**Title**

The Transport of Cosmic Rays and Their Impact on the Thermal Evolution of Galaxy Halos and Clusters

**Author**

Kempski, Philipp Alexander

**Publication Date**

2022

# The Transport of Cosmic Rays and Their Impact on the Thermal Evolution of Galaxy Halos and Clusters

By

Philipp Alexander Kempski

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Astrophysics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Eliot Quataert, Chair
Professor Eugene Chiang
Professor Stuart Bale

Spring 2022

# The Transport of Cosmic Rays and Their Impact on the Thermal Evolution of Galaxy Halos and Clusters

# Abstract

# The Transport of Cosmic Rays and Their Impact on the Thermal Evolution of Galaxy Halos and Clusters

by

Philipp Alexander Kempski

Doctor of Philosophy in Astrophysics

University of California, Berkeley

Professor Eliot Quataert, Chair

Although negligible by number density, relativistic cosmic rays (CRs) are an energetically important component of the plasmas filling the inter-stellar media and halos of galaxies, and the cores of galaxy clusters. As a result, CRs can have a large impact on their host systems, for example by heating diffuse gas or by driving galactic winds. This "CR feedback" is a strong function of how CRs are transported across galactic environments. Despite elaborate efforts to constrain it, the transport of galactic CRs remains uncertain, which makes CR feedback one of the biggest unsolved puzzles in galaxy evolution. The goal of this thesis has been two-fold. A large portion of this work uses theoretically favoured models of CR transport to study how CRs affect the dilute gases in galactic halos and in the intra-cluster medium of galaxy clusters. It is shown that CRs can significantly affect wave propagation by driving rapidly growing fluid instabilities, which has significant implications for the thermal evolution of these systems. This work is complemented by a study that compares existing micro-physical theories of CR transport with detailed measurements of spectra of galactic CRs in the solar neighbourhood. It reveals that existing CR transport models are generally not in good agreement with the data and provides speculations on how this discrepancy could be remedied.

To my Family,

Your unconditional support and encouragement made this thesis possible.

# Contents

# List of Figures

# Acknowledgments

There are many people that have had a profound impact on my life over the past 5 to 27 years whom I feel very indebted to, and I wish I could express my gratitude to them in person more often. In that sense, writing these acknowledgments is perhaps the most important part of my thesis, as it allows me to think back and reflect on how much my life has gained from the great people in it.

First, I am extremely thankful to my advisor Eliot Quataert for his unwavering support throughout my time in graduate school and for teaching me what good science looks like. I am also very grateful for the freedom to pursue random scientific ideas and for his understanding when these ideas turned out borderline crazy. I could not have asked for a better advisor.

I would also like to thank other mentors that have significantly improved my time in graduate school, in particular Matthew Kunz, Jonathan Squire, Siang Peng Oh and Philip Hopkins, as well as the grad students, postdocs, staff members and professors that I had the pleasure of interacting with around Campbell Hall. I am grateful for the many interesting conversations and pieces of advice that have turned me into a better scientist. I would also like to thank the Kavli Institute for Theoretical Physics (KITP) for supporting me on a Graduate Fellowship in the Fall of 2019. Looking back, this was the time in graduate school that was the most fulfilling academically. I learned a lot and I will be forever grateful for this opportunity.

I am thankful to Ralph Schoenrich and David Weinberg for their invaluable mentorship during my undergraduate times. It was their support that motivated me to pursue a degree in astrophysics and led me to graduate school in Berkeley.

I thank my close friends, especially Sevan Adourian, Nick Choksi, Chris Moeckel and Mihai Munteanu. Nick, thank you for being a great housemate and friend over many years, best tennis partner I could ask for, and for introducing me to proper hummus. Sevan, Chris and Mihai, thank you for your emotional support, thought-provoking conversations, top-shelf content and fun trips over all these years. The time when we all overlapped in Berkeley will always be one of my fondest memories.

I would like to thank all my Physcal teammates: Adrian Bayer, Andy Biekert, Kayla Currier, Nicholas Dale, Francisco Machado, Christopher Madsen, Nikola Maksimovic, Aashrita Mangu, Matthew Quenneville, Fabio Lunes Sanches, Max Smiley, Tomo Soejima, Mauricio Ayllon Unzueta and Lorenzo Vergari. Playing soccer with all of you was one of my grad school highlights, and winning the IM summer season together in 2018 continues to be my proudest achievement in Berkeley.

I am also grateful to my good friend and undergraduate flatmate Christoph Weis, who probably

taught me more physics than anyone else. I would also like to thank Michal Dabrowka, Lukasz Falandysz and Michal Wojsznis for their unwavering friendship over all these years.

I thank my partner Aashrita Mangu for being my biggest supporter throughout grad school. Thank you for being there for me in the toughest moments and for cheering me up whenever I was frustrated with work. I am excited for what the future holds for us. I would also like to express my deep gratitude to Shina Mehta, Srinivas Rao, and Animesh Mangu for regularly welcoming me into their home, especially when I was not able to see my family due to COVID.

Finally, though most importantly, I would like to thank my entire family. I thank my parents Danuta and Andrzej for instilling the value of education in me and for their continuous love and support. This thesis is a product of the last 27 years and would not have been possible without them. Thank you, Mamo & Tato, for always being there for me when I need help or advice. I would also like to thank my grandparents, Marylka, Stenia, Janusz and Walerek, as well as my brother Janek and sister-in-law Asia, for their love and encouragement. It is hard to put into words how much my life has been shaped by my big brother. After all, my pursuing a physics degree after high school really was a product of his timely encouragement and recurring gasps about how beautiful physics is (and no, he is not a physicist). After many years of copying him and trying to impress him as a kid, 18-year-old me had no real choice but to abide, which led me to this spot almost a decade later.

# Chapter 1

# Introduction

Galaxies and galaxy clusters are filled with a population of relativistic particles called cosmic rays (CRs), which contain the most energetic particles in the Universe. Ever since their discovery more than a century ago by Victor Hess, continuous observational and theoretical efforts have led to an ever-increasing understanding of how these energetic particles are created and transported across galaxies. Detailed measurements in the solar neighbourhood have revealed that galactic CRs arrive at Earth nearly isotropically, with a spectrum that is an almost perfect power law over many orders of magnitude in energy, with exponent $\approx -2.7$. The steep slope of this spectrum implies that most of the energy is carried by the mildly relativistic ~GeV CR protons, although the most energetic CRs have energies up to $10^{11}$ GeV.[1] The measured galactic CR number density, of order $10^{-9}$ cm$^{-3}$, is much less than typical inter-stellar number densities of thermal particles ($0.001 - 1$ cm$^{-3}$ depending on inter-stellar medium phase). Quite remarkably, however, because of the large energies carried by individual CR particles, the total energy density in CRs is comparable to the energy density of thermal particles in many astrophysical environments, including the Milky Way. It is because of this astonishing fact that CRs can have a large impact on the evolution of galaxies and galaxy clusters. How these astrophysical systems are affected by their most energetic particles is the overarching topic of this thesis.

## 1.1 Cosmic Ray Transport and Feedback in Galaxies and Clusters

CRs are believed to be accelerated and injected into the inter-stellar media (ISM) of galaxies and intra-cluster media (ICM) of galaxy clusters by supernovae (SNe) and active galactic nuclei (AGN) (see, e.g., Drury 2012; Amato 2014 for recent reviews). One would naively expect that CRs produced by SNe and AGN escape their hosts rather quickly, namely at the speed of light. However, the CR lifetime in the Galaxy (>Myr) inferred from their chemical abundances is many orders of magnitude larger than the Milky Way's light crossing time. The long lifetime is believed

---

[1] Remarkably, the kinetic energy of a $10^{11}$ GeV CR particle is comparable to the energy of a tennis ball moving at almost 100 mph!

to be due to (resonant) scattering of CRs by magnetic-field fluctuations that permeate inter-stellar gas (e.g., review by Zweibel 2013). The scattering results in CRs escaping via a random walk, rather than ballistic propagation at the speed of light, and therefore significantly extends their galactic residence time.

The scattering of CRs turns out to have significant implications for the evolution of galaxies and clusters. In particular, the scattering allows CRs to exchange energy and momentum with the ambient magnetic field and inter-stellar gas. This coupling is important because CRs are energetically significant in many astrophysical systems. As a result, CRs can have a big impact on large-scale galactic dynamics and thermodynamics, for example by driving galactic winds and/or heating diffuse gas.

The waves responsible for coupling the energetically important GeV CRs have very short wavelengths, comparable to their gyro-radius, which is of order 1 astronomical unit. This is many orders of magnitude smaller than the kiloparsec scales on which CRs affect their host galaxies. The impact of CRs on galaxy evolution on large scales (e.g., mass outflow rates in galactic winds) is very sensitive to how CRs are scattered by the microscopic gyro-scale fluctuations. How CRs affect galaxies, usually termed "CR feedback", is therefore a remarkable example of cross-talk between micro- and macro-scales, and a challenging multi-scale astrophysics problem.

A good understanding of the waves that scatter CRs is therefore essential for studying CR feedback in galaxies. However, the origin and nature of the waves remain uncertain. As a result, CR feedback is one of the major unsolved puzzles in galaxy evolution. In particular, the two leading theories of CR scattering give rise to two very different types of transport and forms of CR feedback. In "self-confinement" theory, the bulk of CRs are scattered by magneto-hydrodynamic (MHD) Alfvén waves that they themselves excite through the streaming instability (Kulsrud & Pearce 1969). In this picture, on galactic scales the bulk of CRs are well described as a fluid that is streaming at the Alfvén speed relative to the thermal gas (Skilling 1971). The leading alternative to self-confinement theory is "extrinsic turbulence" theory, in which CRs are instead scattered by a pre-existing turbulent cascade of magnetic fluctuations. In this picture, CRs diffuse rather than stream. A variety of simulations have shown that these distinct forms of CR transport yield dramatically different outcomes in galaxy evolution models (e.g., Ruszkowski et al. 2017; Wiener et al. 2017; Farber et al. 2018; Quataert et al. 2022a; Quataert et al. 2022b). In particular, the heating of inter-stellar gas and the launching of galactic winds depend sensitively on the type of CR transport.

## 1.2 Thesis Summary and Guide to Chapters

My journey into CR astrophysics began by studying how streaming CRs affect the propagation and (in)stability of waves in the dilute plasmas that fill the halos of galaxies and the ICM of galaxy clusters. The evolution of these fluctuations directly affects the phase structure of the gas and the transport of energy from one place to another. This has significant implications for the thermal evolution of these systems. The results of this work are presented in Chapters 2, 3 and 4.

By exciting short-wavelength (gyro-resonant) Alfvén waves, streaming CRs heat the ambient thermal gas (Wentzel 1971). In Chapter 2, I study how this heating affects the thermal (in)stability of gas in galactic halos. I first demonstrate that CR heating is plausibly important in galactic halos and may play a role counteracting cooling to keep the hot thermal plasma in these systems in approximate thermal equilibrium. I then proceed by studying how the heating affects local thermal instability, i.e. the condensation of cold gas out of the hot phase. Thermal instability is often invoked to explain the observed cool gas in circum-galactic media (CGM) and in the cores of galaxy clusters (e.g., Nulsen 1986; McCourt et al. 2012; Voit et al. 2015). The presence of this cold gas is important because it provides the fuel for galactic star formation (Tumlinson et al. 2017). Interestingly, I show that even if CRs and CR heating are energetically important, they do not significantly affect the excitation of thermal instability. Heating by streaming CRs turns purely exponential growth into oscillatory growth (overstability), but without significantly affecting the growth rate itself. The surprisingly small impact of CR heating on thermal instability was later confirmed numerically by Butsky et al. (2020).

In Chapters 3 and 4, I study how the presence of streaming CRs affects wave propagation in the dilute ICM of galaxy clusters. The evolution of waves in the ICM is a long-standing problem, with potentially significant implications for its thermal balance. In particular, observations show that despite large cooling losses, the ICM stays in approximate thermal equilibrium. This suggests that there is a source of heating present, which approximately offsets the cooling (Peterson & Fabian 2006). The energy that prevents runaway cooling likely comes from the central AGN and the gigantic radio bubbles inflated by its jet (e.g., Churazov et al. 2000; Bîrzan et al. 2004; Werner et al. 2019). However, how this energy is transported and thermalized throughout the ICM remains an open question. One possibility is that the non-thermally supported radio bubbles inject CRs into the ICM, which heat the gas via streaming (Guo & Oh 2008; Jacob & Pfrommer 2017a; Jacob & Pfrommer 2017b). Alternatively, the buoyantly rising radio bubbles may excite turbulence (e.g., Zhuravleva et al. 2016), and/or weak shocks/sound waves (e.g., Fabian et al. 2003; Sternberg & Soker 2009) which transport the energy from the AGN throughout the ICM. The turbulence/wave models are supported by the observations that cluster cores are turbulent (e.g., Li et al. 2020) and also show signs of powerful sound waves/weak shocks propagating from the AGN out to large distances (Fabian et al. 2003; Fabian et al. 2006). However, sound waves are expected to be strongly damped in the dilute ICM plasma, and so their observed long-term survival has been a theoretical puzzle (e.g., Zweibel et al. 2018). Whether the radio bubbles can indeed efficiently excite turbulence also remains unclear (e.g., Reynolds et al. 2015). In Chapters 3 and 4, I show that the evolution of ICM waves is very sensitive to CR physics, even if the CR energy density is small compared to the thermal energy density. In Chapter 3, I show that in the low-collisionality ICM plasma streaming CRs couple to the thermal gas in a novel way, which drives a rapidly-growing acoustic instability. In Chapter 4, I further show that CR streaming in the ICM also drives a new type of buoyancy instability. The instability is characterised by growth rates of order the local free-fall time for plausible ICM parameters. Both instabilities do not require significant CR pressures when magnetic fields are weak, as is expected in the ICM. Excitation of these instabilities is especially likely in the vicinity of the radio bubbles, where the CR pressure is locally enhanced. My ongoing numerical work on the saturation of the acoustic instability suggests that it evolves into a series of weak shocks. This is broadly consistent with the acoustic disturbances observed in the core of the Perseus cluster (Fabian et al. 2003; Fabian et al. 2006).

The results from Chapters 2, 3 and 4 rely on the assumption that the bulk of CRs stream at the Alfvén speed relative to the thermal gas. In other words, it is assumed that the energetically important GeV CRs are predominantly scattered by self-excited waves and not an ambient turbulent cascade. There is, in fact, a very reasonable theoretical motivation for this choice. First, existing theoretical models of MHD turbulence predict that on small scales the turbulent cascade does not have the properties necessary for efficient scattering of GeV CRs, especially in dilute plasmas like the ICM. At the same time, GeV CRs are the most likely to self-excite Alfvén waves through the streaming instability. This type of argument, which supports the choice of CR transport in Chapters 2, 3 and 4, is arguably quite convincing on a theoretical level. However, it will perhaps not come as a surprise to the reader that issues inevitably arise outside of the theoretical realm. In particular, there are detailed observations of spectra of galactic CRs measured in the solar neighbourhood, which put tight constraints on CR transport in the Galaxy (e.g., Aguilar et al. 2015; Aguilar et al. 2016; Cummings et al. 2016). These turn out to be quite challenging for existing theories of CR transport, a problem I became aware of halfway into graduate school and which quickly became my scientific obsession. My attempts to reconcile CR transport theory with observations resulted in Chapter 5 of this thesis.

To a good approximation, local CR data is consistent with CR scattering by magnetic-field fluctuations with spectrum $\propto k^{-\delta}$ and $\delta \sim [1.4, 1.7]$ (e.g., Amato & Blasi 2018 for a recent review). This corresponds to a CR lifetime in the Galaxy that is energy dependent with rough scaling $\tau \propto E^{-(2-\delta)}$. Phenomenological models of CR transport can successfully reproduce the main trends in local CR data using a combination of isotropic Kolmogorov-like turbulence and waves excited by the CR streaming instability (e.g., Blasi et al. 2012; Aloisio et al. 2015). However, these models are not well motivated theoretically, because MHD turbulence is known to be very different from isotropic Kolmogorov-like turbulence. In particular, incompressible (Alfvénic) MHD turbulence is very anisotropic on small scales (Goldreich & Sridhar 1995), which makes it very inefficient at scattering CRs (Chandran 2000). For this reason, Yan & Lazarian (2004) proposed that CRs are instead scattered by the isotropic turbulent cascade of MHD fast modes. In Chapter 5, I test whether CR observables can be explained using a combination of waves excited by the streaming instability and the weak cascade of fast modes. The key difference relative to models based on isotropic Kolmogorov turbulence is that fast modes are strongly damped on small scales. This makes their effect on CR transport very different from what previous work concluded. While some of the main trends in CR observables can still be reproduced using this combination of scattering mechanisms, I show that it requires a significant amount of fine-tuning of ISM plasma conditions. This is primarily due to the fact that MHD fast modes are strongly damped. I further argue that the weak cascade of fast modes is in fact not well motivated theoretically. This is because the weak cascade is suppressed by wave steepening and weak-shock dissipation even in subsonic turbulence. This raises the significant possibility that fast modes are not important for CR scattering. The significant challenges faced by current micro-physical theories of CR transport suggest that there may be a big gap in our understanding of MHD turbulence and how it scatters CRs. Roughly Alfvénic streaming is still the favored model for low-energy CRs, as assumed in Chapters 2–4. In particular, the breaks in CR spectra around a few hundred GeV suggest that the physics of CR transport changes around that energy (Blasi et al. 2012). A transition from external fluctuations dominating the transport to some degree of self-confinement at lower energies remains the most likely explanation for the origin of the spectral breaks. The main unknown is the source of external scattering waves that could explain the empirically derived transport of higher-energy CRs.

Given how much time and energy was spent on this problem over the last two years, having a single-chapter summary of my efforts does not feel completely adequate. However, although my PhD is coming to an end, my journey with this theoretical mystery continues.

## 1.3 Other Work

In the first two years of graduate school I completed a first-author publication that is unrelated to this thesis. However, it used similar concepts in low-collisionality plasma physics as Chapters 3 and 4 of this dissertation. In Kempski et al. (2019), we studied turbulence in black-hole accretion flows, with a focus on systems in which the mean free paths of the thermal particles are large. This includes Radiatively Inefficient Accretion Flows (RIAFs; Mahadevan & Quataert 1997) onto Supermassive Black Holes (SMBHs), which encompasses the SMBHs targeted by the Event Horizon Telescope (EHT). Simulations with low-collisionality plasma physics are needed to understand the accretion flows in such systems and the generated radiation observed by EHT. However, fully kinetic simulations remain prohibitively expensive and fully collisional MHD simulations are routinely used instead. In Kempski et al. (2019) we attempted to bridge the gap between fully kinetic and fluid simulations by studying the computationally simpler weakly-collisional plasma regime (often referred to as Braginskii MHD; Braginskii 1965) characterized by a large anisotropic viscosity. This weakly-collisional fluid model is the simplest, well-motivated model that captures essential features of kinetic plasma physics on large scales. We employed the pseudo-spectral code Snoopy (Lesur & Longaretti 2007) to simulate the turbulence driven by the magneto-rotational instability (MRI; Balbus & Hawley 1991) in a local patch of the accretion disk using the shearing-box method. We found that the large anisotropic viscous forces significantly change the structure of MRI-driven turbulence and are an important dissipation channel. In particular, we found that anisotropic viscous heating is a major source of plasma heating. At the same time, notably, we found that there are no significant differences in the overall level of angular-momentum transport between our weakly-collisional (Braginskii MHD) and fully-collisional (MHD) simulations.

# Chapter 2

# Thermal Instability of Halo Gas Heated by Streaming Cosmic Rays

An earlier version of this article was previously published as Kempski P., Quataert E., 2020, MNRAS, 493, 1801. Figure 2.2 has been reorganised to accommodate the format of this thesis.

## 2.1 Abstract

Heating of virialized gas by streaming cosmic rays (CRs) may be energetically important in galaxy halos, groups and clusters. We present a linear thermal stability analysis of plasmas heated by streaming CRs. We separately treat equilibria with and without background gradients, and with and without gravity. We include both CR streaming and diffusion along the magnetic-field direction. Thermal stability depends strongly on the ratio of CR pressure to gas pressure, which determines whether modes are isobaric or isochoric. Modes with $\boldsymbol{k} \cdot \boldsymbol{B} \neq 0$ are strongly affected by CR diffusion. When the streaming time is shorter than the CR diffusion time, thermally unstable modes (with $\boldsymbol{k} \cdot \boldsymbol{B} \neq 0$) are waves propagating at a speed $\propto$ the Alfvén speed. Halo gas in photoionization equilibrium is thermally stable independent of CR pressure, while gas in collisional ionization equilbrium is unstable for physically realistic parameters. In gravitationally stratified plasmas, the oscillation frequency of thermally overstable modes can be higher in the presence of CR streaming than the buoyancy/free-fall frequency. This may modify the critical $t_{\rm cool}/t_{\rm ff}$ at which multiphase gas is present. The criterion for convective instability of a stratified, CR-heated medium can be written in the familiar Schwarzschild form $ds_{\rm eff}/dz < 0$, where $s_{\rm eff}$ is an effective entropy involving the gas and CR pressures. We discuss the implications of our results for the thermal evolution and multiphase structure of galaxy halos, groups and clusters.

## 2.2 Introduction

The short radiative cooling times of virialized gas in galaxies and clusters suggest that these systems should contain significantly more cool gas at their centers than is observed (Peterson

& Fabian 2006). This implies that the hot gas surrounding galaxy halos is also heated, which is thought to come from feedback by star formation and central active galactic nuclei (AGN; e.g., Guo et al. 2008). An appreciable fraction of the energy released by AGNs and supernova explosions comes in the form of relativistic cosmic-ray particles (McNamara & Nulsen 2007; Ackermann et al. 2013), which may be important for the dynamics and gas heating in galaxies, halos and clusters (e.g., Breitschwerdt et al. 1991; Loewenstein et al. 1991; Everett et al. 2008; Socrates et al. 2008; Guo & Oh 2008; Zweibel 2013; Ruszkowski et al. 2017; Zweibel 2017; Ehlert et al. 2018).

Cosmic rays are confined in galaxies for times much longer than would be expected from their propagation speed ($\approx$ speed of light), due to scattering off small-scale electromagnetic fluctuations. These fluctuations can be either due to external turbulence, or Alfvén waves generated by the cosmic rays themselves. In the self-excitation scenario, wave growth is driven by the cosmic ray streaming instability (Kulsrud & Pearce 1969): as cosmic rays collectively drift down their pressure gradient, the free energy associated with their velocity anisotropy can excite Alfvén waves. Pitch-angle scattering isotropizes the cosmic rays in the frame comoving with the waves, which, in the absence of strong wave damping, limits the CR drift speed to the local Alfvén speed (by contrast, the cosmic-ray drift speed can be significantly larger than the Alfvén speed in the strong-damping limit; Skilling 1971; Wiener et al. 2013). In a steady state, the streaming-induced wave growth is balanced by wave damping, so that the energy of the cosmic rays is essentially being transferred to the thermal plasma. This couples the background plasma to the cosmic rays, which heat the gas at a rate $-\boldsymbol{v_A} \cdot \boldsymbol{\nabla} p_c$, where $v_A$ is the local Alfvén speed and $p_c$ is the CR pressure (Wentzel 1971).

Guo & Oh (2008), Jacob & Pfrommer (2017a) and Jacob & Pfrommer (2017b) showed that this cosmic-ray heating can suppress the cooling catastrophe in clusters for CR pressures that are consistent with observational bounds. Indeed, they found that the required CR pressure (gradient) is small compared to the gas pressure (gradient), as is also found observationally (e.g., Huber et al. 2013). Whether the same is true in galaxy halos is still unclear (e.g., Hopkins et al. 2020b).

While heating suppresses cooling globally (i.e. on sufficiently long time and length scales) and maintains the hot virialized gas in massive halos in approximate hydrostatic and thermal balance,[1] there is strong observational evidence for cold gas in the halos of galaxies. Cool gas is present both in the circumgalactic medium (CGM) of massive and Milky-Way-like galaxies, and in the intracluster medium (ICM). In the ICM, detailed spatially-resolved observations (that use both atomic and molecular transitions, e.g., Salomé et al. 2006, Cavagnolo et al. 2009) indicate the presence of cold-gas filaments embedded within the otherwise hot, virialized gas, which constitutes most of the intracluster gas mass. In the CGM, the cool-gas morphology is less certain (i.e., it could be filamentary or volume-filling), and the cold gas mass may comprise a significant fraction of the total halo gas mass. Indeed, observations of the CGM using Ly $\alpha$ emission and quasar absorption lines (Werk et al. 2013; Stocke et al. 2013; Cantalupo et al. 2014; Hennawi et al. 2015; Bowen et al. 2016; Cai et al. 2017) suggest the presence of multiphase gas along most lines of sight (suggesting that the cold phase may permeate the CGM, instead of forming a filamentary structure).

The origin of the cold gas remains uncertain. It could be gas elevated into the halo by galac-

---

[1] While the hot virialized gas in clusters has a sufficiently high temperature to be seen directly in emission, the emission from virialized gas in the CGM is too faint for current telescopes. Nevertheless, hot virialized gas is expected to be present in halos of mass $\gtrsim 10^{11.5} M_\odot$ (Birnboim & Dekel 2003; Dekel et al. 2009).

tic winds. However, how the cold gas in high-velocity galactic winds is produced and entrained remains uncertain (Scannapieco & Brüggen 2015; Thompson et al. 2016; Zhang et al. 2017). Alternatively (or, in addition), the cold phase may be produced in situ via thermal instability. Thermal instability is commonly linked to the existence of multiphase gas in the interstellar medium (Field 1965) and has been studied in the context of galaxy halos and clusters using a number of simulations and models (Nulsen 1986; Binney et al. 2009; McCourt et al. 2012; Sharma et al. 2012; Voit et al. 2015; Meece et al. 2015; Voit et al. 2017; Voit 2018). These simulations suggest that the condensation of cold gas via thermal instability can occur if the ratio of the cooling time to the free-fall time is sufficiently small. Typically they find that $t_{\mathrm{cool}}/t_{\mathrm{ff}} \lesssim 10$, however, this value may depend on the size of the initial perturbations (Pizzolato & Soker 2005; Singh & Sharma 2015; Choudhury et al. 2019) and whether magnetic fields are included (Ji et al. 2018). The connection between $t_{\mathrm{cool}}/t_{\mathrm{ff}} \lesssim 10$ and the existence of multiphase structure has been partly born out by the cluster observations of McDonald et al. (2010), but in a more recent sample of 56 clusters observed by the *Chandra X-ray Observatory*, cold gas is present even when $t_{\mathrm{cool}}/t_{\mathrm{ff}} \gtrsim 10$ (Hogan et al. 2017).

The purpose of this paper is to understand the thermal stability of systems heated by streaming cosmic rays, which may be an important heating mechanism in galaxy halos. We first present order-of-magnitude estimates showing that heating due to streaming CRs may be important for a wide range of halo masses. We then perform a linear stability analysis, in which we take into account both CR streaming and diffusion, and we look at equilibria with and without gravity. While we find that explicitly including gravity is not very important for thermal instability growth rates, it can transform thermal instability into a convective instability driven by buoyancy.[2]

The thermal stability of systems with heating by streaming CRs was first considered heuristically in the context of a cooling flow by Loewenstein et al. (1991). Pfrommer (2013) and Wiener et al. (2013) then studied thermal instability with CR heating by assuming that the CR pressure ($p_c$) and gas density ($\rho$) follow the adiabatic relation $p_c \propto \rho^{\gamma_c}$. In this work, we instead explicitly include the evolution equation for the CR pressure, which is in general not consistent with adiabaticity. Cosmic rays are adiabatic only for modes propagating perpendicular to the magnetic field (see Section 2.5.3), but even then we show that correctly perturbing the CR heating produces results that are different from the heuristic calculation in Pfrommer (2013). We also extend previous work by studying the impact of CR diffusion on thermal instability, and we study the instability in different background equilibria.

The remainder of this paper is organised as follows. We introduce the gas–CR equations in Section 2.3. In Section 2.4 we argue that cosmic-ray heating may be important in galactic halos. The linear thermal stability of CR heating is derived in Sections 2.5, 2.6 and 2.7. We solve the perturbed linearised equations in a uniform medium without gravity in Section 2.5. We introduce gas and CR background gradients in Section 2.6 and consider gravitationally stratified equilibria in Section 2.7. In the latter case, we also obtain a criterion for convective instability. We summarize our results and discuss their implications for the multiphase structure of galaxy halos in Section 2.8.

---

[2] This is a rather unsurprising side result of our analysis, because thermal and convective stability are closely linked (Balbus 1995).

We derive estimates for the (global) ratio of CR to thermal pressure in galaxy halos in Appendix 2.A. A heuristic description of the impact of CR diffusion on thermal instability is provided in Appendix 2.B. We show the linearised perturbed equations of a CR-heated background in Appendix 2.C. In Appendix 2.D, we explain why a 1-dimensional calculation (see Section 2.6.2) gives the correct eigenfrequency of the gas entropy mode in a CR-heated background. Finally, we derive an approximate growth rate for the convective instability in a gravitationally stratified medium in Appendix 2.E.

## 2.3 Equations and Timescales

### 2.3.1 Gas–CR Equations

We consider a thermal plasma interacting with a population of relativistic cosmic rays. We model the system by including CR heating and the CR pressure force in the equations of ideal MHD. This results in the following coupled differential equations,

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \boldsymbol{v}) = 0, \tag{2.1}$$

$$\rho \frac{d\boldsymbol{v}}{dt} = -\boldsymbol{\nabla}\left(p_g + p_c + \frac{B^2}{8\pi}\right) + \frac{\boldsymbol{B} \cdot \boldsymbol{\nabla} \boldsymbol{B}}{4\pi} + \rho \boldsymbol{g}, \tag{2.2}$$

$$\frac{\partial \boldsymbol{B}}{\partial t} = \boldsymbol{\nabla} \times (\boldsymbol{v} \times \boldsymbol{B}), \tag{2.3}$$

$$\rho T \frac{ds}{dt} = \mathcal{H} - \boldsymbol{v_A} \cdot \boldsymbol{\nabla} p_c - \rho^2 \Lambda(T), \tag{2.4}$$

$$\frac{dp_c}{dt} = -\frac{4}{3} p_c \boldsymbol{\nabla} \cdot (\boldsymbol{v} + \boldsymbol{v_A}) - \boldsymbol{v_A} \cdot \boldsymbol{\nabla} p_c + \boldsymbol{\nabla} \cdot \left(\kappa \hat{\boldsymbol{b}} \hat{\boldsymbol{b}} \cdot \boldsymbol{\nabla} p_c\right) \tag{2.5}$$

where $\boldsymbol{v}$ is the gas velocity, $\rho$ is the gas density, $p_g$ and $p_c$ are the gas and CR pressures respectively, $\boldsymbol{B}$ is the magnetic field (with unit vector along $\hat{\boldsymbol{b}}$), $\boldsymbol{g}$ is the acceleration due to gravity, $s = k_{\rm B} \ln(p/\rho^\gamma)/(\gamma - 1)m_{\rm H}$ is the gas entropy per unit mass, $\Lambda(T)$ is the temperature-dependent cooling function, and $\mathcal{H}$ is an unspecified heating rate (which we set to 0 except in Section 2.5). $d/dt \equiv \partial/\partial t + \boldsymbol{v} \cdot \boldsymbol{\nabla}$ denotes a total (Lagrangian) time derivative. We assume that cosmic rays stream down their pressure gradient at the Alfvén velocity $\boldsymbol{v_A} = \boldsymbol{B}/\sqrt{4\pi\rho}$, and we also include CR diffusion along the magnetic field, for which we assume a constant diffusion coefficient $\kappa$. We note that formally CRs stream with velocity $\boldsymbol{v}_{\rm st} = -{\rm sgn}(\hat{\boldsymbol{b}} \cdot \boldsymbol{\nabla} p_c)\boldsymbol{v_A}$. This ensures that cosmic rays stream along the magnetic field down their pressure gradient and makes the CR heating term $-\boldsymbol{v}_{\rm st} \cdot \boldsymbol{\nabla} p_c$ positive definite. In our linear stability analysis cosmic rays stream at $\boldsymbol{v_A}$, as we consider background equilibria which satisfy $-\boldsymbol{v_A} \cdot \boldsymbol{\nabla} p_c > 0$ (see footnote 3 in Section 2.5.1 for how this is achieved in a uniform background).

### 2.3.2 CR Transport: Streaming versus Diffusion

The interplay between cosmic-ray streaming and diffusion calls for some further discussion. In the self-confinement picture, the importance of streaming versus diffusion is intimately tied to the saturation of the streaming instability (Kulsrud & Cesarsky 1971; Skilling 1971; Wiener et al. 2013). In the limit of weak damping, the excited Alfvén waves can grow to large amplitudes (compared to when significant damping is present, see next paragraph) until the resultant rapid pitch-angle scattering isotropizes the CRs in the frame of the waves. In this scenario, the cosmic rays are advected down their pressure gradient at the Alfvén speed relative to the thermal plasma, with no diffusive contribution (we neglect diffusion due to external turbulence). This is tantamount to setting $\kappa = 0$ in the above equations. Note that the term $-\boldsymbol{v_A} \cdot \boldsymbol{\nabla} p_c$ in equation 2.4 is then positive definite, because in the limit of self-excited Alfvén waves only (no background turbulence) energy flows from the CRs to the gas (mediated by Alfvén waves), but not vice-versa.

In the opposite limit of significant wave damping, the waves generated by the streaming instability saturate at lower amplitudes. As a result, the CR pitch-angle scattering rate is reduced, and the cosmic-ray momenta do not become fully isotropic in the Alfvén-comoving frame. In this case, the cosmic-ray bulk motion deviates from pure streaming at $\boldsymbol{v_A}$ and $\kappa$ will generally be nonzero. The diffusion coefficient will depend on how the waves are damped. Quite notably, for many of the known damping mechanisms (e.g. turbulent, ambipolar and linear Landau damping), the diffusion term ends up not being diffusive at all (Skilling 1971; Wiener et al. 2013; Wiener et al. 2018). Instead, it has the form of an advective flux (streaming) and the cosmic rays essentially stream down their pressure gradient at super-Alfvénic speeds. This, however, is not always true (e.g., when non-linear Landau damping is dominant and/or if there are external sources of cosmic-ray scattering distinct from self-excited Alfvén waves). For this reason, we keep the diffusion term in our equations (with constant $\kappa$ for simplicity). We do not consider super-Alfvénic streaming in this work, as the dependence of super-Alfvénic streaming velocities on other fluid quantities is uncertain.

### 2.3.3 Dimensionless Parameters and Characteristic Frequencies

We define the ratio of CR pressure to gas pressure,

$$\eta \equiv \frac{p_c}{p_g}, \tag{2.6}$$

and the ratio of thermal to magnetic pressure,

$$\beta \equiv \frac{8\pi p_g}{B^2}. \tag{2.7}$$

We also write the logarithmic slope of the cooling function as

$$\Lambda_T \equiv \frac{\partial \ln \Lambda}{\partial \ln T}. \tag{2.8}$$

There are a number of timescales that characterise the problem. We define the cooling frequency,

$$\omega_c \equiv \frac{\rho^2 \Lambda}{p_g}; \tag{2.9}$$

the wavenumber ($k$) dependent sound frequency (with $c_s$ being the adiabatic sound speed),

$$\omega_s \equiv k c_s; \tag{2.10}$$

the Alfvén *and* CR-heating frequency,

$$\omega_a \equiv \boldsymbol{k} \cdot \boldsymbol{v_A}; \tag{2.11}$$

the cosmic-ray diffusion frequency,

$$\omega_d \equiv \kappa \, (\hat{\boldsymbol{b}} \cdot \boldsymbol{k})^2; \tag{2.12}$$

and the free-fall frequency,

$$\omega_{\rm ff} \equiv \frac{g}{c_s}. \tag{2.13}$$

We stress that $\omega_a$ characterizes both the perturbed magnetic tension (its usual meaning) and the perturbed CR heating $-\boldsymbol{v_A} \cdot \boldsymbol{\nabla} p_c$. Throughout our linear stability calculation in Sections 2.5, 2.6 and 2.7, we focus on local perturbations ($kH \gg 1$, $H$ being a characteristic background length scale), which for our application considered in Section 2.4 corresponds to

$$\omega_s \gg \omega_c, \omega_{\rm ff}, \tag{2.14}$$

and

$$\omega_a \gg \omega_c \tag{2.15}$$

(unless $\mathbf{k} \cdot \mathbf{B} = 0$, in which case eq. 2.15 need not be satisfied). In the CR-heated background, $kH \gtrsim 1$ corresponds to $\omega_a \gtrsim \omega_c \eta^{-1}$ (Section 2.6.1). We find that thermal instability growth rates do not depend significantly on wavenumber $k$, provided that $\omega_a > \omega_c$ ($\omega_a \gtrsim \omega_c \eta^{-1}$) in the uniform (CR-heated) background.

As our fiducial set of parameters, we choose $\omega_a = 10^3 \omega_c$ (which corresponds to fairly high $k$, but such high $k$ is necessary for the CR-heated background if we want to consider $\eta > 0.01$), $\beta = 100$ and, when we include gravity in Section 2.7, $\omega_{\rm ff} = 20\omega_c$. $\omega_{\rm ff} \gtrsim 10\omega_c$ is motivated by observations of hot gas in groups and clusters (McDonald et al. 2010; Hogan et al. 2017), which largely satisfy this constraint. We stress that this choice of $\omega_{\rm ff} \gtrsim 10\omega_c$ is motivated by halo gas specifically, but need not be true in other applications. We show how smaller $\beta$ and $\omega_a$ affect our results in Figure 2.2.

## 2.4 Heating by Cosmic Rays in Galaxy Halos

Before we look at the thermal stability of CR heating, we check under what conditions thermal balance between CR heating and radiative cooling,

$$-\boldsymbol{v_A} \cdot \boldsymbol{\nabla} p_c = \rho^2 \Lambda(T), \tag{2.16}$$

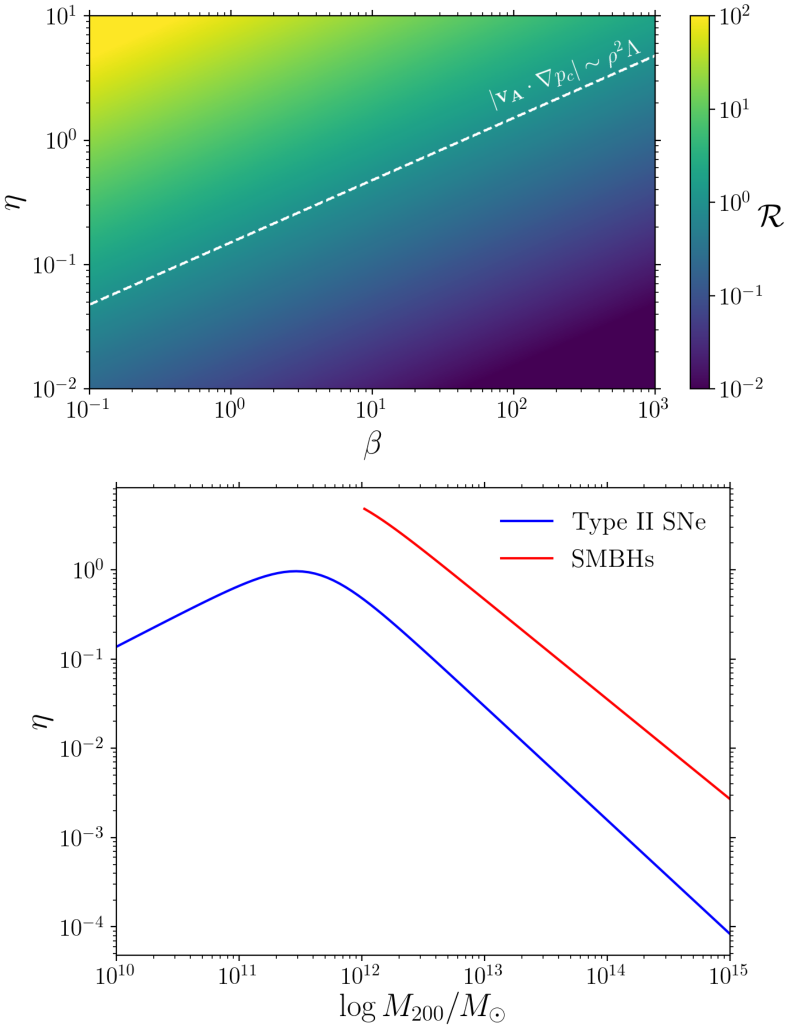


*Figure 2.1*: Top: CR heating versus cooling as a function of $\eta$ (eq. 2.6) and $\beta$ (eq. 2.7). $\mathcal{R}$ is the ratio of CR heating to radiative cooling (eq. 2.18; here we use $\omega_{\rm ff} = 20\omega_c$ and $H_c = 3H$), and increases with increasing CR pressure fraction $\eta$ and with decreasing $\beta$. The white dashed line indicates the approximate region where cosmic-ray heating is comparable to cooling ($\mathcal{R} \sim 1$). Bottom: Order-of-magnitude estimate of the CR pressure fraction as a function of halo mass (see Appendix 2.A). We separately consider cosmic rays injected into the halo by Type II SNe and central SMBHs. We include this plot to motivate that significant cosmic ray pressures are plausible for a wide range of halo masses, especially in $M \approx 10^{12} M_\odot$ halos (large CR pressures, $\eta \gtrsim 1$, have also been found in cosmological simulations, see e.g. Hopkins et al. 2020b). Together, the two panels suggest that there may be significant cosmic-ray heating in galaxy halos for a wide range of halo masses.

may occur in galaxy halos. Our estimates presented in this section suggest that heating by cosmic rays can be important for a broad range of values of $\eta$ and $\beta$ (see Figure 2.1).

We quantify the importance of cosmic-ray heating by defining

$$\mathcal{R} \equiv \frac{|\boldsymbol{v_A} \cdot \boldsymbol{\nabla} p_c|}{\rho^2 \Lambda} \sim \frac{v_A p_c / H_c}{\omega_c p_c \eta^{-1}}, \tag{2.17}$$

where we used our definition of the cooling frequency (2.9). $H_c$ is the CR pressure scale height. Using definition 2.13 and $\beta \approx (c_s/v_A)^2$, we obtain

$$\mathcal{R} \sim \frac{\eta}{\sqrt{\beta}} \frac{H}{H_c} \frac{\omega_{\rm ff}}{\omega_c}, \tag{2.18}$$

where $H = c_s^2/g$. When gas pressure dominates, $H$ is approximately equal to the gas pressure scale height $H_g$.

Our estimate for $\mathcal{R}$ as a function of $\eta$ and $\beta$ is plotted in the top panel of Figure 2.1 for $\omega_{\rm ff}/\omega_c = 20$. We choose a fairly extended CR profile, with $H_c/H = 3$. The dashed white line indicates the approximate region where heating by cosmic rays is comparable to cooling, i.e. $\mathcal{R} \sim 1$.

The top panel of Figure 2.1 suggests that heating by cosmic rays may be important for a wide range of $\eta$ and $\beta$. There is some evidence, from both observations and theory/simulations, suggesting that galaxy halos may often reside above/around the white dashed line (where CR heating is important). While significant CR pressures were measured in the Milky Way (Boulares & Cox 1990) and nearby starburst galaxies (Paglione & Abrahams 2012), observations of cosmic rays and magnetic fields in galaxy halos are challenging and sparse. Nevertheless, there are some constraints that come from synchrotron emission and Faraday rotation measurements along quasar sightlines. Synchrotron emission measurements suggest that cosmic rays and magnetic fields have significantly larger scale heights than the thermal gas (Beck 2015). There is also evidence for strong magnetic fields ($1 - 10\mu$G) that may extend far out (tens of kpc) into the halo (Mora & Krause 2013; Bernet et al. 2013). As a result, it is plausible that there are regions in the halo where $\eta$ is large (e.g., $\sim 1$) and/or $\beta$ is relatively small (e.g., $\lesssim 10$). Under such conditions, equation 2.18 and the top panel of Figure 2.1 suggest that there may be significant CR heating.

Recent cosmological zoom-in simulations with cosmic rays strengthen the claim that CR pressure can be important (even dominant) in galaxy halos (Hopkins et al. 2020b). This is broadly consistent with our estimate for the CR pressure fraction $\eta$, which we show as a function of halo mass in the bottom panel of Figure 2.1 (the calculation can be found in Appendix 2.A). We separately consider the injection of cosmic rays by Type II Supernovae and central Supermassive Black Holes (SMBHs), and we estimate the total energy of cosmic rays out to the virial radius. We assume the (broken power law) stellar mass – halo mass relation from Moster et al. (2013) and the SMBH mass – total stellar mass relation (for ellipticals) from Reines & Volonteri (2015). Comparing the CR energy to the total thermal energy within the virial radius yields the lower panel of Figure 2.1. We find that CR pressure should be significant for a broad range of halo masses and most important in halos of mass $\approx 10^{12} M_\odot$, consistent with Hopkins et al. (2020b).

## 2.5 Cosmic-Ray Thermal Instability in a Uniform Medium

Before we analyse equilibria in which CR heating balances cooling (due to a finite background CR pressure gradient), we look at the simpler case of a uniform background. This setup is particularly relevant for cases where CR heating is not the dominant heating process, but can nevertheless affect the evolution of entropy perturbations (photoionization equilibrium is one such example). The uniform-medium calculation does not capture (slight) modifications to the thermal instability that come from a background CR pressure gradient, but in many ways it produces results that are very similar to the non-uniform medium calculation. For example, the thermal instability growth rates have an almost identical dependence on $\eta$ and CR diffusion. As a result, many of the conclusions drawn here will still be valid in the calculation with background CR heating.

We perform a linear stability calculation of the equations described in Section 2.3. All perturbed quantities are assumed to vary as $\delta Q(\boldsymbol{r}, t) \propto \exp\left[i\boldsymbol{k} \cdot \boldsymbol{r} - i\omega t\right]$. Throughout this (and the next) section, we also ignore gravity, i.e. we set $\boldsymbol{g} = 0$ (we include gravity in Section 2.7).

### 2.5.1 Equilibrium

We consider an equilibrium with

$$\mathcal{H} = \rho^2 \Lambda(T), \tag{2.19}$$

where $\mathcal{H}$ is an unspecified heating rate, which is set to balance cooling (i.e. $\mathcal{H} \gg -\boldsymbol{v_A} \cdot \boldsymbol{\nabla} p_c$). Equilibrium CR heating is considered negligible, and the CR heating term only enters in the perturbed equations. All background fluid variables are assumed to be spatially constant. Without loss of generality, we consider a vertical magnetic field, $\boldsymbol{B} = B\hat{\boldsymbol{z}}$. This equilibrium has the advantage that there are no background gradients in our linear stability analysis.[3] Moreover, treating $\kappa$ as a constant (and not a function of $B$, $p_c$ and other fluid variables) is exact to linear order in a uniform background.

### 2.5.2 Linearised Equations

We ignore perturbations of $\mathcal{H}$, i.e. we set $\delta\mathcal{H} = 0$ (generalization to finite $\delta\mathcal{H}$ is straightforward). We do, however, perturb the cosmic-ray heating term. The linearised perturbed versions of equations 2.1–2.5 are

$$\frac{\delta\rho}{\rho} = -i\boldsymbol{k} \cdot \boldsymbol{\xi}, \tag{2.20}$$

---

[3] For $\nabla p_c$ to have a well-defined sign in our linear stability analysis, so that $-\boldsymbol{v_A} \cdot \boldsymbol{\nabla} p_c$ is positive definite, $p_c$ cannot be completely uniform. We therefore need a small background CR pressure gradient and to this end, we write $-\boldsymbol{v_A} \cdot \boldsymbol{\nabla} p_c = \epsilon \rho^2 \Lambda$. In our linear calculation we can then still drop any background gradients if we adopt the ordering $1 \gg \epsilon \gg \delta Q/Q$ for any quantity $Q$. Under this ordering, we can essentially treat the equlibrium $\rho$, $p_g$ and $p_c$ as uniform. We note, however, that this approach breaks down when $\delta p_c/p_c > (kH_c)^{-1}$, as the perturbations are large enough to flatten out the CR pressure distribution and decouple the cosmic rays from the gas. In the small-background-gradient limit that is the assumption in our uniform medium calculation, this can in practice happen at small $\delta p_c/p_c$.

$$-\rho\omega^2\boldsymbol{\xi} = -i\boldsymbol{k}\Big(\delta p_c + \delta p_g + \frac{\boldsymbol{B}\cdot\boldsymbol{\delta B}}{4\pi}\Big) + i\frac{(\boldsymbol{B}\cdot\boldsymbol{k})\boldsymbol{\delta B}}{4\pi}, \tag{2.21}$$

$$\boldsymbol{\delta B} = i(\boldsymbol{B}\cdot\boldsymbol{k})\boldsymbol{\xi} - i\boldsymbol{B}(\boldsymbol{k}\cdot\boldsymbol{\xi}), \tag{2.22}$$

$$\frac{\delta p_g}{p_g}\Big(\frac{\omega}{\gamma-1} + i\omega_c\Lambda_T\Big) - \omega_a\frac{\delta p_c}{p_g} = \frac{\delta\rho}{\rho}\Big(\omega\frac{\gamma}{\gamma-1} - i\omega_c(2-\Lambda_T)\Big), \tag{2.23}$$

$$\frac{\delta p_c}{p_g}(\omega - \omega_a + i\omega_d) = \frac{\delta\rho}{\rho}\eta\Big(\frac{4}{3}\omega - \frac{2}{3}\omega_a\Big). \tag{2.24}$$

### 2.5.3 Dispersion Relation

We find the exact solutions to (2.20)–(2.24) by numerically solving for the matrix eigenvalues using MATLAB (because the complete dispersion relation is long and not very enlightening, we do not write it down explicitly). We filter out Alfvén waves, which decouple and do not affect thermal instability, by restricting $\boldsymbol{\xi}$, $\boldsymbol{\delta B}$ and $\boldsymbol{k}$ to lie in the $xz$-plane. The exact gas entropy eigenmode that can be derived from (2.20)–(2.24) is necessary for studying thermal instability at low $\beta$. However, we find that our results depend only mildly on $\beta$ for $\beta \gtrsim 3$ (see middle panels of Figure 2.2). In the high-$\beta$ regime the equations simplify considerably, as the CR and gas pressures satisfy the approximate pressure balance $\delta p_c \approx -\delta p_g$. Equations 2.23 and 2.24 then decouple from the rest (cf. thermal instability is associated with the entropy mode in standard hydrodynamics/MHD) and we end up with a quadratic dispersion relation:

$$\begin{gathered} 0 = \eta\Big(\frac{4}{3}\omega - \frac{2}{3}\omega_a\Big)\Big(\frac{3}{2}\omega + \omega_a + i\omega_c\Lambda_T\Big) + \\ \Big(\omega - \omega_a + i\omega_d\Big)\Big(\frac{5}{2}\omega - i\omega_c(2-\Lambda_T)\Big). \end{gathered} \tag{2.25}$$

Note that $\omega_a$ here is due to the perturbed CR heating, and not due to the magnetic tension or pressure forces; the latter are 0 in the approximation used here. We show the calculated growth rate as a function of $\eta$ in the left panels of Figure 2.2, focusing mainly on our fiducial parameters, $\omega_a = 10^3\omega_c$ and $\beta = 100$. The solution of equation 2.25 is not explicitly plotted, as it agrees almost perfectly with the exact solution at $\beta = 100$ and would not be visible (the second solution to equation 2.25, associated with the CR entropy mode, is not shown as it is stable for all $\eta$). The top plot shows the growth rate for different cooling curve slopes $\Lambda_T$. For $\Lambda_T = -1$, we also show how our results change for $\omega_a = 10\omega_c$ (green dashed line; no visible change), $\omega_a = \omega_c$ (green dash-dotted line) and $\omega_a = 0$ (i.e. $\boldsymbol{k}\cdot\boldsymbol{B} = 0$; green dotted line). The middle panel shows how the $\beta = 100$ growth rate ($\approx \beta \to \infty$ growth rate; blue line, $\Lambda_T = -1$) compares to growth rates at smaller $\beta$. We see that the agreement with, e.g., the $\beta = 3$ calculation is still remarkably good. The bottom panel shows the effects of diffusion (again for $\Lambda_T = -1$) for modes with $\omega_d = 0$ (blue), $\omega_a \gg \omega_d \gg \omega_c$ (orange) and $\omega_d \gg \omega_a$ (green). For more discussion of the effects of diffusion, see Section 2.5.6 and Appendix 2.B.

In the case $\omega_a = 0$ (due to $\boldsymbol{k}\cdot\boldsymbol{B} = 0$; green dotted line in top panel) and $\omega_d = 0$ (no diffusion), equation 2.24 reduces to an adiabatic relation between $\delta p_c$ and $\delta\rho$, with adiabatic index 4/3. Our perpendicular-modes calculation is therefore the closest to the calculation in Pfrommer (2013), who assumed an adiabatic relation between CR pressure and gas density. However, our results are different, as the heuristic calculation in Pfrommer (2013) is not accurate: in particular, their perturbed CR heating is not correct.[4]

#### 2.5.3.1 Effect of CR Streaming on Entropy Modes

Before discussing the thermally unstable modes driven by cooling in more detail, we first consider the effect of CR streaming on the entropy modes, which becomes clear if we ignore cooling and CR diffusion in (2.25), i.e. consider $\omega_c = \omega_d = 0$. The dispersion relation then becomes:

$$\eta\Big(\frac{4}{3}\omega - \frac{2}{3}\omega_a\Big)\Big(\frac{3}{2}\omega + \omega_a\Big) + \frac{5}{2}\omega\Big(\omega - \omega_a\Big) = 0. \tag{2.26}$$

This dispersion relation is in fact a statement of pressure balance and can be obtained by setting $\delta p_c + \delta p_g = 0$ (without cooling and CR diffusion). When CR pressure is negligible ($\eta \to 0$), we see that the two solutions are the ordinary MHD gas entropy mode, with $\omega = 0$ (as CR heating is negligible), and the CR entropy mode, which due to the perturbed work done by the CRs ($-\boldsymbol{v_A} \cdot \boldsymbol{\nabla}\delta p_c$)[5] has a frequency $\omega = \omega_a$.

When CR pressure dominates ($\eta \gg 1$), the CR entropy mode frequency is $\omega = \omega_a/2$, as can also be seen directly from equation 2.24 (with $\omega_d = 0$). This comes directly from the CR compressibility term $-(4/3)p_c\boldsymbol{\nabla} \cdot (\boldsymbol{v} + \boldsymbol{\delta v_A})$, which at large CR pressures is more important for the CR entropy mode evolution than the work done by the CRs on the gas (which is related to the term $-\boldsymbol{v_A} \cdot \boldsymbol{\nabla}\delta p_c$). The gas entropy mode at large $\eta$ is isochoric ($|\delta p_c/p_g| \approx |\delta p_g/p_g| \gg |\delta\rho/\rho|$; see Section 2.5.4). CR heating then dominates the evolution of gas-pressure oscillations (LHS of eq. 2.23) and the oscillations occur at a frequency $\omega = -(2/3)\omega_a$.

Thus, CR streaming always gives rise to an oscillatory frequency $O(\omega_a)$ in the CR entropy mode, while in the gas entropy mode CR heating introduces oscillations as long as $\eta$ is finite, and the oscillation frequency approaches $O(\omega_a)$ once $\eta \sim 1$. Note that while in the classic calculations of thermal instability (e.g., Field 1965) the entropy mode is overstable just due to gravity (rather

---

[4] For perpendicular modes ($\omega_a = 0$), the perturbed CR heating is $\boldsymbol{\delta v_A} \cdot \boldsymbol{\nabla} P_c$, which is 0 in a uniform background. In the adiabatic calculation in Pfrommer (2013), the CR heating is incorrectly assumed to scale as $H_{\rm CR} \propto \rho^{\gamma_c + 1/3 - 1/2}$ and contributes to thermal instability as long as there are density perturbations. This heating term is dominated by an assumed dependence $H_{\rm CR} \propto \delta p_c \propto \rho^{\gamma_c}$; in fact, because CRs are adiabatic only for perpendicular modes with $\omega_a = 0$, there is no contribution to $H_{\rm CR}$ from $\delta p_c$. Moreover, for modes with $\boldsymbol{k} \cdot \boldsymbol{B} \neq 0$, perturbations to $H_{\rm CR} \propto \boldsymbol{v_A} \cdot \boldsymbol{\nabla}\delta p_c \propto \omega_a \delta p_c$ primarily contribute to an oscillatory response, not a change to the growth rate. This is also not captured in the heuristic calculation in Pfrommer (2013).

[5] Due to the "-", $-\boldsymbol{v_A} \cdot \boldsymbol{\nabla} p_c$ is actually positive definite, possibly suggesting that the CRs gain energy according to eq. 2.5. However, when the CR energy equation is rewritten in the conservative form,

$$\frac{\partial p_c}{\partial t} + \frac{4}{3}\boldsymbol{\nabla} \cdot \Big((\boldsymbol{v} + \boldsymbol{v_A})p_c\Big) = \frac{1}{3}(\boldsymbol{v} + \boldsymbol{v_A}) \cdot \boldsymbol{\nabla} p_c,$$

it becomes clear that this term is in fact associated with the work done by the CRs on the Alfvén waves (and hence the gas). Note that the CR energy is $E_c = 3p_c$.

than purely growing when there is no gravity), thermal instability modes are overstable even without gravity when there is a finite CR pressure. In particular, in the presence of CR heating thermally unstable modes are waves propagating at a speed $\propto v_A$.

### 2.5.4 Density versus Temperature Perturbations

Equation 2.24 (and $\delta p_c \approx -\delta p_g$) offers insight into the relative importance of $\delta p_g$ and $\delta\rho$ for driving thermal instability. This turns out to depend primarily on the CR pressure fraction $\eta$, due to the coupling of $\delta p_c$ and $\delta\rho$ via equation 2.24. Typically, we have that:[6]

$$|\delta p_g/p_g| \approx |\delta p_c/p_g| \sim \eta|\delta\rho/\rho|, \tag{2.27}$$

so that perturbations are essentially isobaric for $\eta \ll 1$ and isochoric when $\eta \gg 1$ (large CR pressure stiffens the gas). For large $\omega_d$, CR pressure perturbations are suppressed because they are smoothed out by diffusion, and perturbations are isobaric up to larger $\eta$.

### 2.5.5 Asymptotic Limits

We now look back at the dispersion relation in (2.25). How the solutions of (2.25) depend on $\eta$ is particularly transparent. We can read off the solutions in the limits $\eta \to 0$ and $\eta \to \infty$. [7] As $\eta \to 0$, the unstable gas entropy mode is just the standard isobaric thermal instability result, with a small oscillatory part due to the perturbed CR heating:[8]

$$\omega = -\frac{4}{15}\eta\omega_a + \frac{2}{5}i\left(2 - \Lambda_T\right)\omega_c, \tag{2.28}$$

which comes from the isobarically perturbed cooling function, $\delta(-\rho^2\Lambda) = -\omega_c p_g\ (2 - \Lambda_T)\delta\rho/\rho$ (at small $\eta$ we have that $\delta p_g/p_g \ll \delta\rho/\rho$, as discussed in Section 2.5.4). As $\eta \to \infty$, we get an overstable solution:

$$\omega = -\frac{2}{3}\omega_a - \frac{2}{3}i\Lambda_T\omega_c. \tag{2.29}$$

Note that the $-(2/3)\omega_a$ comes from the perturbed CR heating, as discussed in 2.5.3.1, while the $-(2/3)\Lambda_T\omega_c$ growth rate comes from the isochorically perturbed cooling function (recall that in the limit $\eta \to \infty$, $\delta p_g/p_g \gg \delta\rho/\rho$, so that unstable modes are isochoric). CR heating does not directly affect the growth rate. Equations 2.28 and 2.29 are consistent with the low and high $\eta$ limits in Figure 2.2 (left panels).

---

[6] The exceptions to this are if $\omega_d \gg \omega_a$ (so that diffusion wipes out the CR pressure perturbation), or $\omega = \omega_a - i\omega_d$ or $\omega = \omega_a/2$, which are the CR entropy modes at small and large $\eta$ respectively, see Section 2.5.3.1.

[7] Note that in our notation the limits $\eta \to 0$ ($\eta \to \infty$) mean that $\eta$ is much smaller (larger) than any other dimensionless parameter in the problem, e.g. $\omega_d/\omega_a$, $\omega_a/\omega_c$ etc.

[8] The real (oscillatory) part of the solution in eq. 2.28 also assumes $\omega_a > \omega_d$ (for $\omega_d > \omega_a$ the real part vanishes as $\delta p_c$ is suppressed by diffusion).

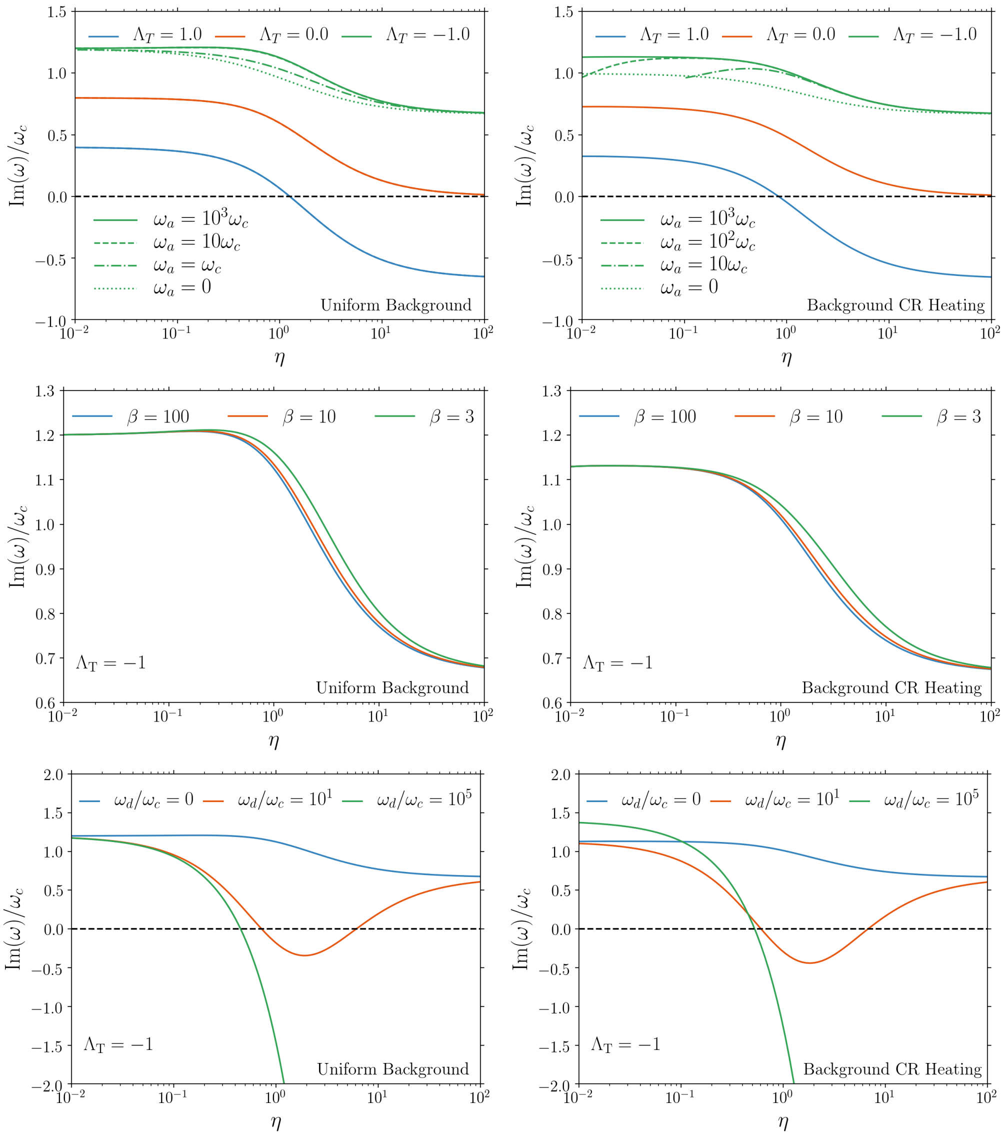


*Figure 2.2*: Thermal instability growth rates as a function of $\eta$. $\mathrm{Im}(\omega) > 0$ corresponds to growing modes. Unless explicitly stated otherwise in the plots, the presented growth rates are for our fiducial parameters ($\omega_a = 10^3\omega_c$ and $\beta = 100$). We consider smaller $\omega_a$ in the top panels (with $\beta = 100$ fixed) and smaller $\beta$ in the middle panels (with $\omega_a = 10^3\omega_c$ fixed). Left panels: Thermal instability in uniform medium. Top: growth/damping rate for different cooling curve slopes. Middle: $\Lambda_T = -1$ growth rate for different $\beta$. For $\beta \gtrsim 3$, the high-$\beta$ result is a very good approximation. Bottom: impact of diffusion on modes with $\boldsymbol{k} \cdot \boldsymbol{B} \neq 0$. Right panels: same as left panels, but for a background with cosmic-ray heating balancing cooling, but no gravity. The small-$\eta$ limit is different for perpendicular modes with $\omega_a = 0$ (dotted green line in top panel; see Section 2.6.3). For $\omega_a = 10\omega_c$ in the top right panel (dash-dotted green line), we only plot the growth rate for $\eta > 0.1$, where our WKB analysis is applicable. In the bottom right plot, the plateau at small $\eta$ depends on whether diffusion is more important than streaming (i.e. whether $\omega_d > \omega_a$ or $\omega_a > \omega_d$).

