting region. The distribution of parcel ages is given by a random variable $\hat{T}$, which is defined so that those parcels of age $\geq \tau_s$ are primordial. Let $q(\tau)$ be the probability density function of $\hat{T}$. Thus a fraction $1 - \int_0^{\tau_s} q(\tau) \, d\tau$ of parcels are primordial.

We track the concentrations of a parent isotope p, a daughter d, and a reference d'. G is the molar fraction of a chemical species that enters the melt, and depends on the melt fraction F and the partition coefficients (3.17). The corresponding concentrations of the three parcel types at the present day are (from (3.27-3.32)):

$$p_{\text{prim}} = \bar{p}, \tag{5.1}$$

$$d_{\text{prim}} = \bar{d}, \tag{5.2}$$

$$d'_{\text{prim}} = \bar{d}'. \tag{5.3}$$

$$p_{\text{melt}}(\tau) = \frac{G_p}{F} \bar{p}, \tag{5.4}$$

$$d_{\text{melt}}(\tau) = \frac{G_d}{F} \bar{d} + \frac{G_p - G_d}{F} \left(e^{\lambda \tau} - 1 \right) \bar{p}, \tag{5.5}$$

$$d'_{\text{melt}}(\tau) = \frac{G_d}{F} \bar{d}'. \tag{5.6}$$

$$p_{\text{res}}(\tau) = \frac{1 - G_p}{1 - F} \bar{p}, \tag{5.7}$$

$$d_{\text{res}}(\tau) = \frac{1 - G_d}{1 - F} \bar{d} - \frac{G_p - G_d}{1 - F} \left(e^{\lambda \tau} - 1 \right) \bar{p}, \tag{5.8}$$

$$d'_{\text{res}}(\tau) = \frac{1 - G_d}{1 - F} \bar{d}', \tag{5.9}$$

56

where $\bar{p}$, $\bar{d}$, and $\bar{d}'$ are the mean concentrations of the isotopes in the box at the present day, and λ is the decay constant.

5.2 Calculating the asymptotic moments

Based on the general analysis of section 3.3.1, introduce new variables $p^\star$, $d^\star$ and $d'^\star$ defined by:

$$p^\star = \frac{1}{\bar{d}'}\left(p - \frac{\bar{p}}{\bar{d}'}d'\right), \tag{5.10}$$

$$d^\star = \frac{1}{\bar{d}'}\left(d - \frac{\bar{d}}{\bar{d}'}d'\right), \tag{5.11}$$

$$d'^\star = \frac{1}{\bar{d}'}\left(d' - \bar{d}'\right). \tag{5.12}$$

In these new variables the different parcel types are given by

$$p^\star_{\text{prim}} = 0, \tag{5.13}$$

$$d^\star_{\text{prim}} = 0, \tag{5.14}$$

$$d'^\star_{\text{prim}} = 0. \tag{5.15}$$

$$p^\star_{\text{melt}}(\tau) = \frac{\bar{p}}{\bar{d}'}\frac{G_p - G_d}{F}, \tag{5.16}$$

$$d^\star_{\text{melt}}(\tau) = \frac{\bar{p}}{\bar{d}'}\frac{G_p - G_d}{F}\left(e^{\lambda\tau} - 1\right), \tag{5.17}$$

$$d'^\star_{\text{melt}}(\tau) = \frac{G_d - F}{F}. \tag{5.18}$$

$$p^\star_{\text{res}}(\tau) = -\frac{\bar{p}}{\bar{d}'}\frac{G_p - G_d}{1 - F}, \tag{5.19}$$

$$d^\star_{\text{res}}(\tau) = -\frac{\bar{p}}{\bar{d}'}\frac{G_p - G_d}{1 - F}\left(e^{\lambda\tau} - 1\right), \tag{5.20}$$

$$d'^\star_{\text{res}}(\tau) = -\frac{G_d - F}{1 - F}. \tag{5.21}$$

Key moments are then

$$\mathbb{E}(\hat{p}^{\star 2}) = \left(\frac{\bar{p}}{\bar{d}'}\right)^2 \frac{(G_p - G_d)^2}{F(1 - F)}\int_0^{\tau_s} q(\tau)\,\mathrm{d}\tau, \tag{5.22}$$

$$\mathbb{E}(\hat{d}^\star \hat{p}^\star) = \left(\frac{\bar{p}}{\bar{d}'}\right)^2 \frac{(G_p - G_d)^2}{F(1 - F)}\int_0^{\tau_s}\left(e^{\lambda\tau} - 1\right) q(\tau)\,\mathrm{d}\tau, \tag{5.23}$$

$$\mathbb{E}(\hat{d}^{\star 2}) = \left(\frac{\bar{p}}{\bar{d}'}\right)^2 \frac{(G_p - G_d)^2}{F(1 - F)}\int_0^{\tau_s}\left(e^{\lambda\tau} - 1\right)^2 q(\tau)\,\mathrm{d}\tau. \tag{5.24}$$

Let $\hat{Z}_d = \sum_i \hat{d}_i / \sum_i \hat{d}'_i$ be the random variable giving the distribution of isotopic ratios d/d' after averaging, and $\hat{Z}_p = \sum_i \hat{p}_i / \sum_i \hat{d}'_i$ be the random variable giving the distribution of isotopic ratios p/d' after averaging. From (3.68) we have that

$$\operatorname{var}\hat{Z}_d \sim \frac{1}{N}\mathbb{E}\left(\hat{d}^{\star 2}\right) = \left(\frac{\bar{p}}{\bar{d}'}\right)^2 \frac{(G_p - G_d)^2}{N F(1 - F)}\int_0^{\tau_s}\left(e^{\lambda\tau} - 1\right)^2 q(\tau)\,\mathrm{d}\tau, \tag{5.25}$$

$$\operatorname{var}\hat{Z}_p \sim \frac{1}{N}\mathbb{E}\left(\hat{p}^{\star 2}\right) = \left(\frac{\bar{p}}{\bar{d}'}\right)^2 \frac{(G_p - G_d)^2}{N F(1 - F)}\int_0^{\tau_s} q(\tau)\,\mathrm{d}\tau. \tag{5.26}$$

(5.25) is the generalisation of (3.76). A similar generalisation for the skew (3.80) can obtained using (3.70).

5.3 Parent-daughter isochrons

It is convenient to introduce $\hat{T}_{\mathrm{m}}$: this is the random variable giving the distribution of parcel ages restricted to those parcels that have passed through the melting region. $q_{\mathrm{m}}(\tau) = q(\tau) / \int_0^{\tau_{\mathrm{s}}} q(\tau)\, \mathrm{d}\tau$ is then the probability density function of $\hat{T}_{\mathrm{m}}$. Suppose we plot d/d' against p/d'. Then the correlation is estimated by (3.72) as

$$r_{pd} \sim \frac{\mathbb{E}(\hat{d}^{\star}\hat{p}^{\star})}{\sqrt{\mathbb{E}(\hat{d}^{\star 2})\mathbb{E}(\hat{p}^{\star 2})}} = \frac{\mathbb{E}(\mathrm{e}^{\lambda \hat{T}_{\mathrm{m}}} - 1)}{\sqrt{\mathbb{E}\left(\mathrm{e}^{\lambda \hat{T}_{\mathrm{m}}} - 1\right)^2}}. \tag{5.27}$$

The slopes of the regression lines are estimated by (3.73), (3.74), and (3.75) as

$$\beta_{pd} \sim \frac{\mathbb{E}(\hat{d}^{\star}\hat{p}^{\star})}{\mathbb{E}(\hat{p}^{\star 2})} = \mathbb{E}\left(\mathrm{e}^{\lambda \hat{T}_{\mathrm{m}}} - 1\right), \tag{5.28}$$

$$\bar{\beta} \sim \sqrt{\frac{\mathbb{E}(\hat{d}^{\star 2})}{\mathbb{E}(\hat{p}^{\star 2})}} = \sqrt{\mathbb{E}\left(\mathrm{e}^{\lambda \hat{T}_{\mathrm{m}}} - 1\right)^2}, \tag{5.29}$$

$$\beta_{dp} \sim \frac{\mathbb{E}(\hat{d}^{\star 2})}{\mathbb{E}(\hat{d}^{\star}\hat{p}^{\star})} = \frac{\mathbb{E}\left(\mathrm{e}^{\lambda \hat{T}_{\mathrm{m}}} - 1\right)^2}{\mathbb{E}\left(\mathrm{e}^{\lambda \hat{T}_{\mathrm{m}}} - 1\right)}. \tag{5.30}$$

The parent-daughter pseudo-isochron age is related to the slope of the regression line by

$$\mathrm{e}^{\lambda \tau_{\mathrm{pdi}}} - 1 = \beta, \tag{5.31}$$

and thus using the geometric mean regression line (5.29) the model parent-daughter pseudo-isochron equation is

$$\left(\mathrm{e}^{\lambda \tau_{\mathrm{pdi}}} - 1\right)^2 = \mathbb{E}\left(\mathrm{e}^{\lambda \hat{T}_{\mathrm{m}}} - 1\right)^2. \tag{5.32}$$

Note that when the decay is linearisable ($\hat{T}_{\mathrm{m}} \ll 1/\lambda$) (5.27) and (5.32) reduce to $r_{pd} = \mathbb{E}\hat{T}_{\mathrm{m}}/\sqrt{\mathbb{E}\hat{T}_{\mathrm{m}}^2}$ and $\tau_{\mathrm{pdi}} = \sqrt{\mathbb{E}\hat{T}_{\mathrm{m}}^2}$.

5.4 Daughter-daughter isochrons

Now suppose we plot $(d/d')_2$ against $(d/d')_1$ for two different isotopic systems 1 and 2. Key moments are then

$$\mathbb{E}(\hat{d}_1^{\star 2}) = \left(\frac{\bar{p}_1}{\bar{d}_1'}\right)^2 \frac{(G_{p1} - G_{d1})^2}{F(1 - F)} \int_0^{\tau_{\mathrm{s}}} \left(\mathrm{e}^{\lambda_1 \tau} - 1\right)^2 q(\tau)\, \mathrm{d}\tau, \tag{5.33}$$

$$\mathbb{E}(\hat{d}_1^{\star}\hat{d}_2^{\star}) = \left(\frac{\bar{p}_1}{\bar{d}_1'}\right)\left(\frac{\bar{p}_2}{\bar{d}_2'}\right) \frac{(G_{p1} - G_{d1})(G_{p2} - G_{d2})}{F(1 - F)} \int_0^{\tau_{\mathrm{s}}} \left(\mathrm{e}^{\lambda_1 \tau} - 1\right)\left(\mathrm{e}^{\lambda_2 \tau} - 1\right) q(\tau)\, \mathrm{d}\tau, \tag{5.34}$$

$$\mathbb{E}(\hat{d}_2^{\star 2}) = \left(\frac{\bar{p}_2}{\bar{d}_2'}\right)^2 \frac{(G_{p2} - G_{d2})^2}{F(1 - F)} \int_0^{\tau_{\mathrm{s}}} \left(\mathrm{e}^{\lambda_2 \tau} - 1\right)^2 q(\tau)\, \mathrm{d}\tau. \tag{5.35}$$

The correlation is then given by

$$r_{12} \sim \frac{\mathbb{E}(e^{\lambda_1 \hat{T}_m} - 1)(e^{\lambda_2 \hat{T}_m} - 1)}{\sqrt{\mathbb{E}\left(e^{\lambda_1 \hat{T}_m} - 1\right)^2 \mathbb{E}\left(e^{\lambda_2 \hat{T}_m} - 1\right)^2}} \operatorname{sgn}\left((G_{p1} - G_{d1})(G_{p2} - G_{d2})\right). \tag{5.36}$$

Note that $|r_{12}| \sim 1$ if the decay is linearisable ($\hat{T}_m \ll 1/\lambda_1$ and $1/\lambda_2$), which is why in Figure 2.4 the model data form almost perfect straight lines. However, when the decay is not linearisable we will not get perfect correlation. The slope of the geometric mean regression line is given by

$$\bar{\beta} \sim \frac{(\bar{p}_2/\bar{d}_2')(G_{p2} - G_{d2})}{(\bar{p}_1/\bar{d}_1')(G_{p1} - G_{d1})} \sqrt{\frac{\mathbb{E}\left(e^{\lambda_2 \hat{T}_m} - 1\right)^2}{\mathbb{E}\left(e^{\lambda_1 \hat{T}_m} - 1\right)^2}}. \tag{5.37}$$

Note that if the decay is linearisable this reduces to

$$\bar{\beta} \sim \frac{(\bar{p}_2/\bar{d}_2')(G_{p2} - G_{d2})\lambda_2}{(\bar{p}_1/\bar{d}_1')(G_{p1} - G_{d1})\lambda_1}, \tag{5.38}$$

which is precisely (3.91). (5.37) is the generalisation of (3.91). Hence for those isotopic systems for which a linear decay approximation is valid the slopes are unchanged.

We are particularly interested in a special case of (5.37). If the parent and daughter elements are the same, and the reference isotope is also the same, then (5.37) reduces to

$$\bar{\beta} \sim \frac{\bar{p}_2}{\bar{p}_1} \sqrt{\frac{\mathbb{E}\left(e^{\lambda_2 \hat{T}_m} - 1\right)^2}{\mathbb{E}\left(e^{\lambda_1 \hat{T}_m} - 1\right)^2}}. \tag{5.39}$$

This is the generalisation of (3.92). The other estimates of the slope of the regression line in this special case are given by

$$\beta_{12} \sim \frac{\bar{p}_2}{\bar{p}_1} \cdot \frac{\mathbb{E}(e^{\lambda_1 \hat{T}_m} - 1)(e^{\lambda_2 \hat{T}_m} - 1)}{\mathbb{E}\left(e^{\lambda_1 \hat{T}_m} - 1\right)^2}, \quad \beta_{21} \sim \frac{\bar{p}_2}{\bar{p}_1} \cdot \frac{\mathbb{E}\left(e^{\lambda_2 \hat{T}_m} - 1\right)^2}{\mathbb{E}(e^{\lambda_1 \hat{T}_m} - 1)(e^{\lambda_2 \hat{T}_m} - 1)}. \tag{5.40}$$

The daughter-daughter pseudo-isochron age is related to the slope of the regression lines by

$$\frac{\bar{p}_2}{\bar{p}_1} \cdot \frac{e^{\lambda_2 \tau_{ddi}} - 1}{e^{\lambda_1 \tau_{ddi}} - 1} = \beta. \tag{5.41}$$

Hence, combining (5.39) and (5.41) the model pseudo-isochron age τ_{ddi} satisfies the simple relationship

$$\frac{(e^{\lambda_2 \tau_{ddi}} - 1)^2}{(e^{\lambda_1 \tau_{ddi}} - 1)^2} = \frac{\mathbb{E}(e^{\lambda_2 \hat{T}_m} - 1)^2}{\mathbb{E}(e^{\lambda_1 \hat{T}_m} - 1)^2}. \tag{5.42}$$

Note that when the decay is linearisable ($\hat{T}_m \ll 1/\lambda_1$ and $1/\lambda_2$) (5.42) reduces to $\tau_{ddi} = \mathbb{E}\hat{T}_m^3/\mathbb{E}\hat{T}_m^2$ (by Taylor series expansion).

5.5 Means

It is important to distinguish between different definitions of the mean age of parcels. Since the heterogeneity we are interested in is generated by fractionation on melting, an important mean age is $\bar{\tau}_{\mathrm{m}} = \mathbb{E}\hat{T}_{\mathrm{m}}$, the mean age of the parcels that have passed through the melting region. Primordial parcels do not contribute to the pseudo-isochron ages. The mean mantle age $\bar{\tau}_{\mathrm{total}}$ is often defined by including the primordial parcels and assigning them an age of τ_{s}, and it is this definition that is used in Xie and Tackley [2004a]. Hence $\bar{\tau}_{\mathrm{total}} \geq \bar{\tau}_{\mathrm{m}}$. We have

$$\bar{\tau}_{\mathrm{total}} = \int_0^{\tau_{\mathrm{s}}} \tau q(\tau)\,\mathrm{d}\tau + \tau_{\mathrm{s}}\left(1 - \int_0^{\tau_{\mathrm{s}}} q(\tau)\,\mathrm{d}\tau\right), \tag{5.43}$$

$$\bar{\tau}_{\mathrm{m}} = \mathbb{E}\hat{T}_{\mathrm{m}} = \int_0^{\tau_{\mathrm{s}}} \tau q_{\mathrm{m}}(\tau)\,\mathrm{d}\tau. \tag{5.44}$$

The parent-daughter pseudo-isochron age τ_{pdi} is an example of a generalised mean [Hardy et al., 1934]. A generalised mean M_ϕ is defined by $M_\phi(\hat{X}) = \phi^{-1}\left(\mathbb{E}\phi(\hat{X})\right)$, where ϕ is a strictly monotonic function, and $\hat{X}$ is a random variable. Commonly encountered examples include $\phi(x) = x$ (arithmetic mean), $\phi(x) = x^r$ (power mean), and $\phi(x) = \log x$ (geometric mean). In the case of the parent-daughter pseudo-isochron (5.32), $\phi(x) = \left(e^{\lambda x} - 1\right)^2$. Generalised means have a number of important properties, but of most interest to us is the notion of 'comparability': whether there is always an inequality between different means regardless of the distribution of the random variable $\hat{X}$. A common example of comparability is the arithmetic mean - geometric mean inequality. There is an important theorem which states that if ψ and χ are monotonically increasing functions, $\phi = \chi\psi^{-1}$, and $\phi'' > 0$, then $M_\psi \leq M_\chi$ (Theorem 96 of Hardy et al. [1934]). By use of this theorem we find the following inequalities are satisfied by the parent-daughter pseudo-isochron age:

$$\min\hat{T}_{\mathrm{m}} \leq \bar{\tau}_{\mathrm{m}} \leq \sqrt{\mathbb{E}\hat{T}_{\mathrm{m}}^2} \leq \tau_{\mathrm{pdi}}(\lambda_1) \leq \tau_{\mathrm{pdi}}(\lambda_2) \leq \max\hat{T}_{\mathrm{m}}, \tag{5.45}$$

where $0 < \lambda_1 < \lambda_2$, and $\min\hat{T}_{\mathrm{m}}$ and $\max\hat{T}_{\mathrm{m}}$ are the smallest and largest ages respectively with any probability mass.

The daughter-daughter pseudo-isochron age τ_{ddi} is not a generalised mean as described by Hardy et al. [1934], but it is an example of a generalised abstracted mean (Definition 2.4 of Qi [2000]). A generalised abstracted mean is defined by $M_{\phi_1,\phi_2}(\hat{X}) = (\phi_1/\phi_2)^{-1}\left(\mathbb{E}\phi_1(\hat{X})/\mathbb{E}\phi_2(\hat{X})\right)$ where ϕ_1/ϕ_2 is a strictly monotonic function. In the case of the daughter-daughter pseudo-isochron equation (5.42), $\phi_1(x) = \left(e^{\lambda_1 x} - 1\right)^2$ and $\phi_2(x) = \left(e^{\lambda_2 x} - 1\right)^2$. The generalised abstracted mean shares many of the properties of the generalised mean of Hardy et al. [1934], and under a suitable transformation of the probability density function can be written in the same form. However, for our purposes what is of most interest is the inequality

$$\tau_{\mathrm{pdi}}(\lambda_2) \leq \tau_{\mathrm{ddi}}(\lambda_1, \lambda_2) \leq \max\hat{T}_{\mathrm{m}}. \tag{5.46}$$

The above inequalities (5.45) and (5.46) are strict inequalities unless all the probability mass is concentrated at a single age. In this case all the inequalities are equalities, and all the means yield this single age.

5.6 Relating melt rate and parcel ages

Suppose melt rate as a function of age τ is $\gamma_{\text{melt}}(\tau)$. Define τ_{melt} to be the melting time scale at the present day, so that $\gamma_{\text{melt}}(0) = 1/\tau_{\text{melt}}$. Let $Q(\tau) = \mathbb{P}(\hat{T} \leq \tau)$ be the proportion of material in the box with age less than τ (the cumulative distribution function), with $1 - Q(\tau_{\text{s}})$ being the proportion of primordial material. $Q(\tau)$ satisfies

$$\frac{\mathrm{d}Q(\tau)}{\mathrm{d}\tau} = \gamma_{\text{melt}}(\tau)\left(1 - Q(\tau)\right), \quad Q(0) = 0, \tag{5.47}$$

and thus

$$Q(\tau) = 1 - \exp\left(-\int_0^\tau \gamma_{\text{melt}}(\tau)\,\mathrm{d}\tau\right). \tag{5.48}$$

Hence the probability density functions are

$$q(\tau) = \frac{\mathrm{d}Q(\tau)}{\mathrm{d}\tau} = \gamma_{\text{melt}}(\tau)\exp\left(-\int_0^\tau \gamma_{\text{melt}}(\tau)\,\mathrm{d}\tau\right), \tag{5.49}$$

$$q_{\mathrm{m}}(\tau) = \frac{q(\tau)}{Q(\tau_{\text{s}})} = \frac{1}{Q(\tau_{\text{s}})}\frac{\mathrm{d}Q(\tau)}{\mathrm{d}\tau}. \tag{5.50}$$

In some cases it is more convenient to work with the cumulative distribution function (cdf) rather than the pdf. Note that if $f(0) = 0$ (as it is for all functions we consider) then

$$\mathbb{E}f(\hat{T}_{\mathrm{m}}) = \int_0^{\tau_{\text{s}}} f'(\tau)(1 - Q_m(\tau))\,\mathrm{d}\tau, \tag{5.51}$$

where the cdf $Q_m(\tau)$ is given by

$$Q_m(\tau) = \int_0^\tau q_m(\tau)\,\mathrm{d}\tau = \frac{Q(\tau)}{Q(\tau_{\text{s}})}. \tag{5.52}$$

Xie and Tackley [2004a] quote results from their numerical simulations in terms of the crustal production rate given by $b\gamma_{\text{melt}}(\tau)$, where b is their basalt fraction ($b = 0.3$ in their standard runs). Their integrated crustal production (integrated forward in time) is given as a function of age by

$$c(\tau) = 1 - \exp\left(-\int_\tau^{\tau_{\text{s}}} b\gamma_{\text{melt}}(\tau)\,\mathrm{d}\tau\right), \tag{5.53}$$

which implies

$$\exp\left(-\int_0^\tau \gamma_{\text{melt}}(\tau)\,\mathrm{d}\tau\right) = \left(\frac{1 - c(0)}{1 - c(\tau)}\right)^{1/b}. \tag{5.54}$$

Hence the cdf $Q_m(\tau)$ (and thus also pseudo-isochron ages) can be calculated directly from their integrated crustal production or their crustal production rate.

