

Year-2 Progress Report on Numerical Methods for BGK-Type Kinetic Equations

Evan Habbershaw[†] and Steven M. Wise^{1†}

[†]Department of Mathematics, The University of Tennessee, Knoxville, TN,
37996

January 29, 2024

¹Corresponding author. Email address: swise1@utk.edu

Abstract

In this second progress report we expand upon our previous report and preliminary work. Specifically, we review some work on the numerical solution of single- and multi-species BGK-type kinetic equations of particle transport. Such equations model the motion of fluid particles via a density field when the kinetic theory of rarefied gases must be used in place of the continuum limit Navier-Stokes and Euler equations. The BGK-type equations describe the fluid in terms of phase space variables, and, in three space dimensions, require 6 independent phase-space variables (3 for space and 3 for velocity) for each species for accurate simulation. This requires sophisticated numerical algorithms and efficient code to realize predictions over desired space and time scales. In particular, stable numerical methods must be designed to handle potential discontinuities (shocks) and rarefaction waves in the solutions coming from conservative advection terms and, in addition, numerical stiffness owing to diffusive particle collision terms. Furthermore, the particle interaction terms are non-local in nature, adding yet another layer of complexity, and the interaction length scales of the non-local terms may be orders of magnitude different, when multiple particle species are involved. In this report, we outline strategies for generating efficient and stable numerical algorithms and code, including the use of (i) stable high-order finite volume methods, (ii) fully implicit and implicit-explicit (IMEX) time integration techniques, (iii) adaptive time-phase-space multi-level methods, (iv) discrete velocity methods, and (v) moment equation methods. The preliminary codes, which will be demonstrated herein, are built in the commercial software package MATLAB for quick and easy prototyping, but will later be translated into production software using modern open languages.

Key words: Boltzmann equation, BGK approximation, multi-species BGK models, finite volume schemes, MUSCL methods, numerically stiff equations, implicit-explicit time stepping strategies, discrete velocity methods, Runge-Kutta methods, moment equations, asymptotic relaxation.

Contents

1	Introduction	2
2	Single Species BGK Kinetic Models	7
2.1	Basic Properties of Solutions	7
2.1.1	Collision Invariants	7
2.1.2	Space-Homogeneous Problem	15
2.1.3	The H-Theorem	17
2.2	Discrete-Velocity BGK Models	23
2.2.1	Discrete-Velocity BGK Model in One Space Dimension	23
2.2.2	Space-Homogeneous Discrete-Velocity Model	31
3	Numerical Approximations for the Single Species Case	34
3.1	Finite Volume Space and Velocity Discretization	34
3.2	Implicit-Explicit Runge Kutta Time Stepping	38
3.3	Fully Discrete Scheme	40
3.4	Poisson Solver for Vlasov-Poisson-BGK Equation	41
3.5	Sample Computations and Accuracy Tests	43
4	Discrete-Velocity Numerical Methods for the Single Species BGK Problem	58
4.1	A Fully Explicit Scheme	58
4.2	An Explicit Scheme for the Space-Homogeneous Case	61
4.3	An Implicit Scheme for the Space Homogeneous Case	63
5	Multispecies BGK Equations	66
5.1	Formulation and Properties	66
5.2	Numerical Approximation	70

6	Asymptotic Relaxation of Multi-Species Moment Equations	72
6.1	Model Equations	73
6.1.1	Model Properties	74
6.1.2	Moment Equations	75
6.2	Properties of the ODE system	76
6.2.1	Positivity of the Temperature	77
6.2.2	Velocity Bounds	79
6.2.3	Existence and Uniqueness	80
6.3	Long-time Behavior	82
6.3.1	Null Spaces of $D - A$, Z , $F - B$, and $\widehat{Z}$	83
6.3.2	Velocity Relaxation Proof	84
6.3.3	Energy Relaxation Proof	86
6.3.4	Temperature Relaxation	88
6.4	Numerical Demonstration	88
6.5	Conclusions	91
7	Progress Summary and Next Steps	94
7.1	Summary	94
7.2	Next Steps	95
8	Acknowledgements	97
A	An Important Integral	98
B	Picard-Lindelöf Theorem	100
C	Bounding the Eigenvalues of Z and $\widehat{Z}$	101
D	Null Space Components of $\widetilde{W}$ and $\widetilde{\xi}$	103
E	Code	104
E.1	Main Driver: vlasovPoissonBGKMain.m	104
E.2	vlasovPoissonBGKSolver.m	111
E.3	BGKCollision.m	115
E.4	divFlux.m	118
E.5	slopeReconstruction.m	120

E.6	minMod.m	122
E.7	sodSoln.m	123
E.8	pStarSolve.m	126
F	Bibliography	127

Chapter 1

Introduction

The Navier-Stokes and Euler equations, with which most computational fluid dynamicists are familiar, are used to describe the evolution of a fluid under the assumption that the constituent particles move in, essentially, lock-step motion. In other words, fluid particles with small space separation are assumed to have nearly identical velocity vectors. To be precise, suppose that the mean free path (diffusion length scale) of particles is denoted λ , and the characteristic spatial size of the problem is denoted L . The *Knudsen number* is defined as $K_n = \lambda/L$. When $K_n \ll 1$, the diffusion length scale is too small to resolve accurately, and, in fact, the individual motions and interactions of constituent particles can be coarse-grained (averaged out) without significant loss of fidelity [29]. Indeed, the Navier-Stokes equation, which is applicable in this physical regime, is a highly successful and accurate model.

However, when $K_n = \mathcal{O}(1)$, the Navier-Stokes equation is no longer valid, and particle interactions must be taken into account. Particles with small space separation could move in entirely contrary directions, and, in this regime, the Boltzmann transport equation is an important model of particle evolution [29]. It describes the distribution of particles as a function of time, space ($d = 3$ dimensions), and velocity ($d = 3$ dimensions). The three dimensions of space and three dimensions of velocity comprise what is known as *phase space*. The Boltzmann equation is complicated not only by the high dimensionality of phase space but also the highly nonlinear, highly nonlocal nature of the collision (particle interaction) operator.

The *Vlasov-Boltzmann equation* for a single species dilute gas is given as follows:

$$\frac{\partial f}{\partial t} + \mathbf{v} \cdot \nabla_{\mathbf{x}} f + \mathbf{a} \cdot \nabla_{\mathbf{v}} f = Q[f](\mathbf{x}, \mathbf{v}, t), \quad (1.1)$$

where f is the density of particles at position $\mathbf{x}$ with velocity $\mathbf{v}$ at time t ; and $\mathbf{a}$ is a particle acceleration determined by an external field, for example, an electric or magnetic field. The *Boltzmann transport equation* results from setting $\mathbf{a} \equiv \mathbf{0}$. The derivation of the equation is a straightforward exercise using the chain rule. In particular, the total time derivative of the distribution f can be realized as

$$\begin{aligned} \frac{d}{dt}f(\mathbf{x}, \mathbf{v}, t) &= \sum_{i=1}^d \frac{\partial f}{\partial x_i} \frac{dx_i}{dt} + \sum_{i=1}^d \frac{\partial f}{\partial v_i} \frac{dv_i}{dt} + \frac{\partial f}{\partial t} \\ &= \mathbf{v} \cdot \nabla_{\mathbf{x}} f + \mathbf{a} \cdot \nabla_{\mathbf{v}} f + \frac{\partial f}{\partial t}. \end{aligned}$$