### 2.5.6 Effect of CR Diffusion

CR diffusion does not suppress the overall excitation of thermal instability. It nevertheless suppresses the growth of some modes which would otherwise be thermally unstable (see, e.g., bottom left panel of Figure 2.2).

We can study the effects of CR diffusion on modes with $\boldsymbol{k} \cdot \boldsymbol{B} \neq 0$ by looking at thermal stability maps in the $(\eta, \omega_d/\omega_a)$ plane. We show this in Figure 2.3. The results of the uniform medium calculation are shown in the top panels, for $\Lambda_T < 0$ (left panel, $\Lambda_T = -1$) and for $2 > \Lambda_T > 0$ (right panel, $\Lambda_T = 1/2$). The blue color corresponds to stable solutions, red denotes growing (i.e. thermally unstable) solutions. We provide approximate boundaries for the region of parameter space where cosmic-ray diffusion can suppress thermal instability (dashed lines). A heuristic derivation of these boundaries can be found in Appendix 2.B.

Here we summarise the main results from Appendix 2.B. For modes with $\omega_d < \omega_a$, diffusion suppresses thermal instability if $\eta$ satisfies:

$$|2 - \Lambda_T| \frac{\omega_c}{\omega_d} \lesssim \eta \lesssim |\Lambda_T|^{-1} \frac{\omega_d}{\omega_c}. \tag{2.30}$$

If $\eta$ is too small for the above condition to be satisfied, the instability is isobaric, with $\mathrm{Im}(\omega) = (2/5)\omega_c\big(2 - \Lambda_T\big)$. If $\eta > |\Lambda_T|^{-1}\omega_d/\omega_c$, the growth rate approaches the asymptotic limit $\mathrm{Im}(\omega) = -(2/3)\Lambda_T\omega_c$ from equation 2.29.

For modes with $\omega_d > \omega_a$, diffusion suppresses thermal instability if $\eta$ satisfies:

$$|2 - \Lambda_T| \frac{\omega_d \omega_c}{\omega_a^2} \lesssim \eta \lesssim |\Lambda_T|^{-1} \frac{\omega_d}{\omega_c}. \tag{2.31}$$

If $\eta < |2 - \Lambda_T|\omega_d\omega_c/\omega_a^2 = |2 - \Lambda_T|\kappa\omega_c/v_A^2$, the instability is again isobaric, with $\mathrm{Im}(\omega) = (2/5)\big(2 - \Lambda_T\big)\omega_c$. When $\eta$ is large, the growth rate again approaches the asymptotic limit $\mathrm{Im}(\omega) = -(2/3)\Lambda_T\omega_c$. Note that in the limit $\kappa \to \infty$, the instability is isobaric for arbitrary $\eta$ because CR diffusion suppresses $\delta p_c$ (recall that in the high-$\beta$ limit we have that $\delta p_c \approx -\delta p_g$, so $\delta p_g \approx 0$ if $\delta p_c$ is suppressed by CR diffusion).

#### 2.5.6.1 Cosmic-Ray Field Length

In Appendix 2.B we show that CR diffusion can suppress thermal instability because it affects the thermal gas in a way akin to thermal conduction (mediated by the perturbed CR heating term, see Appendix 2.B.2). This suggests that there is a CR-diffusion analogue of the Field length for thermal conduction (Field 1965), below which thermal instability is suppressed.

In Appendix 2.B.4 we show that the dimensionless ratio $\kappa\omega_c/(\eta v_A^2)$, the ratio of the cooling rate to the CR-heating rate at high-$k$, determines whether there is a Field length associated with CR diffusion. If $\kappa\omega_c/(\eta v_A^2) \gtrsim 1$ then CR diffusion does not suppress thermal instability of high-$k$ modes ($\omega_d \gg \omega_a$), as the cooling rate exceeds the CR heating rate. There is no “CR Field length" below which thermal instability is completely suppressed. Instead, the instability of high-$k$ modes is isobaric with growth rates $\mathrm{Im}(\omega) = (2/5)\big(2 - \Lambda_T\big)\omega_c$ (as $\delta p_g \approx -\delta p_c$ is suppressed by CR diffusion). On the other hand, if $\kappa \lesssim \eta v_A^2/\omega_c$, there is a maximum $\hat{\boldsymbol{b}} \cdot \boldsymbol{k}$ at which thermal instability occurs (Figure 2.7). The CR Field length is (Appendix 2.B.4):

$$\lambda_{\mathrm{CRF}} \sim \begin{cases} 2\pi|\hat{\boldsymbol{b}} \cdot \hat{\boldsymbol{k}}| \sqrt{\frac{\eta\kappa}{\omega_c}} & \eta < 1 \\ 2\pi|\hat{\boldsymbol{b}} \cdot \hat{\boldsymbol{k}}| \sqrt{\frac{\kappa}{\eta\omega_c}} & \eta > 1. \end{cases} \tag{2.32}$$

Note that the CR Field length is very similar to the classic Field length with the thermal diffusion coefficient replaced by the CR diffusion coefficient. We can estimate $\kappa\omega_c/\eta v_A^2$ for CGMs of Milky-Way-like galaxies:

$$\frac{\kappa\omega_c}{\eta v_A^2} \sim 1 \frac{\kappa}{10^{28}\ \mathrm{cm^2\ s^{-1}}} \frac{\omega_c}{10^{-15}\ \mathrm{s^{-1}}} \left(\frac{\eta}{1}\right)^{-1} \left(\frac{v_A}{3 \times 10^6\ \mathrm{cm\ s^{-1}}}\right)^{-2}. \tag{2.33}$$

We chose $\kappa = 10^{28}$ cm$^2$ s$^{-1}$ motivated by diffusion-only models of CR observations in the Milky Way, which infer $\kappa \sim 10^{28} - 10^{29}$ cm$^2$ s$^{-1}$ depending on the size of the CR halo (e.g., Linden et al. 2010). It is plausible that $\kappa\omega_c/\eta v_A^2 > 1$, so that CR diffusion does not suppress thermal instability at small scales. However, if instead $\kappa\omega_c/\eta v_A^2 < 1$ (e.g., if streaming is the dominant transport process $\kappa$ may be $\ll 10^{28}$ cm$^2$ s$^{-1}$), thermal instability of modes with wavelengths smaller than the CR Field length,

$$\lambda_{\mathrm{CRF}} \sim 7\ \mathrm{kpc}\ |\hat{\boldsymbol{b}} \cdot \hat{\boldsymbol{k}}| \left(\frac{\kappa}{10^{28}\ \mathrm{cm^2\ s^{-1}}}\right)^{1/2} \left(\frac{\omega_c}{10^{-15}\ \mathrm{s^{-1}}},\right)^{-1/2} \tag{2.34}$$

is suppressed by CR diffusion (here we assumed $\eta \sim 1$).

### 2.5.7 Thermal Stability versus Instability

In addition to the slope of the cooling function, $\Lambda_T$, thermal stability clearly also depends on the CR pressure fraction, $\eta$ (which sets whether perturbations are isobaric or isochoric, see Section 2.5.4). We show the “critical" cooling function logarithmic slope, $\Lambda_{T,c}$, demarcating the boundary between thermal stability and instability to any local perturbation (satisfying $\omega_a \gg \omega_c$), in the top panel of Figure 2.4. The solid black line shows the boundary without CR diffusion, i.e. $\kappa = 0$, and the magenta dotted line is a simple broken power-law fit of the form

$$\Lambda_{T,c} = 2[1 + \eta/\eta_*]^q, \tag{2.35}$$

where $\eta_* = 1.62$ and $q = -1.19$ are the best-fit parameters. For a given $\eta$, the system is thermally unstable if $\Lambda_T < \Lambda_{T,c}$. The dotted line is the thermal stability boundary for perpendicular modes only, i.e. for $\omega_a = 0$, for which $p_c$ and $\rho$ follow an adiabatic relation with index 4/3 (note that our thermal stability criterion is not the same as in Pfrommer 2013; see last paragraph of Section 2.5.3). It is notable that the $\omega_a = 0$ and $\omega_a \gg \omega_c$ values of $\Lambda_{T,c}$ in Figure 2.4 are very similar. This again highlights that perturbed CR heating does not significantly affect the growth rates of thermal instability. Instead, it turns a purely growing mode into an overstability (see Section 2.5.3.1).

The dashed horizontal line $\Lambda_{T,c} = 2$ is the thermal stability/instability boundary if CR diffusion is present and $\kappa \gg \eta v_A^2/\omega_c$. $\Lambda_{T,c} = 2$ due to the fact that for $\kappa \gg \eta v_A^2/\omega_c$ high-$k$ perturbations are

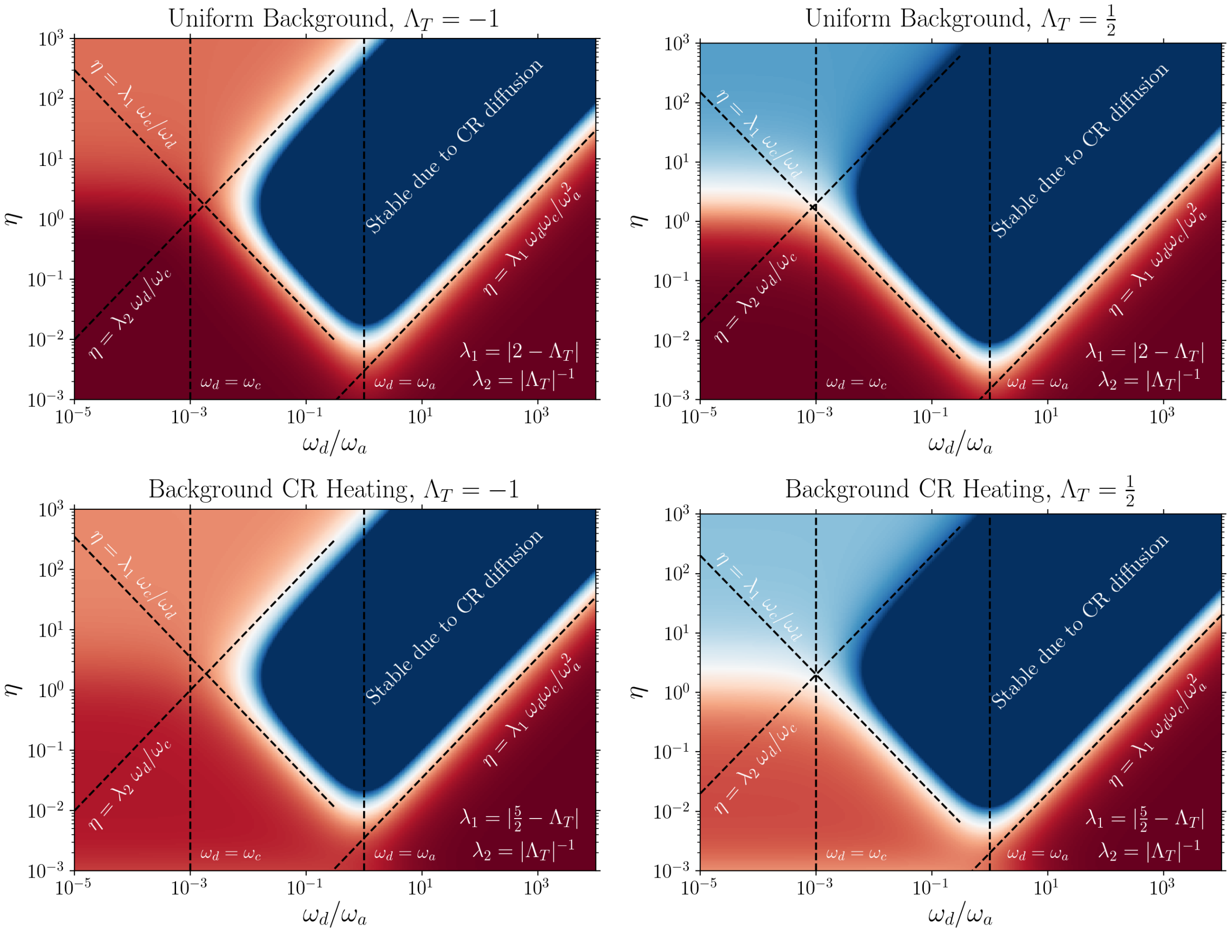


*Figure 2.3*: Effect of CR diffusion on thermal instability. We show thermal stability/instability boundaries of modes with $\boldsymbol{k} \cdot \boldsymbol{B} \neq 0$ in the ($\eta$, $\omega_d/\omega_a$) plane, for $\omega_a = 10^3\omega_c$ and $\beta \to \infty$ (the fiducial $\beta = 100$ case looks the same). $\mathrm{Im}(\omega) > 0$ (growing modes) are shown in red, $\mathrm{Im}(\omega) < 0$ (decaying modes) are shown in blue. Top panels: Stability/instability boundaries in uniform medium. Left: $\Lambda_T = -1$. Right: $\Lambda_T = 1/2$. The dark blue shows the region where thermal instability is suppressed by CR diffusion. The light blue shows thermal stability due to $\Lambda_T > 0$ (eq. 2.39). The approximate boundaries (dashed lines) of the diffusion-affected region are derived in Appendix 2.B. Bottom panels: same as top panels, but in a background with CR heating balancing cooling (but no gravity). Note that the growth rate at small $\eta$ (dark vs light red) now depends on whether $\omega_a > \omega_d$ or $\omega_d > \omega_a$ ($\mathrm{Im}(\omega) = (2/5)\omega_\mathrm{c}\big(11/6 - \Lambda_\mathrm{T}\big)$ and $\mathrm{Im}(\omega) = (2/5)\omega_\mathrm{c}\big(5/2 - \Lambda_\mathrm{T}\big)$ respectively).

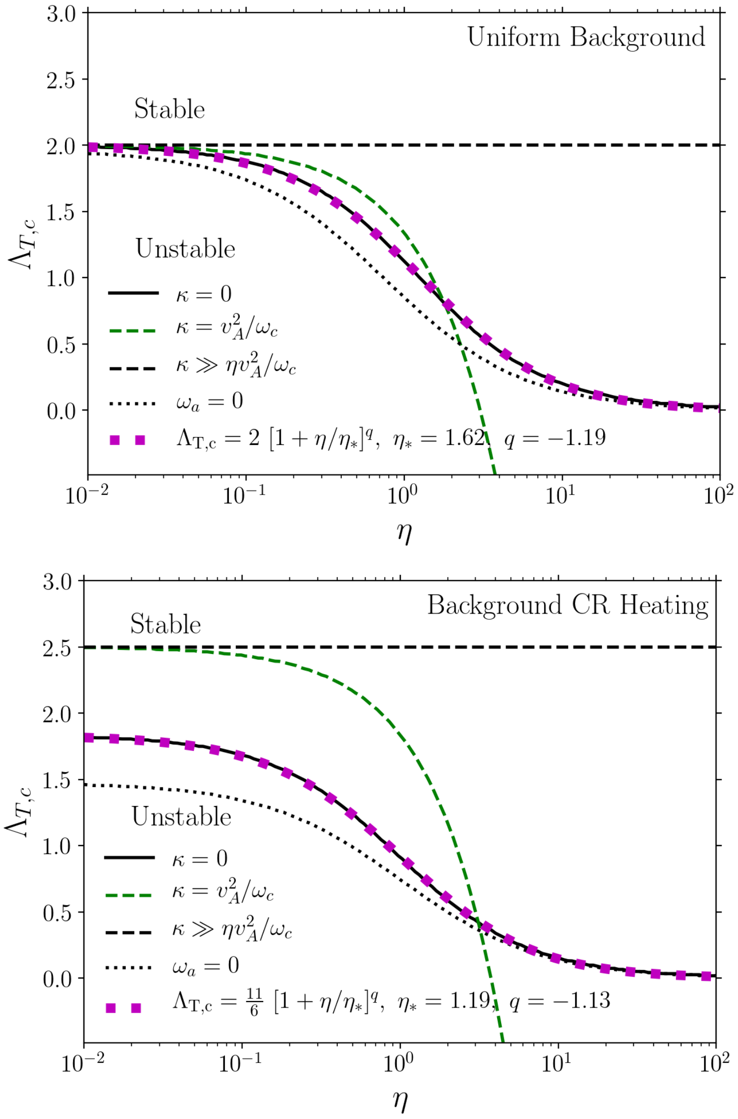


*Figure 2.4*: $\Lambda_{T,c}$ versus $\eta \equiv p_c/p_g$, where $\Lambda_{T,c}$ is the $\partial \ln \Lambda / \partial \ln T$ that defines the boundary between overall thermal stability and instability. For a given $\eta$, thermal instability occurs if $\Lambda_T < \Lambda_{T,c}$. Top: $\Lambda_{T,c}$ in a uniform medium. Bottom: $\Lambda_{T,c}$ in a medium with background CR heating. We use $\beta = 100$ and we include modes that satisfy $\omega_a > 10\omega_c$ ($\omega_a > 10\omega_c$ and $\omega_a > 10\omega_c \eta^{-1}$) in the uniform (CR-heated) background. $\Lambda_{T,c}$ does not change significantly for $\beta \gtrsim 3$, and for $\omega_a \gtrsim \omega_c$ ($\omega_a \gtrsim \omega_c \eta^{-1}$) in the uniform (CR-heated) background absent diffusion (i.e. growth rates are approximately constant for local perturbations satisfying $kH \gtrsim 1$; in Figure 2.2 we show how growth rates depend on $\omega_a$ and $\beta$). The solid black line is for $\kappa = 0$ (no CR diffusion) and the dotted magenta line is a simple broken power-law fit. The horizontal dashed line is the stability/instability boundary when CR diffusion is present and $\kappa \gg \eta v_A^2/\omega_c$, while the green dashed line shows the boundary for $\kappa = v_A^2/\omega_c$. The dotted lines are the thermal instability boundaries for perpendicular modes only (i.e. $\omega_a = \omega_d = 0$).

isobaric, and so high-$k$ modes (with $\boldsymbol{k} \cdot \boldsymbol{B} \neq 0$) always have a growth rate $\mathrm{Im}(\omega) = (2/5)\omega_c\left(2 - \Lambda_T\right)$ (see discussion in Section 2.5.6.1). For $\kappa \lesssim \eta v_A^2/\omega_c$, high-$k$ perturbations are suppressed by CR diffusion. Only modes with wavelengths longer than the CR Field length can be thermally unstable. For $\kappa = v_A^2/\omega_c$ (green dashed line) and $\eta > 1$, the CR Field length is at lower $k$ than the modes used for the stability boundary calculation ($\omega_a > 10\omega_c$) and so CR diffusion suppresses thermal instability of these modes. Note that for $\kappa = v_A^2/\omega_c$ and $\eta \ll 1$ $\Lambda_{T,c} = 2$, as high-$k$ modes are isobaric at low $\eta$. We stress again that perpendicular modes ($\boldsymbol{k} \cdot \boldsymbol{B} = \mathbf{0}$, dotted line) are not affected by CR diffusion.

### 2.5.8 Photoionization Equilibrium

We can easily extend our CR thermal instability analysis to a background in photoionization equilibrium (PIE), with no background CR heating, but where CR heating is still present in the perturbed equations. We can then treat PIE analogously to our uniform background, by absorbing photoionization heating and cooling into an effective cooling function $\Lambda$. In PIE, this effective cooling function satisfies $\Lambda_T > 2$ (e.g., Wiersma et al. 2009), and so such systems are thermally stable for any CR pressure fraction $\eta$.

## 2.6 Equilibrium with Cosmic-Ray Heating Balancing Cooling

### 2.6.1 Equilibrium

In this section we look at equilibria in which cooling is completely balanced by cosmic-ray heating,

$$-\boldsymbol{v_A} \cdot \boldsymbol{\nabla} p_c = \rho^2 \Lambda(T) = \omega_c p_g. \tag{2.36}$$

We still ignore gravity, i.e. we set $\boldsymbol{g} = 0$. Hydrostatic equilibrium implies that the CR pressure gradient is balanced by the gas pressure gradient:

$$\boldsymbol{\nabla} p_c = -\boldsymbol{\nabla} p_g. \tag{2.37}$$

Without loss of generality, we assume that the variation is purely in the vertical direction, i.e. $\boldsymbol{\nabla} p_c = (\partial p_c/\partial z)\hat{\boldsymbol{z}}$. We assume a uniform magnetic field, $\boldsymbol{B} = B \sin\theta_B \hat{\boldsymbol{x}} + B \cos\theta_B \hat{\boldsymbol{z}}$.

We choose the background pressures such that they have a linear profile, i.e. $\partial p_c/\partial z = \mathrm{const}$, so that CR diffusion does not enter in the equilibrium setup. The cosmic-ray pressure equation (2.5) then implies that $p_c \propto \rho^{2/3}$ and $\boldsymbol{\nabla} \cdot \boldsymbol{v_A} = (3/4)\omega_c \eta^{-1}$.

### 2.6.2 Linear Perturbations in 1 Dimension

The background gradients give rise to extra terms in linear perturbation theory, which modify equations 2.20–2.24. We show the linearised equations in a medium with background CR heating in Appendix 2.C, which we again solve using MATLAB (there is again little physical insight gained from explicitly writing down the 6th-order dispersion relation).

In addition to explicitly solving equations 2.74–2.78, we consider the simpler 1-dimensional problem in which $\boldsymbol{B}$, $\boldsymbol{k}$, $\boldsymbol{\xi}$ (as well as the background gradients) are all along $\hat{\boldsymbol{z}}$. This is motivated by the fact that we find that 1D thermal instability growth rates agree essentially perfectly with the more general calculation with $\boldsymbol{B}$, $\boldsymbol{k}$ and $\hat{\boldsymbol{z}}$ not aligned (unless $\boldsymbol{k}\cdot\boldsymbol{B}=0$, which we treat separately in Section 2.6.3). We explain why the 1D calculation correctly predicts the gas entropy mode eigenfrequency in Appendix 2.D. In the high-$\beta$ limit, the 1D dispersion relation simplifies to a quadratic:

$$\begin{aligned}0 = &\Big(\frac{4}{3}\eta\omega - \frac{2}{3}\eta\omega_a + i\omega_c - \omega\omega_c\big(\frac{3}{2}\omega_c\eta^{-1} - i\omega_a\big)^{-1}\Big)\\ &\Big(\frac{3}{2}\omega + \omega_a + i\omega_c\Lambda_T\Big) + \Big(\omega - \omega_a + i\omega_d + i\omega_c\eta^{-1}\Big)\\ &\Big(\frac{5}{2}\omega - i\omega_c\big(\frac{5}{2} - \Lambda_T\big) - \frac{3}{2}\omega\omega_c\big(1 + \frac{5}{2}\eta^{-1}\big)\big(\frac{3}{2}\omega_c\eta^{-1} - i\omega_a\big)^{-1}\Big).\end{aligned} \tag{2.38}$$

Note that $\omega_a$ again shows up due to the perturbed CR heating, and not as a result of the perturbed magnetic field (indeed, $\boldsymbol{\delta B} = 0$ in 1D).

Thermal instability growth rates as a function of $\eta$ in a medium with background CR heating are shown in the right panels of Figure 2.2. Unless explicitly stated otherwise, the growth rates are for our fiducial parameters, $\omega_a = 10^3\omega_c$ and $\beta = 100$. The $\beta \to \infty$ growth rates calculated from (2.38) overlap almost perfectly with the $\beta = 100$ calculation (and are therefore not explicitly plotted). The top right panel shows the growth rates for different cooling curve slopes $\Lambda_T$. We also show how $\Lambda_T = -1$ growth rates change for smaller $\omega_a$: $\omega_a = 10^2\omega_c$, $\omega_a = 10\omega_c$ and $\omega_a = 0$ ($\boldsymbol{k}\cdot\boldsymbol{B} = 0$, see Section 2.6.3). The middle panel shows how the $\beta = 100$ growth rate ($\approx \beta \to \infty$ growth rate) compares to the growth rate at smaller $\beta$ and the same $\Lambda_T$. The bottom panel shows the effects of diffusion (again for $\Lambda_T = -1.0$), in the limits $\omega_d = 0$ (blue), $\omega_a \gg \omega_d \gg \omega_c$ (orange) and $\omega_d \gg \omega_a$ (green). Note that unlike the uniform-medium calculation, the growth rate at small $\eta$ now depends on whether the mode is diffusion dominated (i.e. whether $\omega_d > \omega_a$, see 2.6.2.1).

#### 2.6.2.1 Asymptotic Limits

We now consider the asymptotic limits in the presence of cooling, CR streaming and diffusion. The $\eta \to \infty$ limit is again simple and can be read off directly from (2.38). The solution is overstable,

$$\omega = -\frac{2}{3}\omega_a - \frac{2}{3}i\Lambda_T\omega_c, \tag{2.39}$$

and is identical to the uniform-medium large-$\eta$ result. The $-(2/3)\omega_a$ oscillation frequency again comes from the perturbed CR heating (see Section 2.5.3.1), which turns thermally unstable modes into propagating waves.

The small-$\eta$ limit ($\omega_a \gg \omega_c\eta^{-1} \gg \omega_c$)[9] depends on whether the mode is streaming or diffusion dominated. In the streaming-dominated case (i.e. modes with $\omega_d \ll \omega_a$), the isobaric growth rate,

---

[9] Recall that in our local analysis we only consider $\omega_a > \omega_c\eta^{-1}$. This corresponds to perturbations that satisfy $kH \gtrsim 1$, $H$ being a characteristic background length scale.

$(2/5)\omega_c(2 - \Lambda_T)$, which comes from isobaric perturbations to the cooling function, $\delta(-\rho^2\Lambda) = -\omega_c p_g\,(2 - \Lambda_T)\delta\rho/\rho$, is modified by CR streaming and background heating,[10]

$$\omega = -\frac{4}{15}\eta\omega_a + i\frac{2}{5}\omega_c\Big(\frac{11}{6} - \Lambda_T\Big). \tag{2.40}$$

For modes with $\omega_d \gg \omega_a$, one can show that

$$\omega = i\frac{2}{5}\omega_c\Big(\frac{5}{2} - \Lambda_T\Big). \tag{2.41}$$

In this strong-diffusion limit, the 5/2 (instead of 2) arises from the perturbed CR heating term, $-\boldsymbol{\delta v_A} \cdot \boldsymbol{\nabla} p_c = -(1/2)\omega_c p_g \delta\rho/\rho$. Note that diffusion suppresses CR pressure perturbations, so that $-\boldsymbol{v_A} \cdot \boldsymbol{\nabla}\delta p_c$ is suppressed and does not give rise to gas-entropy oscillations (i.e., the mode is purely growing, unlike eq. 2.40).

These results differ modestly from the uniform medium calculation (see Section 2.5.5 or compare the left and right panels of Figure 2.2), as now the background CR pressure gradient modifies the growth rate.

## 2.6.3 Perpendicular Modes

### 2.6.3.1 Dispersion Relation

The 1D calculation does not apply to modes propagating perpendicular to the magnetic field direction, such that $\boldsymbol{k} \cdot \boldsymbol{B} = 0$. In this case, we can obtain an approximate quadratic dispersion relation by taking the high-$\beta$ limit and dropping advective ($\boldsymbol{\xi} \cdot \boldsymbol{\nabla}$) background-gradient terms:

$$\Big(\omega + i\omega_c\eta^{-1}\Big)\Big(\frac{5}{2}\omega - i\Big(\frac{3}{2} - \Lambda_T\Big)\omega_c\Big) + \Big(\frac{4}{3}\eta\omega + i\omega_c\Big)\Big(\frac{3}{2}\omega + i\omega_c\Lambda_T\Big) = 0. \tag{2.42}$$

Note that $\omega_a = \omega_d = 0$ for modes with $\boldsymbol{k} \cdot \boldsymbol{B} = 0$, and so do not show up in the above dispersion relation.

### 2.6.3.2 Asymptotic Limits

The $\eta \to \infty$ growth rate does not change, and the solution is now a purely growing mode (as $\omega_a = 0$):

$$\omega = -\frac{2}{3}i\Lambda_T\omega_c. \tag{2.43}$$

For $\eta \to 0$, the mode is also purely growing, with

$$\omega = \frac{2}{5}i\Big(\frac{3}{2} - \Lambda_T\Big)\omega_c. \tag{2.44}$$

---

[10] This can be shown by solving eq. 2.38 perturbatively using the ordering $\omega_a \gg \omega_c\eta^{-1} \gg \omega_c$.

This differs from the corresponding limit in Section 2.6.2.1, as for perpendicular modes the perturbed CR heating is $-\boldsymbol{\delta v_A} \cdot \boldsymbol{\nabla} p_c \approx (1/2)\omega_c p_g \delta\rho/\rho$, while the isobarically perturbed cooling function is still $-\omega_c p_g\,(2-\Lambda_T)\delta\rho/\rho$. Note that because $\boldsymbol{k}\cdot\boldsymbol{B} = 0$, there are no entropy oscillations driven by CR heating. The growth rate of perpendicular modes as a function of $\eta$ (for $\Lambda_T = -1$) is plotted as a dotted green line in the top-right panel of Figure 2.2.

### 2.6.4 Effect of CR Diffusion

The bottom panels of Figure 2.3 show stability maps of modes with $\boldsymbol{k}\cdot\boldsymbol{B} \neq 0$ in the $\left(\eta, \omega_d/\omega_a\right)$ plane in a medium with background CR heating. Once again, blue denotes stable solutions, while red denotes growing solutions. The left panel is for $\Lambda_T = -1$ and the right panel is for $\Lambda_T = 1/2$. We again provide approximate boundaries for the region of parameter space where CR diffusion suppresses thermal instability (dashed lines). Note that these *order-of-magnitude* boundaries are essentially the same as in the uniform medium case (see equations 2.30 and 2.31, and Appendix 2.B for a heuristic derivation). (2.30) and (2.31) are only slightly modified to emphasise the extra contribution coming from terms related to background CR heating, and are now:

$$|\frac{5}{2} - \Lambda_T|\frac{\omega_c}{\omega_d} \lesssim \eta \lesssim |\Lambda_T|^{-1}\frac{\omega_d}{\omega_c} \qquad (\omega_d < \omega_a) \tag{2.45}$$

and

$$|\frac{5}{2} - \Lambda_T|\frac{\omega_d\omega_c}{\omega_a^2} \lesssim \eta \lesssim |\Lambda_T|^{-1}\frac{\omega_d}{\omega_c} \qquad (\omega_d > \omega_a) \tag{2.46}$$

respectively. If $\eta$ satisfies the above, thermal instability of modes with the corresponding $\omega_d$ and $\omega_a$ is suppressed by CR diffusion.

As in Section 2.5.6.1, conditions 2.45 and 2.46 can be used to derive a CR Field length below which thermal instability is suppressed by CR diffusion. Like in the uniform medium, if $\kappa\omega_c/(\eta v_A^2) \gtrsim 1$ then CR diffusion does not suppress thermal instability of high-$k$ modes and there is no associated CR Field length. If on the other hand $\kappa\omega_c/(\eta v_A^2) \lesssim 1$, the CR Field length below which CR diffusion suppresses thermal instability is approximately given by (2.32) (ignoring factors of order unity, e.g. $\propto \Lambda_T$). See Appendix 2.B.4 for more discussion.

### 2.6.5 Thermal Stability versus Instability

We show $\Lambda_{T,c}$ (the $\Lambda_T$ that is the boundary between overall thermal stability and instability, to any *local* perturbation satisfying $\omega_a \gg \omega_c\eta^{-1}$) as a function of $\eta$ in the bottom panel of Figure 2.4. The solid black line again shows the boundary for $\kappa = 0$, and the magenta dotted line is a broken power-law fit of the form

$$\Lambda_{T,c} = \frac{11}{6}[1 + \eta/\eta_*]^q, \tag{2.47}$$

with $\eta_* = 1.19$ and $q = -1.13$ being the best-fit parameters. For a given $\eta$, the system is thermally unstable if $\Lambda_T < \Lambda_{T,c}$. The dotted line shows the thermal stability boundary for perpendicular modes only, with $\omega_a = \omega_d = 0$ (note the lower plateau at small $\eta$, see eq. 2.44).

The dashed horizontal line $\Lambda_{T,c} = 5/2$ is the thermal stability/instability boundary if CR diffusion is present and $\kappa \gg \eta v_A^2/\omega_c$. For $\kappa \gg \eta v_A^2/\omega_c$ high-$k$ perturbations are isobaric and have a growth rate $\mathrm{Im}(\omega) = (2/5)\omega_c\big(5/2 - \Lambda_T\big)$ (Section 2.5.6.1). For $\kappa \lesssim \eta v_A^2/\omega_c$, high-$k$ perturbations are suppressed by CR diffusion. Only long-wavelength modes above the CR Field length can be thermally unstable. For $\kappa = v_A^2/\omega_c$ (green dashed line) and $\eta > 1$, the CR Field length is at lower $k$ than the modes used for the stability boundary calculation ($\omega_a > 10\omega_c\eta^{-1}, 10\omega_c$) and so CR diffusion suppresses thermal instability of these modes. Perpendicular modes ($\boldsymbol{k} \cdot \boldsymbol{B} = 0$, dotted line) are not affected by CR diffusion.

## 2.7 CR Heating in Gravitational Field

### 2.7.1 Equilibrium

As in Section 2.6, we consider equilibria in which cooling is completely balanced by cosmic-ray heating (equation 2.16). Throughout this section, we neglect the effects of diffusion. Gravity, $\boldsymbol{g} = -g\hat{\boldsymbol{z}}$, changes the background gas pressure gradient to:

$$\frac{dp_g}{dz} = \frac{dp_c}{dz}\Big(\gamma\frac{\omega_{\rm ff}}{\omega_c}\frac{v_{A,z}}{c_s} - 1\Big), \tag{2.48}$$

where $v_{A,z}$ is the z-component of the Alfvén velocity. We again assume a uniform magnetic field, $\boldsymbol{B} = B\sin\theta_B\hat{\boldsymbol{x}} + B\cos\theta_B\hat{\boldsymbol{z}}$.

### 2.7.2 Thermal Instability

We find numerically (by solving equations 2.74–2.78 in MATLAB) that gravity does not significantly change thermal overstability growth rates for most modes (and it only slightly changes the entropy-mode oscillation frequency, which is dominated by the perturbed CR heating, i.e. $\omega_a$, unless $\omega_a < \omega_{\rm ff}$). In particular, for $\omega_a \gg \omega_{\rm ff}$ we recover the same growth rates as in Section 2.6 and the growth rates obtained from equation 2.38 generally agree well with the exact calculation (which includes gravity). The green line in Figure 2.5 shows this for $\Lambda_T = -1.0$, $\omega_{\rm ff} = 20\omega_c$ and $\omega_a \gg \omega_{\rm ff}$ ($\omega_a = 10^3\omega_c$): the growth rate is again $(2/5)\omega_c\big(11/6 - \Lambda_T\big)$ at small $\eta$ and $-(2/3)\Lambda_T\omega_c$ at large $\eta$.

Gravity is more important when $\omega_a < \omega_{\rm ff}$ (e.g. modes with $\boldsymbol{k} \cdot \boldsymbol{B} = 0$). This is shown by the blue and orange curves in Figure 2.5 (with $\omega_a = 0$ and $\omega_{\rm ff} = 20\omega_c$). At small $\eta$, gravity reduces the thermal instability growth rate by a factor of $\sim 2$ and the real part of the overstable entropy mode is dominated by the free-fall frequency, as has been found in previous work (Field 1965).[11]

---

[11] For $\eta \ll 1$, gravity reduces the thermal instability growth rate to $(1/5)\omega_c\big(5/2 - \Lambda_T\big)$, where the 5/2 (instead of 2) arises from the perturbed CR heating term, $-\boldsymbol{\delta v_A} \cdot \boldsymbol{\nabla} p_c \approx -(1/2)\omega_c p_g \delta\rho/\rho$.

### 2.7.3 Convective Instability

Figure 2.5 also shows that there is a new form of instability occurring at larger $\eta$. We will show below that buoyancy is responsible for the increased growth rate. The buoyancy instability occurs only when $\omega_a \lesssim \omega_{\rm ff}$, which corresponds to approximate adiabaticity. We note that convective behaviour in the presence of cosmic rays has been studied before by Chandran & Dennis (2006), Dennis & Chandran (2009) and Heintz & Zweibel (2018). However, the setup we consider here, i.e. gravitationally stratified media with *background CR heating*, was not part of their calculations. Chandran & Dennis (2006) and Dennis & Chandran (2009) did not include CR heating and focused on the effects of CR diffusion and thermal conduction (which tend to smooth out CR pressure and gas temperature along field lines, so their calculation differs substantially from our Schwarzschild-like calculation below). Heintz & Zweibel (2018) looked at the effect of CR heating on the Parker instability. However, they did not consider *background* CR heating, which is central to our buoyancy-instability calculation. As a result, their instability calculation was different from the setup we consider here.

#### 2.7.3.1 Convective Instability Condition via Schwarzschild Criterion

We can derive a convective stability criterion using the standard picture of a rising blob, which maintains pressure balance with its surroundings and is (approximately) adiabatic. For the latter, we require that $\omega_c < \omega_{\rm ff}$ (typically satisfied in galactic and cluster halos) and that $\omega_a < \omega_{\rm ff}$. The latter inequality is always satisfied for modes propagating perpendicular to the magnetic field, i.e. $\boldsymbol{k} \cdot \boldsymbol{B} = 0$. If both conditions are satisfied, then $\delta \ln(p_c/\rho^{4/3}) \approx 0$ (from the CR pressure equation) and $\delta \ln(p_g/\rho^{5/3}) \approx 0$ (from the gas entropy equation). In the high-$\beta$ limit, pressure balance and adiabaticity imply that:

$$\delta p_g + \delta p_c = \frac{5}{3} p_g \frac{\delta\rho}{\rho} + \frac{4}{3} p_c \frac{\delta\rho}{\rho} = \boldsymbol{\xi} \cdot \boldsymbol{\nabla} p_g + \boldsymbol{\xi} \cdot \boldsymbol{\nabla} p_c. \tag{2.49}$$

The displaced fluid element will be buoyantly unstable if $\delta\rho < \boldsymbol{\xi} \cdot \boldsymbol{\nabla}\rho$, so the condition for instability is

$$\boldsymbol{\xi} \cdot \boldsymbol{\nabla} p_g + \boldsymbol{\xi} \cdot \boldsymbol{\nabla} p_c < \frac{5}{3} \frac{p_g}{\rho} \boldsymbol{\xi} \cdot \boldsymbol{\nabla}\rho + \frac{4}{3} \frac{p_c}{\rho} \boldsymbol{\xi} \cdot \boldsymbol{\nabla}\rho. \tag{2.50}$$

Using (2.48) and that the background density and CR pressure satisfy $\rho \propto p_c^{3/2}$, this can be rewritten as:

$$\eta\Big(\gamma \frac{\omega_{\rm ff}}{\omega_c} \frac{v_{A,z}}{c_s} - 2\Big) > \frac{3}{2}\gamma, \tag{2.51}$$

where $\gamma = 5/3$ is the gas adiabatic index. We derive the same criterion directly from the linearised equations in Appendix 2.E (also assuming adiabaticity). The above condition turns out to be equivalent to

$$\frac{ds_{\rm eff}}{dz} \propto \frac{d}{dz}\Big(\ln \frac{p_g}{\rho^{5/3}} + \eta \ln \frac{p_c}{\rho^{4/3}}\Big) < 0. \tag{2.52}$$

If the above is satisfied, the system is convectively unstable. Condition (2.51) is shown in Figure 2.5 as the dashed vertical line.

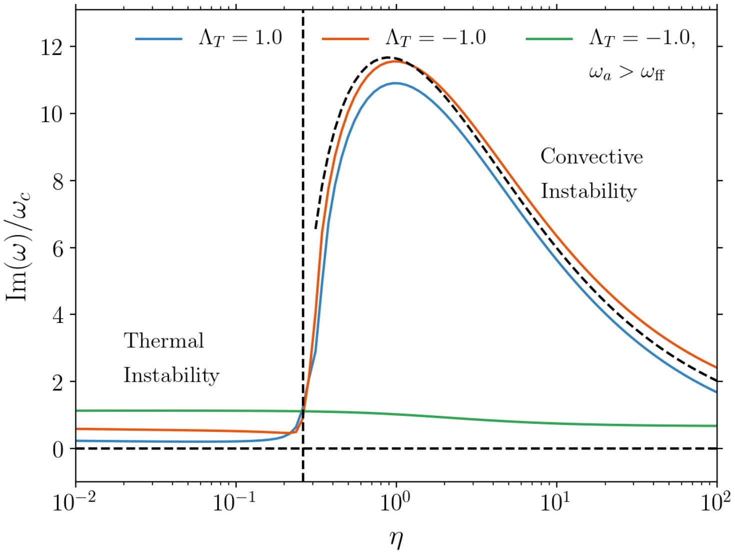


*Figure 2.5*: Thermal and convective instability of gravitationally stratified plasmas with $\omega_{\rm ff} = 20\omega_c, \beta = 10$ and $\omega_a = 0$ (blue and orange lines). Buoyancy-driven instability occurs when $\eta$ satisfies eq. 2.51 (vertical dashed line). The dashed curve shows the approximate growth rate from equation (2.86). At smaller $\eta$ (and $\omega_a < \omega_{\rm ff}$), gravity reduces the thermal instability growth rate by a factor of $\sim$ 2. When $\omega_a > \omega_{\rm ff}$ ($\omega_a = 10^3\omega_c$; green line), perturbations are not adiabatic and we recover the same thermal instability growth rate as without gravity.

Using hydrostatic equilibrium (2.48), we can further rephrase the instability criterion in terms of the CR and gas pressure scale heights ($H_c^{-1} \equiv d\ln p_c/dz$, $H_g^{-1} \equiv d\ln p_g/dz$),

$$\frac{H_c}{H_g} - \eta > \frac{3}{2}\gamma. \tag{2.53}$$

Therefore, a necessary condition for convection is that $H_c/H_g > 5/2$. We show the convective (in)stability in the ($\eta$, $\beta$) plane in Figure 2.6. The system becomes convectively unstable for a larger range of $\eta$ and $\beta$ when $\omega_{\rm ff}\cos\theta_B/\omega_c$ is increased.

We derive an approximate growth rate for the convective instability in the limit $\omega_{\rm ff} \gg \omega_c, \omega_a$ in Appendix 2.E. The approximate growth rate (equation 2.86, which is derived by dropping any dependence on $\Lambda_T$) is shown in Figure 2.5 as the dashed line and agrees well with the exact solution.

## 2.8 Discussion

In this paper, we have studied the linear thermal stability of systems heated by streaming cosmic rays. Streaming cosmic rays can be an important heating mechanism in cluster halos (Guo & Oh 2008; Jacob & Pfrommer 2017a; Jacob & Pfrommer 2017b). Our order-of-magnitude estimates suggest that CR heating may also be important in galactic halos, particularly for Milky Way mass

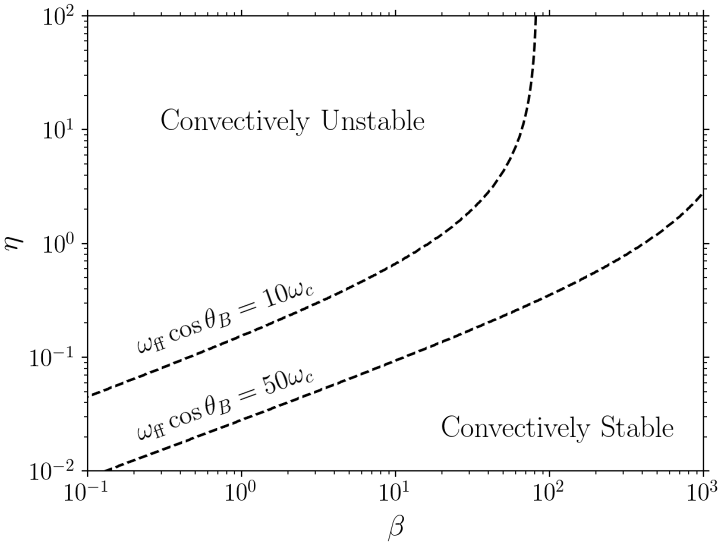


*Figure 2.6*: Convective stability/instability boundary in the ($\beta$, $\eta$) plane for two choices of $\omega_{\rm ff}\cos\theta_B/\omega_c$, where $\theta_B$ is the angle between the z-axis (direction of gravity and pressure gradients) and the magnetic field. For a given choice of $\omega_{\rm ff}\cos\theta_B/\omega_c$, there is a maximum $\beta$ for which convection can occur, as determined by equation (2.51).

systems (see Section 2.4 and Figure 2.1). Thermal instability is a viable mechanism for creating the cold gas that is found in these systems.

We calculated thermal instability growth rates in the presence of CR heating for a wide range of $\eta \equiv p_c/p_g$, in three background equilibria: in a uniform background (where cooling is balanced by an unspecified volumetric heating, but CR heating is present in the perturbed equations; Section 2.5), in a background with CR heating balancing cooling (Section 2.6), and in a gravitationally stratified background heated by CRs (Section 2.7). Gas in PIE in galactic halos is a special case of our uniform background calculation (Section 2.5.8), which is always thermally stable, independent of CR pressure.

The key timescales/frequencies in this problem are summarised in equations 2.9–2.13 and include the cooling, sound, Alfvén (CR heating), CR diffusion and free-fall frequencies. The Alfvén frequency enters the thermal instability calculation primarily via the perturbed CR heating, as the entropy mode is not sensitive to magnetic tension and pressure. We focused on local WKB perturbations, satisfying equations 2.14 and 2.15, in the high-$\beta$ limit. Our fiducial parameter set was $\omega_a = 10^3\omega_c, \beta = 100$, and $\omega_{\rm ff} = 20\omega_c$ when we included gravity (this is well motivated in galactic and cluster halos, where typically $\omega_{\rm ff} \gtrsim 10\omega_c$, but we stress that it need not be true in general astrophysical systems). Our results depend weakly on $\beta$ for $\beta \gtrsim 3$. Thermal instability growth rates also do not depend strongly on wavenumber $k$ for $\omega_a \gtrsim \omega_c$ ($\omega_a \gtrsim \omega_c\eta^{-1}$) in the uniform (CR-heated) background, which corresponds to WKB perturbations satisfying $kH \gtrsim 1$. Figure 2.2 shows this weak dependence for $\omega_a \leq 10^3\omega_c$ (top panels) and $3 \leq \beta \leq 100$ (middle panels).

We have focused on cosmic rays that stream down their pressure gradient at the Alfvén velocity, while also including CR diffusion along the magnetic field. We find that CR diffusion can suppress thermal instability of a subset of modes (bottom panels of Figure 2.2 and Figure 2.3) and modify the overall thermal stability to arbitrary perturbations (Figure 2.4). However, the dominant CR transport process in galactic halos remains unclear and it is uncertain whether CR streaming and diffusion are generally simultaneously relevant. Indeed, a variety of work suggests that if CR scattering is mostly due to self-excited Alfvén waves, then even cosmic rays that are imperfectly coupled to the thermal plasma (where the imperfect coupling is what gives rise to a possibly diffusive behaviour) are not truly diffusive (Skilling 1971; Wiener et al. 2013; Wiener et al. 2018), and instead stream at super-Alvénic speeds. Super-Alfvénic streaming does not, however, increase the rate at which CRs heat the gas (i.e. the CRs still heat the gas at a rate $-\boldsymbol{v_A} \cdot \boldsymbol{\nabla} p_c$). Moreover, how super-Alfvénic streaming speeds depend on other fluid quantities is not well known. As a result, we did not include super-Alfvénic streaming in our calculation.

The different background equilibria we have considered allowed us to disentangle how CR physics affects thermal instability. Independent of background, thermal instability growth rates depend strongly on $\eta$, which determines whether the instability is isobaric (small $\eta$) or isochoric (large $\eta$); see equation 2.27 and associated discussion. The perturbed CR heating also introduces high-frequency oscillations (order $\omega_a$ for $\eta \sim 1$, independent of background) in the gas entropy mode (see Section 2.5.3.1), so that thermal instability is formally an overstability with an oscillation frequency comparable to or larger than its growth rate. This is true even in a uniform medium, where thermal instability is normally a purely growing mode. CR heating does not, however, modify isobaric (small $\eta$) or isochoric (large $\eta$) thermal instability growth rates, $(2/5)\omega_c(2 - \Lambda_T)$ and $-(2/3)\Lambda_T\omega_c$ respectively, in a uniform background (Section 2.5.5 and left panels of Figure 2.2). Background CR heating does slightly change isobaric thermal instability growth rates at small $\eta$ (Sections 2.6.2 and 2.6.3, and right panels of Figure 2.2). Incorporating gravity in our analysis (Section 2.7) did not significantly affect thermal instability growth rates, but it allowed us to determine under what conditions a gravitationally stratified, CR-heated medium is buoyantly unstable.

Thermal instability growth rates as a function of $\eta$ are plotted in Figure 2.2 for different cooling curve slopes $\Lambda_T$ (top panels), plasma-$\beta$ (middle panels) and for different CR-diffusion frequencies $\omega_d$ (bottom panels; there is no CR diffusion present in the top and middle panels). The left panels show the uniform-background (Section 2.5) calculation results, the right panels show the corresponding results in a background in which cosmic-ray heating balances cooling (Section 2.6). As already mentioned in the previous paragraph, thermal instability growth rates depend strongly on $\eta$. Growth rates do not depend strongly on $\beta$ for $\beta \gtrsim 3$, as is expected in galaxy halos, groups and clusters.

Figure 2.4 shows the boundary between thermal stability and instability to arbitrary (WKB and high-$\beta$) perturbations in a uniform background (top) and the CR-heated background (bottom). We formulate this in terms of the critical cooling curve slope $\Lambda_{T,c} = \partial \ln \Lambda / \partial \ln T$ above which all perturbations are thermally stable. When there is no CR diffusion present, the “critical" cooling curve slope $\Lambda_{T,c}$ as a function of $\eta$ is well described by a broken power law. If CR diffusion is present and $\kappa\omega_c/\eta v_A^2 \gg 1$, the stability/instability boundary is simple and independent of $\eta$: CR diffusion renders high-$k$ (large $\omega_d$, $\boldsymbol{k} \cdot \boldsymbol{B} \neq 0$) perturbations isobaric, so that $\Lambda_{T,c} = 2$ in a uniform background (eq. 2.28) and $\Lambda_{T,c} = 5/2$ in the CR-heated background (eq. 2.41). If $\kappa\omega_c/\eta v_A^2 \lesssim 1$, CR diffusion introduces a Field length below which thermal instability is suppressed (eq. 2.32). This affects the thermal stability/instability boundary in Figure 2.4 (e.g., green dashed line for $\kappa = v_A^2/\omega_c$). We stress, however, that unlike thermal conduction, there is only an effective CR Field length for particular CR diffusion coefficients, namely $\kappa \lesssim \eta v_A^2/\omega_c$. For $\kappa \to \infty$ CRs have no effect on thermal instability because the CR pressure is essentially uniform.

Modes with $\omega_{\rm ff} > \omega_{\rm a}$ (i.e. nearly adiabatic modes) can further be convectively unstable (driven by buoyancy). Convective instability occurs if equation 2.51 is satisfied (see also Figure 2.6). By defining $s_{\rm eff} \propto \ln p_g/\rho^{5/3} + \eta \ln p_c/\rho^{4/3}$, the criterion for convective instability can be written in the form $ds_{\rm eff}/dz < 0$. In our setup, this turns out to be satisfied if the ratio of the CR pressure scale height to the gas pressure scale height is sufficiently large (equation 2.53). We also derive an approximate expression for the growth rate of the convective instability for perpendicular modes (see equation 2.86 and the dashed curve in Figure 2.5). Our calculation differs from previous work that considered buoyancy instabilities in the presence of cosmic rays, which did not consider *background* CR heating (Chandran & Dennis 2006; Dennis & Chandran 2009; Heintz & Zweibel 2018). Background CR heating is essential in our calculation, as it is the background gas-pressure gradient, set by hydrostatic equilibrium and $\nabla p_c$ (which is set by cooling), that drives convection.

Our calculations show that systems heated by cosmic rays are likely thermally unstable for temperature ranges relevant to galactic halos ($10^5$K $\lesssim T \lesssim 10^7$K, where $\Lambda_T \lesssim 0$; Draine 2011). In halos that are in PIE, however, $\Lambda_T$ is large ($> 2$) and the gas is thermally stable for any $\eta$. In cluster halos, where the temperature can exceed $\approx 10^7 K$ and thermal Bremsstrahlung is the dominant radiative cooling process (with $\Lambda_T = 0.5$), CR heating could lead to thermal stability if CR pressure dominates (i.e. $\eta \gtrsim 1$) and CR streaming dominates over diffusion. If CR diffusion is important (and $\kappa \gg \eta v_A^2/\omega_c$), however, it eliminates CR pressure perturbations and cooling by Bremsstrahlung would be thermally unstable. Moreover, $\eta \gtrsim 1$ in cluster halos is disfavoured observationally (e.g., Huber et al. 2013). It is thus likely that all halo gas in CIE is thermally unstable in the presence of CR heating.

It remains to be seen how CR heating affects the nonlinear evolution of the thermal instability and the resulting multiphase structure of halo gas. In particular, are there significant differences introduced by the $O(\omega_a)$ entropy oscillations introduced by the CR heating term? This heating frequency can be larger than the free-fall frequency, and it is plausible that this may change the effect of buoyant oscillations on the saturation of thermal instability. However, we note the caveat that long-wavelength modes tend to dominate the nonlinear saturation of thermal instability, for which $\omega_a > \omega_{\rm ff}$ is not necessarily satisfied. A sufficiently small ratio of the cooling time to the free-fall time, $t_{\rm cool}/t_{\rm ff} \lesssim 10$, has been identified as crucial for the development of multiphase gas in hydro simulations (e.g. Sharma et al. 2012). Ji et al. (2018) showed that magnetic fields enhance thermal instability by suppressing buoyant oscillations via magnetic tension. Future simulations will address how entropy oscillations driven by CR heating (which also occur at ~ the Alfvén frequency) affect this evolution and the creation of multiphase gas. In particular, it seems plausible that the dimensionless ratios $t_{\rm cool}/t_{\rm A}$ (with $t_{\rm A} \equiv H/v_A$) and $\eta = p_c/p_g$, which are related to the propagation speed of thermally unstable modes, may be important for the nonlinear evolution of thermal instability.

# Appendix

## 2.A Cosmic-Ray Pressure Fraction

In this section we derive order-of-magnitude estimates of the CR pressure fraction, $\eta$, which we used to create the bottom panel of Figure 2.1. We consider the injection of cosmic rays by Type II Supernovae (Section 2.A.2) and AGNs (Section 2.A.3). This enables us to estimate the (global) cosmic ray energy budget, which is related to a spatially-averaged CR pressure fraction $\eta$. We find that both mechanisms can in principle produce a significant cosmic-ray pressure, with the caveat that we treat the system in a globally-averaged sense.

### 2.A.1 Gas Thermal Energy

We want to compare the total cosmic ray energy to the total thermal energy of the gas within the halo. In what follows, we estimate the thermal energy of the gas filling the galaxy out to the virial radius $R_{\rm vir}$. We define $R_{\rm vir}$ as the radius within which the mean matter density is 200 times the cosmic critical density, i.e. $\langle\rho_m\rangle_{R_{\rm vir}} = 200\rho_c$. We can approximate the total thermal energy within the virial radius as

$$E_{\rm th} \sim \frac{3}{2}\int_0^{R_{\rm vir}} \frac{\rho_g k_B T}{m_H} 4\pi r^2 dr \tag{2.54}$$

where $\rho_g$ is the thermal gas density. We assume an isothermal profile with virialized $k_BT = GM_{200}m_H/(3R_{\rm vir})$:

$$E_{\rm th} \sim 2\pi R_{\rm vir}^3 \frac{\langle\rho_g\rangle_{R_{\rm vir}} k_B T}{m_H} \sim \frac{\langle\rho_g\rangle_{R_{\rm vir}}}{\langle\rho_m\rangle_{R_{\rm vir}}}\frac{GM_{200}^2}{2R_{\rm vir}}, \tag{2.55}$$

where $M_{200} = (4\pi/3)\langle\rho_m\rangle R_{vir}^3$. For $\langle\rho_g\rangle_{R_{\rm vir}} = (x\Omega_B/\Omega_M)\langle\rho_m\rangle_{\rm vir}$, where $x$ accounts for missing baryonic mass relative to the cosmic mean, we obtain

$$E_{\rm th} \sim \frac{x\Omega_B}{\Omega_M}\frac{GM_{200}^2}{2R_{\rm vir}} \sim 3.6\times 10^{58}\ \text{ergs}\left(\frac{x}{1.0}\right)\left(\frac{M_{200}}{10^{12}M_\odot}\right)^{5/3}. \tag{2.56}$$

### 2.A.2 Energy of Cosmic Rays: Injection by Type II SNe

We first consider cosmic rays injected by Type II Supernovae. We assume that there is a core-collapse Supernova for every $100M_\odot$ of stars formed, so that $N_{\rm II} = M_*/100M_\odot$ is the total number of Type II SNe in a galaxy with stellar mass $M_*$. We define $\chi$ as the ratio of the number of cosmic rays still present in the halo (out to the virial radius) to the total number produced throughout the galaxy's lifetime. The total cosmic ray energy in the halo is then

$$E_c = \chi f_{\rm II} E_{\rm II} N_{\rm II}. \tag{2.57}$$

In the above expression, $f_{\rm II} E_{\rm II}$ is the typical CR energy injected by a single type II Supernova ($f_{\rm II}$ is the fraction of supernova energy released as cosmic rays). As is commonly assumed, we take $f_{\rm II} \approx 0.1$ (Zweibel 2017), so that for a typical non-neutrino energy of $10^{51}$ ergs released by Type II Supernovae, $f_{\rm II} E_{\rm II} \approx 10^{50}$ergs. This implies that

$$E_c \sim 10^{59} \text{ ergs} \left(\frac{\chi}{1.0}\right)\left(\frac{M_*}{10^{11} M_\odot}\right), \tag{2.58}$$

where our choice of $\chi = 1.0$ reflects the possibility that in massive galaxies with large $R_{\rm vir}$, the escape time of cosmic rays may be of order the galaxy age. Comparing (2.58) to (2.56) gives

$$\eta \sim \frac{E_c}{2E_{\rm th}} \sim 1.4 \left(\frac{\chi}{1.0}\right) \left(\frac{x}{1.0}\right)^{-1} \left(\frac{M_*}{10^{11} M_\odot}\right)\left(\frac{M_{200}}{10^{12} M_\odot}\right)^{-5/3}. \tag{2.59}$$

### 2.A.3 Energy of Cosmic Rays: Injection by SMBHs

We now consider cosmic rays that are created by SMBHs. We can estimate the CR energy content by assuming that a fraction $f_{\rm BH}$ of the SMBH luminosity goes into cosmic rays, i.e. the SMBH injects cosmic-ray energy at a rate $\epsilon f_{\rm BH} \dot{M} c^2$, where $\epsilon$ is the black hole's radiative efficiency and $\dot{M}$ is its mass accretion rate. This gives:

$$E_c \sim \chi \epsilon f_{\rm BH} M_{\rm BH} c^2, \tag{2.60}$$

where $\chi$ is again defined as the ratio of the number of cosmic rays still present in the halo to the total number produced throughout the galaxy's lifetime. The total CR energy is approximately

$$E_c \sim 18 \times 10^{59} \text{ergs} \left(\frac{\chi}{1.0}\right) \left(\frac{\epsilon f_{\rm BH}}{10^{-3}}\right)\left(\frac{M_{\rm BH}}{10^9 M_\odot}\right). \tag{2.61}$$

Comparing this to the total thermal energy in equation (2.56), we find that

$$\eta \sim \frac{E_c}{2E_{th}} \sim 25 \left(\frac{\chi}{1.0}\right) \left(\frac{\epsilon f_{\rm BH}}{10^{-3}}\right) \left(\frac{x}{1.0}\right)^{-1} \left(\frac{M_{\rm BH}}{10^9 M_\odot}\right)\left(\frac{M_{200}}{10^{12} M_\odot}\right)^{-5/3}. \tag{2.62}$$

## 2.B Cosmic-Ray Diffusion versus Thermal Instability

We now provide a short, heuristic derivation for conditions (2.30) and (2.31). For simplicity, here we consider the case of a uniform background (Section 2.5). An analogous calculation for the CR-heated background gives the very similar conditions (2.45) and (2.46) (see also Figure 2.3). In 2.B.4 we show that CR diffusion can introduce a Field length, below which thermal instability is suppressed.

### 2.B.1 Modes with $\omega_d \ll \omega_c$

In the limit where $\omega_d = 0$ or $\omega_d \ll \omega_c$, diffusion is negligible for any $\eta$ and the $\eta \to 0$, $\eta \to \infty$ limits are connected smoothly at intermediate $\eta$ (as in Figure 2.2).

### 2.B.2 Modes with $\omega_c \ll \omega_d \lesssim \omega_a$

In this limit, (2.24) gives an approximate leading-order relation between $\delta p_c$ and $\delta\rho$:

$$\frac{\delta p_c}{p_g} \sim \eta\frac{\delta\rho}{\rho}\Big(1 + i\frac{\omega_d}{\omega_a}\Big). \tag{2.63}$$

Inserting this approximate relation into equation 2.23 gives

$$\frac{\delta p_g}{p_g}\Big(\frac{\omega}{\gamma - 1} + i\omega_c\Lambda_T\Big) - \frac{\delta\rho}{\rho}\Big(\frac{\gamma\omega}{\gamma - 1} - i\omega_c(2 - \Lambda_T)\Big) \sim \frac{\delta\rho}{\rho}\eta(\omega_a + i\omega_d). \tag{2.64}$$

For $\eta \ll 1$, i.e. $\delta p_g/p_g \approx -\delta p_c/p_g \ll \delta\rho/\rho$, the CR-diffusion term introduced by the perturbed CR heating essentially acts like a thermal-conduction term with a thermal diffusion coefficient $\sim \eta\kappa$. The diffusive term $\propto i\eta\omega_d\delta\rho/\rho$ acts to oppose the perturbed cooling term $\propto i\omega_c(2-\Lambda_T)\delta\rho/\rho$ which drives thermal instability. Diffusion suppresses thermal instability if:

$$\eta\omega_d \gtrsim |2 - \Lambda_T|\, \omega_c, \tag{2.65}$$

which gives the lower bound in (2.30).

This suppression of thermal instability by CR diffusion is not present at large $\eta$, when the second term on the LHS of (2.64), $\propto \delta\rho/\rho$, is negligible (thermal instability is isochoric). Using $\delta p_g \approx -\delta p_c$ and (2.63) in eq. 2.64 one can show that CR diffusion is unimportant when

$$\eta \gtrsim \frac{\omega_d}{\omega_c}|\Lambda_T|^{-1}, \tag{2.66}$$

at which point we recover the $\eta \to \infty$ (isochoric) asymptotic growth rate (eq. 2.39).

### 2.B.3 Modes with $\omega_d \gg \omega_a$

In this limit, (2.24) gives:

$$\frac{\delta p_c}{p_g} \sim \frac{\delta\rho}{\rho}i\eta\frac{\omega_a}{\omega_d}. \tag{2.67}$$

Inserting this into equation 2.23 gives

$$\frac{\delta p_g}{p_g}\Big(\frac{\omega}{\gamma - 1} + i\omega_c\Lambda_T\Big) - \frac{\delta\rho}{\rho}\Big(\frac{\gamma\omega}{\gamma - 1} - i\omega_c(2 - \Lambda_T)\Big) \sim \frac{\delta\rho}{\rho}i\eta\frac{\omega_a^2}{\omega_d}. \tag{2.68}$$

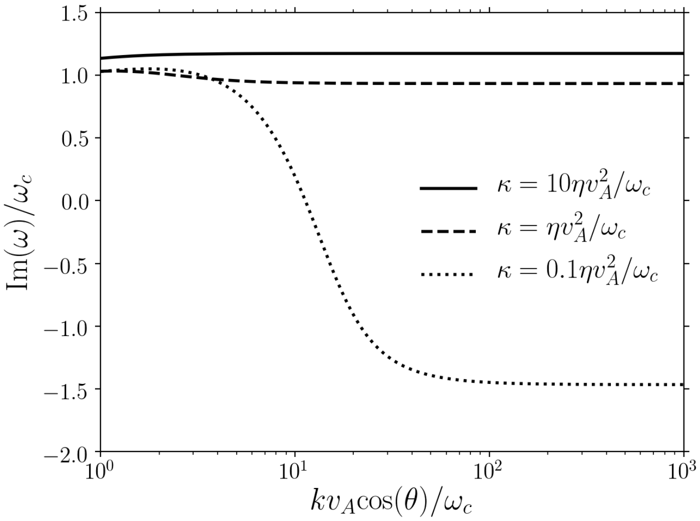


*Figure 2.7*: Thermal instability growth rates as a function of wavenumber $k$ for $\Lambda_T = -1, \beta = 100$ and $\eta = 1$ in a uniform medium. $\theta$ is the angle between $\boldsymbol{k}$ and the background magnetic field. We show growth rates for different CR diffusion coefficients $\kappa$. For $\kappa \gg \eta v_A^2/\omega_c$ (solid line) CR diffusion does not affect thermal instability growth rates at high $k$. For $\kappa \ll \eta v_A^2/\omega_c$ (dotted line) diffusion introduces a CR Field length below which thermal instability is suppressed. In all cases, growth/damping rates are constant at high $k$, as the perturbed CR heating is scale-independent for $\omega_d \gg \omega_a$.