For a constant melt rate $\gamma_{\text{melt}}(\tau) = 1/\tau_{\text{melt}}$ and $\hat{T}$ is an exponential random variable with parameter $1/\tau_{\text{melt}}$. The corresponding probability density functions are

$$q(\tau) = \mathrm{e}^{-\tau/\tau_{\text{melt}}}/\tau_{\text{melt}}, \tag{5.55}$$

$$q_{\mathrm{m}}(\tau) = \mathrm{e}^{-\tau/\tau_{\text{melt}}}/\left(\tau_{\text{melt}}\left(1 - \mathrm{e}^{-\tau_{\text{s}}/\tau_{\text{melt}}}\right)\right), \tag{5.56}$$

and corresponding means are

$$\bar{\tau}_{\text{total}} = \tau_{\text{melt}}(1 - e^{-\tau_s/\tau_{\text{melt}}}), \tag{5.57}$$

$$\bar{\tau}_{\text{m}} = \mathbb{E}\hat{T}_{\text{m}} = \tau_{\text{melt}} \frac{1 - e^{-\tau_s/\tau_{\text{melt}}}(1 + \tau_s/\tau_{\text{melt}})}{1 - e^{-\tau_s/\tau_{\text{melt}}}}, \tag{5.58}$$

$$\sqrt{\mathbb{E}\hat{T}_{\text{m}}^2} = \sqrt{2}\tau_{\text{melt}} \sqrt{\frac{1 - e^{-\tau_s/\tau_{\text{melt}}}\left(1 + \tau_s/\tau_{\text{melt}} + \frac{1}{2}\left(\tau_s/\tau_{\text{melt}}\right)^2\right)}{1 - e^{-\tau_s/\tau_{\text{melt}}}}}. \tag{5.59}$$

5.7 Constant melt rate: Parent-daughter isochrons

For constant melt rate the parent-daughter pseudo-isochron equation is

$$(e^{\lambda\tau_{\text{pdi}}} - 1)^2 = \frac{\int_0^{\tau_s}(e^{\lambda\tau} - 1)^2 e^{-\tau/\tau_{\text{melt}}}\,d\tau}{\tau_{\text{melt}}\left(1 - e^{-\tau_s/\tau_{\text{melt}}}\right)}. \tag{5.60}$$

To gain some insight into the behaviour of this equation we now consider some simple asymptotics.

5.7.1 Asymptotics when τ_{melt} or $\tau_s \ll 1/\lambda$

Since we often study slowly decaying isotopes, the most important asymptotics are when τ_{melt} or $\tau_s \ll 1/\lambda$. In this limit the correlation becomes

$$r_{pd} = \frac{1 - e^{-\tau_s/\tau_{\text{melt}}}(1 + \tau_s/\tau_{\text{melt}})}{\sqrt{2\left(1 - e^{-\tau_s/\tau_{\text{melt}}}\right)\left(1 - e^{-\tau_s/\tau_{\text{melt}}}\left(1 + \tau_s/\tau_{\text{melt}} + \frac{1}{2}\left(\tau_s/\tau_{\text{melt}}\right)^2\right)\right)}}. \tag{5.61}$$

Furthermore, if $\tau_s/\tau_{\text{melt}} \gg 1$ then $r_{pd} \sim 1/\sqrt{2} \approx 0.71$. Alternatively if $\tau_s/\tau_{\text{melt}} \ll 1$ then $r_{pd} \sim \sqrt{3}/2 \approx 0.87$. For linearisable decay r_{pd} always lies between these two values. It is important to note that there is thus never a perfect correlation between d/d' and p/d'.

When τ_{melt} or $\tau_s \ll 1/\lambda$ the parent-daughter pseudo-isochron age becomes simply $\tau_{\text{pdi}} = \sqrt{\mathbb{E}\hat{T}_{\text{m}}^2}$, which is given for constant melt rate by (5.59). If $\tau_s/\tau_{\text{melt}} \gg 1$ (5.59) simplifies to $\tau_{\text{pdi}} \sim \sqrt{2}\tau_{\text{melt}}$, which determines the slope of curves near the origin in Figures 4.3b and 4.4b, and the asymptotes for large τ_s in Figure 4.3a. On the other hand, if $\tau_s/\tau_{\text{melt}} \ll 1$ (5.59) simplifies to $\tau_{\text{pdi}} \sim \tau_s/\sqrt{3}$, which determines the slope of curves near the origin in Figures 4.3a and 4.4a, and the asymptotes for large τ_{melt} in Figure 4.3b.

5.7.2 Asymptotics when $\tau_s \ll \tau_{\text{melt}}$

When the decay is not linearisable, asymptotics based solely on $\tau_s \ll \tau_{\text{melt}}$ can be found. The pseudo-isochron equation (5.60) becomes

$$(e^{\lambda\tau_{\text{pdi}}} - 1)^2 = \frac{2\lambda\tau_s + 3 - 4e^{\lambda\tau_s} + e^{2\lambda\tau_s}}{2\lambda\tau_s}, \tag{5.62}$$

which is independent of τ_{melt}. This equation determines the $\tau_{\text{melt}} = \infty$ Ga curve in Figures 4.3a and 4.4a, and asymptotes for large τ_{melt} in Figures 4.3b and 4.4b.

5.7.3 Asymptotics when $\tau_s \gg \tau_{\text{melt}}$

(5.60) has two regimes of asymptotic behaviour when $\tau_{\text{s}}/\tau_{\text{melt}} \gg 1$. If $\lambda\tau_{\text{melt}} > 1/2$ then

$$e^{2\lambda\tau_{\text{pdi}}} = \frac{e^{(2\lambda\tau_{\text{melt}}-1)\tau_{\text{s}}/\tau_{\text{melt}}}}{2\lambda\tau_{\text{melt}} - 1}, \tag{5.63}$$

whereas if $\lambda\tau_{\text{melt}} < 1/2$ then

$$(e^{\lambda\tau_{\text{pdi}}} - 1)^2 = \frac{2(\lambda\tau_{\text{melt}})^2}{(1 - \lambda\tau_{\text{melt}})(1 - 2\lambda\tau_{\text{melt}})}. \tag{5.64}$$

Note that (5.63) depends on τ_{s} whereas (5.64) is independent of τ_{s}. This is why in Figure 4.4a curves with $\tau_{\text{melt}} < 1/2\lambda$ flatten out for large τ_{s}, while curves with $\tau_{\text{melt}} > 1/2\lambda$ grow linearly for large τ_{s} with slope $1 - 1/2\lambda\tau_{\text{melt}}$.

5.8 Constant melt rate: Daughter-daughter isochrons

The constant melt rate daughter-daughter pseudo-isochron equation is

$$\frac{(e^{\lambda_2\tau_{\text{ddi}}} - 1)^2}{(e^{\lambda_1\tau_{\text{ddi}}} - 1)^2} = \frac{\int_0^{\tau_{\text{s}}}(e^{\lambda_2\tau} - 1)^2 e^{-\tau/\tau_{\text{melt}}}\,d\tau}{\int_0^{\tau_{\text{s}}}(e^{\lambda_1\tau} - 1)^2 e^{-\tau/\tau_{\text{melt}}}\,d\tau}. \tag{5.65}$$

Again, insights into the behaviour of (5.65) can be gained by some simple asymptotics, although linearising the decay is not as relevant here. We will assume without loss of generality that $\lambda_2 > \lambda_1$ for the subsequent asymptotics.

5.8.1 Asymptotics when $\tau_s \ll \tau_{\text{melt}}$

When $\tau_s \ll \tau_{\text{melt}}$ the pseudo-isochron equation (5.65) becomes

$$\frac{(e^{\lambda_2\tau_{\text{ddi}}} - 1)^2}{(e^{\lambda_1\tau_{\text{ddi}}} - 1)^2} = \frac{\lambda_1}{\lambda_2} \cdot \frac{2\lambda_2\tau_{\text{s}} + 3 - 4e^{\lambda_2\tau_{\text{s}}} + e^{2\lambda_2\tau_{\text{s}}}}{2\lambda_1\tau_{\text{s}} + 3 - 4e^{\lambda_1\tau_{\text{s}}} + e^{2\lambda_1\tau_{\text{s}}}}. \tag{5.66}$$

Note that this equation is independent of τ_{melt}. This equation determines the asymptotes for large τ_{melt} in Figure 4.5b, and the $\tau_{\text{melt}} = \infty$ Ga curve in Figure 4.5a. Furthermore, if it is also the case that $\lambda\tau_s \ll 1$, we can Taylor expand both sides to find

$$\frac{\lambda_2^2}{\lambda_1^2}\left(1 + (\lambda_2 - \lambda_1)\tau_{\text{ddi}} + ...\right) = \frac{\lambda_2^2}{\lambda_1^2}\left(1 + \frac{3(\lambda_2 - \lambda_1)}{4}\tau_s + ...\right), \tag{5.67}$$

and thus get the simple result that $\tau_{\text{ddi}} \sim \frac{3}{4}\tau_s$, which determines the slope of curves near the origin in Figure 4.5a. If instead $\lambda\tau_s \gg 1$ then we can approximate both sides as

$$e^{2(\lambda_2 - \lambda_1)\tau_{\text{ddi}}} = \frac{\lambda_1}{\lambda_2}e^{2(\lambda_2 - \lambda_1)\tau_{\text{s}}}, \tag{5.68}$$

which demonstrates that the slope of the $\tau_{\text{melt}} = \infty$ Ga curve in Figure 4.5a will approach 1 for large τ_s.

5.8.2 Asymptotics when $\tau_s \gg \tau_{\text{melt}}$

We have three different regimes of asymptotic behaviour for (5.65) when $\tau_s \gg \tau_{\text{melt}}$. If $\tau_{\text{melt}} > 1/2\lambda_1$ the pseudo-isochron equation is

$$\mathrm{e}^{2(\lambda_2-\lambda_1)\tau_{\text{ddi}}} = \mathrm{e}^{2(\lambda_2-\lambda_1)\tau_s}\frac{2\lambda_1\tau_{\text{melt}} - 1}{2\lambda_2\tau_{\text{melt}} - 1}, \tag{5.69}$$

if $1/2\lambda_2 < \tau_{\text{melt}} < 1/2\lambda_1$ then

$$\mathrm{e}^{2(\lambda_2-\lambda_1)\tau_{\text{ddi}}} = \frac{\mathrm{e}^{(2\lambda_2\tau_{\text{melt}}-1)\tau_s/\tau_{\text{melt}}}(1 - \lambda_1\tau_{\text{melt}})(1 - 2\lambda_1\tau_{\text{melt}})}{(\lambda_1\tau_{\text{melt}})^2(2\lambda_2\tau_{\text{melt}} - 1)}, \tag{5.70}$$

and if $\tau_{\text{melt}} < 1/2\lambda_2$ then

$$\frac{(\mathrm{e}^{\lambda_2\tau_{\text{ddi}}} - 1)^2}{(\mathrm{e}^{\lambda_1\tau_{\text{ddi}}} - 1)^2} = \frac{\lambda_2^2(1 - \lambda_1\tau_{\text{melt}})(1 - 2\lambda_1\tau_{\text{melt}})}{\lambda_1^2(1 - \lambda_2\tau_{\text{melt}})(1 - 2\lambda_2\tau_{\text{melt}})}. \tag{5.71}$$

The most important feature of (5.71) is that it is independent of τ_s, whereas in the other two asymptotic regimes there is a dependence on τ_s. This is why for values of $\tau_{\text{melt}} < 1/2\lambda_2$ in Figure 4.5a the curve flattens out for large τ_s whereas in the other regimes they grow linearly for large τ_s with slopes $(\lambda_2 - 1/2\tau_{\text{melt}})/(\lambda_2 - \lambda_1)$ and 1 respectively. (5.71) can be further simplified if $\lambda\tau_{\text{melt}} \ll 1$. Both sides can be Taylor expanded to yield

$$\frac{\lambda_2^2}{\lambda_1^2}\left(1 + (\lambda_2 - \lambda_1)\tau_{\text{ddi}} + ...\right) = \frac{\lambda_2^2}{\lambda_1^2}\left(1 + 3(\lambda_2 - \lambda_1)\tau_{\text{melt}} + ...\right), \tag{5.72}$$

leading to the simple result that $\tau_{\text{ddi}} \sim 3\tau_{\text{melt}}$, which determines the slope of curves near the origin in Figure 4.5b.

5.9 Relationship to linear evolution models

There is an important connection between our statistical box model and the pseudo-isochron equations derived from linear evolution models by previous authors [Albarède, 2001, Allègre and Lewin, 1995a, Donnelly et al., 2004]. Suppose we divide the box into two based on parcel type. Let all the residue parcels and a fraction $1-F$ of the primordial parcels be called reservoir 1, and all the melt parcels and a fraction F of the primordial parcels be reservoir 2. These two reservoirs do not change in size over time. Reservoir 1 is a fraction $1-F$ of the box, and reservoir 2 a fraction F. At an age τ_s before the present all parcels are primordial and thus both reservoirs have the same uniform isotopic concentrations. The mean concentrations of p, d and d' at the present in each reservoir can be calculated by integrating (5.1-5.9) over the age distribution of parcels. For the isochron calculation it is simplest to consider instead the $p^\star$ and $d^\star$ values. Let $\hat{p}_1^\star$ be the random variable giving the distribution of $p^\star$ in reservoir 1, and $\hat{d}_1^\star$, $\hat{p}_2^\star$ and $\hat{d}_2^\star$ be defined similarly. Note that the subscripts 1 and 2 now refer to the different reservoirs, rather than the different isotopic systems as in earlier sections. Integration of (5.13-5.21) yields

the mean values as

$$\mathbb{E}\hat{p}_1^\star = -\frac{\bar{p}}{\bar{d'}}\frac{G_p - G_d}{1 - F}\int_0^{\tau_s} q(\tau)\,\mathrm{d}\tau, \tag{5.73}$$

$$\mathbb{E}\hat{d}_1^\star = -\frac{\bar{p}}{\bar{d'}}\frac{G_p - G_d}{1 - F}\int_0^{\tau_s}\left(\mathrm{e}^{\lambda\tau} - 1\right) q(\tau)\,\mathrm{d}\tau, \tag{5.74}$$

$$\mathbb{E}\hat{p}_2^\star = \frac{\bar{p}}{\bar{d'}}\frac{G_p - G_d}{F}\int_0^{\tau_s} q(\tau)\,\mathrm{d}\tau, \tag{5.75}$$

$$\mathbb{E}\hat{d}_2^\star = \frac{\bar{p}}{\bar{d'}}\frac{G_p - G_d}{F}\int_0^{\tau_s}\left(\mathrm{e}^{\lambda\tau} - 1\right) q(\tau)\,\mathrm{d}\tau. \tag{5.76}$$

By multiplying these expressions by the fraction of the box each reservoir occupies these can be converted into molar values $n^\star$ and $m^\star$ ((4.21) and (4.22)) as

$$n_1^\star = (1 - F)\,\mathbb{E}\hat{p}_1^\star = -\frac{\bar{p}}{\bar{d'}}\left(G_p - G_d\right)\int_0^{\tau_s} q(\tau)\,\mathrm{d}\tau, \tag{5.77}$$

$$m_1^\star = (1 - F)\,\mathbb{E}\hat{d}_1^\star = -\frac{\bar{p}}{\bar{d'}}\left(G_p - G_d\right)\int_0^{\tau_s}\left(\mathrm{e}^{\lambda\tau} - 1\right) q(\tau)\,\mathrm{d}\tau. \tag{5.78}$$

where $n_2^\star = -n_1^\star$ and $m_2^\star = -m_1^\star$. (5.77) and (5.78) are solutions of (4.25) and (4.26), and are the backward in time versions of (A1-4) and (A1-5) of Allègre and Lewin [1995a] with $\Delta\langle\alpha\rangle = 0$. The number of moles of each isotope species in the two reservoirs must be modelled by the linear evolution equations (4.10-4.15); the only question is finding expressions for the residence times. Here the relaxation times for parent and daughter are both given by $1/\gamma_{\mathrm{melt}}$, since they are both fractionated by the same melting process on the same timescale. The individual residence times for each reservoir are determined by the fraction of moles G that enter the melt for each chemical species as

$$\tau_1 = \frac{1}{G_p\gamma_{\mathrm{melt}}}, \quad \tau_2 = \frac{1}{(1 - G_p)\,\gamma_{\mathrm{melt}}}, \tag{5.79}$$

$$\theta_1 = \frac{1}{G_d\gamma_{\mathrm{melt}}}, \quad \theta_2 = \frac{1}{(1 - G_d)\,\gamma_{\mathrm{melt}}}. \tag{5.80}$$

Consider a pseudo-isochron defined by the where the two reservoirs and the whole box lie on the isochron diagram. The slope in the parent-daughter isochron diagram is given by

$$\frac{(m/s)_1 - (m/s)_{1+2}}{(n/s)_1 - (n/s)_{1+2}} = \frac{m_1^\star}{n_1^\star} = \frac{\int_0^{\tau_s}(\mathrm{e}^{\lambda\tau} - 1)q(\tau)\,\mathrm{d}\tau}{\int_0^{\tau_s} q(\tau)\,\mathrm{d}\tau} = \mathbb{E}\left(\mathrm{e}^{\lambda\hat{T}_m} - 1\right). \tag{5.81}$$

Hence the corresponding pseudo-isochron equations are

$$\mathrm{e}^{\lambda\tau_{\mathrm{pdil}}} - 1 = \mathbb{E}(\mathrm{e}^{\lambda\hat{T}_m} - 1), \tag{5.82}$$

$$\frac{\mathrm{e}^{\lambda_2\tau_{\mathrm{ddil}}} - 1}{\mathrm{e}^{\lambda_1\tau_{\mathrm{ddil}}} - 1} = \frac{\mathbb{E}(\mathrm{e}^{\lambda_2\hat{T}_m} - 1)}{\mathbb{E}(\mathrm{e}^{\lambda_1\hat{T}_m} - 1)}, \tag{5.83}$$

which are just (5.32) and (5.42) with squareds removed, and hence we will refer to these as the linear pseudo-isochron equations.

An alternative reservoir representation is to consider the box split into three reservoirs, one with all the melt parcels, one with all the residue parcels, and one with all the primordial parcels. These three reservoirs change in size over time, with the melt and residue reservoirs growing at the expense of the primordial reservoir. This is analogous to

model I of Jacobsen and Wasserburg [1979] where their depleted mantle reservoir 2 and crust reservoir 3 grow from an homogeneous undepleted mantle reservoir 1. Compare the statistical box model result, obtained by integrating (5.4-5.9),

$$\frac{\mathbb{E}\hat{d}_{\text{res}}}{\mathbb{E}\hat{d}'_{\text{res}}} - \frac{\bar{d}}{\bar{d}'} = -\frac{\bar{p}}{\bar{d}'}\frac{G_p - G_d}{1 - G_d}\mathbb{E}\left(\text{e}^{\lambda\hat{T}_{\text{m}}} - 1\right), \tag{5.84}$$

$$\frac{\mathbb{E}\hat{d}_{\text{melt}}}{\mathbb{E}\hat{d}'_{\text{melt}}} - \frac{\bar{d}}{\bar{d}'} = \frac{\bar{p}}{\bar{d}'}\frac{G_p - G_d}{G_d}\mathbb{E}\left(\text{e}^{\lambda\hat{T}_{\text{m}}} - 1\right), \tag{5.85}$$

with (19) of Jacobsen and Wasserburg [1979], rewritten using an integration by parts,

$$\frac{N_{d,j}(\tau)}{N_{s,j}(\tau)} - \frac{N_{d,1}(\tau)}{N_{s,1}(\tau)} = \frac{N_{r,1}(\tau)}{N_{s,1}(\tau)}f_j^{r/s}\int_0^\tau\left(\text{e}^{\lambda(\tau-\xi)} - 1\right)\frac{1}{M_j(\tau)}\frac{\text{d}M_j(\xi)}{\text{d}\xi}\,\text{d}\xi. \tag{5.86}$$

(5.84) and (5.85) are equivalent to (5.86), although note that (5.86) is written with τ running forward in time. The fractionation factors relate through $1+f_2 = (1-G_p)/(1-G_d)$ and $1 + f_3 = G_p/G_d$. The corresponding probability density function of parcels ages is $(1/M_j(\tau))\text{d}M_j(\xi)/\text{d}\xi$ (compare with (5.50)). Note that the relationship between melt rate and age distribution will be slightly different between the two models, since there is recycling in the statistical box model (a melt parcel may become a residue parcel and vice-versa), but not in model I. The pseudo-isochron equations that result from where the three reservoirs plot on the isochron diagram are also (5.82) and (5.83).