Thus, the Vlasov-Boltzmann equation may be constructed via the following law: the total time rate of change of the distribution is equal to the collision (frequency) operator. In other words,

$$\frac{d}{dt}f(\mathbf{x}, \mathbf{v}, t) = Q[f](\mathbf{x}, \mathbf{v}, t).$$

The *Boltzmann collision operator*, $Q[f]$, requires much more physical insight for a clear derivation [29], involving certain conservation principles, in particular. Therefore, we will content ourselves by only stating its generic form (for $d = 3$ dimensions):

$$Q[f](\mathbf{x}, \mathbf{v}, t) = \int_{\mathbb{R}^3 \times \mathbb{S}^2} [f(\mathbf{x}, \mathbf{v}', t) f(\mathbf{x}, \mathbf{v}_*, t) - f(\mathbf{x}, \mathbf{v}, t) f(\mathbf{x}, \mathbf{v}_*, t)] B(|\mathbf{v} - \mathbf{v}_*|, \boldsymbol{\sigma}) d\boldsymbol{\sigma} d\mathbf{v}_*,$$

where B is the collision kernel describing interactions between particles, and $\boldsymbol{\sigma}$ is the unit vector in the scattering direction $\mathbf{v} - \mathbf{v}_*$. The velocities of the two interacting particles before collision, $\mathbf{v}'$ and $\mathbf{v}_*$, can be expressed in terms of the velocities of the particles after collision, $\mathbf{v}$ and $\mathbf{v}_*$, via the expressions

$$\mathbf{v}' = \frac{\mathbf{v} + \mathbf{v}_*}{2} + \frac{|\mathbf{v} - \mathbf{v}_*|}{2} \boldsymbol{\sigma}, \quad \mathbf{v}_* = \frac{\mathbf{v} + \mathbf{v}_*}{2} - \frac{|\mathbf{v} - \mathbf{v}_*|}{2} \boldsymbol{\sigma}.$$

The computation of the Boltzmann collision operator is very expensive. For the 3D case, the (5 dimensional) integral must be computed at every value of $(\mathbf{x}, \mathbf{v})$ in phase space. Additionally, the integral cannot be evaluated analytically, except for the simplest of cases (e.g., Maxwell molecules, with carefully prepared initial conditions). Thus, computation of the collision operator is usually the most expensive part of computing numerical solutions of the Boltzmann equation. For this reason, Monte Carlo methods are generally preferred for numerical simulation [24]. Unfortunately, these are slow to converge, as is well known. What

is worse, for the N_s -species case, there are N_s^2 collision operators that must be considered. Thus, a simpler, less costly approximation is typically desired and implemented.

The *Vlasov-BGK equation* — BGK stands for the names Bhatnagar, Gross, and Krook, who introduced it 1954 [1] — is a model derived by approximating the Boltzmann collision operator with a simpler nonlinear, nonlocal relaxation operator of the form

$$Q^{BGK}[f](\mathbf{x}, \mathbf{v}, t) = \lambda(M_f(\mathbf{x}, \mathbf{v}, t) - f(\mathbf{x}, \mathbf{v}, t)), \quad (1.2)$$

where $\lambda = \frac{1}{\tau}$ is the collision frequency between particles, $\tau > 0$ is a characteristic time, and M_f is the *Maxwellian*, which is defined using the moments

$$n(\mathbf{x}, t) := \int_{\mathbb{R}^d} f \, d\mathbf{v}, \quad \mathbf{u}(\mathbf{x}, t) := \frac{1}{n} \int_{\mathbb{R}^d} f \mathbf{v} \, d\mathbf{v}, \quad \theta(\mathbf{x}, t) := \frac{1}{nd} \int_{\mathbb{R}^d} f |\mathbf{v} - \mathbf{u}|^2 \, d\mathbf{v}, \quad (1.3)$$

as

$$M_f(\mathbf{x}, \mathbf{v}, t) := M[f](\mathbf{x}, \mathbf{v}, t) := n \left(\frac{1}{2\pi\theta} \right)^{\frac{d}{2}} \exp \left(-\frac{|\mathbf{v} - \mathbf{u}|^2}{2\theta} \right), \quad (1.4)$$

Here we have normalized the model so that the mass of the particles is unity. The space-time fields n , $\mathbf{u}$, and θ , are the macroscopic (coarse grained) number density, bulk velocity, and temperature, respectively, and the energy density is given by $E(\mathbf{x}, t) := \frac{1}{2} \int_{\mathbb{R}^d} f |\mathbf{v}|^2 \, d\mathbf{v}$. Using this approximation drastically reduces the computational cost of the simulation of dilute gases, and recovers both equilibrium and streaming behavior of the Boltzmann equation in collision-dominated, and collision-free limits. The BGK collision operator also satisfies conservation and entropy properties of the Boltzmann operator, as we show below. Of course, the basic design and principle of the BGK approximation is that the density, f , should relax over time toward the Maxwellian, and, clearly, at this state, the collision operator gives a zero contribution.

However, the BGK collision operator does not capture some important properties. First, it fails to capture the correct Prandtl number (essentially the ratio of viscosity to thermal conductivity), largely because the collision rate is velocity independent. As a result, the model may not agree with the compressible Navier-Stokes Equations that are derived from the Boltzmann equation in high collision regimes. A number of generalizations of the BGK model have been proposed to deal with this shortcoming. These include the ES-BGK [12] and Shakov [28] models, which incorporate extra degrees of freedom. The model of Mieussens and Struchtrup [18] incorporates a velocity dependent collision rate, which improves the capture

of the correct Prandtl number. These models give a more physically realistic model, but cause a substantial increase to computational costs.

To summarize, the single-species Vlasov-BGK equation has the form

$$\frac{\partial f}{\partial t} + \mathbf{v} \cdot \nabla_{\mathbf{x}} f + \mathbf{a} \cdot \nabla_{\mathbf{v}} f = \lambda(M_f(\mathbf{x}, \mathbf{v}, t) - f(\mathbf{x}, \mathbf{v}, t)), \quad (1.5)$$

and, in analogy to the previous setting, the *single-species BGK equation* results from setting $\mathbf{a} \equiv \mathbf{0}$. When, in equation (1.5), the acceleration is determined by an electric field, according to the model

$$\mathbf{a}(\mathbf{x}, t) = -\chi \nabla_{\mathbf{x}} \Phi(\mathbf{x}, t),$$

where $\chi > 0$ is a constant, and

$$-\Delta \Phi(\mathbf{x}, t) = n(\mathbf{x}, t),$$

the resulting model is known as the *Vlasov-Poisson-BGK equation*.

Numerical methods for the BGK family of equations have been proposed and analyzed extensively. The discretization of phase space is costly, and there are challenges that occur in high-collision regimes. First, the BGK operator becomes stiff as the collision frequency becomes large, so an implicit approach is desired for this term, to avoid unacceptably small time steps. In a seminal paper [5], it was shown that a Backward Euler step could be applied in an explicit manner, allowing stable time stepping for a wide range of collisional regimes. Taking advantage of this, the paper [24] introduced an implicit-explicit (IMEX) Runge-Kutta scheme based on [20], treating the convection term explicitly and the collision term implicitly. Second, stable and accurate high-order conservation schemes are required for the convection terms. For example, Pieraccini and Puppo [24] use a high-order weighted essentially non-oscillatory (WENO) finite volume scheme, though high-order flux-limited schemes perform well, as we show, and are easily scaled to higher dimensions.