Note that the CR heating term $\propto \eta\omega_a^2/\omega_d = \eta v_A^2/\kappa$ is scale independent. CR diffusion again acts to oppose the perturbed cooling term $\propto i\omega_c(2 - \Lambda_T)\delta\rho/\rho$ which drives thermal instability. For $\delta p_g/p_g \ll \delta\rho/\rho$ (for $\eta \ll \omega_d/\omega_a$) CR diffusion suppresses thermal instability if

$$\eta\frac{\omega_a^2}{\omega_d} \sim \frac{\eta v_A^2}{\kappa} \gtrsim |2 - \Lambda_T|\omega_c. \tag{2.69}$$

This is the lower bound in (2.31). Note that there is no scale dependence. As a result, if $\kappa\omega_c/(\eta v_A^2) \gg 1$ then short-wavelength modes with $\omega_d \gg \omega_a$ are not suppressed by CR diffusion. If, however, $\kappa\omega_c/(\eta v_A^2) \ll 1$, CR diffusion instead leads to the decay of high-$k$ gas-entropy modes (see Figure 2.7).

As described before in 2.B.2, when $\eta$ is sufficiently large for the second term on the LHS to be negligible (thermal instability is isochoric), CR diffusion does not affect the TI growth rate. CR diffusion is unimportant when:

$$\eta \gtrsim \frac{\omega_d}{\omega_c}|\Lambda_T|^{-1}. \tag{2.70}$$

This is the upper bound in (2.31).

### 2.B.4 CR-Diffusion Field Length

We can rephrase conditions 2.30 (eq. 2.65 and eq. 2.66) and 2.31 (eq. 2.69 and eq. 2.70) in terms of length scales at which CR diffusion suppresses thermal instability. In particular, in Appendix 2.B.2 we demonstrate that CR diffusion can play a role similar to thermal conduction. This suggests that there is a CR-diffusion analogue of the Field length for thermal conduction (Field 1965) below which thermal instability is suppressed. Thermal instability of long-wavelength modes with $\hat{\boldsymbol{b}} \cdot \boldsymbol{k} < v_A/\kappa$ ($\omega_d < \omega_a$) is suppressed by CR diffusion if

$$(\hat{\boldsymbol{b}} \cdot \boldsymbol{k})^2 \gtrsim \max\Big(\frac{\omega_c}{\eta\kappa}|2 - \Lambda_T|, \frac{\eta\omega_c}{\kappa}|\Lambda_T|\Big). \tag{2.71}$$

The above is derived from and equivalent to equation 2.30. Thermal instability of short-wavelength modes with rapid CR diffusion, $\hat{\boldsymbol{b}} \cdot \boldsymbol{k} > v_A/\kappa$ ($\omega_d > \omega_a$), is suppressed by CR diffusion if

$$\frac{\kappa\omega_c}{\eta v_A^2}|2 - \Lambda_T| \lesssim 1 \;\text{ and }\; (\hat{\boldsymbol{b}} \cdot \boldsymbol{k})^2 \gtrsim \frac{\eta\omega_c}{\kappa}|\Lambda_T|. \tag{2.72}$$

The above is equivalent to equation 2.31.

Note that if $\kappa \gtrsim \eta v_A^2/\omega_c$ ($\omega_d \gtrsim \eta\omega_a^2/\omega_c$) then CR diffusion does not suppress thermal instability of high-$k$ modes ($\omega_d \gg \omega_a$), even though $\omega_d$ is large. There is therefore no "CR Field length" below which thermal instability is completely suppressed. Instead, the instability of high-$k$ modes is isobaric with growth rates $\mathrm{Im}(\omega) = (2/5)\big(2 - \Lambda_T\big)\omega_c$. Cosmic rays have no effect on thermal instability as the rate at which they heat the gas at high $k$ is less than the cooling rate ($\kappa\omega_c/\eta v_A^2$ is the ratio of the cooling rate to the CR heating rate at high $k$, see eq. 2.68).

Conversely, if $\kappa \lesssim \eta v_A^2/\omega_c$, the cosmic-ray heating rate at high $k$ exceeds the gas cooling rate. CR diffusion then suppresses thermal instability of high-$k$ gas-entropy modes. In other words, when $\kappa \lesssim \eta v_A^2/\omega_c$ there is a maximum $\hat{\boldsymbol{b}} \cdot \boldsymbol{k}$ at which thermal instability occurs. Using (2.71) and (2.72) and dropping order unity $\sim \Lambda_T$ factors one can show that the CR Field length is given by:

$$\lambda_{\mathrm{CRF}} \sim \begin{cases} 2\pi|\hat{\boldsymbol{b}} \cdot \hat{\boldsymbol{k}}|\sqrt{\frac{\eta\kappa}{\omega_c}} & \eta < 1 \\ 2\pi|\hat{\boldsymbol{b}} \cdot \hat{\boldsymbol{k}}|\sqrt{\frac{\kappa}{\eta\omega_c}} & \eta > 1. \end{cases} \tag{2.73}$$

This is the CR-diffusion analogue of the Field length (Field 1965). We stress again that this CR Field length exists only if $\kappa \lesssim \eta v_A^2/\omega_c$. Figure 2.7 shows how the value of $\kappa$ determines the stability/instability of high-$k$ modes.

## 2.C Linearised Equations with Background Cosmic-Ray Heating

When there is background cosmic-ray heating (balancing cooling), the linearised equations are

$$\frac{\delta\rho}{\rho} - \frac{3}{2}\eta^{-1}\omega_c\frac{\xi_z}{v_{A,z}} = -i\boldsymbol{k} \cdot \boldsymbol{\xi}, \tag{2.74}$$

$$-\rho\omega^2\boldsymbol{\xi} = -i\boldsymbol{k}\Big(\delta p_c + \delta p_g + \frac{\boldsymbol{B}\cdot\boldsymbol{\delta B}}{4\pi}\Big) + i\frac{(\boldsymbol{B}\cdot\boldsymbol{k})\boldsymbol{\delta B}}{4\pi} - \omega_{\rm ff} c_s \delta\rho\,\hat{\boldsymbol{z}} \tag{2.75}$$

$$\boldsymbol{\delta B} = i(\boldsymbol{B}\cdot\boldsymbol{k})\boldsymbol{\xi} - i\boldsymbol{B}(\boldsymbol{k}\cdot\boldsymbol{\xi}), \tag{2.76}$$

$$\begin{aligned}\frac{\delta p_g}{p_g}\Big(\frac{\omega}{\gamma-1} + i\omega_c\Lambda_T\Big) + \frac{\omega\omega_c\xi_z}{(\gamma-1)v_{A,z}}\Big(1-\gamma\frac{\omega_{\rm ff}}{\omega_c}\frac{v_{A,z}}{c_s}\Big) = \omega_a\frac{\delta p_c}{p_g}& \\ - \omega_a\omega_c\frac{\xi_z}{v_{A,z}} - i\omega_c\frac{\delta\rho}{\rho}\Big(\frac{5}{2} - \Lambda_T\Big) - \Big(\frac{\gamma}{\gamma-1}\omega + i\omega_c\Big)i\boldsymbol{k}\cdot\boldsymbol{\xi},&\end{aligned} \tag{2.77}$$

$$\begin{aligned}\frac{\delta p_c}{p_g}(\omega - \omega_a + i\omega_d + i\omega_c\eta^{-1}) - \omega\omega_c\frac{\xi_z}{v_{A,z}} = \frac{\delta\rho}{\rho}\Big(-\frac{2}{3}\eta\omega_a + i\omega_c\Big)& \\ - \frac{4}{3}i\eta\omega\boldsymbol{k}\cdot\boldsymbol{\xi} + i\kappa\hat{\boldsymbol{b}}\cdot\boldsymbol{k}(\boldsymbol{\delta}\hat{\boldsymbol{b}}\cdot\boldsymbol{\nabla}p_c) + i\kappa\boldsymbol{k}\cdot\boldsymbol{\delta}\hat{\boldsymbol{b}}(\hat{\boldsymbol{b}}\cdot\boldsymbol{\nabla}p_c),&\end{aligned} \tag{2.78}$$

where $\boldsymbol{\delta}\hat{\boldsymbol{b}} = \boldsymbol{\delta B}/B - \delta B/B\,\hat{\boldsymbol{b}}$. For the calculations in Section 2.6, where we ignore gravity, we set $\omega_{\rm ff} = 0$ in the above equations.

## 2.D Validity of the 1D Thermal Instability Calculation

For a high-$\beta$ uniform medium, the dispersion relation of the thermally unstable entropy mode can be derived simply by imposing pressure balance $\delta p_c \simeq -\delta p_g$ and combining the CR and gas energy equations. This leads directly to equation 2.25. The same approach does not work in the presence of background cosmic-ray heating and/or gravity because then the cosmic-ray and gas energy equations have terms proportional to the fluid displacement $\xi$ (see Appendix 2.C) and so imposing $\delta p_c \simeq -\delta p_g$ is not sufficient to uniquely determine the entropy mode properties. In this Appendix, we discuss the approximations that successfully reproduce the entropy mode in this limit. In particular, we explain why the 1D calculation in Section 2.6.2, which assumes $\boldsymbol{\xi} \parallel \boldsymbol{B}$, is a reasonable approximation for the thermal instability eigenfrequency.

We note from the start that the usual Boussinesq approximation, $\boldsymbol{k}\cdot\boldsymbol{\xi} = 0$, often utilized to impose pressure balanced fluctuations, is not appropriate for this problem. Instead, for $\omega_a \gg \omega_c$, pressure balance, $\delta p_c + \delta p_g \approx 0$ ($\omega \ll kc_s$ implies that $\delta p_c + \delta p_g \ll \delta\rho c_s^2$), simply determines the leading-order gas entropy frequency (eq. 2.26), which satisfies $\omega < \omega_a$ for all $\eta$. For $\eta \lesssim 1$, $\omega \ll \omega_a$ and the induction equation implies that $\delta B/B \ll \delta\rho/\rho$. Moreover, one can show that

$$\frac{\xi_\perp}{\xi_\parallel} = \frac{k_\perp k_\parallel \omega^2}{\omega^2 k_\parallel^2 - \omega_a^2 k^2} \sim \frac{\omega^2}{\omega_a^2}, \tag{2.79}$$

where $\xi_\perp$ and $\xi_\parallel$ are the fluid displacements perpendicular and parallel to the magnetic field, respectively. So, for $\eta \lesssim 1$, where $\omega \ll \omega_a$, we can restrict our analysis to field-aligned perturbations $\boldsymbol{\xi} \parallel \boldsymbol{B}$, for which $\boldsymbol{\delta B} = 0$ (just like in the 1D calculation). It turns out that $\boldsymbol{\xi} \parallel \boldsymbol{B}$ yields the same dispersion relation as the 1D calculation (2.38), independent of propagation direction (for a fixed $\omega_a$).

For $\eta \gg 1$ thermal instability becomes isochoric. As a result, the 1D dispersion relation still gives the correct thermal-instability eigenfrequency, even though $\boldsymbol{\xi} \parallel \boldsymbol{B}$ is not strictly true ($\xi_\parallel$ still exceeds $\xi_\perp$ by a factor of a few). When $\delta p_c/p_g \gg \delta\rho/\rho$, $\boldsymbol{\xi}$ is not important for thermal instability, as it is tied to density perturbations.

The 1D calculation also works well in the limit of strong diffusion, $\omega_d \gg \omega_a$. In this case, the $\omega$ of the gas entropy mode also never exceeds $\omega_a$. This, again, is determined by $\delta p_c + \delta p_g \approx 0$, which to leading order gives the quadratic:

$$\eta\Big(\frac{4}{3}\omega - \frac{2}{3}\omega_a\Big)\Big(\frac{3}{2}\omega + \omega_a\Big) + \frac{5}{2}\omega\Big(\omega - \omega_a + i\omega_d\Big) = 0. \tag{2.80}$$

Small deviations (primarily in the real part) between the 1D thermal instability eigenfrequency and the exact solution occur only when $\eta \sim \omega_c\omega_d/\omega_a$, where thermal instability is most strongly damped by diffusion. The CR-diffusion induced damping rate is $O(\omega_a)$, while $\delta p_c/p_g \sim \delta\rho/\rho$, and so the assumption that $\boldsymbol{\xi} \parallel \boldsymbol{B}$ is only approximately well motivated. We stress, however, that the deviations (which mainly affect the oscillation frequency) occur only in the case where thermal instability is very rapidly damped.

## 2.E Convective Instability: Growth-Rate Derivation from Linearised Equations

In this section we derive an approximate growth rate for the CR convective instability in the limit $\omega_{\rm ff} \gg \omega_c, \omega_a$. The growth rate is not exact, as we drop any dependence that is $O(\omega_c) \ll \omega_{\rm ff}, \omega$. We consider the simplest case of a purely vertical magnetic field and horizontal propagation: $\boldsymbol{B} = B\hat{\boldsymbol{z}}$ and $\boldsymbol{k} = k\hat{\boldsymbol{x}}$, for which $\omega_a = 0$.

In this limit, one can derive the approximate dispersion relation

$$\omega^4 + b\omega^2 + c = 0, \tag{2.81}$$

where

$$b = -\Big(\omega_s^2 + \frac{4}{3\gamma}\eta\omega_s^2 + \frac{3}{2}\eta^{-1}\omega_c\omega_{\rm ff}\frac{c_s}{v_{A,z}}\Big), \tag{2.82}$$

$$c = -\frac{\omega_s^2\omega_c\omega_{\rm ff}}{2\gamma\eta}\frac{c_s}{v_{A,z}}\Big[2\eta\Big(\gamma\frac{\omega_{\rm ff}v_{A,z}}{\omega_c c_s} - 2\Big) - 3\gamma\Big]. \tag{2.83}$$

This has solutions of the form:

$$\omega^2 = \frac{-b \pm \sqrt{b^2 - 4c}}{2} \tag{2.84}$$

which will have an unstable branch if $c < 0$, i.e.

$$2\eta\Big(\gamma\frac{\omega_{\rm ff}}{\omega_c}\frac{v_{A,z}}{c_s} - 2\Big) > 3\gamma, \tag{2.85}$$

which is the same condition as obtained in the main text using the Schwarzschild criterion (equation 2.51). Equations (2.82)–(2.84) can be combined to give an expression for the growth rate. In the common limit $b^2 \gg c$ (large $\omega_s$ limit), the growth rate simplifies to

$$\omega \approx i\sqrt{\frac{c}{b}} \approx i\sqrt{\omega_c \omega_{\rm ff}\frac{c_s}{v_{A,z}}}\left[\frac{\frac{1}{\gamma}\left(\gamma\frac{\omega_{\rm ff} v_{A,z}}{\omega_c c_s} - 2\right) - \frac{3}{2}\eta^{-1}}{1 + 4\eta/(3\gamma)}\right]^{1/2} \tag{2.86}$$

Note that as $\eta \to \infty$, the growth rate goes to 0. The dashed curve in Figure 2.5 shows the approximate growth rate from equation (2.86), which agrees well with the exact calculation.

# Chapter 3

# Sound-Wave Instabilities in Dilute Plasmas with Cosmic Rays: Implications for Cosmic-Ray Confinement and the Perseus X-ray Ripples

An earlier version of this article was previously published as Kempski P., Quataert E., and Squire J., 2020, MNRAS, 493, 5323 with a correction published in Kempski P., Quataert E., and Squire J., 2021, MNRAS, 500, 1231. Some text, figures and equations in this chapter are updated versions of the originally published ones to include the correction.

## 3.1 Abstract

Weakly collisional, magnetised plasmas characterised by anisotropic viscosity and conduction are ubiquitous in galaxies, halos and the intracluster medium (ICM). Cosmic rays (CRs) play an important role in these environments as well, by providing additional pressure and heating to the thermal plasma. We carry out a linear stability analysis of weakly collisional plasmas with cosmic rays using Braginskii MHD for the thermal gas. We assume that the CRs stream at the Alfvén speed, which in a weakly collisional plasma depends on the pressure anisotropy ($\Delta p$) of the thermal plasma. We find that this $\Delta p$-dependence introduces a phase shift between the CR-pressure and gas-density fluctuations. This drives a fast-growing acoustic instability: CRs offset the damping of acoustic waves by anisotropic viscosity and give rise to wave growth when the ratio of CR pressure to gas pressure is $\gtrsim \alpha\beta^{-1/2}$, where $\beta$ is the ratio of thermal to magnetic pressure, and $\alpha$, typically $\sim 1$, depends on other dimensionless parameters. In high-$\beta$ environments like the ICM, this condition is satisfied for small CR pressures. We speculate that the instability studied here may contribute to the scattering of high-energy CRs and to the excitation of sound waves in galaxy-halo, group and cluster plasmas, including the long-wavelength X-ray fluctuations in *Chandra* observations of the Perseus cluster. It may also be important in the vicinity of shocks

in dilute plasmas (e.g., cluster virial shocks or galactic wind termination shocks), where the CR pressure is locally enhanced.

## 3.2 Introduction

The interstellar medium (ISM), the intracluster medium (ICM), and the halos of galaxy groups and Milky-Way-like galaxies are filled with hot and dilute gas, in which the electron/ion mean free paths along the magnetic field greatly exceed the particle gyroradii. Under such conditions, transport of heat and momentum is anisotropic and happens preferentially in the direction of the local magnetic field. The particle mean free path in these tenuous plasma environments can be large (i.e. the plasma is weakly collisional). As a result, anisotropic transport is efficient and can significantly affect the thermal and dynamical evolution of the gas.

The importance of anisotropic conduction and viscosity in cluster environments has been underpinned by a variety of analytic theory and simulations. Anisotropic transport is an efficient driver of buoyancy instabilities (Balbus 2000; Quataert 2008; Kunz et al. 2012) and significantly affects the gas dynamics in cluster simulations (e.g., Ruszkowski & Oh 2010; Parrish et al. 2012; Yang & Reynolds 2016; Barnes et al. 2019; Kingsland et al. 2019).

The ISM, galaxy halos, groups and the ICM are also permeated by a non-thermal population of relativistic particles known as cosmic rays (CRs). Even though they essentially propagate at the speed of light, their lifetime in galactic discs and halos can be quite long due to scattering off electromagnetic fluctuations. The waves responsible for the scattering can be Alfvén waves generated by the cosmic rays themselves through the streaming instability (Kulsrud & Pearce 1969). Pitch-angle scattering by the excited waves isotropises the cosmic rays in the frame of the Alfvén waves. In this so-called self-confinement picture, cosmic rays are scattered towards isotropy in the Alfvén frame and collectively drift down their pressure gradient at the Alfvén speed, provided that the pitch-angle scattering is sufficiently rapid. For slower pitch-angle scattering rates, the CR transport deviates from pure streaming at the Alfvén speed, but its exact nature remains uncertain (cosmic rays are believed to either diffuse or stream at super-Alfvénic speeds, or both; Skilling 1971, Wiener et al. 2013; Amato & Blasi 2018). The self-confinement picture is in contrast to the extrinsic turbulence picture, where CRs are scattered primarily by extrinsic fluctuations that are not excited by the particles themselves. In this case, cosmic rays generally do not stream at Alfvénic speeds, even in the limit of fast scattering. In this work, we focus on self-confined cosmic rays.

The additional pressure force ($-\boldsymbol{\nabla} p_c$) and gas heating ($-\mathbf{v_A}\cdot\boldsymbol{\nabla} p_c$; Wentzel 1971) provided by the cosmic rays can be important for the dynamics and thermal evolution of gas in galaxies, halos and clusters (e.g., Breitschwerdt et al. 1991; Loewenstein et al. 1991; Everett et al. 2008; Socrates et al. 2008; Guo & Oh 2008; Zweibel 2013; Ruszkowski et al. 2017; Zweibel 2017; Jacob & Pfrommer 2017a; Jacob & Pfrommer 2017b; Ehlert et al. 2018; Farber et al. 2018; Kempski & Quataert 2020). Cosmic rays can also directly affect MHD waves. For example, Begelman & Zweibel (1994), hereafter BZ94, showed that CR heating can drive an acoustic instability in low-$\beta$ plasmas ($\beta \lesssim 1$).

The purpose of this work is to study sound waves in the presence of cosmic rays in magnetised, weakly collisional plasmas, i.e. plasmas with large anisotropic viscosity and conduction. We use the Braginskii MHD closure for weakly collisional plasmas (Braginskii 1965) with anisotropic conduction and anisotropic pressure (the latter acts as an anisotropic viscosity), coupled to a 1-moment fluid equation for the cosmic-ray pressure. The cosmic rays are assumed to stream at the Alfvén speed $\mathrm{v}_{\mathrm{A},\Delta p}$, which in a weakly collisional plasma depends on the pressure anisotropy of the thermal plasma, $\Delta p$. We find that this dependence of $\mathrm{v}_{\mathrm{A},\Delta p}$ on $\Delta p$, which is not present in standard high-collisionality MHD, gives rise to a rapidly growing acoustic instability (i.e. instability of the fast magnetosonic mode). The instability is driven by a phase shift between the CR pressure and gas density. Unlike the acoustic instability in BZ94, the Cosmic Ray Acoustic Braginskii (CRAB) instability that we find here is not driven by CR heating and does not require low $\beta$. In fact, the CRAB instability exists even at small CR pressures and has faster growth rates in high-$\beta$ systems. It is thus likely important in the ICM, in galactic halos and in the hot ISM.

The remainder of this work is organised as follows. We present the CR–gas equations, examine the validity of our model and introduce characteristic timescales in Section 3.3. We describe the CR-driven acoustic instability in Section 3.4 and consider possible astrophysical implications in Section 3.5. In Section 3.5.1 we speculate on the potential connection between the acoustic instability and the X-ray surface-brightness fluctuations observed in galaxy clusters such as Perseus (Fabian et al. 2003). In 3.5.2 we hypothesise that the instability is likely important close to shocks, including the vicinity of the virial radius. We discuss the potential contribution of the long-wavelength waves generated by the instability to the scattering of high-energy cosmic rays in Section 3.5.3. We summarise our results in Section 3.6.

## 3.3 Equations

We model the dilute, weakly-collisional plasmas filled with cosmic rays by using the Braginskii MHD equations coupled to a cosmic-ray pressure,

$$\frac{\partial \rho}{\partial t} + \boldsymbol{\nabla} \cdot (\rho \mathbf{v}) = 0 \tag{3.1}$$

$$\rho \frac{d\mathbf{v}}{dt} = -\boldsymbol{\nabla}\Big(p_\perp + p_c + \frac{B^2}{8\pi}\Big) + \frac{\boldsymbol{B} \cdot \boldsymbol{\nabla} \boldsymbol{B}}{4\pi} + \boldsymbol{\nabla} \cdot (\hat{\boldsymbol{b}}\hat{\boldsymbol{b}}\Delta p) \tag{3.2}$$

$$\frac{\partial \boldsymbol{B}}{\partial t} = \boldsymbol{\nabla} \times (\mathbf{v} \times \boldsymbol{B}) \tag{3.3}$$

$$\rho T \frac{ds}{dt} = -\mathbf{v}_{\mathbf{A},\boldsymbol{\Delta p}} \cdot \boldsymbol{\nabla} p_c + \mathcal{H} - \mathcal{C} - \boldsymbol{\nabla} \cdot \Big(\boldsymbol{\Pi} \cdot \mathbf{v}\Big) - \boldsymbol{\nabla} \cdot \boldsymbol{Q} \tag{3.4}$$

$$\frac{dp_c}{dt} = -\frac{4}{3} p_c \boldsymbol{\nabla} \cdot (\mathbf{v} + \mathbf{v}_{\mathbf{A},\boldsymbol{\Delta p}}) - \mathbf{v}_{\mathbf{A},\boldsymbol{\Delta p}} \cdot \boldsymbol{\nabla} p_c + \boldsymbol{\nabla} \cdot (\kappa \hat{\boldsymbol{b}}\hat{\boldsymbol{b}} \cdot \boldsymbol{\nabla} p_c), \tag{3.5}$$

where $\mathbf{v}$ is the gas velocity, $\rho$ is the gas density, $p_g$ and $p_c$ are the gas and CR pressures respectively, $\boldsymbol{B}$ is the magnetic field (with unit vector $\hat{\boldsymbol{b}}$), and $s = k_{\mathrm{B}} \ln(p_g/\rho^\gamma)/(\gamma - 1)m_{\mathrm{H}}$ is the gas entropy per unit mass. $d/dt \equiv \partial/\partial t + \mathbf{v} \cdot \boldsymbol{\nabla}$ denotes a total (Lagrangian) time derivative. $\mathcal{H}$ and $\mathcal{C}$ are arbitrary volumetric heating and cooling rates. The Braginskii MHD pressure anisotropy (with viscosity $\nu_{\rm B}$) is

$$\Delta p = p_\perp - p_\parallel = 3\rho\nu_{\rm B}\big(\hat{\boldsymbol{b}}\hat{\boldsymbol{b}} : \nabla \mathbf{v} - \frac{1}{3}\nabla \cdot \mathbf{v}\big) = 3\rho\nu_{\rm B}\frac{d}{dt}\ln\frac{B}{\rho^{2/3}}, \tag{3.6}$$

where $\perp$ and $||$ denote the directions perpendicular and parallel to the magnetic field (Braginskii 1965). $p_\perp$ and $p_\parallel$ are related to the total thermal pressure by

$$p_\perp = p_g + \frac{1}{3}\Delta p. \tag{3.7}$$

The viscous stress tensor in the gas-entropy equation depends on the pressure anisotropy and is given by

$$\mathbf{\Pi} = -\Delta p\Big(\hat{\boldsymbol{b}}\hat{\boldsymbol{b}} - \frac{\boldsymbol{\mathcal{I}}}{\mathbf{3}}\Big). \tag{3.8}$$

Note that in the absence of background flow, the perturbed $\nabla \cdot (\mathbf{\Pi} \cdot \mathbf{v})$ in the gas-entropy equation is second-order and does not contribute in our linear analysis. $\boldsymbol{Q}$ in equation 3.4 is the anisotropic thermal heat flux,

$$\boldsymbol{Q} = -\kappa_{\rm B}\hat{\boldsymbol{b}}\hat{\boldsymbol{b}} \cdot \nabla T, \tag{3.9}$$

where $\kappa_{\rm B}$ is the thermal conductivity.[1] $\mathrm{v}_{\mathrm{A},\Delta p}$ in equations 3.4 and 3.5 is the Alfvén speed in the presence of pressure anisotropy,

$$\mathbf{v}_{\mathbf{A},\boldsymbol{\Delta p}} = \frac{\boldsymbol{B}}{\sqrt{4\pi\rho}}\Big(1 + \frac{4\pi\Delta p}{B^2}\Big)^{1/2}. \tag{3.10}$$

We assume that cosmic rays stream down their pressure gradient at the Alfvén velocity $\mathbf{v}_{\mathbf{A},\boldsymbol{\Delta p}}$ and we also include CR diffusion along the magnetic field, for which we assume a constant diffusion coefficient $\kappa$. We note that formally CRs stream with velocity $\mathbf{v}_{\mathbf{st}} = -\mathrm{sgn}(\hat{\boldsymbol{b}} \cdot \nabla p_c)\mathbf{v}_{\mathbf{A},\boldsymbol{\Delta p}}$. This ensures that cosmic rays stream along the magnetic field down their pressure gradient and makes the CR heating term $-\mathbf{v}_{\mathbf{st}} \cdot \nabla p_c$ positive definite. In our linear stability analysis cosmic rays stream at $\mathbf{v}_{\mathbf{A},\boldsymbol{\Delta p}}$, as we consider background equilibria which satisfy $-\mathbf{v}_{\mathbf{A},\boldsymbol{\Delta p}} \cdot \nabla p_c > 0$.[2]

---

[1] While in this work we assume that the heat transport is diffusive, we note that recent particle-in-cell simulations suggest that the transport may also occur down the temperature gradient at the whistler phase speed (Roberg-Clark et al. 2018).

[2] We do not explicitly include background gradients in our linear stability calculation, and so we treat the background as effectively uniform. However, for $\nabla p_c$ to have a well-defined sign in the linear stability calculation, so that $-\mathbf{v}_{\mathbf{A},\boldsymbol{\Delta p}} \cdot \nabla p_c$ is positive definite, $p_c$ cannot be exactly uniform. A background CR pressure gradient is necessary. However, we can neglect terms associated with background gradients as long as the additional timescale introduced by a spatially varying background is much longer than the timescale associated with the acoustic instability considered here. This is well-motivated given the fast growth rates of the instability, which can be comparable to the sound oscillation frequency. We can, for example, consider an equilibrium with $-\mathbf{v}_{\mathbf{A},\boldsymbol{\Delta p}} \cdot \nabla p_c = C$, where $C$ is a cooling rate with an associated cooling frequency $\omega_c$. Our linear stability calculation can neglect background gradients provided that $\omega_s, \mathrm{Im}(\omega) \gg \omega_c$. If this is satisfied, the equlibrium $\rho$, $p_g$ and $p_c$ can be treated as uniform without significantly changing the results.

### 3.3.1 MHD Waves in Weakly Collisional Plasmas

In this section we ignore cosmic rays and review how standard MHD waves are modified at low collisionality. The pressure anisotropy changes the Alfvén speed (eq. 3.10) because it modifies the effective magnetic tension, as can be seen by rewriting eq. 3.2:

$$\rho\frac{d\mathbf{v}}{dt} = -\boldsymbol{\nabla}(p_\perp + p_c + \frac{B^2}{8\pi}) + \boldsymbol{\nabla}\cdot\Big(\frac{\boldsymbol{B}\boldsymbol{B}}{4\pi}(1+\frac{4\pi\Delta p}{B^2})\Big). \tag{3.11}$$

Note that the factor $(1 + 4\pi\Delta p/B^2)$ enters the effective magnetic tension term (which is responsible for Alfvén waves). The dispersion relation for Alfvén waves can then be easily derived by assuming wave perturbations $\propto \exp\big(i\boldsymbol{k}\cdot\boldsymbol{r} - i\omega t\big)$, crossing the momentum equation twice with $\boldsymbol{k}$ and noting that $\delta\Delta p = 0$ for Alfvénic perturbations (which are incompressible and do not change the $B$-field strength, see eq. 3.6). From this, equation 3.10 follows.

The pressure anisotropy has a different effect on the slow and fast modes, which are viscously damped in Braginskii MHD (still ignoring cosmic rays). By inserting eq. 3.6 into eq. 3.2 (and noting eq. 3.7), we obtain:

$$\rho\frac{d\mathbf{v}}{dt} = ... + \boldsymbol{\nabla}\cdot\Big(3\rho\nu_{\mathrm{B}}(\hat{\boldsymbol{b}}\hat{\boldsymbol{b}} - \frac{\mathcal{I}}{3})(\hat{\boldsymbol{b}}\hat{\boldsymbol{b}} : \boldsymbol{\nabla}\mathbf{v} - \frac{1}{3}\boldsymbol{\nabla}\cdot\mathbf{v})\Big). \tag{3.12}$$

This diffusion operator associated with the Braginskii viscosity damps the fast and slow magnetosonic waves, because they involve perturbations that linearly generate $\delta\Delta p$, unlike the linearly undamped Alfvén waves for which $\delta\Delta p = 0$. In the weak damping limit, the fast and slow modes are damped at a rate (Braginskii 1965; Parrish et al. 2012):

$$\omega_\nu = \frac{\nu_{\mathrm{B}}k^2}{6}\Big((\hat{\boldsymbol{k}}\cdot\hat{\mathbf{v}}) - 3(\hat{\boldsymbol{b}}\cdot\hat{\boldsymbol{k}})(\hat{\boldsymbol{b}}\cdot\hat{\mathbf{v}})\Big)^2, \tag{3.13}$$

where $\hat{\mathbf{v}}$ is the unit vector in the direction of the mode's perturbed velocity. We will show that in the presence of cosmic rays this is strongly modified, and sound waves can instead be linearly unstable.

### 3.3.2 Validity of the Model

Our CR–Braginskii MHD fluid model requires that the CR scattering rate is fast, so that eq. 3.5 appropriately describes the CR pressure evolution. It also requires that the collision time of the thermal ions is short compared to the macroscopic timescales of interest (so that a weakly collisional, rather than collisionless, treatment is appropriate for the thermal plasma). In what follows, we check the validity of the CR–Braginskii MHD fluid model, focusing on the ICM and the hot phase of the ISM.

For the CR pressure equation (eq. 3.5) to be a good model of the cosmic rays, the GeV CR collision frequency must be large. It is the GeV CRs that are important, as they dominate the bulk

CR energy. The CR collision frequency is the rate at which the pitch angle changes by order unity, due to scattering by EM fluctuations of magnitude $\delta B_\perp$ at the resonant wavelength:

$$\nu_{\rm CR} \sim \Omega\Big(\frac{\delta B_\perp}{B}\Big)^2 \sim 10^{-8}\ {\rm s}^{-1}\left(\frac{\gamma_c}{1}\right)^{-1}\frac{B}{1\ \mu{\rm G}}\Big(\frac{\delta B_\perp/B}{10^{-3}}\Big)^2, \tag{3.14}$$

where $\gamma_c$ is the CR Lorentz factor and $\delta B_\perp$ is evaluated for fluctuations whose wavelength parallel to the mean B-field is of order the Larmor radius of the GeV particles. Models of CR observations in the Milky Way based on pure diffusion infer a CR diffusion coefficient $\kappa \sim 10^{28} - 10^{29}$ cm$^2$ s$^{-1}$ depending on assumptions about the CR halo size (e.g., Linden et al. 2010). This motivates the choice of $\delta B_\perp/B \sim 10^{-3}$ used in (3.14), as $\delta B_\perp/B \sim 10^{-3}$ corresponds to a GeV CR diffusion coefficient $\kappa \sim c^2/\nu_{\rm CR} \sim 10^{29}$ cm$^2$ s$^{-1}$ in a 1 $\mu$G field. However, this observationally inferred CR diffusion coefficient is not necessarily appropriate if CR streaming is the dominant transport process, as is theoretically favoured for the low-energy CRs that dominate the total energy density (these low-energy cosmic rays are the most likely to be adequately described by the fluid model used in this paper; Blasi et al. 2012). In the case of streaming transport the diffusion coefficient may be $\ll 10^{28} - 10^{29}$ cm$^2$ s$^{-1}$. In particular, in the hot ISM and ICM damping processes are weaker than in the cold/neutral ISM and so the streaming instability can grow to large amplitudes (e.g., Figure 1 in Amato & Blasi 2018). For example, if $\delta B_\perp/B \sim 10^{-2}$ then $\nu_{\rm CR} \sim 10^{-6}$ s$^{-1}$ and the CR diffusion coefficient is significantly smaller, $\kappa \sim 10^{27}$ cm$^2$ s$^{-1}$.

The weakly collisional fluid model used in this paper requires that the ion-ion collision frequency is larger than the rate of change of all fields, $\omega \lesssim \nu_{\rm ii}$. We note that $\omega \ll \nu_{\rm ii}$ is formally required in deriving the equations of Braginskii MHD. We will consider $\omega_s \lesssim \nu_{\rm ii}$ in our calculations (where $\omega_s$ is the adiabatic sound frequency and typically the largest frequency in the problem), but our main conclusions do not change if we choose a smaller upper limit on $\omega_s$. We now separately estimate the ion collision rates in the ICM and the hot ISM. We will show that the CR scattering rate is much higher than the ion-ion collision frequency in both the ICM and hot ISM.

#### 3.3.2.1 ICM

Under typical ICM conditions, the plasma is magnetised and the collisionality is low. For representative ICM temperatures and densities, the ion-ion collision frequency is

$$\nu_{\rm ii} \sim \frac{n_{\rm i} e^4 \pi \ln\Lambda}{m_{\rm i}^{1/2}(k_{\rm B}T)^{3/2}} \sim 8\times10^{-14}\ {\rm s}^{-1}\left(\frac{T}{5\times10^7\ {\rm K}}\right)^{-3/2}\frac{n_{\rm i}}{0.01\ {\rm cm}^{-3}}, \tag{3.15}$$

for a Coulomb logarithm $\ln\Lambda \approx 38$. This corresponds to a collision time of approximately 0.4 Myrs. We note that $\nu_{\rm CR} \gg \nu_{\rm ii}$ (see eq. 3.14) and so the CR–Braginskii MHD fluid model is a good description for the ICM if we consider fields that vary at a frequency $\omega \ll \nu_{\rm ii}$ (see Section 3.3.3 for how this translates into constraints on the characteristic frequencies in our problem).

#### 3.3.2.2 Hot ISM

The plasma filling the hot ISM is cooler, so that the ion-ion collision frequency is larger than in the ICM:

$$\nu_{\rm ii} \sim 3\times10^{-11}\ {\rm s}^{-1}\left(\frac{T}{10^6\ {\rm K}}\right)^{-3/2}\frac{n_{\rm i}}{0.01\ {\rm cm}^{-3}} \tag{3.16}$$

(for $\ln\Lambda \approx 32$). This ion-ion collision frequency is still, however, significantly smaller than the CR collision frequency (order unity pitch angle change; see eq. 3.14). Just like in the ICM, the CR–Braginskii MHD formulation is therefore well motivated in the hot ISM as long as we consider $\omega \ll \nu_{\rm ii}$.

### 3.3.3 Dimensionless Parameters and Characteristic Timescales

We define the ratio of CR pressure to gas pressure,

$$\eta \equiv \frac{p_c}{p_g}, \tag{3.17}$$

and the ratio of thermal to magnetic pressure,

$$\beta \equiv \frac{8\pi p_g}{B^2}. \tag{3.18}$$

The key frequencies in this problem are the gas sound frequency (with $c_s$ being the adiabatic gas sound speed),

$$\omega_s \equiv kc_s; \tag{3.19}$$

the Alfvén *and* CR-heating frequency,

$$\omega_a \equiv \boldsymbol{k}\cdot\mathbf{v_A}; \tag{3.20}$$

the cosmic-ray diffusion frequency,

$$\omega_d \equiv \kappa\,(\hat{\boldsymbol{b}}\cdot\boldsymbol{k})^2; \tag{3.21}$$

the Braginskii viscous frequency,

$$\omega_{\rm B} \equiv \nu_{\rm B}(\hat{\boldsymbol{b}}\cdot\boldsymbol{k})^2 \approx \frac{p_g}{3\rho\nu_{\rm ii}}(\hat{\boldsymbol{b}}\cdot\boldsymbol{k})^2; \tag{3.22}$$

and the conductive frequency

$$\omega_{\rm cond} \equiv \chi_{\rm B}(\hat{\boldsymbol{b}}\cdot\boldsymbol{k})^2, \tag{3.23}$$

where $\chi_{\rm B} = \kappa_{\rm B}/nk_{\rm B}$ is the thermal diffusion coefficient. We define the Braginskii viscous scale,

$$l_{\nu_{\rm B}} \equiv \frac{\nu_{\rm B}}{c_s} \sim l_{\rm mfp}, \tag{3.24}$$

where $l_{\rm mfp}$ in the last step is the ion mean free path. We can relate the diffusive timescales by defining the thermal Prandtl number,

$$\mathrm{Pr} \equiv \frac{\nu_{\rm B}}{\chi_{\rm B}}, \tag{3.25}$$

and the ratio of the CR diffusion coefficient to the Braginskii viscosity,

$$\Phi \equiv \frac{\kappa}{\nu_{\rm B}}. \tag{3.26}$$

It is commonly assumed that the heat flow is dominated by electrons, such that for a typical plasma $\mathrm{Pr} \sim 10^{-2}$ (set by the ion-to-electron mass ratio). This assumption is, however, not well motivated when the timescales of interest are shorter than the ion-electron temperature equilibration time (which is longer than the ion-ion collision time by a square root of the ion to electron mass ratio). This is the case in this work, where we consider sound waves at low collisionalities. A more accurate calculation should therefore consist of two entropy equations and two heat fluxes, one for each species. We avoid this complication in the main text of this paper by considering a single heat flux with varying conductivity: $\omega_{\rm cond} = \omega_{\rm B}$ ($\mathrm{Pr} = 1$, $\sim$ heat flux carried by ions) and $\omega_{\rm cond} = 100\omega_{\rm B}$ ($\mathrm{Pr} = 0.01$, $\sim$ heat flux carried by electrons). We show in Appendix 3.A and Figure 3.7 that our conclusions do not change when a two-fluid electron-ion system is considered instead, and that $\mathrm{Pr} \sim 1$ is a somewhat better approximation to the two-fluid results (a similar two-fluid electron-ion system was used in the context of cluster sound waves by Zweibel et al. 2018).

$\Phi$ in eq. 3.26 relates the Braginskii viscous frequency to the CR diffusion frequency,

$$\Phi = \frac{\omega_d}{\omega_{\rm B}} \sim \frac{c^2}{c_s^2}\frac{\nu_{\rm ii}}{\Omega(\delta B_\perp/B)^2}, \tag{3.27}$$

where $c$ is the speed of light. For typical ICM parameters,

$$\Phi \sim 2\left(\frac{T}{5\times10^7\ \mathrm{K}}\right)^{-5/2}\frac{n_{\rm i}}{0.01\ \mathrm{cm}^{-3}}\left(\frac{B}{1\ \mu\mathrm{G}}\right)^{-1}\left(\frac{\delta B_\perp/B}{10^{-3}}\right)^{-2}. \tag{3.28}$$

This suggests that $\Phi \sim 1$ in the ICM (or $\Phi \ll 1$, if $\delta B_\perp/B \gg 10^{-3}$). $\Phi \gg 1$ for typical temperatures in the hot ISM, unless $\delta B_\perp/B \gg 10^{-3}$ (which is plausible, see discussion in Section 3.3.2). Motivated by these results, we will focus primarily on $\Phi = 0$ ($\omega_d = 0$), $\Phi = 1$ ($\omega_d = \omega_{\rm B}$) and $\Phi = 10$ ($\omega_d = 10\omega_{\rm B}$).

$\omega_s$ is the largest characteristic frequency in the $\beta > 1$ plasmas that we focus on. We require that $\omega_s \lesssim \nu_{\rm ii}$ so that the weakly collisional description is appropriate (see Section 3.3.2), which translates into

$$\frac{\omega_s}{\nu_{\rm ii}} \sim \frac{\omega_{\rm B}}{\omega_s} \lesssim 1. \tag{3.29}$$

The ICM is of primary interest in this work and so we will focus mainly on the high-$\beta$ limit ($\beta \sim 100$ unless specified otherwise).

### 3.3.4 Linearised Equations

We consider a uniform and static background equilibrium with $\mathcal{H} = C$, i.e. all background fluid variables are assumed to be spatially constant. Thus, there are no background gradients in the linear stability analysis (see the comment regarding the CR pressure gradient in Footnote 2). Without loss of generality, we consider a vertical magnetic field, $\boldsymbol{B} = B\hat{\boldsymbol{z}}$.

We carry out a linear stability calculation of the CR–gas equations (see Section 3.3). All perturbed quantities are assumed to vary as $\delta X(\boldsymbol{r}, t) \propto \exp\left(i\boldsymbol{k}\cdot\boldsymbol{r} - i\omega t\right)$. Without loss of generality, we take $\boldsymbol{k}$ in the $xz$-plane, $\boldsymbol{k} = k\sin\theta\hat{\boldsymbol{x}} + k\cos\theta\hat{\boldsymbol{z}}$. Alfvén waves can be isolated as described in Section 3.3.1, which remains valid in the presence of CRs. The remaining modes can be found by considering all linearised equations excluding the $y$-component of the momentum and induction equations:

$$\omega\frac{\delta\rho}{\rho} = \boldsymbol{k}\cdot\mathbf{v}, \tag{3.30}$$

$$\omega \mathrm{v}_x = k_x\frac{c_s^2}{\gamma}\frac{\delta p_g}{p_g} - \omega_a \mathrm{v}_\mathrm{A}\frac{\delta B_x}{B} + k_x \mathrm{v}_\mathrm{A}^2\frac{\delta B_z}{B} + \frac{2}{3}i\frac{k_x}{k_z}\omega_\mathrm{B}\mathrm{v}_z - \frac{1}{3}i\frac{k_x^2}{k_z^2}\omega_\mathrm{B}\mathrm{v}_x + \eta k_x\frac{c_s^2}{\gamma}\frac{\delta p_c}{p_c}, \tag{3.31}$$

$$\omega \mathrm{v}_z = k_z\frac{c_s^2}{\gamma}\frac{\delta p_g}{p_g} - \omega_a \mathrm{v}_\mathrm{A}\frac{\delta B_z}{B} + k_z \mathrm{v}_\mathrm{A}^2\frac{\delta B_z}{B} - \frac{4}{3}i\omega_\mathrm{B}\mathrm{v}_z + \frac{2}{3}i\frac{k_x}{k_z}\omega_\mathrm{B}\mathrm{v}_x + \eta k_z\frac{c_s^2}{\gamma}\frac{\delta p_c}{p_c}, \tag{3.32}$$

$$\omega\frac{\delta B_x}{B} = -k_z\mathrm{v}_x, \tag{3.33}$$

$$\omega\frac{\delta B_z}{B} = k_x\mathrm{v}_x, \tag{3.34}$$

$$\omega\frac{\delta p_g}{p_g} = \gamma\boldsymbol{k}\cdot\mathbf{v} - i(\gamma-1)\omega_\mathrm{cond}\left(\frac{\delta p_g}{p_g} - \frac{\delta\rho}{\rho}\right) + \eta(\gamma-1)\omega_a\frac{\delta p_c}{p_c}, \tag{3.35}$$

$$\omega\frac{\delta p_c}{p_c} = \frac{4}{3}\boldsymbol{k}\cdot\mathbf{v} - \frac{2}{3}\omega_a\frac{\delta\rho}{\rho} + \frac{4}{3}i\frac{\omega_\mathrm{B}}{\omega_a}k_z\mathrm{v}_z - \frac{2}{3}i\frac{\omega_\mathrm{B}}{\omega_a}k_x\mathrm{v}_x + (\omega_a - i\omega_d)\frac{\delta p_c}{p_c}, \tag{3.36}$$

where $\gamma = 5/3$ is the gas adiabatic index. We find the exact eigenmodes by solving the full matrix eigenvalue problem using MATLAB.

## 3.4 The Cosmic-Ray Acoustic Instability in Braginskii MHD

Before we show growth rates and simplified dispersion relations, we discuss the physical mechanism that drives the sound-wave instability.

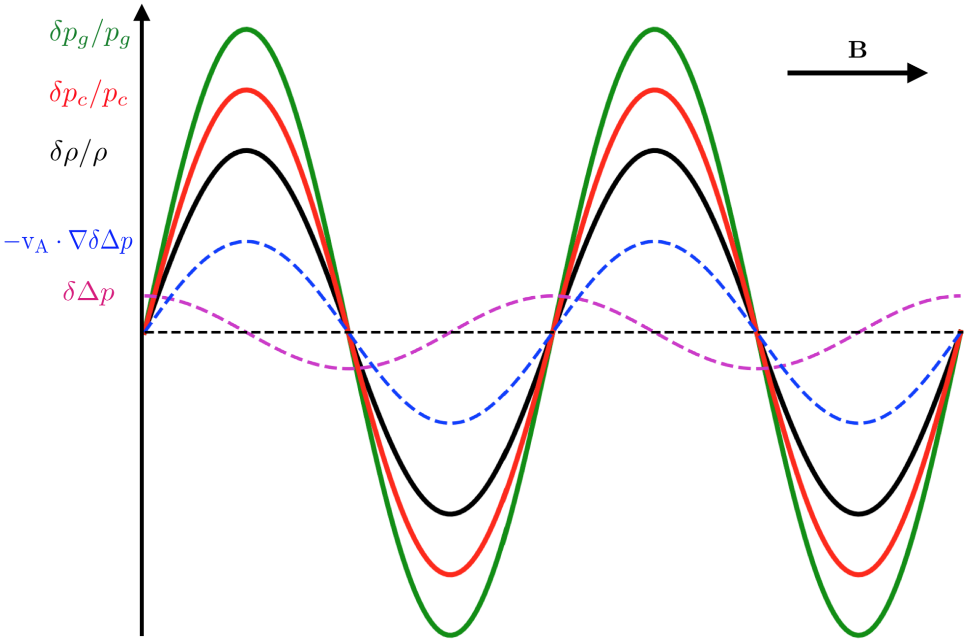


*Figure 3.1*: Schematic of the mechanism driving the acoustic instability. The solid waveforms show the leading-order adiabatic gas-density, gas-pressure and CR-pressure perturbations in the frame comoving with the sound wave at the phase speed $\mathrm{v_{ph}}$ in the $\boldsymbol{B}$-direction. The pressure-anisotropy perturbation $\delta\Delta p \sim \nu_B \mathrm{v_{ph}} d\delta\rho/dz$ (eq. 3.6 in the moving frame; dashed magenta line) has a 90° phase shift relative to $\delta\rho/\rho$, $\delta p_c/p_c$ and $\delta p_g/p_g$. Without CRs this phase shift leads to the well-known damping of acoustic waves by anisotropic viscosity. In the presence of cosmic rays, the work done by $\delta\Delta p$ on the CRs (dashed blue line and eq. 3.37) is positive in regions where $\delta p_c > 0$: it therefore amplifies $\delta p_c$ in the regions where $\delta p_c > 0$ and reduces $\delta p_c$ where $\delta p_c < 0$. This drives the perturbations and leads to wave growth.

### 3.4.1 Driving Mechanism and Negative Effective Viscosity from Cosmic Rays

The instability is driven by a phase shift between the CR-pressure and the gas-density perturbations, which comes from the dependence of the Alfvén speed on $\Delta p$ (see eq. 3.39 in Section 3.4.2). Such phase shifts generally occur in the presence of diffusion operators (e.g., CR diffusion also leads to a phase shift between $\delta p_c$ and $\delta \rho$). However, these tend to damp the perturbations instead of driving instabilities. The phase shift introduced by $\mathrm{v}_{\mathrm{A},\Delta p}$ gives rise to an instability because it introduces an additional diffusion operator in the momentum equation (eq. 3.2) which can have a negative diffusivity (negative viscosity) and thus generate wave growth.[3]

For standard, collisional MHD sound waves, the CR pressure responds essentially adiabatically to density fluctuations in the limit $\omega_a,\ \omega_d \ll \omega_s$ (otherwise the CR response is generally non-adiabatic, see eq. 3.36 with $\omega_{\mathrm{B}} = 0$). In weakly collisional plasmas the CR pressure also responds to changes in the pressure anisotropy, which in turn depends on the rate of change of $\delta\rho$. This phase shift (in addition to the adiabatic response) provides a driving force to the wave, which can win over the damping by anisotropic viscosity and give rise to instability.

The key term for driving the instability is the compression work done on the cosmic rays by the pressure anisotropy,

$$\frac{d\delta p_c}{dt} = -\frac{2p_c}{3\rho \mathrm{v}_{\mathrm{A}}^2}\mathbf{v}_{\mathbf{A}} \cdot \boldsymbol{\nabla}\delta\Delta p + \dots , \tag{3.37}$$

which comes from the $\boldsymbol{\nabla} \cdot \mathbf{v}_{\mathbf{A},\boldsymbol{\Delta p}}$ term in equation 3.5. To see what this term does to the sound wave, it is useful to consider the frame comoving with the wave in the $\boldsymbol{B}$-direction. In this frame, moving at a phase speed $v_{\mathrm{ph}}$, the wave profile is stationary to leading order (i.e. ignoring the growth or damping of the wave) and is shown in Figure 3.1. $\delta\Delta p \sim \nu_{\mathrm{B}}\mathrm{v}_{\mathrm{ph}} d\delta\rho/dz$ (eq. 3.6 in the moving frame) has a 90° phase shift relative to $\delta\rho/\rho$, $\delta p_c/p_c$ and $\delta p_g/p_g$, and without cosmic rays this phase shift leads to wave damping. However, the work done by $\Delta p$ on the CRs (eq. 3.37) is positive in regions where $\delta p_c > 0$, as shown in Figure 3.1. This process amplifies $\delta p_c$ in the regions where $\delta p_c > 0$ and reduces $\delta p_c$ where $\delta p_c < 0$. This drives the perturbations and leads to wave growth.

This driving manifests itself mathematically as a negative effective diffusion coefficient (i.e. negative viscosity) introduced by the cosmic rays in the momentum equation. This can be demonstrated by inserting equation 3.5 into equation 3.2 and assuming wave perturbations proportional to $f(\boldsymbol{k} \cdot \boldsymbol{r} - \omega t)$ propagating at the sound speed. Ignoring all other non-diffusive terms in the momentum equation, this gives:

$$\begin{aligned}\rho\frac{d\mathbf{v}}{dt} = \dots &+ \boldsymbol{\nabla} \cdot \Big(3\rho\nu_{\mathrm{B}}(\hat{\boldsymbol{b}}\hat{\boldsymbol{b}} - \frac{\mathcal{I}}{3})(\hat{\boldsymbol{b}}\hat{\boldsymbol{b}} : \boldsymbol{\nabla}\mathbf{v} - \frac{1}{3}\boldsymbol{\nabla}\cdot\mathbf{v})\Big)\\ &- \frac{2\eta\sqrt{\beta}}{3\sqrt{2\gamma}}(\hat{\boldsymbol{b}}\cdot\hat{\boldsymbol{k}})\boldsymbol{\nabla}\Big(3\rho\nu_{\mathrm{B}}(\hat{\boldsymbol{b}}\hat{\boldsymbol{b}} : \boldsymbol{\nabla}\mathbf{v} - \frac{1}{3}\boldsymbol{\nabla}\cdot\mathbf{v})\Big).\end{aligned} \tag{3.38}$$

The first term is the damping by Braginskii viscosity, the second term is the additional diffusive term that comes from $\delta p_c$. Longitudinal acoustic waves approximately satisfy $\mathbf{v} \parallel \boldsymbol{k}$. If $\hat{\boldsymbol{b}} \cdot \hat{\boldsymbol{k}} > 0$,

---

[3] A negative diffusion coefficient can be thought of as standard diffusion reversed in time.

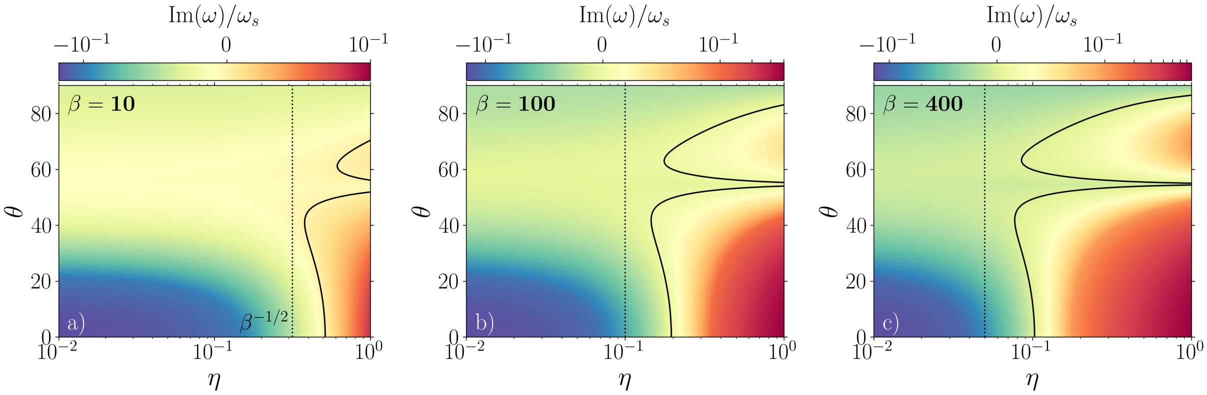


*Figure 3.2*: Growth rates of the CR-driven acoustic instability in the ($\eta = p_c/p_g, \theta$) plane for $\beta = 8\pi p_g/B^2 = 10, 100, 400$, Pr = 1, $\Phi = 0$ and $\nu_{\rm B} k^2 = 0.2\omega_s$ ($kl_{\nu_{\rm B}} = 0.2 \sim kl_{\rm mfp}$, where $l_{\rm mfp}$ is the ion mean free path; see eq. 3.24). Im($\omega$) > 0 corresponds to wave growth and the contour lines show the boundary between damping by anisotropic viscosity and the growth driven by the CRs. Marginal stability occurs first for oblique modes, but otherwise parallel propagating modes are fastest-growing. The dotted vertical lines show $\eta = \beta^{-1/2}$, which is the approximate instability-threshold scaling at high $\beta$ (e.g., eq. 3.43). Even for $\eta$ well above the instability threshold (shown by the contour line), there is a ridge of stability around $\theta = 55°$. For $\theta \lesssim 55°$, the mode with Re($\omega$) $\approx \omega_s$ is unstable, while for $\theta \gtrsim 55°$, the counterpropagating mode with Re($\omega$) $\approx -\omega_s$ is unstable (see Section 3.4.1). All colormaps in this work have log-linear scales that are linear between −0.1 and 0.1.

the last term acts as a diffusion operator with negative viscosity if $\hat{\boldsymbol{b}}\hat{\boldsymbol{b}} : \boldsymbol{\nabla}\mathbf{v} - \frac{1}{3}\boldsymbol{\nabla}\cdot\mathbf{v} \gtrsim 0$, i.e. $\cos^2\theta \gtrsim 1/3$ ($\theta \lesssim 55°$), where $\theta$ is the angle between $\boldsymbol{k}$ and $\boldsymbol{B}$. For $\cos^2\theta \lesssim 1/3$ ($\theta \gtrsim 55°$), it acts as a diffusion operator with negative viscosity for longitudinal acoustic modes propagating in the opposite direction, $\hat{\boldsymbol{b}}\cdot\hat{\boldsymbol{k}} < 0$.[4]

The transition at $\theta \approx 55°$ is clearly present in Figure 3.2, where we show growth rates of the sound-wave instability in the ($\eta, \theta$) plane, for $\nu_{\rm B} k^2 = 0.2\omega_s$ ($kl_{\nu_{\rm B}} = 0.2$) and $\beta = 10, 100, 400$. Even for $\eta$ well above the instability threshold (shown by the contour line), there is a ridge of stability around $\theta = 55°$. For $\theta \lesssim 55°$, the mode with Re($\omega$) $\approx \omega_s$ is unstable, while for $\theta \gtrsim 55°$, the counterpropagating mode with Re($\omega$) $\approx -\omega_s$ is unstable.

### 3.4.2 1D Dispersion Relation

Because sound waves are primarily longitudinal, it is instructive and also physically well motivated to look at the instability in the 1-dimensional case. This also turns out to be sufficient to predict the approximate growth rate of the fastest growing mode in most cases, as fastest growth typically occurs for propagation parallel to $\boldsymbol{B}$. As we will show, this is not true for $\eta$ just above

---

[4] Note that the magnetic-field direction in $\hat{\boldsymbol{b}}\cdot\hat{\boldsymbol{k}}$ in eq. 3.38 comes from the direction of CR streaming (which occurs in the $\hat{\boldsymbol{b}}$-direction due to our assumption that the background CR pressure decreases in the direction of the magnetic field).

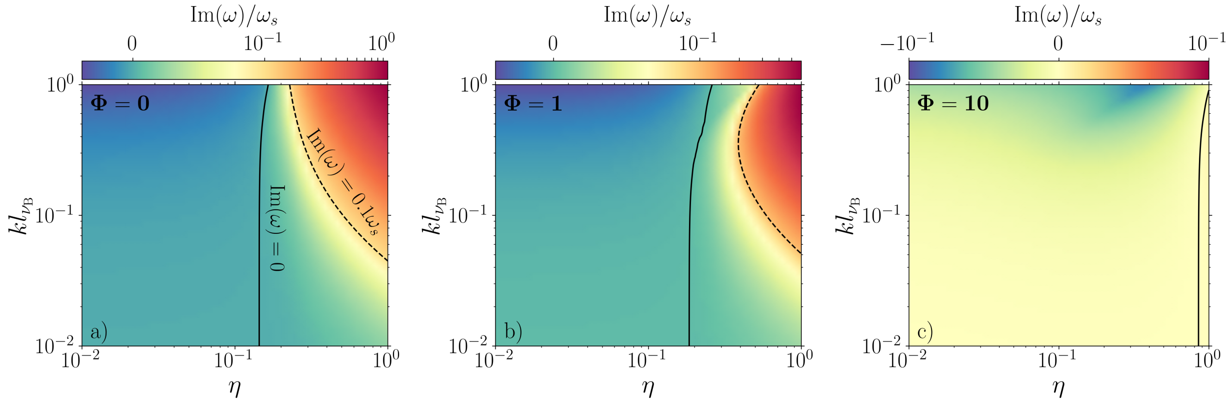


*Figure 3.3*: Wavenumber dependence of growth rates of the acoustic instability as a function of $\eta = p_c/p_g$ ($\beta = 100$ and $\mathrm{Pr} = 1$ in all panels). On the y-axis, $l_{\nu_B} \equiv \nu_B/c_s \sim l_{\mathrm{mfp}}$, where $l_{\mathrm{mfp}}$ is the ion mean free path (i.e. the y-axis can be written as $kl_{\nu_B} = \nu_B k^2/\omega_s$). We consider different CR diffusion coefficients (see eq. 3.26 for the definition of $\Phi$). The solid line corresponds to $\mathrm{Im}(\omega) = 0$, the dashed line is $\mathrm{Im}(\omega) = 0.1\omega_s$. At each $k$, we plot the fastest growth rate (i.e. fastest growing mode across all directions of propagation, $\theta$). See Sections 3.4.2.1, 3.4.2.2 and 3.4.2.3 for more discussion of how $\Phi$ affects the growth rates.

marginal stability, where fastest growth can occur at finite $\theta$, and when the CR diffusion coefficient is large.

For simplicity, we consider sound waves in the high-$\beta$ limit, such that $\omega \sim \omega_s \gg \omega_a$. Equation 3.36 then simplifies to

$$\frac{\delta p_c}{p_c} = \frac{\delta\rho}{\rho}\left(\frac{4}{3} + \frac{4}{3}i\frac{\omega_B}{\omega_a}\right)\left(1 + i\frac{\omega_d}{\omega}\right)^{-1}. \tag{3.39}$$

The phase shift between the CR pressure and gas density introduced by $\Delta p$ (the first bracket multiplying $\delta\rho/\rho$) is what destabilises the wave. In contrast, the phase shift introduced by CR diffusion (second term in the second bracket) acts as a damping.

In the high-$\beta$ limit ($\omega_s \gg \omega_a$), the 1D dispersion relation for sound waves is given by,

$$\begin{aligned} 0 = \omega^2 - \frac{\omega_s^2}{\gamma}\frac{\gamma\omega + i(\gamma-1)\omega_{\mathrm{cond}}}{\omega + i(\gamma-1)\omega_{\mathrm{cond}}} + \frac{4}{3}i\omega_B\omega \\ - \eta\frac{\omega_s^2}{\gamma}\left(\frac{4}{3} + \frac{4}{3}i\frac{\omega_B}{\omega_a}\right)\left(1 + i\frac{\omega_d}{\omega}\right)^{-1}. \end{aligned} \tag{3.40}$$

The second term represents the standard sound-wave frequency in the presence of anisotropic conduction (adiabatic without conduction, isothermal in the limit of rapid conduction), the third term is the damping by anisotropic viscosity and the fourth term is the additional pressure response that comes from the cosmic rays, which can be destabilising.

We first consider equation 3.40 without CR diffusion, i.e. $\omega_d = 0$. We then look at the impact of CR diffusion in Section 3.4.2.3.

### 3.4.2.1 Nearly Isothermal Sound Waves

In the limit of rapid conduction, $\omega_{\rm cond} \gg \omega_s$ (heat conduction carried by electrons and equilibrated with the ions, i.e. Pr $\ll 1$), the dispersion relation is (in the absence of CR diffusion)

$$\omega^2 - \frac{\omega_s^2}{\gamma} + \frac{4}{3} i \omega_{\rm B} \omega - \eta \frac{\omega_s^2}{\gamma}\left(\frac{4}{3} + \frac{4}{3} i \frac{\omega_{\rm B}}{\omega_a}\right) = 0. \tag{3.41}$$

Driving by $\delta p_c$ (Section 3.4.1) wins over damping by anisotropic viscosity when

$$\eta \frac{4\omega_s^2 \omega_{\rm B}}{3\gamma\omega_a} \gtrsim \frac{4}{3}\omega_{\rm B}\omega \approx \frac{4\omega_{\rm B}\omega_s}{3\sqrt{\gamma}}, \tag{3.42}$$

where we ignored $O(\eta)$ corrections to the sound speed due to the cosmic rays. The condition for instability can be written in terms of $\eta$ and $\beta$ as (in 1D):

$$\eta \gtrsim \beta^{-1/2} \qquad \text{(nearly isothermal).} \tag{3.43}$$

Note that the instability threshold is independent of $\omega_{\rm B}$, as $\omega_{\rm B}$ is the characteristic frequency of both anisotropic viscous damping and the driving by $\delta p_c$. We will show that the instability threshold is generally at slightly lower $\eta$ if oblique propagation is included.