The linear parent-daughter pseudo-isochron age τ_{pdil} is also an example of a generalised mean, with $\phi(x) = \text{e}^{\lambda x} - 1$. The linear pseudo-isochron ages satisfy

$$\min\hat{T}_{\text{m}} \le \bar{\tau}_{\text{m}} \le \tau_{\text{pdil}}(\lambda_1) \le \tau_{\text{pdil}}(\lambda_2) \le \tau_{\text{ddil}}(\lambda_1, \lambda_2) \le \max\hat{T}_{\text{m}}, \tag{5.87}$$

where $0 < \lambda_1 < \lambda_2$. For linearisable decay ($\hat{T}_{\text{m}} \ll 1/\lambda$) the linear parent-daughter pseudo-isochron equation (5.82) reduces to $\tau_{\text{pdil}} = \mathbb{E}\hat{T}_{\text{m}} = \bar{\tau}_{\text{m}}$, and the linear daughter-daughter pseudo-isochron equation (5.83) to $\tau_{\text{ddil}} = \mathbb{E}\hat{T}_{\text{m}}^2/\mathbb{E}\hat{T}_{\text{m}}$. It can be shown that the linear ages are always less than the corresponding quadratic ages, namely that $\tau_{\text{pdil}}(\lambda) \le \tau_{\text{pdi}}(\lambda)$, and $\tau_{\text{ddil}}(\lambda_1, \lambda_2) \le \tau_{\text{ddi}}(\lambda_1, \lambda_2)$.

For the case of constant melt rate the linear parent-daughter pseudo-isochron equation is

$$\begin{aligned}
\text{e}^{\lambda\tau_{\text{pdil}}} - 1 &= \frac{\int_0^{\tau_{\text{s}}}(\text{e}^{\lambda\tau} - 1)\text{e}^{-\tau/\tau_{\text{melt}}}\,\text{d}\tau}{\tau_{\text{melt}}(1 - \text{e}^{-\tau_{\text{s}}/\tau_{\text{melt}}})} \\
&= \frac{\lambda\tau_{\text{melt}} + (1 - \lambda\tau_{\text{melt}})\text{e}^{-\tau_{\text{s}}/\tau_{\text{melt}}} - \text{e}^{(\lambda\tau_{\text{melt}}-1)\tau_{\text{s}}/\tau_{\text{melt}}}}{(1 - \lambda\tau_{\text{melt}})\left(1 - \text{e}^{-\tau_{\text{s}}/\tau_{\text{melt}}}\right)},
\end{aligned} \tag{5.88}$$

which is the parent-daughter pseudo-isochron equation (A2-4) of Allègre and Lewin [1995a]. This should be compared with our (5.60). The corresponding linear daughter-daughter pseudo-isochron equation is

$$\begin{aligned}
\frac{\text{e}^{\lambda_2\tau_{\text{ddil}}} - 1}{\text{e}^{\lambda_1\tau_{\text{ddil}}} - 1} &= \frac{\int_0^{\tau_{\text{s}}}(\text{e}^{\lambda_2\tau} - 1)\text{e}^{-\tau/\tau_{\text{melt}}}\,\text{d}\tau}{\int_0^{\tau_{\text{s}}}(\text{e}^{\lambda_1\tau} - 1)\text{e}^{-\tau/\tau_{\text{melt}}}\,\text{d}\tau} \\
&= \frac{1 - \lambda_1\tau_{\text{melt}}}{1 - \lambda_2\tau_{\text{melt}}}\cdot\frac{(1 - \lambda_2\tau_{\text{melt}})\text{e}^{-\tau_{\text{s}}/\tau_{\text{melt}}} + \lambda_2\tau_{\text{melt}} - \text{e}^{(\lambda_2\tau_{\text{melt}}-1)\tau_{\text{s}}/\tau_{\text{melt}}}}{(1 - \lambda_1\tau_{\text{melt}})\text{e}^{-\tau_{\text{s}}/\tau_{\text{melt}}} + \lambda_1\tau_{\text{melt}} - \text{e}^{(\lambda_1\tau_{\text{melt}}-1)\tau_{\text{s}}/\tau_{\text{melt}}}},
\end{aligned} \tag{5.89}$$

which is precisely the daughter-daughter pseudo-isochron equation (A2-9) of Allègre and Lewin [1995a]. This should be compared with our (5.65). Furthermore, when $\tau_{\text{melt}} \ll \tau_{\text{s}}$ and $\lambda\tau_{\text{melt}} < 1$ (5.88) and (5.89) become

$$e^{\lambda\tau_{\text{pdil}}} - 1 = \frac{\lambda\tau_{\text{melt}}}{1 - \lambda\tau_{\text{melt}}} \implies \tau_{\text{pdil}} = -\frac{1}{\lambda}\log\left(1 - \lambda\tau_{\text{melt}}\right), \tag{5.90}$$

$$\frac{e^{\lambda_2\tau_{\text{ddil}}} - 1}{e^{\lambda_1\tau_{\text{ddil}}} - 1} = \frac{\lambda_2\left(1 - \lambda_1\tau_{\text{melt}}\right)}{\lambda_1\left(1 - \lambda_2\tau_{\text{melt}}\right)}, \tag{5.91}$$

which are the pseudo-isochron equations (A2-6) and (A2-10) of Allègre and Lewin [1995a], (31) and (32) of Albarède [2001], and (7) and (9) of Donnelly et al. [2004]. The corresponding equations in our model are (5.64) and (5.71). The $\tau_{\text{melt}} \ll \tau_{\text{s}}$ and $\lambda\tau_{\text{melt}} < 1$ limit is precisely the same as the conditions for secular equilibrium in Albarède [2001] and Donnelly et al. [2004]. Allègre and Lewin [1995a] explores various other asymptotic regimes for (5.88) and (5.89), as we have for our constant melt rate equations in sections 5.7 and 5.8.

Chapter 6

Future generalisations

The statistical box model of chapters 2 to 5 is consistent with many features of the observed isotopic heterogeneity of MORB, but not all. This chapter examines where the model fails, and suggests useful future generalisations.

6.1 Decorrelations

The most striking discrepancy is that the present model has all the isotopic ratios correlating well with each other. In fact, in the large N limit the model predicts perfect correlation (except when the half-lives are short, which only significantly affects $^{207}\mathrm{Pb}/^{204}\mathrm{Pb}$). In contrast, the measured MORB data does not correlate perfectly, and there is a particularly noticeable decorrelation between the Pb isotopic ratios and the others. While there is a little decorrelation in the model due to finite N, it is not sufficient to account for the observed decorrelations.

The most natural way of causing decorrelation is through differences in the melting process. In the model as it stands, melting is highly simplified: there is just one melt region, with one melt fraction, and one set of partition coefficients. Real mantle melting takes place over a wide range of melt fractions, and in different melting environments. A simple generalisation of the model can be considered to investigate this.

Suppose the different melting environments are labelled by a number u, say between 0 and 1, and suppose there is a corresponding random variable $\hat{U}$ giving the probability of being in such an environment. Let the partition coefficients and melt fractions be functions of this random variable, so that we now have the functions $F(u)$, $G_p(u)$ and $G_d(u)$. Assuming that sampling of the different components is still unbiased, (5.34) generalises to

$$\mathbb{E}(\hat{d}_1^\star \hat{d}_2^\star) = \left(\frac{\bar{p}_1}{\bar{d}_1'}\right)\left(\frac{\bar{p}_2}{\bar{d}_2'}\right) \int_0^1 \frac{(G_{p1}(u) - G_{d1}(u))(G_{p2}(u) - G_{d2}(u))}{F(u)\,(1 - F(u))}\, q_{\hat{U}}(u)\,\mathrm{d}u$$
$$\times \int_0^{\tau_\mathrm{s}} \left(\mathrm{e}^{\lambda_1 \tau} - 1\right)\left(\mathrm{e}^{\lambda_2 \tau} - 1\right) q_{\hat{T}}(\tau)\,\mathrm{d}\tau, \qquad (6.1)$$

where $q_{\hat{U}}(u)$ is the probability density function of $\hat{U}$, and $q_{\hat{T}}(\tau)$ the probability density function of $\hat{T}$. The correlation r_{12} between two isotopic ratios d/d' can then be written for large N as

$$r_{12} \sim r_\mathrm{melt} r_\mathrm{decay} \qquad (6.2)$$

where

$$r_{\text{decay}} = \frac{\mathbb{E}(e^{\lambda_1 \hat{T}_{\text{m}}} - 1)(e^{\lambda_2 \hat{T}_{\text{m}}} - 1)}{\sqrt{\mathbb{E}\left(e^{\lambda_1 \hat{T}_{\text{m}}} - 1\right)^2 \mathbb{E}\left(e^{\lambda_2 \hat{T}_{\text{m}}} - 1\right)^2}}, \tag{6.3}$$

$$r_{\text{melt}} = \frac{\mathbb{E} f_1(\hat{U}) f_2(\hat{U})}{\sqrt{\mathbb{E}\left(f_1(\hat{U})\right)^2 \mathbb{E}\left(f_2(\hat{U})\right)^2}}, \tag{6.4}$$

and

$$f_1(u) = \frac{G_{p1}(u) - G_{d1}(u)}{\sqrt{F(u)\,(1 - F(u))}}, \tag{6.5}$$

$$f_2(u) = \frac{G_{p2}(u) - G_{d2}(u)}{\sqrt{F(u)\,(1 - F(u))}}. \tag{6.6}$$

(6.2) expresses the idea that decorrelation can come from two sources: either from the melting process, or from differing decay constants. For the long-lived isotopes $r_{\text{decay}} \sim 1$, and only the differences in melting can cause the decorrelation. The form of the functions $f_1(u)$ and $f_2(u)$ determine the amount of decorrelation. If $f_1(u)$ and $f_2(u)$ are the same or proportional, there will be perfect correlation. Generally, if the relative compatibilities are similar in behaviour then there will be good correlation.

It is straightforward to envisage a model situation where the lead isotopic ratios decorrelate from the others. Suppose we have just two melting environments. Suppose that in one environment Nd is fractionated from Sm, Hf from Lu, and Rb from Sr; but Th, U and Pb are not fractionated. In the other environment suppose we have the reverse: Th and U are fractionated from Pb, but the others are not fractionated. Suppose system 1 is ^{143}Nd/^{144}Nd and system 2 is ^{208}Pb/^{204}Pb. Then $f_2(u) = 0$ in the first environment, and $f_1(u) = 0$ in the second environment. Thus $f_1(u)f_2(u) \equiv 0$, and from (6.4) there is thus zero correlation between ^{143}Nd/^{144}Nd and ^{208}Pb/^{204}Pb. On the other hand, the correlation will be near perfect between ^{143}Nd/^{144}Nd, ^{87}Sr/^{86}Sr and ^{177}Hf/^{176}Hf. The correlation will also be very good between the Pb isotopic ratios, although some small decorrelation is expected due to the shorter half-lives.

This simple model situation suggests that the most likely explanation for the decorrelation of the lead isotopic ratios is that there is some process that fractionates Pb from U and Th, but does not fractionate the other parents from their daughters. The question is now a geochemical one, as to what this process might be. The geochemical behaviour of lead is still not fully understood, and it is thought that some of the strange behaviour is due to partitioning in sulphide phases.

Note that this generalisation does not change the pseudo-isochron results, and the pseudo-isochron ages will be still be given by the same expressions in terms of $\hat{T}_{\text{m}}$. However, more care is needed in interpreting the age distribution. It is best interpreted in terms of a time scale for the destruction of heterogeneity. For example, again suppose we have two melting environments, one which fractionates parent from daughter, the other which does not. By the way the model is defined, both environments destroy heterogeneity, but only one of them creates it. Thus heterogeneity is now created on a different time scale to its destruction, but it is the time scale of destruction that the pseudo-isochron age relates to. This is somewhat analogous to the definition of residence time in reservoir modelling: it is the time scale for exiting the reservoir, not entering, that determines the residence time. Of course, the overall variances of the isotopic ratios will be reduced by having an environment which only destroys heterogeneity and does not fractionate.

Having different melting environments overcomes another shortcoming of the model. Most melting of the mantle occurs at mid-ocean ridges, and indeed this is where we are sampling from. However, melting at mid-ocean ridges involves large melt fractions, and thus does not create significant fractionation. Melting at ocean islands, on the other hand, does involve small melt fractions, and can cause the necessary fractionation. These two types of melting environment could be described by this generalisation.

There are other ways in which melt modelling could be improved. Taking account of the variation in melting behaviour with depth is likely to be important. At present, the partition coefficients used for the single melt region are an average of the garnet and spinel partition coefficients. Better modelling the range of melt fractions and behaviour with depth may have significant effect: particularly for the Lu-Hf system, where Lu has significantly different partition coefficients depending on whether garnet or spinel is the stable mineral.

6.2   Ocean island basalt

No attempt has yet been made to model the complex isotopic systematics of ocean island basalt (OIB) with the statistical box model. First it is important to note that the mean of the isotopic ratios for OIB is slightly different to that of MORB (OIB is more towards the enriched end: lower $^{143}Nd/^{144}Nd$). This either means OIB taps a different source to MORB, or that there is some difference in the melting process that causes some components to more readily enter the melt than others. The second of these suggestions, a system of preferential melting, could be modelled by a biased system of sampling of the components [Ito and Mahoney, 2005]. For example, melt components may be more likely to melt again, and thus more likely to be sampled. Such preferential melting effects will be more important at ocean islands, where only a small fraction of the underlying mantle melts, than at the mid-ocean ridges, where there is much more melting.

A further complication with OIB is that it varies significantly between different islands, whereas MORB is much more similar between different ridges. This may be a result of different islands tapping different sources, or of the more variable melting processes occurring under the different islands. As such, OIB should not be considered a single population. Different ocean islands are likely to have different amounts of preferential melting, and different amounts of mixing before sampling (different N) leading to the different means and variances. It would also be worthwhile to examine further how well MORB can be considered as a single population: investigating the small differences between ridges of different spreading rates. These can probably be modelled by different amounts of mixing before sampling (different N). For example, slower spreading ridges have smaller ranges of isotopic ratios which probably reflects greater amounts of mixing before sampling (larger N).

6.3   Continental crust

The statistical box model makes no attempt to model the extraction of the continental crust. The depleted mantle is simply treated as one statistical reservoir made up of enriched, depleted and primordial components; and interpreted in terms of oceanic crust recycling. As discussed in the introduction, the continental crust has formed over Earth's history, and indeed this is why the depleted mantle is depleted. An important future investigation would be to add statistical distributions to a mean reservoir model that is compatible with our understanding of the mean isotopic evolution of the Earth.

A simple generalisation can be motivated by the work of Albarède [2005]. Suppose that in a melting event we now permanently remove a fraction ζ of the atoms of a given element, and recycle the remaining $1 - \zeta$ back into the box. This fraction ζ can be thought of as going to the continental crust, and this will lead to a corresponding depletion of the box over time. Assuming there is no significant mass loss due to the formation of continental crust, the mantle parcels will remain of the same size and have the same age distribution.

Albarède [2005] derives residence time distributions for atoms in the mantle reservoir in such a situation. It should be noted that these are different, but related, to the distribution of parcel ages. Recall that the cumulative distribution function (cdf) for the distribution of parcels ages satisfies (section 5.6)

$$\frac{\mathrm{d}Q(\tau)}{\mathrm{d}\tau} = \gamma_{\mathrm{melt}}(\tau)\left(1 - Q(\tau)\right), \quad Q(0) = 0, \tag{6.7}$$

and thus

$$Q(\tau) = 1 - \mathrm{e}^{-\Gamma_{\mathrm{melt}}(\tau)}, \tag{6.8}$$

where

$$\Gamma_{\mathrm{melt}}(\tau) = \int_0^\tau \gamma_{\mathrm{melt}}(\tau)\,\mathrm{d}\tau. \tag{6.9}$$

$Q(\tau_s)$ is the fraction of parcels that are primordial. Now, instead of considering parcels of mantle material, consider atoms of a given element. The corresponding cdf $Q_a(\tau)$ for the atomic age distribution satisfies the similar equation

$$\frac{\mathrm{d}Q_a(\tau)}{\mathrm{d}\tau} = (1 - \zeta)\gamma_{\mathrm{melt}}(\tau)\left(1 - Q_a(\tau)\right), \quad Q_a(0) = 0, \tag{6.10}$$

which takes account of the fact that only a fraction $1 - \zeta$ of atoms survive a melting event. Thus

$$Q_a(\tau) = 1 - \mathrm{e}^{-(1-\zeta)\Gamma_{\mathrm{melt}}(\tau)}. \tag{6.11}$$

For constant melt rate, $\Gamma_{\mathrm{melt}}(\tau) = \tau/\tau_{\mathrm{melt}}$ and thus

$$Q_a(\tau) = 1 - \mathrm{e}^{-(1-\zeta)\tau/\tau_{\mathrm{melt}}}, \tag{6.12}$$

which is the analogue of (13) of Albarède [2005]. Note that Albarède [2005] works with the fraction of recycled atoms $\omega = 1 - \zeta$, and gives the result for the fraction of atoms of age greater than τ rather than the cdf (the fraction of atoms of age less than τ). Note also that the mass fraction of the mantle that is primordial is $\mathrm{e}^{-\tau_s/\tau_{\mathrm{melt}}}$, but the fraction of atoms of a given element that are primordial is $\mathrm{e}^{-(1-\zeta)\tau_s/\tau_{\mathrm{melt}}}$. For example, if all of an element is extracted to the continents in a melting event, then $\zeta = 1$ and thus all the atoms of that element in the mantle box must be primordial. Alternatively, if there is no removal in a melting event, then $\zeta = 0$ and the age distribution of parcels is the same as age distribution of atoms.

The mean isotopic evolution of the box is given by

$$\left(\frac{\bar{p}}{\bar{d'}}\right)_\tau = \frac{\bar{p}}{\bar{d'}}\mathrm{e}^{\lambda\tau + (\zeta_p - \zeta_d)\Gamma_{\mathrm{melt}}(\tau)}, \tag{6.13}$$

$$\left(\frac{\bar{d}}{\bar{d'}}\right)_\tau = \frac{\bar{d}}{\bar{d'}} - \frac{\bar{p}}{\bar{d'}}\int_0^\tau \lambda\mathrm{e}^{\lambda\tau + (\zeta_p - \zeta_d)\Gamma_{\mathrm{melt}}(\tau)}\,\mathrm{d}\tau, \tag{6.14}$$

where $\bar{d}/\bar{d}'$ is the present day value, and $\left(\bar{d}/\bar{d}'\right)_\tau$ is the value at an age τ before the present. For constant melt rate these become

$$\left(\frac{\bar{p}}{\bar{d}'}\right)_\tau = \frac{\bar{p}}{\bar{d}'}e^{(\lambda+k)\tau}, \tag{6.15}$$

$$\left(\frac{\bar{d}}{\bar{d}'}\right)_\tau = \frac{\bar{d}}{\bar{d}'} - \frac{\bar{p}}{\bar{d}'}\frac{\lambda}{\lambda+k}\left(e^{(\lambda+k)\tau} - 1\right), \tag{6.16}$$

where $k = (\zeta_p - \zeta_d)/\tau_{\text{melt}}$. These are the same equations derived by Allègre [1969], and have commonly been used in constructing depleted mantle models. For example, Workman and Hart [2005] model the evolution of the depleted mantle by continuous extraction of continental crust over the last 3.0 Ga using the above equations. To generalise the Workman and Hart [2005] model to a statistical box model, we just need the appropriate melting law. The simplest case would be to let melt parcels have concentrations $(G-\zeta)/F$ and the residue parcels $(1-G)/(1-F)$ relative to the source values. This reflects the depletion of the recycled melt due to continent formation. Note that $0 \leq \zeta \leq G \leq 1$. Explicit expressions for the concentrations of the different parcels types can then be written, but are rather cumbersome.

Generalisations to take into account the evolution of the depleted mantle are likely to affect the isotopic heterogeneity of some systems more than others. For example, the Sm/Nd ratio of the depleted mantle changes very little over time with the extraction of continental crust. Since the parent/daughter ratio is a key control on the variance of the isotopic ratios we might expect the variance of ^{143}Nd/^{144}Nd to be little affected by continent formation. On the other hand, Rb/Sr is quite drastically affected by continent formation, and so this could substantially alter the variance of ^{87}Sr/^{86}Sr. However, it is important to note that the isotopic heterogeneity only records a portion of the crustal evolution history, which depends on the melting time scale τ_{melt}. If τ_{melt} is very small then only the very recent history of continental formation can affect the spread of the isotopic ratios. Since there is thought to be much less exchange of material between the crust and the mantle at present, this may mean the effects of continental formation on isotopic heterogeneity are not that significant.

There have been a wide variety of different models for the mean isotopic evolution of the depleted mantle, and there has yet to be any clear consensus. For example, it is known that models which have simple unidirectional transport of material from mantle to crust are not compatible with observations of ϵ_{Nd} for older rocks. It seems the mantle acquired a significant positive ϵ_{Nd} very early on in the Earth's history, and the evolution of ϵ_{Nd} in the depleted mantle has been fairly linear since. Unidirectional transport models cannot satisfy this [Galer et al., 1989], and thus there must have been some recycling of the continental crust.