Most of the numerical schemes for the BGK equation, including the one that we focus on here from [24], lose the exact conservation of mass, momentum and energy at the discrete level, and, additionally, the entropy dissipation at the discrete level. This shortcoming is overcome in the work by Mieussens [16, 17]. Exact conservation is obtained by computing a discrete equilibrium function, which requires the solution of a nonlinear system of 5 equations for the BGK model and a nonlinear system of 10 equations for the ES-BGK model at each grid point in space. Another challenge is to recover consistent numerical solutions of the

Euler and Navier-Stokes Equations for compressible flows, which can be derived from the BGK model using a Chapman-Enskog expansion [29].

In some situations, an explicit treatment of the advection terms in the BGK equation — even while giving the collision operator an implicit treatment — can lead to a method that requires excessively small time steps for stability and accuracy [7]. This can happen when there are long time scales that lead to incompressible equations in the high collision limit, or in problems for which the maximum velocity in the computational domain is significantly larger than the fluid speed of sound. A fully implicit approach can be taken to address these issues. This is common for time dependent kinetic equations in radiation transport contexts, and has been considered for electron transport problems [7]. Fully implicit methods for dilute gases and collisionless plasmas have been proposed in [7]. These approaches use sophisticated iterative methods to manage the cost and memory requirements of the implicit update.

In this progress report, some background for BGK-type models is presented. We do not conduct an extensive review but give the reader (especially those unfamiliar with kinetic equations) a gentle, albeit brief, introduction. In Section 2.1 we introduce the single species BGK Equation. We describe some basic theory (Sections 2.1.1 – 2.1.3), deriving conservation and (mathematical) entropy dissipation properties. Finite volume implicit-explicit (IMEX) Runge-Kutta (RK) numerical methods are presented in Section 3. We give some preliminary simulation results in Section 3, for problems including the Sod shock tube benchmark and two-stream instability. In Section 5, we give a brief introduction to multispecies BGK models in order to explain some numerical challenges for such equations. Section 6 describes asymptotic dynamics of the space homogeneous multispecies problem. We conclude the report with a brief summary of preliminary work and near future work in Section 7. A prototype 1x1v MATLAB code for the Sod shock tube problem is contained in Appendix E.

Chapter 2

Single Species BGK Kinetic Models

2.1 Basic Properties of Solutions

In this section, we describe some basic theory for the single species BGK equation [1, 29], namely,

$$\frac{\partial f}{\partial t} + \mathbf{v} \cdot \nabla_{\mathbf{x}} f = \lambda (M_f - f). \quad (2.1)$$

The existence of nonnegative solutions to the (single species) BGK equation was proved for $\mathbf{x} \in \mathbb{R}^d$ by Perthame (1989) [22], and on bounded domains by Ringeissen (1991) [26]. Uniqueness of mild solutions for the (single species) periodic (in $\mathbf{x}$) case was shown by Perthame and Pulvirenti (1993) [21], and extended to the full space, $\mathbf{x} \in \mathbb{R}^d$, by Mischler (1996) [19].

Beyond questions about existence and uniqueness of solutions, it is important to have a firm grasp of the properties of solutions to BGK-type equations, since it is vital to build numerical approximation schemes that respect analogous features at the fully discrete level. Herein, we review conservation and entropy dissipation solution properties, and, later, we discuss how these are used in the design of numerical schemes.

2.1.1 Collision Invariants

Let us begin with an important property for the well-definedness of the model. In particular, the temperature is non-negative, since f is nonnegative. To see this, consider

Proposition 2.1.1. *If $f : \Omega \times \mathbb{R}^d \times [0, \infty) \rightarrow \mathbb{R}$ is non-negative, then n is nonnegative.*

Consequently, since

$$\frac{d}{2}n(\mathbf{x}, t)\theta(\mathbf{x}, t) = \frac{1}{2} \int_{\mathbb{R}^d} |\mathbf{v} - \mathbf{u}(\mathbf{x}, t)|^2 f(\mathbf{x}, \mathbf{v}, t) d\mathbf{v} \geq 0,$$

the temperature is also nonnegative. Furthermore, the temperature is related to the other moments via the relation

$$E(\mathbf{x}, t) - \frac{1}{2}n(\mathbf{x}, t)|\mathbf{u}(\mathbf{x}, t)|^2 = \frac{d}{2}n(\mathbf{x}, t)\theta(\mathbf{x}, t).$$

Proof. The first assertion is clear from the definition of n . The second assertion is clear from the fact that n and f are nonnegative, and by the definition of the temperature. Next, expanding the right-hand side in the definition of the temperature, we have

$$\begin{aligned} \frac{d}{2}n(\mathbf{x}, t)\theta(\mathbf{x}, t) &= \frac{1}{2} \int_{\mathbb{R}^d} |\mathbf{v} - \mathbf{u}(\mathbf{x}, t)|^2 f(\mathbf{x}, \mathbf{v}, t) d\mathbf{v} \\ &= \frac{1}{2} \int_{\mathbb{R}^d} |\mathbf{v}|^2 f(\mathbf{x}, \mathbf{v}, t) d\mathbf{v} - \mathbf{u}(\mathbf{x}, t) \cdot \int_{\mathbb{R}^d} \mathbf{v} f(\mathbf{x}, \mathbf{v}, t) d\mathbf{v} \\ &\quad + \frac{|\mathbf{u}(\mathbf{x}, t)|^2}{2} \int_{\mathbb{R}^d} f(\mathbf{x}, \mathbf{v}, t) d\mathbf{v} \\ &= E(\mathbf{x}, t) - \mathbf{u}(\mathbf{x}, t) \cdot \mathbf{u}(\mathbf{x}, t)n(\mathbf{x}, t) + \frac{1}{2}n(\mathbf{x}, t)|\mathbf{u}(\mathbf{x}, t)|^2 \\ &= E(\mathbf{x}, t) - \frac{1}{2}n(\mathbf{x}, t)|\mathbf{u}(\mathbf{x}, t)|^2. \end{aligned}$$

The proof is complete. ■

For convenience, we define $\langle \cdot \rangle : L^1(\mathbb{R}^d; \mathbb{R}^p) \rightarrow \mathbb{R}^p$ by

$$\langle \mathbf{g} \rangle = \int_{\mathbb{R}^d} \mathbf{g}(\mathbf{v}) d\mathbf{v}.$$

Therefore,

$$\begin{bmatrix} n(\mathbf{x}, t) \\ n(\mathbf{x}, t)\mathbf{u}(\mathbf{x}, t) \\ E(\mathbf{x}, t) \end{bmatrix} = \left\langle f(\mathbf{x}, \mathbf{v}, t) \begin{bmatrix} 1 \\ \mathbf{v} \\ \frac{1}{2}|\mathbf{v}|^2 \end{bmatrix} \right\rangle.$$

Thus, the macroscopic density, n , momentum, $\mathbf{p} := n\mathbf{u}$, and energy density, E , may be viewed as the first three moments of the particle distribution function f . The first three moments of the collision operator are equal to zero. This gives us some conservation properties, as we show below.