### 3.4.2.2 Nearly Adiabatic Sound Waves

If the thermal Prandtl number is not set by electron conduction and we instead have Pr $\sim 1$, the appropriate limit to consider is $\omega_s \gg \omega_{\rm B} \sim \omega_{\rm cond}$. The dispersion relation is then approximately given by

$$\omega^3 - \omega_s^2\omega + i\omega_s^2\frac{(\gamma-1)^2}{\gamma}\omega_{\rm cond} + \frac{4}{3} i \omega_{\rm B}\omega^2 - \eta\frac{\omega_s^2\omega}{\gamma}\left(\frac{4}{3} + \frac{4}{3} i \frac{\omega_{\rm B}}{\omega_a}\right) \approx 0. \tag{3.44}$$

Now the driving from the $\delta p_c$ response has to compete against damping by both anisotropic conduction and viscosity (third and fourth terms, respectively). For $\omega_{\rm cond} \sim \omega_{\rm B}$ (Pr $\sim 1$), however, the correction to the instability threshold is at most order unity, and $\eta \gtrsim \beta^{-1/2}$ (eq. 3.43) is still the approximate instability condition.

Figure 3.2 shows that this $\eta \gtrsim \beta^{-1/2}$ scaling for instability works well for a wide range of $\beta$ ($\eta = \beta^{-1/2}$ is shown by the dotted vertical lines). The plots are for $\nu_{\rm B} k^2 = 0.2\omega_s$ and Pr $= 1$, i.e. $\omega_s \gg \omega_{\rm B}, \omega_{\rm cond}$. The contours show the transition from damping by Braginskii viscosity to growth driven by the cosmic rays. Marginal stability occurs first for oblique modes, but otherwise parallel propagating modes are fastest-growing.

Figure 3.3a shows how growth rates depend on wavenumber $k$, for Pr $= 1$, $\beta = 100$ and $\Phi = 0$ (no CR diffusion). The solid line corresponds to $\mathrm{Im}(\omega) = 0$, the dashed line is $\mathrm{Im}(\omega) = 0.1\omega_s$. At each $k$, we plot the fastest growth rate (i.e. fastest growing mode across all directions of propagation, $\theta$). The instability threshold is nearly independent of $k$, as the damping rates by conduction and viscosity are comparable at all $k$, so that $\eta \gtrsim \beta^{-1/2}$ is sufficient for instability across the entire range in $k$ (equation 3.43 and discussion in the paragraph following eq. 3.44). However, the growth rates generally increase with increasing $k$.

#### 3.4.2.3 Effect of CR Diffusion

In the limit where CR diffusion is slow compared to the sound frequency, $\omega_d \ll \omega_s$ (this corresponds to $\Phi \ll (kl_{\rm mfp})^{-1}$), the CR term driving the instability in equation 3.40 is mildly reduced (compared to the $\omega_d = 0$ case):

$$\eta\frac{\omega_s^2}{\gamma}\frac{4}{3}i\frac{\omega_{\rm B}}{\omega_a} \to \eta\frac{\omega_s^2}{\gamma}\Big(\frac{4}{3}i\frac{\omega_{\rm B}}{\omega_a} - \frac{4}{3}i\frac{\omega_d}{\omega}\Big). \tag{3.45}$$

CR diffusion acts to oppose the $\Delta p$ perturbations in the CR pressure equation that drive the instability, and as a result shifts the instability threshold to larger $\eta$ (compared to, e.g., eq. 3.43). This shift is small, however, if $\omega_d/\omega_s \ll \omega_B/\omega_a$, i.e. $\Phi \ll \sqrt{\beta}$ (as well as $\Phi \ll (kl_{\rm mfp})^{-1}$, i.e. the weak diffusion limit).

In the limit $\omega_d \gg \omega \sim \omega_s$, the CR term in eq. 3.40 is:

$$\eta\frac{\omega_s^2}{\gamma}\Big(\frac{4}{3} + \frac{4}{3}i\frac{\omega_{\rm B}}{\omega_a}\Big)\Big(1 + i\frac{\omega_d}{\omega}\Big)^{-1} \approx \eta\Big( -\frac{4}{3}i\frac{\omega}{\omega_d} + \frac{4}{3}\frac{\omega_{\rm B}}{\omega_a}\frac{\omega}{\omega_d}\Big). \tag{3.46}$$

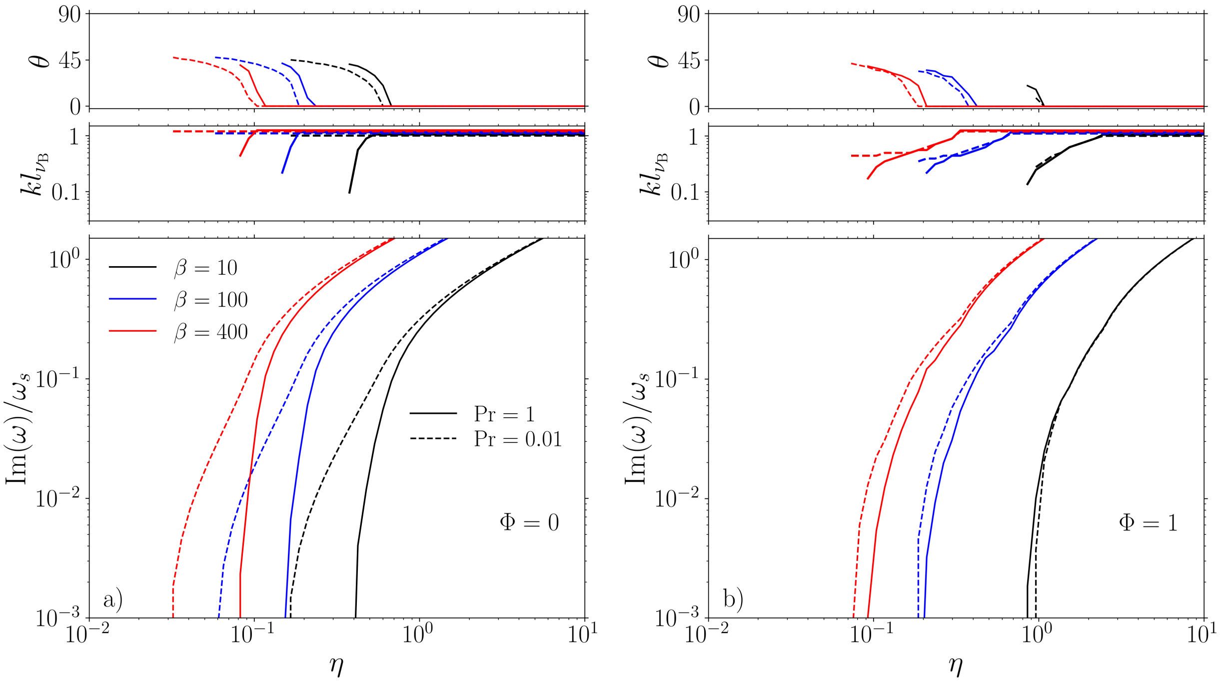
The driving by CR pressure is completely shut off as $\delta p_c$ is suppressed by diffusion. In the 1D case considered here, for $\omega_d \gg \omega_s$ instability can only occur if $\omega \gg \omega_s$, i.e. the sound speed is much larger than the thermal adiabatic sound speed. This occurs at $\eta \gg 1$, when the CRs set the sound speed (the CR sound speed is $\sqrt{4p_c/3\rho}$ ).

Note, however, that even if $\omega_d \gg \omega_s$ for parallel propagation, $\omega_d$ will be less than $\omega_s$ at the same $k$ for $\theta$ close to 90 degrees. As a result, for $\eta \lesssim 1$ and large CR diffusion coefficients, $\Phi \gg 1$, the short-wavelength perturbations with $\kappa k^2 \gg kc_s$ can still be unstable for oblique propagation (this can, e.g., be seen in Figure 3.4c).

The effects of CR diffusion as a function of wavenumber $k$ and CR pressure fraction $\eta$ are shown in Figure 3.3b and Figure 3.3c. As before, the solid line corresponds to $\mathrm{Im}(\omega) = 0$, the dashed line is $\mathrm{Im}(\omega) = 0.1\omega_s$. At each $k$, we plot the fastest growth rate across all propagation angles. All parameters are the same as in Figure 3.3a, except for $\Phi$, which now is $\Phi = 1$ in 3.3b and $\Phi = 10$ in 3.3c. The $\Phi = 1$ growth rates are quite similar to $\Phi = 0$ (no diffusion). Noticeable differences occur primarily at high $k$, so that the overall instability threshold is not significantly changed. When CR diffusion is strong ($\Phi = 10$), significantly larger $\eta$ is required for instability.

### 3.4.3 Stability versus Instability & Maximum Growth Rate

We show the fastest growing mode as a function of $\eta$, restricting to modes with $kl_{\nu_{\rm B}} \leq 1$, in Figure 3.4. We select the mode with the maximum $\mathrm{Im}(\omega)$, but in the plots we normalise its growth rate using the $\omega_s$ at the $k$ at which the maximum growth occurs. At each $\eta$, we also show the $\theta$ and $k$ at which the fastest growth occurs. The top panel is for $\Phi = 0$ (no CR diffusion), the middle panel is for $\Phi = 1$ ($\omega_d = \omega_{\rm B}$) and the bottom panel is for $\Phi = 10$ ($\omega_d = 10\omega_{\rm B}$). We see that the minimum CR pressure fraction ($\eta$) required for instability is lowest for small thermal Prandtl numbers and no CR diffusion (see also Figure 3.3). While the instability threshold is not significantly modified when $\Phi = 1$, it occurs at significantly larger $\eta$ in the limit of strong CR diffusion, $\Phi = 10$.



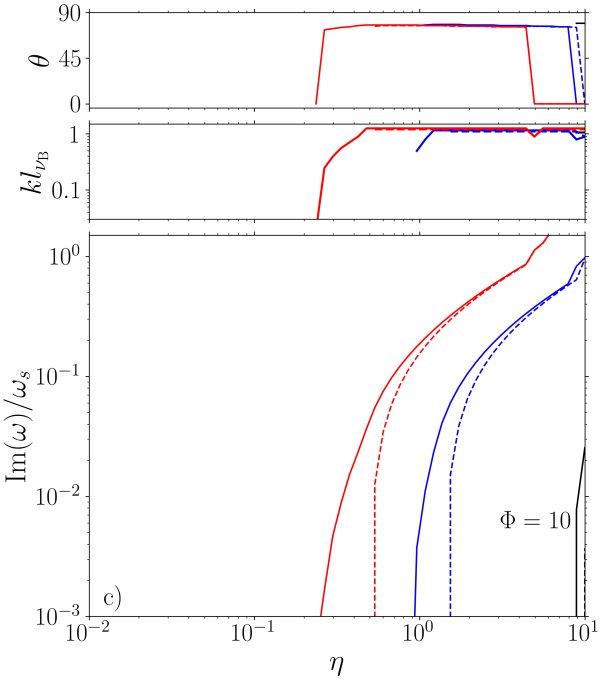


*Figure 3.4*: Fastest growing mode of the acoustic instability as a function of $\eta$. We consider wavelengths that satisfy $kl_{\nu_{\rm B}} \leq 1$, i.e. $\nu_{\rm B}k^2 \leq \omega_s$ ($kl_{\rm mfp} \lesssim 1$). We select the mode with the maximum $\mathrm{Im}(\omega)$, but in the plots we normalise its growth rate using the adiabatic sound frequency at the $k$ where the maximum growth occurs. Panel a) is for $\Phi = 0$ (no CR diffusion), panel b) is for $\Phi = 1$ ($\omega_d = \omega_{\rm B}$) and Panel c) is for $\Phi = 10$ ($\omega_d = 10\omega_{\rm B}$). In each of the three panels, we also show the wavenumber $k$ and direction of propagation $\theta$ of the fastest growing mode (at $kl_{\nu_{\rm B}} = 1$ the lines are slightly displaced for visualisation purposes). Instability occurs for smaller $\eta$ when the thermal Prandtl number Pr is small and when there is no CR diffusion. When CR diffusion is strong ($\Phi = 10$) significantly larger $\eta$ are required for instability. See Section 3.4.3 for more discussion.

In the absence of CR diffusion (Figure 3.4a), fastest growth occurs at the highest $k$ and $\theta = 0$, except when $\eta$ is just above marginal stability. CR diffusion often shifts the fastest growing mode to lower $k$ (middle and bottom panels). However, even in the presence of CR diffusion, when $\eta$ is sufficiently above threshold, fastest growth again occurs at the highest $k$ and $\theta = 0$.

We conclude by stressing that for the wide range of parameters (Pr, $\Phi$, $\beta$) considered here, the instability and fast growth rates $\sim O(\omega_s)$ occur even for small $\eta$ in high-$\beta$ evironments like the ICM. We also note that while we have focused on the simple case of a background equilibrium with $\Delta p = 0$, the acoustic instability will not be significantly affected by a finite background $\Delta p$ as long as the timescale over which the background $\Delta p$ changes is slow compared to the growth rate of the instability.[5]

### 3.4.4 Short Wavelengths and the Collisionless Limit

For the acoustic instability considered in this work, the Braginskii MHD model of the thermal plasma is only valid for timescales longer than the ion-ion collision time, i.e. wavelengths longer than the ion mean free path. To examine the acoustic instability below the mean-free-path scale (but at scales sufficiently large for the cosmic rays to be coupled to the gas), a collisionless description of the thermal plasma is necessary. Preliminary calculations using the CGL and Landau-fluid closures of the kinetic MHD equations (Chew et al. 1956; Snyder et al. 1997) suggest that the instability still exists below mean-free-path scales and has growth rates that are faster than in the weakly collisional limit. The mechanism driving the instability is somewhat different from the weakly-collisional regime illustrated in Figure 3.1: the predominant driver in the collisionless limit is that at high $\beta$, the pressure anisotropy can turn cosmic rays into a fluid with ~negative effective adiabatic index, thus rendering sound waves unstable.[6]

We note, however, that the collisionless description of the thermal plasma coupled to a CR-pressure equation (eq. 3.5) is itself valid only at sufficiently large scales. It breaks down on small scales below the CR mean free path, where the CRs are no longer coupled to the thermal plasma

---

[5] When the background $\Delta p$ is spatially varying, there will be an extra timescale, $\tau$, in our problem. However, as long as $\mathrm{Im}(\omega)\tau \gg 1$, which is reasonable for short-wavelength sound waves, the instability will not be significantly affected by the background $\Delta p$. The effect of a spatially constant $\Delta p$ is to modify the effective magnetic-tension and CR-heating terms, i.e. terms that are $O(\omega_a)$ and negligible for our acoustic instability at high $\beta$.

[6] We consider the 1D case in which $\boldsymbol{\delta B} = 0$ and for simplicity ignore the effect of heat fluxes on the pressure anisotropy (i.e. CGL closure). In the collisionless limit $\Delta p$ approximately satisfies,

$$\frac{1}{p_g}\frac{d\Delta p}{dt} \sim -\frac{1}{\rho}\frac{d\rho}{dt}, \tag{3.47}$$

so that $\delta\Delta p/p_g \sim -\delta\rho/\rho$. In contrast to the weakly collisional case, the relative phase shift between $\delta\Delta p$ and $\delta\rho$ is $\pi$ instead of $\pi/2$ (Figure 3.1). Assuming $\omega \approx \omega_s \approx \sqrt{\beta}\omega_a$ and $\beta \gg 1$, $\delta p_c$ and $\delta\rho$ then roughly satisfy:

$$\delta p_c/p_c \sim -\sqrt{\beta}\delta\rho/\rho. \tag{3.48}$$

Cosmic rays thus behave like a fluid with large negative ($\sim -\sqrt{\beta}$) adiabatic index. This can destabilise the sound wave.

(i.e. the CR scattering rate is no longer the fastest timescale in the problem). We defer a more detailed treatment of the collisionless limit to future work.

### 3.4.5 Relation to BZ94 Acoustic Instability

The CRAB instability is very different from the low-$\beta$ acoustic instability driven by CR heating found in BZ94. BZ94 considered high-collisionality MHD, not the Braginskii MHD limit we have focused on. Moreover, the CRAB instability is not driven by CR heating, but by the work done on the cosmic rays by the pressure anisotropy of the thermal plasma (and is more unstable at high $\beta$).

Nevertheless, at low $\beta$ ($\beta < 1$) we do also find the BZ94 acoustic instability, albeit diminished by the damping by anisotropic viscosity and conduction (most strongly at short wavelengths). In addition to the BZ94 acoustic instability, at $\beta < 1$ there are still unstable modes driven by the pressure anisotropy.

### 3.4.6 Role of Plasma Microinstabilities

Future simulations will shed light on the long-term evolution of the CRAB instability. Nevertheless, we can already anticipate that plasma microinstabilities growing at the ion gyroscale, such as the mirror and firehose instabilities, may significantly affect the instability at large amplitudes.

Both the mirror (Barnes 1966; Hasegawa 1969) and firehose (Rosenbluth 1956; Chandrasekhar et al. 1958; Parker 1958) instabilities are excited when the pressure anisotropy becomes comparable to the magnetic pressure: the mirror instability is excited when $\Delta p \gtrsim B^2/8\pi$, while the firehose instability is excited when $\Delta p \lesssim -B^2/4\pi$. Kinetic simulations have shown that these instabilities tend to pin the pressure anisotropy near the instability thresholds via increased scattering of particles through wave-particle interactions (Kunz et al. 2014).

When the acoustic waves grow to large amplitudes and the microinstabilities become important ($\delta\Delta p \sim B^2/8\pi$), $\Delta p$ will no longer be set just by the fluid flow (i.e. the sound wave). Instead, it will be set by the plasma microinstabilities, which will act to pin $\Delta p$ near marginal stability. Recall that the work done by $\Delta p$ on the cosmic rays is the driver of the acoustic instability. It thus seems plausible that the role of the gyroscale microinstabilities will be to slow down (and/or perhaps ultimately suppress) the acoustic instability.

At what sound-wave amplitudes do the plasma microinstabilities become important? For simplicity, consider an acoustic wave with $\delta\rho/\rho \gg \delta B/B$ (as is the case for the rapidly growing mode propagating parallel to $\boldsymbol{B}$). The pressure anisotropy is given by

$$\delta\Delta p = 3\rho\nu_{\rm B}\frac{d}{dt}\ln\frac{B}{\rho^{2/3}} \sim \rho\nu_{\rm B}\omega_s\frac{\delta\rho}{\rho}. \tag{3.49}$$

$\delta\Delta p \sim B^2/8\pi$ when

$$\frac{\delta p_g}{p_g} \sim \frac{\omega_s}{\omega_{\rm B}}\beta^{-1} \sim \frac{1}{kl_{\rm mfp}}\beta^{-1}. \tag{3.50}$$

In high-$\beta$ systems, it is therefore the long-wavelength modes that can grow to large amplitudes without exciting kinetic microinstabilities. Short-wavelength perturbations ($kl_{\rm mfp} \sim 1$), which tend to be the fastest growing modes, are affected by pressure-anisotropy-driven microinstabilities at smaller amplitudes than the long-wavelength modes.

## 3.5 Applications

In this section we speculate on example astrophysical applications of the CR-driven acoustic instability. We first consider the impact of the CRAB instability on sound waves propagating through galaxy clusters (Section 3.5.1). This is motivated by large-amplitude surface-brightness fluctuations observed in the Perseus cluster (Fabian et al. 2003), often interpreted to be long-wavelength sound waves. In 3.5.2 we speculate that cosmic rays may efficiently excite sound waves in the vicinity of shocks and in the outskirts of galaxy and cluster halos close to the virial radius. We also argue that the sound waves excited by the low-energy GeV cosmic rays may be important for the scattering of higher-energy cosmic rays (Section 3.5.3).

### 3.5.1 X-Ray Ripples in Perseus

*Chandra* X-ray observations have revealed long-wavelength, $\mathcal{O}(10\ \mathrm{kpc})$, surface-brightness ripples in the Perseus cluster (Fabian et al. 2003; Fabian et al. 2006). The inferred $\mathcal{O}(10\%)$ density fluctuations are believed to be sound waves propagating through the cluster. More generally, it is believed that sound waves excited by time-variable AGN activity are important for heating cluster plasmas (e.g., Li et al. 2015; Bambic & Reynolds 2019). The gas in Perseus and other clusters is weakly collisional and is likely also filled with cosmic rays. Thus, it is plausible that these sound waves are affected by the CRAB instability described in this paper.

The CR pressure fraction in Perseus and other cluster environments is constrained to be of order a few percent to a few tens of percent.[7] For a gas temperature of $5 \times 10^7$K and number density $0.03\ \mathrm{cm}^{-3}$ appropriate for Perseus (Fabian et al. 2006), the ion mean free path is of order 0.1 kpc. This translates into $\omega_\mathrm{B}/\omega_s \sim kl_\mathrm{mfp} \sim 0.1$ for a ~10 kpc wavelength. By how much can this wave be amplified through the CRAB instability?

We show growth rates of a $kl_{\nu_\mathrm{B}} = 0.1$ (a wavelength of order $\lambda \sim 10$ kpc) acoustic wave in $\mathrm{Myrs}^{-1}$ in the $(\eta, \beta)$ plane in Figure 3.5a. We use $\mathrm{Pr} = 1$ and assume no CR diffusion, $\Phi = 0$. To clearly show where the instability becomes important, we explicitly show contours where the growth rates are $0, 0.02$ and $0.1\ \mathrm{Myrs}^{-1}$.

The sound speed in Perseus is of order $\sim 10^8\ \mathrm{cm\ s}^{-1}$, so that waves propagate a distance 50 kpc (say) in ~50 Myrs. For the wave to undergo at least one e-folding in that time, the required growth rate is $\mathrm{Im}(\omega) \gtrsim 0.02\ \mathrm{Myrs}^{-1}$. This is satisfied if $\eta$ and $\beta$ are sufficiently large, see Figure 3.5a. Whether rapid growth of long-wavelength waves does indeed occur in Perseus is somewhat unclear, due to the lack of sufficiently good constraints on the cluster magnetic-field strength and cosmic-ray pressure (there are also uncertainties in the particle mean free path due to the role of

---

[7] A few percent according to Aleksić et al. (2010) and Aleksić et al. (2012), but their study uses primarily high-energy CRs. The upper limit on the total CR pressure in Perseus – dominated by the low-energy CRs that are the most important for this work – is significantly larger in Huber et al. (2013).

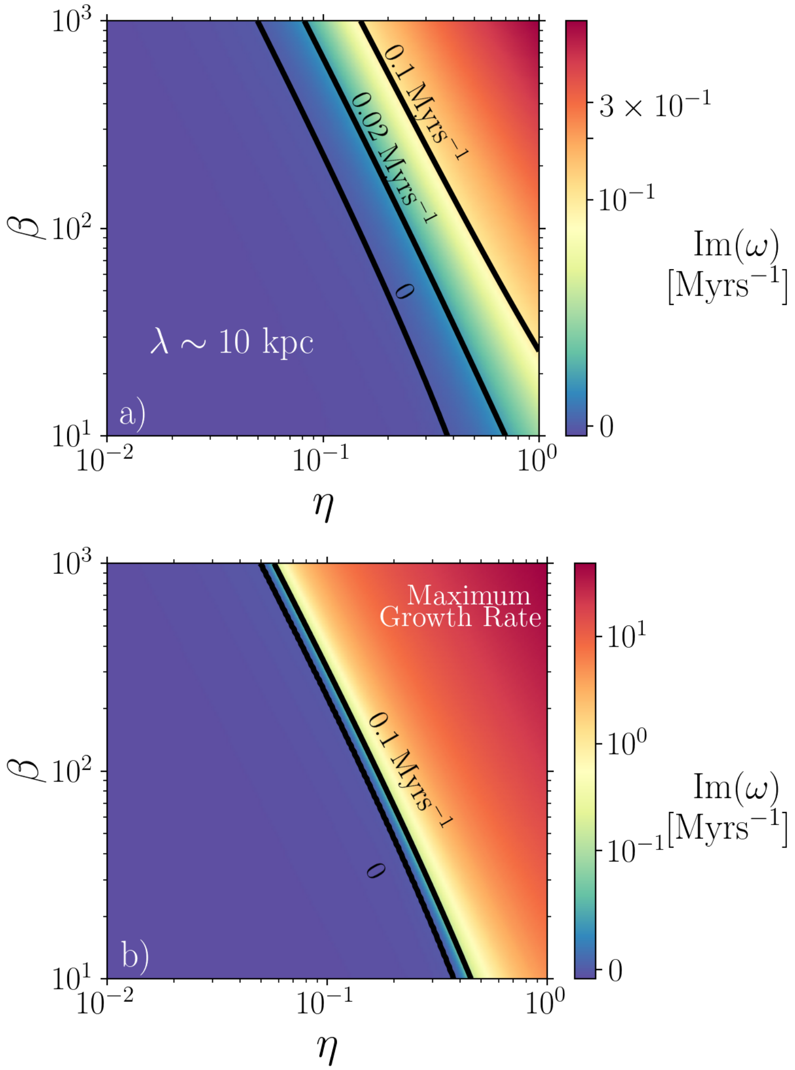


*Figure 3.5*: Growth rate of the acoustic instability as a function of $\eta$ and $\beta$ for Perseus-like parameters: $T \sim 5 \times 10^7$K and $n_\mathrm{i} \sim 0.03$ cm$^{-3}$ (Fabian et al. 2006), so that the ion mean free path $l_\mathrm{mfp}$ is of order 0.1 kpc. Here we assume $\beta = 100$, Pr = 1 and no CR diffusion. Panel a) shows the maximum growth rate of a $kl_{\nu_\mathrm{B}} = 0.1$ sound wave ($\lambda \sim 10$ kpc; 10 kpc corresponds to the approximate wavelength of the X-ray surface-brightness fluctuations in Perseus as observed by *Chandra*). Note that these growth rates are larger at larger distances from the cluster center, where the density is lower and the mean free path is larger. Panel b) shows the maximum growth rate when we consider all wavelengths that satisfy $kl_\mathrm{mfp} \leq 1$. Significant amplification over timescales of order 10 Myrs (timescale for sound waves to propagate tens of kpc) is plausible for realistic values of $\eta$ and $\beta$.

plasma microinstabilities). However, it seems possible at least in localised regions with sufficiently large $\eta$ and $\beta$ (see also Section 3.5.2). We also note that for $\beta \sim 100$, a $kl_{\rm mfp} = 0.1$ ($\sim 10$ kpc) wave can grow to fairly large amplitudes, $\delta p_g / p_g \sim \delta\rho/\rho \sim 10\%$, before pressure-anisotropy microinstabilities become important which likely slow down and/or ultimately suppress the instability (see Section 3.4.6). This is consistent with the $\mathcal{O}(10\%)$ density fluctuations inferred in Perseus (Fabian et al. 2003; Fabian et al. 2006). Finally, we note that at larger distances from the cluster core the density is lower and the mean free path is larger. As a result, long-wavelength ($\lambda \sim 10$ kpc) modes will have faster growth rates at large distances from the cluster center.

In Figure 3.5b we do not restrict our attention to $\lambda \sim 10$ kpc wavelengths, and instead show the overall maximum growth rates in the $(\eta, \beta)$ plane. We use the same Perseus temperatures and densities as before, such that $l_{\rm mfp} \sim 0.1$ kpc, and we consider wavelengths satisfying $kl_{\rm mfp} \leq 1$ ($kl_{\nu_{\rm B}} \leq 1$). The CRAB instability occurs and has fast growth rates for a wide range of realistic cluster values of $\eta$ and $\beta$. We therefore conclude that cosmic rays likely lead to large amplifications of kpc-scale sound waves propagating in dilute cluster plasmas.

### 3.5.2 Sound-Wave Excitation in the Vicinity of (Virial) Shocks

The CRAB instability is particularly important at high cosmic-ray pressures, i.e. large $\eta$. This suggests that the instability is easily excited in the vicinity of shocks that are responsible for CR acceleration, i.e. where $\eta$ is typically much higher than its average value in the ambient medium. This may be relevant for shocks in supernova remnants and shocks driven by galactic winds or AGN jets in galaxy halos and clusters.

In addition, cluster simulations that include the production of cosmic rays in structure-formation shocks find that the CR pressure fraction is higher close to the virial radius (virial shock) than in the central regions of the cluster (Pfrommer et al. 2008). It seems possible that sound waves excited close to the virial radius through the CRAB instability discussed in this work can then propagate in towards the cluster core. Modes with longer wavelengths, $\sim$ 10s of kpc, will grow much faster at large radii near the viral radius than in the cluster core because of the much lower density and larger ion mean free path at these radii.

### 3.5.3 Scattering of High-Energy Cosmic Rays

The overstable sound waves found in this paper have long wavelengths ($\gtrsim 1$ kpc in the ICM and $\gtrsim 1$ pc in the hot ISM) and can have growth rates that are significant compared to the oscillation frequency. The purpose of this section is to point out that the growth rates of the sound-wave instability are significantly larger than the growth rates of the Alfvén waves excited by high-energy CRs through the streaming instability. The streaming-instability growth rate is given by (Kulsrud & Pearce 1969; Zweibel 2013),

$$\Gamma_k \sim \Omega_0 \frac{n_{\rm CR}(p > p_{\rm min})}{n_{\rm i}} \frac{{\rm v_D} - {\rm v_A}}{{\rm v_A}}, \tag{3.51}$$

where $\rm v_D$ is the CR drift speed, $\Omega_0$ is the nonrelativistic gyrofrequency and $n_{\rm i}$ is the thermal ion number density. $n_{\rm CR}(p > p_{\rm min})$ is the number density of CRs that can resonate with a wave with

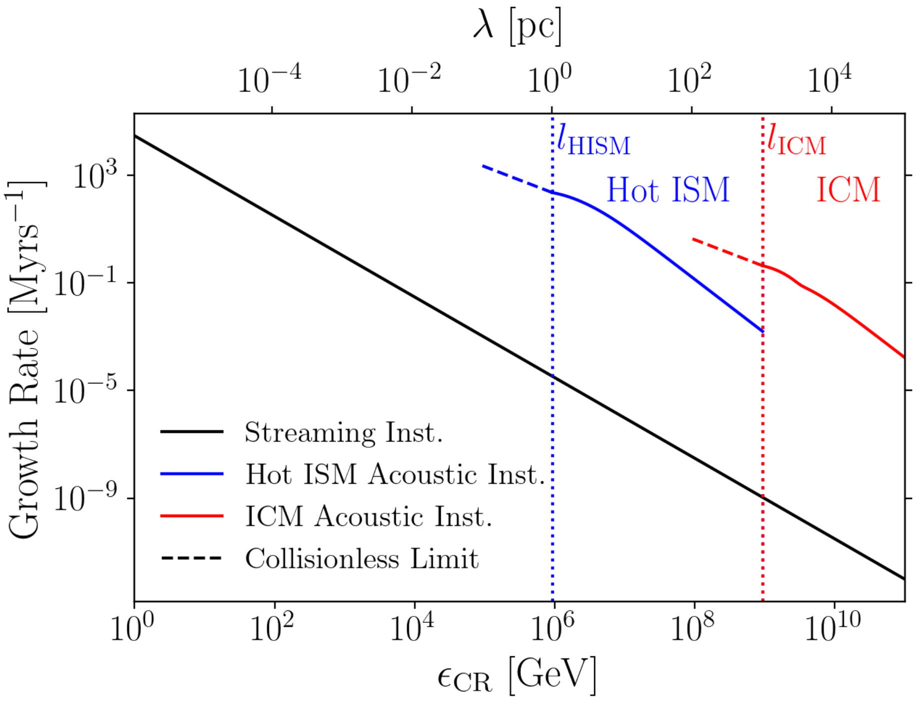


*Figure 3.6*: Schematic growth-rate comparison of the gyroresonant streaming instability of Alfvén waves and the long-wavelength acoustic instability excited by the GeV CR fluid coupled to the thermal plasma. The growth rate is plotted against CR energy (bottom horizontal axis) as well as wavelength (top horizontal axis; the two are related by the resonance condition $r_L/\lambda \sim \epsilon_{\rm CR}/\lambda eB \sim 1$, where $r_L$ is the CR gyroradius). For the streaming-instability growth rates we use a single CR spectral slope, $\alpha = 4.5$, a 1 $\mu$G magnetic field, $n_{\rm CR}/n_{\rm i} = 10^{-7}$ and $(v_D - v_A)/v_A = 1$ (eq. 3.51). The acoustic instability is plotted for Pr = 1 and $\Phi = 0$ (no CR diffusion), and wavelengths larger than the ion mean free path, $l_{\rm mfp}/\lambda \leq 1$ (we assume $l_{\rm mfp} = 1$ pc in the hot ISM and $l_{\rm mfp} = 1$ kpc in the ICM). We use $\eta = 1$, $\beta = 10$ in the hot ISM and $\eta = 0.2$, $\beta = 100$ in the ICM (these values of $\eta$ correspond to $n_{\rm CR}/n_{\rm i} \sim \mathcal{O}(10^{-7})$ for typical hot ISM and ICM temperatures). At long wavelengths, the growth rates of the acoustic instability are orders of magnitude larger than the streaming-instability growth rates. In principle, the sound-wave instability grows sufficiently fast to contribute to the scattering of higher-energy (~PeV and ~EeV) cosmic rays. The dashed blue and red lines represent sub-mean-free-path scales, where the thermal plasma is collisionless. We defer a detailed treatment of this regime to future work, but preliminary calculations suggest that the instability is still present in the collisionless limit.

wavenumber $k$, and $p_{\rm min} = m\Omega_0/k$. Because the CR spectrum is steep, the number of high-energy cosmic rays resonating with long-wavelength modes is very small. This leads to very small $\Gamma_k$ for modes with wavelengths that can scatter and confine the high-energy CRs: $f(p) \propto p^{-\alpha}$ with $\alpha \approx 4.5$, so $n_{\rm CR}(p > p_{\rm min}) \propto p_{\rm min}^{3-\alpha} \propto k^{\alpha-3}$, which decays rapidly with CR energy. As a result, high-energy cosmic rays are not able to confine themselves. Here we inspect the possibility that the acoustic instability excited by the GeV cosmic-ray fluid can scatter and at least partially confine higher-energy cosmic rays.

Figure 3.6 shows a schematic growth-rate comparison of the gyroresonant streaming instability of Alfvén waves and the CRAB instability considered in this work. The growth rate is plotted against CR energy (bottom horizontal axis) as well as wavelength (top horizontal axis). The wavelength and CR energy are related by the resonance condition $\epsilon_{\rm CR} \sim \lambda e B$. We assume a CR spectral slope $\alpha = 4.5$ and $({\rm v_D} - {\rm v_A})/{\rm v_A} = 1$. For the acoustic instability we consider wavelengths larger than the ion mean free path, $l_{\rm mfp}/\lambda \leq 1$. For the hot ISM, we assume an ion mean free path $l_{\rm mfp} = 1$ pc and for the ICM we assume an ion mean free path $l_{\rm mfp} = 1$ kpc. The Braginskii MHD description of the thermal plasma is appropriate above the ion mean-free-path scale. However, preliminary calculations using collisionless fluid closures suggest that the instability also exists below the mean-free-path scale (see Section 3.4.4). We show this using the dashed blue and red lines. We stress again that the collisionless description of the thermal plasma coupled to a CR-pressure equation breaks down at small scales where the CRs are no longer coupled to the thermal plasma. The growth rates are not plotted below this scale in Figure 3.6 (the CR mean free path is somewhat uncertain and for this reason we extend growth rates only one order of magnitude below the ion mean-free-path scale; however, this range might be significantly larger, e.g. in the ICM where the ion mean free path is large).

Figure 3.6 shows that the growth rates of the CRAB instability are orders of magnitude faster than the growth rates of the streaming instability excited by the high-energy CRs. The growth rate is relatively independent of propagation angle for $\theta \lesssim 55^\circ$ (Figure 3.2), so modes with appreciable $\delta B_\perp/B$ can be excited. Sound waves may, in principle, grow sufficiently fast to reach large amplitudes and efficiently scatter high-energy cosmic rays. The CR scattering rate is proportional to $\Omega(\delta B_\perp/B)^2$. If the acoustic waves destabilised by the GeV CRs saturate at sufficiently large $\delta B_\perp/B$, the acoustic instability identified here may significantly affect cosmic-ray confinement. While large $\delta B_\perp/B$ seem possible given the fast growth rates, future simulations will be necessary to study the saturation of the instability and address the efficiency of scattering high-energy CRs. Finally, we note that turbulence will likely be produced in the gas as a result of the CRAB instability. This may significantly affect the scattering and transport properties of intermediate-energy ($\lesssim$ PeV) cosmic rays, whose gyroradii are too small to directly resonate with linearly unstable acoustic waves.

## 3.6 Conclusions

The interstellar, circum-galactic and intracluster media are filled with dilute, weakly-collisional plasmas characterised by anisotropic viscosity and conduction. Without cosmic rays, these anisotropic transport properties lead to the well-known damping of sound waves (the slow and fast magnetosonic modes). In this paper we have shown that when cosmic rays are present, sound waves can instead grow exponentially in time, even for small CR pressures. We have termed this the Cosmic Ray Acoustic Braginskii (CRAB) instability.

We model the dilute plasmas filled with cosmic rays by using the Braginskii MHD closure for weakly collisional plasmas (Braginskii 1965) coupled to a pressure equation for the cosmic rays (Section 3.3). The cosmic rays are assumed to stream at the Alfvén speed $\mathrm{v}_{\mathrm{A},\Delta p}$, which in a weakly collisional plasma depends on the pressure anisotropy $\Delta p$ (eq. 3.10). We also include CR diffusion along the magnetic-field direction.

The key frequencies and dimensionless parameters in our problem are summarised in Section 3.3.3. We focus on high-$\beta$ ($\beta = 8\pi p_g/B^2 \sim 100$) plasmas, as is appropriate for the ICM. The Braginskii MHD model is valid provided that the timescales of interest are longer than the ion-ion collision time. We impose this by constraining the anisotropic-viscous (Braginskii) frequency, $\omega_B$, to be smaller than the sound frequency, $\omega_s$ (see Section 3.3.3).

The CRAB instability is driven by a phase shift between the CR-pressure and gas-density perturbations. This phase shift is introduced by the dependence of the Alfvén speed on $\Delta p$ (eq. 3.10). The physical mechanism driving the instability is sketched out in Figure 3.1: work done by the pressure anisotropy on the cosmic rays enhances regions of larger than average CR pressure, leading to a positive feedback loop. Sound waves are unstable if $\eta = p_c/p_g \gtrsim \alpha\beta^{-1/2}$, where $\alpha$ depends on the thermal Prandtl number and the CR diffusion coefficient. We find that $\alpha$ is typically of order 1 (unless the CR diffusion coefficient is much larger than the thermal-plasma anisotropic viscosity, in which case $\alpha > 1$; see bottom panel of Figure 3.4). Thus, even small CR pressures are sufficient for instability in high-$\beta$ plasmas such as the ICM. We find that the acoustic instability is characterised by large growth rates, comparable to the sound-wave oscillation frequency (Figure 3.4).

The growth rates absent CR diffusion are not a strong function of propagation angle relative to $\boldsymbol{B}$ for $\theta \lesssim 55^\circ$ (Figure 3.2). However, the fastest growing mode is typically propagating parallel to the magnetic-field direction (except for $\eta$ just above marginal stability or when CR diffusion is strong, see Figure 3.2 and Figure 3.4). This result motivated a simplified 1D derivation of the dispersion relation, which we show in (3.40). Growth rates are typically largest at the highest $k$, except at small $\eta$ just above the instability threshold or when CR diffusion is significant (Figure 3.3 and Figure 3.4).

We considered astrophysical implications of the CRAB instability in Section 3.5. In Section 3.5.1 we argue that the instability is likely important for amplifying sound waves propagating through galaxy cluster and group environments. This includes the Perseus cluster, where long-wavelength, large-amplitude X-ray surface-brightness fluctuations observed by *Chandra* are often interpreted as sound waves. We show instability growth rates as a function of $\eta$ and $\beta$ for Perseus-like parameters in Figure 3.5. In Section 3.5.2 we hypothesise that the acoustic instability is likely important near shocks, where the CR pressure is large. This includes the outskirts of galactic and cluster halos, i.e. regions close to the virial shock, as well as shocks associated with supernovae, galactic winds, or AGN winds/jets propagating into the hot ISM or halo environments. In Section 3.5.3 we speculate that the long-wavelength acoustic modes excited by the GeV cosmic-ray fluid can contribute to the scattering of higher-energy cosmic rays. In Figure 3.6 we show that the long-wavelength acoustic modes grow orders of magnitude faster than the Alfvén waves excited by the high-energy CRs through the gyroresonant streaming instability. It remains to be seen, however, whether the sound waves grow to large enough amplitudes and/or generate smaller-scale fluctuations through turbulence to efficiently scatter TeV to EeV cosmic rays.

Future simulations will address the saturation of the CRAB instability. They will show whether the excited sound waves can grow to large enough amplitudes to efficiently scatter high-energy cosmic rays. Global simulations that include both Braginskii MHD and cosmic rays will shed light on the importance of the acoustic instability for the evolution of gas and the propagation of sound waves in the ISM, galactic halos and the ICM. We also plan to extend the CR–Braginskii MHD fluid model to collisionless models of the thermal plasma.

Finally, we note the caveat that the dominant transport process of cosmic rays through galaxies and clusters remains uncertain (e.g., Amato & Blasi 2018). For example, even cosmic rays that are not strongly coupled to the thermal plasma (i.e. not locked to the Alfvén frame, as a result of a low pitch-angle scattering rate, e.g. due to wave damping) may actually not be diffusing under certain conditions, but instead streaming at super-Alfvénic speeds (Skilling 1971; Wiener et al. 2013). The development of more accurate fluid models of cosmic rays is therefore a high priority.

# Acknowledgements

We thank R. Farber, Y. Jiang, M. Kunz, S. P. Oh, C. Pfrommer, M. Ruszkowski, A. Spitkovsky, J. Stone, I. Zhuravleva & E. Zweibel for enlightening discussions. This research was supported in part by the Heising-Simons Foundation, the Simons Foundation, and National Science Foundation Grant No. NSF PHY-1748958, NSF grant AST-1715070, and a Simons Investigator award from the Simons Foundation. PK would like to thank the Kavli Institute for Theoretical Physics for their hospitality and support offered via the Graduate Fellowship Program. Support for J.S. was provided by Rutherford Discovery Fellowship RDF-U001804 and Marsden Fund grant UOO1727, which are managed through the Royal Society Te Apārangi.

# Appendix

## 3.A Acoustic Instability in Two-Fluid Plasma

In Section 3.3.3 we pointed out that the ion-electron temperature equilibration timescale is longer in the regime of interest than the relevant sound timescale. Using a single-fluid approach with heat flow carried by the electrons is then not correct. Instead, separate entropy equations and heat fluxes should be used for each species. In the main text, we only considered a single thermal fluid and a single heat flux for simplicity (with varying thermal Prandtl number), and we demonstrated that our results do not depend strongly on the value of the chosen thermal diffusivity. Here we show that our conclusions do not change in a more accurate two-fluid model, when separate electron and ion pressure equations are included.

In the two-fluid model we consider, the continuity, induction and CR pressure equations (3.1, 3.3 and 3.5) remain unchanged (we assume quasi-neutrality, $n_\mathrm{i} = n_\mathrm{e}$). The remaining equations that we need to solve are the momentum equation, and the ion and electron entropy equations:

$$\rho\frac{d\mathbf{v}}{dt} = -\boldsymbol{\nabla}\left(p_i + p_e + p_c + \frac{B^2}{8\pi}\right) + \frac{\boldsymbol{B}\cdot\boldsymbol{\nabla}\boldsymbol{B}}{4\pi} + \boldsymbol{\nabla}\cdot(\hat{\boldsymbol{b}}\hat{\boldsymbol{b}}\Delta p) - \frac{1}{3}\boldsymbol{\nabla}\Delta p, \tag{3.52}$$

$$\frac{1}{\gamma-1}\frac{dp_i}{dt} = -\frac{\gamma}{\gamma-1}p_i\boldsymbol{\nabla}\cdot\mathbf{v} - \alpha_1\mathbf{v}_{\mathbf{A},\boldsymbol{\Delta p}}\cdot\boldsymbol{\nabla}p_c - \alpha_2\boldsymbol{\nabla}\cdot\left(\boldsymbol{\Pi}\cdot\mathbf{v}\right) - \boldsymbol{\nabla}\cdot\boldsymbol{Q_i} - \frac{p_i - p_e}{\tau_\mathrm{eq}}, \tag{3.53}$$

$$\frac{1}{\gamma-1}\frac{dp_e}{dt} = -\frac{\gamma}{\gamma-1}p_e\boldsymbol{\nabla}\cdot\mathbf{v} - (1-\alpha_1)\mathbf{v}_{\mathbf{A},\boldsymbol{\Delta p}}\cdot\boldsymbol{\nabla}p_c - (1-\alpha_2)\boldsymbol{\nabla}\cdot\left(\boldsymbol{\Pi}\cdot\mathbf{v}\right) - \boldsymbol{\nabla}\cdot\boldsymbol{Q_e} + \frac{p_i - p_e}{\tau_\mathrm{eq}}. \tag{3.54}$$

$p_i$ and $p_e$ are the ion and electron pressures, respectively. $\tau_\mathrm{eq}$ is the timescale over which the electrons and ions come into thermal equilibrium. This timescale is long compared to the electron-electron and ion-ion collision times, $\tau_\mathrm{ee} \sim \sqrt{m_\mathrm{e}/m_\mathrm{i}}\tau_\mathrm{ii} \sim (m_\mathrm{e}/m_\mathrm{i})\tau_\mathrm{eq}$. The smallness of the equilibration term in a weakly collisional plasma is what motivates the 2-fluid model and so we will drop the terms $\propto \tau_\mathrm{eq}^{-1}$ in equations 3.53 and 3.54 in this section. $\alpha_1$ and $\alpha_2$ are parameters which set how much of the CR and viscous heating goes into the ions vs. electrons. Viscous heating does not enter in our analysis to linear order, and so $\alpha_2$ can be ignored. We choose $\alpha_1 = 0.5$, but our results do not depend on it, as the instability is ultimately not driven by CR heating at high $\beta$. In

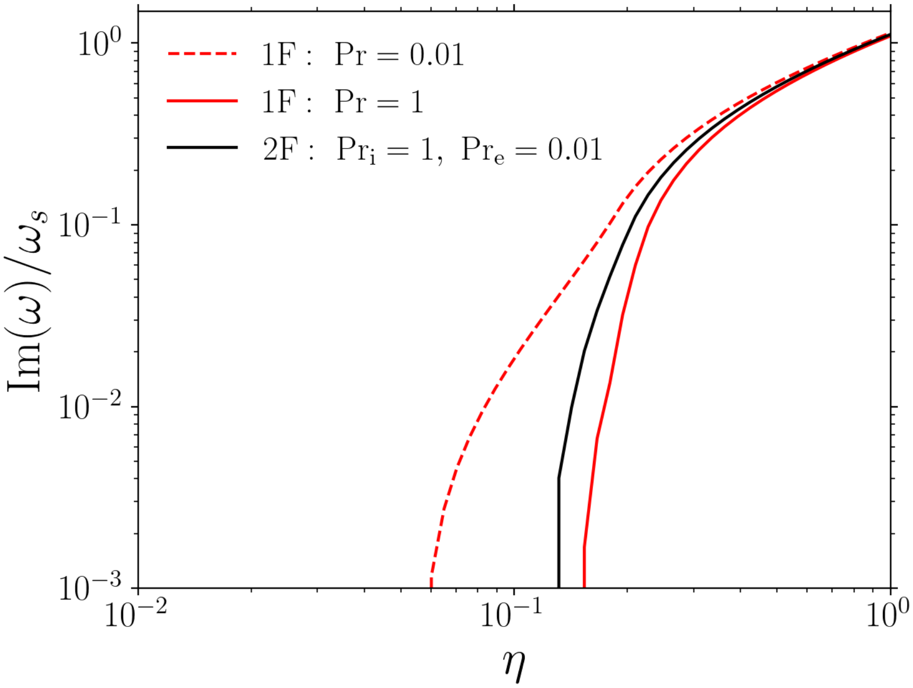


*Figure 3.7*: Comparison of the maximum sound-wave growth rates for the single-fluid (1F) and two-fluid (2F) ion-electron plasma as a function of $\eta$, for $\beta = 100$ and $\omega_d = 0$ (no CR diffusion). This figure is analogous to Figure 3.4. We find that our results are not significantly affected by the extension to two entropy equations and heat fluxes. The two-fluid growth rate for $\nu_B/\chi_i = 1$ and $\nu_B/\chi_e = 0.01$ (black line) is essentially in between the 1-fluid prediction with Pr = 1 and Pr = 0.01 (red lines).

(3.53) and (3.54), $\boldsymbol{Q_i} = -n_\mathrm{i} k_\mathrm{B} \chi_i \hat{\boldsymbol{b}}\hat{\boldsymbol{b}} \cdot \boldsymbol{\nabla} T_i$ is the ion heat flux ($\chi_i$ is the ion thermal diffusivity) and $\boldsymbol{Q_e} = -n_\mathrm{e} k_\mathrm{B} \chi_e \hat{\boldsymbol{b}}\hat{\boldsymbol{b}} \cdot \boldsymbol{\nabla} T_e$ is the electron heat flux ($\chi_e$ is the electron thermal diffusivity).

The linearised versions of (3.53) and (3.54) are:

$$\omega \frac{\delta p_i}{p_i} = \gamma \boldsymbol{k} \cdot \mathbf{v} + 2\alpha_1 (\gamma - 1)\omega_a \eta \frac{\delta p_c}{p_c} - i(\gamma - 1)\omega_\mathrm{cond,i}\Big(\frac{\delta p_i}{p_i} - \frac{\delta \rho}{\rho}\Big), \tag{3.55}$$

$$\omega \frac{\delta p_e}{p_e} = \gamma \boldsymbol{k} \cdot \mathbf{v} + 2(1 - \alpha_1)(\gamma - 1)\omega_a \eta \frac{\delta p_c}{p_c} - i(\gamma - 1)\omega_\mathrm{cond,e}\Big(\frac{\delta p_e}{p_e} - \frac{\delta \rho}{\rho}\Big), \tag{3.56}$$

where we defined the ion and electron thermal diffusion frequencies, $\omega_\mathrm{cond,i/e} = \chi_{i/e}(\hat{\boldsymbol{b}} \cdot \boldsymbol{k})^2$.

We assume an equilibrium with $p_i = p_e$. As in Section 3.4.2, we can derive a 1D dispersion relation for the two-fluid acoustic instability (again assuming high $\beta$, $\omega_s \gg \omega_a$):

$$\begin{aligned} 0 =& \omega^2 - \frac{\omega_s^2}{2\gamma}\Big(\frac{\gamma\omega + i(\gamma - 1)\omega_\mathrm{cond,i}}{\omega + i(\gamma - 1)\omega_\mathrm{cond,i}} + \frac{\gamma\omega + i(\gamma - 1)\omega_\mathrm{cond,e}}{\omega + i(\gamma - 1)\omega_\mathrm{cond,e}}\Big) \\ &+ \frac{4}{3} i\omega\omega_\mathrm{B} - \eta\frac{\omega_s^2}{\gamma}\Big(\frac{4}{3} + \frac{4}{3} i\frac{\omega_\mathrm{B}}{\omega_a}\Big)\Big(1 + i\frac{\omega_d}{\omega}\Big)^{-1}. \end{aligned} \tag{3.57}$$

Since electron conduction is rapid, we can consider the regime $\omega_\mathrm{cond,e} \gg \omega_s$, such that the electrons are essentially isothermal. The above then simplifies to:

$$\begin{aligned} 0 =& \omega^2 - \frac{\omega_s^2}{2\gamma}\Big(\frac{\gamma\omega + i(\gamma - 1)\omega_\mathrm{cond,i}}{\omega + i(\gamma - 1)\omega_\mathrm{cond,i}} + 1\Big) + \frac{4}{3} i\omega\omega_\mathrm{B} \\ &- \eta\frac{\omega_s^2}{\gamma}\Big(\frac{4}{3} + \frac{4}{3} i\frac{\omega_\mathrm{B}}{\omega_a}\Big)\Big(1 + i\frac{\omega_d}{\omega}\Big)^{-1} \end{aligned} \tag{3.58}$$

Note that (3.57) and (3.58) are very similar to the dispersion relation in (3.40). As a result, we find that our results are not significantly affected by the extension to two entropy equations and heat fluxes. This is confirmed in Figure 3.7, where we see that the two-fluid fast magnetosonic growth rate for $\nu_\mathrm{B}/\chi_i = 1$ and $\nu_\mathrm{B}/\chi_e = 0.01$ is essentially in between the single-fluid prediction with $\mathrm{Pr} = 1$ and $\mathrm{Pr} = 0.01$.

# Chapter 4

# A New Buoyancy Instability in Galaxy Clusters due to Streaming Cosmic Rays

A version of this article is to be submitted to MNRAS by Kempski P., Quataert E., and Squire J.

## 4.1 Abstract

Active Galactic Nuclei (AGN) are believed to provide the energy that prevents runaway cooling of gas in the cores of galaxy clusters. However, how this energy is transported and thermalized throughout the Intracluster Medium (ICM) remains unclear. In recent work we showed that streaming cosmic rays (CRs) destabilise sound waves in dilute ICM plasmas. Here we show that CR streaming in the presence of gravity also destabilises a pressure-balanced wave. We term this new instability the CR buoyancy instability (CRBI). In stark contrast to standard results without CRs, the pressure-balanced mode is highly compressible at short wavelengths due to CR streaming. Growth rates at short wavelengths are of order $(p_c/p_g)\beta^{1/2}\omega_{\rm ff}$, where $p_c/p_g$ is the ratio of CR pressure to thermal gas pressure, $\beta$ is the ratio of thermal to magnetic pressure and $\omega_{\rm ff}$ is the free-fall frequency. The CRBI operates alongside buoyancy instabilities driven by background heat fluxes, i.e. the heat-flux-driven buoyancy instability (HBI) and the magneto-thermal instability (MTI). When the thermal mean free path $l_{\rm mfp}$ is $\ll$ the gas scale height $H$, the HBI/MTI set the growth rate on large scales, while the CRBI sets the growth rate on small scales. Conversely, when $l_{\rm mfp} \sim H$ and $(p_c/p_g)\beta^{1/2} \gtrsim 1$, the CRBI has growth rates that exceed HBI/MTI growth rates even on large scales. Our results suggest that CR-driven instabilities may be partially responsible for the sound waves/weak shocks and turbulence observed in galaxy clusters. CR-driven instabilities generated near radio bubbles may also play an important role redistributing AGN energy throughout clusters.

## 4.2 Introduction

The cores of galaxy clusters are filled with virialized, hot gas, with typical temperatures exceeding $10^7$ K. The X-ray luminosities of most cluster cores imply cooling times that are significantly shorter than the ages of these systems. Without a source of heating, this hot gas is expected to cool, sink to the center and form stars at a high rate. However, observations find significantly smaller star formation rates and cold gas masses than are predicted by the “cooling flow" model (e.g., Peterson & Fabian 2006). This suggests that there is a source of heating present that keeps the gas in cluster cores in approximate thermal balance.

Central Active Galactic Nuclei (AGN) and the interaction of their jets with the Intracluster Medium (ICM) are believed to play an important role in providing the energy that prevents runaway cooling of ICM gas. In particular, observations suggest that energy is carried away from the central AGN by jet-inflated bubbles of relativistic plasma that buoyantly rise into the ICM. There is a strong correlation between the power needed to inflate the bubbles and the radiative losses of the hot gas (Churazov et al. 2000; Bîrzan et al. 2004; Rafferty et al. 2006; Nulsen et al. 2009; Hlavacek-Larrondo et al. 2012; see Werner et al. 2019 for a recent review).

How this energy is subsequently transported and thermalized throughout cluster cores remains an open question. It is possible that the buoyantly rising radio bubbles stir turbulence by exciting internal gravity waves (IGWs; e.g., Zhuravleva et al. 2016, Zhang et al. 2018), launch sound waves and/or weak shocks (e.g., Fabian et al. 2003, Fabian et al. 2006, Sternberg & Soker 2009), and/or inject cosmic rays (CRs) into the ICM (e.g., Guo & Oh 2008, Jacob & Pfrommer 2017a, Jacob & Pfrommer 2017b). These processes can plausibly occur to some extent simultaneously, but it is unclear which (if any) one is the dominant channel for ICM heating.

Relativistic CRs from both star formation and AGN may play an important role in the evolution of gas in clusters by driving outflows and heating diffuse gas (e.g., Breitschwerdt et al. 1991; Loewenstein et al. 1991; Everett et al. 2008; Socrates et al. 2008; Guo & Oh 2008 Zweibel 2013; Ruszkowski et al. 2017; Zweibel 2017; Jacob & Pfrommer 2017a; Jacob & Pfrommer 2017b; Ehlert et al. 2018; Farber et al. 2018; Kempski & Quataert 2020; Quataert et al. 2022b; Quataert et al. 2022a). CRs couple to the thermal gas by scattering from small-scale magnetic fluctuations. In self-confinement theory, cosmic rays are scattered by Alfvén waves propagating down the CR pressure gradient, which they themselves excite through the streaming instability (Kulsrud & Pearce 1969). Pitch-angle scattering by the excited Alfvén waves isotropises the cosmic rays in the frame of the waves. In the absence of damping of the self-excited waves, this results in CR streaming relative to the thermal gas at the local Alfvén speed. If damping is present, CR propagation deviates from pure Alfvénic streaming. The form of the transport correction is, however, quite peculiar, as it corresponds to neither streaming nor diffusion (Skilling 1971, Wiener et al. 2013; Kempski & Quataert 2021).

In recent work, we showed that streaming cosmic rays drive a rapidly growing acoustic instability in dilute ICM plasmas, the Cosmic Ray Acoustic Braginskii (CRAB) instability (Kempski et al. 2020). This suggests that the different channels for transferring and thermalizing energy in the ICM (waves, CRs, turbulence...), usually considered separately in theoretical models, may in fact be closely related. Here, we demonstrate that in the presence of gravity, streaming cosmic rays also destabilise a pressure-balanced wave, more specifically the CR entropy mode modified by gravity. We term this instability the CR buoyancy instability (CRBI) because CRs and buoyancy (gravity) are critical for setting its properties. The growth rates of the CRBI are of order the natural buoyancy frequency (the local free-fall frequency) for plausible ICM parameters. Our work demonstrates the potential physical richness of CR feedback in dilute plasmas: both the CRBI considered in this work and the CRAB instability in Kempski et al. (2020) are driven by CR streaming at the Alfvén speed, which in a weakly collisional plasma depends on the pressure anisotropy of the thermal gas (the pressure is anisotropic because of the large thermal-particle mean free path in the ICM). This dependence introduces a new form of coupling between the CRs and the thermal gas, which is very unstable.

The remainder of this work is organised as follows. We present our model of cosmic rays coupled to a low-collisionality plasma in Section 4.3. In Section 4.4 we provide a physical overview of the instability. We derive a dispersion relation and asymptotic growth rates in Section 4.5. In Section 4.6 we present numerical solutions to the linearised system of equations for an isothermal background. We discuss the relationship to other buoyancy instabilities and thermal instability, and the importance of CR diffusion in Section 4.7. We summarise our results in Section 4.8.

## 4.3 Model

### 4.3.1 Equations

We consider a low-collisionality plasma coupled to streaming cosmic rays. We use the weakly-collisional Braginskii MHD model to describe the thermal gas (Braginskii 1965). The equations for the gas and cosmic rays are,

$$\frac{\partial \rho}{\partial t} + \boldsymbol{\nabla} \cdot (\rho \mathbf{v}) = 0, \tag{4.1}$$

$$\rho \frac{d\mathbf{v}}{dt} = -\boldsymbol{\nabla}\Big(p_\perp + p_c + \frac{B^2}{8\pi}\Big) + \frac{\boldsymbol{B} \cdot \boldsymbol{\nabla} \boldsymbol{B}}{4\pi} + \boldsymbol{\nabla} \cdot (\hat{\boldsymbol{b}}\hat{\boldsymbol{b}} \Delta p) + \rho \boldsymbol{g}, \tag{4.2}$$

$$\frac{\partial \boldsymbol{B}}{\partial t} = \boldsymbol{\nabla} \times (\mathbf{v} \times \boldsymbol{B}), \tag{4.3}$$

$$\rho T \frac{ds}{dt} = -\mathbf{v}_{\mathbf{st},\Delta p} \cdot \boldsymbol{\nabla} p_c - \boldsymbol{\nabla} \cdot \Big(\boldsymbol{\Pi} \cdot \mathbf{v}\Big) - \boldsymbol{\nabla} \cdot \boldsymbol{Q} - \rho^2 \Lambda(T), \tag{4.4}$$

$$\frac{dp_c}{dt} = -\frac{4}{3} p_c \boldsymbol{\nabla} \cdot (\mathbf{v} + \mathbf{v}_{\mathbf{st},\Delta p}) - \mathbf{v}_{\mathbf{st},\Delta p} \cdot \boldsymbol{\nabla} p_c + \boldsymbol{\nabla} \cdot (\kappa \hat{\boldsymbol{b}}\hat{\boldsymbol{b}} \cdot \boldsymbol{\nabla} p_c), \tag{4.5}$$

where $\mathbf{v}$ is the gas velocity, $\rho$ is the gas density, $\boldsymbol{B}$ is the magnetic field (with unit vector $\hat{\boldsymbol{b}}$), $s = k_{\mathrm{B}} \ln(p_g/\rho^\gamma)/(\gamma - 1)m_{\mathrm{H}}$ is the gas entropy per unit mass, $\Lambda(T)$ is the temperature-dependent cooling function and $p_c$ is the CR pressure. $d/dt \equiv \partial/\partial t + \mathbf{v} \cdot \boldsymbol{\nabla}$ denotes a total (Lagrangian) time derivative. $\Delta p = p_\perp - p_\parallel$ is the gas pressure anisotropy, where $p_\perp$ and $p_\parallel$ denote the pressures in the directions perpendicular and parallel to the magnetic field, respectively. $p_\perp$ and $p_\parallel$ are related to $p_g$ in eq. (4.4) by

$$p_\perp = p_g + \frac{1}{3}\Delta p. \tag{4.6}$$

The pressure anisotropy in Braginskii MHD (with viscosity $\nu_\mathrm{B} \approx p_g/\rho\nu_\mathrm{ii}$, where $\nu_\mathrm{ii}$ is the ion-ion collision rate; Braginskii 1965) is given by

$$\Delta p = p_\perp - p_\parallel = 3\rho\nu_\mathrm{B}(\hat{\boldsymbol{b}}\hat{\boldsymbol{b}} : \nabla\mathbf{v} - \frac{1}{3}\nabla \cdot \mathbf{v}). \tag{4.7}$$

The viscous stress tensor in the gas-entropy equation is

$$\boldsymbol{\Pi} = -\Delta p\Big(\hat{\boldsymbol{b}}\hat{\boldsymbol{b}} - \frac{\boldsymbol{\mathcal{I}}}{\mathbf{3}}\Big). \tag{4.8}$$

In the absence of background $\Delta p$, the perturbed $\nabla \cdot (\boldsymbol{\Pi} \cdot \mathbf{v})$ in the gas-entropy equation is second-order and does not contribute in our linear analysis. $\boldsymbol{Q}$ in equation (4.4) is the anisotropic thermal heat flux,

$$\boldsymbol{Q} = -\kappa_\mathrm{B}\hat{\boldsymbol{b}}\hat{\boldsymbol{b}} \cdot \nabla T, \tag{4.9}$$

where $\kappa_\mathrm{B}$ is the thermal conductivity.

$\mathrm{v}_{\mathrm{st},\Delta p}$ in equation (4.5) is the CR streaming speed. We assume that cosmic rays stream down their pressure gradient at the Alfvén speed, which in low-collisionality plasmas depends on the thermal-gas pressure anisotropy,[1]

$$\mathbf{v}_{\mathbf{st},\boldsymbol{\Delta p}} = \chi\frac{\boldsymbol{B}}{\sqrt{4\pi\rho}}\Big(1 + \frac{4\pi\Delta p}{B^2}\Big)^{1/2}, \tag{4.10}$$

where $\chi \equiv -\hat{\boldsymbol{b}} \cdot \nabla p_c/|\hat{\boldsymbol{b}} \cdot \nabla p_c| = \pm 1$ ensures that the cosmic rays stream along $\boldsymbol{B}$ down their pressure gradient and makes the CR heating term $-\mathbf{v}_\mathbf{st} \cdot \nabla p_c$ in the gas energy equation (4.4) positive definite. We also include CR diffusion along the magnetic field in eq. (4.5).