More complicated reservoir models do involve crustal recycling, but there are further problems. The lead isotopic systems in particular cause difficulties, and these difficulties are referred to as the lead paradoxes. Solutions to these paradoxes usually rely on a hidden reservoir in the Earth that cannot be sampled, e.g. the lower continental crust, or the D" layer. While hidden reservoirs can solve the problem, they are poorly constrained. The mean isotopic evolution of the Earth may need to be better understood before we add the complexities associated with the statistical distributions.

An important generalisation when considering crustal evolution would be to examine non-Poisson models. Reservoir models are based on a Poisson process assumption, whereby each atom in a reservoir has an equal probability of being extracted. It would be interesting to consider transport models that do not make this assumption. For example, the probability of extraction could be a function of the time an atom spends in a reservoir.

This could be used to make older crust less likely to recycle than younger crust, and could be important in understanding the age distribution of continental rocks.

6.4 Helium

Isotopes of noble gases, such as helium, can place further constraints on mantle evolution, particularly of the history of degassing. ^{4}He is produced through the alpha decay of ^{232}Th, ^{238}U and ^{235}U. ^{3}He is a stable reference isotope. The isotopic ratio ^{4}He/^{3}He is thus of a similar form to the other isotopic ratios we have studied except that it has three different parent isotopes instead of just one. The other key difference is that helium is a gas, and is lost to the atmosphere after melt comes to the surface. Traditionally, the isotopic ratio is reported upside down as $R = {}^3$He/^{4}He, and is reported relative to the atmospheric value $R_A = 1.4 \times 10^{-6}$.

MORB records a fairly narrow range of ^{3}He/^{4}He values, around 7-9 R/R_A. OIB records a much wider range of values, with some very large values (e.g. 35 R/R_A reported in Iceland) and some smaller values (e.g. 4 R/R_A reported in the Azores). The high values of ^{3}He/^{4}He have traditionally been associated with tapping of primordial material, or at least relatively less degassed material. An alternative viewpoint is that the source of MORB and OIB are the same, and the differences in ^{3}He/^{4}He result from differences in the sampling process [Meibom and Anderson, 2004]. For example, the extreme values could be the result of a less averaged distribution, or result from preferential extraction of certain components.

Extensions of the statistical box model may be able to help tackle this problem. The outgassing of helium could be modelled by the generalisation discussed in the previous section, by removing most of the helium from the melt in a melting event. Unfortunately there are many uncertainties in the helium system, and the overall helium budget is not well constrained. The initial parent/daughter ratios such as ^{238}U/^{3}He are not know, nor is the partitioning behaviour of helium well understood (in particular, it is not known for sure whether He is more or less compatible than Th and U). The mantle convection simulations of Xie and Tackley [2004b] have attempted to model the distributions of helium isotopes. Preliminary results from a generalisation of the statistical box model suggest that good agreement can be found with the distributions obtained from their convection simulations. Thus it seems that the key features of the ^{3}He/^{4}He distribution do not depend strongly on the details of the underlying flow, and the statistical box model may provide a neater way of attacking the problem, and a faster way of exploring parameter space.

6.5 Spatial structure

The statistical box model takes no account of spatial structure. The stirring of mantle convection reduces the length scales of heterogeneity, and redistributes heterogeneity. So far, a highly simplified approach to this has been taken: modelling melting as a Poisson process. The melting before sampling has a mixing effect, and the simplified modelling of this has been the averaging over N components. Both of these modelling approaches deserve careful further study.

The validity of Poisson process approximations are probably best assessed by comparison with numerical mantle convection simulations. Results so far seem to indicate that Poisson process models capture a lot of the important detail. However, the comparisons in this thesis were made indirectly without access to the actual simulations, and a much more detailed study should be made by those running the simulations. An intermediate

approach might be useful too: studying geochemical tracers in idealised flows for which the stirring behaviour is well understood.

Sampling is currently modelled fairly crudely by the single mixing parameter N. More sophisticated sampling algorithms have been developed by other authors. For example, the Kellogg et al. [2002] model takes into account the fact that older heterogeneities are likely to have shorter length scales due to larger amounts of stretching and folding over their history. Mixing in the Kellogg et al. [2002] model is parameterised in terms of a stirring time τ_{stir} and a sampling length scale l_s. It will be interesting to investigate how the schemes compare, particularly in the limiting case of large amounts of mixing (large N, large l_s). In this limit, both schemes approach normal distributions, so it is likely this would help justify the more simplistic approach taken here.

A perhaps more fruitful approach to understanding the length scales of mantle heterogeneity is to examine the observed spatial distribution of isotopic data. High quality isotopic data with accurate geographical sampling co-ordinates is now available for large regions of the mid-ocean ridge system [Agranier et al., 2005, Graham et al., 2006]. This may potentially provide clues as to length scales of mantle heterogeneity and is an area of active research. How to best model this data is not yet clear. It should be reiterated that we cannot observe the length scales of heterogeneity of the mantle directly, and can observe only through the filter of melting and mixing. This makes the problem of understanding length scales of mantle heterogeneity much more challenging.

6.6 Statistical methodology

There are a number of problems where a similar statistical methodology might be applied. For example, more can be made of the U-series disequilibria data than just the mean values: there are again variances, correlations, and distributions that can be exploited. There have been several fluid dynamical approaches to modelling U-series disequilibria [McKenzie, 1985, Richardson and McKenzie, 1994, Spiegelman and Elliott, 1993, Williams and Gill, 1989], but it would be interesting to see how far simple statistical approaches can go. The U-series problem is essentially one of residence times: disequilibrium values of $(^{230}\mathrm{Th}/^{238}\mathrm{U}) > 1$ arise from Th having a shorter residence time than U in melt extraction. Isochron dating can also be applied to activity ratios, and this suggests that the pseudo-isochron work might be relevant.

This thesis has focused on a rather narrow set of geochemical observations, those of certain radiogenic isotopes. Much more information is likely to be gained from using combinations of other geochemical observables such as major elements, trace elements, and U-series disequilibria. Unfortunately it not always the case that all measurements are made on all samples, but this is beginning to change. Correlations between isotopic ratios and trace elements were exploited by McKenzie et al. [2004], and it is likely there is more to be gained by further statistical modelling of different combinations of geochemical data.

Chapter 7

Conduit waves

7.1 Introduction

An important problem in geophysics is to understand the processes whereby melt, produced at depths of around 100 km in the Earth's mantle, reaches the surface. Geochemical measurements, notably of uranium series disequilibria, show that the melt migration process has to happen quickly (~ 100 years). Furthermore, there must be little interaction of the melt with the surrounding matrix as it travels to the surface, in order to preserve this disequilibrium signal.

Early physical models of melt migration considered transport by diffuse porous flow along grain boundaries; but this neither transports melt fast enough, nor is likely to preserve the disequilibrium signal. To overcome this difficulty it was proposed that the melt is organised into a network of high porosity conduits which allows much faster transport [Kelemen et al., 1997]. This motivated the theoretical study of the behaviour of an open melt conduit in a deformable porous matrix by Richardson et al. [1996], and this chapter revisits their work. The conduit was shown to support solitary waves of elevation with a region of trapped melt travelling with the wave. However, the analysis presented here shows that in some parameter regimes the conduit will only support solitary waves of depression. These parameter regimes include that thought appropriate to melt migration in the Earth's mantle. The waves of depression do not contain trapped melt and may have slightly different geochemical consequences.

The analysis largely follows that of Richardson et al. [1996] with two important modifications. Instead of assuming a constant bulk viscosity we assume a power law behaviour with porosity (this turns out to be of particular importance when considering waves of depression). We also do not assume that the porosities are small, as this does not significantly complicate the analysis. A list of the notation used is given at the end of this chapter.

7.2 Compaction equations

Consider an axisymmetric conduit of radius $a(z, t)$. The conduit is pure melt of viscosity μ and density ρ_f. Outside the conduit is a deformable porous matrix. In this section we study the dynamics of the deformable porous matrix in order to find expressions for the pressure and leakage flux across the conduit wall. In the following section the dynamics inside the conduit is studied using these expressions.

The behaviour of melt and matrix outside the conduit is described by the compaction equations [McKenzie, 1984, Spiegelman, 1993]. In their most general form with no melting

the compaction equations are

$$\frac{\partial (\rho_f \phi)}{\partial t} + \nabla \cdot (\rho_f \phi \mathbf{v}) = 0, \tag{7.1}$$

$$\frac{\partial (\rho_s (1 - \phi))}{\partial t} + \nabla \cdot (\rho_s (1 - \phi) \mathbf{V}) = 0, \tag{7.2}$$

$$\phi (\mathbf{v} - \mathbf{V}) = -\frac{k_\phi}{\mu} \nabla \mathcal{P}, \tag{7.3}$$

$$\nabla \cdot \boldsymbol{\sigma} = \nabla \mathcal{P} + (1 - \phi) \Delta \rho g \hat{\mathbf{z}}, \tag{7.4}$$

$$\boldsymbol{\sigma} = 2\eta_\phi \mathbf{E} + \left(\zeta_\phi - \tfrac{2}{3}\eta_\phi\right) \mathcal{C} \mathbf{I}, \tag{7.5}$$

$$\mathbf{E} = \tfrac{1}{2} \left(\nabla \mathbf{V} + (\nabla \mathbf{V})^T\right), \tag{7.6}$$

$$\mathcal{C} = \nabla \cdot \mathbf{V}. \tag{7.7}$$

(7.1) and (7.2) describe conservation of mass for melt and matrix, where ρ_f, ρ_s and $\mathbf{v}$, $\mathbf{V}$ are the densities and velocities of melt and matrix respectively. ϕ is the porosity, the volume fraction of melt. (7.3) is the appropriate form of Darcy's law, where k_ϕ is the permeability, and $\mathcal{P}$ is the piezometric pressure (the pressure in excess of hydrostatic pressure for the melt $\mathcal{P} = p + \rho_f g z$). (7.4) is the momentum equation where $\Delta \rho = \rho_s - \rho_f$. Finally (7.5) gives the constitutive law for the behaviour of the matrix where (7.6) and (7.7) define the matrix rate of strain tensor $\mathbf{E}$ and the compaction rate $\mathcal{C}$ respectively. To complete this set of equations constitutive laws are needed expressing the dependence on porosity ϕ of the matrix shear viscosity η_ϕ, the matrix bulk viscosity ζ_ϕ, and the permeability k_ϕ. We will assume $\eta_\phi = \eta$ is constant, and that

$$k_\phi = k_0 \left(\frac{\phi}{\phi_0}\right)^n, \quad \zeta_\phi + \tfrac{4}{3}\eta = \left(\zeta_0 + \tfrac{4}{3}\eta\right) \left(\frac{\phi}{\phi_0}\right)^{-m}, \tag{7.8}$$

where k_0, ζ_0 and ϕ_0 are the respective constant background state values. Typical values of the exponents are $n = 3$ and $m = 1$. Assuming further that ρ_f, ρ_s and μ are constants, the compaction equations then simplify to

$$\nabla \mathcal{P} = -\frac{\mu}{k_\phi} \mathbf{q}, \tag{7.9}$$

$$\nabla \cdot \mathbf{q} = -\mathcal{C}, \tag{7.10}$$

$$\nabla \cdot \mathbf{V} = \mathcal{C}, \tag{7.11}$$

$$\frac{\partial \phi}{\partial t} + \mathbf{V} \cdot \nabla \phi = (1 - \phi) \mathcal{C}, \tag{7.12}$$

$$\nabla \mathcal{P} = -\eta \nabla \times \boldsymbol{\omega} + \nabla \left(\zeta_\phi + \tfrac{4}{3}\eta\right) \mathcal{C} - (1 - \phi) \Delta \rho g \hat{\mathbf{z}}, \tag{7.13}$$

where the matrix vorticity $\boldsymbol{\omega} = \nabla \times \mathbf{V}$ and the Darcy flux $\mathbf{q} = \phi (\mathbf{v} - \mathbf{V})$.

Far from the conduit we assume the background steady state solution of uniform buoyancy driven flow $\phi = \phi_0$, $\mathbf{V} = 0$, $\mathbf{q} = \Delta \rho g k_0 (1 - \phi_0) \hat{\mathbf{z}}/\mu$, $\mathcal{C} = 0$, $\mathcal{P} = \mathcal{P}_\infty \equiv -(1 - \phi_0) \Delta \rho g z$. To simply further, we assume that the compaction driven flow is purely radial (this will turn out to be a long wavelength assumption). The validity of this approximation has to be checked after the analysis. Letting $P = \mathcal{P} - \mathcal{P}_\infty = p + \bar{\rho}_0 g z$ (the pressure in excess of hydrostatic pressure for the background melt and matrix, which has mean density $\bar{\rho}_0 = \phi_0 \rho_f + (1 - \phi_0) \rho_s$) the above equations then reduce to the 1-D compaction

equations

$$\frac{\partial P}{\partial r} = -\frac{\mu}{k_\phi} q_r, \tag{7.14}$$

$$\frac{1}{r}\frac{\partial}{\partial r}(rq_r) = -\mathcal{C}, \tag{7.15}$$

$$\frac{1}{r}\frac{\partial}{\partial r}(rV_r) = \mathcal{C}, \tag{7.16}$$

$$\frac{\partial \phi}{\partial t} + V_r \frac{\partial \phi}{\partial r} = (1 - \phi)\,\mathcal{C}, \tag{7.17}$$

$$P = \left(\zeta_\phi + \tfrac{4}{3}\eta\right)\mathcal{C}, \tag{7.18}$$

where q_r and V_r are the radial components of the Darcy flux $\mathbf{q}$ and matrix velocity $\mathbf{V}$ respectively.

On the conduit wall the matrix velocity must match the expansion rate of the conduit. This gives the boundary condition

$$V_r = \frac{\partial a}{\partial t}, \qquad (r = a). \tag{7.19}$$

The melt pressure must balance the matrix stress on the conduit wall, which from (7.5) gives

$$0 = 2\eta \frac{\partial V_r}{\partial r} + (\zeta_\phi - \tfrac{2}{3}\eta)\mathcal{C}, \qquad (r = a). \tag{7.20}$$

Using (7.16) and (7.18) this can be simplified to

$$P = \frac{2\eta}{a}\frac{\partial a}{\partial t}, \qquad (r = a). \tag{7.21}$$

The final boundary condition is to match the background state far from the conduit

$$P \to 0, \qquad (r \to \infty). \tag{7.22}$$

Along with initial conditions for ϕ, the three boundary conditions (7.19), (7.21) and (7.22) and the governing equations (7.14-7.18) completely describe the dynamics outside the conduit.

7.2.1 Non-dimensional equations

Introduce the non-dimensional parameters

$$\delta_0^2 = \frac{k_0\left(\zeta_0 + \tfrac{4}{3}\eta\right)}{\mu a_0^2}, \quad \mathcal{B}_0 = \frac{\eta}{\zeta_0 + \tfrac{4}{3}\eta}. \tag{7.23}$$

δ_0 is the compaction length of the background state non-dimensionalised on the undisturbed conduit radius a_0. $\mathcal{B}_0$ is a ratio of matrix viscosities for the background state. Note that $0 \leq \mathcal{B}_0 \leq 3/4$. Introduce scaled variables $a = a_0 a'$, $r = a_0 r'$, $P = P_0 P'$, $V_r = V_0 V_r'$, $q_r = V_0 q_r'$, $\mathcal{C} = \mathcal{C}_0 \mathcal{C}'$, and $\phi = \phi_0 \phi'$ where

$$P_0 = \frac{\zeta_0 + \tfrac{4}{3}\eta}{a_0}\frac{\partial a}{\partial t}, \quad \mathcal{C}_0 = \frac{1}{a_0}\frac{\partial a}{\partial t}, \quad V_0 = \frac{\partial a}{\partial t}. \tag{7.24}$$

With this rescaling the problem in the original variables (r, t) becomes one in the new variables (r', a') with governing equations

$$\frac{\partial P'}{\partial r'} = -\frac{1}{\delta_0^2}\phi'^{-n}q_r',$$

(7.25)

$$\frac{1}{r'}\frac{\partial}{\partial r'}\left(r'q_r'\right) = -\mathcal{C}',$$

(7.26)

$$\frac{1}{r'}\frac{\partial}{\partial r'}\left(r'V_r'\right) = \mathcal{C}',$$

(7.27)

$$\frac{\partial\phi'}{\partial a'} + V_r'\frac{\partial\phi'}{\partial r'} = \frac{1 - \phi_0\phi'}{\phi_0}\mathcal{C}',$$

(7.28)

$$\mathcal{C}' = P'\phi'^m,$$

(7.29)

and boundary and initial conditions

$$V_r' = 1, \quad P' = \frac{2\mathcal{B}_0}{a'}, \quad (r' = a'),$$

(7.30)

$$P' \to 0 \quad (r' \to \infty),$$

(7.31)

$$\phi' = 1 \quad (a' = 1).$$

(7.32)

These equations are straightforward to solve numerically. (7.28) can be marched forward or backward from the initial condition (7.32) at $a' = 1$ (the undisturbed radius) where at each a'-step the two point boundary value problem specified by the remaining equations is solved.

Of particular interest for the next section is the leakage flux of melt from the conduit $F(a) = q_r(r = a)$. Letting $f(a') = q_r'(r' = a')$ be found by the solution of the above, we have that

$$F(a) = \frac{\partial a}{\partial t}f\left(a/a_0\right).$$

(7.33)

Note that the form of the non-dimensional function $f(a')$ depends on the five non-dimensional parameters δ_0, $\mathcal{B}_0$, ϕ_0, n and m.

7.2.2 Porosity on the conduit wall

It is interesting to note that we can directly find the porosity of the matrix on the conduit wall using (7.28), (7.29) and (7.30). Let $\phi_i'(a) = \phi'(a, a)$ be the scaled porosity on the wall, and $\phi_i = \phi_0\phi_i'$ the corresponding unscaled porosity. Then

$$\frac{d\phi_i'}{da'} = \frac{2\mathcal{B}_0}{a'}\frac{1 - \phi_0\phi_i'}{\phi_0}\phi_i'^m.$$

(7.34)

For $m = 0$ (as used by Richardson et al. [1996]) this can be integrated to give

$$\phi_i'(a) = 1 + \frac{1 - \phi_0}{\phi_0}\left(1 - a'^{-2\mathcal{B}_0}\right).$$

(7.35)

This has the problem that ϕ_i becomes negative as $a' \to 0$, which is unphysical: it implies that the matrix continues to compact even when all the melt has been squeezed out. Richardson et al. [1996] only considered waves of elevation, where $a' \geq 1$ and so this did not pose a problem. However, for a wave of depression $a' \leq 1$ and this will be important. For $m \geq 1$ this problem is avoided, e.g. $m = 1$ gives

$$\phi_i'(a) = \left(\phi_0 + (1 - \phi_0)a'^{-2\mathcal{B}_0/\phi_0}\right)^{-1}.$$

(7.36)

This is well behaved both as $a' \to 0$ and as $a' \to \infty$ with $\phi_i \to 0$ and $\phi_i \to 1$ respectively.

7.3   Conduit equations

Inside the conduit the behaviour is governed by the following three equations

$$Q = -\frac{\pi a^4}{8\mu}\frac{\partial \mathcal{P}}{\partial z}, \tag{7.37}$$

$$\frac{\partial Q}{\partial z} = -2\pi a\left(\frac{\partial a}{\partial t} + F(a)\right) = -2\pi a\frac{\partial a}{\partial t}\left(1 + f\left(a/a_0\right)\right), \tag{7.38}$$

$$\mathcal{P} = P(r = a) + \mathcal{P}_\infty = \frac{2\eta}{a}\frac{\partial a}{\partial t} - (1 - \phi_0)\Delta\rho g z. \tag{7.39}$$

(7.37) is Poiseuille's law for flow in a pipe. (7.38) expresses conservation of mass: the flux in the conduit changes due to changes in the conduit radius and through leakage into the pores of the matrix. Finally (7.39) gives that the pressure in the conduit must match that on the conduit walls, using the boundary condition (7.21).