For convenience, let us define

$$\boldsymbol{\rho}(\mathbf{x}, t) := [n, n\mathbf{u}, E]^\top : \mathbb{R}^d \times \mathbb{R}^+ \rightarrow \mathbb{R}^{d+2}; \quad (2.2)$$

$$\mathbf{m}(\mathbf{v}) := \left[1, \mathbf{v}, \frac{1}{2}|\mathbf{v}|^2 \right]^\top : \mathbb{R}^d \rightarrow \mathbb{R}^{d+2}; \quad (2.3)$$

and

$$\boldsymbol{\alpha}(\mathbf{x}, t) := \left[\ln \left(\frac{n}{(2\pi\theta)^{d/2}} \right) - \frac{|\mathbf{u}|^2}{2\theta}, \frac{\mathbf{u}}{\theta}, -\frac{1}{\theta} \right]^\top : \mathbb{R}^d \times \mathbb{R}^+ \rightarrow \mathbb{R}^{d+2}. \quad (2.4)$$

Then, succinctly,

$$\boldsymbol{\rho} = \langle f \mathbf{m} \rangle, \quad \text{and} \quad M_f = \exp(\boldsymbol{\alpha} \cdot \mathbf{m}(\mathbf{v})). \quad (2.5)$$

Indeed,

$$\begin{aligned} \exp(\boldsymbol{\alpha} \cdot \mathbf{m}) &= \exp \left(\ln \left(\frac{n}{(2\pi\theta)^{d/2}} \right) - \frac{|\mathbf{u}|^2}{2\theta} \right) \exp \left(\frac{\mathbf{u} \cdot \mathbf{v}}{\theta} \right) \exp \left(-\frac{1}{2\theta} |\mathbf{v}|^2 \right) \\ &= n \left(\frac{1}{2\pi\theta} \right)^{d/2} \exp \left(-\frac{|\mathbf{u}|^2}{2\theta} \right) \exp \left(\frac{\mathbf{u} \cdot \mathbf{v}}{\theta} \right) \exp \left(-\frac{1}{2\theta} |\mathbf{v}|^2 \right) \\ &= n \left(\frac{1}{2\pi\theta} \right)^{d/2} \exp \left(-\frac{1}{2\theta} |\mathbf{u} - \mathbf{v}|^2 \right) \\ &= M_f. \end{aligned}$$

Lemma 2.1.2. *The following equality holds:*

$$\langle (M_f - f) \mathbf{m} \rangle = \left\langle \begin{bmatrix} 1 \\ \mathbf{v} \\ \frac{|\mathbf{v}|^2}{2} \end{bmatrix} (M_f - f) \right\rangle = \mathbf{0}. \quad (2.6)$$

Proof. 1. First we show that

$$\int_{\mathbb{R}^d} M_f \, d\mathbf{v} = \int_{\mathbb{R}^d} f \, d\mathbf{v}. \quad (2.7)$$

By definition,

$$\begin{aligned} \int_{\mathbb{R}^d} M_f \, d\mathbf{v} &= \int_{\mathbb{R}^d} \frac{n}{(2\pi\theta)^{\frac{d}{2}}} \exp \left(-\frac{|\mathbf{v} - \mathbf{u}|^2}{2\theta} \right) \, d\mathbf{v} \\ &= \frac{n}{(2\pi\theta)^{\frac{d}{2}}} \int_{\mathbb{R}^d} \exp \left(-\left| \frac{\mathbf{v} - \mathbf{u}}{(2\theta)^{\frac{1}{2}}} \right|^2 \right) \, d\mathbf{v}. \end{aligned} \quad (2.8)$$

Consider the substitution

$$\mathbf{s} = \frac{\mathbf{v} - \mathbf{u}}{(2\theta)^{\frac{1}{2}}} \implies d\mathbf{s} = \left(\frac{1}{(2\theta)^{\frac{1}{2}}} \right)^d d\mathbf{v} \implies d\mathbf{v} = (2\theta)^{\frac{d}{2}} d\mathbf{s}. \quad (2.9)$$

The integral becomes

$$\begin{aligned} \int_{\mathbb{R}^d} M_f d\mathbf{v} &= \frac{n}{(2\pi\theta)^{\frac{d}{2}}} \int_{\mathbb{R}^d} \exp(-|\mathbf{s}|^2) (2\theta)^{\frac{d}{2}} d\mathbf{s} \\ &= \frac{n(2\theta)^{\frac{d}{2}}}{\pi^{\frac{d}{2}} (2\theta)^{\frac{d}{2}}} \int_{\mathbb{R}^d} e^{-|\mathbf{s}|^2} d\mathbf{s} \\ &= \frac{n}{\pi^{\frac{d}{2}}} \cdot \pi^{\frac{d}{2}} \\ &= \int_{\mathbb{R}^d} f d\mathbf{v}, \end{aligned} \quad (2.10)$$

as desired. Note that we have used the definition $n = \int_{\mathbb{R}^d} f d\mathbf{v}$, and the fact that in d dimensions, the integral $\int_{\mathbb{R}^d} e^{-|\mathbf{x}|^2} d\mathbf{x} = \pi^{\frac{d}{2}}$. We will use this frequently.

2. Next we show that

$$\int_{\mathbb{R}^d} \mathbf{v} M_f d\mathbf{v} = \int_{\mathbb{R}^d} \mathbf{v} f d\mathbf{v}. \quad (2.11)$$

Following a similar line of work, we have

$$\begin{aligned} \int_{\mathbb{R}^d} \mathbf{v} M_f d\mathbf{v} &= \int_{\mathbb{R}^d} \frac{n}{(2\pi\theta)^{\frac{d}{2}}} \mathbf{v} \exp\left(-\frac{|\mathbf{v} - \mathbf{u}|^2}{2\theta}\right) d\mathbf{v} \\ &= \frac{n}{(2\pi\theta)^{\frac{d}{2}}} \int_{\mathbb{R}^d} \mathbf{v} \exp\left(-\left|\frac{\mathbf{v} - \mathbf{u}}{(2\theta)^{\frac{1}{2}}}\right|^2\right) d\mathbf{v}. \end{aligned} \quad (2.12)$$

Using the substitution in (2.9), the integral becomes

$$\begin{aligned} \int_{\mathbb{R}^d} \mathbf{v} M_f d\mathbf{v} &= \frac{n}{(2\pi\theta)^{\frac{d}{2}}} \int_{\mathbb{R}^d} \left((2\theta)^{\frac{1}{2}} \mathbf{s} + \mathbf{u} \right) \exp(-|\mathbf{s}|^2) \cdot (2\theta)^{\frac{d}{2}} d\mathbf{s} \\ &= \frac{n(2\theta)^{\frac{d}{2}}}{\pi^{\frac{d}{2}} (2\theta)^{\frac{d}{2}}} \left[(2\theta)^{\frac{1}{2}} \int_{\mathbb{R}^d} \mathbf{s} e^{-|\mathbf{s}|^2} d\mathbf{s} + \mathbf{u} \int_{\mathbb{R}^d} e^{-|\mathbf{s}|^2} d\mathbf{s} \right] \\ &= \frac{n}{\pi^{\frac{d}{2}}} \left[(2\theta)^{\frac{1}{2}} \int_{\mathbb{R}^d} \mathbf{s} e^{-|\mathbf{s}|^2} d\mathbf{s} + \mathbf{u} \int_{\mathbb{R}^d} e^{-|\mathbf{s}|^2} d\mathbf{s} \right]. \end{aligned} \quad (2.13)$$