### 4.3.2 Dimensionless Parameters and Characteristic Frequencies

We define the ratio of CR pressure to gas pressure,

$$\eta \equiv \frac{p_c}{p_g}, \tag{4.11}$$

and the ratio of thermal to magnetic pressure,

$$\beta \equiv \frac{8\pi p_g}{B^2}. \tag{4.12}$$

---

[1] We note that the original expression for the Alfvén speed in Kempski et al. (2020) was incorrect, as the 1/2 exponent in the $\Delta p$ term in (4.10) was by accident omitted. This was corrected in Kempski et al. (2021). However, the conclusions of Kempski et al. (2020) are not affected by this change.

The relevant frequencies are the gas sound frequency (with $c_s$ being the isothermal gas sound speed),

$$\omega_s \equiv k c_s; \tag{4.13}$$

the Alfvén *and* CR-streaming frequency,

$$\omega_{\rm A} \equiv \boldsymbol{k} \cdot \mathbf{v_A}; \tag{4.14}$$

the CR diffusion frequency,

$$\omega_d \equiv \kappa \, (\hat{\boldsymbol{b}} \cdot \boldsymbol{k})^2; \tag{4.15}$$

the free-fall frequency,

$$\omega_{\rm ff} \equiv \frac{g}{c_s} = \frac{c_s}{H}, \tag{4.16}$$

where $H$ is the gas scale height; the cooling frequency,

$$\omega_c \equiv \frac{\rho^2 \Lambda}{p_g}; \tag{4.17}$$

the ion-ion collsion frequency $\nu_{\rm ii}$ and the associated Braginskii viscous frequency,

$$\omega_{\rm B} \equiv \nu_{\rm B} (\hat{\boldsymbol{b}} \cdot \boldsymbol{k})^2 \approx \frac{p_g}{\rho \nu_{\rm ii}} (\hat{\boldsymbol{b}} \cdot \boldsymbol{k})^2; \tag{4.18}$$

and the conductive frequency,

$$\omega_{\rm cond} \equiv \chi_{\rm B} (\hat{\boldsymbol{b}} \cdot \boldsymbol{k})^2, \tag{4.19}$$

where $\chi_{\rm B} = \kappa_{\rm B}/n k_{\rm B}$ is the thermal diffusion coefficient. We can relate the diffusive timescales by defining the thermal Prandtl number,

$$\mathrm{Pr} \equiv \frac{\nu_{\rm B}}{\chi_{\rm B}}. \tag{4.20}$$

We use $\mathrm{Pr} = 0.02$ as the default thermal Prandtl number in this work ($\approx$ square root of the electron-to-ion mass ratio), i.e. heat conduction due to electrons operates on a much shorter timescale than viscous forces due to the ions. We also define the ratio of the CR diffusion coefficient to the Braginskii viscosity,

$$\Phi \equiv \frac{\kappa}{\nu_{\rm B}}, \tag{4.21}$$

which turns out to be an important parameter quantifying the suppression of the CRBI and CRAB instability by CR diffusion (see also Kempski et al. 2020).

### 4.3.3 Validity of the Model

#### 4.3.3.1 Thermal Gas

The CR entropy mode describes the response of the two-fluid CR–thermal gas system to a CR pressure perturbation. Since CRs propagate along field lines at the Alfvén speed, this mode is characterised by a frequency of order $\omega_{\rm A}$ (absent CR diffusion). The collisional (Braginskii MHD) regime then requires that $\omega_{\rm A} \ll \nu_{\rm ii}$. To ignore the effect of heat fluxes in the calculation of the pressure anisotropy we also require that $\omega_{\rm A} \lesssim \nu_{\rm ii}\beta^{-1/2}$. The Braginskii MHD model is thus valid for

$$\omega_{\rm A} \lesssim \nu_{\rm ii}\beta^{-1/2} \Longleftrightarrow \frac{\omega_{\rm B}}{\omega_{\rm A}} \lesssim \beta^{1/2}, \tag{4.22}$$

or equivalently

$$kl_{\rm mfp} \lesssim 1. \tag{4.23}$$

In the dilute and hot ICM, the ion-ion collision rate is small

$$\nu_{\rm ii} \sim \frac{n_{\rm i}e^4\pi \ln\Lambda}{m_{\rm i}^{1/2}(k_{\rm B}T)^{3/2}} \sim 8\times10^{-14}\ {\rm s}^{-1}\left(\frac{T}{5\times10^7\ {\rm K}}\right)^{-3/2}\frac{n_{\rm i}}{0.01\ {\rm cm}^{-3}}, \tag{4.24}$$

for a Coulomb logarithm $\ln\Lambda \approx 38$. The collision rate in (4.24) corresponds to a mean free path of order 0.1 kpc.

#### 4.3.3.2 Cosmic Rays

The CR pressure equation (eq. 4.5) is a good model for the cosmic rays if the collision frequency of the energetically important GeV CRs is much larger than any other timescale of interest. As pointed out in Kempski et al. (2020), the GeV CR collision frequency (the rate at which the pitch angle changes by order unity, due to scattering by magnetic fluctuations) is likely much higher than the thermal ion-ion collision frequency in the ICM:

$$\nu_{\rm CR} \sim \Omega_0\left(\frac{\delta B_\perp}{B}\right)^2 \sim 10^{-8}\ {\rm s}^{-1}\left(\frac{\delta B_\perp/B}{10^{-3}}\right)^2, \tag{4.25}$$

where $\Omega_0$ is the non-relativistic gyro-frequency and $\delta B_\perp/B$ is evaluated for fluctuations whose wavelength parallel to the mean B-field is of order the Larmor radius of the GeV particles. The above collision frequency corresponds to a CR mean free path of order 1 pc, which approximately corresponds to the empirically derived CR mean free path in the Milky Way (e.g., Amato & Blasi 2018). This suggests that treating cosmic rays as collisional on thermal-ion-mean-free-path scales is a reasonable model for the ICM.

In the limit of good coupling between the GeV cosmic rays and the self-excited Alfvén waves (large pitch-angle scattering rate), CR transport is to leading order described by Alfvénic streaming. Damping of the self-excited Alfvén waves introduces corrections to Alfvénic streaming. We model this by also including CR diffusion in our linear analysis. As we will show, significant CR diffusion suppresses the CRBI on small scales.

### 4.3.4 Background Equilibrium

We consider static background equilibria with constant $\boldsymbol{B} = (B_x, 0, B_z)$, in which the CR and gas pressure gradients balance gravity, $\boldsymbol{g} = -g\hat{\boldsymbol{z}}$,

$$\frac{d}{dz}(p_g + p_c) = -\rho g. \tag{4.26}$$

If CR diffusion does not affect the background state (either because $\kappa = 0$ or $dp_c/dz = \text{const}$), CRs are in equilibrium according to (4.5) if $p_c \propto \mathrm{v_A}^{-4/3}$, or equivalently,

$$p_c \propto \rho^{2/3}. \tag{4.27}$$

In this work we perturb equilibria that satisfy (4.26) and (4.27), although these equations do not yet constrain the background temperature $T(z)$. We will consider different temperature profiles in Sections 4.6 and 4.7.

### 4.3.5 Linearised Equations

We carry out a linear stability calculation of the CR–thermal gas equations (Section 4.3.1) for backgrounds that satisfy (4.26) and (4.27). We focus on short-wavelength modes satisfying $kH \gg 1$ and perform a local WKB calculation in which all perturbed quantities are assumed to vary as $\delta X(\boldsymbol{r}, t) \propto \exp\left(i\boldsymbol{k}\cdot\boldsymbol{r} - i\omega t\right)$. We consider two coordinate systems in this work. We will mostly use the usual cartesian coordinate system with the $z$-axis anti-parallel to $\boldsymbol{g}$ and $x$ and $y$ defining the plane perpendicular to gravity. We define associated polar angles such that $B_x = B\sin\theta_\mathrm{B}$, $B_z = B\cos\theta_\mathrm{B}$, $k_x = k\sin\theta_k\cos\phi_k$, $k_y = k\sin\theta_k\sin\phi_k$ and $k_z = k\cos\theta_k$. We use $\phi_k = 0$ in all the figures except Figure 4.5, because $\phi_k = 0$ captures the fastest-growing mode. For analytic purposes, it is convenient to also use a coordinate system aligned with the direction of the magnetic field. We define $\perp$ and $||$ to denote the directions perpendicular and parallel to the magnetic field. We will use this coordinate system in Section 4.5 when we derive an approximate dispersion relation, as it makes the analytics more tractable. However, when we plot numerically calculated growth rates of the instability in Figures 4.2–4.9, we adopt the more standard notation with directions/angles defined relative to the positive z-direction, i.e. $\cos\theta_B \equiv -\hat{\boldsymbol{b}}\cdot\hat{\boldsymbol{g}}$ and $\cos\theta_k \equiv -\hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{g}}$.

The linearised equations are

$$\omega\frac{\delta\rho}{\rho} = \boldsymbol{k}\cdot\mathbf{v} - i\mathrm{v}_z\frac{d\ln\rho}{dz}\times\mathrm{BG}, \tag{4.28}$$

$$\omega\mathbf{v} = \boldsymbol{k}c_s^2\Big(\frac{\delta p_\perp}{p_g} + \eta\frac{\delta p_c}{p_c} + \frac{\mathrm{v_A^2}}{c_s^2}\frac{\delta B_{||}}{B}\Big) - \omega_\mathrm{A}\mathrm{v_A}\frac{\boldsymbol{\delta B}}{B} - \hat{\boldsymbol{b}}(\hat{\boldsymbol{b}}\cdot\boldsymbol{k})c_s^2\frac{\delta\Delta p}{p_g} + i\boldsymbol{g}\frac{\delta\rho}{\rho}, \tag{4.29}$$

$$\omega\frac{\boldsymbol{\delta B}_\perp}{B} = -k_{||}\mathbf{v}_\perp, \tag{4.30}$$

$$\omega\frac{\delta B_{||}}{B} = \boldsymbol{k}_\perp\cdot\mathbf{v}_\perp, \tag{4.31}$$

$$\begin{aligned}\omega\frac{\delta p_g}{p_g} =& \gamma\boldsymbol{k}\cdot\mathbf{v} - i(\gamma-1)\omega_\mathrm{cond}\Big(\frac{\delta p_g}{p_g} - \frac{\delta\rho}{\rho}\Big) + \chi\eta(\gamma-1)\omega_\mathrm{A}\frac{\delta p_c}{p_c} - i\omega_c\frac{\partial\ln\Lambda}{\partial\ln T}\frac{\delta p_g}{p_g} - i\omega_c\Big(2 - \frac{\partial\ln\Lambda}{\partial\ln T}\Big)\frac{\delta\rho}{\rho}\\ &+ \mathrm{BG}\times\Big[-i\mathrm{v}_z\frac{d\ln p_g}{dz} - i\chi\eta(\gamma-1)\delta\mathrm{v}_{\mathrm{A},z}\frac{d\ln p_c}{dz} + \frac{i(\gamma-1)}{T}\boldsymbol{\nabla}\cdot(\chi_\mathrm{B}\boldsymbol{\delta}\hat{\mathbf{b}}\hat{\mathbf{b}}\cdot\boldsymbol{\nabla}T + \chi_\mathrm{B}\hat{\mathbf{b}}\boldsymbol{\delta}\hat{\mathbf{b}}\cdot\boldsymbol{\nabla}T)\Big],\end{aligned} \tag{4.32}$$

$$\frac{\delta \Delta p}{p_g} = i\frac{\rho \nu_{\mathrm{B}}}{p_g}(2k_{\parallel}\mathrm{v}_{\parallel} - \boldsymbol{k}_{\perp} \cdot \mathbf{v}_{\perp}). \tag{4.33}$$

$$\begin{aligned}\omega\frac{\delta p_c}{p_c} =& \frac{4}{3}\boldsymbol{k} \cdot \mathbf{v} - \frac{2}{3}\chi\omega_{\mathrm{A}}\frac{\delta\rho}{\rho} + \frac{1}{3}\beta\chi\omega_{\mathrm{A}}\frac{\delta\Delta p}{p_g} + (\chi\omega_{\mathrm{A}} - i\omega_d)\frac{\delta p_c}{p_c} + \mathrm{BG} \times \Big[ - i\mathrm{v}_z\frac{d\ln p_c}{dz} \\ &+ \frac{2}{3}i\chi\mathrm{v}_{\mathrm{A,z}}\frac{d\ln\rho}{dz}\Big(\frac{\delta p_c}{p_c} - \frac{\delta\rho}{\rho}\Big) + \frac{i}{p_c}\nabla\cdot\Big(\kappa\boldsymbol{\delta}\hat{\mathbf{b}}\hat{\mathbf{b}} \cdot \boldsymbol{\nabla}p_c + \kappa\hat{\mathbf{b}}\boldsymbol{\delta}\hat{\mathbf{b}} \cdot \boldsymbol{\nabla}p_c\Big)\Big],\end{aligned} \tag{4.34}$$

where BG in equations (4.28)–(4.34) is a flag equal to 1 or 0. It specifies whether gradients from the background equilibrium in (4.26) and (4.27), i.e. terms of the form $\mathrm{v}_z d\rho/dz$, are included in the calculation. Without the background gradients equilibrium is not satisfied and so the BG terms should formally be kept in the linear perturbation analysis. However, for analytic simplicity, we set BG = 0 when we discuss the physics of the instability and derive approximate dispersion relations in Sections 4.4.2, 4.4.3 and 4.5. This turns out to be a reasonable approximation, as the CRBI is not directly driven by the background gradients. We do, however, include background gradients, i.e. we set BG = 1, when we solve for the exact eigenmodes numerically (Section 4.6 and all the figures presented in this paper).

## 4.4 Physical Overview of the Instability

In this section, we provide an overview of the key physics describing the CRBI. To start, we give a brief summary of entropy and gravity waves in stratified CR MHD (4.4.1). This will set the stage for our analysis and elucidate how the CRBI is related to other instabilities that may operate in the ICM (such as thermal instability or other buoyancy instabilities).

### 4.4.1 Gravity and Entropy Modes in Stratified Media

#### 4.4.1.1 MHD Modes

In a gravitationally stratified medium without magnetic fields, the hydrodynamic entropy mode becomes an internal gravity wave characterised by the Brunt–Väisälä frequency,

$$N^2 = \frac{g}{\gamma}\Big(\frac{d\ln p_g}{dz} - \gamma\frac{d\ln\rho}{dz}\Big). \tag{4.35}$$

In MHD, gravitational stratification affects the MHD slow magnetosonic modes, as both buoyancy and magnetic tension act as the modes' restoring forces. The resulting mode resembles the hydro IGW at long wavelengths, and the standard MHD slow mode (absent gravity) at short wavelengths. In contrast to hydrodynamics, in MHD there is also a mode that is unaffected by buoyancy despite a non-zero density perturbation. We refer to this mode as the MHD entropy mode.

We now compute these modes more explicitly in high-$\beta$ collisional MHD, first assuming that no cosmic rays are present. For simplicity, we shall consider the 2D case with $\boldsymbol{B} = B\hat{\boldsymbol{z}}$, $\boldsymbol{g} = -g\hat{\boldsymbol{z}}$

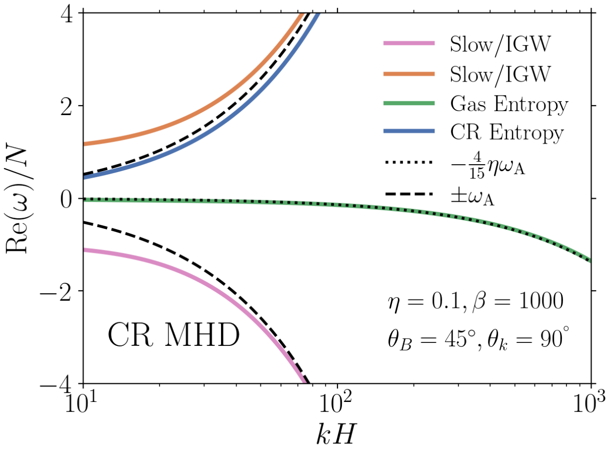


*Figure 4.1*: Oscillation frequencies of gravity/slow and entropy modes in stratified, collisional CR MHD. The frequencies are normalised by the Brunt–Väisälä frequency. Here and in other figures (except Figure 4.5), we use $\phi_k = 0$. The two modes that are characterised by the buoyancy frequency at long wavelengths (orange and pink lines) become the MHD slow magnetosonic modes at short wavelengths. The oscillation frequency of the gas-entropy mode (green line) is due to CR heating. The blue line shows the CR-entropy mode with frequency $\approx \omega_{\rm A}$, with the small deviation arising from finite $\eta$. We will show that gravity destabilises the CR-entropy mode in low-collisionality MHD. We term this instability the CR buoyancy instability (CRBI).

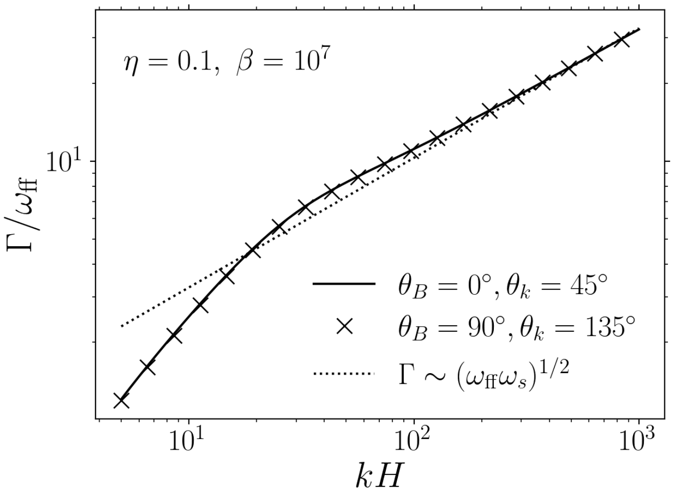


*Figure 4.2*: Growth rates of the CRBI for $\beta = 10^7$, $\eta = 0.1$, and $H = 1000 l_{\rm mfp}$. The angles $\theta_B$ and $\theta_k$ are the directions of $\boldsymbol{B}$ and $\boldsymbol{k}$ with respect to the positive $z$-direction ($\boldsymbol{g} = -g\hat{\boldsymbol{z}}$). At high $\beta$, damping by anisotropic pressure is negligible over a range of $k$ and the instability is well described by the simple model in Section 4.4.3. The dotted line labelled $\sim \sqrt{\omega_s \omega_{\rm ff}}$ shows the predicted growth rate from eq. (4.48). This elucidates the physics of the CRBI in its simplest form. Figures 4.3 – 4.9 show results for more realistic ICM conditions.

and $\boldsymbol{k}$ in the $x-z$ plane. We start with the momentum equation,

$$\omega \mathbf{v} = \boldsymbol{k}\Big(\frac{\delta p_g}{p_g} + \frac{B^2}{4\pi}\frac{\delta B}{B}\Big) - \omega_{\mathrm{A}} \mathrm{v}_{\mathrm{A}} \frac{\boldsymbol{\delta B}}{B} - i\frac{\delta\rho}{\rho} g\hat{\boldsymbol{z}}. \tag{4.36}$$

Crossing the momentum equation twice with $\boldsymbol{k}$ and taking the z-component we find,

$$\omega^2 (k_z k_x \mathrm{v}_x - k_x^2 \mathrm{v}_z) = \omega_{\mathrm{A}} \mathrm{v}_{\mathrm{A}} k^2 k_x \mathrm{v}_x + i\omega g k_x^2 \frac{\delta\rho}{\rho}. \tag{4.37}$$

In the Boussinesq limit ($\delta p_g/p_g \ll \delta\rho/\rho$), the adiabatic gas entropy equation $ds/dt = 0$ (eq. 4.32 without CRs, cooling and conduction) implies that

$$\omega\frac{\delta\rho}{\rho} = \frac{i}{\gamma}\mathrm{v}_z\Big(\frac{d\ln p_g}{dz} - \gamma\frac{d\ln\rho}{dz}\Big) = i\frac{\mathrm{v}_z}{g}N^2 \tag{4.38}$$

Using incompressibility $k_x \mathrm{v}_x \approx -k_z \mathrm{v}_z$ we find the dispersion relation for MHD slow modes modified by gravity,

$$\omega^2 k^2 = \omega_{\mathrm{A}} \mathrm{v}_{\mathrm{A}} k^2 k_z + g k_x^2 \frac{\gamma - 1}{\gamma}\frac{d\ln s}{dz}, \tag{4.39}$$

or,

$$\omega^2 = \omega_{\mathrm{A}}^2 + \frac{k_x^2}{k^2}N^2. \tag{4.40}$$

For subdominant magnetic tension (first term on RHS) we get the usual dispersion relation for the two hydrodynamic IGWs, $\omega = \pm N k_x/k$. At short wavelengths where magnetic tension dominates, the two gravity waves satisfy the standard dispersion relation for the MHD slow modes at high $\beta$.

Unlike the hydrodynamic entropy mode, the MHD entropy mode does not pick up a buoyancy response. To see why this is the case, we first note that a mode with $\omega = 0$ and $\mathrm{v} = 0$, but finite $\delta\rho/\rho$, satisfies the continuity (eq. 4.28), induction (eq. 4.30–4.31), as well as the gas-entropy (eq. 4.32 absent heating, cooling and conduction) equations. In hydrodynamics, however, this mode does not satisfy the momentum equation (see eq. 4.36 without the perturbed magnetic field), as the direction of gravity and $\boldsymbol{k}$ are generally not co-linear. By contrast, in MHD the perturbed pressure and magnetic-tension terms (first and second terms on the RHS of eq. 4.36) are mutually orthogonal and can exactly cancel the perturbed gravitational force. As a result, while a mode with $\omega = 0$ and $\mathrm{v} = 0$, but finite $\delta\rho/\rho$, is not an eigenmode in stratified hydrodynamics, it is an eigenmode in stratified MHD and involves a finite $\boldsymbol{\delta B}$ perturbation. When magnetic tension is negligible, i.e. in the hydrodynamic limit with $\omega_{\mathrm{A}} \ll N$ (long wavelengths or $\beta \to \infty$), the $\omega = 0$ mode can only be satisfied if $\delta B/B \gg \delta\rho/\rho$. The hydrodynamic variables are therefore essentially unperturbed in this limit, which is consistent with the result in stratified hydrodynamics.

#### 4.4.1.2 CR MHD Modes

When Alfvénically streaming CRs are present, the CR pressure equation (eq. 4.5) introduces a new mode, which we refer to as the CR entropy mode. Because CRs are assumed to stream at the

Alfvén speed along field lines, the CR entropy mode is characterised by the Alfvén frequency, $\omega = \chi \boldsymbol{k} \cdot \mathbf{v_A} = \chi\omega_{\rm A}$ (although the eigenfrequency can deviate appreciably from the Alfvén frequency if CR diffusion is important or if $\eta \gtrsim 1$ due to the fact that the CR entropy mode is then associated with significant density fluctuations). The $\chi$ factor in front of $\omega_{\rm A}$ reflects the fact that the CR entropy mode propagates down the CR pressure gradient. For $\eta \to 0$ the impact of CRs on the thermal gas is small, and so the CR entropy mode does not perturb the thermal gas. Using that $2d\ln\rho/dz = 3d\ln p_c/dz$ in equilibrium, (4.34) becomes,

$$\left(\omega - \chi\omega_{\rm A} + i\omega_d - i\chi {\rm v}_{{\rm A},z}\frac{d\ln p_c}{dz}\right)\frac{\delta p_c}{p_c} \approx 0, \tag{4.41}$$

with solution,

$$\omega = \chi\omega_{\rm A} - i\omega_d + i\chi {\rm v}_{{\rm A},z}\frac{d\ln p_c}{dz}. \tag{4.42}$$

Diffusive corrections to CR streaming act to damp the mode. The CR background gradient term also introduces an imaginary part that looks like damping. However, the more accurate interpretation of this term is that as the mode propagates down the CR pressure gradient, the perturbation amplitude normalized by the local CR pressure, $\delta p_c/p_c(z)$, remains constant.

Because streaming CRs heat the gas at a rate $-\chi\mathbf{v_A}\cdot\boldsymbol{\nabla} p_c$, they also modify the gas-entropy mode by giving it a real (oscillatory) frequency, which at small $\eta$ is $\omega \approx -4\eta\omega_{\rm A}/15$ (Kempski & Quataert 2020). Importantly, to leading order CR heating does not significantly affect the growth/damping of the gas-entropy mode (e.g., due to thermal instability), just its real frequency.

We summarise the discussion above by showing the oscillation frequencies of gravity and entropy modes in stratified, collisional CR MHD in Figure 4.1. The frequencies are normalised by the Brunt–Väisälä frequency. The two modes that are characterised by the buoyancy frequency at long wavelengths (orange and pink lines) become the MHD slow magnetosonic modes at short wavelengths. The gas-entropy mode (green line) is unaffected by buoyancy, and its oscillation frequency is due to CR heating. The blue line shows the CR-entropy mode with frequency $\approx \omega_{\rm A}$. Figure 4.1 suggests that the CR-entropy mode is not significantly affected by buoyancy in collisional MHD. However, we will show that gravity destabilises the mode in low-collisionality MHD (i.e., on small scales), which we term the CR buoyancy instability (CRBI).

### 4.4.2 Compressible CR Entropy Mode due to Streaming

Surprisingly, the CR entropy mode becomes highly compressible at short wavelengths due to the influence of a finite mean free path in the background plasma, i.e. due to $\Delta p$. At high $\beta$, pressure balance implies $\delta p_c + \delta p_\perp \approx 0$. In the isothermal limit due to rapid conduction (Pr $\ll$ 1), and neglecting CR diffusion, this can be rewritten as (see eq. 4.34):

$$\frac{\delta\rho}{\rho} + \eta\frac{\frac{4}{3}\boldsymbol{k}\cdot\mathbf{v} - \frac{2}{3}\chi\omega_{\rm A}\frac{\delta\rho}{\rho} + \frac{1}{3}\beta\omega_{\rm A}\chi\frac{\delta\Delta p}{p_g}}{\omega - \chi\omega_{\rm A}} \approx 0. \tag{4.43}$$

In the limit of long wavelengths/high collisionality such that $\delta\Delta p$ is negligible, pressure balance is achieved if $\delta\rho/\rho \to 0$, i.e. $k_x{\rm v}_x + k_y{\rm v}_y + k_z{\rm v}_z = 0$ or $\boldsymbol{k}_\perp\cdot\mathbf{v}_\perp + k_\parallel{\rm v}_\parallel = 0$. This is the standard result that pressure-balanced modes are nearly incompressible at high $\beta$.

At shorter wavelengths and high $\beta$, the $\delta\Delta p/p_g$ term is dominant for the CR entropy mode with $\omega \approx \omega_{\mathrm{A}}$. Pressure balance then requires that the pressure anisotropy is minimised, $\delta\Delta p \approx 0$. This condition leads to the unusual requirement that the pressure-balanced mode is highly compressible. In Braginskii MHD $\Delta p \propto (2k_{\|}\mathrm{v}_{\|} - \boldsymbol{k}_{\perp} \cdot \mathbf{v}_{\perp})$ so that (4.43) is satisfied when,

$$\boldsymbol{k}_{\perp} \cdot \mathbf{v}_{\perp} \approx 2k_{\|}\mathrm{v}_{\|}. \tag{4.44}$$

The exact relation satisfied by $\mathrm{v}_{\perp}$ and $\mathrm{v}_{\|}$ can be different in the collisionless regime, i.e. below the ion mean free path, where Braginskii MHD is no longer valid. However, the qualitative conclusion remains the same and the mode is very compressible. We next show how the pressure-balanced compressible mode is destabilised by gravity.

### 4.4.3 How Gravity Destabilises the Compressible CR Entropy Mode

To show how gravity destabilises the compressible CR entropy mode discussed above, we consider a simple model in which we ignore magnetic tension and damping by anisotropic pressure (viscosity). As in Section 4.4.1, we consider the case $\boldsymbol{B} = B\hat{\boldsymbol{z}}$, $\boldsymbol{g} = -g\hat{\boldsymbol{z}}$, and consider a mode with $k_x\mathrm{v}_x = \alpha k_z\mathrm{v}_z$ imposed by pressure balance. As described in Section 4.4.2, $\alpha = -1$ in standard MHD, while $\alpha = 2$ in Braginskii MHD with CRs. The momentum equations are,

$$\omega\mathrm{v}_x = \frac{k_x}{\rho}\delta P_{\mathrm{tot}}, \tag{4.45}$$

$$\omega\mathrm{v}_z = \frac{k_z}{\rho}\delta P_{\mathrm{tot}} - i\omega_{\mathrm{ff}}c_s\frac{(1+\alpha)k_z\mathrm{v}_z}{\omega}, \tag{4.46}$$

where in the $\mathrm{v}_z$ equation we used that $\omega\delta\rho/\rho \approx (1+\alpha)k_z\mathrm{v}_z$. Multiplying (4.45) by $k_x$ and subtracting from (4.46) times $\alpha k_z$, we find that

$$\frac{\delta P_{\mathrm{tot}}}{\rho}(\alpha k_z^2 - k_x^2) = i\alpha(1+\alpha)k_z c_s\omega_{\mathrm{ff}}\frac{k_z\mathrm{v}_z}{\omega}. \tag{4.47}$$

Using this expression for $\delta P_{\mathrm{tot}}$ back in equations (4.45) and (4.46) gives a simple dispersion relation:

$$\omega \approx \frac{1}{\sqrt{2}}(1+i)\Big[(1+\alpha)k_z c_s\omega_{\mathrm{ff}}\frac{kx^2}{\alpha k_z^2 - k_x^2}\Big]^{1/2} \sim \sqrt{\omega_s\omega_{\mathrm{ff}}}. \tag{4.48}$$

In this simplified picture gravity leads to growth rates that are of order $\sim \sqrt{\omega_{\mathrm{ff}}\omega_s} \gg \omega_{\mathrm{ff}}$. The above analysis can be easily repeated for the case $\boldsymbol{g} \perp \boldsymbol{B}$ (i.e. horizontal magnetic field in a vertical gravitational field), with very similar results. We stress that the scaling $\sim \sqrt{\omega_{\mathrm{ff}}\omega_s}$ in (4.48) does not describe compressible sound waves, which are approximately longitudinal at high $\beta$: $\alpha \approx k_x^2/k_z^2$ and the denominator in (4.48) is approximately zero.

Akin to standard buoyancy instabilities, the mode found here is destabilised by the unbalanced gravitational force acting on the mode's density fluctuations. However, in contrast to standard buoyancy instabilities, such as thermal convection in stars or the magneto-thermal instability (MTI;

Balbus 2000) and the heat-flux-driven buoyancy instability (HBI; Quataert 2008) in clusters, the density fluctuations in the CR-driven instability are not due to the background stratification of the plasma. The density fluctuations at short wavelengths are instead due to CR streaming, independent of the background stratification.

In subsequent sections we show that while magnetic tension does not affect the growth rate significantly, the effect of damping by anisotropic pressure should generally be retained. However, at sufficiently high $\beta$ there is a range of scales for which the simple model considered in this section provides a good picture of the instability (see eq. 4.52 below), and the growth rates are indeed $\sim \sqrt{\omega_s \omega_{\rm ff}}$. This regime is shown in Figure 4.2. The solid black line shows the instability growth rate, computed from numerical solutions of the full set of linear equations in 4.3.5, as a function of wavenumber for $\beta = 10^7$, $\eta = 0.1$, $\boldsymbol{B}$ antiparallel to $\boldsymbol{g}$ and $k_x = k_z$. The 'x' markers show the growth rate for $\boldsymbol{B}$ perpendicular to $\boldsymbol{g}$ and $k_x = -k_z$. The dotted line labelled $\sim \sqrt{\omega_s \omega_{\rm ff}}$ shows the predicted growth rate from (4.48), which matches the exact solution very well over a wide range in $k$.

We now discuss the instability in the astrophysically more relevant regimes where (4.48) is not as accurate.

## 4.5 Dispersion Relation and Growth Rates of Short-Wavelength Modes

In 4.4.3 we neglected damping by anisotropic pressure (viscosity), which generally changes the growth rates relative to those predicted in (4.48) (except for certain asymptotic limits, e.g. very large $\beta$ as in Figure 4.2). In this section we present a more accurate calculation and derive an approximate dispersion relation.

For simplicity, we here ignore background gradient terms (i.e. we set BG = 0 in equations 4.28–4.34) in our analytic derivation of the growth rates. Ignoring background gradients is a reasonable simplification because we find that short-wavelength modes (which are the fastest growing modes) are not significantly affected by explicitly including background gradients (even though they should formally be included). This is because the CRBI is due to streaming-induced compressibility and not background stratification.[2] This is in contrast to thermal convection in stars or the MTI/HBI in clusters, which are driven by heat conduction and background temperature gradients. Our analytic results from this section will be supplemented by numerical solutions of the full system of linearised equations (equations 4.28-4.34) including background gradients in Section 4.6.

---

[2] However, we note that a nonzero CR pressure gradient along the magnetic field is generally necessary for equilibrium and to couple CRs to the thermal gas.

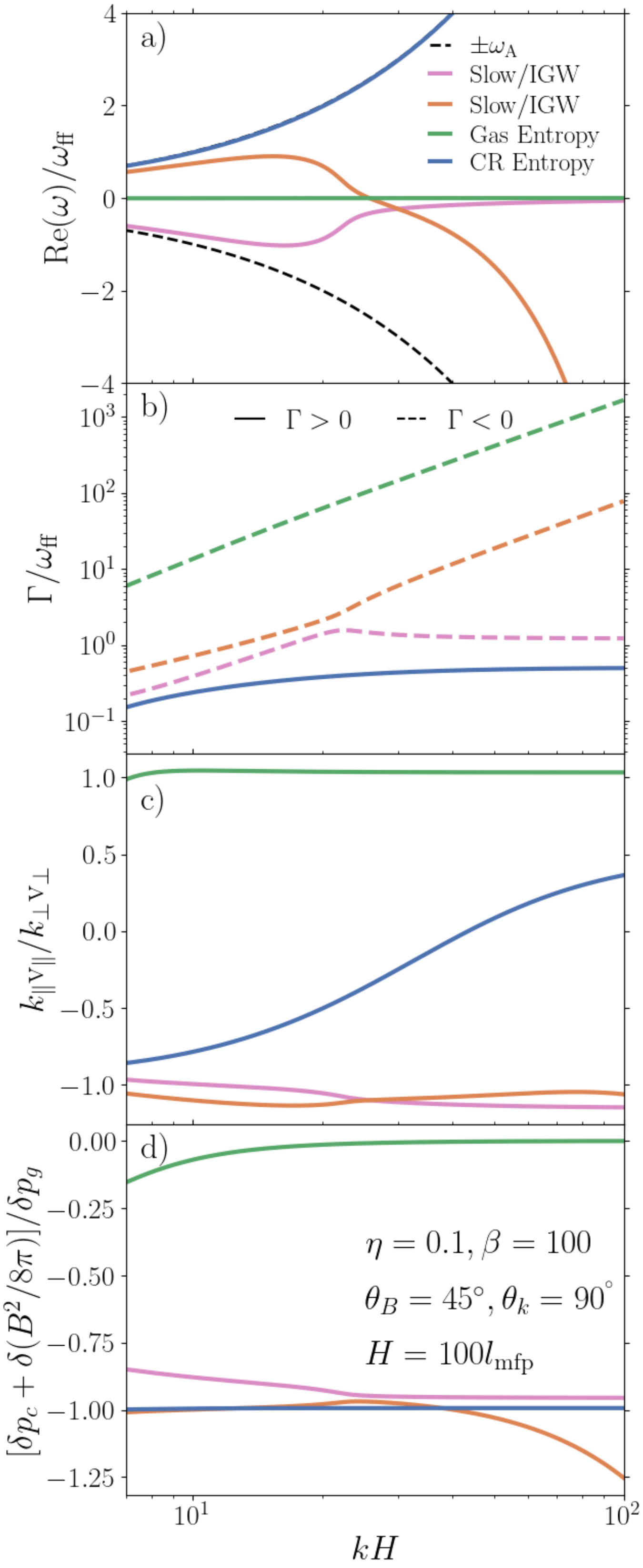


*Figure 4.3*: Mode properties for $\eta = 0.1$, $\beta = 100$ and $H = 100 l_{\rm mfp}$. $\theta_B = 45^\circ$ and $\theta_k = 90^\circ$ are the directions of $\boldsymbol{B}$ and $\boldsymbol{k}$ with respect to $\hat{\boldsymbol{z}}$ ($\boldsymbol{g} = -g\hat{\boldsymbol{z}}$), while $\perp$ and $\parallel$ are defined w.r.t. $\boldsymbol{B}$. We plot oscillation frequencies in panel a), growth rates in b) ($\Gamma > 0$ corresponds to growth), $k_\parallel v_\parallel / k_\perp v_\perp$, which quantifies the compressibility of the mode, in c), and $[\delta p_c + \delta(B^2/8\pi)]/\delta p_g$, which quantifies the degree of pressure balance, in d). The blue line shows the unstable CR entropy mode, the orange and pink lines are the MHD slow modes and the green line is the gas-entropy mode. All modes except the CR entropy mode are damped by low-collisionality effects. Long-wavelength CR entropy modes are approximately incompressible and the growth rates are small. High-$k$ CR entropy modes become compressible due to CR streaming while maintaining pressure balance, and the growth rate reaches the plateau given by eq. (4.54).

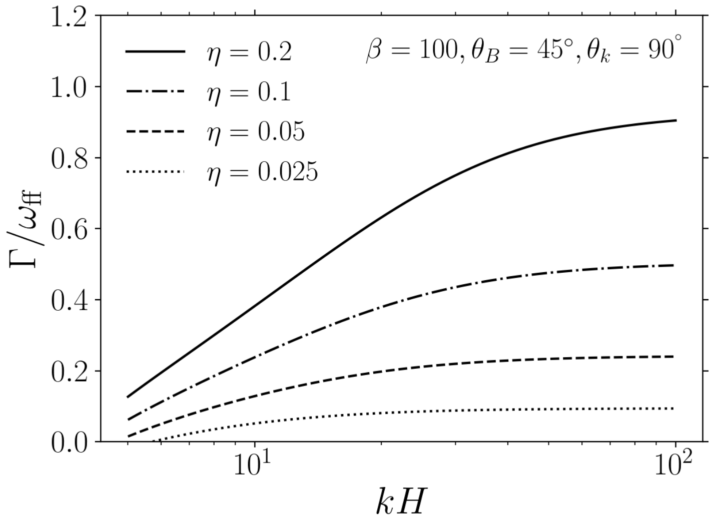


*Figure 4.4*: Growth rates of the CRBI for $\beta = 100$, $H = 100 l_{\rm mfp}$ and different ratios of CR pressure to gas pressure, $\eta$. The other parameters are set to the fiducial values (Section 4.6). The angles $\theta_B = 45°$ and $\theta_k = 90°$ are the directions of $\boldsymbol{B}$ and $\boldsymbol{k}$ with respect to the positive $z$-direction ($\boldsymbol{g} = -g\hat{\boldsymbol{z}}$). Even at very small CR pressures (small $\eta$), the instability still exists, but with reduced growth rates (eq. 4.54).

### 4.5.1 Dispersion Relation

In this section, it is convenient to work in a coordinate system aligned with the magnetic field: $\boldsymbol{B} = (0, 0, B)$, $\boldsymbol{k} = (k_\perp, 0, k_\parallel)$ and $\boldsymbol{g} = (g_{\perp;1}, g_{\perp;2}, g_\parallel)$. In this coordinate system, and if we neglect background-gradient terms (which we do here), Alfvénic fluctuations of the form $\boldsymbol{\delta B} = (0, \delta B, 0)$ and $\boldsymbol{\delta \mathrm{v}} = (0, \delta \mathrm{v}, 0)$ decouple, which leaves velocity/magnetic field fluctuations in the $\boldsymbol{k} - \boldsymbol{B}$ plane for the remaining modes. If background gradients are kept, this is strictly true only if gravity is coplanar with $\boldsymbol{k}$ and $\boldsymbol{B}$. In the high-$\beta$ and isothermal limit ($\omega_{\rm cond} \gg \omega$), the third order dispersion relation for the slow-magnetosonic and CR-entropy modes can be found by crossing the momentum equation (4.29) twice with $\boldsymbol{k}$, taking the component parallel to $\boldsymbol{B}$ and using (4.43):

$$\begin{aligned} 0 = {} & \omega^2\Big(\frac{2}{3}i\chi\eta\frac{\omega_B}{\omega_{\rm A}}\omega(2k_\parallel^2 - k_\perp^2) + k^2\Delta\Big) + 3ik_\perp^2\omega_B\omega\Delta \\ & + 2\chi\eta\frac{\omega_B}{\omega_{\rm A}}k_\parallel\omega_{\rm ff}c_s\omega\Big(-k^2(\hat{\boldsymbol{b}}\cdot\hat{\boldsymbol{g}}) + (\boldsymbol{k}\cdot\hat{\boldsymbol{g}})k_\parallel\Big) \\ & - \omega_{\rm A}\mathrm{v}_{\rm A}k^2k_\parallel\Big(\frac{4}{3}i\eta\chi\frac{\omega_B}{\omega_{\rm A}}\omega + \Delta\Big), \end{aligned} \tag{4.49}$$

where,

$$\Delta = \omega - \chi\omega_{\rm A} + i\omega_d + \frac{4}{3}\eta\omega - \frac{2}{3}\eta\chi\omega_{\rm A}. \tag{4.50}$$

The 3 solutions of the cubic dispersion relation in (4.49) are the two slow magnetosonic waves and the CR entropy mode. The gas entropy mode is not present because our calculation assumed that modes are isothermal, $\delta p_g/p_g = \delta\rho/\rho$, which eliminates the gas entropy mode (i.e. the gas entropy equation acts as a constraint). The first term in (4.49) comes from crossing the velocity perturbation twice with $\boldsymbol{k}$, the second term is the damping by Braginskii viscosity, the third term is the gravitational force and the last term comes from the perturbed magnetic tension. We note that the gravitational force term, which drives the instability, is zero if $\boldsymbol{g}$ is perpendicular to the $\boldsymbol{k}-\boldsymbol{B}$ plane.

As explained in Section 4.4.3, the instability is driven by gravity mediated by compressibility induced by CR streaming. The compressibility is a consequence of the perturbed pressure anisotropy $\delta\Delta p$, characterised by the Braginskii viscous frequency $\omega_{\mathrm{B}}$, in the CR entropy equation. This suggests that compressibility effects are most important in the short-wavelength limit $\omega_B/\omega_{\mathrm{A}} \gg 1$. In this limit, the dispersion relation can be simplified to

$$\begin{aligned} 0 = {} & \omega^2\frac{2}{3}i\chi\eta\frac{\omega_B}{\omega_{\mathrm{A}}}(2k_\parallel^2 - k_\perp^2) + 3ik_\perp^2\omega_B\Delta + 2\chi\eta\frac{\omega_B}{\omega_{\mathrm{A}}}k_\parallel\omega_{\mathrm{ff}}c_s\Big(-k^2(\hat{\boldsymbol{b}}\cdot\hat{\boldsymbol{g}}) + (\boldsymbol{k}\cdot\hat{\boldsymbol{g}})k_\parallel\Big) \\ & - \omega_{\mathrm{A}}\mathrm{v}_{\mathrm{A}}k^2k_\parallel\frac{4}{3}i\eta\chi\frac{\omega_B}{\omega_{\mathrm{A}}}. \end{aligned} \tag{4.51}$$

### 4.5.2 Growth Rates of Short-Wavelength Modes

We split the eigenmode frequency into real and imaginary parts, $\omega = \omega_{\mathrm{R}} + i\Gamma$, such that $\Gamma > 0$ corresponds to exponential growth. We note that if magnetic tension and damping by pressure anisotropy in eq. (4.51) can be ignored (4th and 2nd term, respectively), we recover the growth rate that was derived in Section 4.4.3, i.e. $\Gamma \sim \sqrt{\omega_s\omega_{\mathrm{ff}}}$. From (4.51) we see that damping by anisotropic viscosity (pressure) can be ignored if the third term $\propto \omega_{\mathrm{ff}}$ is much larger than the damping term, or

$$\beta \gg \eta^{-2}\Big(\frac{\omega_s}{\omega_{\mathrm{ff}}}\Big). \tag{4.52}$$

At sufficiently high $\beta$ there is therefore a finite range of scales where the simple model from Section 4.4.3 and eq. (4.48) correctly predict the solution (see Figure 4.2).

We proceed by solving (4.51) in the limit $\eta \ll 1$, i.e. for small CR pressure fractions. To leading order, $\Delta = 0$ is a solution, which, ignoring CR diffusion, implies

$$\omega \approx \chi\omega_{\mathrm{A}}. \tag{4.53}$$

This mode is the CR entropy mode. We stress again that the dependence on $\omega_{\mathrm{A}}$ in the CR entropy mode does not come from the perturbed magnetic tension, but from CR streaming along field lines at the Alfvén speed, which also has characteristic frequency $\omega_{\mathrm{A}}$. The growth rate of the mode can be found at first order in $\eta$ and is approximately given by,

$$\Gamma \approx \chi\frac{\sqrt{2}}{3}\eta\beta^{1/2}\omega_{\mathrm{ff}}\frac{-k^2\hat{\boldsymbol{b}}\cdot\hat{\boldsymbol{g}} + \boldsymbol{k}\cdot\hat{\boldsymbol{g}}k_\parallel}{k_\perp^2}, \tag{4.54}$$

where $\chi = \pm 1$ is the parameter that ensures that CR stream down their pressure gradient (eq. 4.10). The growth rate increases with increasing $\beta$, as at higher $\beta$ the pressure anisotropy is better

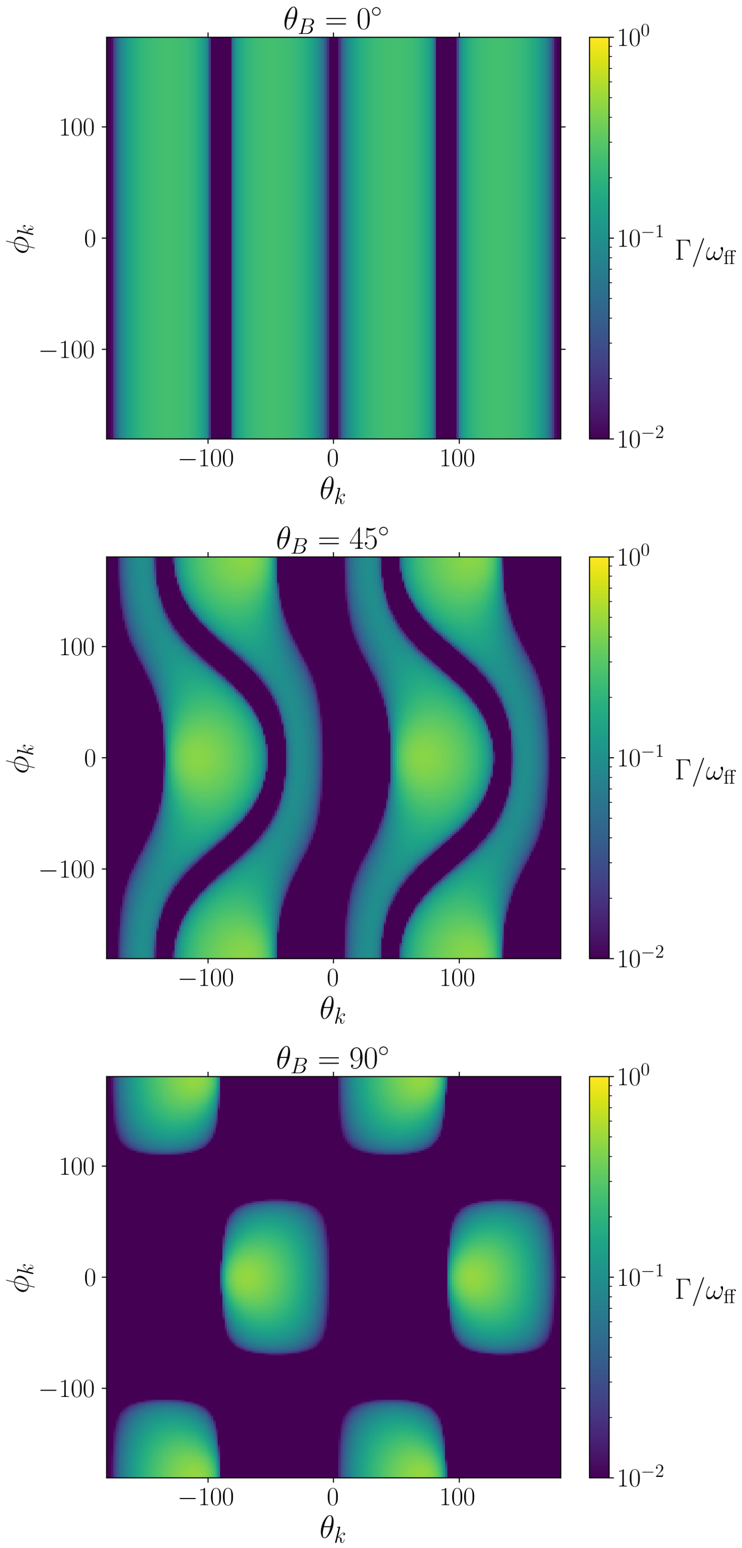


*Figure 4.5*: CRBI growth rates as a function of propagation direction for $\eta = 0.1$, $\beta = 100$, $H = 100 l_{\rm mfp}$ and fixed $kH = 20$. The angles $\theta_k$ and $\phi_k$ are defined such that $k_x = k \sin\theta_k \cos\phi_k$, $k_y = k \sin\theta_k \sin\phi_k$ and $k_z = k \cos\theta_k$ (all the other figures in this paper use $\phi_k = 0$). The three panels show growth rates for different orientations of the background magnetic field relative to gravity, which is in the $-\hat{z}$ direction. The growth rates have a qualitatively different angular dependence for the different field geometries (see equations 4.54, 4.55 and 4.56).

minimised (eq. 4.43) and anisotropic viscous damping is reduced. There is no unstable growth if gravity is normal to the $\boldsymbol{k}-\boldsymbol{B}$ plane. In the case of a magnetic field that is antiparallel to $\boldsymbol{g}$,

$$\Gamma \sim \chi\eta\beta^{1/2}\omega_{\mathrm{ff}}, \quad \boldsymbol{B} = B\hat{\boldsymbol{z}},\ \boldsymbol{g} = -g\hat{\boldsymbol{z}}. \tag{4.55}$$

For a horizontal magnetic field along $x$,

$$\Gamma \sim -\chi\eta\beta^{1/2}\omega_{\mathrm{ff}}\frac{k_z k_x}{k_z^2+k_y^2}, \quad \boldsymbol{B} = B\hat{\boldsymbol{x}},\ \boldsymbol{g} = -g\hat{\boldsymbol{z}}. \tag{4.56}$$

While for $\boldsymbol{B} = B\hat{\boldsymbol{z}}$ and $\chi > 0$ all modes are unstable (except when $\boldsymbol{k}$ is approximately parallel to or perpendicular to $\boldsymbol{B}$), for $\boldsymbol{B} = B\hat{\boldsymbol{x}}$ growth occurs if $\chi k_x k_z < 0$. Unless $\boldsymbol{B} \parallel \boldsymbol{g}$ and $\chi\boldsymbol{B}\cdot\boldsymbol{g} > 0$ ($dp_c/dz > 0$, which is unlikely), there exists a region in $k$-space where there is wave growth. We note that for horizontal magnetic fields growth rates can be higher than for vertical magnetic fields because of the extra factor that depends on the direction of propagation, although growth rates with $k_x \gg k_y, k_z$ do not diverge because the ordering used to derive (4.54)–(4.56) breaks down.

The above solutions are for pure CR streaming and also do not include the impact of background gradients on the growth rates. By neglecting CR diffusion we have assumed that CRs are perfectly coupled to the self-excited Alfvén waves and so stream at the Alfvén speed. We include CR diffusion in our analysis to relax the assumption of pure Alfvénic streaming, which certainly breaks down on small scales. From eq. (4.49), or from the fact that in the presence of diffusion the CRs' natural frequency is $\omega \approx \omega_{\mathrm{A}} - i\omega_d$ (see eq. 4.42), we can estimate that CR diffusion suppresses the instability when

$$\omega_d \gtrsim \eta\beta^{1/2}\omega_{\mathrm{ff}} \tag{4.57}$$

or equivalently

$$kl_{\mathrm{mfp}} \gtrsim \left(\Phi^{-1}\frac{l_{\mathrm{mfp}}}{H}\eta\beta^{1/2}\right)^{1/2}, \tag{4.58}$$

where we used the parameter $\Phi$ defined in eq. (4.21).

(4.42) also allows us to estimate how background gradients affect the growth rate in (4.54). The background CR pressure gradient modifies the mode's imaginary part by $\mathrm{v}_{\mathrm{A},z}d\ln p_c/dz \sim \mathrm{v}_{\mathrm{A},z}/H \sim \omega_{\mathrm{ff}}/\beta^{1/2}$, which encapsulates that as the perturbation propagates down the CR pressure gradient, $\delta p_c/p_c(z)$ remains constant in collisional MHD without CR diffusion. The background therefore significantly modifies the growth rate if,

$$\frac{\omega_{\mathrm{ff}}}{\beta^{1/2}} \gtrsim \eta\beta^{1/2}\omega_{\mathrm{ff}} \implies \eta \lesssim \beta^{-1}. \tag{4.59}$$

We stress again that for $\eta \lesssim \beta^{-1}$ the mode is not damped in the usual sense, because it maintains approximately constant $\delta p_c/p_c(z)$ as it propagates down the CR pressure gradient.

## 4.6 CR Buoyancy Instability in an Isothermal Atmosphere

We now complement the analytics of Section 4.5 with numerical solutions of equations (4.28)-(4.34) including background gradients. In this section, we consider an isothermal atmosphere in

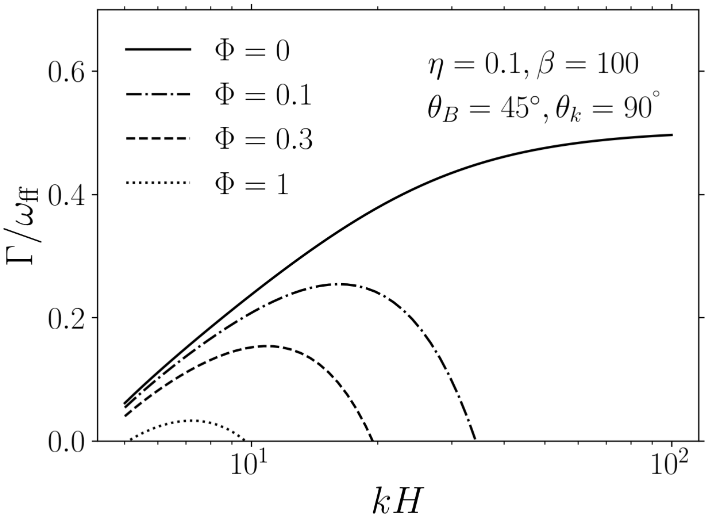


*Figure 4.6*: CRBI growth rates as a function of wavenumber and the CR diffusion coefficient (quantified using the parameter $\Phi$ defined in eq. 4.21). We assume $\eta = 0.1, \beta = 100$, $H = 100 l_{\rm mfp}$ and the isothermal background described in Section 4.6. CR diffusion suppresses growth at high $k$ (eq. 4.58) and can completely shut off the instability if sufficiently large (eq. 4.60).

order to isolate the CRBI from the HBI/MTI, which require background temperature gradients. As our fiducial set of parameters, we use $\eta = 0.1, \beta = 100$ and $H = 100 l_{\rm mfp}$. The background is in the hydrostatic equilibrium described by (4.26) and (4.27). We assume that the background CR heating is balanced by an unspecified cooling function, which we do not perturb in our linear analysis, except in Section 4.7.2 (see Kempski & Quataert 2020 for a discussion of thermal instability with streaming CRs). Unless specified otherwise, we will consider wavevectors $\boldsymbol{k}$ in the $\boldsymbol{B} - \boldsymbol{g}$ plane, i.e. $\phi_k = 0$ (see Section 4.3.5), motivated by the fact that the instability is not present if $\boldsymbol{g}$ is perpendicular to the $\boldsymbol{B} - \boldsymbol{k}$ plane. We show how growth rates depend on $\phi_k$ in Figure 4.5.

The physics of the instability described in Section 4.4 becomes apparent by plotting the properties of the CRBI. We show the mode properties for $\eta = 0.1$ and $\beta = 100$ in Figure 4.3. We plot the oscillation frequency in panel a), the growth rate in panel b), $k_\parallel {\rm v}_\parallel / k_\perp {\rm v}_\perp$, which quantifies the compressibility of the mode (= −1 if incompressible), in panel c), and $[\delta p_c + \delta(B^2/8\pi)]/\delta p_g$, which quantifies the degree of pressure balance (= −1 if pressure balanced, i.e. $\delta p_c + \delta(B^2/8\pi) + \delta p_g = 0$), in panel d). The blue line shows the unstable CR entropy mode. For completeness, we also plot the MHD slow modes and the gas-entropy mode. Panel b) shows that all modes except the CR entropy mode are strongly damped by low-collisionality physics (viscosity and conduction). At small $k$, the CR entropy mode is approximately incompressible ($k_\parallel {\rm v}_\parallel \approx -k_\perp {\rm v}_\perp$) and the growth rates are $\ll \omega_{\rm ff}$. At high $k$, the mode becomes compressible due to CR streaming, approaches $k_\perp {\rm v}_\perp = 2 k_\parallel {\rm v}_\parallel$, and the growth rate reaches the plateau given by eq. (4.54). We note that $k_\perp {\rm v}_\perp$ does not quite reach $2 k_\parallel {\rm v}_\parallel$ because we limit the x-axis to $k l_{\rm mfp} < 1$, where the Braginskii MHD model is valid. The oscillation frequency $\approx \omega_A$ is set by the characteristic frequency of CR streaming.

In Figure 4.4 we show growth rates of the CRBI for the fiducial parameters and different ratios of CR pressure to gas pressure, $\eta$. The instability exists even for small CR pressures, but with reduced growth rates. In the analytics in Section 4.5 without background gradients, the instability exists for arbitrarily small CR pressures. Here we find that with background gradients, within our local WKB framework, the instability exists for any $\eta \gtrsim \beta^{-1}$ (ignoring CR diffusion), which is consistent with the discussion at the end of Section 4.5.2.

We show growth rates as a function of propagation direction at fixed $kH = 20$ in Figure 4.5 for different orientations of the background magnetic field. The growth rates in the three panels have a qualitatively different angular dependence, consistent with equations (4.55) and (4.56). For horizontal magnetic fields growth rates are larger than for vertical magnetic fields, but growth occurs in a smaller region of $k$-space.

In Section 4.5.2 we noted that significant CR diffusion suppresses the CRBI at short wavelengths. We now show this explicitly in Figure 4.6 for different values of $\Phi$, which quantifies the strength of CR diffusion (eq. 4.21).[3] As predicted by (4.58), diffusion suppresses the instability at high $k$.

Given that the CRBI typically becomes important on scales $kH \gg 1$ and that at longer wavelengths buoyancy instabilities such as the HBI and MTI are generally more important (see Section 4.7.1 and Figure 4.7), we can rephrase (4.58) in terms of a rough overall criterion for the suppression of the CRBI by CR diffusion (however, we note that for large thermal mean free paths, as in cluster outskirts, the CR-driven instability is important on scales $kH \sim 1$, see Figure 4.9). Setting $k_{\rm min}H$ as the largest scale on which the CRBI operates and using (4.58), we find that CR diffusion suppresses the instability if,

$$\Phi \gtrsim 0.01\eta\beta^{1/2}\frac{H}{l_{\rm mfp}}\left(\frac{k_{\rm min}H}{10}\right)^{-2}. \tag{4.60}$$

For $\eta = 0.1$, $\beta = 100$, $k_{\rm min}H = 10$ and $H = 100l_{\rm mfp}$, CR diffusion suppresses the instability for $\Phi \gtrsim 1$, roughly consistent with Figure 4.6.

## 4.7 Discussion

### 4.7.1 CRBI versus HBI/MTI

We now consider the relationship between the CRBI and previously identified buoyancy instabilities driven by background temperature gradients and heat fluxes, i.e. the MTI (Balbus 2000) and the HBI (Quataert 2008). How CRs may affect these buoyancy instabilities has been considered in previous work (e.g., Chandran & Dennis 2006, Dennis & Chandran 2009, Sharma et al. 2009). However, they did not use streaming CR transport and so the CRBI was not included in their calculation.

---

[3] We assume that CR diffusion does not affect the background equilibrium, i.e. we assume an approximately linear CR pressure profile, $\kappa d^2 p_c/dz^2 \approx 0$.

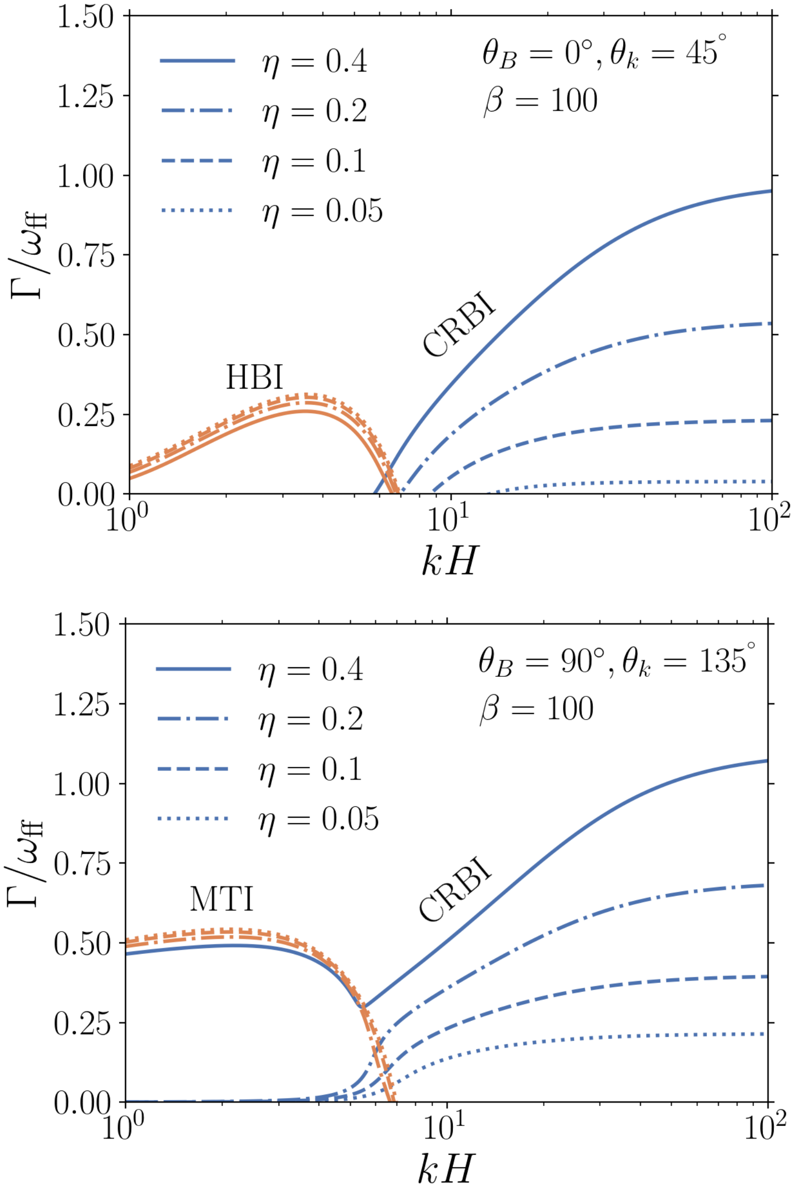


*Figure 4.7*: CRBI versus buoyancy instabilities driven by background stratification and anisotropic heat conduction, the HBI (top) and MTI (bottom). We use $\beta = 100$, $H = 100 l_{\rm mfp}$ and different $\eta$. Long-wavelength modes are destabilised by thermal conduction, as in the MTI/HBI. Short-wavelength modes are stable to the HBI/MTI due to the stabilising effect of magnetic tension. Short-wavelength modes are, however, destabilised by the compressibility induced by CR streaming. The CRBI therefore operates alongside the long-wavelength HBI and MTI. For $\eta = 0.4$, the CRBI and the MTI are destabilising the same mode (also true for the HBI at higher $\beta$).