It is more convenient to work with the conduit area rather than the conduit radius. Let $b = \pi a^2$ be the conduit area. The conduit equations then become

$$Q = -\frac{b^2}{8\pi\mu}\frac{\partial \mathcal{P}}{\partial z}, \tag{7.40}$$

$$\frac{\partial Q}{\partial z} = -\frac{\partial b}{\partial t}\left(1 + f\left(\left(b/\pi a_0^2\right)^{1/2}\right)\right), \tag{7.41}$$

$$\mathcal{P} = \frac{\eta}{b}\frac{\partial b}{\partial t} - (1 - \phi_0)\Delta\rho g z. \tag{7.42}$$

7.3.1   Non-dimensional equations

Let $z = z_0 z'$, $t = t_0 t'$, $b = \pi a_0^2 b'$, $Q = Q_0 Q'$, $\mathcal{P} = \mathcal{P}_0 \mathcal{P}'$ where

$$z_0 = a_0\left(\frac{\eta}{8\left(1 + f(1)\right)\mu}\right)^{1/2}, \tag{7.43}$$

$$t_0 = \frac{\left(8\left(1 + f(1)\right)\eta\mu\right)^{1/2}}{a_0(1 - \phi_0)\Delta\rho g}, \tag{7.44}$$

$$c_0 = \frac{z_0}{t_0} = \frac{a_0^2\left(1 - \phi_0\right)\Delta\rho g}{8\mu\left(1 + f(1)\right)}, \tag{7.45}$$

$$Q_0 = \frac{\pi a_0^4\left(1 - \phi_0\right)\Delta\rho g}{8\mu}, \tag{7.46}$$

$$\mathcal{P}_0 = (1 - \phi_0)\Delta\rho g z_0. \tag{7.47}$$

Note that this is a different non-dimensionalisation to that used by Richardson et al. [1996], but allows for easier comparison with the standard conduit equations. The conduit equations reduce to

$$Q' = -b'^2\frac{\partial \mathcal{P}'}{\partial z'}, \tag{7.48}$$

$$\frac{\partial Q'}{\partial z'} = -h(b')\frac{\partial b'}{\partial t'}, \tag{7.49}$$

$$\mathcal{P}' = \frac{1}{b'}\frac{\partial b'}{\partial t'} - z', \tag{7.50}$$

where

$$h(b') = \frac{1 + f\left(b'^{1/2}\right)}{1 + f(1)}. \tag{7.51}$$

Note that $h(1) = 1$. These equations can be combined to give a single nonlinear wave equation for the non-dimensional conduit area

$$h(b')\frac{\partial b'}{\partial t'} + \frac{\partial}{\partial z'}\left(b'^2\left(1 - \frac{\partial}{\partial z'}\left(\frac{1}{b'}\frac{\partial b'}{\partial t'}\right)\right)\right) = 0. \tag{7.52}$$

If $h(b') \equiv 1$ then this is the standard conduit solitary wave equation [Helfrich and White-head, 1990, Olson and Christensen, 1986, Whitehead and Helfrich, 1986]. This same equation also governs porosity solitary waves ('magmons') and has been studied extensively by many previous authors [Barcilon and Lovera, 1989, Barcilon and Richter, 1986, Richter and McKenzie, 1984, Scott and Stevenson, 1984, 1986, Scott et al., 1986]. Our interest is in how the form of $h(b')$ changes the problem. From now on we will drop the primes and unless otherwise stated are working with non-dimensional variables.

7.4 Solitary waves

The governing equation (7.52) can be written as

$$bb_{zzt} - b_{zz}b_t - h(b)b_t - 2bb_z = 0. \tag{7.53}$$

This equation admits small amplitude linear wave solutions $b = 1 + \epsilon e^{i(kx-\omega t)}$, $\epsilon \ll 1$ which are dispersive with dispersion relationship

$$\omega = \frac{2k}{1 + k^2}. \tag{7.54}$$

Note that long wavelength waves ($k \ll 1$) have phase and group velocity equal to 2.

Of greater interest is that (7.53) also admits nonlinear solitary wave solutions. We seek solutions of the form $b(z,t) = b(\xi)$ where $\xi = z - ct$. Then (7.53) becomes

$$bb_{\xi\xi\xi} - b_{\xi\xi}b_\xi - h(b)b_\xi + \frac{2}{c}bb_\xi = 0. \tag{7.55}$$

This can be integrated once to give

$$bb_{\xi\xi} - b_\xi^2 - H(b) + \frac{1}{c}b^2 = M, \tag{7.56}$$

and integrated a second time to give

$$\frac{b_\xi^2}{2b^2} - \frac{G(b)}{2b^2} + \frac{1}{c}\log b = -\frac{M}{2b^2} + N, \tag{7.57}$$

where

$$H(b) = \int_1^b h(\beta)\,d\beta, \quad G(b) = 2b^2\int_1^b \beta^{-3}H(\beta)\,d\beta. \tag{7.58}$$

Note that for the standard conduit equations $h(b) = 1$, $H(b) = b - 1$ and $G(b) = (b-1)^2$. The integration constants M and N depend on boundary conditions. We impose the standard solitary wave boundary conditions

$$b = 1, b_\xi = 0, b_{\xi\xi} = 0, \quad (\xi \to \pm\infty), \tag{7.59}$$

$$b = B, b_\xi = 0, \quad (\xi = 0). \tag{7.60}$$

Here B is the wave amplitude. The boundary condition (7.59) implies $M = c^{-1}$, $N = (2c)^{-1}$. (7.57) then becomes

$$b_\xi^2 = G(b) - \frac{1}{c} \left(2b^2 \log b - b^2 + 1 \right). \tag{7.61}$$

(7.60) then determines the dispersion relation of the solitary wave, relating wave speed to wave amplitude as

$$c = c(B) = \frac{2B^2 \log B - B^2 + 1}{G(B)}. \tag{7.62}$$

In terms of the dispersion relationship (7.62), (7.61) can be rewritten as

$$b_\xi^2 = \left(2b^2 \log b - b^2 + 1 \right) \left(\frac{1}{c(b)} - \frac{1}{c(B)} \right). \tag{7.63}$$

This equation can be integrated numerically to find the mode shape $b(\xi)$. In order for solutions of this form to exist $b_\xi^2 \geq 0$. Since $\left(2b^2 \log b - b^2 + 1 \right) \geq 0$ for $b \geq 0$, this implies that in terms of the dispersion relationship we require $c(B) \geq c(b)$ for all values of b on the wave. Hence from a plot of the dispersion relationship the existence of solitary wave solutions of a given amplitude can readily be determined.

7.5 Streamfunction

The method of Helfrich and Whitehead [1990] can be used to find the streamfunction inside the conduit. Denote the components of velocity inside the channel by u_r, u_z. We are assuming Poiseuille flow in the conduit, which in dimensional form is

$$u_z = \frac{2Q}{\pi a^4} \left(a^2 - r^2 \right). \tag{7.64}$$

Non-dimensionalising $u_z = u_z' c_0$ using (7.45), and noting that for the solitary wave $Q = 1 + cH(b)$ yields the non-dimensional equation

$$u_z = \frac{2}{b^2} \left(1 + f(1) \right) \left(1 + cH(b) \right) \left(b - r^2 \right). \tag{7.65}$$

Introduce the streamfunction ψ in the frame moving with the wave as

$$u_r = -\frac{1}{r} \frac{\partial \psi}{\partial \xi}, \quad u_z - c = \frac{1}{r} \frac{\partial \psi}{\partial r}. \tag{7.66}$$

The Poiseuille velocity field (7.65) can then be integrated to find the streamfunction as

$$\psi = \frac{2}{b^2} \left(1 + f(1) \right) \left(1 + cH(b) \right) \left(\frac{br^2}{2} - \frac{r^4}{4} \right) - \frac{cr^2}{2}. \tag{7.67}$$

We are particularly interested in the stagnation points of this flow. Since $u_r = -b_\xi \psi_b / r$, then the radial component of velocity is zero when $b_\xi = 0$, in particular at $b = 1$ and $b = B$. Stagnation points will occur when we also have $u_z = c$. Thus at

$$b = 1, \quad r^2 = 1 - \frac{c}{2 \left(1 + f(1) \right)}, \tag{7.68}$$

$$b = B, \quad r^2 = B - \frac{cB^2}{2 \left(1 + f(1) \right) \left(1 + cH(B) \right)}, \tag{7.69}$$

there will be stagnation points, where these roots exist. There are possibly other stagnation points in the flow where $\psi_b = 0$ and $u_z = c$. Of particular interest is the stagnation point (7.69). If this stagnation point is a centre there will be a region of closed streamlines, and thus a region of trapped melt will be transported along with the wave. If the stagnation point is a saddle there will be no such region of trapped melt.

7.6 Small amplitude solitary waves

The equations of the previous sections can be used to numerically calculate the shape and streamline pattern for solitary waves of given amplitude. Further insight can be gained by looking at the particular limit of small amplitude solitary waves $B = 1 + \epsilon$, $\epsilon \ll 1$. By Taylor expanding the dispersion relationship (7.62) we find

$$c = c(1 + \epsilon) = 2 + \frac{2}{3}\left(1 - h'(1)\right)\epsilon + \mathcal{O}(\epsilon^2). \tag{7.70}$$

Note that the condition for existence of solutions $c(B) \geq c(b)$ implies that for a small amplitude wave of elevation ($\epsilon > 0$) we require $h'(1) < 1$ whereas for a wave of depression ($\epsilon < 0$) we require $h'(1) > 1$. Note also that in terms of $f(a)$

$$h'(1) = \frac{f'(1)}{2\left(1 + f(1)\right)}. \tag{7.71}$$

Furthermore, for small amplitudes the governing equation (7.53) can be approximated by the Korteweg-deVries equation [Whitehead and Helfrich, 1986]. Letting $b = 1 + \epsilon b_1 + \epsilon^2 b_2 + ...$, and introducing stretched time and length scales as $T = |\epsilon|^{3/2}t$, $Z = |\epsilon|^{1/2}(z - 2t)$ then the leading order governing equation is

$$b_{1T} + 2(1 - h'(1))\operatorname{sgn}(\epsilon)b_1 b_{1Z} + 2b_{1ZZZ} = 0, \tag{7.72}$$

which is of KdV form. We seek solitary wave solutions with $b_1(Z, T) = b_1(\chi)$ where $\chi = Z - CT$. The boundary conditions are

$$b_1 = 0, b_{1\chi} = 0, b_{1\chi\chi} = 0, \quad (\chi \to \pm\infty), \tag{7.73}$$
$$b_1 = 1, b_{1\chi} = 0, \quad (\chi = 0), \tag{7.74}$$

which imply the familiar sech^2 soliton solution

$$C = \frac{2}{3}\left(1 - h'(1)\right)\operatorname{sgn}(\epsilon), \tag{7.75}$$

$$b_1 = \operatorname{sech}^2\left(\frac{1}{2}\left(\frac{(1 - h'(1))\operatorname{sgn}(\epsilon)}{3}\right)^{1/2}\chi\right). \tag{7.76}$$

Hence the approximate solution for b is

$$b \sim 1 + \epsilon\operatorname{sech}^2\left(\frac{1}{2}\left(\frac{(1 - h'(1))\epsilon}{3}\right)^{1/2}(z - ct)\right). \tag{7.77}$$

The stagnation points found in (7.68) and (7.69) are located to $\mathcal{O}(\epsilon)$ at

$$b = 1, \quad r^2 = \frac{f(1)}{1 + f(1)} + \frac{h'(1) - 1}{3\left(1 + f(1)\right)}\epsilon + \mathcal{O}\left(\epsilon^2\right), \tag{7.78}$$

$$b = 1 + \epsilon, \quad r^2 = \frac{f(1)}{1 + f(1)} + \left(1 + \frac{h'(1) - 1}{3\left(1 + f(1)\right)}\right)\epsilon + \mathcal{O}\left(\epsilon^2\right). \tag{7.79}$$

For $f(1) > 0$ and ϵ sufficiently small, these stagnation points always exist. Furthermore, the character of the stagnation points can be classified in terms of the determinant $D = \psi_{\xi\xi}\psi_{rr} - \psi_{\xi r}^2$. If $D > 0$ then the stagnation point is a centre, whereas if $D < 0$ then it is a saddle. When $b_\xi = 0$, $D = b_{\xi\xi}\psi_b\psi_{rr}$. For small amplitude waves $D = -4b_{\xi\xi}(f(1))^2 + \mathcal{O}(\epsilon)$.

Now for a wave of elevation, $\epsilon > 0$ and $b_{\xi\xi} < 0$ at $b = B$, and thus the stagnation point there is a centre and there is a region of trapped melt. However, for a wave of depression, $\epsilon < 0$ and $b_{\xi\xi} > 0$ at $b = B$, and thus the stagnation point is a saddle and there is no trapped melt.

It is important to note that in practice the values of ϵ for which the above approximations are reasonable may be quite small. For example, it will depend on how large higher derivatives of $h(b)$ are. However, for given values of the five non-dimensional parameters a small enough ϵ can always be found so that the approximation is good. While this small amplitude analysis may not always be a good approximation in practice, it nevertheless provides good insight into the features of the solution we expect to see.

7.6.1 Series expansion to find $f(1)$ and $h'(1)$

To calculate the key values $f(1)$ and $h'(1)$ we return to the 1-D compaction equations outside the conduit. Let $a = 1 + \epsilon$ and seek a solution of the form

$$\phi(r, a) = 1 \qquad + \epsilon\phi_1(r) + \epsilon^2\phi_2(r) + \cdots , \tag{7.80}$$

$$P(r, a) = P_0(r) + \epsilon P_1(r) + \epsilon^2 P_2(r) + \cdots , \tag{7.81}$$

$$q_r(r, a) = q_0(r) + \epsilon q_1(r) + \epsilon^2 q_2(r) + \cdots , \tag{7.82}$$

$$V_r(r, a) = V_0(r) + \epsilon V_1(r) + \epsilon^2 V_2(r) + \cdots . \tag{7.83}$$

Note that since $f(a) = q_r(a, a)$ we have that

$$f(1) = q_r(1, 1) = q_0(1), \tag{7.84}$$

$$f'(1) = \left.\frac{\mathrm{d}q_r(a, a)}{\mathrm{d}a}\right|_{a=1} = q_0'(1) + q_1(1), \tag{7.85}$$

$$h'(1) = \frac{q_0'(1) + q_1(1)}{2\left(1 + q_0(1)\right)}. \tag{7.86}$$

From the governing equations (7.25-7.28) the leading order problem is

$$\frac{\mathrm{d}P_0}{\mathrm{d}r} + \frac{1}{\delta_0^2}q_0 = 0, \tag{7.87}$$

$$\frac{1}{r}\frac{\mathrm{d}}{\mathrm{d}r}\left(rq_0\right) + P_0 = 0, \tag{7.88}$$

$$\frac{1}{r}\frac{\mathrm{d}}{\mathrm{d}r}\left(rV_0\right) - P_0 = 0, \tag{7.89}$$

$$\phi_1 = \frac{1 - \phi_0}{\phi_0}P_0, \tag{7.90}$$

with boundary conditions from (7.30) and (7.31) as

$$V_0(1) = 1, \quad P_0(1) = 2\mathcal{B}_0, \quad P_0 \to 0 \text{ as } r \to \infty. \tag{7.91}$$

This has an explicit solution in terms of the modified Bessel function of the second kind as

$$P_0(r) = 2\mathcal{B}_0\frac{K_0(r/\delta_0)}{K_0(1/\delta_0)}, \tag{7.92}$$

$$q_0(r) = \frac{2\mathcal{B}_0 K_1(r/\delta_0)}{(1/\delta_0)K_0(1/\delta_0)}, \tag{7.93}$$

$$V_0(r) = \frac{1}{r} + \frac{2\mathcal{B}_0}{(1/\delta_0)K_0(1/\delta_0)}\left(\frac{K_1(1/\delta_0)}{r} - K_1(r/\delta_0)\right). \tag{7.94}$$

Hence

$$f(1) = q_0(1) = \frac{2\mathcal{B}_0 K_1(1/\delta_0)}{(1/\delta_0)K_0(1/\delta_0)}. \tag{7.95}$$

Note that $f(1) \geq 0$, and that $f(1) = 0$ only if $\mathcal{B}_0 = 0$ or $\delta_0 = 0$ (the impermeable limit). Note also that $f(1) \sim 2\mathcal{B}_0\delta_0$ for $\delta_0 \ll 1$ and that $f(1) \sim 2\mathcal{B}_0\delta_0^2/\log\delta_0$ for $\delta_0 \gg 1$.

To find q_1 we need to solve the next order problem

$$\frac{\mathrm{d}P_1}{\mathrm{d}r} + \frac{1}{\delta_0^2}q_1 = \frac{n(1-\phi_0)}{\delta_0^2\phi_0}q_0 P_0, \tag{7.96}$$

$$\frac{1}{r}\frac{\mathrm{d}}{\mathrm{d}r}(rq_1) + P_1 = -\frac{m(1-\phi_0)}{\phi_0}P_0^2, \tag{7.97}$$

with boundary conditions

$$P_1(1) = -P_0'(1) - 2\mathcal{B}_0, \quad P_1 \to 0 \text{ as } r \to \infty. \tag{7.98}$$

This problem can be solved numerically, and $h'(1)$ calculated. While an explicit expression for $h'(1)$ has not been found, asymptotic expressions can be found. For $\delta_0 \ll 1$, $h'(1)$ is given asymptotically by

$$h'(1) \sim \left(-\mathcal{B}_0 + \frac{2}{3}\mathcal{B}_0^2(n+m)\frac{1-\phi_0}{\phi_0}\right)\delta_0 + \mathcal{O}\left(\delta_0^2\right). \tag{7.99}$$

For $\delta_0 \gg 1$ there is the corresponding expression

$$h'(1) \sim -1 + \frac{n\mathcal{B}_0}{2}\frac{1-\phi_0}{\phi_0} + \mathcal{O}\left(\frac{1}{\log\delta_0}\right), \tag{7.100}$$

although this is not as useful in practice due to the slow decaying $1/\log\delta_0$ terms.

Of particular interest is finding the parameter values for which $h'(1) = 1$. These divide parameter space into those regions which allow small amplitude waves of elevation and those which allow small amplitude waves of depression. A calculation of this line for particular parameter values is shown in Figures 7.1 and 7.2.

7.7 Results

We now turn to numerical solutions of the equations. Figures 7.3 and 7.4 plot the dispersion relationship (7.62) relating wave speed to wave amplitude for some example parameter values. These dispersion curves demonstrate that in some parameter regimes waves of elevation are admissible solutions whereas in other regimes waves of depression are.

Figure 7.5 shows streamline plots inside the conduit for different parameter values. These are plotted in the frame moving with the wave and so are also particle paths. In each case the background matrix bulk viscosity is assumed to be the same as the shear viscosity, namely $\mathcal{B}_0 = 3/7$. The key difference between the figures on the left on those on the right is the value of the background porosity: $\phi_0 = 0.01$ for the figures on the left, $\phi_0 = 0.3$ for those on the right. These differences in porosity lead to waves of depression (here $A = 0.95$) being possible in the first case, and waves of elevation (here $A = 1.05$) possible in the second.

A number of the features seen in the streamline plots are as expected from the small amplitude analysis of the previous section. Figure 7.1 shows that we expect small amplitude waves of elevation and depression in the parameter regimes we see them. Moreover
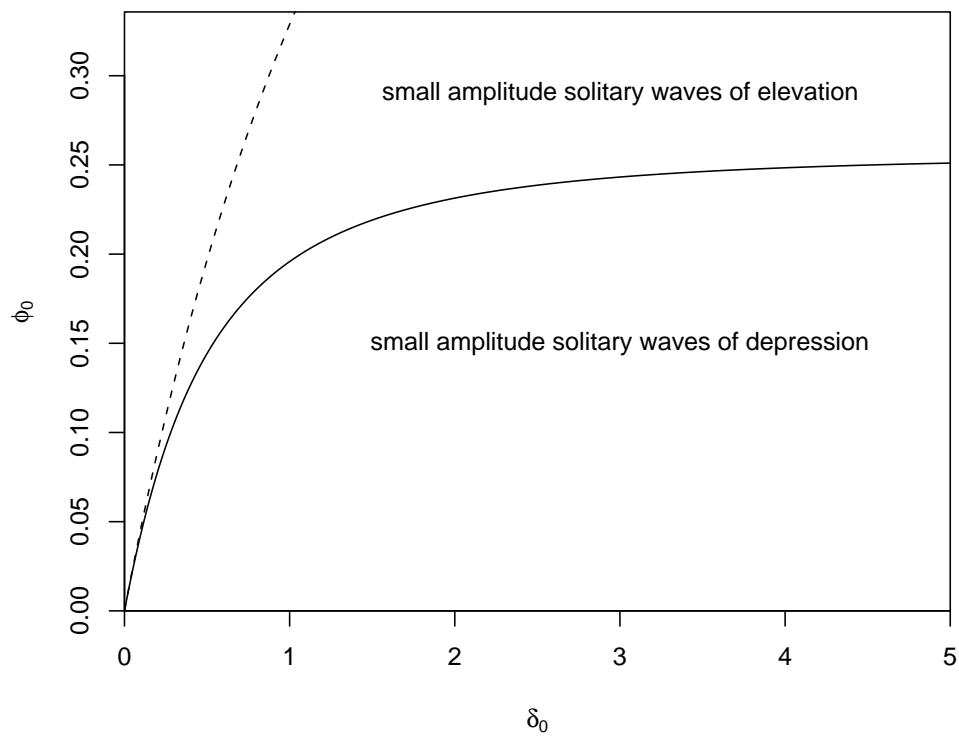
Figure 7.1: Critical line where $h'(1) = 1$ for $n = 3$, $m = 1$ and $\mathcal{B}_0 = 3/7$. Above the solid line $h'(1) < 1$ and there are small amplitude solitary waves of elevation. Below the solid line $h'(1) > 1$ and there are small amplitude solitary waves of depression. The dotted line shows the small δ_0 asymptotic result given by (7.99).
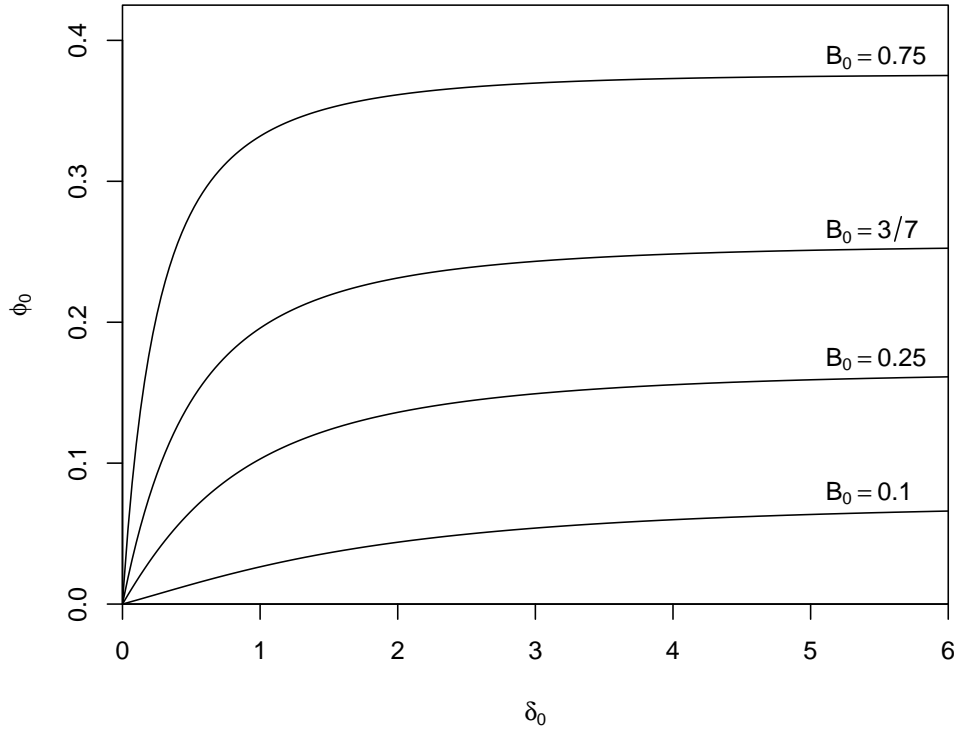
Figure 7.2: As Figure 7.1 showing the critical lines where $h'(1) = 1$ for $n = 3$, $m = 1$, but for different values of $\mathcal{B}_0$.

we expect the waves of elevation to have a centre stagnation point on $\xi = 0$, and the waves of depression to have a saddle stagnation point on $\xi = 0$. As δ_0 is increased the stagnation point at $\xi = 0$ moves towards the conduit wall as expected from (7.68) and (7.69), although note that it never moves outside the conduit. It is also worth noting that the small amplitude analysis is a good approximation for the waves of elevation, but less so for the waves of depression. The reason for this is that $h(b)$ grows much more rapidly for the waves of depression due to the large changes in permeability.