Let us deal with the first term in (2.13). One can easily show that

$$\int_{-\infty}^{\infty} x e^{-x^2} dx = 0 \implies \int_{\mathbb{R}^d} \mathbf{s} e^{-|\mathbf{s}|^2} d\mathbf{s} = \mathbf{0}.$$

Therefore, we have

$$\int_{\mathbb{R}^d} \mathbf{v} M_f d\mathbf{v} = \frac{n}{\pi^{\frac{d}{2}}} \cdot \mathbf{u} \int_{\mathbb{R}^d} e^{-|\mathbf{s}|^2} d\mathbf{s} = \frac{n\mathbf{u}}{\pi^{\frac{d}{2}}} \cdot \pi^{\frac{d}{2}} = n\mathbf{u} = \int_{\mathbb{R}^d} \mathbf{v} f d\mathbf{v}.$$

3. Finally, we show that

$$\int \frac{|\mathbf{v}|^2}{2} M_f d\mathbf{v} = \int \frac{|\mathbf{v}|^2}{2} f d\mathbf{v}. \quad (2.14)$$

First, observe that

$$\int_{\mathbb{R}^d} |\mathbf{s}|^2 e^{-|\mathbf{s}|^2} d\mathbf{s} = \frac{d}{2} \pi^{\frac{d}{2}}. \quad (2.15)$$

(To see how to prove such results, see Appendix A). Now, using the substitution in (2.9), and the same integration techniques on the LHS of (2.14),

$$\begin{aligned} \int \frac{|\mathbf{v}|^2}{2} M_f d\mathbf{v} &= \frac{1}{2} \frac{n}{(2\pi\theta)^{\frac{d}{2}}} \int |\mathbf{v}|^2 e^{-\frac{|\mathbf{v}-\mathbf{u}|^2}{2\theta}} d\mathbf{v} \\ &= \frac{1}{2} \frac{n}{(2\pi\theta)^{\frac{d}{2}}} \int |\mathbf{u} + (\mathbf{v} - \mathbf{u})|^2 e^{-\frac{|\mathbf{v}-\mathbf{u}|^2}{2\theta}} d\mathbf{v} \\ &= \frac{n}{2(2\pi\theta)^{\frac{d}{2}}} \int (-|\mathbf{u}|^2 + 2\mathbf{u}^\top \mathbf{v} + |\mathbf{v} - \mathbf{u}|^2) e^{-\frac{|\mathbf{v}-\mathbf{u}|^2}{2\theta}} d\mathbf{v} \\ &= \frac{-n|\mathbf{u}|^2}{2(2\pi\theta)^{\frac{d}{2}}} (2\theta)^{\frac{d}{2}} \int_{\mathbb{R}^d} e^{-|\mathbf{s}|^2} d\mathbf{s} + \frac{2n(2\theta)^{\frac{d}{2}}}{2(2\pi\theta)^{\frac{d}{2}}} \mathbf{u}^\top \int_{\mathbb{R}^d} [\mathbf{u} + (2\theta)^{\frac{1}{2}} \mathbf{s}] e^{-|\mathbf{s}|^2} d\mathbf{s} \\ &\quad + \frac{n(2\theta)^{\frac{d}{2}+1}}{2(2\pi\theta)^{\frac{d}{2}}} \int_{\mathbb{R}^d} |\mathbf{s}|^2 e^{-|\mathbf{s}|^2} d\mathbf{s} \\ &= \frac{-n|\mathbf{u}|^2}{2\pi^{\frac{d}{2}}} \pi^{\frac{d}{2}} + \frac{n|\mathbf{u}|^2}{\pi^{\frac{d}{2}}} \pi^{\frac{d}{2}} + 0 + \frac{2n\theta}{2\pi^{\frac{d}{2}}} \frac{d}{2} \pi^{\frac{d}{2}} \\ &= \frac{1}{2} n|\mathbf{u}|^2 + \frac{d}{2} n\theta \\ &= \int_{\mathbb{R}^d} \frac{|\mathbf{v}|^2}{2} f d\mathbf{v}, \end{aligned} \quad (2.16)$$

where the final line follows from Proposition 2.1.1 and the definition of the energy density, E .

The proof is complete. ■

Definition 2.1.3. We say that a function $g(\mathbf{v})$ is **collision invariant** iff

$$\langle g(\mathbf{v}) (M_f - f) \rangle = \int_{\mathbb{R}^d} g(\mathbf{v}) (M_f - f) d\mathbf{v} = 0.$$

From the last Lemma, we observe that 1 , $\mathbf{v}$ and $|\mathbf{v}|^2$ are collision invariant. Of course, any linear combination of these functions will also be collision invariant. We will use this fact later.

Let us introduce some further notation.

Definition 2.1.4. Suppose that $\mathbf{p} \in \mathbb{R}^s$ and $\mathbf{q} \in \mathbb{R}^t$. Then the vector outer product $\mathbf{p} \otimes \mathbf{q}$ is a matrix in $\mathbb{R}^{s \times t}$ whose components are precisely

$$[\mathbf{p} \otimes \mathbf{q}]_{i,j} = p_i q_j.$$

Multiplying the BGK Equation by the vector $\mathbf{m} = [1, \mathbf{v}, \frac{1}{2}|\mathbf{v}|^2]^\top$ and integrating, we get expressions for the conservation of mass, momentum, and energy, respectively.

Lemma 2.1.5. Suppose that f is a non-negative solution to the BGK equation, that is, the Vlasov-BGK equation with $\mathbf{a} \equiv \mathbf{0}$. The following conservation equations hold:

$$\frac{\partial}{\partial t} n + \nabla_{\mathbf{x}} \cdot (n\mathbf{u}) = 0, \quad (2.17)$$

$$\frac{\partial}{\partial t} (n\mathbf{u}) + \nabla_{\mathbf{x}} \cdot (n\mathbf{u} \otimes \mathbf{u} + \mathbf{P}) = 0, \quad (2.18)$$

$$\frac{\partial}{\partial t} E + \nabla_{\mathbf{x}} \cdot (E\mathbf{u} + \mathbf{P}\mathbf{u} + \mathbf{q}) = 0, \quad (2.19)$$

where

$$\mathbf{P} := \int_{\mathbb{R}^d} (\mathbf{v} - \mathbf{u}) \otimes (\mathbf{v} - \mathbf{u}) f d\mathbf{v}$$

is the pressure tensor and

$$\mathbf{q} := \frac{1}{2} \int_{\mathbb{R}^d} |\mathbf{v} - \mathbf{u}|^2 (\mathbf{v} - \mathbf{u}) f d\mathbf{v}$$

is the heat flux.

Proof. 1. Multiplying the BGK equation by 1 and integrating, we have

$$\int_{\mathbb{R}^d} \left[\frac{\partial f}{\partial t} + \nabla_{\mathbf{x}} \cdot (\mathbf{v}f) \right] d\mathbf{v} = \int_{\mathbb{R}^d} \lambda(M_f - f) d\mathbf{v} \quad (2.20)$$

$$\iff \frac{\partial}{\partial t} \left(\int_{\mathbb{R}^d} f d\mathbf{v} \right) + \nabla_{\mathbf{x}} \cdot \left(\int_{\mathbb{R}^d} \mathbf{v}f d\mathbf{v} \right) = 0 \quad (2.21)$$

$$\iff \frac{\partial}{\partial t} n + \nabla_x \cdot (n\mathbf{u}) = 0. \quad (2.22)$$

This verifies the first equation, which is an expression for the conservation of mass.