Instead of an isothermal atmosphere as in Section 4.6, we here consider a background temperature that increases with height, as is the case in cluster cores:

$$\frac{d}{dz}(p_g + p_c) = -\rho g, \quad p_c \propto \rho^{2/3}, \quad \frac{d\ln T}{dz} = H^{-1}, \tag{4.61}$$

where $H = c_s^2/g = \omega_{\rm ff}/c_s$. This equilibrium with vertical magnetic field and $T$ increasing with height is unstable to the HBI at high $\beta$ (driven by the background anisotropic heat flux). We also consider a background with $dT/dz < 0$ and a horizontal magnetic field, $\boldsymbol{B} = B\hat{\boldsymbol{x}}$, which is unstable to the MTI. To study the MTI we consider the following background,

$$\frac{dp_g}{dz} = -\rho g, \quad \rho = \text{const}, \quad \frac{d\ln T}{dz} = -H^{-1} = -\frac{\omega_{\rm ff}}{c_s}. \tag{4.62}$$

For the MTI, we assume that $|\hat{\boldsymbol{b}} \cdot \boldsymbol{\nabla} p_c|/p_c \ll H^{-1}$, so that the cosmic rays are coupled but their background gradient is sufficiently small to be ignored (which is consistent with our choice of $\rho \approx$ const). We stress that this choice is made for the sake of simplicity and is not necessarily representative of cluster conditions.

We show growth rates for the backgrounds described by equations (4.61) and (4.62) in Figure 4.7, for $\beta = 100$, $H/l_{\rm mfp} = 100$ and different values of $\eta$. At small $k$, growth rates are dominated by the HBI/MTI. At high $k$, the HBI is partially stabilized by anisotropic viscosity (Kunz 2011), and the MTI and HBI are completely suppressed by magnetic tension. The HBI is also partially suppressed by the CR pressure gradient at long wavelengths (for large $\eta$), where CRs are approximately adiabatic, $\omega_{\rm A} < \omega_{\rm ff}$ (Sharma et al. 2009). Short-wavelength modes are destabilised by compressibility induced by CR streaming. We note that in the top panel there is a range of wavelengths where our calculation does not predict unstable growth at small $\eta$, which is not the case in the bottom panel. This is due to the effect of the background CR pressure gradient on the growth rate, explained in Section 4.5.2, which is not present in the equilibrium used to study the MTI (eq. 4.62). Finally, we note that for $\eta = 0.4$ the CRBI and the MTI are in fact driving the same mode (this is also true for the HBI for $\beta$ slightly larger than used in Figure 4.7).

The CRBI considered in this work therefore operates alongside standard buoyancy instabilities driven by background gradients, such as the HBI/MTI. Both types of instabilities are driven by gravity acting on density fluctuations. At long wavelengths, for which the HBI/MTI operate, the density fluctuations that introduce unstable buoyancy are due to a combination of background density stratification, rapid heat conduction and pressure balance. At short wavelengths, the unstable density fluctuations are due to CR streaming and pressure balance, independent of the background stratification.

The transition from heat-flux-driven growth to CR-driven growth in Figure 4.7 occurs around $kH \sim 5$. The exact value is sensitive to our choice of parameters, such as $\eta$ or $\beta$. It also depends strongly on the thermal mean free path, more specifically the ratio $H/l_{\rm mfp}$, which sets the range of $k$ for which CR streaming drives the mode away from incompressibility. In particular, for $H \sim l_{\rm mfp}$ the CRBI can have faster growth rates than the HBI/MTI at long wavelengths (see Figure 4.9) for plausible parameters. We discuss the dependence of the CRBI on the value of the thermal particle mean free path in more detail in Section 4.7.3.

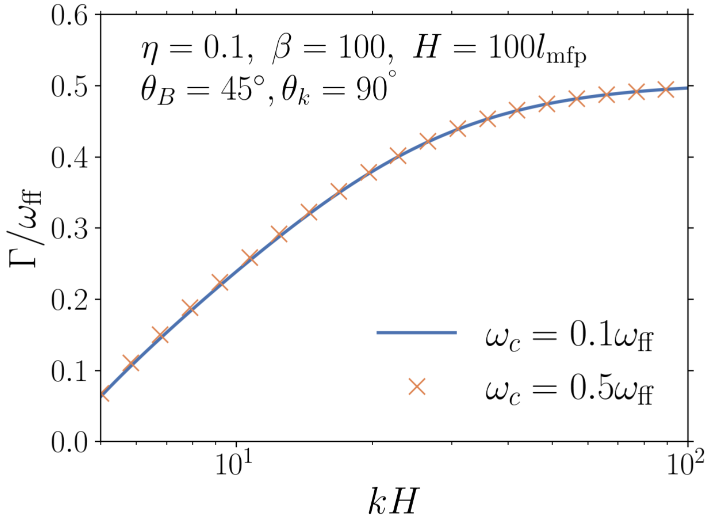


*Figure 4.8*: Growth rates of the CRBI are not significantly affected by cooling even when the cooling rate $\omega_c$ is comparable to the growth rate. This is because the mode is isothermal due to rapid conduction (the wavelengths shown are below the Field length where thermal instability is suppressed by conduction; Field 1965). For this plot, we assume thermal Bremsstrahlung to be the dominant radiative cooling process.

### 4.7.2 Impact of Cooling

We have ignored cooling throughout this work. Given that the unstable short-wavelength CR entropy modes have significant density fluctuations due to CR streaming, cooling could in principle have an impact on the instability. However, because the unstable wavelengths are characterised by thermal-conduction times that are much shorter than the cooling time, the dominant response of the gas is simply that it is isothermal, even in the presence of cooling and large CR-driven density fluctuations. The perturbed cooling therefore has no significant effect on the CRBI even when the cooling rate $\omega_c$ is comparable to the growth rate, as we show in Figure 4.8.

### 4.7.3 Dilute Cluster Outskirts and the Collisionless Regime

In Figures 4.3–4.8 we used a fixed $H/l_{\rm mfp} = 100$. While $H/l_{\rm mfp} \gg 1$ is representative of the conditions in the inner regions of galaxy clusters, $H/l_{\rm mfp}$ is likely smaller in the outskirts, where the ICM plasma density is significantly reduced. A larger mean free path implies that CR-streaming-induced compressibility effects become important on larger scales; the growth rates of long-wavelength modes will thus be enhanced relative to the results from Figures 4.3–4.8.

Considering a larger mean free path runs into the issue that the range of scales for which both the Braginskii MHD model and the WKB approximation are valid ($kl_{\rm mfp} \ll 1$ and $kH \gg 1$, respectively) becomes very limited. To alleviate this issue, we here consider a different model for the low-collisionality thermal plasma. We use the kinetic MHD equations (Chew et al. 1956) with a “Landau-fluid" prescription for the heat fluxes, i.e. the heat fluxes are constrained by the

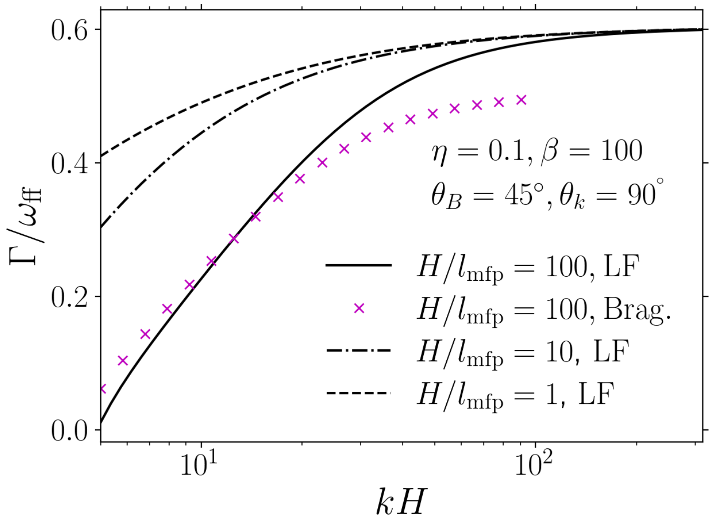


*Figure 4.9*: Growth rates of the CRBI, calculated using the Landau-fluid model (black lines), for $\eta = 0.1$, $\beta = 100$ and different $H/l_{\rm mfp}$. The Landau-fluid model allows us to compute approximate growth rates for both the collisional and collisionless regimes. For larger $l_{\rm mfp}$, as is likely the case in dilute cluster outskirts, the growth rates of long-wavelength modes are enhanced, as CR streaming leads to significant density fluctuations on large scales. There is good agreement between Landau-fluid and Braginskii MHD (magenta crosses) predictions in the collisional regime, as expected.

requirement that the fluid equations approximately match the linear response of the kinetic thermal plasma (Snyder et al. 1997). We use the heat fluxes from Snyder et al. (1997) that depend on the collision rate, allowing for a smooth transition between the weakly-collisional (Braginskii MHD) and collisionless regimes. The equations of the Landau-fluid model are provided in Appendix 4.A. For simplicity, we ignore electron physics and only consider the thermal ions, which dominate the pressure anisotropy because the ion collision rate is much smaller than the electron collision rate. Ignoring the effect of the electron heat flux on the ions is generally not rigorous (in our Braginskii calculation the heat flux was due to electrons, hence $\mathrm{Pr} = 0.02$). However, the electron-ion thermal equilibration rate is slower than the ion-ion collision frequency by the square root of the electron-to-ion mass ratio, $\tau_{\rm eq}^{-1} \sim (m_e/m_i)^{1/2}\nu_{\rm ii} \sim \nu_{\rm ii}/40$, which means that ions and electrons are thermally decoupled for $\omega \sim \omega_{\rm A}$ modes if $kl_{\rm mfp} \gtrsim \beta^{1/2}/40$. For $\beta \sim 100$ and $kl_{\rm mfp} \gtrsim 1$ electrons and ions are then approximately thermally decoupled. It is therefore reasonable to neglect the electron heat flux in the collisionless regime, which is of primary interest in this section.

We show growth rates of the CRBI, calculated using the Landau-fluid model (black lines), in Figure 4.9 for $\eta = 0.1, \beta = 100$ and different $H/l_{\rm mfp}$. As expected, for larger $l_{\rm mfp}$ the growth rates of long-wavelength modes are enhanced, as CR streaming leads to larger density fluctuations on large scales. The asymptotic growth rate at high $k$ is independent of the mean free path (absent CR diffusion). We also compare the Landau-fluid results to the Braginskii MHD calculation for the fiducial case $H/l_{\rm mfp} = 100$ (solid line and magenta crosses in Figure 4.9). There is good agreement between the two models in the collisional regime, which shows that instability growth rates are not very sensitive to the thermal Prandtl number, as the Braginskii MHD model has Pr = 0.02 while the Landau-fluid model has Pr ∼ 1 (this is consistent with our finding that Pr = 0.02 and Pr ∼ 1 yield similar results in Braginskii MHD). The asymptotic high-$k$ growth rates in the two models are also remarkably similar. This is because the qualitative physical picture of the instability does not change between the collisional and collisionless regimes (although the exact relationship satisfied by $k_\parallel \mathrm{v}_\parallel$ and $k_\perp \mathrm{v}_\perp$ at high $k$ is different in the two models, and in the Landau-fluid model $k_\parallel \mathrm{v}_\parallel \gg k_\perp \mathrm{v}_\perp$).

### 4.7.4 Diffusive Correction to CR streaming

Figure 4.6 shows that significant CR diffusion can suppress the CRBI. The magnitude of the diffusive correction to Alfvénic streaming is therefore critical. The diffusive correction depends on the damping of the Alfvén waves excited by the CR streaming instability. In the ICM, the dominant damping mechanisms are nonlinear Landau damping $\sim k\mathrm{v}_{\rm th}(\delta B/B)^2$ (Kulsrud 2005) and linear Landau damping of Alfvén waves in a turbulent background (Wiener et al. 2018). For turbulence injected on ∼ 10 kpc (a common scale for the radio bubbles) with perturbations comparable to the Alfvén speed, the linear-Landau damping rate of $k \sim r_L^{-1}$ Alfvén waves excited by GeV CRs (where $r_L$ is the GeV CR gyroradius) is

$$\Gamma_{\rm L} \sim \frac{0.4\mathrm{v}_{\rm th}}{(r_L L_{\rm turb})^{1/2}} \sim 10^{-10}\ \mathrm{s}^{-1}\frac{\mathrm{v}_{\rm th}}{10^8 \mathrm{cm\ s}^{-1}}\Big(\frac{L_{\rm turb}}{10\ \mathrm{kpc}}\Big)^{-1/2}\Big(\frac{B}{1\ \mu\mathrm{G}}\Big)^{1/2}. \tag{4.63}$$

We compute the correction to Alfvénic streaming for a combination of linear and nonlinear damping mechanisms in Appendix 4.B. The resulting diffusion coefficient is a function of the background CR pressure gradient. We split the total diffusion coefficient $\kappa$ into two components, $\kappa(\nabla p_c) = \kappa_{\rm diff}(\nabla p_c) + \kappa_{\rm st}(\nabla p_c)$, where $\kappa_{\rm st}$ is the part of the diffusion coefficient that scales as $\kappa_{\rm st} \propto (\nabla p_c)^{-1}$ and therefore does not result in real diffusive behaviour (as needed to suppress the CRBI). For purely linear damping mechanisms, $\kappa = \kappa_{\rm st}$ (Skilling 1971). Diffusive behaviour in the form of a finite $\kappa_{\rm diff}$ comes from non-zero nonlinear damping.

$\kappa_{\rm diff}$ is plotted for different linear damping strengths in Figure 4.10 as a function of the CR pressure gradient, normalized using $p_c = 10^{-12}$ ergs/cm$^3$ and a scale height $H_c = 10$ kpc. In addition to linear damping, the waves excited by the streaming instability are damped by nonlinear Landau damping. We plot $\kappa_{\rm diff}$ rather than the total $\kappa = \kappa_{\rm diff} + \kappa_{\rm st}$ because $\kappa_{\rm diff}$ is the component that acts as a diffusion term (see Appendix 4.B). The non-diffusive correction to Alfvénic streaming $\kappa_{\rm st}$ likely does not suppress the CRBI and is a small correction to Alfvénic streaming for the GeV CRs in a steady state (Kempski & Quataert 2021).

The horizontal dotted line in Figure 4.10 shows the Braginskii viscosity of the thermal plasma for $l_{\rm mfp} = 0.2$kpc and $T = 3 \times 10^7$K, and is larger than $\kappa_{\rm diff}$ in most of the parameter space. Figure 4.10 therefore shows that $\Phi < 1$ is plausible in the ICM, and so the CRBI is not completely suppressed by CR diffusion.

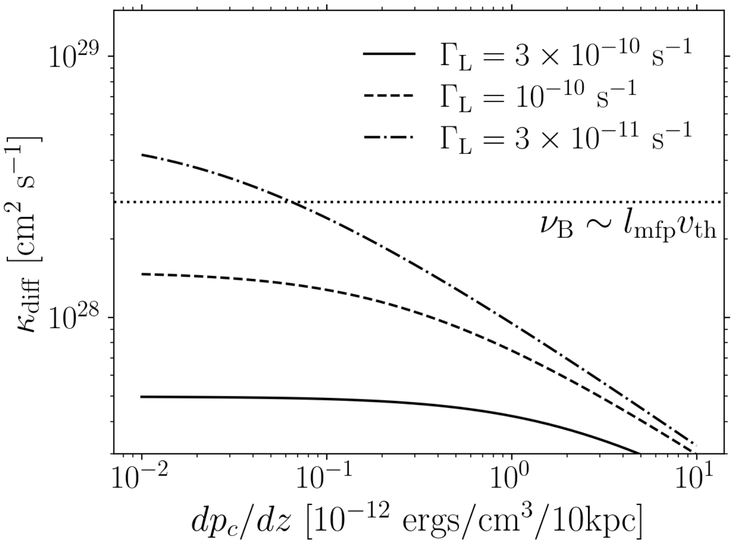


*Figure 4.10*: The diffusive correction to Alfvénic streaming calculated from eq. (4.71) as a function of the CR pressure gradient, for different magnitudes of the linear damping of Alfvén waves excited by the CR streaming instability. The horizontal dotted line is the anisotropic viscosity of the thermal gas for $l_{\rm mfp} \sim 0.2$ kpc and $T = 3 \times 10^7$K. Thus, $\Phi < 1$ (see eq. 4.21) is plausible in cluster cores and the CRBI is likely only partially suppressed by CR diffusion (Figure 4.6).

## 4.8 Conclusions

In Kempski et al. (2020) we showed that streaming CRs destabilise sound waves in the low-collisionality ICM. The instability arises because the Alfvén speed in low-collisionality plasmas depends on the pressure anisotropy of the thermal gas (eq. 4.10). This introduces a new unstable form of coupling between CRs and the thermal plasma. In this work, we showed that Alfvénically streaming CRs in a gravitationally stratified medium also destabilise a pressure-balanced mode, more specifically the CR entropy mode. We term this the Cosmic Ray buoyancy instability (CRBI) because it is the combined action of CR streaming and gravity (buoyancy) that drives the instability. CR entropy modes are highly compressible on small scales (Figure 4.3), which drives them unstable in a gravitational field. In the limit of pure CR streaming (no diffusion), there likely is no threshold for the CRBI (see discussion in Sections 4.5.2 and 4.6). The fastest growth occurs at short wavelengths, where the mode is highly compressible, with growth rates of order $\eta\beta^{1/2}\omega_{\rm ff}$ (eq. 4.54) where $\eta = p_c/p_g$, $\beta = 8\pi p_g/B^2$ and $\omega_{\rm ff}$ is the free-fall frequency. Our results show that CR streaming in cluster plasmas is a dramatically unstable process and that CR physics is important for understanding wave propagation in the ICM, even for subdominant CR pressures.

We gave a physical overview of the CRBI in Section 4.4. Instability arises due to gravity acting on the mode's density fluctuations. In standard buoyancy instabilities, such as thermal convection in stars or the magneto-thermal instability (MTI; Balbus 2000) and the heat-flux-driven buoyancy instability (HBI; Quataert 2008) in clusters, the density fluctuations are due to the background stratification of the plasma. Notably, in the CRBI the density fluctuations at short wavelengths are due to the combined action of CR streaming and pressure balance, independent of the background stratification. We complemented the qualitative physical picture from Section 4.4 with a quantitative dispersion-relation calculation in Section 4.5, and showed growth rates and mode properties for a wide range of physical parameters in Figures 4.3–4.9.

Previous work on dilute cluster plasmas showed that anisotropic conduction leads to buoyancy instabilities, the MTI and HBI. Figure 4.7 shows that these instabilities dominate growth rates at long wavelengths even in the presence of CRs, if the gas scale height is significantly larger than the thermal mean free path, as is the case in cluster cores. The CR-driven instability operates on small scales, precisely where the heat-flux-driven buoyancy instabilities are stable due to magnetic tension. The MTI/HBI and the CRBI of this paper can thus operate simultaneously in cluster plasmas. However, we note that the scale separation between the MTI/HBI and the CRBI is not always so clear: in the more dilute cluster outskirts, where the thermal mean free path is significantly larger, the CRBI can have significant growth rates (of order the free-fall frequency for plausible parameters) even for long-wavelength $kH \sim 1$ modes (Figure 4.9).

In Figure 4.11, we summarise how the CRBI and the Cosmic Ray Acoustic Braginskii (CRAB) instability from Kempski et al. (2020) compare to previously identified instabilities that may operate in ICM plasmas. We sketch representative growth rates due to the different instabilities as a function of $\eta$. For small $\eta$, the HBI/MTI are the fastest growing instabilities operating in the ICM (at large CR pressures we use a dashed line for HBI/MTI because CRs may suppress the HBI, and the CRBI and HBI/MTI can be associated with the same mode; see Figure 4.7). For $\eta \gtrsim \beta^{-1/2}$ (recall that $\beta \gg 1$ in the ICM), the growth rate of short-wavelength CR entropy modes driven compressible by CR streaming becomes comparable to or larger than $\omega_{\rm ff}$. For $\eta \gtrsim \beta^{-1/2}$ the CRAB instability of sound waves is also excited (Kempski et al. 2020). The impact of CRs on thermal-instability (TI) growth rates is modest (Kempski & Quataert 2020). The CRAB instability generally drives the fastest growing mode. This, however, does not necessarily mean that for large $\eta$ the nonlinear dynamics are dominated by the CRAB instability, as the saturation of both CR-driven instabilities remains unclear and is the subject of ongoing work. In particular, while the unstable CR entropy modes have smaller growth rates, they also have smaller group speeds and so remain in the region in which they are excited for longer. This is especially true at high $\beta$: waves propagating at the Alfvén speed with growth rates of order $\omega_{\rm ff}$ undergo several e-foldings over the distance of one gas scale height.

Heating by streaming CRs may balance cooling in the inner regions of cluster cores (Guo & Oh 2008; Jacob & Pfrommer 2017a; Jacob & Pfrommer 2017b). For a cooling rate $\omega_c$, this requires CR pressures of order (Kempski & Quataert 2020),

$$\eta \sim \beta^{1/2}\frac{\omega_c}{\omega_{\rm ff}}. \tag{4.64}$$

The CRAB instability and the CRBI become important for $\eta\beta^{1/2} \gtrsim 1$ and therefore destabilise a CR-heated medium if,

$$\eta\beta^{1/2} \sim \beta\frac{\omega_c}{\omega_{\rm ff}} \gtrsim 1. \tag{4.65}$$

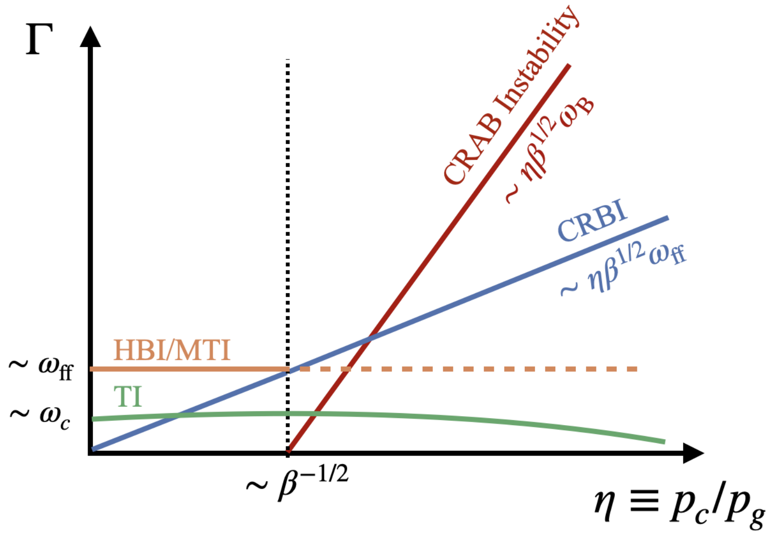


*Figure 4.11*: Overview of instabilities in dilute ICM plasmas as a function of $\eta = p_c/p_g$. For small $\eta$, the HBI/MTI have the highest growth rates. For $\eta \gtrsim \beta^{-1/2}$, the CRBI has growth rates larger than the HBI/MTI, and the CRAB instability of sound waves is excited. We include approximate growth rates of the CRBI and CRAB instability ($\Gamma \sim \eta\beta^{1/2}\omega_{\rm B}$ for the CRAB instability is valid above the instability threshold and as long as $\Gamma \ll \omega_s$; Kempski et al. 2020). CRs do not significantly affect thermal-instability (TI) growth rates for $\eta \lesssim 1$ (Kempski & Quataert 2020). The CRAB instability drives the fastest growing mode for $\eta \gtrsim \beta^{-1/2}$. However, while the CR entropy modes have smaller growth rates, they also have smaller group speeds and so remain in the region in which they are excited for longer, potentially leading to larger overall amplification.

Observations suggest that $\omega_{\rm ff}/\omega_c \gtrsim 10$ in cluster cores (e.g., McDonald et al. 2010, Hogan et al. 2017). A CR-heated medium is therefore plausibly unstable to the CRAB instability and the CRBI for $\beta \gtrsim 10$, a condition that is likely satisfied in the ICM.

We also note that the CRAB and CR buoyancy instabilities may have, to some extent, similar observational appearances. In particular, although CR entropy modes are pressure-balanced, they are compressible and involve finite gas-pressure fluctuations (balanced by CR-pressure fluctuations). CR entropy modes may therefore masquerade as sound waves if only the thermal-gas fluctuations are measured. Moreover, due to their compressible nature both the CRAB instability and the CRBI may evolve into shock-like structures that resemble the weak shocks observed in the Perseus cluster (Fabian et al. 2003; Fabian et al. 2006).

Both the CRBI considered in this work and the CRAB instability in Kempski et al. (2020) are driven by CR streaming at the Alfvén speed. The instabilities therefore operate only if the bulk of CRs in the ICM are self-confined, rather than scattered by an extrinsic turbulent cascade of magnetic fluctuations. According to current theoretical models of MHD turbulence, CR scattering by Alfvénic turbulence is likely negligible, due to the anisotropy of the cascade (Chandran 2000).

The MHD weak cascade of fast modes may be isotropic and more efficient at scattering CRs, and has been proposed as an alternative to self-confinement (Yan & Lazarian 2004). However, because the weak cascade of fast modes is strongly damped in dilute high-$\beta$ plasmas, and may generally be suppressed by wave steepening (Kadomtsev & Petviashvili 1973; Kempski & Quataert 2021), scattering of energetically important GeV CRs by fast modes is likely suppressed in the high-$\beta$ ICM. Self-confinement and streaming transport are therefore plausible.

The finite CR mean free path in the frame moving with the self-excited Alfvén waves necessarily implies a correction to the pure Alfvénic streaming model. Significant CR diffusion resulting from this correction can suppress the CRBI (Figure 4.6). However, the magnitude and nature of the correction to Alfvénic streaming remains uncertain. In particular, the form of the transport correction turns out to be rather peculiar, as it corresponds to neither streaming nor diffusion (Skilling 1971, Wiener et al. 2013; Kempski & Quataert 2021). We attempted to quantify the magnitude of the *diffusive* part of the transport correction, i.e. the contribution that may suppress the CRBI, in Section 4.7.4, which was based on the calculation from Appendix 4.B. Figure 4.10 shows that the CRBI is usually not suppressed by CR diffusion for expected ICM conditions. It would also be valuable to carry out a more complete calculation – based on CR kinetic theory – to test the conclusions of our simplified fluid treatment (although it is worth noting that existing theories of CR transport are quite uncertain and have difficulties explaining CR measurements in the Milky Way; e.g., Kempski & Quataert 2021, Hopkins et al. 2021b).

In the standard picture, AGN in cool cluster cores excite sound waves and internal gravity waves via the time dependence of the AGN jet and the buoyant motion of radio bubbles into the ICM. Kempski et al. (2020) and this work suggest that waves can also be excited by the CR pressure gradient that the bubbles provide. Future simulations will address the nonlinear evolution of the CR-driven instabilities. There are two important stages that are crucial for the nonlinear evolution and saturation: when the amplitudes become large enough to locally flatten the CR pressure gradient and shut off CR streaming ($\delta p_c/p_c \sim 1/kH$; though this does not necessarily shut off the instability, see Hin Navin Tsung et al. 2021) and when the amplitudes become large enough for the pressure anisotropy to excite kinetic microinstabilities such as the mirror (Barnes 1966; Hasegawa 1969) and firehose (Rosenbluth 1956; Chandrasekhar et al. 1958; Parker 1958) instabilities (which occur when $|\Delta p| \sim B^2/4\pi$). Upcoming work will address how this additional physics, which is not part of the linear analysis presented here, affects the evolution of the CR-driven instabilities and their impact on the ICM.

## Acknowledgements

We thank Matthew Kunz for useful comments. This research was supported in part by NSF grant AST-2107872 and a Simons Investigator award from the Simons Foundation. Support for J.S. was provided by Rutherford Discovery Fellowship RDF-U001804, which is managed through the Royal Society Te Apārangi.

# Appendix

## 4.A Landau-Fluid Closure for Low-Collisionality Plasmas

Here we provide the kinetic MHD equations and the Landau-fluid closure for the heat fluxes used in Section 4.7.3 and Figure 4.9. The kinetic MHD evolution equations for the pressures perpendicular and parallel to the magnetic field are (Chew et al. 1956),

$$\frac{\partial p_\perp}{\partial t} + \boldsymbol{\nabla} \cdot (p_\perp \mathbf{v}) + p_\perp \boldsymbol{\nabla} \cdot \mathbf{v} + \boldsymbol{\nabla} \cdot (q_\perp \hat{\boldsymbol{b}}) + q_\perp \boldsymbol{\nabla} \cdot \hat{\boldsymbol{b}} = p_\perp \hat{\mathbf{b}}\hat{\mathbf{b}} : \nabla \mathrm{v} - \frac{1}{3}\nu_{\mathrm{ii}} \Delta p, \tag{4.66}$$

$$\frac{\partial p_\parallel}{\partial t} + \boldsymbol{\nabla} \cdot (p_\parallel \mathbf{v}) + \boldsymbol{\nabla} \cdot (q_\parallel \hat{\boldsymbol{b}}) - 2q_\perp \boldsymbol{\nabla} \cdot \hat{\boldsymbol{b}} = -\,2p_\parallel \hat{\mathbf{b}}\hat{\mathbf{b}} : \nabla \mathrm{v} + \frac{2}{3}\nu_{\mathrm{ii}} \Delta p - 3(\gamma - 1)\mathbf{v_{st}} \cdot \boldsymbol{\nabla} p_c, \tag{4.67}$$

where we made the somewhat uncertain assumption that CR heating is predominantly in the direction parallel to the magnetic field. This is motivated by the fact that CR heating is due to the excitation of parallel-propagating modes, although we note that this is not true if damping by Alfvénic turbulence dominates, which acts to shear the waves to high $k_\perp$. This choice does not, however, significantly affect the results. The above equations are not yet complete, as the heat fluxes are still undetermined. In the Landau-fluid closure, the heat fluxes are set such that the linear behaviour of the fluid model approximately matches the linear response of the fully kinetic thermal plasma (Snyder et al. 1997). The Landau-fluid closure has been popular for modeling collisionless plasmas, as it recovers the fully kinetic linear damping rates (e.g. linear Landau damping of ion acoustic waves) and instabilities (e.g. MRI) of all MHD modes. A convenient form for the heat fluxes, which recovers Braginskii MHD in the collisional limit, is given by (Snyder et al. 1997),

$$q_\perp = -\frac{2c_{s\parallel}^2}{\sqrt{2\pi}|k_\parallel|c_{s\parallel} + \nu_{\mathrm{ii}}}\left[\rho\nabla_\parallel\left(\frac{p_\perp}{\rho}\right) - p_\perp\left(1 - \frac{p_\perp}{p_\parallel}\right)\frac{\nabla_\parallel B}{B}\right], \tag{4.68}$$

$$q_\parallel = -\frac{8c_{s\parallel}^2}{\sqrt{8\pi}|k_\parallel|c_{s\parallel} + (3\pi - 8)\nu_{\mathrm{ii}}}\rho\nabla_\parallel\left(\frac{p_\parallel}{\rho}\right), \tag{4.69}$$

where $c_{s\parallel} = \sqrt{p_\parallel/\rho}$. In Section 4.7.3 and Figure 4.9 we use the linearised versions of (4.66)–(4.69) instead of the linearised Braginskii MHD equations (4.32) and (4.33).

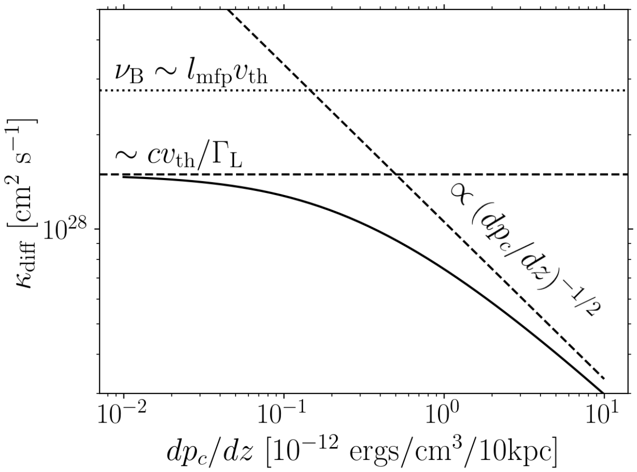


*Figure 4.12*: $\kappa_{\rm diff}$ as a function of the CR pressure gradient for a linear damping rate $\Gamma_{\rm L} = 10^{-10}\ s^{-1}$, where $\kappa_{\rm diff}$ is the component of the diffusion coefficient $\kappa$ that gives rise to actually diffusive behaviour. We estimate $\kappa_{\rm diff}$ from self-confinement theory by computing the total diffusion coefficient $\kappa$ from eq. 4.71 and subtracting $\kappa_{\rm st}$ (eq. 4.74) in order to not include the non-diffusive correction when linear damping dominates. While this method of computing $\kappa_{\rm diff}$ is not exact, it correctly recovers the diffusive correction in the two asymptotic shown by the dashed lines (equations 4.75 and 4.76). This suggests that the solid line in Figure 4.12 is a reasonable approximation of the diffusive correction to Alfvénic streaming.

## 4.B CR Diffusion Coefficient in Self-Confinement Theory

In this section we provide a heuristic calculation of the CR diffusion coefficient in self-confinement theory (a similar calculation can be found in Hopkins et al. 2020a). One challenge in this calculation is that leading-order corrections to Alfvénic streaming are often not diffusive. Instead they are better described by a (super-Alfvénic) streaming or sink term (this is the case when linear damping of Alfvén waves dominates; Skilling 1971; Wiener et al. 2018; Kempski & Quataert 2021). For our linear analysis calculation, we are mainly interested in the leading-order *diffusive* correction, which is more likely to suppress the instability than a streaming/sink term.

We calculate the amplitude of waves excited by the CR streaming instability, and the resulting CR scattering frequency, by equating Alfvén-wave growth and damping. We consider a steady state with

$$(\Gamma_{\rm L} + \Gamma_{\rm NL})\frac{\delta B^2}{4\pi} = |\mathbf{v_A} \cdot \boldsymbol{\nabla} p_c|, \tag{4.70}$$

where we split the wave damping into a linear and nonlinear part ($\propto \delta B^2$). $\Gamma_{\rm L}$ is the sum of all linear damping contributions, turbulent (Farmer & Goldreich 2004), linear-Landau (Wiener et al. 2018), ion-neutral and dust (Squire et al. 2021) damping, although the latter two are likely not important in the hot and dilute ICM. $\Gamma_{\rm NL}$ is the nonlinear Landau damping rate (Kulsrud 2005), $\Gamma_{\rm NL} = \gamma_{\rm NL}(\delta B/B)^2$, where $\gamma_{\rm NL} \sim k {\rm v}_{\rm th}$, $k$ is the wavenumber of Alfvén waves resonant with ∼ GeV CRs and ${\rm v}_{\rm th}$ is the ion thermal speed. Equation 4.70 becomes a quadratic equation for the wave amplitude

$$\left(\frac{\delta B^2}{B^2}\right)^2 + \frac{\Gamma_{\rm L}}{\gamma_{\rm NL}}\frac{\delta B^2}{B^2} - \frac{\Gamma_{\rm A}}{2\gamma_{\rm NL}}\frac{p_c}{\epsilon_{\rm B}} = 0, \tag{4.71}$$

where $\Gamma_{\rm A} \equiv |\mathbf{v_A} \cdot \nabla p_c|/p_c$ is the inverse of the Alfvén crossing time and $\epsilon_{\rm B}$ is the magnetic-field energy density. We first consider the the limit $\Gamma_{\rm L} \gg \Gamma_{\rm NL}$. (4.71) can then be solved perturbatively to yield

$$\frac{\delta B^2}{B^2} \approx \frac{\Gamma_{\rm A} p_c}{2\Gamma_{\rm L}\epsilon_{\rm B}}\left(1 - \frac{\gamma_{\rm NL}}{2\Gamma_{\rm L}}\frac{\Gamma_{\rm A} p_c}{2\Gamma_{\rm L}\epsilon_{\rm B}}\right). \tag{4.72}$$

The pitch-angle scattering rate of GeV CRs is $\nu_{\rm CR} \sim \Omega_0 \delta B^2/B^2$, which corresponds to a diffusion coefficient

$$\kappa \sim \frac{c^2}{\nu_{\rm CR}} \approx \kappa_{\rm st} + \kappa_{\rm diff}, \tag{4.73}$$

where,

$$\kappa_{\rm st} = \frac{c^2}{\Omega_0}\frac{2\Gamma_{\rm L}\epsilon_{\rm B}}{\Gamma_{\rm A} p_c}, \tag{4.74}$$

and

$$\kappa_{\rm diff} = \frac{c^2}{\Omega_0}\frac{\gamma_{\rm NL}}{\Gamma_{\rm L}} \sim \frac{c v_{\rm th}}{\Gamma_{\rm L}}. \tag{4.75}$$

$\kappa_{\rm st} \propto |\mathbf{v_A} \cdot \nabla p_c|^{-1}$ reflects the well-known result that for purely linear damping rates the diffusion coefficient is inversely proportional to the CR pressure gradient, so that the diffusion term $\nabla \cdot (\kappa \hat{\boldsymbol{b}}\hat{\boldsymbol{b}} \cdot \nabla p_c)$ ends up not being diffusive at all and is better described as a (super-Alfvénic) streaming or sink term (Skilling 1971; Wiener et al. 2013; Kempski & Quataert 2021). By contrast, $\kappa_{\rm diff}$ is independent of the CR pressure gradient and is therefore a regular diffusion coefficient.

Conversely, if nonlinear Landau damping dominates, the CR diffusion coefficient is to leading order (from eq. 4.71):

$$\kappa_{\rm NL} \approx \kappa_{\rm diff} \approx \frac{c^2}{\Omega_0}\left(\frac{2\gamma_{\rm NL}\epsilon_{\rm B}}{\Gamma_{\rm A} p_c}\right)^{1/2}. \tag{4.76}$$

$\kappa_{\rm NL} \propto |\mathbf{v_A} \cdot \nabla p_c|^{-1/2}$ and so we end up with a term that is again not diffusive in the usual sense. However, in linear theory with a background CR pressure gradient, $\nabla \cdot (\kappa \hat{\boldsymbol{b}}\hat{\boldsymbol{b}} \cdot \nabla p_c)$ still gives a term $\propto \kappa k^2$ (where $\kappa$ depends on the background gradient) and is therefore linearly diffusive.

In the high-$\beta$ ICM, linear Landau damping (Wiener et al. 2018) is likely the most important linear damping rate. For turbulence injected on $\sim$ 10 kpc (common scale of the radio bubbles) with perturbations comparable to the Alfvén speed, the damping rate of $k \sim r_L^{-1}$ Alfvén waves excited by GeV CRs (where $r_L$ is the GeV CR gyroradius) is

$$\Gamma_{\rm L} \sim \frac{0.4 v_{\rm th}}{(r_L L_{\rm turb})^{1/2}} \sim 10^{-10}\ {\rm s}^{-1}\frac{v_{\rm th}}{10^8 {\rm cm\ s}^{-1}}\left(\frac{L_{\rm turb}}{10\ {\rm kpc}}\right)^{-1/2}\left(\frac{B}{1\ \mu{\rm G}}\right)^{1/2}. \tag{4.77}$$

We plot $\kappa_{\rm diff}$, i.e. the component of the diffusion coefficient $\kappa$ that gives rise to actually diffusive behaviour, as a function of the CR pressure gradient in Figure 4.12 (see also Figure 4.10 for a different version of this plot). We calculate $\kappa_{\rm diff}$ by computing the total diffusion coefficient $\kappa$ from eq. 4.71 and subtracting $\kappa_{\rm st}$ (eq. 4.74) in order to not include the non-diffusive correction when linear damping dominates. Simply subtracting $\kappa_{\rm st}$ to obtain the diffusive correction is not exact. However, it correctly recovers the two asymptotic limits (equations 4.75 and 4.76 and the dashed lines in Figure 4.12). This suggests that the solid line in Figure 4.12 is a reasonable approximation of the diffusive correction to Alfvénic streaming.

For large CR pressure gradients, the streaming instability reaches large amplitudes (for a fixed linear damping rate) and nonlinear Landau damping is more important than linear damping mechanisms. The resulting diffusion coefficients are $\propto (dp_c/dz)^{-1/2}$. For small CR pressure gradients, linear damping dominates as the amplitudes reached by the streaming instability are not large enough for nonlinear Landau damping to be important. The diffusion coefficient is constant and approximately given by eq. 4.75. The horizontal dotted line in Figure 4.12 is the anisotropic viscosity of the thermal gas for $l_{\rm mfp} \sim 0.2$ kpc and $T = 3 \times 10^7$ K. It is therefore plausible to expect $\Phi < 1$ in cluster cores and the CRBI remains active, though is likely partially suppressed (Figure 4.6).

# Chapter 5

# Reconciling Cosmic-Ray Transport Theory with Phenomenological Models Motivated by Milky-Way Data

An earlier version of this article by Kempski P., and Quataert E. has been accepted for publication in MNRAS.

## 5.1 Abstract

Phenomenological models of cosmic-ray (CR) transport in the Milky Way (MW) can reproduce a wide range of observations assuming that CRs scatter off of magnetic-field fluctuations with spectrum $\propto k^{-\delta}$ and $\delta \sim [1.4, 1.67]$. We study the extent to which such models can be reconciled with current microphysical theories of CR transport, specifically self-confinement due to the streaming instability and/or extrinsic turbulence due to a cascade of MHD fast modes. We first review why it is that on their own neither theory is compatible with observations. We then highlight that CR transport is a strong function of local plasma conditions in the multi-phase interstellar medium (ISM), and may be diffusive due to turbulence in some regions and streaming due to self-confinement in others. A multi-phase combination of scattering mechanisms can in principle reproduce the main trends in the proton spectrum and the boron-to-carbon ratio (B/C). However, models with a combination of scattering by self-excited waves and fast-mode turbulence require significant fine-tuning due to fast-mode damping, unlike phenomenological models that assume undamped Kolmogorov turbulence. The assumption that fast modes follow a weak cascade is also not well justified theoretically, as the weak cascade is suppressed by wave steepening and weak-shock dissipation even in subsonic turbulence. These issues suggest that there may be a significant theoretical gap in our understanding of MHD turbulence. We discuss a few topics at the frontier of MHD turbulence theory that bear on this (possible) gap and that may be relevant for CR scattering.

## 5.2 Introduction

Cosmic rays may play an important role in the evolution of galaxies and diffuse gas in galaxy halos (see Zweibel 2017 for a recent review). However, the impact that CRs have on their host environment is a strong function of the adopted transport model (e.g., Ruszkowski et al. 2017; Farber et al. 2018; Hopkins et al. 2020b; Quataert et al. 2022a; Quataert et al. 2022b). As a result, the uncertainties in CR feedback are primarily driven by uncertainties in CR transport.

The CR lifetime in galaxies is much longer than the light-crossing time. It is widely accepted that the long CR confinement time is due to (resonant) scattering by small-scale electromagnetic fluctuations. Progress in understanding this scattering has occurred on both observational and theoretical fronts. There are now detailed measurements of CR spectra in the solar neighbourhood (e.g., Stone et al. 2013; Aguilar et al. 2015; Aguilar et al. 2016; Cummings et al. 2016), which put strong constraints on CR propagation models. On the theoretical side, there is growing understanding of how CRs are scattered in pre-existing MHD turbulence (Chandran 2000; Yan & Lazarian 2004; Yan & Lazarian 2008; Xu & Lazarian 2018; Lazarian & Xu 2021; Fornieri et al. 2021) and/or by self-excited waves (Kulsrud & Pearce 1969; Skilling 1971; Felice & Kulsrud 2001; Farmer & Goldreich 2004; Bai et al. 2019; Squire et al. 2021; Bai 2021).

In the self-excitation scenario, waves are generated by the CR streaming instability (Kulsrud & Pearce 1969): cosmic rays excite Alfvén waves if they collectively drift down their pressure gradient at speeds exceeding the Alfvén speed. The excited waves pitch-angle scatter cosmic rays towards isotropy in the wave frame. In the absence of damping of the self-excited waves, this limits the CR drift speed to the local Alfvén speed. In this limit, all CRs stream at the Alfvén speed, and so CR transport and the lifetime of CRs in the galaxy is energy-independent. In the presence of wave damping, the CRs are no longer fully isotropic in the frame of the self-excited Alfvén waves and CR transport exceeds the Alfvén speed by an amount that depends on the damping strength (e.g., Skilling 1971; Wiener et al. 2013). Importantly, the transport correction due to damping introduces energy dependence. However, due to the peculiar form of the correction term introduced by damping, which is neither truly diffusive nor streaming in nature,[1] its consequences for CR transport remain somewhat unclear.

The alternative to self-confinement is that CRs are scattered by a pre-existing turbulent MHD cascade. Phenomenological models of CR transport often assume that the cascade is isotropic, undamped and follows the Kolmogorov $k^{-5/3}$ (or close to Kolmogorov, e.g. $\propto k^{-3/2}$) spectral scaling, as in hydrodynamics. The resulting CR diffusion coefficient $\kappa_{\rm turb} \propto E^{\delta}$ with $\delta \sim 0.3 - 0.6$ turns out to match CR observables in the Milky Way remarkably well (e.g., Trotta et al. 2011; Gaggero et al. 2014; Hopkins et al. 2021a); this includes the spectra of secondary-to-primary CRs (e.g. the B/C ratio), which under the assumption of diffusive CR transport directly probe CR transport independent of injection physics (this is not true for streaming transport, as we show in Section 5.5). A subset of the phenomenological literature uses a combination of isotropic undamped Kolmogorov turbulence and waves excited by the CR streaming instability, which are assumed to cascade in $k_{\parallel}$

---

[1] For example, for linear damping mechanisms the term is independent of the CR distribution function. The term has been often interpreted, not entirely correctly, as super-Alfvénic streaming, with a streaming speed correction that is inversely proportional to the CR distribution function.

just like the background turbulence, to model the break in CR spectra around a few hundred GeV (e.g., Blasi et al. 2012; Aloisio & Blasi 2013 Aloisio et al. 2015). However, these phenomenological models are not justified theoretically, as MHD turbulence is known to be very different from hydrodynamic Kolmogorov-like turbulence. In MHD, the turbulent cascade of Alfvén and slow waves does have a Kolmogorov spectrum, but only in directions perpendicular to the magnetic field (the spectrum may be slightly shallower than Kolmogorov, Boldyrev 2006). The spectrum along the local magnetic-field direction, which is relevant for scattering, is $k_\parallel^{-2}$ and thus steeper than Kolmogorov. Moreover, on small scales the cascade is highly anisotropic with $k_\perp \gg k_\parallel$ (Goldreich & Sridhar 1995), which is very inefficient at scattering cosmic rays (Chandran 2000). As a result, existing models of Alfvénic turbulence predict negligible CR confinement. The compressible fast-mode cascade may be isotropic (Cho & Lazarian 2003), and may have a spectral slope that is not too far from Kolmogorov, $\propto k^{-3/2}$ (Zakharov & Sagdeev 1970; Cho & Lazarian 2003). But fast modes are subject to strong damping and the physics of their cascade (and thus spectral slope) remains uncertain. As a result, reconciling CR observables with existing theories of MHD turbulence remains an open problem (see Fornieri et al. 2021 for a recent attempt that uses only fast-mode turbulence).

In this paper we attempt to reconcile microscopic CR transport theory with phenomenological models based on MW data and highlight some of the main issues that existing theories face. We prioritise building physical intuition rather than deriving exact results. For this reason, we use simplified CR transport models to calculate order-of-magnitude estimates of MW observables. We first give a pedagogical review of the issue that existing theories of self-confinement and scattering by ambient turbulence cannot, on their own, reproduce observed CR spectra. Instead, a combination of self-confinement and turbulence may be needed to explain CR observables. We then argue that the phase structure of the ISM is important because CR transport depends on the local plasma conditions. In particular, the streaming instability is suppressed by strong damping unless the plasma is well ionized. And how efficiently turbulence can cascade to small scales to scatter cosmic rays depends on the local plasma conditions. The interstellar medium of star-forming galaxies is multi-phase with most of the volume near the mid plane dominated by the warm and hot ISM, while a significant fraction of the mass is in denser phases (e.g., McKee & Ostriker 1977). At larger heights above the midplane, the hot ISM increasingly becomes the dominant ISM component by volume and perhaps by mass (e.g., Kim & Ostriker 2017). We show that as a result of the dramatically changing conditions throughout the ISM, CR transport may be either diffusive due to turbulence or streaming due to self-confinement, depending on the ISM phase.

The paper is structured as follows. We first give an overview of CR self-confinement theory and derive associated steady-state solutions in Section 5.3. We then review CR scattering by MHD turbulence, specifically the cascade of MHD fast modes, in Section 5.4. Many of the results in Sections 5.3 and 5.4, e.g. the conclusion that neither self-confinement nor extrinsic-turbulence theory can on their own explain the CR data in the Milky Way, are not new (e.g., Kulsrud & Cesarsky 1971; Farmer & Goldreich 2004; Blasi et al. 2012; Fornieri et al. 2021), but we repeat them for pedagogical purposes (for self-confinement our approach is also quite different from what is usually done in the literature). In Section 5.5 we consider CR scattering by a combination of self-excited Alfvén waves and fast-mode turbulence. In Section 5.5.1 we show that CR transport may be multi-phase, i.e. streaming or diffusion depending on the ISM phase. We consider a particular multi-phase transport model and compare it to observations in Section 5.5.2. Our calculations show that a multi-phase combination of scattering by self-excited waves and a weak fast-mode cascade can in principle reproduce the main trends in the proton spectrum and the boron-to-carbon ratio (B/C), but that requires a significant amount of fine-tuning. We note that there already exists literature that tries to combine streaming and turbulence to explain CR observables (e.g., Blasi et al. 2012; Aloisio & Blasi 2013; Aloisio et al. 2015). However, these models assume an undamped Kolmogorov cascade. We instead focus on the theoretically better motivated interplay of streaming instability and fast-mode turbulence. The key difference is that fast modes are damped on scales $\sim$ the Larmor radius of $\sim 100-1000$ GeV particles. This makes their effect on CR transport and their interaction with CR-streaming-unstable Alfvén waves very different from what previous work concluded, which focused on undamped Kolmogorov fluctuations. We discuss uncertainties in the physics of MHD fast-mode turbulence in Section 5.6, which have significant implications for CR transport, but have not been taken into account in previous work (e.g. Yan & Lazarian 2004; Yan & Lazarian 2008; Fornieri et al. 2021). In Section 5.7 we speculate about additional uncertainties and ongoing developments from the field of MHD turbulence that may be relevant for CR transport. We summarise our results in Section 5.8.

## 5.3 Self-Confinement Theory and Steady-State Solutions

As a simple model of CR transport in galaxies, we consider 1D propagation away from the galactic disk. CRs are assumed to be injected in the disk by supernovae at a rate $2Q(p)$. We consider a uniform vertical magnetic field, $\boldsymbol{B} = B\hat{\boldsymbol{z}}$. In a steady state, $(\partial f/\partial t = 0)$ the CR distribution function as a function of momentum $p$ satisfies

$$\begin{aligned}(\mathrm{u}+\mathrm{v}_{\mathrm{st}})\frac{\partial f}{\partial z} =& \frac{1}{3}p\frac{\partial f}{\partial p}\frac{\partial}{\partial z}(\mathrm{u}+\mathrm{v}_{\mathrm{st}}) + \frac{\partial}{\partial z}\Big(\kappa\frac{\partial f}{\partial z}\Big)\\ &+ 2Q\delta(z) - 2\frac{f}{\tau_{\mathrm{loss}}(p)}h\delta(z).\end{aligned} \tag{5.1}$$

The terms in the first line describe CR transport by advection and diffusion. The first term in the second line is the source function, while the second term represents losses due to interactions with ISM material in the dense galactic mid-plane of half-thickness $h$. We ignore hadronic pion-producing collisions, as measurements in the Milky Way show that the Galaxy loses its cosmic rays through escape and not hadronic losses (Strong et al. 2010; Lacki et al. 2011). We use the expressions in Schlickeiser (2002) to evaluate $\tau_{\mathrm{loss}}(p)$ for ionisation and Coulomb losses. For both loss mechanisms, $\tau_{\mathrm{loss}}(p) \propto p^3$ for sub-relativistic CRs and $\tau_{\mathrm{loss}}(p) \propto p$ for super-relativistic CRs (see also eq. 5.51). Due to the strong dependence on momentum, the loss term is negligible for ultra-relativistic CRs ($E \gg$GeV) and so we ignore it for much of the discussion in Sections 5.3 and 5.4. However, the loss term is important for trans- and sub-relativistic CRs. We note that the loss term is more correctly described by a flux in momentum space. For the CR protons and the purposes of this paper, however, the approximate form in (5.1) is a reasonable approximation and we use it to compare our model predictions to observed proton spectra at low energies. By contrast, for boron nuclei it is important to use the flux form of the energy loss term, as we discuss in Appendix 5.C. For this reason, including energy losses in the calculations of B/C spectra is beyond the scope of this paper.

In this section, we consider CR scattering by self-excited waves generated by the streaming instability (Kulsrud & Pearce 1969). We consider CR scattering by extrinsic turbulence in Section 5.4. For self-confinement, $\mathrm{v}_{\mathrm{st}} = \mathrm{v}_{\mathrm{A}}\mathrm{sgn}(z)$ and the diffusive correction to streaming, $\kappa$, is calculated by equating the Alfvén wave damping rate and the growth rate due to the streaming instability (see, e.g., Skilling 1971). In this work, we ignore the motion of the thermal gas, i.e. we set $\mathrm{u} = 0$. However, our results for $\mathrm{v}_{\mathrm{st}} \neq 0$ easily generalise to the case with finite u, so our streaming results also mimic gas advection in, e.g., supernova-heated flows. For a constant $\mathrm{v}_{\mathrm{A}}$ and ignoring the energy loss term, we can integrate (5.1) to find that $f$ satisfies,

$$\mathrm{v}_{\mathrm{A}} f(z) - \kappa \frac{\partial f}{\partial z} = \mathrm{const} \approx Q. \tag{5.2}$$

The interpretation of equation 5.2 is simple. The rate of escape of CRs (i.e. the flux) balances the injection rate at the central source. The conserved flux on the right hand side is exactly $Q$ for scattering by turbulence and $\approx Q$ for self-confinement, with a fractional error that does not qualitatively affect the results in this paper. For a spatially varying Alfvén speed, (5.2) is valid for $z \lesssim H_A$, where $H_A$ is the ~Alfvén scale height.[2]

### 5.3.1 Single-Phase Steady-State Solutions

We assume that the diffusion coefficient $\kappa = \kappa(f_p)$ is a function of the CR *proton* distribution function only, as protons are the most abundant CR species. $\kappa$ in self-confinement theory is calculated by balancing the damping rate of Alfvén waves and the growth rate due to the CR streaming instability. As a result, the “diffusive" correction to Alfvénic streaming in (5.2) strongly depends on the mechanisms that damp the excited Alfvén waves, which in turn depend on the ISM phase.

#### 5.3.1.1 Linear Damping Mechanisms

For linear damping mechanisms, the resulting diffusion coefficient has the form (Skilling 1971),

$$\kappa_{\mathrm{sc}} = S|\hat{\boldsymbol{b}} \cdot \boldsymbol{\nabla} f_p|^{-1}, \tag{5.4}$$

---

[2] For a CR distribution function that is a power-law in momentum, $f \sim p^{-\alpha}$, the steady-state solution of (5.1) for pure Alfvénic streaming is,

$$f(z,p) = \left[\frac{\mathrm{v}_{\mathrm{A}}(0)}{\mathrm{v}_{\mathrm{A}}(z)}\right]^{\alpha/3} \frac{3Q}{\alpha \mathrm{v}_{\mathrm{A}}(0)}, \tag{5.3}$$

and so the CR scale height is comparable to the Alfvén scale height. For a constant Alfvén speed, the exact conserved flux in (5.2) is $[Q - \mathrm{v}_{\mathrm{A}} f(0)(\alpha - 3)/3]$, which gives $f = 3Q/\alpha \mathrm{v}_{\mathrm{A}}$ instead of the $f = Q/\mathrm{v}_{\mathrm{A}}$ implied by (5.2). The correction is therefore at most order unity (as $\alpha \approx 4.5$) and is essentially degenerate with the choice of $\mathrm{v}_{\mathrm{A}}$. Because $\alpha$ is an a priori unknown function of CR momentum, using the exact conserved flux unnecessarily obscures the essential physics of the calculation by making it more mathematically complicated (especially when diffusive corrections are included, e.g. in Section 5.5.1). For pedagogical purposes, we therefore use the approximation in (5.2), which does not affect our results qualitatively.

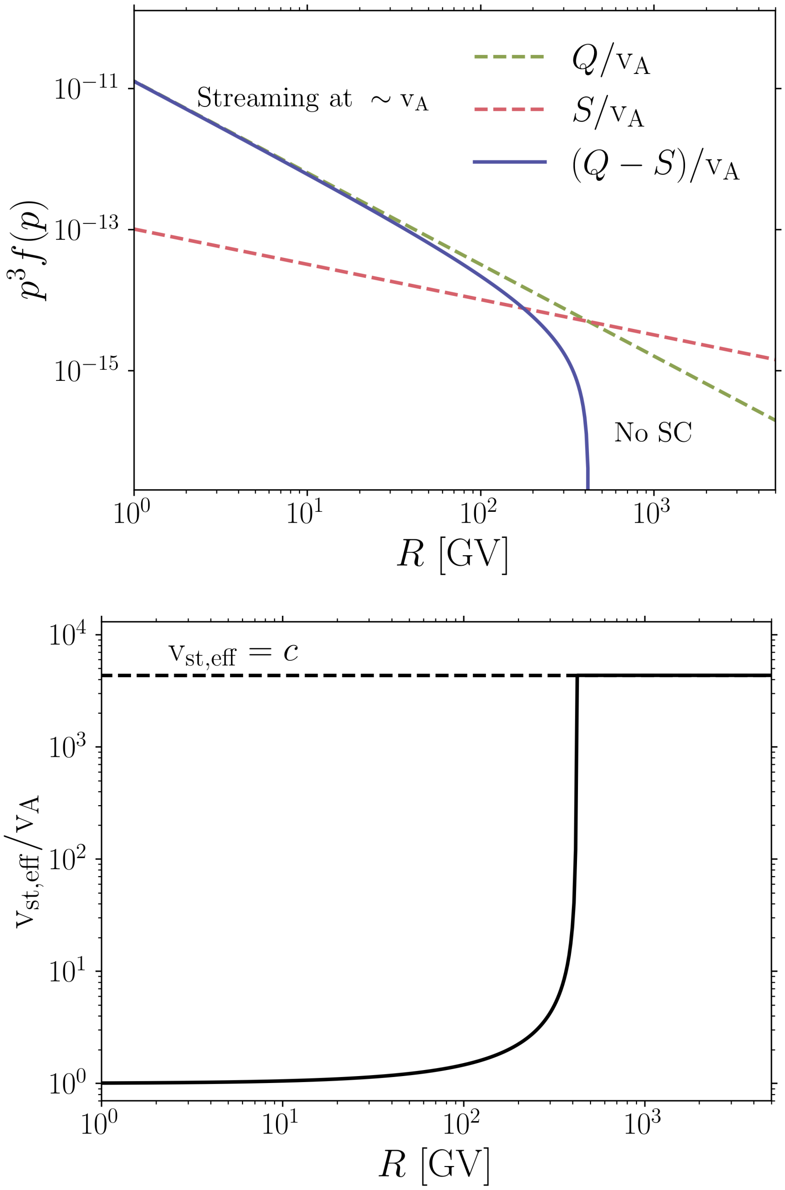


*Figure 5.1*: Top: steady-state solution in self-confinement theory with linear damping mechanisms (eq. 5.5). Here we use damping by Alfvénic turbulence with $\Gamma \propto k^{1/2}$. For a wide range of rigidities the solution (blue solid line) agrees well with the solution absent any corrections to Alfvénic streaming (green dashed line). This is followed by a sharp cutoff and at higher energies CRs are not self-confined. Bottom: effective streaming speed associated with the solution in the top panel (eq. 5.8). To reasonable approximation, for all linear damping mechanisms self-confinement theory predicts that CRs stream at either the Alfvén speed or the speed of light.

where $S$ depends on the properties of the plasma and the damping. Because $\kappa_{\rm sc} \propto |\hat{\boldsymbol{b}} \cdot \boldsymbol{\nabla} f_p|^{-1}$, the "diffusion" term ends up not being diffusive at all and is in fact independent of the CR distribution function. This peculiar result implies that linear damping mechanisms lead to a transport correction that in a steady state effectively acts like a sink of CRs, as can be seen from (5.2) combined with (5.4),

$$f \approx \frac{Q - S}{v_{\rm A}}. \tag{5.5}$$

The two terms in the numerator are functions of CR momentum, $Q \sim Q_0(p/p_0)^{-\gamma_{\rm inj}}$ with $\gamma_{\rm inj} \gtrsim 4$ and $S \propto p^{-3-a}$ for a damping $\Gamma \propto k^a$ (Skilling 1971). Thus, the two terms in the numerator have a different power-law dependence on CR momentum, and for known linear damping mechanisms, $S$ will generally have a harder spectrum than Q. For example, for linear damping by ambient Alfvénic turbulence with $\Gamma \sim k^{1/2}$ (Farmer & Goldreich 2004), $S \sim p^{-3.5}$, for damping by charged interstellar dust grains with $\Gamma \sim k^{3/4}$ (Squire et al. 2021), $S \sim p^{-3.75}$, and for ion-neutral damping with $\Gamma \sim k^0$, $S \sim p^{-3}$. We can therefore define a cutoff momentum,

$$Q(p_c) = S(p_c). \tag{5.6}$$

For $p > p_c$, $S > Q$ which gives a negative solution for $f$: CRs are unable to confine themselves. On the other hand, for almost all $p < p_c$ we have $Q \gg S$ and

$$f \approx \frac{Q}{v_{\rm A}} \qquad p \ll p_c. \tag{5.7}$$

Almost all self-confined CRs therefore stream at essentially the Alfvén speed. Thus, the primary role of linear damping in self-confinement theory is to set the energy range of CRs that are able to self-confine. An example of this is provided in Figure 5.1. In the top panel, we plot the steady-state solution from equation 5.5 (blue line). For the CR source term (= $2Q$, see eq. 5.1) we use a supernova rate of 1 per 100 years, each supernova injecting $10^{50}$ ergs in cosmic rays with a spectrum $Q \propto p^{-4.3}$. For $S$, we use $\Gamma \propto k^{1/2}$ (e.g., damping by ambient Alfénic turbulence) with $\Gamma = 10^{-11}$ s$^{-1}$ at scales resonant with 1 GeV CRs. We use a $1\mu$G magnetic field and thermal-gas density $n = 0.001$ cm$^{-3}$. We see that for a wide range of energies the solution agrees well with the solution absent any corrections to Alfvénic streaming (green dashed line). This is followed by a sharp cutoff and at higher energies CRs are not self-confined. We can define an effective streaming speed,

$$\frac{Q}{v_{\rm st,eff}} \equiv \frac{Q - S}{v_{\rm A}}, \tag{5.8}$$

which we plot in the bottom panel of Figure 5.1. The effective streaming speed is essentially the Alfvén speed for a wide range of energies. It sharply transitions to propagation at the speed of light at the cutoff energy for self-confinement. The above analysis and Figure 5.1 show that linear damping corrections to Alfvénic streaming cannot produce energy-dependent transport like that needed in phenomenological models. Instead, the linear damping of Alfvén waves just sets a maximum CR energy above which the streaming instability ceases to operate.

The above calculation can be easily extended to include the proton energy-loss term from eq. 5.1 by modifying the RHS of eq. 5.2 to $Q - f(0)h/\tau_{\rm loss}(p)$. One can then show that

$$f(0) = \frac{Q - S}{{\rm v_A}}\Big(1 + \frac{h}{{\rm v_A}\tau_{\rm loss}({\rm p})}\Big)^{-1}. \tag{5.9}$$

At high energies $h \ll {\rm v_A}\tau_{\rm loss}$ and we recover eq. 5.7, i.e. the steady state is set by CR escape from the Galaxy. At low energies $h \gg {\rm v_A}\tau_{\rm loss}$ so that $f(0) = [Q - S(0)]\tau_{\rm loss}/h$ and CRs are in the loss-dominated regime.

#### 5.3.1.2 Nonlinear Damping Mechanisms

If instead nonlinear Landau damping (NLLD; $\Gamma \sim k{\rm v_{th}}(\delta B/B)^2$; Lee & Völk 1973, Kulsrud 2005) is the dominant damping mechanism for self-excited waves, the diffusion coefficient in self-confinement theory can be calculated analogously to the linear damping case in Skilling (1971). We find,

$$\kappa_{\rm sc} \approx X|\hat{\boldsymbol{b}} \cdot \boldsymbol{\nabla} f_p|^{-1/2}, \tag{5.10}$$

with $X$ given by,

$$X \sim \frac{{\rm v_p}^2}{2\pi^2\Omega_0}\Big(\frac{8{\rm v_{th}}\Omega_0 B^2}{m_{\rm p}{\rm v_A}{\rm v_p}^3 p^3}\Big)^{1/2}, \tag{5.11}$$

where ${\rm v_p} \approx c$ is the speed of individual CR protons, $\Omega_0$ is the non-relativistic gyro-frequency, and the other symbols have the usual meaning.