Figure 7.6 shows the corresponding contour plots of porosity outside the conduit. The porosity profiles for the waves of depression are quite different from those of elevation. Notably for the waves of depression there is a large central region where the porosity varies very little with ξ. This is due to the fact that as the matrix compacts it becomes more difficult to extract further melt as the bulk viscosity of the matrix increases. As δ_0 is increased there is remarkably little qualitative change in the porosity profiles, although the changes in porosity do extend further into the matrix as expected.

Of particular interest is the behaviour for $\delta_0 \gg 1$, since it is thought that in the mantle regime the conduits will have radius much smaller than the compaction length. Figure 7.7 plots streamlines inside the conduit and melt flow lines outside the conduit for $\delta_0 = 100$. The pictures remain qualitatively similar for larger values of δ_0. Note that the stagnation points lie very close to the edge of the conduit. Figure 7.8 shows the corresponding porosity profiles.

The flows shown in Figure 7.7 can be divided into three distinct regions. First is the fast flowing inner region in the centre of the conduit which has no interaction with the surrounding matrix. Second is the middle region where melt flows between the conduit and the matrix. Third is the slow flowing outer region which is solely in the matrix.
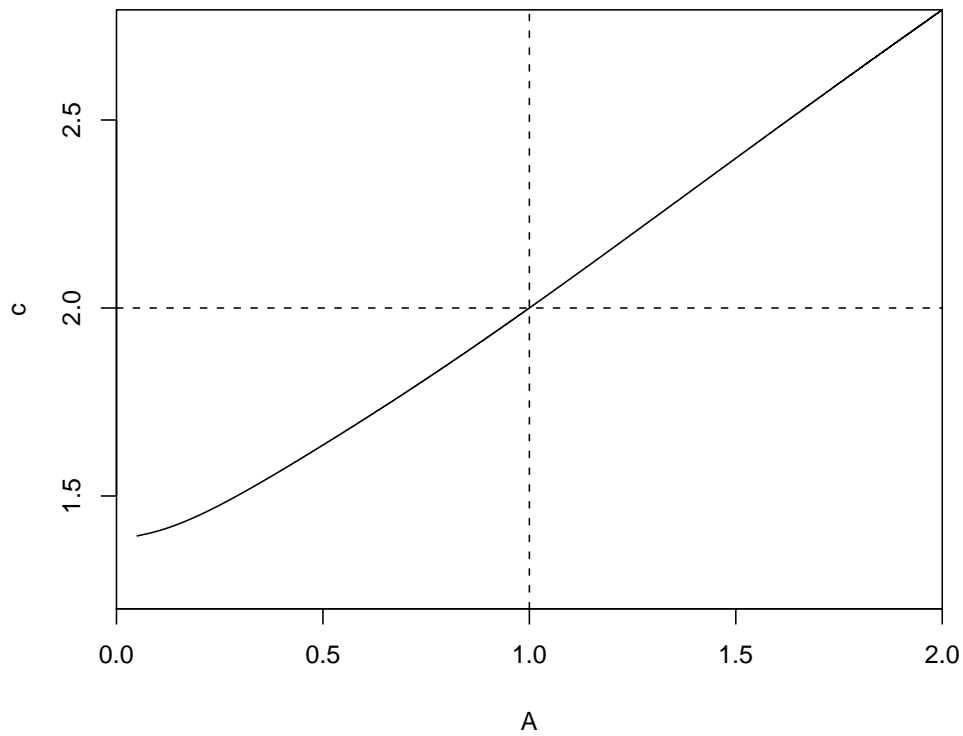
Figure 7.3: Plot of the dispersion relationship (7.62), where $\phi_0 = 0.3$, $\delta_0 = 1.0$, $\mathcal{B}_0 = 3/7$, $n = 3$ and $m = 1$. Here c is plotted as a function of A the non-dimensional conduit radius rather than the non-dimensional conduit area $B = A^2$. The condition $c(B) \geq c(b)$ for all values of b on the wave implies that waves of elevation (on this plot $1.0 < A < 2.0$) are admissible solutions, but waves of depression (on this plot $0.0 < A < 1.0$) are not.
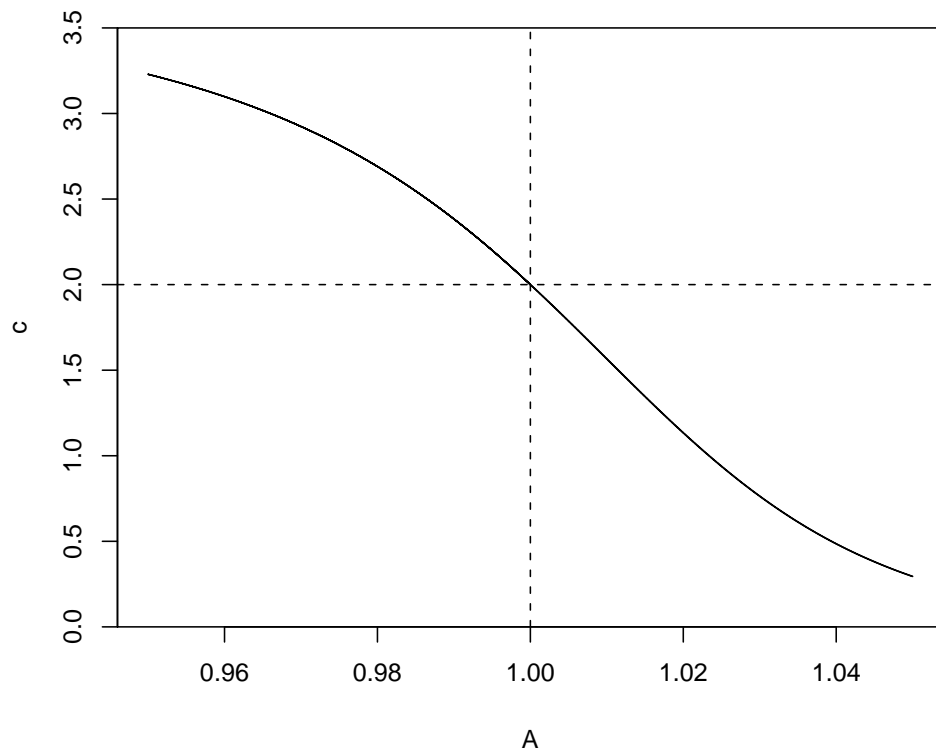
Figure 7.4: Plot of the dispersion relationship (7.62) as in Figure 7.3, except now $\phi_0 = 0.01$. The condition $c(B) \geq c(b)$ for all values of b on the wave now implies that waves of depression (on this plot $0.95 < A < 1.0$) are admissible solutions, but waves of elevation (on this plot $1.0 < A < 1.05$) are not.
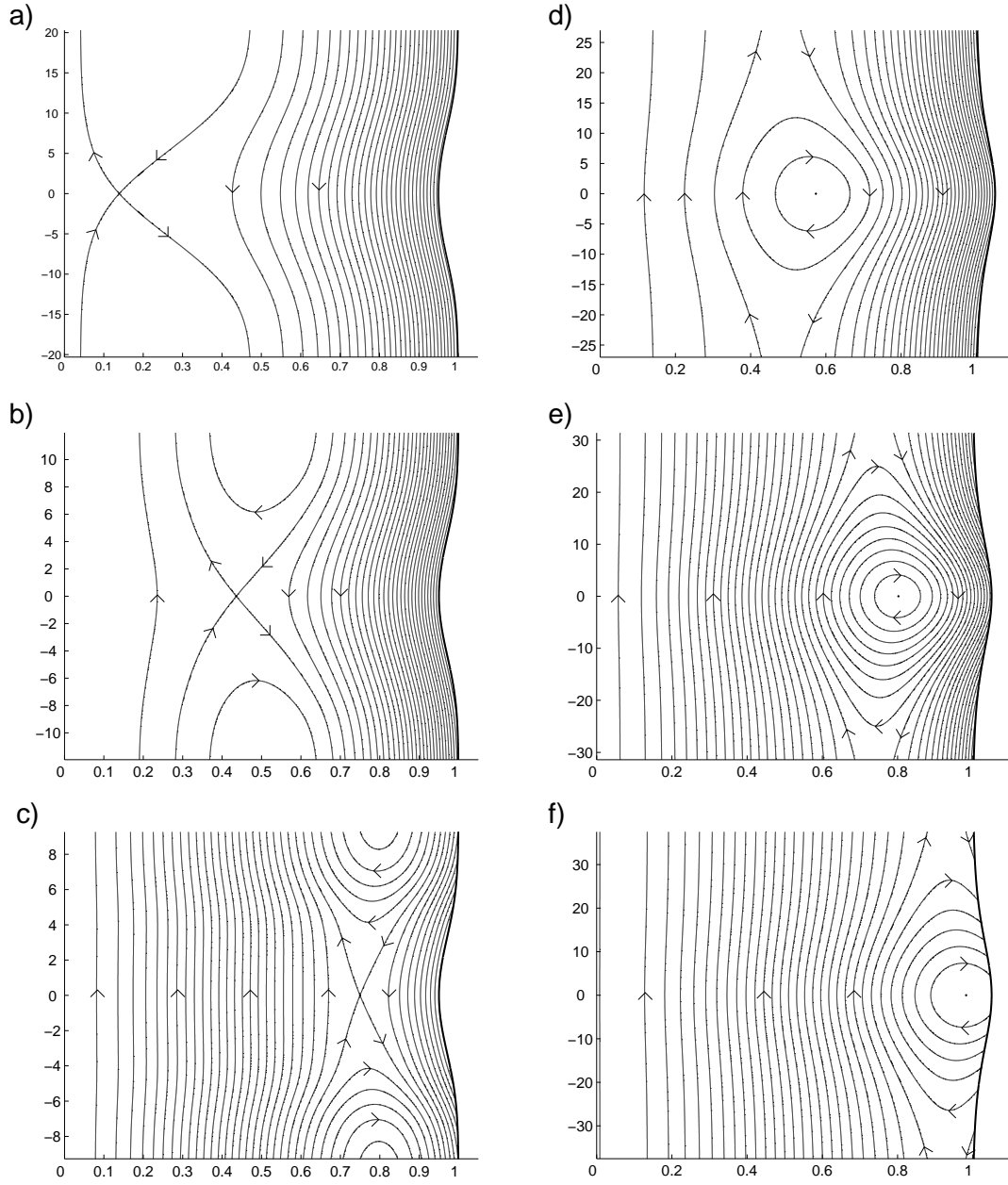
Figure 7.5: Streamlines inside the conduit in the frame moving with the wave. The thicker lines show the outline of each wave. Note that the r (horizontal) and ξ (vertical) axes are non-dimensionalised on a_0 and z_0 respectively. In all cases $n = 3$, $m = 1$, and $\mathcal{B}_0 = 3/7$. a), b) and c) show waves of depression where $\phi_0 = 0.01$ and amplitude $A = 0.95$. d), e) and f) show waves of elevation where $\phi_0 = 0.3$ and amplitude $A = 1.05$. a) and d) have $\delta_0 = 0.333$, b) and e) have $\delta_0 = 1.0$, and c) and f) have $\delta_0 = 3.33$. (cf. Richardson et al. [1996] Figure 3.)
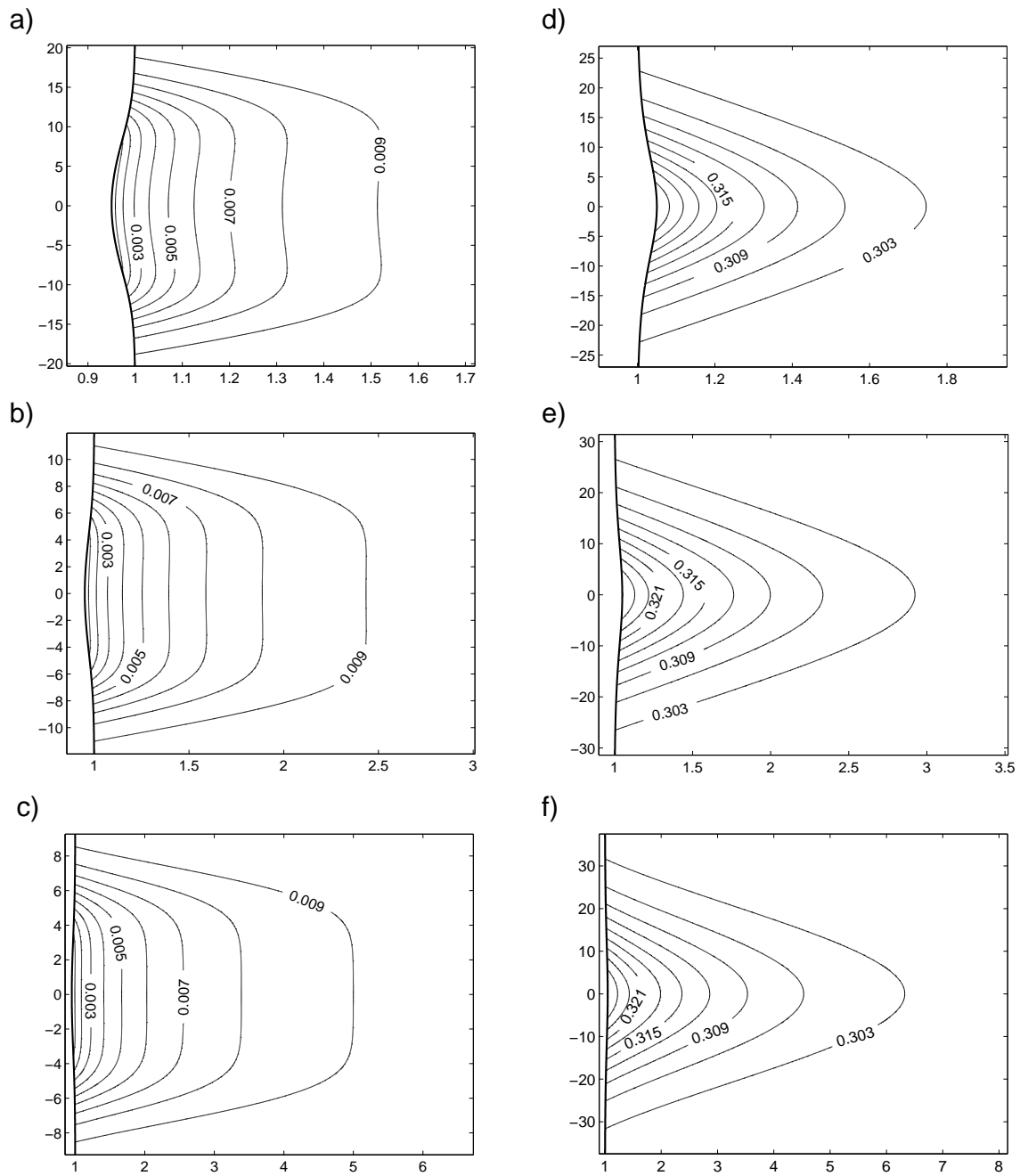
Figure 7.6: Unscaled porosity profiles outside the conduit for parameter values given in Figure 7.5. Diagrams on the left have a background porosity $\phi_0 = 0.01$, those on the right $\phi_0 = 0.3$. Note that waves of depression are surrounded by a region of lower porosity relative to the background value, whereas waves of elevation have a region of higher porosity.
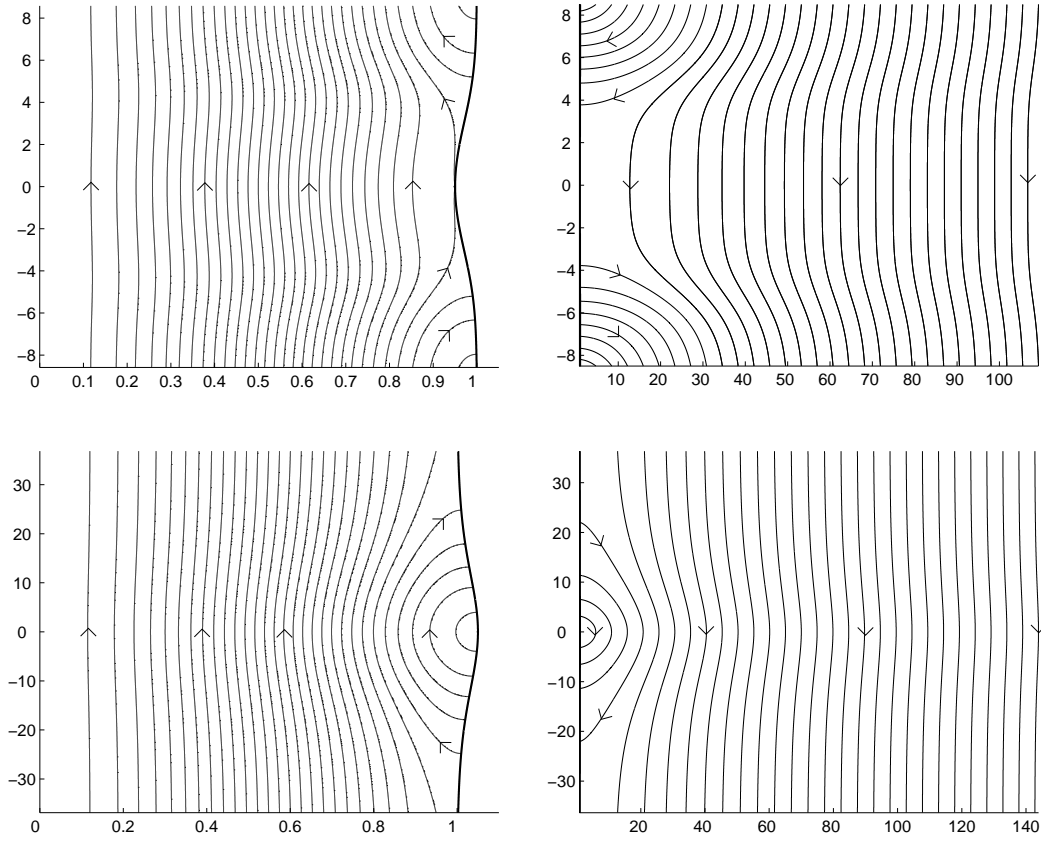
Figure 7.7: Streamlines inside the conduit (left diagrams) and melt flow lines outside the conduit (right diagrams) in the frame moving with the wave for $n = 3$, $m = 1$, $\mathcal{B}_0 = 3/7$, and $\delta_0 = 100$. The melt flow lines outside the conduit are found by directly following the velocity field $\mathbf{v}$. The top diagrams show a wave of depression with $\phi_0 = 0.01$, $A = 0.95$, and the bottom diagrams a wave of elevation with $\phi_0 = 0.3$, $A = 1.05$. (cf. Richardson et al. [1996] Figure 4.)

The important difference between the waves of elevation and depression only occurs in this second region. For a wave of elevation there are closed loops with melt constantly circulating between the conduit and the matrix, and thus a region of conserved fluid travelling with the wave. There are no such loops for a wave of depression, and each flow line crosses the boundary of the conduit only once.