2. Next, multiplying by $\mathbf{v}$ and integrating, we have

$$\int_{\mathbb{R}^d} \left[\mathbf{v} \frac{\partial f}{\partial t} + \nabla_x \cdot (\mathbf{v} \otimes \mathbf{v} f) \right] d\mathbf{v} = \int_{\mathbb{R}^d} \mathbf{v} \lambda (M_f - f) d\mathbf{v} \quad (2.23)$$

$$\iff \frac{\partial}{\partial t} \left(\int_{\mathbb{R}^d} \mathbf{v} f d\mathbf{v} \right) + \nabla_x \cdot \left(\int_{\mathbb{R}^d} \mathbf{v} \otimes \mathbf{v} f d\mathbf{v} \right) = \mathbf{0}. \quad (2.24)$$

The first term is equal to $\partial_t(n\mathbf{u})$, as desired. It remains to show that

$$\int_{\mathbb{R}^d} \mathbf{v} \otimes \mathbf{v} f d\mathbf{v} = n\mathbf{u} \otimes \mathbf{u} + \mathbf{P}.$$

Writing $\mathbf{v} = \mathbf{u} + (\mathbf{v} - \mathbf{u})$, gives

$$\begin{aligned} \int_{\mathbb{R}^d} \mathbf{v} \otimes \mathbf{v} f d\mathbf{v} &= \int_{\mathbb{R}^d} [\mathbf{u} + (\mathbf{v} - \mathbf{u})] \otimes [\mathbf{u} + (\mathbf{v} - \mathbf{u})] f d\mathbf{v} \\ &= \int_{\mathbb{R}^d} [\mathbf{u} \otimes \mathbf{u} + \mathbf{u} \otimes (\mathbf{v} - \mathbf{u}) + (\mathbf{v} - \mathbf{u}) \otimes \mathbf{u} + (\mathbf{v} - \mathbf{u}) \otimes (\mathbf{v} - \mathbf{u})] f d\mathbf{v} \\ &= n\mathbf{u} \otimes \mathbf{u} + \mathbf{P} + \int_{\mathbb{R}^d} \mathbf{u} \otimes (\mathbf{v} - \mathbf{u}) f d\mathbf{v} + \int_{\mathbb{R}^d} (\mathbf{v} - \mathbf{u}) \otimes \mathbf{u} f d\mathbf{v}. \end{aligned} \quad (2.25)$$

The last two terms in the above expression are equal to zero:

$$\mathbf{u} \otimes \int_{\mathbb{R}^d} (\mathbf{v} - \mathbf{u}) f d\mathbf{v} = \mathbf{u} \otimes \left[\int_{\mathbb{R}^d} \mathbf{v} f d\mathbf{v} - \mathbf{u} \int_{\mathbb{R}^d} f d\mathbf{v} \right] = \mathbf{u} \otimes [n\mathbf{u} - n\mathbf{u}] = \mathbf{0}.$$

Putting it all together, we obtain the second equation, an expression for the conservation of momentum.

3. Finally, multiplying by $\frac{1}{2}|\mathbf{v}|^2$ and integrating, gives

$$\int_{\mathbb{R}^d} \frac{1}{2} |\mathbf{v}|^2 \left(\frac{\partial f}{\partial t} + \nabla_x \cdot (\mathbf{v} f) \right) d\mathbf{v} = \int_{\mathbb{R}^d} \frac{1}{2} |\mathbf{v}|^2 \lambda (M_f - f) d\mathbf{v} \quad (2.26)$$

$$\iff \frac{\partial}{\partial t} \left(\int_{\mathbb{R}^d} \frac{1}{2} |\mathbf{v}|^2 f d\mathbf{v} \right) + \nabla_x \cdot \left(\int_{\mathbb{R}^d} \frac{1}{2} |\mathbf{v}|^2 \mathbf{v} f d\mathbf{v} \right) = \mathbf{0}. \quad (2.27)$$

The first term is equal to $\partial_t E$, as desired. It remains to show that

$$\int_{\mathbb{R}^d} \frac{1}{2} |\mathbf{v}|^2 \mathbf{v} f \, d\mathbf{v} = E\mathbf{u} + \mathbf{P}\mathbf{u} + \mathbf{q}.$$

Writing $\mathbf{v} = \mathbf{u} + (\mathbf{v} - \mathbf{u})$, gives

$$\begin{aligned} \int_{\mathbb{R}^d} \frac{1}{2} |\mathbf{v}|^2 \mathbf{v} f \, d\mathbf{v} &= \frac{1}{2} \int_{\mathbb{R}^d} |\mathbf{v}|^2 [\mathbf{u} + (\mathbf{v} - \mathbf{u})] f \, d\mathbf{v} \\ &= \mathbf{u} \int_{\mathbb{R}^d} \frac{1}{2} |\mathbf{v}|^2 f \, d\mathbf{v} + \frac{1}{2} \int_{\mathbb{R}^d} |\mathbf{v}|^2 (\mathbf{v} - \mathbf{u}) f \, d\mathbf{v} \\ &= E\mathbf{u} + \frac{1}{2} \int_{\mathbb{R}^d} [\mathbf{u} + (\mathbf{v} - \mathbf{u})]^\top [\mathbf{u} + (\mathbf{v} - \mathbf{u})] (\mathbf{v} - \mathbf{u}) f \, d\mathbf{v} \\ &= E\mathbf{u} + \frac{1}{2} \int_{\mathbb{R}^d} [|\mathbf{u}|^2 + 2\mathbf{u}^\top (\mathbf{v} - \mathbf{u}) + |\mathbf{v} - \mathbf{u}|^2] (\mathbf{v} - \mathbf{u}) f \, d\mathbf{v} \\ &= E\mathbf{u} + \frac{|\mathbf{u}|^2}{2} \int_{\mathbb{R}^d} (\mathbf{v} - \mathbf{u}) f \, d\mathbf{v} + \int_{\mathbb{R}^d} \mathbf{u}^\top (\mathbf{v} - \mathbf{u}) (\mathbf{v} - \mathbf{u}) f \, d\mathbf{v} + \frac{1}{2} \int_{\mathbb{R}^d} |\mathbf{v} - \mathbf{u}|^2 (\mathbf{v} - \mathbf{u}) f \, d\mathbf{v} \\ &= E\mathbf{u} + \mathbf{q} + \frac{1}{2} |\mathbf{u}|^2 \left(\int_{\mathbb{R}^d} \mathbf{v} f \, d\mathbf{v} - \mathbf{u} \int_{\mathbb{R}^d} f \, d\mathbf{v} \right) + \int_{\mathbb{R}^d} \mathbf{u}^\top (\mathbf{v} - \mathbf{u}) (\mathbf{v} - \mathbf{u}) f \, d\mathbf{v} \\ &= E\mathbf{u} + \mathbf{q} + \frac{1}{2} |\mathbf{u}|^2 (n\mathbf{u} - \mathbf{u}n) + \int_{\mathbb{R}^d} \mathbf{u}^\top (\mathbf{v} - \mathbf{u}) (\mathbf{v} - \mathbf{u}) f \, d\mathbf{v} \\ &= E\mathbf{u} + \mathbf{q} + \int_{\mathbb{R}^d} \mathbf{u}^\top (\mathbf{v} - \mathbf{u}) (\mathbf{v} - \mathbf{u}) f \, d\mathbf{v}. \end{aligned} \tag{2.28}$$