Using the diffusion coefficient due to NLLD in (5.10) and assuming $\partial f/\partial z < 0$ above the galactic plane, equation 5.2 for CR protons becomes,

$${\rm v_A} f_p + X\Big(-\frac{\partial f_p}{\partial z}\Big)^{1/2} \approx Q. \tag{5.12}$$

This differential equation is separable with solution,

$$f_p(z) = \frac{Q}{{\rm v_A}} + \frac{X^2}{{\rm v_A}^2(z + C)}. \tag{5.13}$$

$C$ is a yet undetermined integration constant, for which we need to specify a boundary condition. In phenomenological models of CR transport one traditionally specifies a “CR halo size", such that $f_p(z_H) = 0$. It may seem unphysical to introduce an ad hoc cutoff, especially if it is at a small distance above the disk (e.g. a few kpc, which is difficult to reconcile with more extended CR synchrotron emission). This boundary condition also still requires specifying a value for $z_H$, which is a commonly encountered ambiguity in the CR literature. In particular, there turns out to be a degeneracy between the CR diffusion coefficient and $z_H$ when trying to infer the CR diffusion coefficient from local measurements. This degeneracy is lifted once $z_H \gtrsim r \sim 10$kpc, i.e. of order the CR injection length scale in the disk (see, e.g., Figure 10 in Linden et al. 2010). The physical reason is that for $z \gtrsim r$, CRs start propagating spherically away from the galaxy and have a smaller chance of making it back to the disk. Whether the same remains true for streaming is unclear. However, given the lack of unambiguously better alternatives, we adopt $f_p(z = r) = 0$ as our boundary condition here. We then have,

$$f_p(z) = \frac{Q}{\mathrm{v_A}} + \frac{X^2}{\mathrm{v_A}^2(z - r - X^2/(Q\mathrm{v_A}))}. \tag{5.14}$$

For small $X$ (negligible diffusion set by NLLD), we have escape via Alfvénic streaming and $f_p \approx Q/\mathrm{v_A} \propto p^{-\gamma_{\rm inj}}$. At higher energies ($X^2 \gg Q\mathrm{v_A}r$), the NLLD term becomes dominant and $f_p$ is given by $f_p(z) \approx (r - z)Q^2/X^2 \propto p^{-5.4}$, for $\gamma_{\rm inj} = 4.2$. In this regime, the CR escape time from the Galaxy therefore has a strong energy dependence, $\tau_{\rm esc} \propto p^{-\gamma_{\rm inj}+3} \propto p^{-1.2}$(see also Ptuskin et al. 1997). We note that in a steady state with CR injection balancing escape, the energy dependence of CR transport due to nonlinear Landau damping is significantly stronger than the scaling usually quoted in the literature, where the measured MW CR spectrum is used to calculate the CR diffusion coefficient (e.g., $\tau_{\rm esc} \propto p^{0.75}$ in Kulsrud 2005 assuming $f_p \propto p^{-4.5}$; $\tau_{\rm esc} \propto p^{0.85}$ in Blasi 2019 assuming $f_p \propto p^{-4.7}$).

We stress that eq. 5.13 is only valid when $\Gamma_{\rm lin} < \Gamma_{\rm NLLD}$. The nonlinear Landau damping rate can be approximately expressed as,

$$\Gamma_{\rm NLLD} \sim \left(\frac{n_{\rm CR}}{n_{\rm th}}\frac{\mathrm{v_{th}}}{\mathrm{v_A}}\Omega_0\frac{c}{L_{\rm CR}}\right)^{1/2} R_{\rm GV}^{-(\alpha-3)/2}, \tag{5.15}$$

where $n_{\rm CR}$ is the total CR number density, $R_{\rm GV}$ is the CR rigidity in GV and $\alpha$ is the slope of the CR spectrum. For $\alpha \approx 4.7$, $\Gamma_{\rm NLLD} \sim R^{-0.85}$, which is similar to the energy dependence for damping by charged interstellar dust grains and stronger than the energy dependence for damping by ambient Alfvénic turbulence ($\Gamma \sim R^{-0.5}$). Thus, even if NLLD is the dominant damping mechanism at low energies, linear damping likely becomes dominant above a rigidity $R$ for which $\Gamma_{\rm lin}(R) \sim \Gamma_{\rm NLLD}(R)$.

We can again include the proton energy-loss term from (5.1) in our solution through the transformation $Q \to Q - f_p(0)h/\tau_{\rm loss}$ in (5.14). Evaluating the result at $z = 0$ then involves a quadratic equation for $f_p(0)$. However, because the term associated with NLLD (second term on the RHS of eq. 5.14) becomes important at ultra-relativistic energies where losses are unimportant, to a very good approximation it is sufficient to carry out the transformation only on the term associated with Alfvénic streaming (first term on the RHS of eq. 5.14). As in (5.9), to include losses we therefore should multiply the solution absent losses (eq. 5.14) by $(1 + h/\mathrm{v_A}\tau_{\rm loss})^{-1}$:

$$f_p(0) \approx \left[\frac{Q}{\mathrm{v_A}} - \frac{X^2}{\mathrm{v_A}^2(r + X^2/(Q\mathrm{v_A}))}\right]\left(1 + \frac{h}{\mathrm{v_A}\tau_{\rm loss}(\mathrm{p})}\right)^{-1}. \tag{5.16}$$

### 5.3.2 Self-Confinement in a Stratified Galaxy

In Section 5.3.1 we showed that predictions from self-confinement theory are not compatible with phenomenological models of CR transport based on MW data. In particular, for both linear and nonlinear damping mechanisms, the energy dependence of CR transport is very different from the empirically derived $\kappa \sim E^{\delta}$ with $\delta \sim 0.3 - 0.7$ (see, e.g., Blasi et al. 2012, or Hopkins et al.

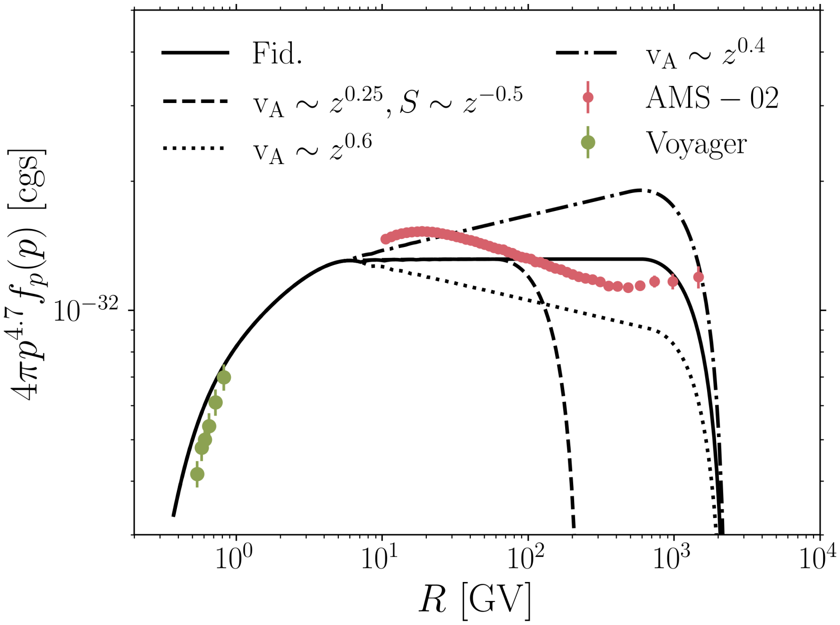


*Figure 5.2*: Comparison of self-confinement proton spectra in a stratified galaxy with AMS-02 and Voyager data (Aguilar et al. 2015; Cummings et al. 2016). In this calculation, waves excited by the streaming instability are damped linearly by the ambient Alfvénic turbulence. The CR spectrum is steepened by transport in a stratified medium, in which the Alfvén speed and the damping of self-excited Alfvén waves depend on the distance from the disk, $z$. The spectrum is steepened relative to the injection spectrum despite roughly Alfvénic streaming in most of the volume where CRs are confined. The spectral slope depends sensitively on the spatial profiles of the plasma properties that affect the Alfvén speed and the damping strength. We show how variations around the fiducial scalings (black solid line; $v_A \sim z^{0.5}$, $S \sim z^{-1}$, where $S$ characterises the spatial variation of the linear damping) affect the spectrum. See section 5.3.2 for details of the calculation. The galactic phase structure has significant implications for the interpretation of CR observables.

2021a who favour $\delta \sim 0.5-0.6$). The discrepancy between theoretical predictions and observations is particularly strong for linear damping mechanisms: based on Figure 5.1 it may seem hopeless to reconcile the theoretical transport prediction with empirical scalings.

However, before we completely discard transport corrections due to linear damping mechanisms, we note that an energy dependence in CR observables can also be introduced in rather unconventional ways. As an example, in this section we will show how roughly Alfvénic streaming, a process usually considered to be energy independent, in a stratified galaxy with linear damping of self-excited Alfvén waves can mimic energy-dependent transport with the appropriate scaling. While our calculation is not much more than a toy model, it illustrates the importance of the ISM phase structure, and its spatial variation, for interpreting CR observables. We note that CR transport in a stratified galaxy has been studied in recent years by numerous authors (see, e.g., Ptuskin et al. 1997, Recchia et al. 2016, Evoli et al. 2018 for CR transport with advection by winds or Alfvén waves, and Tomassetti 2012 for pure diffusion with vertical variation), who showed that features in CR spectra (e.g. hardening) can be the product of vertical variations in CR transport. However, our calculation is significantly different from their models. In particular, the energy dependence in CR transport that we derive is a product of the spatial variation of the linear damping rate of self-excited Alfvén waves, which can shut off the streaming instability (not considered in the works referenced above), and the simultaneous spatial variation of the Alfvén speed. Our calculation is thus similar to the models considered in Holmes (1974) and Holmes (1975), who considered exponentially varying ion-neutral damping rates and Alfvén speeds. The calculation presented here is valid for arbitrary linear damping mechanisms, including damping due ambient Alfvénic turbulence, which unlike ion-neutral damping operates in the hot ionized phases of the ISM. These occupy most of the galactic volume and are most important for setting CR observables.

In reality, galaxies and their halos are vertically stratified. In particular, the Alfvén speed likely increases, while the damping strength of self-excited Alfvén waves decreases, with increasing distance from the galactic disk. As a result of the decreasing damping, the range of CR energies that are self-confined changes with distance from the disk. And because of the simultaneous increase in Alfvén speed, CRs of different energies sample different effective escape speeds from the galaxy, even though their transport is approximately Alfvénic as discussed in Section 5.3.1.1. In a stratified medium, the distribution function for self-confined CR protons roughly satisfies (cf. equation 5.5),

$$f(p,z) \sim \frac{Q(p) - S(p,z)}{\mathrm{v_A}(z)} \qquad z \gtrsim z_{\rm sc}, \tag{5.17}$$

where we have introduced the “self-confinement height" $z_{\rm sc}(p)$, the height beyond which CRs of a given momentum are self-confined, $Q(p) = S(p, z_{\rm sc})$, and $S$ is the correction term that comes from an arbitrary linear damping mechanism of self-excited Alfvén waves. For $z < z_{\rm sc}$ CRs are not self-confined and the distribution function is approximately given by equations 5.18 and 5.19 below. We shall consider cases in which the spatial variation of $S$ is stronger than the spatial variation of $\mathrm{v_A}$. We note that (5.17) is an approximation, as we used (5.5), which assumes a constant $\mathrm{v_A}$. In a vertically stratified galaxy, with $\mathrm{v_A} \neq$ const, the Alfvénic loss term on the RHS of (5.1) is not zero and so (5.2), used to derive (5.5), is not correct. Equation 5.17 is, however, a reasonable approximation in the region $(z - z_{\rm sc})/H_A \lesssim 1$, where $H_A$ is the Alfvén scale height.

The approximate solution at some small height above the disk, $z = z_0$, for low-energy CRs with $Q \gg S(z_0)$ is found by directly evaluating (5.17), which yields the very weak energy dependence of CR transport described in Section 5.3.1.1 with $f \sim Q/\mathrm{v_A}(z_0)$. The more interesting energy dependence comes from CR momenta $p > p_*$ which are unable to self-confine at small $z$ due to large damping rates, $S(p > p_*, z = z_0) > Q(p > p_*)$. While the injection rate $Q$ is fixed, $S(z)$ is plausibly a decreasing function of $z$, and so for sufficiently low-energy CRs there may be a $z_{\rm sc}(p)$ where $Q(p) = S(p, z_{\rm sc})$ and CRs start to self-confine. We stress that $z_{\rm sc}(p)$ is a function of CR momentum, which will give rise to energy dependence in the CR distribution function.

What are the consequences for CR observables close to the disk? Because close to the disk CRs with $p > p_*$ are not self-confined, one might first guess the free-streaming solution $f(p > p_*) \sim Q/c$ at small $z$. However, if CRs do confine themselves above some height and $f(z \sim z_{\rm sc}) \sim Q/\mathrm{v_A}$, the free-streaming solution is unphysical as it corresponds to CRs streaming up their pressure gradient. Instead, the solution relaxes to a steady state with a flat spatial profile in regions where self-confinement does not operate,

$$f(p > p_*, z_0 \le z \le z_{\rm sc}) \sim \max\Big[\frac{Q(p > p_*) - S(p > p_*, z)}{\mathrm{v_A}(z)}\Big]_z, \tag{5.18}$$

i.e. $f$ is independent of $z$ for $z < z_{\rm sc}$. To get a sense of the energy scaling implied by (5.18), it is useful to further approximate the above by,

$$f(p > p_*, z_0 \le z \le z_{\rm sc}) \sim \frac{Q(p > p_*)}{\mathrm{v_A}(z_{\rm sc}(p))} \sim p^{-\gamma_{\rm inj} - \lambda_1\lambda_3}, \tag{5.19}$$

where we used our assumption that the spatial variation of $S$ is stronger than the spatial variation of $\mathrm{v_A}$ and in the last step we assumed that the spatial variations are well described by simple power laws, $\mathrm{v_A} \sim z^{\lambda_1}$, $S \sim z^{-\lambda_2}$, and as a result $z_{\rm sc} \sim p^{\lambda_3}$. For $\lambda_1 > 0$ and $\lambda_3 > 0$, the CR spectrum is steepened. For a linear damping $\Gamma \propto k^a$, $S \sim z^{-\lambda_2} p^{-3-a}$ and one can show that the self-confinement height $z_{\rm sc}$ is given by,

$$z_{\rm sc} \sim z_0 p^{\lambda_3} \sim z_0 p^{\frac{\gamma_{\rm inj} - (3+a)}{\lambda_2}}, \tag{5.20}$$

and (5.19) becomes,

$$f(p > p_*, z_0 \le z \le z_{\rm sc}) \sim p^{-\gamma_{\rm inj} - \lambda_1(\gamma_{\rm inj} - 3 - a)/\lambda_2}. \tag{5.21}$$

For $\lambda_1 > 0$ ($\mathrm{v_A}$ increases with increasing $z$), $\lambda_2 > 0$ ($S$ decreases with increasing $z$) and $a < \gamma_{\rm inj} - 3$ (true for all known linear damping mechanisms except ion-neutral damping at long wavelengths), the spectrum is steepened relative to the injection spectrum.

So far we have kept the discussion fairly general and considered arbitrary linear damping mechanisms. Let us now provide a more concrete solution for the steady-state $f$. We consider CR propagation in a turbulent inner galactic halo, so that the linear damping is due to the shearing of self-excited waves by the ambient Alfvénic turbulence. Thus, $\Gamma \propto k^{0.5}$ and for turbulence injected with Ma $\sim 1$ on scales $L$ (Skilling 1971; Farmer & Goldreich 2004),

$$S \sim p^{-3.5} \frac{L^{-0.5} B_{\rm tot} B_z}{4\pi^3 (m\Omega_0)^{0.5}} \propto B_{\rm tot}^{0.5} B_z L^{-0.5} \propto z^{-\lambda_2}, \tag{5.22}$$

where $B_{\rm tot}$ is the total magnetic field (as opposed to just the vertical component). Suppose we take $v_A \sim z^{0.5}$ i.e. $\lambda_1 = 0.5$ (e.g. $\rho \sim z^{-1}$ and constant $B_z$). Let us further assume $S \sim z^{-1}$, i.e. $\lambda_2 = 1$. For $\gamma_{\rm inj} = 4.3$, $z_{\rm sc} \sim p^{0.8}$ per equation 5.20, i.e. $\lambda_3 = 0.8$. Then, $f(p > p_*, z \leq z_{\rm sc})$ per equations 5.19 or 5.21 scales as $f \sim p^{-4.3-0.4} \sim p^{-4.7}$, approximately consistent with observations. This is despite the fact that CRs are advected at $\approx v_A$ in most of the volume in which they are self-confined. Roughly Alfvénic streaming in a stratified medium with the right spatial variations can therefore mimic energy-dependent streaming $\propto E^{0.3-0.7}$.

In Figure 5.2 we show example proton spectra from the stratified-halo calculation and compare them to AMS-02 and Voyager data (Aguilar et al. 2015; Cummings et al. 2016; for AMS-02 we only consider $E \gtrsim 10$GeV, as measurements of lower-energy CR protons are strongly affected by solar modulation). We take $z_0 = 500$pc as our base height. For the CR source term ($= 2Q$, see eq. 5.1) we take a supernova rate of 1 per 100 years (every supernova injecting $10^{50}$ ergs of CR energy) with a spectral slope $\gamma_{\rm inj} = 4.3$. We assume that the spatial variation of $f$ below $z_0$ is small, so that $f(z_0) \approx f(0)$. At the base we take $n_{\rm th} = 0.005{\rm cm}^{-3}$ and $B = 1\mu$G. The domain of our calculation extends from $z_0$ to $z_{\rm max} = 20$ kpc. We include ionisation losses as described in Section 5.3.1 and eq. 5.9, which we assume occur in a disk of neutral-gas density 1.25 cm$^{-3}$ and half-thickness 200 pc. Turbulence at the base is injected on scales $L = z_0 = 500$pc. As our fiducial spatial scalings we take $v_A \sim z^{0.5}$ ($\lambda_1 = 0.5$) and $S \sim z^{-1}$ ($\lambda_2 = 1$), as described in the paragraph above. We show the proton spectrum for the fiducial parameters and scalings as the solid line. We also show how variations around the fiducial scalings affect the spectrum. Interestingly, according to this mechanism of spectral steepening, small changes in the spectral slope as a function of CR energy in MW data are the result of changes in the spatial profiles with increasing distance from the galaxy.

The cutoff in the spectra in Figure 5.2 is set by $z_{\rm max} = 20$kpc. We could in principle shift the cutoff to higher energies by increasing $z_{\rm max}$ (or $r_{\rm max}$ in spherical geometry), but our adopted scalings would not hold indefinitely and would likely be different in the CGM relative to the near-disk environment. Thus, in the stratified-halo model it remains true that self-confinement cannot explain CR observations at high energies and there remains a need for additional scattering.

### 5.3.3 The Need for Extrinsic Waves

Sections 5.3.1 and 5.3.2 showed the well-known result that higher-energy CRs are unable to self-confine due to the linear damping of waves excited by the streaming instability. Moreover, even in the limit of negligible linear damping, nonlinear Landau damping predicts a proton spectrum that is asymptotically significantly steeper than the observed spectrum in the solar neighbourhood (Section 5.3.1.2). Thus, it transpires that self-confinement alone cannot explain observations of CRs with energies $\gtrsim$TeV. There must exist a different scattering mechanism that is necessarily dominant at high energies, and potentially dominant at low energies. A natural way of generating waves over a wide range of spatial scales, that can scatter CRs over a wide range of energies, is via a turbulent cascade. Currently, the most likely candidate for efficient CR scattering is the cascade of fast modes (Yan & Lazarian 2004; Yan & Lazarian 2008).

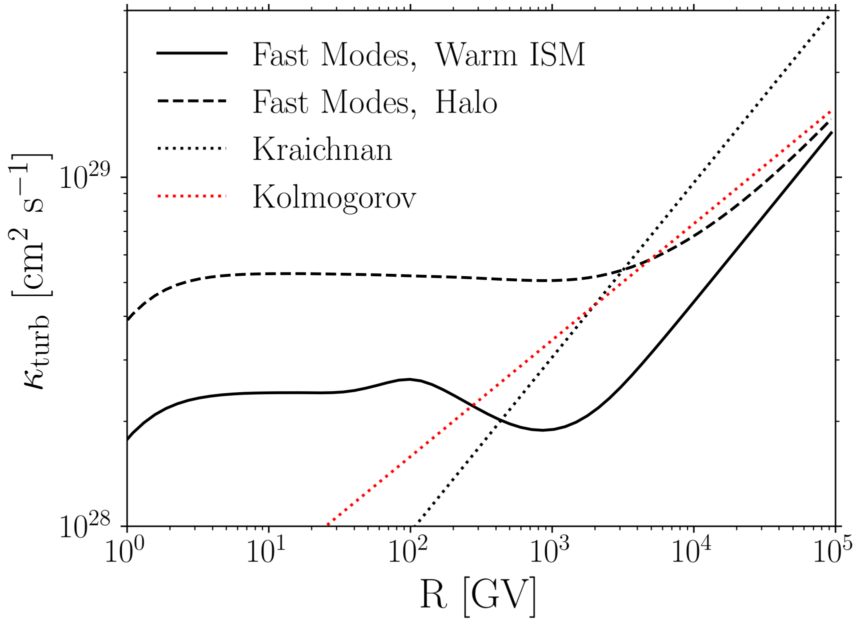


*Figure 5.3*: Diffusion coefficients due to the weak-turbulence MHD fast-mode cascade in a low-$\beta$ warm ISM (black solid line) and a low-$\beta$ halo (black dashed line; see main text in Section 5.4 for parameters). In the halo, the diffusion coefficient has a very weak energy dependence due to the collisionless damping of the cascade. In the warm ISM, the thermal mean free path is small, and the fast-mode cascade is isotropic (with $\kappa_{\rm turb} \propto p^{0.5}$) until viscous damping becomes important on scales resonant with $\sim 10^3$ GeV CRs. For comparison, we show the scaling of $\kappa_{\rm turb}$ for undamped Kolmogorov ($\propto R^{1/3}$) and Kraichnan-like ($\propto R^{1/2}$) cascades (often assumed in phenomenological models), which have a very different energy dependence from the MHD fast-mode calculation at low energies.

## 5.4 Scattering by Turbulence

When CR scattering is due to an extrinsic (balanced) turbulent cascade, $v_{\rm st} = 0$ in (5.1) and (5.2), and $\kappa = \kappa_{\rm turb}$ is specified by the properties of the turbulent cascade. For $\kappa_{\rm turb}$ independent of position, the solution to (5.2) is particularly simple,

$$f(z) = \frac{Q}{\kappa_{\rm turb}}(H - z) + f(H), \qquad z < H, \tag{5.23}$$

where we used $f = f(H)$ at $z = H$ as an outer boundary condition.

The main theoretical challenge with (5.23) lies in specifying $\kappa_{\rm turb}$. The MHD turbulent cascades of Alfvén and slow magnetosonic waves are believed to be inefficient at scattering CRs due to the fact that eddies are highly elongated along the local magnetic field (Chandran 2000). For this reason, Yan & Lazarian (2004) proposed that CRs are scattered by the MHD fast-mode cascade, which they took to be isotropic and obey a weak-turbulence $\propto k^{-3/2}$ spectrum based on the theory put forward by Zakharov & Sagdeev (1970) and the numerical work by Cho & Lazarian (2003).

Following Yan & Lazarian (2004) and Yan & Lazarian (2008), in this section we review the calculation of $\kappa_{\rm turb}$ using the weak-turbulence formalism for fast modes. In Section 5.6 we discuss some of the significant uncertainties in this calculation.

We include both collisionless damping (Ginzburg 1961) on scales smaller than the thermal-particle mean free path and anisotropic viscous damping (Braginskii 1965) on scales larger than the mean free path ($\sim 10^{13}$ cm in the warm ISM, comparable to the Larmor radius of $\sim 10$ GeV CRs). At low $\beta$, where $\beta = 8\pi p_g/B^2$ is the ratio of thermal to magnetic pressure, the damping makes the fast-mode cascade highly anisotropic below the viscous scale because parallel propagating modes are damped least efficiently. In particular, there is a scale-dependent critical wave pitch angle $\theta_c(k)$ for which the cascade rate,

$$\tau_{\rm casc}^{-1} \sim \frac{k\delta{\rm v}^2}{{\rm v_{ph}}} \sim \left(\frac{k}{L}\right)^{1/2}\frac{\delta V^2}{{\rm v_{ph}}} \tag{5.24}$$

($\delta$v denotes the amplitude at scale $k$, while $\delta V$ is the amplitude at the injection scale) is equal to the wave damping rate. For viscous damping of quasi-parallel modes at low $\beta$ the damping rate is approximately (Braginskii 1965),

$$\Gamma(k_{\|}l_{\rm mfp}{\rm v_A}/{\rm v_{th}} \ll 1) \sim \frac{\nu_{\rm B}k^2}{6}\theta^2 \qquad \theta \ll 1, \tag{5.25}$$

and for collisionless damping of quasi-parallel modes at low $\beta$ it is approximately (Ginzburg 1961),

$$\Gamma(k_{\|}l_{\rm mfp}{\rm v_A}/{\rm v_{th}} \gg 1) \sim \frac{\sqrt{\pi\beta}\theta^2}{4}\left(\frac{m_e}{m_i}\right)^{1/2}k{\rm v_A} \qquad \theta \ll 1. \tag{5.26}$$

We provide more general expressions for the damping rates in Appendix 5.A (eq. 5.45 for collisional and eq. 5.46 for collisionless damping, respectively). Equating the cascade rate (eq. 5.24) and the angle-dependent damping rates (eq. 5.45 and 5.46) gives the scale-dependent critical wave pitch angle $\theta_c(k)$. Modes with $\theta < \theta_c$ are not strongly damped. Yan & Lazarian 2004 and Yan & Lazarian 2008 assume that fast modes with $\theta < \theta_c$ continue cascading to smaller scales unaffected by the damping, while the remaining modes are fully damped. Here we use the same assumption, but we stress that it is quite uncertain. When we evaluate the cascade and damping rates in (5.24) and (5.45)–(5.46), we take the average of the cascade/damping rate in the interval $(\theta - \delta\theta, \theta + \delta\theta)$, where $\delta\theta$ is the spread in mode pitch angle experienced by a fast mode during one cascade time due to turbulent magnetic-field-line wandering. The field-line wandering due to ambient Alfvénic turbulence with ${\rm Ma_{Alf}} \sim 1$ experienced by quasi-parallel fast modes with wavenumber $k$ is,

$$\frac{\delta B}{B} \sim \left({\rm Ma}^2(kL)^{1/2}\right)^{-1/2}. \tag{5.27}$$

Throughout this work ${\rm Ma_{Alf}}$ is the turbulent amplitude (normalised by ${\rm v_A}$) at the injection scale in the Alfvénic cascade, while Ma corresponds to the amplitude (normalised by ${\rm v_A}$) in the fast-mode branch. In our calculation, we also assume that the medium is sufficiently ionized (neutral fraction $\lesssim 1\%$) to ignore ion-neutral damping (see Xu et al. 2016 for a recent discussion of MHD turbulence in partially ionized media). In the warm phase of the ISM, however, this assumption is likely only valid in a fraction of the volume. We provide a more complete summary of the $\kappa_{\rm turb}$ calculation in Appendix 5.A.

We show example diffusion coefficients in fast-mode turbulence in Figure 5.3, for turbulence injected with Mach number Ma = 1 on scales $L = 100$pc. The warm ISM ($B = 6\mu$G, $n_{\rm th} = 0.1{\rm cm}^{-3}$, $T = 2 \times 10^4$K) is shown as the solid line. We also show the diffusion coefficient for a low-$\beta$ halo for comparison ($B = 6\mu$G, $n_{\rm th} = 0.001{\rm cm}^{-3}$, $T = 2 \times 10^6$K). The diffusion coefficient in the halo has a very weak energy dependence, due to the influence of collisionless damping on the cascade. In the warm ISM with a short thermal mean free path, the fast-mode cascade is unaffected by damping (with $\kappa_{\rm turb} \propto p^{0.5}$) down to scales resonant with $\sim 10^3$GV CRs, where viscous damping becomes important. Below a few GV, $\kappa_{\rm turb}$ decreases with decreasing rigidity as CRs become trans-relativistic. For comparison we show the scaling of $\kappa_{\rm turb}$ for the undamped Kolmogorov ($\kappa_{\rm turb} \propto p^{0.33}$) and Iroshnikov-Kraichnan ($\kappa_{\rm turb} \propto p^{0.5}$) phenomenologies commonly used in the CR transport literature, which differ significantly from the fast-mode diffusion coefficients at low energies.

We can explain the main trends in Figure 5.3 heuristically. On large scales the cascade is isotropic and unaffected by the damping, with $(\delta B/B)^2 \sim k^{-1/2}$ and $\kappa \sim {\rm v}^2/\nu^G_{\Gamma=0} \sim {\rm v}^2/(\Omega(\delta B/B)^2) \sim R^{0.5}$, where $\nu^G_{\Gamma=0}$ is the gyroresonant scattering rate absent any damping. Recall that $R \sim k^{-1}$ for gyro-resonance. For collisionless damping, equating (5.24) and (5.26) gives $\theta_c^2 \sim k^{-1/2}$, i.e. surviving fast modes cover a solid angle $\sim \theta_c^2$ that shrinks $\propto k^{-1/2}$. Thus, the wave power available to scatter CRs decreases with decreasing spatial scale. Gyroresonant scattering then scales as $\nu^G \sim \nu^G_{\Gamma=0}\theta_c^2 \sim R^{1/2-1/2} \sim R^0$. Gyroresonant scattering is therefore rigidity independent in the collisionlessly damped regime. Since scattering by transit time damping (TTD; see Appendix 5.A) is also energy independent, this results in a roughly energy independent CR diffusion coefficient, as shown by the black dashed line at small rigidities in Figure 5.3. The analogous calculation for viscous damping gives $\theta_c^2 \sim k^{-3/2}$ and $\nu^G \sim \nu^G_{\Gamma=0}\theta_c^2 \sim R^1$. Thus, gyroresonant scattering decreases with decreasing rigidity, while the TTD scattering is energy-independent. As a result, the total $\kappa_{\rm turb}$ in the warm ISM, shown as the solid line in Figure 5.3, increases slowly with decreasing rigidity in the viscous regime (roughly $10^2 - 10^3$ GV), followed by weak energy dependence in the collisionless regime (energies $\lesssim$ 100 GV).

The low-energy scattering frequency associated with the diffusion coefficients in Figure 5.3, $\nu \gtrsim 10^{-8}{\rm s}^{-1}$, likely sufficiently isotropises the CR distribution function to suppress the excitation of the streaming instability. The CR anisotropy $\sim {\rm v_D}/c$, where ${\rm v_D}$ is the CR drift speed and $c$ is the speed of light, is (Skilling 1975)

$$\frac{{\rm v_D}}{c} \sim \frac{\kappa/L_{\rm CR}}{c} \sim \frac{c/L_{\rm CR}}{\nu} \sim 10^{-4}\,\frac{10\ {\rm kpc}}{L_{\rm CR}}\,\frac{10^{-8}\ {\rm s}^{-1}}{\nu}, \tag{5.28}$$

where $L_{\rm CR}$ is the CR scale height and $\nu$ is the scattering rate ($10^{-8}{\rm s}^{-1}$ corresponds to a CR mean free path of $\sim$ 1pc and a diffusion coefficient of $\sim 10^{29}$ cm$^2$s$^{-1}$). If ${\rm v_D} < {\rm v_A}$, the streaming instability is suppressed. At low $\beta$ with ${\rm v_A} > c_s$, ${\rm v_D} \sim 3 \times 10^6$ cm s$^{-1}$ is comparable to (less than) the Alfvén speed in the warm ISM (the halo/hot ISM). The CR anisotropy in low-$\beta$ turbulence is therefore plausibly small enough that it does not excite the streaming instability (especially if damping of Alfvén waves is also taken into account), consistent with external turbulence dominating the transport. Interestingly, for a plausible $L_{\rm CR} \sim 10$kpc, equation 5.28 suggests that it is implausible that CR diffusion coefficients due to turbulence are $\gg 10^{29}$ cm$^2$s$^{-1}$ for the bulk of the CRs: in such a regime, the streaming instability is excited, which increases the scattering rate on top of the turbulent scattering rate, thus lowering the effective diffusion coefficient (at higher energies the streaming instability is fully damped and so $\kappa_{\rm turb}$ can in principle be arbitrarily large).

Observations suggest a CR diffusion coefficient that approximately scales as $\kappa \propto E^{0.3-0.7}$ down to ~GeV energies. Thus, the very different scaling of $\kappa_{\rm turb}$ for fast modes at low energies in Figure 5.3 is not observed.

## 5.5 Self-Excited Waves + Fast-Mode Turbulence and Multi-Phase Models of CR Transport

Sections 5.3 and 5.4, and Figures 5.1–5.3 suggest that existing theories of self-confinement and scattering by weak fast-mode turbulence cannot, on their own, explain the rather consistent scaling with energy of the CR diffusion coefficient inferred from observations, $\kappa \propto E^{0.3-0.7}$ (see also Kulsrud & Cesarsky 1971; Farmer & Goldreich 2004; Fornieri et al. 2021). This suggests that 1) there is a different, yet unknown, universal process/cascade that efficiently scatters CRs with the right energy dependence, or that 2) a combination of scattering by self-excited waves and fast-mode turbulence conspires to mimic the empirically derived CR diffusion coefficient. At present, no MHD cascade satisfies the properties required for scenario 1). We present some speculative suggestions that might remedy this in Section 5.7. In what follows, we consider the alternative possibility that CRs are scattered by a combination of fast modes and self-excited waves. We shall assume that the MHD-fast mode cascade is isotropic and well-described by weak-turbulence theory (Section 5.4), but we stress that this assumption is very uncertain, as we discuss in Sections 5.6.1 and 5.6.2.

Our calculation is the first attempt to combine microphysical theories of CR self-confinement and scattering by MHD fast-mode turbulence. It will highlight the important issue that this theoretically motivated combination of scattering mechanisms can in principle reproduce rough trends in CR spectra, but this requires significant fine-tuning of plasma parameters. The fine-tuning issue arises primarily because the MHD fast-mode cascade is significantly damped on small scales corresponding to CR energies $\lesssim$ TeV (Figure 5.3). As a result, the CR diffusion coefficient due to fast-mode turbulence deviates strongly from the empirically derived $E^{0.3-0.7}$ scaling. This is not captured in phenomenological models of CR transport based on undamped isotropic (~ isotropic Alfvénic) turbulence.

### 5.5.1 Two-Phase Model of CR Transport

We now calculate CR spectra (the proton spectrum and the B/C spectrum) using a combination of self-excited Alfvén waves and the fast-mode turbulence model from Section 5.4.

Because fast modes are strongly damped in dilute high-$\beta$ environments (which likely make up a large fraction of the volume), while in low-$\beta$ regions they may be very efficient scatterers

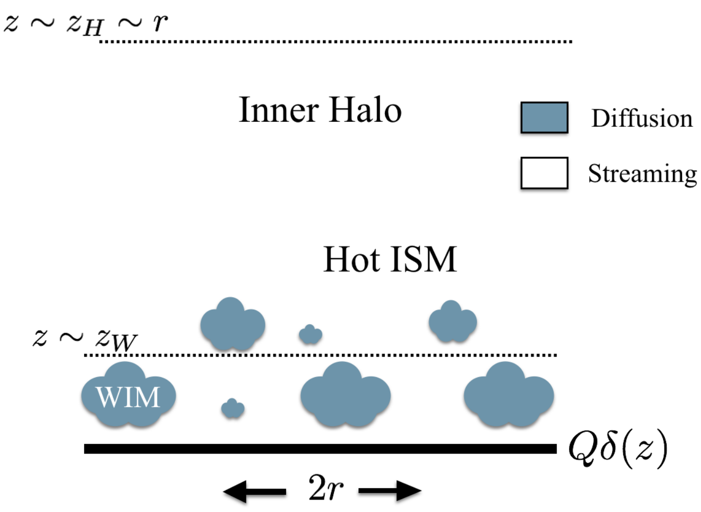


*Figure 5.4*: Example model for multi-phase CR transport in the Milky Way. Cosmic rays stream/diffuse away from the galactic disk where they are injected by supernovae and/or spallation reactions ($Q$ is the CR injection rate per unit area). In the hot ionized medium and halo, CRs are self-confined and streaming. In the warm ionized medium (WIM), where it is plausible that $\beta \equiv 8\pi p_g/B^2 < 1$, CRs are scattered by fast-mode turbulence. This results in diffusive transport.

and suppress the streaming instability (assuming the weak-turbulence calculation of Figure 5.3 correctly describes the fast-mode cascade), a possible picture emerges that CRs are scattered by only fast modes or only self-excited Alfvén waves, depending on the ISM phase. This motivates treating CR propagation as a multi-phase problem. We show an illustration of what multi-phase CR propagation may look like in Figure 5.4. There is observational evidence that the warm phases of the ISM in spiral galaxies are magnetically dominated (e.g., Beck 2015), i.e. $\beta = 8\pi p_g/B^2 \lesssim 1$. This may also be true in the inner CGM (e.g., Beck 2015). The transport model in Figure 5.4 has a low-$\beta$ warm ISM in which CRs diffuse due to scattering by fast-mode turbulence. If $\beta > 1$, however, fast modes are damped and low-energy CRs are likely self-confined. We thus assume that cosmic rays are self-confined and streaming in the remaining regions of the galaxy, i.e. the hot ISM and inner halo. We note that different ISM phase structures are also plausible and the depiction in Figure 5.4 should be regarded as one particular example of a broader class of multi-phase propagation models. We focus on the particular model illustrated in Figure 5.4 because we find that it can, in principle, match observations. This is important in terms of assessing whether existing theories of CR transport are at all compatible with observations.

The warm ionized medium in our 1D formalism is modelled as a thin layer in the disk, i.e. we assume that the region between $z = 0$ and $z = z_W$ is entirely filled by the warm ISM (this 1D picture could, for example, approximate a more realistic multi-phase ISM where the warm ISM fills 50% of the volume up to $z = 2z_W$, as in Fig. 5.4). We treat CRs as self-confined in the coronal regions and the inner halo of the galaxy, $z > z_W$. We now consider solutions to (5.2) for the two regions with different transport mechanisms, starting with the self-confinement region.

#### 5.5.1.1 Self-Confinement Regions with NLLD: the Proton Spectrum

For simplicity, we first consider a constant Alfvén speed throughout the self-confined coronal regions and inner halo, so that (5.2) is the adequate CR transport equation (as we discuss below, relaxing this assumption does not change our conclusions qualitatively). Since the “diffusive" correction due to linear damping mechanisms effectively acts like a sink of CRs, which is $\ll Q$ for most self-confined CRs and thus does not introduce energy-dependence of the right form (eq. 5.5; although we again point the reader to Section 5.3.2 for an interesting caveat), we here focus on nonlinear Landau damping. Including linear damping mechanisms in addition to nonlinear Landau damping does not significantly affect our conclusions as long as the cutoff rigidity imposed by linear damping $R_{\rm cutoff} \gg 100$GV (see Figure 5.5). In the hot and ionized halo, the relevant linear damping is set by the ambient Alfvénic turbulence (Farmer & Goldreich 2004), so we essentially assume a weakly turbulent halo, so that NLLD is more important than linear damping (this requires fairly small amplitudes at the turbulence injection scale, $\delta {\rm V}/{\rm v_A} \sim \mathcal{O}(0.1)$). Under these assumptions, per equations 5.1 and 5.10 the CR flux in the self-confined halo, $z > z_W$, satisfies,

$$
{\rm v_A} f_p + X\Big(-\frac{\partial f_p}{\partial z}\Big)^{1/2} = {\rm const} = -\kappa_{\rm turb}\Big[\frac{\partial f_p}{\partial z}\Big]_{z\to z_W^-} = Q, \tag{5.29}
$$

where we imposed flux continuity at $z = z_W$ and in the last step we used the solution in (5.23) to evaluate the CR flux $-\kappa_{\rm turb}\partial f_p/\partial z$ in the warm ISM in which CRs are scattered by ambient fast-mode turbulence. In the self-confinement region, the CR distribution function therefore satisfies the same equation as for the single-phase medium in Section 5.3.1.2 (eq. 5.12). Thus, the steady-state CR proton distribution function for $z > z_W$ is in our model given by equation 5.14.

In deriving eq. 5.14 we made the assumption that ${\rm v_A}$ is constant and one may wonder how much our results depend on it. For negligible NLLD and for an Alfvén speed that increases with $z$, the solution to (5.1) is $f_p \approx Q/{\rm v_A}\exp(-z/H_A)$, where $H_A$ is the $\sim$ Alfvén scale height. One can show that in this case NLLD becomes important when $X^2 \gtrsim Q{\rm v_A}{\rm H_A}$. Previously, in (5.14) we found that NLLD becomes important when $X^2 \gtrsim Q{\rm v_A}{\rm r}$. At high energies (large $X$) our results are independent of ${\rm v_A}$ in both cases. So, a spatially-varying ${\rm v_A}$ merely changes the CR energy at which NLLD becomes important by order unity (since $H_A$ is expected to be quite large in the halo, $\gtrsim 5$kpc). Importantly, the exact spatial profile of ${\rm v_A}$ does not qualitatively change our results.

The solution in (5.14) shows that the proton spectrum depends sensitively on the injection spectrum $Q$. The slope of the injection spectrum remains uncertain, which is a limitation for putting tight constraints on CR transport using just the proton flux. In this work we will assume that the spectral slope at injection is around $\gamma_{\rm inj} \approx 4.2 \pm 0.1$, motivated by acceleration at strong shocks and FERMI data on luminous starbursts (Ackermann et al. 2012), where gamma-ray emission likely traces the injection spectrum without any energy-dependent losses (e.g., Lacki et al. 2011).

#### 5.5.1.2 Self-Confinement Regions with NLLD: the Boron-to-Carbon Ratio

Cleaner constraints on CR transport can usually be obtained using the boron-to-carbon ratio (B/C). While C nuclei are CR primaries, i.e. they are injected at the sources, B nuclei are secondaries created by spallation reactions of C nuclei in the ISM. The B/C ratio is therefore a direct probe of the average column density / grammage traversed by CRs during their lifetime in the galaxy. Importantly, for diffusion in turbulence B/C is independent of the injection spectrum. However, there is significant dependence on the injection spectrum if one considers self-confinement (see eq. 5.30 below).

In self-confinement regions ($z > z_W$), B nuclei are passively scattered by the Alfvén waves excited by the protons via the streaming instability. By imposing flux continuity at $z = z_W$ as in Section 5.5.1.1 and equation 5.29, we find that B nuclei satisfy equation 5.2, with $\kappa = \kappa(f_p)$ according to (5.10),[3] and $f_p$ given by (5.14). As the production mechanism for B is different, however, we need to replace the proton source term $Q$ with $f_C h\delta(z)/\tau_{\mathrm{spall}}$, where $f_C$ is the carbon distribution function, $h$ is the thickness of the dense thin disk in which boron is assumed to be produced, and $\tau_{\mathrm{spall}}$ is the spallation reaction timescale in the disk (we assume the cross section to be constant for relativistic CRs, = 61mB). The differential equation is solvable, with

$$\frac{f_B(z,R)}{f_C(0,R)} = \frac{h}{\mathrm{v_A}\tau_{\mathrm{spall}}}\frac{(r - z + X^2/(Q\mathrm{v_A}))^{\mathrm{v_p}/\mathrm{v_B}} - (X^2/(Q\mathrm{v_A}))^{\mathrm{v_p}/\mathrm{v_B}}}{(r - z + X^2/(Q\mathrm{v_A}))^{\mathrm{v_p}/\mathrm{v_B}}}, \tag{5.30}$$

where $\mathrm{v_B}$ and $\mathrm{v_p}$ are the B and proton speeds at the same rigidity $R$ ($\mathrm{v_p}/\mathrm{v_B} \approx 1$ for $R \gg 1$ GV), and $Q$ and $X$ are evaluated for proton momenta corresponding to rigidity $R$. We note that $f_C$ in the denominator is evaluated at $z = 0$ due to the delta function that appears in the B source term. At low energies (small $X$), $f_B$ is set by Alfvénic escape, while at higher energies NLLD becomes important. At high energies with $\mathrm{v_p} \approx \mathrm{v_B}$ and $r\mathrm{v_A} \ll X^2/Q$, $f_B(z = 0)/f_C(z = 0) \propto Q/X^2 \propto p^{-\gamma_{\mathrm{inj}}+3}$, and so B/C in self-confinement theory depends strongly on the injection spectrum.

#### 5.5.1.3 Diffusion in Turbulence in the Warm ISM

We now complete our solution for B/C and the proton spectrum by considering CR diffusion in the warm ISM layer ($z < z_W$). Because protons and B nuclei are both passively scattered by the turbulent cascade, they are described by the same equation. Assuming $\kappa_{\mathrm{turb}} = \mathrm{const}$, the solution to (5.2) is,

$$f(z) = \frac{Q'}{\kappa_{\mathrm{turb}}}(z_W - z) + f(z_W), \qquad z < z_W, \tag{5.31}$$

where $f$ can denote any CR species and $Q'$ is the source term for either protons or B nuclei. We imposed continuity of $f$ between the warm ISM and the coronal regions as the outer boundary condition, i.e. we use (5.14) and (5.30) for the protons and boron nuclei, respectively (evaluated at $z = z_W$). Because $f(z_W)$ is set by the transport in the halo, eq. 5.31 highlights a key property of our multi-modal CR transport models: different transport mechanisms set observables at $z = 0$ depending on CR energy and the ISM phase structure. In particular, the transport in the warm ISM (assumed diffusive here) sets observables in the disk if

$$\frac{Q' z_W}{\kappa_{\mathrm{turb}}} \gtrsim f(z_W). \tag{5.32}$$

---

[3] For B nuclei with the same rigidity as protons, we have to multiply the proton diffusion coefficient by the speed ratio $\mathrm{v_B}(R)/\mathrm{v_p}(R)$ to get the boron diffusion coefficient at the same rigidity.

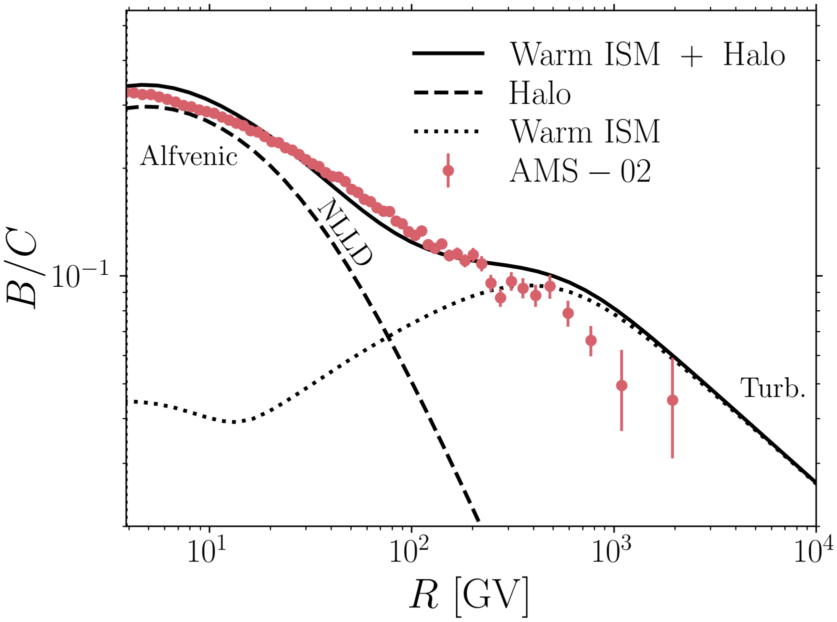


*Figure 5.5*: The boron-to-carbon ratio (B/C) for our fiducial two-phase ISM model. The red points are AMS-02 measurements (Aguilar et al. 2016) and the black line is the spectrum obtained using (5.30) and (5.31) evaluated at $z = 0$. The dotted and dashed lines show the individual contributions from the turbulent warm ISM, and the self-confined hot ISM and halo (first and second terms in eq. 5.31, respectively). There is good agreement between the model and the data. For $R \lesssim 100$ GV, self-confinement in the diffuse hot ISM and halo sets the local B/C. At GeV energies, CR transport has a weak energy dependence as CRs stream at the Alfvén speed. At intermediate energies, NLLD is important and introduces a strong dependence on energy. Above a few hundred GeV, diffusion in the turbulent warm ISM becomes the escape-rate-limiting step and sets the local ($z = 0$) CR observables. We show rigidities that correspond to kinetic energies > 1 GeV/nucleon, for which we find that escape dominates over ionisation losses, i.e. $\tau_{\rm ion} \gg h/{\rm v_A}$, consistent with our model assumptions.

In the opposite limit, it is the transport in the halo (assumed to be streaming here) that sets the local observables. We note that the above inequality is similar, though not equivalent, to the common assumption that low-energy CRs are scattered by self-excited waves, while higher-energy CRs are scattered by turbulence. In fact, in the low-$\beta$ warm ISM CRs of all energies are scattered by the fast-mode cascade (see Figure 5.3). For the model presented here, the more accurate interpretation is that *the escape-rate-limiting step* is self-confinement for low-energy CRs and fast-mode turbulence for higher-energy CRs.

#### 5.5.1.4 Proton Ionisation Losses

CRs are also subject to ionisation losses in the dense disk in which boron is produced (here we assume that the disk is mostly neutral, so that ionisation losses dominate over Coulomb losses). The effect of ionisation losses on the CR proton spectrum can be included analogously to equations 5.9 and 5.16 by transforming the solution in (5.31),

$$f(0) \rightarrow f(0)\Big(1 + \frac{h}{\mathrm{v_A}\tau_{\mathrm{loss}}} + \frac{h}{\tau_{\mathrm{loss}}}\frac{z_W}{\kappa_{\mathrm{turb}}}\Big)^{-1}. \tag{5.33}$$

### 5.5.2 Two-Phase Model Versus Observations

We now compare our multi-phase model to local CR measurements. Finding a “best-fit" set of parameters would be overkill given the simplicity of our model and the theoretical uncertainties. Instead, we show that there exists a plausible set of physical parameters for which we can recover the normalisation and essential trends of the proton and B/C spectra. However, our results also show that the spectra calculated using a combination of self-excited Alfvén waves and fast-mode turbulence are extremely sensitive to local plasma conditions. As a result, generating CR spectra with approximately constant power-law slope using a combination of self-excited waves and fast-mode turbulence requires significant fine-tuning, unlike phenomenological models of CR transport based on self-generated waves and undamped isotropic (typically Kolmogorov) turbulence.

We use the following fiducial parameters: a supernova rate of 1 per century, injecting $10^{50}$ ergs of CRs per supernova with a spectrum $Q \propto p^{-4.3}$. In the hot ISM we use $n_{\mathrm{th}} = 0.003\ \mathrm{cm}^{-3}$, $T = 3 \times 10^6$ K and $B = 1\mu$G. For the warm ISM, we use a density $n_{\mathrm{th}} = 0.2\ \mathrm{cm}^{-3}$, temperature $T = 10^4$ K and magnetic field $B = 6\ \mu$G. This corresponds to $\beta \approx 0.2$ in the warm ISM and $\beta \approx 30$ in the hot ISM. We assume that the transition between the warm ISM and the coronal regions occurs at $z_W = 500$ pc above the disk and we use $r = 10$ kpc as the characteristic CR injection length scale in the disk. We assume that fast-mode turbulence in the warm ISM is injected with $\mathrm{Ma} = 0.6$ on scales of order $L = 50$ pc. For the boron source term and ionisation loss term, we assume a thin disk of half-thickness 200 pc and density $\approx 1.25\ \mathrm{cm}^{-3}$. We stress that this is not a constrained, unique set of parameters. It is merely a set of physical parameters that matches the normalisation and main spectral trends in the observed data. We show some variation about these parameters in Figure 5.6 discussed below.

The B/C spectrum for our fiducial parameters is shown in Figure 5.5. The red points are AMS-02 measurements (Aguilar et al. 2016) and the black solid line is the spectrum obtained

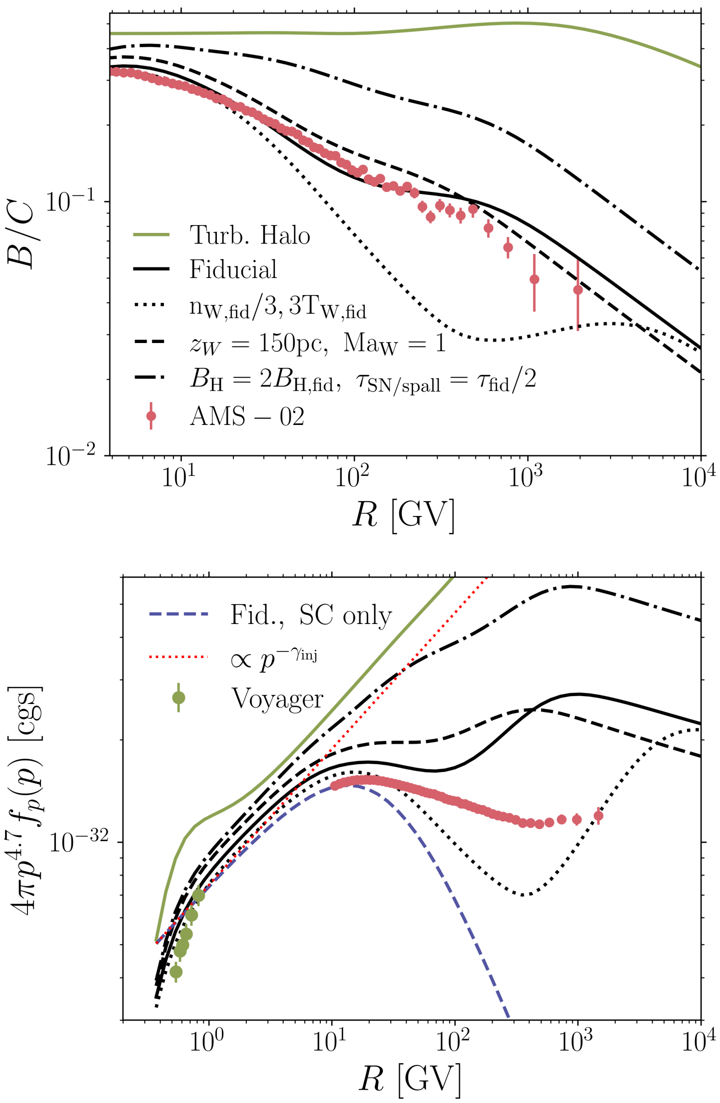


*Figure 5.6*: Dependence of spectra (top: B/C, bottom: protons) on our choice of parameters and propagation model. Red points are AMS-02 measurements, green points are Voyager measurements. The black lines show the effect of variations in the parameters for our fiducial two-phase propagation model (Figure 5.4). Modifying the parameters (black lines) changes the quality of the fit, but some qualitative trends remain the same. The proton spectrum is significantly steeper than the injection spectrum (red dotted line) for energies $\gtrsim$ 10GeV. However, even small changes can introduce significant deviations from a power-law (e.g., the dotted line), which is incompatible with the observations. Due to the wide range of plasma conditions present in a realistic multi-phase galaxy, CR transport is likely some convolution of our model's results for different parameters. A multi-phase combination of scattering mechanisms can in principle reproduce rough trends in CR spectra, but this requires significant fine-tuning. The green line is a different model with scattering by fast-mode turbulence everywhere, i.e. no self-confinement; it shows qualitatively very different behaviour. Due to the weak energy dependence of $\kappa_\mathrm{turb}$ in the halo (Figure 5.3), B/C is essentially flat and the proton spectrum $\sim p^{-\gamma_\mathrm{inj}}$, inconsistent with observations. The blue dashed line in the bottom panel shows the spectrum calculated using the self-confined halo only (no contribution from the warm ISM). Without scattering by fast-mode turbulence, the spectrum at high energies is too steep unless the transport region is stratified as in Figure 5.2.

using (5.30) and (5.31) evaluated at $z = 0$. The dotted and dashed lines show the contributions from the turbulent warm ISM and self-confined halo, respectively. At rigidities $\lesssim$ 100 GV, self-confinement in the diffuse hot ISM and halo sets the local B/C. At GeV energies, CRs stream at the Alfvén speed. At higher energies, but still in the self-confined limit, NLLD introduces a stronger dependence on energy. Above a few hundred GeV, diffusion in the turbulent warm ISM becomes the escape-rate-limiting step and sets the local CR observables. There is good agreement between our solution and the observations, especially given the simplicity of our model. As we do not include ionisation losses in the B/C calculation, we do not compare our model predictions to low-energy B/C data measured by Voyager (Cummings et al. 2016), which probe CRs in the energy-loss dominated regime. In Figure 5.5 we instead only show rigidities that correspond to kinetic energies > 1 GeV/nucleon, for which we find that escape dominates over ionisation losses, i.e. $\tau_{\rm ion} \gg h/{\rm v_A}$, consistent with our model assumptions (at these energies solar modulation also does not significantly affect B/C; e.g. Aloisio et al. 2015, Bresci et al. 2019).

While Figure 5.5 suggests that there exist parameters for which there is good agreement between the model and the data, this is not generally the case. We consider variations in our parameters and the propagation model in Figure 5.6. We show results for B/C in the top panel, and for the proton spectrum in the bottom panel. We include proton ionisation losses using eq. 5.33 to enable comparison to both AMS-02 and low-energy Voyager measurements (Aguilar et al. 2015; Cummings et al. 2016). We again do not include energy losses in our calculation of the B/C spectra, as our simple implementation of the loss term in (5.1) is not appropriate for B particles (see Appendix 5.C). We therefore compare our calculated B/C only to AMS-02 data points with $E > 1$ GeV/nucleon. The black lines all correspond to the fiducial propagation model (Figure 5.4) for different choices of parameters (for example, $n_{\rm W,fid}/3$ in the legend means that the density in the warm ISM is decreased by a factor of 3 relative to the fiducial parameter). Importantly, the proton spectrum is significantly steeper than the injection spectrum (red dotted line) for energies $\gtrsim$ 10GeV. There is a break in the spectra between $10^2$ and $10^3$ GeV, which comes from the transition between $f_p(0)$ being set by self-confinement in the halo and $f_p(0)$ being set by turbulent diffusion in the warm ISM. The transition between self-confinement and extrinsic turbulence dominating the transport also gives rise to a spectral break in phenomenological models based on undamped Kolmogorov-like turbulence (e.g., Blasi et al. 2012; Aloisio & Blasi 2013; Aloisio et al. 2015). However, unlike these phenomenological models, modest changes in our model parameters can introduce significant deviations from a power-law (e.g., the dotted line), which is incompatible with the observations. The real ISM samples a wide range of plasma conditions and so CR transport likely is some convolution of our toy model's results for different parameters. It is unclear what the resulting CR spectrum would be in a more realistic multiphase model.

The green line in Figure 5.6 shows the prediction for a different propagation model, in which CRs are scattered by fast-mode turbulence throughout the entire volume of the galaxy, including the halo and hot ISM (i.e. no self-confinement at all). We use the diffusion coefficients from Figure 5.3 and a 5kpc halo size to adjust the normalisation. The green line shows qualitatively very different behaviour from our multiphase model. Due to the weak energy dependence of $\kappa_{\rm turb}$ in the halo (Figure 5.3), B/C is essentially flat and the proton spectrum $\sim p^{-\gamma_{\rm inj}}$. The blue dashed line in the bottom panel of Figure 5.6 is the proton spectrum calculated using the self-confined halo only (no warm ISM contribution; same as the dashed line in Figure 5.5 for B/C). Without scattering by fast-mode turbulence, the spectrum at high energies is too steep (see discussion under eq. 5.14).

Figure 5.6 thus shows that neither fast-mode turbulence nor self-confinement alone give reasonable agreement with the observations. The combination in principle can, but it requires a significant amount of fine-tuning of plasma conditions. This includes our rather uncertain assumption of a weakly turbulent halo, such that linear damping of self-excited Alfvén waves by ambient Alfvénic turbulence does not inhibit self-confinement at $R \sim 100$ GV. We briefly note that this fine-tuning issue is not a consequence of multi-phase CR propagation, but is also present if CRs are scattered by self-excited Alfvén waves and fast-mode turbulence in the same ISM phase, as we discuss in Appendix 5.B. It is also worth stressing that models that use self-confinement + damped fast-mode turbulence require significantly more fine-tuning than phenomenological models based on self-confinement + undamped Kolmogorov turbulence. In these phenomenological models, CR injection rates and streaming speeds are constrained by observables at low energies where self-confinement dominates the transport, while the turbulence strength is set by CR observables at high energies where self-confinement no longer operates. Combined, these two pieces of physics constrain CR transport. By contrast, the MHD fast-mode cascade is very sensitive to local plasma conditions, which strongly affect the damping rates. This is best illustrated by the solid and dotted black lines in Figure 5.6. The only parameters that are different between these two models are the warm ISM density and temperature. The turbulence strength at the outer scale is unchanged and yet the two models have very different predictions for the CR proton and B/C spectra. In addition to being theoretically unsatisfying, the fine-tuning required in Figures 5.5 and 5.6 may also face challenges explaining the relatively small spatial variations in CR spectra in the MW inferred from synchrotron and gamma-ray data (e.g., Miville-Deschênes et al. 2008; Acero et al. 2016; Yang et al. 2016).

## 5.6 Uncertainties in Fast-Mode Turbulence

We now discuss significant uncertainties in the physics of MHD fast-mode turbulence, which are usually not taken into account in the CR literature (e.g. Yan & Lazarian 2004; Yan & Lazarian 2008; Fornieri et al. 2021), but may strongly affect the CR diffusion coefficients from Figure 5.3 and the predicted spectra in Section 5.5. These uncertainties appear to primarily *increase* the discrepancy between phenomenological CR scattering models and weak fast-mode turbulence predictions, as we discuss below.

### 5.6.1 Suppression of the Weak Cascade by Wave Steepening

The governing principles of the MHD fast-mode cascade are still up for debate. It is uncertain what happens to the cascade below the viscous scale at low $\beta$, where anisotropic damping becomes important for non-parallel propagating modes. Moreover, the calculation from Yan & Lazarian (2008), used in Sections 5.4 and 5.5 and in Figures 5.3–5.6, hinges on the assumption that MHD fast modes (or hydro sound waves) indeed follow a weak-turbulence cascade as argued by Zakharov & Sagdeev (1970). However, the weak turbulence assumption is uncertain. Kadomtsev & Petviashvili (1973) instead argued that for sufficiently small viscosities sound waves inevitably steepen to form (weak) shocks. Indeed, even for low-Ma turbulence the sound-wave steepening timescale is significantly shorter than the weak-turbulence nonlinear interaction (cascading) timescale,

$$\frac{\tau_{\rm steepen}^{-1}}{\tau_{\rm casc}^{-1}} \sim \frac{k\delta {\rm v}}{k\delta {\rm v}^2/{\rm v}_{\rm ph}} \sim \frac{{\rm v}_{\rm ph}}{\delta {\rm v}} \gg 1, \tag{5.34}$$

where ${\rm v}_{\rm ph}$ is the phase speed of sound waves. In the Kadomtsev & Petviashvili (1973) picture steepened fast modes produce weak shocks and then follow a $k^{-2}$ spectrum. This spectrum was indeed observed by Kowal & Lazarian (2010) in their compressible MHD simulations and, more recently, by Makwana & Yan (2020) in their sub-sonic turbulence simulations.

After sound waves steepen and form weak shocks, they dissipate on a timescale of order the steepening timescale (Landau & Lifshitz 1959),

$$\tau_{\rm diss}^{-1} \sim k\delta {\rm v}. \tag{5.35}$$

Both the steepening and weak-shock dissipation timescales are thus shorter than the weak-cascade timescale by a factor $\delta {\rm v}/{\rm v}_{\rm ph}$ (eq. 5.34). This factor is $\ll 1$ deep inside the inertial range on small scales resonant with $\lesssim$ PeV CRs. It is therefore likely that wave steepening suppresses the weak cascade and so the $k^{-3/2}$ spectrum.

CR scattering in a field of weak shocks is qualitatively different from standard pitch-angle diffusion in wave turbulence. In the latter case, CRs undergo frequent uncorrelated small changes in pitch angle through wave-particle interactions. By contrast, in a field of weak shocks CRs are likely scattered by single strong events at the shock discontinuities, if the width of the shock is $\ll$ the CR gyroradius. Thus, for CRs with sufficiently large gyroradii (energies), the scattering mean free path is constant and set by the separation of the shocks, which is likely of the same order of magnitude as the turbulence injection scale. The width of the shock is of the order of the ion mean free path in the collisional case ($w \sim l_{\rm mfp}{\rm v}_{\rm th}/\delta V$), and much smaller than the mean free path in the collisionless case ($\sim$ ion gyroradius). In the warm ionized medium the mean free path is of order $10^{13}$ cm, and so CRs with with energies $E \gtrsim$ 10GeV can be scattered by this mechanism. We note, however, that it is not fully clear how the steepening of sound waves in a low-collisionality plasma progresses once the spatial scales approach the ion mean free path.The perturbation might transition to a collisional shock or in some cases generate collisionless velocity space instabilities that alter the subsequent evolution from what would be predicted by fluid theory. While this weak-shock-mediated CR transport could be tuned to have the right normalisation, the transport is energy independent, in disagreement with the empirical $\kappa \propto E^{0.3-0.7}$ (see blue dotted line in Figure 5.7). This is a problem especially at high energies where self-confinement certainly no longer operates. At lower energies, $\lesssim$TeV, there can in principle still be an energy dependence to CR transport due to scattering by a combination of self-excited waves and extrinsic weak shocks.