   Richardson et al. [1996] did not find solitary waves of depression, despite discussing mantle parameter values $n = 3$, $m = 0$, $\mathcal{B}_0 = 3/7$, $\phi_0 = 0.01$, $\delta_0 \gg 1$ in which the present analysis shows are the solutions. Two reasons can be suggested for this. First they found a linear solution whereby compaction was allowed to take place, but the variation in porosity that this generates was ignored. This is essentially the same as setting $n = m = 0$, and the linear solution takes on a similar form to the solution of the leading order problem (7.92-7.94). In particular, this linear solution has $h'(1) \leq 0$, and only allows waves of elevation. However, solitary waves are fundamentally a nonlinear phenomenon, and even for a solitary wave of infinitesimal amplitude the solution depends crucially on the details of this nonlinearity. The linear approximation would be fine were we just considering linear waves (7.54), but for the nonlinear solitary waves it is certainly not.
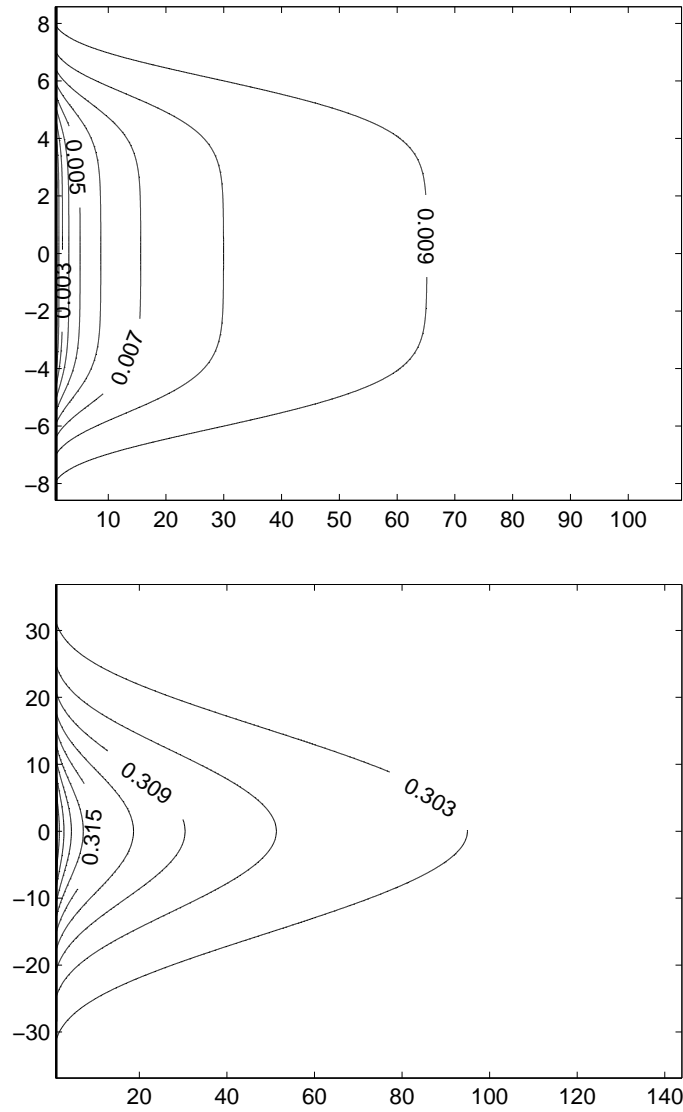
Figure 7.8: Unscaled porosity profiles for the parameter values given in Figure 7.7. The top diagram has background porosity $\phi_0 = 0.01$, the bottom diagram has $\phi_0 = 0.3$. (cf. Richardson et al. [1996] Figure 6.)

Secondly Richardson et al. [1996] claimed to have solved the full problem including this nonlinearity numerically, and yet found very similar results to their linear solution, and did not find the waves of depression. The reason for this is likely to be a belief that the porosity ϕ should simply scale with ϕ_0 as it does for the similar problem of porosity solitary waves. We assume this to be the case since the numerical value of ϕ_0 is not mentioned until near the end of the paper, and their porosity profiles (their Figure 6) are claimed to be contoured in levels of ϕ_0. In fact, as this analysis has shown, there is no such simple scaling and the problem depends nonlinearly on ϕ_0. Since the problem was assumed to scale with ϕ_0 it is likely that the numerical simulations simply did not include this as a parameter, effectively setting its value to 1. For large values of the background porosity ϕ_0 we are in the regime where only waves of elevation are possible (Figure 7.1) which explains why these were the only waves they found. This is seen in the $\phi_0 = 0.3$ plots, which are very similar to those in Richardson et al. [1996]. $\phi_0 = 0.3$ is rather high to be observed in practice: at this porosity the matrix will start to disaggregate, and the model will break down. It is possible to have more realistic background porosities and still get waves of elevation, but only if $\mathcal{B}_0$ is much smaller (Figure 7.2). However, this will produce waves of elevation that are more similar to the impermeable case with the circulating region lying further inside the conduit.

It is important to note that this transition from waves of elevation to waves of depression is found in many other solitary wave problems, including Russell's original problem of solitary waves in a channel. By increasing the surface tension of the fluid in the channel, solitary waves of depression rather than elevation can be found [Miles, 1980].

7.8 Conclusions

There are a number of pieces of analysis that should still be done. Firstly, it is important to check the conditions under which the present analysis holds, and particularly the long wavelength approximation. This approximation is justified provided the non-dimensional parameter μ/η is sufficiently small, but it is important to determine how sufficiently small in terms of the other non-dimensional parameters. This isn't quite as straightforward as proposed by Richardson et al. [1996] as in general it depends on the amplitude of the wave and some of the nonlinear details, for example the wavelength of small amplitude waves depends on $h'(1)$ (see (7.77)). However, (7.77) also shows that for sufficiently small amplitude waves, the wavelength of the waves will be sufficiently long to justify the approximation. As Richardson et al. [1996] mention, an interesting problem for further study would be where this approximation breaks down. Here we expect the conduit to become a less important feature, and the waves to become more like the porosity solitary waves ('magmons').

There may be some useful analysis of large amplitude waves that can be done. It is interesting to note that the dispersion relation comes to a peak for small values of a in some parameter regimes, leading to a minimum amplitude for the solitary wave. Small values of a are fairly difficult to tackle numerically because the very small porosities involved cause the numerical scheme to fail. It may be possible to analyse this situation more closely by considering similarity solutions.

The time dependent problem is also worth further study. For the standard conduit equations, an initial disturbance in the channel breaks up into a train of solitary waves, and it will be interesting to see if this also occurs in the modified problem. Certainly for small amplitude disturbances we expect a train of solitary waves, since the governing equations can be approximated by the KdV equations. However, for larger amplitude disturbances

the behaviour may well be different, and it should be straightforward to develop numerical simulations to test this.

It is not known whether the open conduits that have been studied here exist in the mantle. An alternative is localised vertical regions of high porosity. These would be expected to behave similarly, but without doing the full analysis of this situation we do not know in which parameter regimes we expect waves of depression or elevation. The relevant important non-dimensional parameter may be the relative porosity of the high porosity region to the low porosity region, rather than simply the background porosity ϕ_0 as it is for the pure melt conduits. In which case, we might expect waves of elevation to be more likely if the porosity contrast is less.

The main geochemical result of Richardson et al. [1996] is unchanged by this new possibility of solitary waves of depression. It is still the case that there is fast flowing central region of the conduit which will not interact greatly with the surrounding matrix. This allows a radioactive disequilibrium signal to propagate unaffected to the Earth's surface in a short time, which we know from the observations must be the case. Of secondary importance is the trapped region of melt which travels with the wave, which it now appears may not exist.

The study of the dynamics of melt conduits perhaps misses the more important physical questions: how does melt begin to separate from its residue, and what are the processes that lead to the network of channels? Here it is necessary to investigate physics on the grain scale, which are averaged out when considering the compaction equations.

Notation

A	non-dimensional wave amplitude (radius)		-
a	conduit radius		m
a_0	steady conduit radius		m
$\mathcal{B}_0$	background matrix viscosity ratio	(7.23)	-
B	non-dimensional wave amplitude (area)	$= A^2$	-
b	conduit area		m^2
b_0	steady conduit area		m^2
$\mathcal{C}$	matrix compaction rate	$= \nabla \cdot \mathbf{V}$	s^{-1}
$\mathcal{C}_0$	compaction scale	(7.24)	s^{-1}
c	wave velocity		m s^{-1}
c_0	velocity scale	(7.45)	m s^{-1}
$\mathbf{E}$	matrix rate of strain tensor	(7.6)	s^{-1}
g	acceleration due to gravity		m s^{-2}
k	linear wave number		m^{-1}
k_ϕ	permeability		m^2
k_0	background permeability		m^2
n	permeability exponent	(7.8)	-
m	matrix bulk viscosity exponent	(7.8)	-
$\mathcal{P}$	piezometric pressure	$= p + \rho_f g z$	Pa
$\mathcal{P}_0$	piezometric pressure scale	(7.47)	Pa
P	excess pressure	$= p + \bar{\rho}_0 g z$	Pa
P_0	excess pressure scale	(7.24)	Pa
Q	conduit volume flux		m^3 s^{-1}
Q_0	volume flux scale	(7.46)	m^3 s^{-1}
$\mathbf{q}$	Darcy flux	$= \phi(\mathbf{v} - \mathbf{V})$	m s^{-1}

r	radius		m
t	time		s
t_0	time scale	(7.44)	s
u_r	conduit radial velocity		$\mathrm{m\,s^{-1}}$
u_z	conduit vertical velocity		$\mathrm{m\,s^{-1}}$
$\mathbf{V}$	matrix velocity		$\mathrm{m\,s^{-1}}$
V_0	matrix velocity scale	(7.24)	$\mathrm{m\,s^{-1}}$
$\mathbf{v}$	melt velocity		$\mathrm{m\,s^{-1}}$
z	vertical co-ordinate		m
z_0	vertical length scale	(7.43)	m
$\Delta\rho$	density difference	$= \rho_s - \rho_f$	$\mathrm{kg\,m^{-3}}$
δ_0	compaction length / a_0	(7.23)	-
ζ_ϕ	matrix bulk viscosity		Pa s
ζ_0	background matrix bulk viscosity		Pa s
η	matrix shear viscosity		Pa s
μ	melt viscosity		Pa s
ξ	moving frame co-ordinate	$= z - ct$	m
ρ_f	melt density		$\mathrm{kg\,m^{-3}}$
ρ_s	matrix density		$\mathrm{kg\,m^{-3}}$
$\bar\rho_0$	mean background density	$= \phi_0\rho_f + (1 - \phi_0)\rho_s$	$\mathrm{kg\,m^{-3}}$
$\boldsymbol{\sigma}$	matrix stress tensor		Pa
ϕ	porosity		-
ϕ_0	background porosity		-
$\boldsymbol{\omega}$	matrix vorticity	$= \nabla \times \mathbf{V}$	$\mathrm{s^{-1}}$
ω	linear wave frequency		$\mathrm{s^{-1}}$

Chapter 8

Conclusions

This thesis has successfully tackled a few narrow problems associated with the study of the Earth's mantle. The work has been driven by the geochemical, geophysical and geological observations, and in particular by radiogenic isotope geochemistry. The goal has been to better understand the physical processes of melting, melt migration, and mantle convection. Key to this has been the construction of highly idealised and analytically tractable models to explore the essential physical processes involved without including unnecessary detail.

Statistical modelling has formed the core of the thesis, and has proved a powerful tool for understanding the isotopic heterogeneity of mid-ocean ridge basalt. The observations have been used to constrain aspects of mantle convection and melt generation. For example, a 0.5 Ga time scale for remelting was inferred from the observed 2.0 Ga lead-lead pseudo-isochron, and small melt fractions were shown to be necessary to generate the observed heterogeneity. The statistical box model has led to a new interpretation of mantle pseudo-isochrons in terms of age distributions, and has also helped unify the different approaches to modelling mantle geochemistry, from simple geochemical reservoir models to complex geodynamic convection simulations. There are many possibilities for future work based on this statistical approach.

Continuum mechanics formed the basis for the last part of the thesis, and in particular the fluid dynamics of two phase flow was studied to model melt migration in the partially molten regions of the mantle. Knowledge of fluid dynamics is vital to understanding the physical processes that occur in the Earth. Indeed, much of our current understanding of mantle convection, from the fact the convection actually occurs, to the existence of plumes, comes from simple fluid dynamical considerations.

While the problems focused on in this thesis have been rather narrow and specialised, the techniques and methods used have much wider application. Isotope geochemistry pervades all modern Earth science. Indeed, as was remarked by Karl K. Turekian, "No field is immune to the invasion of isotope geochemistry". While the geochemistry of the solid earth has been the focus of this thesis, similar geochemical arguments are exploited in a diverse range of disciplines such as cosmology, oceanography and climate science. A number of these disciplines are, like mantle geochemistry, observation-rich and model-poor, and there is plenty of scope for some good applied mathematics to be done. The powerful techniques from continuum mechanics and statistics can be brought to bear on so many different problems, that no field is immune to the invasion of applied mathematics.

Appendix A

Data sources

Table A.1 gives the raw data from which some of the data in Table 2.1 was calculated. The weighted partition coefficients P and D given in Table 2.1 are a linear combination of 60% the garnet and 40% the spinel coefficients given in Table A.1.

The MORB data set was a compilation by A. Stracke from the work of Bach et al. [1994], Chauvel and Blichert-Toft [2001], Dosso et al. [1988, 1991, 1993, 1999], Douglass et al. [1999], Fontignie and Schilling [1996], Frey et al. [1993], Hamelin and Allègre [1985], Hamelin et al. [1986], Hanan et al. [1986], Hegner and Tatsumoto [1987], Ito et al. [1987], Kempton et al. [2000], Klein et al. [1988, 1991], MacDougall and Lugmair [1986], Mahoney et al. [1989, 1994, 2002], Mertz and Haase [1997], Mertz et al. [1991], Michard et al. [1986], Newsom et al. [1986], Price et al. [1986], Pyle et al. [1992], Rehkämper and Hofmann [1997], Salters [1996], Salters and Hart [1991], Salters and White [1998], Schilling et al. [1994, 1999], Shirey et al. [1987], Sims et al. [2002], Vlastelic et al. [1999], Wendt et al. [1999], White et al. [1987], Yu et al. [1997].

Table A.1: Elemental concentrations, garnet/spinel partition coefficients and isotopic abundances.

	^{147}Sm	^{87}Rb	^{176}Lu	^{232}Th	^{238}U	^{235}U
parent isotope p	^{147}Sm	^{87}Rb	^{176}Lu	^{232}Th	^{238}U	^{235}U
daughter isotope d	^{143}Nd	^{87}Sr	^{176}Hf	^{208}Pb	^{206}Pb	^{207}Pb
reference isotope d'	^{144}Nd	^{86}Sr	^{177}Hf	^{204}Pb	^{204}Pb	^{204}Pb
parent element conc. /ppm[a]	0.272	0.0930	0.0550	0.00420	0.00170	0.00170
daughter element conc. /ppm[a]	0.677	14.0	0.237	0.0150	0.0150	0.0150
par./dau. elemental ratio[a]	0.402	0.00664	0.232	0.280	0.113	0.113
p isotope abundance /%[b]	15.0	27.84	2.60	100.0	99.27	0.72
d' isotope abundance /%[b]	23.8	9.86	18.6	1.40	1.40	1.40
parent D_{spinel}[a]	0.03477	0.00039	0.05077	0.00016	0.00012	0.00012
parent P_{spinel}[a]	0.10634	0.00056	0.12193	0.00052	0.00036	0.00036
parent D_{garnet}[a]	0.04760	0.00039	0.66733	0.00028	0.00118	0.00118
parent P_{garnet}[a]	0.15028	0.00062	2.18507	0.00091	0.00391	0.00391
daughter D_{spinel}[a]	0.02778	0.03454	0.02888	0.03198	0.03198	0.03198
daughter P_{spinel}[a]	0.08592	0.10403	0.08878	0.08393	0.08393	0.08393
daughter D_{garnet}[a]	0.02801	0.02277	0.06939	0.02186	0.02186	0.02186
daughter P_{garnet}[a]	0.08745	0.06666	0.22516	0.05348	0.05348	0.05348

(a) McKenzie and O'Nions [1991]
(b) Lide [2003]

References

A. Agranier, J. Blichert-Toft, D. Graham, V. Debaille, P. Schiano, and F. Albarède. The spectra of isotopic heterogeneities along the mid-Atlantic Ridge. *Earth Planet. Sci. Lett.*, 238:96–109, 2005. doi:10.1016/j.epsl.2005.07.011.

F. Albarède. The survival of mantle geochemical heterogeneities. In R. D. van der Hilst, J. D. Bass, J. Matas, and J. Trampert, editors, *Earth's Deep Mantle: Structure, Composition, and Evolution*, volume 160, pages 27–46. AGU Geophysical Monograph Series, 2005. ISBN 0-87590-425-2.

F. Albarède. Radiogenic ingrowth in systems with multiple reservoirs: applications to the differentiation of the mantle-crust system. *Earth Planet. Sci. Lett.*, 189:59–73, 2001. doi:10.1016/S0012-821X(01)00350-8.

F. Albarède. *Geochemistry: An Introduction*. Cambridge University Press, 2003. ISBN 0-521-81468-5. doi:10.2277/0521891485.

C. J. Allègre. Comportement des systemes U-Th-Pb dan le manteau superieur et modele d'evolution de ce dernier au cours des temps geologiques. *Earth Planet. Sci. Lett.*, 5: 261–269, 1969. doi:10.1016/S0012-821X(68)80050-0.

C. J. Allègre and E. Lewin. Isotopic systems and stirring times of the earth's mantle. *Earth Planet. Sci. Lett.*, 136:629–646, 1995a. doi:10.1016/0012-821X(95)00184-E.

C. J. Allègre and E. Lewin. Scaling laws and geochemical distributions. *Earth Planet. Sci. Lett.*, 132:1–13, 1995b. doi:10.1016/0012-821X(95)00049-I.

R. L. Armstrong. A model for the evolution of strontium and lead isotopes in a dynamic earth. *Rev. Geophys.*, 6:175–199, 1968.

W. Bach, E. Hegner, J. Erzinger, and M. Satir. Chemical and isotopic variations along the superfast spreading East Pacific Rise from 6°S to 30°S. *Contrib. Mineral. Petr.*, 116:365–380, 1994.

V. Barcilon and O. M. Lovera. Solitary waves in magma dynamics. *J. Fluid Mech.*, 204: 121–133, 1989. doi:10.1017/S0022112089001680.

V. Barcilon and F. M. Richter. Nonlinear waves in compacting media. *J. Fluid Mech.*, 164:429–448, 1986. doi:10.1017/S0022112086002628.

S. Blinnikov and R. Moessner. Expansions for nearly gaussian distributions. *Astron. Astrophys. Suppl. Ser.*, 130:193–205, 1998. doi:10.1051/aas:1998221.

J. Blundy and B. Wood. Partitioning of trace elements between crystals and melts. *Earth Planet. Sci. Lett.*, 210:383–397, 2003. doi:10.1016/S0012-821X(03)00129-8.

C. Brooks, S. R. Hart, A. Hofmann, and D. E. James. Rb-Sr mantle isochrons from oceanic regions. *Earth Planet. Sci. Lett.*, 32:51–61, 1976. doi:10.1016/0012-821X(76)90184-9.

C. Chauvel and J. Blichert-Toft. A hafnium isotope and trace element perspective on melting of the depleted mantle. *Earth Planet. Sci. Lett.*, 190:137–151, 2001. doi:10.1016/S0012-821X(01)00379-X.

U. R. Christensen and A. W. Hofmann. Segregation of subducted oceanic crust in the convecting mantle. *J. Geophys. Res.*, 99:19867–19884, 1994. doi:10.1029/94JB03403.

G. F. Davies. Stirring geochemistry in mantle convection models with stiff plates and slabs. *Geochim. Cosmochim. Acta*, 66:3125–3142, 2002. doi:10.1016/S0016-7037(02)00915-8.

V. Debaille, J. Blichert-Toft, A. Agranier, R. Doucelance, P. Schiano, and F. Albarède. Geochemical component relationships in MORB from the Mid-Atlantic Ridge, $22-35°N$. *Earth Planet. Sci. Lett.*, 241:844–862, 2006. doi:10.1016/j.epsl.2005.11.004.

A. P. Dickin. *Radiogenic isotope geochemistry*. Cambridge University Press, 2nd edition, 2005. ISBN 0-521-53017-2. doi:10.2277/0521530172.

K. E. Donnelly, S. L. Goldstein, C. H. Langmuir, and M. Spiegelman. Origin of enriched ocean ridge basalts and implications for mantle dynamics. *Earth Planet. Sci. Lett.*, 226: 347–366, 2004. doi:10.1016/j.epsl.2004.07.019.

L. Dosso, H. Bougault, P. Beuzart, J. Y. Calvez, and J. L. Joron. The geochemical structure of the Southeast Indian Ridge. *Earth Planet. Sci. Lett.*, 88:47–59, 1988. doi:10.1016/0012-821X(88)90045-3.

L. Dosso, B. B. Hanan, H. Bougault, J. G. Schilling, and J. L. Joron. Sr-Nd-Pb geochemical morphology between 10°N and 17°N on the Mid-Atlantic Ridge - A new MORB isotope signature. *Earth Planet. Sci. Lett.*, 106:29–43, 1991. doi:10.1016/0012-821X(91)90061-L.