It should be easy to see that

$$\mathbf{u}^\top (\mathbf{v} - \mathbf{u}) (\mathbf{v} - \mathbf{u}) = [(\mathbf{v} - \mathbf{u}) \otimes (\mathbf{v} - \mathbf{u})] \mathbf{u}.$$

Thus, the final term in the above Equation is

$$\begin{aligned} \int_{\mathbb{R}^d} \mathbf{u}^\top (\mathbf{v} - \mathbf{u}) (\mathbf{v} - \mathbf{u}) f \, d\mathbf{v} &= \int_{\mathbb{R}^d} [(\mathbf{v} - \mathbf{u}) \otimes (\mathbf{v} - \mathbf{u})] \mathbf{u} f \, d\mathbf{v} \\ &= \left(\int_{\mathbb{R}^d} (\mathbf{v} - \mathbf{u}) \otimes (\mathbf{v} - \mathbf{u}) f \, d\mathbf{v} \right) \mathbf{u} \\ &= \mathbf{P}\mathbf{u}. \end{aligned} \tag{2.29}$$

Putting it all together gives the third equation, which is an expression for the conservation of energy.

The proof is complete. ■

The conservation laws/properties above are reminiscent of those involved with the deriva-

tions of the Navier-Stokes and Euler equations. In fact, using the Chapman-Enskog expansion method one can show that in the limit as $\lambda \rightarrow \infty$, or $\tau \rightarrow 0$, one recovers macroscopic Navier-Stokes and/or Euler equations as formal limits, under certain assumptions [27, 32, 31, 29]. These limits can even guide in the design of stable numerical methods for the macroscopic models [32].

The following is an even simpler consequence of the collision invariance of $\mathbf{m}$ stated in Lemma 2.1.2. Its proof is omitted.

Proposition 2.1.6. *Suppose that f is a non-negative solution to the BGK equation. Then,*

$$\partial_t \langle f \mathbf{m} \rangle + \nabla_{\mathbf{x}} \langle \mathbf{v} \otimes \mathbf{m} f \rangle = \mathbf{0}.$$

2.1.2 Space-Homogeneous Problem

Suppose that the density function is spatially homogeneous and/or particle advection may be neglected in the system. In this case, the distribution function f satisfies the space-homogeneous problem

$$\frac{\partial f}{\partial t} = \lambda(M_f - f). \quad (2.30)$$

This is a first order integro-differential equation (IDE), which can be solved using the integrating factor method.

Lemma 2.1.7. *Suppose that f is a solution to the space homogeneous problem (2.30). Then*

$$f(\mathbf{x}, \mathbf{v}, t) = e^{-\lambda t} f(\mathbf{x}, \mathbf{v}, 0) + (1 - e^{-\lambda t}) M_f(\mathbf{x}, \mathbf{v}, 0). \quad (2.31)$$

Proof. We begin by proving that $\frac{\partial n}{\partial t} = 0 = \frac{\partial \theta}{\partial t}$, $\frac{\partial \mathbf{u}}{\partial t} = \mathbf{0}$, and $\frac{\partial M_f}{\partial t} = 0$. We will make frequent use of Lemma 2.1.2.

1. First, we show that $\frac{\partial n}{\partial t} = 0$. Utilizing the IDE,

$$\frac{\partial}{\partial t} n = \frac{\partial}{\partial t} \int_{\mathbb{R}^d} f \, d\mathbf{v} = \int_{\mathbb{R}^d} \frac{\partial f}{\partial t} \, d\mathbf{v} = \int_{\mathbb{R}^d} \lambda(M_f - f) \, d\mathbf{v} \stackrel{2.1.2}{=} 0. \quad (2.32)$$

2. Next, we show that $\frac{\partial \mathbf{u}}{\partial t} = \mathbf{0}$. Using the quotient rule on the definition of $\mathbf{u} = \frac{1}{n} \int_{\mathbb{R}^d} f \, d\mathbf{v}$, and the fact that $\frac{\partial n}{\partial t} = 0$ gives

$$\frac{\partial \mathbf{u}}{\partial t} = \frac{\partial}{\partial t} \frac{1}{n} \int_{\mathbb{R}^d} f \mathbf{v} \, d\mathbf{v} = \frac{(n) \int_{\mathbb{R}^d} \frac{\partial f}{\partial t} \mathbf{v} \, d\mathbf{v} - 0}{n^2} = \frac{1}{n} \int_{\mathbb{R}^d} \lambda(M_f - f) \mathbf{v} \, d\mathbf{v} \stackrel{2.1.2}{=} \mathbf{0}. \quad (2.33)$$

3. Next, we show that $\frac{\partial \theta}{\partial t} = 0$. Recall that $\frac{d}{dt} n\theta = E - \frac{1}{2}n|\mathbf{u}|^2 \iff \theta = \frac{2E}{nd} - \frac{1}{d}|\mathbf{u}|^2$. Using the quotient rule, and the fact that $\frac{\partial n}{\partial t} = 0$ and $\frac{\partial \mathbf{u}}{\partial t} = \mathbf{0}$, gives

$$\frac{\partial \theta}{\partial t} = \frac{2}{d} \frac{\left(\frac{\partial E}{\partial t}\right)(n) - 0}{n^2} - \frac{2}{d} \mathbf{u}^\top \frac{\partial \mathbf{u}}{\partial t} = \frac{2}{nd} \int_{\mathbb{R}^d} \frac{|\mathbf{v}|^2}{2} \lambda(M_f - f) d\mathbf{v} \stackrel{2.1.2}{=} 0. \quad (2.34)$$

4. Next, to show $\frac{\partial M_f}{\partial t} = 0$, we use the temporal invariance of n , $\mathbf{u}$, and θ , as shown above,

$$\begin{aligned} \frac{\partial}{\partial t} M_f(\mathbf{x}, \mathbf{v}, t) &= (2\pi)^{-\frac{d}{2}} \frac{\partial}{\partial t} \left[n(\mathbf{x}, t) \theta(\mathbf{x}, t)^{-\frac{d}{2}} \exp \left(-\frac{1}{2} |\mathbf{v} - \mathbf{u}(\mathbf{x}, t)|^2 \theta(\mathbf{x}, t)^{-1} \right) \right] \\ &= (2\pi)^{-\frac{d}{2}} \left[\frac{\partial n}{\partial t} \theta^{-\frac{d}{2}} \exp \left(-\frac{1}{2} |\mathbf{v} - \mathbf{u}|^2 \theta^{-1} \right) \right. \\ &\quad + n(\mathbf{x}, t) \left(-\frac{d}{2} \right) \theta^{-\frac{d+2}{2}} \frac{\partial \theta}{\partial t} \exp \left(-\frac{1}{2} |\mathbf{v} - \mathbf{u}|^2 \theta^{-1} \right) \\ &\quad \left. + n\theta^{-\frac{d}{2}} \exp \left(-\frac{1}{2} |\mathbf{v} - \mathbf{u}|^2 \theta^{-1} \right) \frac{\partial}{\partial t} \left[-\frac{1}{2} |\mathbf{v} - \mathbf{u}|^2 \theta^{-1} \right] \right] \\ &= 0. \end{aligned} \quad (2.35)$$