### 5.6.2 Anisotropic Fast-Mode Turbulence at Low $\beta$

Yan & Lazarian (2004) and Yan & Lazarian (2008) assume that in the absence of damping the

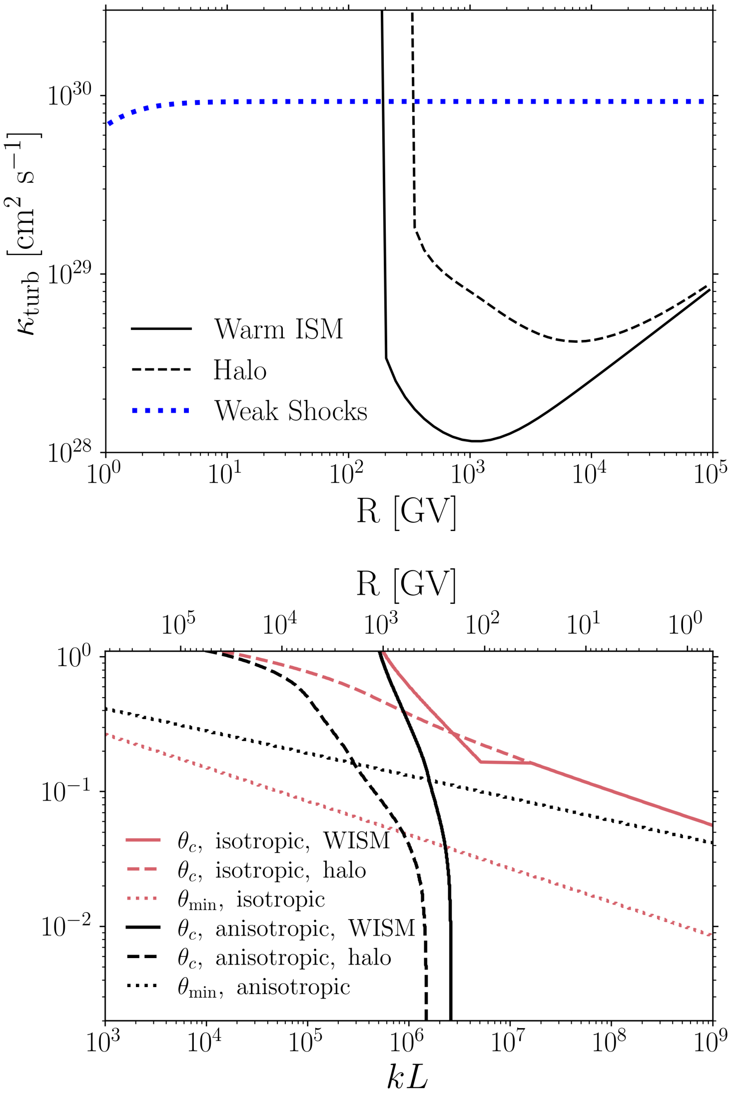


*Figure 5.7*: Uncertainties in MHD fast-mode turbulence significantly affect CR transport. Top: the blue dotted line shows a roughly energy-independent CR diffusion coefficient if wave steepening suppresses the weak cascade of fast modes (Section 5.6.1). We assume $O(1)$ shocks separated by $\sim$ 10 pc. The black lines are CR diffusion coefficients calculated using the low-$\beta$ anisotropic scalings for a weak fast-mode cascade from (5.36)–(5.39) for the same warm ISM and halo plasma conditions as in Figure 5.3. Due to the slow cascade rate of quasi-parallel modes (eq. 5.36) and field-line wandering (eq. 5.39), CR scattering is completely suppressed for rigidities $\lesssim$ 100 GV. Bottom: the critical mode pitch angle $\theta_c$ as a function of spatial scale. Fast modes with $\theta > \theta_c(k)$ are fully damped. The red solid and dashed lines show $\theta_c$ for the isotropic fast-mode scalings in Section 5.4 and Figure 5.3. The black solid and dashed lines are for the anisotropic fast-mode scalings from equations 5.36–5.39 (and correspond to the black solid and dashed lines in the top panel). In the anisotropic case $\theta_c$ shrinks significantly faster with increasing $k$. On scales $kL \gtrsim 10^6$ (CR rigidities $\lesssim$ few hundred GV) $\theta_c$ is smaller than $\theta_{\rm min}$, where $\theta_{\rm min}$ is the minimum well-defined pitch angle a fast mode can have over one cascade timescale due to field-line wandering. As a result, the cascade is fully damped for $kL \gtrsim 10^6$.

fast-mode power spectrum is isotropic, $P \propto k^{-3/2}$, which is largely based on the numerical results in Cho & Lazarian (2003). While fast-modes are very likely isotropic at high $\beta$, where they behave essentially like hydro sound waves, this is probably not true at low $\beta$ (which is the most important regime for CR scattering). In particular, at low-$\beta$ quasi-parallel fast modes are significantly less compressive than oblique fast modes. At low $\beta$, $\delta \mathrm{v}_{\|}/\delta \mathrm{v}_{\perp} \sim \beta \sin\theta \cos\theta$ (where $\|$ and $\perp$ denote the velocity-fluctuation components parallel and perpendicular to $\boldsymbol{B}$, and $\theta$ is the angle between $\boldsymbol{k}$ and $\boldsymbol{B}$). This means that for quasi-parallel modes $\delta \mathrm{v}_{\|}/\delta \mathrm{v}_{\perp} \sim \beta \sin\theta \ll 1$ and $\boldsymbol{k} \cdot \boldsymbol{\delta \mathrm{v}} \sim k \delta \mathrm{v} \sin\theta \ll k\delta \mathrm{v}$. Quasi-parallel fast modes therefore interact significantly more weakly than quasi-perpendicular modes. This plausibly generates anisotropy in the cascade. In particular, the cascade timescale becomes angle-dependent,

$$\tau_{\mathrm{casc}}^{-1} \sim k \frac{\delta \mathrm{v}^2}{\mathrm{v_{ph}}} \sin^2\theta, \tag{5.36}$$

(cf. equation 5.24). If turbulence is forced isotropically at the outer scale (the energy transfer rate $\epsilon$ is isotropic), then

$$\frac{\delta \mathrm{v}}{\mathrm{v_{ph}}} \sim \mathrm{Ma}(kL)^{-1/4}(\sin\theta)^{-1/2}, \tag{5.37}$$

where Ma is defined as $\delta \mathrm{v}/\mathrm{v_{ph}}$ at the injection scale for $\sin\theta \sim 1$. This corresponds to a power spectrum,

$$P \propto k^{-3/2}/\sin\theta \tag{5.38}$$

(see also Chandran 2005 for a more rigorous weak-turbulence-theory derivation of this result). The power does not diverge as $\theta \to 0$ due to the effect of field-line random walk. Before it cascades, a fast mode with wavenumber $k$ travels a distance $\mathrm{v_{ph}}\tau_{\mathrm{casc}}(k, \theta)$. A quasi-parallel mode which cascades at a rate given by (5.36) experiences field-line wandering due to $\mathrm{Ma_{Alf}} \sim 1$ Alfvénic turbulence of order,

$$\frac{\delta B}{B} \sim \left(\mathrm{Ma}^2 (kL)^{1/2} \sin\theta\right)^{-1/2} \tag{5.39}$$

(cf. equation 5.27). A mode has a well-defined pitch-angle over a cascade time only if $\theta \gtrsim \delta\theta \sim \delta B/B$. As in Section 5.4, when we evaluate the cascade (eq. 5.36) and damping (eq. 5.45 and 5.46) rates, we take the average of the cascade/damping rate in the interval $(\theta - \delta\theta, \theta + \delta\theta)$. According to (5.39), $\delta\theta$ diverges as $\theta \to 0$. This is, however, unphysical given the presence of finite field-line wandering. We remedy this by imposing a lower limit on $\theta$ in (5.39), $\theta_{\mathrm{min}}(k)$, defined such that $\delta\theta(k, \theta_{\mathrm{min}}) = \theta_{\mathrm{min}}$. $\theta_{\mathrm{min}}$ is therefore the smallest well-defined pitch-angle a fast mode can have. For an isotropic fast-mode cascade, $\theta_{\mathrm{min}}$ is simply equal to the pitch-angle spread due to field-line wandering in (5.27), $\theta_{\mathrm{min}} \sim \delta B/B$.

In the top panel of Figure 5.7, the black lines show CR diffusion coefficients calculated using the anisotropic scalings of the fast-mode cascade from (5.36)–(5.39) for the same warm ISM and halo plasma conditions as in Figure 5.3. CR scattering is completely suppressed for rigidities $\lesssim 100$ GV. The suppression of CR scattering at small energies relative to Figure 5.3 is due to the slower cascade rate of quasi-parallel modes (eq. 5.36) and increased field-line wandering experienced by quasi-parallel modes during one cascade time (eq. 5.39). The solid and dashed lines in the bottom panel of Figure 5.7 show how the critical mode pitch angle $\theta_c$ depends on spatial scale. Fast modes with $\theta > \theta_c(k)$ are fully damped. The solid and dashed lines therefore show how the cone of undamped fast modes shrinks with spatial scale. The red solid and dashed lines show $\theta_c$ for the isotropic fast-mode scalings in Section 5.4 and Figure 5.3. On collisionless scales $\theta_c \propto k^{-1/4}$ (see Section 5.4), and because $\theta_{\rm min}$ (red dotted line) has the same scaling with $k$ (eq. 5.27), the fast-mode cascade is not fully damped. This is in contrast to the results for the anisotropic fast-mode scalings from equations 5.36–5.39. In this case $\theta_c$ (black solid and dashed lines) shrinks significantly faster with increasing $k$ than $\theta_{\rm min}$ (black dotted line). On scales $kL \gtrsim 10^6$ (corresponding to CR energies of a few hundred GeV) $\theta_c$ is smaller than $\theta_{\rm min}$ and the cascade is fully damped.

We also note that the $\theta_{\rm min}$ lines in Figure 5.7 are assumed to be due to Alfvénic turbulence. There is, however, also the additional field-line wandering generated by the cascade of fast modes. This may truncate the fast-mode cascade on even larger scales. The field-line wandering experienced by a quasi-parallel fast-mode during one cascade time due to undamped fast-mode turbulence is roughly,

$$\frac{\delta B}{B} \sim \left(\frac{\mathrm{Ma}^2}{(kL)^{1/2}\sin^3\theta}\right)^{1/4} \tag{5.40}$$

(for an isotropic fast-mode cascade, the $\sin^{-3}\theta$ factor should be dropped). Due to the weak dependence on $k$, $\propto (kL)^{-1/8}$, weak fast-mode turbulence can be the dominant source of field-line wandering experienced by high-$k$ modes. The weaker dependence on $k$ relative to eq. 5.39 is due to the fact that the weak-turbulence fast-mode spectrum ($\propto k^{-3/2}$) is shallower than the Alfvénic-turbulence spectrum in the parallel direction ($\propto k_\parallel^{-2}$). Importantly, the field-line wandering experienced by high-$k$ modes is due to longer-wavelength modes on scales $\mathrm{v_{ph}}\tau_{\rm casc}(k) \gg k^{-1}$ which may not be affected by damping. As a result, assuming an undamped cascade in (5.40) turns out to be a reasonable assumption in e.g. the warm ISM.

As we noted in Section 5.6.1, it is unclear whether weak-turbulence theory is at all applicable to the MHD fast-mode cascade given that wave steepening and weak-shock dissipation occur on a shorter timescale than the weak-turbulence cascade.[4] In this section we showed that if theoretically motivated cascade anisotropies at $\beta \lesssim 1$ are taken into account, then even within the weak-turbulence theory framework CR scattering by fast modes is completely suppressed at energies $\lesssim$100 GeV. The anisotropy of weak fast-mode turbulence at low $\beta$ has not yet been observed in simulations (e.g. Cho & Lazarian 2003). However, this is not very surprising given that the weak cascade itself remains elusive. Moreover, close to the injection scale field-line-wandering effects are significant and act to suppress the anisotropy of the cascade.

## 5.7 Discussion

The good agreement between the multi-phase model and observations in Section 5.5.2 may be viewed as a success. After all, there appears to be *a* set of simplified yet plausible conditions where

---

[4] This is also true if fast-mode anisotropies at low $\beta$ are taken into account. For $\beta \ll 1$, fast modes steepen at a rate $\tau_{\rm steepen}^{-1} \sim k\delta\mathrm{v}\sin\theta$. At low $\beta$ quasi-parallel fast modes therefore steepen at rate that is slower than at high $\beta$, but the steepening rate is still faster than the weak-turbulence cascade rate (eq. 5.36) by a factor $\mathrm{v_{ph}}/(\delta\mathrm{v}\sin\theta)$.

existing theories of CR propagation meet observations. In that sense, our results can be interpreted as a proof of concept for existing theoretical models of CR transport. However, the success of a theory depends not only on its ability to match the data, but also on its robustness. For this reason, the good fit in Section 5.5.2 should not overshadow the uncertain micro-physical building blocks of our model (e.g., the significant uncertainties in the MHD fast-mode cascade discussed in Sections 5.6.1 and 5.6.2). The fine-tuning of parameters that is necessary to create smooth spectra with the right energy dependence may also, understandably, leave a bittersweet taste. Especially since the fine-tuning carries a message of non-universality: in this model other Milky-Way-like galaxies, even ones without calorimetric CR losses, may have very different CR spectra. And one might expect rather different spectra in different regions of the Milky Way, which is naively not consistent with synchrotron and gamma-ray data (e.g., Miville-Deschênes et al. 2008; Acero et al. 2016; Yang et al. 2016).

In this section we discuss some speculative extensions of self-confinement and MHD turbulence theory that might improve agreement between CR scattering theory and phenomenological models. We begin by noting that alternative nonlinear damping mechanisms of self-excited waves would modify the energy dependence predicted by self-confinement theory, potentially yielding better agreement with observations (Section 5.7.1). A more theoretically attractive alternative is scattering of CRs by a single mechanism, which would be devoid of the fine-tuning problem discussed above, e.g., scattering by a turbulent cascade with a universal spectral index. This is in part why the phenomenological Kolomogorov-turbulence model has been so successful in explaining CR observations. However, the Kolmogorov phenomenology is not theoretically well motivated and at present no MHD cascade is known to have the desired properties over the entire range of CR energies. Still, MHD turbulence remains sufficiently uncertain to not discard this option. In Sections 5.7.3–5.7.5 we discuss a few topics at the frontier of MHD turbulence theory that may be relevant for a universal, turbulence-based theory of CR transport. In Section 5.7.2 we also briefly consider the possibility that the waves that scatter CRs are *excited* on a wide range of scales by an external driving process (e.g., stellar feedback), instead of only being generated via a turbulent cascade from large scales.

### 5.7.1 Nonlinearities in Self-Confinement Theory

We showed in Section 5.3.1.2 that nonlinear damping of self-excited Alfvén waves gives rise to a smooth transition from Alfvénic streaming to energy-dependent transport with constant spectral index (in contrast to linear damping). However, the spectral index for nonlinear Landau damping is too steep relative to observations (Section 5.3.1.2). This is not necessarily true for nonlinear damping mechanisms that are different from $\Gamma_{\rm NLLD} \sim k v_{\rm th}(\delta B/B)^2$. For example, Alfvén waves on scales larger than the thermal-particle mean free path are nonlinearly damped by pressure-anisotropy effects at a rate $\Gamma \sim (k l_{\rm mfp})\Gamma_{\rm NLLD}$ (Squire et al. 2017). With this extra $k$ dependence, nonlinear damping yields energy-dependent transport very close to the observationally inferred $E^{0.3-0.7}$ scaling. However, this particular physical mechanism is only valid above the thermal-particle mean free path, which is very large ($\gtrsim$pc) in the dilute and hot plasmas in which these nonlinear damping mechanisms are important. This particular damping is thus not applicable for CRs below PeV energies. Nevertheless, this example illustrates that other nonlinear damping mechanisms can in principle give rise to appropriate energy dependence. The nonlinear damping would need to be strong, and linear damping negligible, so as to not introduce a sharp cutoff as in Figure 5.1.

### 5.7.2 External Driving on a Range of Scales

It is also possible that the waves that scatter CRs over a wide range of energies are directly excited by some external driving process. This could be due to fluid instabilities acting on a range of scales, e.g. Kelvin-Helmholtz or dust-driven instabilities (Squire & Hopkins 2018). Alternatively, it is possible that stellar feedback (winds, supernovae, etc.) drives not only large-scale motions (which then cascade down to smaller scales), but also fluctuations on a wide range of smaller scales. Driving on smaller scales leads to less anisotropy in the Alfvénic cascade and thus more efficient CR scattering. For example, the periods of the Alfvén waves resonant with GeV CRs are comparable to the variability timescale of solar-type stars, and so it is in principle possible that stellar winds interacting with the magnetized ISM can contribute to some scattering of lower-energy CRs. The power required to excite waves with amplitudes consistent with phenomenological CR transport models is $\ll$ the total power from stellar feedback. In particular, the power required to excite $\delta B/B \sim 10^{-3}$ waves (rough amplitude suggested by phenomenological models) in a cylindrical galaxy with radius $R = 10$kpc and height $H = 1$kpc is $P_{\rm scatt} \sim 4 \times 10^{36}$ ergs/s, assuming $B \sim 1\mu$G and a plausible wave damping rate $\Gamma \sim 10^{-11}{\rm s}^{-1}$. This is indeed orders of magnitude smaller than the total power injected by supernovae, $P_{\rm SN} \sim 3 \times 10^{41}$ ergs/s for a supernova rate of 1 per century. However, this speculation faces a severe fine-tuning problem, as it is unclear how waves excited on a range of scales would conspire to have a spectrum that resembles an undamped turbulent cascade.

### 5.7.3 Cascade Damping: Role of the Plasma Echo

As shown in Figure 5.3 and Figure 5.7, damping prevents a weak-turbulence cascade of fast modes from maintaining a universal $k^{-3/2}$ power-law at high $k$ (or $\kappa_{\rm turb} \propto E^{1/2}$ at scales resonant with low-energy CRs). The damping sets in at particularly large scales in dilute and hot plasmas like the hot ISM or the galactic halo, which constitute most of the confining volume. In these dilute systems, linear compressive waves are Landau-damped, which transfers energy from spatial fluctuations to ever-finer structures in velocity space in the thermal-particle distribution function, a process called phase-mixing.

Our calculation (based on the calculation in Yan & Lazarian 2008) assumes that compressive fluctuations in a turbulent cascade are damped at their linear damping rates. However, recent theoretical and numerical progress suggests that in the presence of nonlinear interactions and sufficiently small particle collisionalities (so that phase mixing can be reversible), fluctuations that have phase-mixed away to fine velocity scales can be brought back by coupling to other perturbations, a process called the “stochastic plasma echo", which effectively reverses phase-mixing (Schekochihin et al. 2016; Meyrand et al. 2019). Landau damping is suppressed and the collisionless plasma behaves more fluid-like. The fluidization of plasma turbulence appears consistent with the seemingly undamped cascade of compressive fluctuations in the solar wind (in this case, slow modes; Chen 2016; Verscharen et al. 2017). An analogous fluidization of fast-mode turbulence would allow $\kappa_{\rm turb}$ in the halo (Figure 5.3) to continue its $R^{0.5}$ scaling to lower CR energies.

One issue with this idea is that the plasma echo can anti-phase-mix fluctuations before they are damped by collisions only if the nonlinear timescale is sufficiently short. This in turn requires large amplitudes for the turbulent fluctuations, which then leads to small CR diffusion coefficients if the desired $E^{0.5}$ scaling continues to small energies. For example, using the results by Adkins & Schekochihin (2018), we estimate that for the stochastic echo to operate in the hot ISM, amplitudes of order Ma $\gtrsim 1$ on the outer scale appear necessary, leading to diffusion coefficients of order $10^{27}$ cm$^2$ s$^{-1}$ at a GeV, far too small compared to observationally inferred CR diffusion coefficients. However, the theory of the stochastic echo in a fast-mode cascade has not been developed and clearly merits more work. Moreover, if the conditions to realise the stochastic echo only occur in a small fraction of the ISM volume, this would reduce the effective volume-averaged diffusion coefficient.

### 5.7.4 Reconnection and the Anisotropy of Alfvénic Turbulence

An alternative to scattering by fast-mode turbulence is that there is a different, yet unknown, regime of Alfvénic turbulence that can efficiently scatter CRs. Though critical balance as the main governing principle of MHD turbulence stands on firm ground, it is possible that additional physics on small scales can change the cascade physics. For example, the large aspect ratios of structures formed in strong turbulence on small scales plausibly become unstable to instabilities such as tearing and/or Kelvin-Helmholtz. In particular, Alfvénic turbulence may enter a new regime of tearing-mediated turbulence on scales much larger than one would naively expect, i.e. $\gg$ the resistive scale (Loureiro & Boldyrev 2017; Mallet et al. 2017a; Mallet et al. 2017b). There is some numerical evidence/theory that turbulence in reconnecting layers may obey different anisotropy scalings than standard Alfvénic turbulence, e.g. $k_\parallel \sim k_\perp$ (Huang & Bhattacharjee 2016; Boldyrev & Loureiro 2019), or even $k_\parallel \sim k_\perp^{1.2}$ (Yang et al. 2020). This has the potential for more efficient CR scattering and the latter scaling could even give rise to an energy dependence of CR scattering that is roughly in the right ballpark. However, applying the theory of Mallet et al. (2017a), we estimate that tearing-mediated turbulence is expected to become important only on scales $\ll$ the gyroradius of GeV CRs. If this is generally true, it would not be important for CR scattering.

### 5.7.5 Balanced versus Imbalanced Turbulence

Numerical studies of MHD turbulence usually assume approximate equipartition between kinetic and magnetic energies, and roughly equal energies in forward and backward propagating Alfvén waves. While this setup is numerically convenient, it is generally not realised in nature. Many of the cascade properties, including the spectral slopes and anisotropy, of imbalanced Alfvénic turbulence are quite different from the usual balanced case (e.g., Beresnyak & Lazarian 2009 or Section 7 in the review by Schekochihin 2020 for a summary). However, it is believed that the parallel spectrum remains steep, with a spectral slope $\lesssim -2$ and $k_\perp \gg k_\parallel$. This still means inefficent CR scattering. Nevertheless, imbalanced MHD turbulence remains a fairly unexplored terrain compared to its balanced counterpart, so there is still room for surprises there.

## 5.8 Summary

Neither confinement of CRs by self-excited Alfvén waves nor scattering by an ambient fast-mode cascade can plausibly explain CR observations over the entire range of CR energies (e.g., Kulsrud & Cesarsky 1971; Farmer & Goldreich 2004; Fornieri et al. 2021). We provide a pedagogical review of this important result in Sections 5.3 and 5.4. Due to the strong damping of fast modes on small scales, the energy dependence of the CR diffusion coefficient in fast-mode turbulence is incompatible with observations at CR energies $\lesssim 10^3$ GeV (Figure 5.3 and Figure 5.7). The discrepancy is significantly larger than in phenomenological models which assume CR scattering by an undamped isotropic turbulent cascade (e.g., Trotta et al. 2011; Gaggero et al. 2014; Hopkins et al. 2021a). In self-confinement theory, higher-energy CRs are unable to self-confine due to the linear damping of Alfvén waves excited by the streaming instability. Indeed, for self-excited waves damped by linear damping mechanisms, self-confined CRs stream, to reasonable approximation, at either the Alfvén speed or the speed of light (Figure 5.1), which is naively incompatible with the smooth energy dependence measured empirically. Non-linear damping of self-excited waves (e.g., non-linear Landau damping) introduces a smoother energy-dependence to CR transport but existing theoretical models predict an energy dependence that is stronger than that observed (see Section 5.3.1.2 and the dashed line in Figure 5.5).

CR transport is theoretically predicted to depend sensitively on spatial variations in the plasma properties because local plasma conditions change the strength of ambient turbulence, the efficiency of fast-mode damping and the damping of self-excited Alfvén waves. This strong dependence of CR transport on local plasma conditions motivates our considering how such variations might affect CR observables at Earth. To start, we have shown that even pure Alfvénic streaming can in fact produce energy dependent transport and energy dependent observables if both the Alfvén speed and the region in which CRs of a given energy are self-confined vary with height above the galactic disk (see also Holmes 1974). These two conditions mean that CRs of different energy effectively sample different Alfvén speeds, leading to energy dependent transport that can in principle be similar to that observed (Section 5.3.2 and Figure 5.2). Even in this model, however, self-excited Alfvén waves are eventually fully damped at sufficiently high energies and a separate CR scattering mechanism is required.

The fact that neither self-confinement theory nor scattering by weak fast-mode turbulence can, on their own, explain CR spectra in the Milky Way suggests two possibilities. There is either a yet unidentified process that efficiently scatters CRs with the right energy dependence, or, a combination of scattering by self-excited waves and fast-mode turbulence conspires to mimic the empirically derived CR diffusion coefficient. In Section 5.5 we considered the latter possibility. The multi-phase nature of the ISM of galaxies independently suggests that CR transport may be multi-modal, i.e. a mix of streaming and diffusion in turbulence. Self-confinement likely operates in regions where the turbulent cascade of scattering waves (e.g., fast modes) is strongly damped, whereas in regions with efficient scattering by turbulence the streaming instability is suppressed. In Section 5.5.1, we have considered a particular model in which CRs diffuse in the warm ISM due to scattering by the MHD fast-mode cascade, and are self-confined in the coronal regions and the halo (Figure 5.4). We have assumed that turbulence in the halo is weak, so that nonlinear Landau damping is the dominant damping of Alfvén waves excited by the streaming instability. We stress that this particular model belongs to a broader class of possible multi-phase CR propagation models. In particular, the exact filling fractions of different phases is likely a function of a number of properties of the galaxy, such as the local gas surface density and star formation rate surface density, which set the cooling rates and supernovae heating rate. This will in turn modify CR propagation. The purpose of the model considered in this work is to demonstrate that a relatively simple model of multi-phase CR transport can in principle explain some of the main trends observed in CR spectra in the solar neighbourhood (Figure 5.5 and Figure 5.6), but that significant fine-tuning of plasma parameters is needed. These trends are difficult to explain using only self-confinement theory (Section 5.3 and, e.g., dashed line in Figure 5.5) or only scattering by MHD turbulence (Section 5.4, Figure 5.3 and the green line in Figure 5.6).

The calculation in Section 5.5 is the first attempt to combine microphysical theories of CR self-confinement and scattering by MHD fast-mode turbulence. While there already exists literature that tries to combine streaming and turbulence to explain CR spectra measured in the MW (e.g. Blasi et al. 2012; Aloisio & Blasi 2013; Aloisio et al. 2015), these models assume undamped Kolmogorov-like ($\sim$ isotropic Alfvénic) turbulence. In this work, we instead considered the theoretically better motivated interplay of the streaming instability and fast-mode turbulence. Because fast modes are damped on scales $\sim$ the Larmor radius of $\lesssim$ TeV particles, their impact on CR transport and their interaction with self-excited Alfvén waves are very different from results based on an undamped Kolmogorov cascade (see also Appendix 5.B). In particular, Section 5.5 highlights the important issue that, due to fast-mode damping, CR propagation models based on a combination of self-excited Alfvén waves and MHD fast-mode turbulence require a significant amount of fine-tuning of plasma parameters to recover the almost-pure-power-law CR spectra measured in the Milky Way. This issue is not captured in phenomenological models based on undamped isotropic Kolmogorov-like turbulence.

Models of CR scattering by MHD fast-mode turbulence in the literature also rely on the uncertain assumption that fast modes follow a weak isotropic cascade in the absence of wave damping (e.g. Yan & Lazarian 2004; Yan & Lazarian 2008; Xu & Lazarian 2018; Fornieri et al. 2021). As we discussed in Section 5.6, weak fast-mode turbulence is theoretically expected to be anisotropic at low $\beta$ (Section 5.6.2), and is probably completely suppressed by wave steepening (Section 5.6.1). This raises the significant possibility that fast modes are unimportant at scattering CRs.

Interpreting CR data using a combination of scattering mechanisms in different phases of the ISM is physically well-motivated, but will inevitably require some fine-tuning (especially if the turbulent cascade is strongly damped). Moreover, CR scattering by fast-mode turbulence is likely inefficient due to wave steepening. An attractive alternative is that CR scattering is dominated by a currently unidentified source of a roughly Kolmogorov-like cascade, as is often assumed in phenomenological models. Such a cascade is not well motivated by current theoretical models of MHD turbulence. In Section 5.7, we discussed some of the uncertain aspects of MHD turbulence theory that could bear on the presence of such a cascade and thus on CR scattering. In this context it is worth noting that the energetic requirements for confining CRs by scattering are very small compared to the overall energetics of interstellar turbulence (Section 5.7.2). Thus, an energetically minor component of the cascade or turbulent driving by an energetically subdominant channel could nonetheless dominate CR scattering.

In this work we have focused on CR transport along field lines. There is a separate question of how the geometry of magnetic-field lines, e.g. turbulent field-line wandering, affects CR transport perpendicular to the mean magnetic field. We have ignored this aspect of CR transport in this work because the perpendicular diffusion time is expected to be orders of magnitude longer than the parallel diffusion time (e.g., Giacalone & Jokipii 1999; Desiati & Zweibel 2014; Dundovic et al. 2020). Nevertheless, given that the magnetic field is on average close to being planar in the galactic disk and that perpendicular diffusion of CRs remains quite uncertain, it is possible that perpendicular transport is important for regulating CR escape from the Galaxy and setting CR observables.

In the course of this work, we became aware of similar efforts to reconcile CR transport theories with MW observables by Hopkins et al. (2021b). Using galaxy-formation simulations combined with a broad selection of CR transport models motivated by microphysics, Hopkins et al. (2021b) also conclude that self-confinement theory and "extrinsic turbulence" theory alone cannot qualitatively reproduce CR spectra observed in the solar neighbourhood. As we showed in Section 5.3.2 and Section 5.5, the exact ISM phase structure and spatial variations of plasma properties can significantly affect CR observables and there exist multi-phase solutions (combinations of self-confinement and extrinsic turbulence) that can, in principle, qualitatively reproduce local CR spectra (Figure 5.6). These solutions require a particular set of fine-tuned ISM plasma conditions, which are not realised in the simulations of Hopkins et al. (2021b). Hopkins et al. (2021b) instead elucidate the form of driving/damping necessary for extrinsic turbulence and/or self-confinement to be compatible with MW measurements (analogous to our Section 5.7).

## Acknowledgements

We thank Philip Hopkins and Jonathan Squire for many enlightening discussions and helpful comments on earlier drafts of this manuscript. We are grateful to the referee for very useful feedback on the paper. We also thank Susan Clark, Brandon Hensley, Chang-Goo Kim, Matthew Kunz, Eve Ostriker, Alexander Schekochihin, and Anatoly Spitkovsky for useful conversations. This research was supported by NSF grant AST-1715070 and a Simons Investigator award from the Simons Foundation.

# Appendix

## 5.A CR Diffusion Coefficient in Weak Fast-Mode Turbulence

To calculate the CR diffusion coefficient in isotropic weak fast-mode turbulence, we use the results of Yan & Lazarian (2008). We use the same notation as Yan & Lazarian (2008) and define the cosine of the wave pitch angle $\xi = \cos\theta$, the dimensionless wavenumber $x = kL$, the dimensionless rigidity $R = \mathrm{v}/(L\Omega)$ and the perpendicular wavenumber normalised by the CR gyroradius, $w = k_\perp r_L = x_\perp R(1-\mu^2)^{1/2}$, where $\mathrm{v} \approx \mathrm{c}$ is the speed of the CR particle, $L$ is the outer scale of the turbulence and $\Omega$ is the relativistic gyrofrequency. The CR pitch-angle diffusion coefficient due to gyroresonant scattering is then (Yan & Lazarian 2008),

$$D^G_{\mu\mu} = \mathrm{Ma}^2 \frac{\mathrm{v}\pi^{1/2}(1-\mu^2)}{2LR^2} \int_0^1 d\xi \int_1^{k_{\max}(\xi)L} dx \frac{x^{-5/2}\xi}{\Delta\mu} [J_1'(w)]^2 \exp\Big(-\frac{(\mu - (x\xi R)^{-1})^2}{\Delta\mu^2}\Big) \tag{5.41}$$

where $J_1$ is the Bessel function of the first kind and the prime indicates a derivative, $\Delta\mu^2 = \mathrm{Ma}(1-\mu^2)$ and the exponential reflects the broadening of the gyro-resonance condition $k_\parallel \mathrm{v}_\parallel \approx \Omega$ due to large-scale magnetosonic fluctuations (Völk 1973, Voelk 1975, Yan & Lazarian 2008). The CR pitch angle diffusion coefficient due to nonresonant scattering by transit-time damping (TTD) is,

$$D^T_{\mu\mu} = \mathrm{Ma}^2 \frac{\mathrm{v}\pi^{1/2}(1-\mu^2)}{2LR^2} \int_0^1 d\xi \int_1^{k_{\max}(\xi)L} dx \frac{x^{-5/2}\xi}{\Delta\mu} [J_1(w)]^2 \exp\Big(-\frac{(\mu - \mathrm{v_A}/(\mathrm{v}\xi))^2}{\Delta\mu^2}\Big). \tag{5.42}$$

The spatial diffusion coefficient due to gyroresonant and TTD scattering is (see, e.g., Zweibel 2017),

$$\kappa = \frac{\mathrm{v}^2}{4} \int_0^1 d\mu \frac{(1-\mu^2)^2}{D^G_{\mu\mu} + D^T_{\mu\mu}}. \tag{5.43}$$

Because nonresonant scattering by TTD is dominated by large-scale modes, $D^T_{\mu\mu}$ is essentially energy independent. Thus, energy-dependence in the CR diffusion coefficient comes from gyroresonant scattering, i.e. $D^G_{\mu\mu}$. The impact of wave damping on the cascade is reflected by the upper bound $= k_{\max}(\xi)L$ in the $x$ integral in equations 5.41 and 5.42. In particular, the angle-dependent cutoff $k_{\max}(\xi)L$ corresponds to the scale $k(\xi = \cos\theta_c)$ where the cascade timescale is equal to the wave damping timescale. As in Yan & Lazarian 2004 and Yan & Lazarian 2008, we here assume that fast modes with $\theta < \theta_c$ continue cascading to smaller scales unaffected by the damping, while the remaining modes are fully damped. However, we stress again that this assumption is quite uncertain. For example, non-linearities might transfer energy from small $\theta$ to larger $\theta$ where energy is dissipated, thus removing energy from the otherwise weakly damped parallel propagating modes.

For a fully ionized plasma, the main damping mechanisms are viscous damping on scales larger than the thermal ion mean free path and collisionless damping on scales smaller than the mean free path. $k_{\max}(\xi)L$ is then found by equating the cascade rate,

$$\tau_{\rm casc}^{-1} \sim \left(\frac{k}{L}\right)^{1/2} \frac{\delta V^2}{V_{\rm ph}} \tag{5.44}$$

($\delta V$ is the amplitude at the injection scale) and the relevant damping rate. For $\beta \ll 1$ in the collisional limit, i.e. $k_{\|} l_{\rm mfp} {\rm v_A}/{\rm v_{th}} \ll 1$, the damping rate is (Braginskii 1965)

$$\begin{aligned} \Gamma\left(k_{\|} l_{\rm mfp} \frac{\rm v_A}{\rm v_{th}} \ll 1\right) &= \frac{\nu_{\rm B} k^2}{6}(1-\xi^2) \\ &\sim \frac{\nu_{\rm B} k^2}{6}\theta^2 \qquad \theta \ll 1, \end{aligned} \tag{5.45}$$

where $\nu_{\rm B} \sim l_{\rm mfp} v_{\rm th}$ is the anisotropic Braginskii viscosity and in the second line we took the small-angle limit. For $\beta \ll 1$ and $k_{\|} l_{\rm mfp} {\rm v_A}/{\rm v_{th}} \gg 1$, the collisionless damping rate is (Ginzburg 1961)

$$\begin{aligned} \Gamma\left(k_{\|} l_{\rm mfp} \frac{\rm v_A}{\rm v_{th}} \gg 1\right) &= \frac{(\pi\beta)^{1/2} \sin^2\theta}{4\cos\theta} k {\rm v_A} \Big[ \left(\frac{m_e}{m_i}\right)^{1/2} \exp\left(-\frac{m_e}{\beta m_i \cos^2\theta}\right) + 5\exp\left(-\frac{1}{\beta \cos^2\theta}\right) \Big] \\ &\sim \frac{\sqrt{\pi\beta}\theta^2}{4}\left(\frac{m_e}{m_i}\right)^{1/2} k {\rm v_A} \qquad \theta \ll 1, \end{aligned} \tag{5.46}$$

where in the last step we again took the small-angle limit. We note that both damping mechanisms → 0 for modes propagating parallel to the local magnetic field ($\theta \approx 0$) We calculate $k_{\max}(\xi)L$ by equating the cascade rate in (5.44) with the damping rates in (5.45) and (5.46), and then evaluate the integrals in (5.41) and (5.42) numerically. To evaluate the cascade and damping rates in (5.44)–(5.46), we take the average of the cascade/damping rate in the interval $(\theta - \delta\theta, \theta + \delta\theta)$, where $\delta\theta$ is the spread in mode pitch angle experienced by a fast mode during one cascade time due to turbulent magnetic-field-line wandering. The field-line wandering due to ambient Alfvénic turbulence with ${\rm Ma_{Alf}} \sim 1$ experienced by fast modes with wavenumber $k$ and pitch angle $\theta$ can be roughly approximated as,

$$\frac{\delta B}{B} \sim \left[\frac{\cos\theta}{{\rm Ma}^2 (kL)^{1/2}} + \left(\frac{\sin\theta}{{\rm Ma}^2 (kL)^{1/2}}\right)^{2/3}\right]^{1/2}. \tag{5.47}$$

We stress that we include collisionless damping only on scales $< l_{\rm mfp} {\rm v_A}/{\rm v_{th}}$ and collisional damping only on scales $> l_{\rm mfp} {\rm v_A}/{\rm v_{th}}$. For example, if equating (5.44) and (5.46) yields $k_{\max,\|}(\xi) {\rm v_A}/{\rm v_{th}} < l_{\rm mfp}^{-1}$ for a $\xi$ that was not damped in the viscous regime, we correct it by setting $k_{\max,\|} = l_{\rm mfp}^{-1} {\rm v_{th}}/{\rm v_A}$, as collisionless damping is not the appropriate damping on scales $k_{\|} l_{\rm mfp} {\rm v_A}/{\rm v_{th}} < 1$. Figure 5.3 shows the results of this calculation.

## 5.B Simultaneous Scattering by Self-Excited Waves and Weak Fast-Mode Turbulence

In some phenomenological models of CR transport it is assumed that the self-excited waves cascade (diffuse in $k$-space) due to the ambient Kolmogorov-like turbulence, which produces a spectrum of waves slightly different from Kolmogorov in the wavenumber range affected by the streaming instability (e.g. Blasi et al. 2012; Aloisio & Blasi 2013; Aloisio et al. 2015). The self-excited Alfvén waves can in principle interact with either the Alfvénic component of the turbulence or the fast-mode component. While three-wave interactions between high-frequency self-excited Alfvén waves and turbulent fast modes can in principle occur at low $\beta$ (as the modes can satisfy frequency-matching conditions; Chandran 2005), the fast-mode cascade is significantly affected by damping in the wavelength regime where self-excitation is important (Figure 5.3). As a result, the fast-mode cascade differs significantly from Kolmogorov (Figure 5.3; it may in fact be completely suppressed on small scales, see Figure 5.7). Moreover, even in the absence of fast-mode damping, self-excited Alfvén waves are more likely to be sheared away to high $k_\perp$ by the background Alfvénic turbulence (Farmer & Goldreich 2004; Lazarian 2016). To see this, we compare the cascade time of the self-excited Alfvén waves due to fast modes, $\tau_{\rm casc}$, to the shearing time set by background Alfvénic turbulence, $\tau_{\rm shear}$. At low $\beta$ and for turbulence injected with $\mathrm{Ma_{Alf}} < 1$,

$$\frac{\tau_{\rm casc}^{-1}}{\tau_{\rm shear}^{-1}} \sim \Big(\frac{\mathrm{Ma}}{\mathrm{Ma_{Alf}}}\Big)^2, \tag{5.48}$$

while for $\mathrm{Ma_{Alf}} > 1$,

$$\frac{\tau_{\rm casc}^{-1}}{\tau_{\rm shear}^{-1}} \sim \Big(\frac{\mathrm{Ma}}{\mathrm{Ma_{Alf}}}\Big)^{3/2}\mathrm{Ma}^{1/2}, \tag{5.49}$$

Most of the turbulent energy is expected to reside in the Alfvénic branch, $\mathrm{Ma_{Alf}} > \mathrm{Ma}$, and so $\tau_{\rm shear} < \tau_{\rm casc}$, i.e., the cascading of self-excited Alfvén waves to higher $k_\parallel$ by fast-mode turbulence is likely ineffective. We reiterate that damping reduces the power in fast-mode turbulence on small scales, further increasing $\tau_{\rm casc}/\tau_{\rm shear}$ and reinforcing our conclusion.

This implies that self-excited Alfvén waves and weak-fast mode turbulence act as independent scatterers of CRs. Their amplitudes and thus CR scattering rates are set by very different physics (different driving, cascading and damping mechanisms). Combined with the non-uniform scaling with energy of $\kappa_{\rm turb}$ in Figure 5.3, this implies that generating CR spectra with a roughly constant power-law slope over a wide range of energies using a combination of these scattering mechanisms requires significant fine-tuning of plasma conditions, regardless of whether CR propagation is single- or multi-phase. This is in contrast to phenomenological models based on an undamped Kolmogorov cascade, which do not face such significant fine-tuning issues.

## 5.C Impact of Ionisation Losses on the B/C Spectrum

As we pointed out in Section 5.3, the energy loss term from eq. 5.1 is more correctly written as a flux in momentum space. Using the correct form turns out to be particularly important for B nuclei at low energies, as we now show. In the loss regime, B nuclei satisfy,

$$-\frac{1}{p^2}\frac{\partial}{\partial p}\Big(p^2\frac{p}{\tau_{\rm loss}(p)}f_B\Big) = \frac{f_c}{\tau_{\rm spall}} = n\sigma_{\rm spall}{\rm v}f_C, \tag{5.50}$$

where the energy loss rate of a nucleus with mass $m_B$ and charge $Z$ due to ionisation of ISM material is (Schlickeiser 2002),

$$\begin{aligned}\tau_{\rm loss}^{-1} = \tau_{\rm ion}^{-1} &\approx 2\times 10^{-16}\ s^{-1}\ Z^2\frac{n}{1\ {\rm cm}^{-3}}\Big(\frac{\rm v}{c}\Big)^{-2}\Big(\frac{p}{{\rm GeV}/c}\Big)^{-1}\\ &\approx L\Big(\frac{p}{p_0}\Big)^{-3}, \quad p \ll m_B c,\end{aligned} \tag{5.51}$$

where $L$ in the last step is a constant that depends on the density of the medium. (5.50) and (5.51) imply that in the energy-loss dominated regime, the proton spectrum is $f_p \sim Q_p\tau_{\rm loss} \propto p^{-\gamma_{\rm inj}+3}$ and similarly the carbon spectrum is $f_C \sim Q_C\tau_{\rm loss} \propto p^{-\gamma_{\rm inj}+3}$. One might further guess that the boron spectrum is given by, $f_B \sim f_C\tau_{\rm loss}/\tau_{\rm spall} \sim Q_C\tau_{\rm loss}^2/\tau_{\rm spall} \sim p^{-\gamma_{\rm inj}+7}$. However, this turns out to not be a valid solution of (5.50). In particular, the absolute value of the energy flux due to ionisation losses associated with this spectrum decreases with decreasing momentum (the LHS of eq. 5.50 is negative becasue $\partial f_B/\partial p > 0$). Equivalently, a given momentum shell is populated from higher momenta at a rate that is faster than the rate at which it loses particles to lower momenta. The correct solution of equation 5.50 is instead given by,

$$f_B = A - \frac{1}{7-\gamma_{\rm inj}}\Big(\frac{p}{p_0}\Big)^3 L^{-1}\tau_{\rm spall}^{-1}f_C \approx A \tag{5.52}$$

where $A$ is a constant. In the last step we used the requirement $f_B \geq 0$, which implies that the B spectrum is set by $A$ at most momenta in the energy-loss dominated regime, due to the strong momentum dependence of the second term. The constant $A$ can be found by imposing continuity between the loss-dominated and escape-dominated regimes. The loss-dominated B/C spectrum is then,

$$\frac{f_B}{f_C} = \frac{A}{f_C} - \frac{1}{7-\gamma_{\rm inj}}\Big(\frac{p}{p_0}\Big)^3 L^{-1}\tau_{\rm spall}^{-1} \approx \frac{A}{f_C} \propto p^{\gamma_{\rm inj}-3}, \tag{5.53}$$

which is $\propto p^{1.3}$ for $\gamma_{\rm inj} = 4.3$.

# Bibliography

Acero, F., Ackermann, M., Ajello, M., et al. 2016, ApJS, 223, 26

Ackermann, M., Ajello, M., Allafort, A., et al. 2012, ApJ, 755, 164

—. 2013, Science, 339, 807

Adkins, T., & Schekochihin, A. A. 2018, Journal of Plasma Physics, 84, 905840107

Aguilar, M., Aisa, D., Alpat, B., et al. 2015, Phys. Rev. Lett., 114, 171103

Aguilar, M., Ali Cavasonza, L., Ambrosi, G., et al. 2016, Phys. Rev. Lett., 117, 231102

Aleksić, J., Antonelli, L. A., Antoranz, P., et al. 2010, ApJ, 710, 634

Aleksić, J., Alvarez, E. A., Antonelli, L. A., et al. 2012, A&A, 541, A99

Aloisio, R., & Blasi, P. 2013, J. Cosmology Astropart. Phys., 2013, 001

Aloisio, R., Blasi, P., & Serpico, P. D. 2015, A&A, 583, A95

Amato, E. 2014, International Journal of Modern Physics D, 23, 1430013

Amato, E., & Blasi, P. 2018, Advances in Space Research, 62, 2731

Bai, X.-N. 2021, arXiv e-prints, arXiv:2112.14782

Bai, X.-N., Ostriker, E. C., Plotnikov, I., & Stone, J. M. 2019, ApJ, 876, 60

Balbus, S. A. 1995, in Astronomical Society of the Pacific Conference Series, Vol. 80, The Physics of the Interstellar Medium and Intergalactic Medium, ed. A. Ferrara, C. F. McKee, C. Heiles, & P. R. Shapiro, 328

Balbus, S. A. 2000, ApJ, 534, 420

Balbus, S. A., & Hawley, J. F. 1991, ApJ, 376, 214

Bambic, C. J., & Reynolds, C. S. 2019, arXiv e-prints, arXiv:1906.03272

Barnes, A. 1966, The Physics of Fluids, 9, 1483

Barnes, D. J., Kannan, R., Vogelsberger, M., et al. 2019, MNRAS, 488, 3003

Beck, R. 2015, The Astronomy and Astrophysics Review, 24, 4

Beck, R. 2015, A&A Rev., 24, 4

Begelman, M. C., & Zweibel, E. G. 1994, ApJ, 431, 689

Beresnyak, A., & Lazarian, A. 2009, ApJ, 702, 460

Bernet, M. L., Miniati, F., & Lilly, S. J. 2013, ApJ, 772, L28

Binney, J., Nipoti, C., & Fraternali, F. 2009, MNRAS, 397, 1804

Birnboim, Y., & Dekel, A. 2003, MNRAS, 345, 349

Bîrzan, L., Rafferty, D. A., McNamara, B. R., Wise, M. W., & Nulsen, P. E. J. 2004, ApJ, 607, 800

Blasi, P. 2019, Galaxies, 7, 64

Blasi, P., Amato, E., & Serpico, P. D. 2012, Phys. Rev. Lett., 109, 061101

Boldyrev, S. 2006, Phys. Rev. Lett., 96, 115002
Boldyrev, S., & Loureiro, N. F. 2019, Physical Review Research, 1, 012006
Boulares, A., & Cox, D. P. 1990, ApJ, 365, 544
Bowen, D. V., Chelouche, D., Jenkins, E. B., et al. 2016, ApJ, 826, 50
Braginskii, S. I. 1965, Rev. Plasma Phys., 1, 205
Breitschwerdt, D., McKenzie, J. F., & Voelk, H. J. 1991, A&A, 245, 79
Bresci, V., Amato, E., Blasi, P., & Morlino, G. 2019, MNRAS, 488, 2068
Butsky, I. S., Fielding, D. B., Hayward, C. C., et al. 2020, ApJ, 903, 77
Cai, Z., Fan, X., Bian, F., et al. 2017, ApJ, 839, 131
Cantalupo, S., Arrigoni-Battaia, F., Prochaska, J. X., Hennawi, J. F., & Madau, P. 2014, Nature, 506, 63
Cavagnolo, K. W., Donahue, M., Voit, G. M., & Sun, M. 2009, ApJS, 182, 12
Chandran, B. D., & Dennis, T. J. 2006, ApJ, 642, 140
Chandran, B. D. G. 2000, Phys. Rev. Lett., 85, 4656
Chandran, B. D. G. 2005, Phys. Rev. Lett., 95, 265004
Chandrasekhar, S., Kaufman, A. N., & Watson, K. M. 1958, Proc. R. Soc. London A, 245, 435
Chen, C. H. K. 2016, Journal of Plasma Physics, 82, 535820602
Chew, C. F., Goldberger, M. L., & Low, F. E. 1956, Proc. R. Soc. London A, 236, 112
Cho, J., & Lazarian, A. 2003, MNRAS, 345, 325
Choudhury, P. P., Sharma, P., & Quataert, E. 2019, arXiv e-prints, arXiv:1901.02903
Churazov, E., Forman, W., Jones, C., & Böhringer, H. 2000, A&A, 356, 788
Cummings, A. C., Stone, E. C., Heikkila, B. C., et al. 2016, ApJ, 831, 18
Dekel, A., Birnboim, Y., Engel, G., et al. 2009, Nature, 457, 451
Dennis, T. J., & Chandran, B. D. G. 2009, ApJ, 690, 566
Desiati, P., & Zweibel, E. G. 2014, ApJ, 791, 51
Draine, B. T. 2011, Physics of the Interstellar and Intergalactic Medium. Princeton University Press, NJ.
Drury, L. O. C. 2012, Astroparticle Physics, 39, 52
Dundovic, A., Pezzi, O., Blasi, P., Evoli, C., & Matthaeus, W. H. 2020, Phys. Rev. D, 102, 103016
Ehlert, K., Weinberger, R., Pfrommer, C., Pakmor, R., & Springel, V. 2018, MNRAS, 481, 2878
Everett, J. E., Zweibel, E. G., Benjamin, R. A., et al. 2008, ApJ, 674, 258
Evoli, C., Blasi, P., Morlino, G., & Aloisio, R. 2018, Phys. Rev. Lett., 121, 021102
Fabian, A. C., Sanders, J. S., Allen, S. W., et al. 2003, MNRAS, 344, L43
Fabian, A. C., Sanders, J. S., Taylor, G. B., et al. 2006, MNRAS, 366, 417
Farber, R., Ruszkowski, M., Yang, H. Y. K., & Zweibel, E. G. 2018, ApJ, 856, 112
Farmer, A. J., & Goldreich, P. 2004, ApJ, 604, 671
Felice, G. M., & Kulsrud, R. M. 2001, ApJ, 553, 198
Field, G. B. 1965, ApJ, 142, 531
Fornieri, O., Gaggero, D., Cerri, S. S., De La Torre Luque, P., & Gabici, S. 2021, MNRAS, 502, 5821
Gaggero, D., Maccione, L., Grasso, D., Di Bernardo, G., & Evoli, C. 2014, Phys. Rev. D, 89, 083007

Giacalone, J., & Jokipii, J. R. 1999, ApJ, 520, 204

Ginzburg, V. L. 1961, Propagation of electromagnetic waves in plasma (New York: Gordon & Breach)

Goldreich, P., & Sridhar, S. 1995, ApJ, 438, 763

Guo, F., & Oh, S. P. 2008, MNRAS, 384, 251

Guo, F., Oh, S. P., & Ruszkowski, M. 2008, ApJ, 688, 859

Hasegawa, A. 1969, The Physics of Fluids, 12, 2642

Heintz, E., & Zweibel, E. G. 2018, ApJ, 860, 97

Hennawi, J. F., Prochaska, J. X., Cantalupo, S., & Arrigoni-Battaia, F. 2015, Science, 348, 779

Hin Navin Tsung, T., Oh, S. P., & Jiang, Y.-F. 2021, arXiv e-prints, arXiv:2107.07543

Hlavacek-Larrondo, J., Fabian, A. C., Edge, A. C., et al. 2012, MNRAS, 421, 1360

Hogan, M. T., McNamara, B. R., Pulido, F. A., et al. 2017, ApJ, 851, 66

Holmes, J. A. 1974, MNRAS, 166, 155

—. 1975, MNRAS, 170, 251

Hopkins, P. F., Butsky, I. S., Panopoulou, G. V., et al. 2021a, arXiv e-prints, arXiv:2109.09762

Hopkins, P. F., Squire, J., Butsky, I. S., & Ji, S. 2021b, arXiv e-prints, arXiv:2112.02153

Hopkins, P. F., Squire, J., Chan, T. K., et al. 2020a, arXiv e-prints, arXiv:2002.06211

Hopkins, P. F., Chan, T. K., Garrison-Kimmel, S., et al. 2020b, MNRAS, 492, 3465

Huang, Y.-M., & Bhattacharjee, A. 2016, ApJ, 818, 20

Huber, B., Tchernin, C., Eckert, D., et al. 2013, A&A, 560, A64

Jacob, S., & Pfrommer, C. 2017a, MNRAS, 467, 1449

—. 2017b, MNRAS, 467, 1478

Ji, S., Oh, S. P., & McCourt, M. 2018, MNRAS, 476, 852

Kadomtsev, B. B., & Petviashvili, V. I. 1973, Doklady Akademii Nauk, 208, 794

Kempski, P., & Quataert, E. 2020, MNRAS, 493, 1801

—. 2021, arXiv e-prints, arXiv:2109.10977

Kempski, P., Quataert, E., & Squire, J. 2020, MNRAS

—. 2021, MNRAS, 500, 1231

Kempski, P., Quataert, E., Squire, J., & Kunz, M. W. 2019, MNRAS, 486, 4013

Kim, C.-G., & Ostriker, E. C. 2017, ApJ, 846, 133

Kingsland, M., Yang, H. Y. K., Reynolds, C. S., & Zuhone, J. Z. 2019, arXiv e-prints, arXiv:1909.01339

Kowal, G., & Lazarian, A. 2010, ApJ, 720, 742

Kulsrud, R., & Pearce, W. P. 1969, ApJ, 156, 445

Kulsrud, R. M. 2005, Plasma physics for astrophysics

Kulsrud, R. M., & Cesarsky, C. J. 1971, Astrophysical Letters, 8, 189

Kunz, M. W. 2011, MNRAS, 417, 602

Kunz, M. W., Bogdanović, T., Reynolds, C. S., & Stone, J. M. 2012, ApJ, 754, 122

Kunz, M. W., Schekochihin, A. A., & Stone, J. M. 2014, Phys. Rev. Lett., 112, 205003

Lacki, B. C., Thompson, T. A., Quataert, E., Loeb, A., & Waxman, E. 2011, ApJ, 734, 107

Landau, L. D., & Lifshitz, E. M. 1959, Fluid mechanics

Lazarian, A. 2016, ApJ, 833, 131

Lazarian, A., & Xu, S. 2021, ApJ, 923, 53
Lee, M. A., & Völk, H. J. 1973, Ap&SS, 24, 31
Lesur, G., & Longaretti, P. Y. 2007, MNRAS, 378, 1471
Li, Y., Bryan, G. L., Ruszkowski, M., et al. 2015, ApJ, 811, 73
Li, Y., Gendron-Marsolais, M.-L., Zhuravleva, I., et al. 2020, ApJ, 889, L1
Linden, T., Profumo, S., & Anderson, B. 2010, Phys. Rev. D, 82, 063529
Loewenstein, M., G. Zweibel, E., & C. Begelman, M. 1991, The Astrophysical Journal, 377
Loureiro, N. F., & Boldyrev, S. 2017, Phys. Rev. Lett., 118, 245101
Mahadevan, R., & Quataert, E. 1997, ApJ, 490, 605
Makwana, K. D., & Yan, H. 2020, Physical Review X, 10, 031021
Mallet, A., Schekochihin, A. A., & Chandran, B. D. G. 2017a, Journal of Plasma Physics, 83, 905830609
Mallet, A., Schekochihin, A. A., & Chandran, B. D. G. 2017b, MNRAS, 468, 4862
McCourt, M., Sharma, P., Quataert, E., & Parrish, I. J. 2012, MNRAS, 419, 3319
McDonald, M., Veilleux, S., Rupke, D. S. N., & Mushotzky, R. 2010, ApJ, 721, 1262
McKee, C. F., & Ostriker, J. P. 1977, ApJ, 218, 148
McNamara, B. R., & Nulsen, P. E. J. 2007, ARA&A, 45, 117
Meece, G. R., O'Shea, B. W., & Voit, G. M. 2015, ApJ, 808, 43
Meyrand, R., Kanekar, A., Dorland, W., & Schekochihin, A. A. 2019, Proceedings of the National Academy of Science, 116, 1185
Miville-Deschênes, M. A., Ysard, N., Lavabre, A., et al. 2008, A&A, 490, 1093
Mora, S. C., & Krause, M. 2013, A&A, 560, A42
Moster, B. P., Naab, T., & White, S. D. M. 2013, MNRAS, 428, 3121
Nulsen, P., Jones, C., Forman, W., et al. 2009, in American Institute of Physics Conference Series, Vol. 1201, American Institute of Physics Conference Series, ed. S. Heinz & E. Wilcots, 198
Nulsen, P. E. J. 1986, MNRAS, 221, 377
Paglione, T. A. D., & Abrahams, R. D. 2012, ApJ, 755, 106
Parker, E. N. 1958, Phys. Rev., 109, 1874
Parrish, I. J., McCourt, M., Quataert, E., & Sharma, P. 2012, MNRAS, 422, 704
Peterson, J. R., & Fabian, A. C. 2006, Phys. Rep., 427, 1
Pfrommer, C. 2013, ApJ, 779, 10
Pfrommer, C., Enßlin, T. A., & Springel, V. 2008, MNRAS, 385, 1211
Pizzolato, F., & Soker, N. 2005, ApJ, 632, 821
Ptuskin, V. S., Voelk, H. J., Zirakashvili, V. N., & Breitschwerdt, D. 1997, A&A, 321, 434
Quataert, E. 2008, ApJ, 673, 758
Quataert, E., Jiang, F., & Thompson, T. A. 2022a, MNRAS, 510, 920
Quataert, E., Thompson, T. A., & Jiang, Y.-F. 2022b, MNRAS, 510, 1184
Rafferty, D. A., McNamara, B. R., Nulsen, P. E. J., & Wise, M. W. 2006, ApJ, 652, 216
Recchia, S., Blasi, P., & Morlino, G. 2016, MNRAS, 462, 4227
Reines, A. E., & Volonteri, M. 2015, ApJ, 813, 82
Reynolds, C. S., Balbus, S. A., & Schekochihin, A. A. 2015, ApJ, 815, 41
Roberg-Clark, G. T., Drake, J. F., Reynolds, C. S., & Swisdak, M. 2018, Phys. Rev. Lett., 120,

035101
Rosenbluth, M. N. 1956, Los Alamos Sci. Lab. Rep., LA-2030
Ruszkowski, M., & Oh, S. P. 2010, ApJ, 713, 1332
Ruszkowski, M., Yang, H. Y. K., & Zweibel, E. 2017, ApJ, 834, 208
Salomé, P., Combes, F., Edge, A. C., et al. 2006, A&A, 454, 437
Scannapieco, E., & Brüggen, M. 2015, ApJ, 805, 158
Schekochihin, A. A. 2020, arXiv e-prints, arXiv:2010.00699
Schekochihin, A. A., Parker, J. T., Highcock, E. G., et al. 2016, Journal of Plasma Physics, 82, 905820212
Schlickeiser, R. 2002, Cosmic Ray Astrophysics
Sharma, P., Chandran, B. D. G., Quataert, E., & Parrish, I. J. 2009, ApJ, 699, 348
Sharma, P., McCourt, M., Quataert, E., & Parrish, I. J. 2012, MNRAS, 420, 3174
Singh, A., & Sharma, P. 2015, MNRAS, 446, 1895
Skilling, J. 1971, ApJ, 170, 265
—. 1975, MNRAS, 172, 557
Snyder, P. B., Hammett, G. W., & Dorland, W. 1997, Physics of Plasmas, 4, 3974
Socrates, A., Davis, S. W., & Ramirez-Ruiz, E. 2008, ApJ, 687, 202
Squire, J., & Hopkins, P. F. 2018, ApJ, 856, L15
Squire, J., Hopkins, P. F., Quataert, E., & Kempski, P. 2021, MNRAS, 502, 2630
Squire, J., Schekochihin, A. A., & Quataert, E. 2017, New Journal of Physics, 19, 055005
Sternberg, A., & Soker, N. 2009, MNRAS, 395, 228
Stocke, J. T., Keeney, B. A., Danforth, C. W., et al. 2013, ApJ, 763, 148
Stone, E. C., Cummings, A. C., McDonald, F. B., et al. 2013, Science, 341, 150
Strong, A. W., Porter, T. A., Digel, S. W., et al. 2010, ApJ, 722, L58
Thompson, T. A., Quataert, E., Zhang, D., & Weinberg, D. H. 2016, MNRAS, 455, 1830
Tomassetti, N. 2012, ApJ, 752, L13
Trotta, R., Jóhannesson, G., Moskalenko, I. V., et al. 2011, ApJ, 729, 106
Tumlinson, J., Peeples, M. S., & Werk, J. K. 2017, ARA&A, 55, 389
Verscharen, D., Chen, C. H. K., & Wicks, R. T. 2017, ApJ, 840, 106
Voelk, H. J. 1975, Reviews of Geophysics and Space Physics, 13, 547
Voit, G. M. 2018, ApJ, 868, 102
Voit, G. M., Donahue, M., Bryan, G. L., & McDonald, M. 2015, Nature, 519, 203
Voit, G. M., Meece, G., Li, Y., et al. 2017, ApJ, 845, 80
Völk, H. J. 1973, Ap&SS, 25, 471
Wentzel, D. G. 1971, ApJ, 163, 503
Werk, J. K., Prochaska, J. X., Thom, C., et al. 2013, ApJS, 204, 17
Werner, N., McNamara, B. R., Churazov, E., & Scannapieco, E. 2019, Space Sci. Rev., 215, 5
Wiener, J., Oh, S. P., & Guo, F. 2013, MNRAS, 434, 2209
Wiener, J., Pfrommer, C., & Oh, S. P. 2017, MNRAS, 467, 906
Wiener, J., Zweibel, E. G., & Oh, S. P. 2013, The Astrophysical Journal, 767, 87
Wiener, J., Zweibel, E. G., & Oh, S. P. 2018, MNRAS, 473, 3095
Wiersma, R. P. C., Schaye, J., & Smith, B. D. 2009, MNRAS, 393, 99

Xu, S., & Lazarian, A. 2018, ApJ, 868, 36
Xu, S., Yan, H., & Lazarian, A. 2016, ApJ, 826, 166
Yan, H., & Lazarian, A. 2004, ApJ, 614, 757
—. 2008, ApJ, 673, 942
Yang, H. Y. K., & Reynolds, C. S. 2016, ApJ, 818, 181
Yang, L., Li, H., Guo, F., et al. 2020, ApJ, 901, L22
Yang, R., Aharonian, F., & Evoli, C. 2016, Phys. Rev. D, 93, 123007
Zakharov, V. E., & Sagdeev, R. Z. 1970, Soviet Physics Doklady, 15, 439
Zhang, C., Churazov, E., & Schekochihin, A. A. 2018, MNRAS, 478, 4785
Zhang, D., Thompson, T. A., Quataert, E., & Murray, N. 2017, MNRAS, 468, 4801
Zhuravleva, I., Churazov, E., Arévalo, P., et al. 2016, MNRAS, 458, 2902
Zweibel, E. G. 2013, Physics of Plasmas, 20, 055501
—. 2017, Physics of Plasmas, 24, 055402
Zweibel, E. G., Mirnov, V. V., Ruszkowski, M., et al. 2018, ApJ, 858, 5