L. Dosso, H. Bougault, and J. L. Joron. Geochemical morphology of the North Mid-Atlantic Ridge, 10°-24°N - Trace element-isotope complementarity. *Earth Planet. Sci. Lett.*, 120:443–462, 1993. doi:10.1016/0012-821X(93)90256-9.

L. Dosso, H. Bougault, C. Langmuir, C. Bollinger, O. Bonnier, and J. Etoubleau. The age and distribution of mantle heterogeneity along the Mid-Atlantic Ridge (31-41°N). *Earth Planet. Sci. Lett.*, 170:269–286, 1999. doi:10.1016/S0012-821X(99)00109-0.

J. Douglass, J. G. Schilling, and D. Fontignie. Plume-ridge interactions of the Discovery and Shona mantle plumes with the Southern Mid-Atlantic Ridge (40°-55°S). *J. Geophys. Res.-Solid.*, 104:2941–2962, 1999. doi:10.1029/98JB02642.

G. Faure. *Principles of Isotope Geology*. Wiley, 2nd edition, 1986. ISBN 0-471-86412-9.

S. Ferrachat and Y. Ricard. Mixing properties in the earth's mantle: Effects of the viscosity stratification and of oceanic crust segregation. *Geochem. Geophys. Geosyst.*, 2, 2001. doi:10.1029/2000GC000092.

D. Fontignie and J. G. Schilling. Mantle heterogeneities beneath the South Atlantic: A Nd-Sr-Pb isotope study along the Mid-Atlantic Ridge (3°S-46°S). *Earth Planet. Sci. Lett.*, 142:209–221, 1996. doi:10.1016/0012-821X(96)00079-9.

F. Frey, N. Walker, D. Stakes, S. Hart, and R. Nielsen. Geochemical characteristics of basaltic glasses from the Amar and FAMOUS Axial Valleys, Mid-Atlantic Ridge (36°-37°N) - petrogenetic implications. *Earth Planet. Sci. Lett.*, 115:117–136, 1993. doi:10.1016/0012-821X(93)90217-W.

S. J. G. Galer and R. K. O'Nions. Residence time of thorium, uranium and lead in the mantle with implications for mantle convection. *Nature*, 316:778–782, 1985. doi:10.1038/316778a0.

S. J. G. Galer, S. L. Goldstein, and R. K. O'Nions. Limits on chemical and convective isolation in the Earth's interior. *Chem. Geol.*, 75:257–290, 1989. doi:10.1016/0009-2541(89)90001-6.

D. W. Graham, J. Blichert-Toft, C. J. Russo, K. H. Rubin, and F. Albarède. Cryptic striations in the upper mantle revealed by hafnium isotopes in southeast Indian ridge basalts. *Nature*, 440:199–202, 2006. doi:10.1038/nature04582.

T. H. Green. Experimental studies of trace-element partitioning applicable to igneous petrogenesis: Sedona 16 years later. *Chem. Geol.*, 117:1–36, 1994. doi:10.1016/0009-2541(94)90119-8.

B. Hamelin and C. J. Allègre. Large-scale regional units in the depleted upper mantle revealed by an isotope study of the southwest Indian Ridge. *Nature*, 315:196–199, 1985. doi:10.1038/315196a0.

B. Hamelin, B. Dupre, and C. J. Allègre. Pb-Sr-Nd isotopic data of Indian-Ocean Ridges - New evidence of large-scale mapping of mantle heterogeneities. *Earth Planet. Sci. Lett.*, 76:288–298, 1986. doi:10.1016/0012-821X(86)90080-4.

B. B. Hanan, R. H. Kingsley, and J. G. Schilling. Pb isotope evidence in the South-Atlantic for migrating ridge hotspot interactions. *Nature*, 322:137–144, 1986. doi:10.1038/322137a0.

G. N. Hanson. Geochemical variation of sub-oceanic mantle. *J. Geol. Soc. Lond.*, 134: 235–53, 1977.

G. H. Hardy, J. E. Littlewood, and G. Pólya. *Inequalities*. Cambridge University Press, 1934.

E. Hegner and M. Tatsumoto. Pb, Sr, and Nd isotopes in basalts and sulphides from the Juan-de-Fuca Ridge. *J. Geophys. Res.-Solid.*, 92:11380–11386, 1987.

K. R. Helfrich and J. A. Whitehead. Solitary waves on conduits of buoyant fluid in a more viscous fluid. *Geophys. Astrophys. Fluid Dyn.*, 51:35–52, 1990.

A. W. Hofmann. Sampling mantle heterogeneity through oceanic basalts: Isotopes and trace elements. In H. Holland and K. K. Turekian, editors, *Treatise on Geochemistry*, volume 2, chapter 3, pages 61–101. Elsevier-Pergamon, Oxford, 2003. ISBN 0-08-043751-6.

E. Ito, W. M. White, and C. Gopel. The O, Sr, Nd and Pb isotope geochemistry of MORB. *Chem. Geol.*, 62:157–176, 1987. doi:10.1016/0009-2541(87)90083-0.

G. Ito and J. J. Mahoney. Flow and melting of a heterogeneous mantle: 1. Method and importance to the geochemistry of ocean island and mid-ocean ridge basalts. *Earth Planet. Sci. Lett.*, 230:29–46, 2005. doi:10.1016/j.epsl.2004.10.035.

S. B. Jacobsen and G. J. Wasserburg. The mean age of mantle and crustal reservoirs. *J. Geophys. Res.*, 84:7411–7427, 1979.

P. B. Kelemen, G. Hirth, N. Shimizu, M. Spiegelman, and H. J. B. Dick. A review of melt migration processes in the adiabatically upwelling mantle beneath oceanic spreading ridges. *Phil. Trans. R. Soc. Lond. A*, 355:283–318, 1997. doi:10.1098/rsta.1997.0010.

P. B. Kelemen, M. Braun, and G. Hirth. Spatial distribution of melt conduits in the mantle beneath oceanic spreading ridges: Observations from the Ingalls and Oman ophiolites. *Geochem. Geophys. Geosys.*, 1:844–862, 2000. doi:10.1029/1999GC000012.

J. B. Kellogg. *Towards an understanding of chemical and isotopic heterogeneity in the Earth's mantle.* PhD thesis, Harvard University, 2004.

J. B. Kellogg, S. B. Jacobsen, and R. J. O'Connell. Modelling the distribution of isotopic ratios in geochemical reservoirs. *Earth Planet. Sci. Lett.*, 204:183–202, 2002. doi:10.1016/S0012-821X(02)00981-0.

P. D. Kempton, J. G. Fitton, A. D. Saunders, G. M. Nowell, R. N. Taylor, B. S. Hardarson, and G. Pearson. The Iceland plume in space and time: a Sr-Nd-Pb-Hf study of the North Atlantic rifted margin. *Earth Planet. Sci. Lett.*, 177:255–271, 2000. doi:10.1016/S0012-821X(00)00047-9.

E. M. Klein, C. H. Langmuir, A. Zindler, H. Staudigel, and B. Hamelin. Isotope evidence of a mantle convection boundary at the Australian-Antarctic discordance. *Nature*, 333: 623–629, 1988. doi:10.1038/333623a0.

E. M. Klein, C. H. Langmuir, and H. Staudigel. Geochemistry of basalts from the Southeast Indian Ridge, 115°E-138°E. *J. Geophys. Res.-Solid.*, 96:2089–2107, 1991.

D. R. Lide. *CRC Handbook of Chemistry and Physics.* CRC Press, 84th edition, 2003. ISBN 0-849-30484-9.

J. D. MacDougall and G. W. Lugmair. Sr and Nd isotopes in basalts from the East Pacific Rise - significance for mantle heterogeneity. *Earth Planet. Sci. Lett.*, 77:273–284, 1986. doi:10.1016/0012-821X(86)90139-1.

J. J. Mahoney, J. H. Natland, W. M. White, R. Poreda, S. H. Bloomer, R. L. Fisher, and A. N. Baxter. Isotopic and geochemical provinces of the Western Indian-Ocean spreading centers. *J. Geophys. Res.-Solid.*, 94:4033–4052, 1989.

J. J. Mahoney, J. M. Sinton, M. D. Kurz, J. D. MacDougall, K. J. Spencer, and G. W. Lugmair. Isotope and trace-element characteristics of a super-fast spreading ridge - East Pacific Rise, 13-23°S. *Earth Planet. Sci. Lett.*, 121:173–193, 1994. doi:10.1016/0012-821X(94)90039-6.

J. J. Mahoney, D. W. Graham, D. M. Christie, K. T. M. Johnson, L. S. Hall, and D. L. Vonderhaar. Between a hotspot and a cold spot: Isotopic variation in the Southeast Indian Ridge asthenosphere, 86°E-118°E. *J. Petrol.*, 43:1155–1176, 2002. doi:10.1093/petrology/43.7.1155.

D. McKenzie. The generation and compaction of partially molten rock. *J. Petrol.*, 25: 713–765, 1984. doi:10.1093/petrology/25.3.713.

D. McKenzie. ^{230}Th-^{238}U disequilibrium and the melting processes beneath ridge axes. *Earth Planet. Sci. Lett.*, 72:149–157, 1985. doi:10.1016/0012-821X(85)90001-9.

D. McKenzie and R. K. O'Nions. Partial melt distributions from inversion of rare earth element concentrations. *J. Petrol.*, 32:1021–1091, 1991. doi:10.1093/petrology/32.5.1021.

D. McKenzie, A. Stracke, J. Blichert-Toft, F. Albarède, K. Grönvold, and R. K. O'Nions. Source enrichment processes responsible for isotopic anomalies in oceanic island basalts. *Geochim. Cosmochim. Acta*, 68:2699–2724, 2004. doi:10.1016/j.gca.2003.10.029.

A. Meibom and D. L. Anderson. The statistical upper mantle assemblage. *Earth Planet. Sci. Lett.*, 217:123–139, 2004. doi:10.1016/S0012-821X(03)00573-9.

D. F. Mertz and K. M. Haase. The radiogenic isotope composition of the high-latitude North Atlantic mantle. *Geology*, 25:411–414, 1997. doi:10.1130/0091-7613(1997)025⟨0411:TRICOT⟩2.3.CO;2.

D. F. Mertz, C. W. Devey, W. Todt, P. Stoffers, and A. W. Hofmann. Sr-Nd-Pb isotope evidence against plume asthenosphere mixing north of Iceland. *Earth Planet. Sci. Lett.*, 107:243–255, 1991. doi:10.1016/0012-821X(91)90074-R.

A. Michard, R. Montigny, and R. Schlich. Geochemistry of the mantle beneath the Rodriguez triple junction and the Southeast Indian Ridge. *Earth Planet. Sci. Lett.*, 78: 104–114, 1986. doi:10.1016/0012-821X(86)90176-7.

J. W. Miles. Solitary waves. *Ann. Rev. Fluid Mech.*, 12:11–43, 1980. doi:10.1146/annurev.fl.12.010180.000303.

D. T. Murphy, B. S. Kamber, and K. D. Collerson. A refined solution to the first terrestrial Pb-isotope paradox. *J. Petrol.*, 44:39–53, 2003. doi:10.1093/petrology/44.1.39.

H. E. Newsom, W. M. White, K. P. Jochum, and A. W. Hofmann. Siderophile and chalcophile element abundances in oceanic basalts, Pb-isotope evolution and growth of the Earth's core. *Earth Planet. Sci. Lett.*, 80:299–313, 1986. doi:10.1016/0012-821X(86)90112-3.

S. Y. Novak. On the distribution of the ratio of sums of random variables. *Theory Probab. Appl.*, 41:479–503, 1996. doi:10.1137/S0040585X97975228.

P. Olson and U. Christensen. Solitary wave propagation in a fluid conduit within a viscous matrix. *J. Geophys. Res.*, 91:6367–6374, 1986.

C. Patterson. Age of meteorites and the earth. *Geochim. Cosmochim. Acta*, 10:230–237, 1956. doi:10.1016/0016-7037(56)90036-9.

R. C. Price, A. K. Kennedy, M. Riggssneeringer, and F. A. Frey. Geochemistry of basalts from the Indian-Ocean triple junction - Implications for the generation and evolution of Indian-Ocean ridge basalts. *Earth Planet. Sci. Lett.*, 78:379–396, 1986. doi:10.1016/0012-821X(86)90005-1.

D. G. Pyle, D. M. Christie, and J. J. Mahoney. Resolving an isotopic boundary within the Australian-Antarctic discordance. *Earth Planet. Sci. Lett.*, 112:161–178, 1992. doi:10.1016/0012-821X(92)90014-M.

F. Qi. Generalized abstracted mean values. *J. Inequal. Pure and Appl. Math.*, 1:Art. 4, 2000.

R Development Core Team. *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria, 2006. URL http://www. R-project.org. ISBN 3-900051-07-0.

M. Rehkämper and A. W. Hofmann. Recycled ocean crust and sediment in Indian Ocean MORB. *Earth Planet. Sci. Lett.*, 147:93–106, 1997. doi:10.1016/S0012-821X(97)00009-5.

C. Richardson and D. McKenzie. Radioactive disequilibria from 2D models of melt generation by plumes and ridges. *Earth Planet. Sci. Lett.*, 128:425–437, 1994. doi:10.1016/0012-821X(94)90160-0.

C. N. Richardson, J. R. Lister, and D. McKenzie. Melt conduits in a viscous porous matrix. *J. Geophys. Res.*, 101:20423–20432, 1996. doi:10.1029/96JB01212.

F. Richter and D. McKenzie. Dynamical models for melt separation from a matrix. *J. Geol.*, 92:729–740, 1984.

K. H. Rubin, I. van der Zander, M. C. Smith, and E. C. Bergmanis. Minimum speed limit for ocean ridge magmatism from ^{210}Pb-^{226}Ra-^{230}Th disequilibria. *Nature*, 437:534–538, 2005. doi:10.1038/nature03993.

V. J. M. Salters. The generation of mid-ocean ridge basalts from the Hf and Nd isotope perspective. *Earth Planet. Sci. Lett.*, 141:109–123, 1996. doi:10.1016/0012-821X(96)00070-2.

V. J. M. Salters and S. R. Hart. The mantle sources of ocean ridges, islands and arcs - The Hf-isotope connection. *Earth Planet. Sci. Lett.*, 104:364–380, 1991. doi:10.1016/0012-821X(91)90216-5.

V. J. M. Salters and A. Stracke. Composition of the depleted mantle. *Geochem. Geophys. Geosyst.*, 5, 2004. doi:10.1029/2003GC000597.

V. J. M. Salters and W. M. White. Hf isotope constraints on mantle evolution. *Chem. Geol.*, 145:447–460, 1998. doi:10.1016/S0009-2541(97)00154-X.

J. G. Schilling, B. B. Hanan, B. McCully, R. H. Kingsley, and D. Fontignie. Influence of the Sierra-Leone mantle plume on the equatorial Mid-Atlantic Ridge : A Nd-Sr-Pb isotopic study. *J. Geophys. Res.-Solid.*, 99:12005–12028, 1994. doi:10.1029/94JB00337.

J. G. Schilling, R. Kingsley, D. Fontignie, R. Poreda, and S. Xue. Dispersion of the Jan Mayen and Iceland mantle plumes in the Arctic: A He-Pb-Nd-Sr isotope tracer study of basalts from the Kolbeinsey, Mohns, and Knipovich Ridges. *J. Geophys. Res.-Solid.*, 104:10543–10569, 1999. doi:10.1029/95JB00668.

D. R. Scott and D. J. Stevenson. Magma solitons. *Geophys. Res. Lett.*, 11:1161–1164, 1984.

D. R. Scott and D. J. Stevenson. Magma ascent by porous flow. *J. Geophys. Res.*, 91: 9283–9296, 1986.

D. R. Scott, D. J. Stevenson, and J. A. Whitehead Jr. Observations of solitary waves in a viscously deformable pipe. *Nature*, 319:759–761, 1986. doi:10.1038/319759a0.

D. M. Shaw. Trace element fractionation during anatexis. *Geochim. Cosmochim. Acta*, 34:237–243, 1970. doi:10.1016/0016-7037(70)90009-8.

S. J. Sheather and M. C. Jones. A reliable data-based bandwidth selection method for kernel density estimation. *J. Roy. Statist. Soc. B*, 53:683–690, 1991.

S. Shirey, J. Bender, and C. Langmuir. 3-Component isotopic heterogeneity near the Oceanographer transform, Mid-Atlantic Ridge. *Nature*, 325:217–223, 1987. doi:10.1038/325217a0.

B. W. Silverman. *Density Estimation for Statistics and Data Analysis*, volume 26 of *Monographs on Statistics and Applied Probability*. Chapman and Hall, London, 1986. ISBN 0-412-24620-1.

K. W. W. Sims, S. J. Goldstein, J. Blichert-Toft, M. R. Perfit, P. Kelemen, D. J. Fornari, P. Michael, M. T. Murrell, S. R. Hart, D. J. DePaolo, G. Layne, L. Ball, M. Jull, and J. Bender. Chemical and isotopic constraints on the generation and transport of magma beneath the East Pacific Rise. *Geochim. Cosmochim. Acta*, 66:3481–3504, 2002. doi:10.1016/S0016-7037(02)00909-2.

L. Slater, D. McKenzie, K. Grönvold, and N. Shimizu. Melt generation and movement beneath Theistareykir, NE Iceland. *J. Petrol.*, 42:321–354, 2001. doi:10.1093/petrology/42.2.321.

M. Spiegelman. Flow in deformable porous media. Part 1 Simple analysis. *J. Fluid Mech.*, 247:17–38, 1993. doi:10.1017/S0022112093000369.

M. Spiegelman and T. Elliott. Consequences of melt transport for uranium series disequilibrium in young lavas. *Earth Planet. Sci. Lett.*, 118:1–20, 1993. doi:10.1016/0012-821X(93)90155-3.

M. Tatsumoto. Isotopic composition of lead in oceanic basalt and its implication to mantle evolution. *Earth Planet. Sci. Lett.*, 38:63–87, 1978. doi:10.1016/0012-821X(78)90126-7.

P. E. van Keken, E. H. Hauri, and C. J. Ballentine. Mantle mixing: The generation, preservation, and destruction of chemical heterogeneity. *Annu. Rev. Earth Planet. Sci.*, 30:493–525, 2002. doi:10.1146/annurev.earth.30.091201.141236.

I. Vlastelic, D. Aslanian, L. Dosso, H. Bougault, J. L. Olivet, and L. Geli. Large-scale chemical and thermal division of the Pacific mantle. *Nature*, 399:345–350, 1999. doi:10.1038/20664.

J. I. Wendt, M. Regelous, Y. L. Niu, R. Hekinian, and K. D. Collerson. Geochemistry of lavas from the Garrett Transform Fault: insights into mantle heterogeneity beneath the Eastern Pacific. *Earth Planet. Sci. Lett.*, 173:271–284, 1999. doi:10.1016/S0012-821X(99)00236-8.

W. M. White. Sources of oceanic basalts: Radiogenic isotopic evidence. *Geology*, 13: 115–118, 1985. doi:10.1130/0091-7613(1985)13⟨115:SOOBRI⟩2.0.CO;2.

W. M. White. *Geochemistry: An on-line textbook*. John Hopkins University Press, 2005. URL http://www.geo.cornell.edu/geology/classes/geo455/Chapters.HTML.

W. M. White. ^{238}U/^{204}Pb in MORB and open system evolution of the depleted mantle. *Earth Planet. Sci. Lett.*, 115:211–226, 1993. doi:10.1016/0012-821X(93)90223-V.

W. M. White, A. W. Hofmann, and H. Puchelt. Isotope geochemistry of Pacific mid-ocean ridge basalt. *J. Geophys. Res.-Solid.*, 92:4881–4893, 1987.

J. A. Whitehead and K. R. Helfrich. The Korteweg-deVries equation from laboratory conduit and magma migration equations. *Geophs. Res. Lett.*, 13:545–546, 1986.

R. W. Williams and J. B. Gill. Effects of partial melting on the uranium decay series. *Geochim. Cosmochim. Acta*, 53:1607–1619, 1989. doi:10.1016/0016-7037(89)90242-1.

R. K. Workman and S. R. Hart. Major and trace element composition of the depleted MORB mantle (DMM). *Earth Planet. Sci. Lett.*, 231:53–72, 2005. doi:10.1016/j.epsl.2004.12.005.

S. Xie and P. J. Tackley. Evolution of U-Pb and Sm-Nd systems in numerical models of mantle convection and plate tectonics. *J. Geophys. Res.*, 109:B11204, 2004a. doi:10.1029/2004JB003176.

S. Xie and P. J. Tackley. Evolution of helium and argon isotopes in a convecting mantle. *Phys. Earth Planet. Inter.*, 146:417–439, 2004b. doi:10.1016/j.pepi.2004.04.003.

D. Yu, D. Fontignie, and J. G. Schilling. Mantle plume-ridge interactions in the Central North Atlantic: A Nd isotope study of Mid-Atlantic Ridge basalts from 30°N to 50°N. *Earth Planet Sci. Lett.*, 146:259–272, 1997. doi:10.1016/S0012-821X(96)00221-X.

A. Zindler and S. R. Hart. Chemical geodynamics. *Ann. Rev. Earth Planet. Sci.*, 14: 493–571, 1986. doi:10.1146/annurev.ea.14.050186.002425.

H. Zou. Trace element fractionation during modal and nonmodal dynamic melting and open-system melting: A mathematical treatment. *Geochim. Cosmochim. Acta*, 62:1937–1945, 1998. doi:10.1016/S0016-7037(98)00115-X.