5. So far, we have $\frac{\partial n}{\partial t} = 0 = \frac{\partial \theta}{\partial t}$, $\frac{\partial \mathbf{u}}{\partial t} = \mathbf{0}$, and $\frac{\partial M_f}{\partial t} = 0$. To finish, we use these properties to solve the IDE. Rearranging (2.30), gives

$$\frac{df}{dt} + \lambda f = \lambda M.$$

Multiplying by the integrating factor, $e^{\lambda t}$, and applying the product rule in reverse, gives

$$\frac{d}{dt} (e^{\lambda t} f) = \lambda M_f e^{\lambda t}.$$

Integrating with respect to t , we have

$$e^{\lambda t} f = C + \lambda \int e^{\lambda t} M_f(\mathbf{x}, \mathbf{v}, t) dt,$$

Since $M_f = M_f(\mathbf{x}, \mathbf{v}, t)$ is constant with respect to t , $M_f(\mathbf{x}, \mathbf{v}, t) = M_f(\mathbf{x}, \mathbf{v}, 0)$. Hence, it can be pulled through the integral sign:

$$\begin{aligned} e^{\lambda t} f &= C + \lambda M_f(\mathbf{x}, \mathbf{v}, 0) \int e^{\lambda t} dt \\ &= C + M_f(\mathbf{x}, \mathbf{v}, 0) e^{\lambda t}. \end{aligned} \quad (2.36)$$

Plugging in $t = 0$, and rearranging gives an expression for the constant term:

$$e^0 f(\mathbf{x}, \mathbf{v}, 0) = C + e^0 M(\mathbf{x}, \mathbf{v}, 0) \quad (2.37)$$

$$\iff C = f(\mathbf{x}, \mathbf{v}, 0) - M(\mathbf{x}, \mathbf{v}, 0). \quad (2.38)$$

Putting everything together, and multiplying by $e^{-\lambda t}$, we have the following solution to the space homogeneous problem:

$$\begin{aligned} f(\mathbf{x}, \mathbf{v}, t) &= e^{-\lambda t} (f(\mathbf{x}, \mathbf{v}, 0) - M_f(\mathbf{x}, \mathbf{v}, 0)) + M_f(\mathbf{x}, \mathbf{v}, 0) \\ &= e^{-\lambda t} f(\mathbf{x}, \mathbf{v}, 0) + (1 - e^{-\lambda t}) M_f(\mathbf{x}, \mathbf{v}, 0). \end{aligned} \quad (2.39)$$

■

Since we have the true solution to this space homogeneous problem, this allows us to test the accuracy of the code on the right hand side source term. Initial tests are performed in Section 3.5.

2.1.3 The H-Theorem

In this section, we discuss a very important solution property for the BGK equation, namely the entropy dissipation property. This is a key stability concept that should, in some way, be preserved in numerical approximations. We start off this section with a definition.

Definition 2.1.8. Suppose that $f : \Omega \times \mathbb{R}^d \times [0, \infty) \rightarrow [0, \infty)$ is a particle density function. The object

$$H[f](\mathbf{x}, t) := \int_{\mathbb{R}^d} f(\mathbf{x}, \mathbf{v}, t) \ln(f(\mathbf{x}, \mathbf{v}, t)) d\mathbf{v}$$

is called the **H functional** or **kinetic entropy**. We say that a flow is **entropy-dissipative** or **entropy-stable** iff, for every $0 \leq t_1 \leq t_2 < \infty$,

$$\int_{\Omega} H[f](\mathbf{x}, t_2) d\mathbf{x} \leq \int_{\Omega} H[f](\mathbf{x}, t_1) d\mathbf{x}.$$

We will need the following technical lemma.

Lemma 2.1.9. For any $x, y \in (0, \infty)$

$$(\ln(x) - \ln(y))(x - y) \geq 0.$$

Proof. Observe that the function $q(x) = x \ln(x)$ is strictly convex on $[0, \infty)$. In fact, for any $x \in (0, 1)$,

$$q''(x) = \frac{1}{x} > 0.$$

By Taylor's theorem, for any $x, y \in (0, \infty)$,

$$x \ln(x) = y \ln(y) + (\ln(y) + 1)(x - y) + \frac{1}{2} \frac{1}{\xi} (x - y)^2 \geq y \ln(y) + (\ln(y) + 1)(x - y),$$

for some ξ between x and y . The inequality above can be rewritten as

$$x \ln(x) - x \ln(y) \geq x - y.$$

Reversing the roles of x and y , we have

$$y \ln(y) - y \ln(x) \geq y - x.$$

Adding the inequalities, we have

$$(x - y)(\ln(x) - \ln(y)) \geq 0,$$

which is the desired result. ■

Lemma 2.1.10. *The function $\ln(M_f)$ is a collision invariant, that is*

$$\int_{\mathbb{R}^d} \ln(M_f) (M_f - f) d\mathbf{v} = 0.$$

Proof. We already know that 1 , $\mathbf{v}$, and $|\mathbf{v}|^2$ are collision invariants, as are any linear combinations of these functions. Using (2.3), (2.4), and (2.5), we have

$$\ln(M_f) = \ln(\exp(\boldsymbol{\alpha} \cdot \mathbf{m})) = \boldsymbol{\alpha} \cdot \mathbf{m} = \ln\left(\frac{n}{(2\pi\theta)^{d/2}}\right) - \frac{|\mathbf{v} - \mathbf{u}|^2}{2\theta},$$

and it follows that $\ln(M_f)$ is also a collision invariant. ■

Theorem 2.1.11. *Suppose that $f : \Omega \times \mathbb{R}^d \times [0, \infty) \rightarrow [0, \infty)$ is a solution to the spatially homogeneous BGK problem, that is,*

$$\frac{\partial f}{\partial t} = \frac{1}{\tau} (M_f - f),$$

where $\tau > 0$ is a constant. In particular, let us assume that f has no variation with respect to $\mathbf{x}$, i.e., $\nabla_{\mathbf{x}} f = \mathbf{0}$. Then

$$\frac{d}{dt} H[f] \leq 0.$$

Proof. Observe that

$$\begin{aligned} \frac{d}{dt} H[f](t) &= \int_{\mathbb{R}^d} \frac{\partial f}{\partial t} (\ln(f) + 1) \, d\mathbf{v} \\ &= \frac{1}{\tau} \int_{\mathbb{R}^d} (M_f - f) (\ln(f) + 1) \, d\mathbf{v} \\ &= \frac{1}{\tau} \int_{\mathbb{R}^d} \ln(f) (M_f - f) \, d\mathbf{v} \end{aligned}$$

To finish the proof, we use that fact that $\ln(M_f)$ is a collision invariant, that is

$$\int_{\mathbb{R}^d} \ln(M_f) (M_f - f) \, d\mathbf{v} = 0.$$

Thus

$$\begin{aligned} \frac{d}{dt} H[f](t) &= \frac{1}{\tau} \int_{\mathbb{R}^d} (\ln(f) - \ln(M_f)) (M_f - f) \, d\mathbf{v} \\ &= -\frac{1}{\tau} \int_{\mathbb{R}^d} (\ln(M_f) - \ln(f)) (M_f - f) \, d\mathbf{v} \leq 0, \end{aligned}$$

where in the last step we used the fact that

$$(\ln(x) - \ln(y)) (x - y) \geq 0, \quad \forall x, y \in (0, \infty).$$

■

More generally, we have

Theorem 2.1.12. Suppose that $f : \Omega \times \mathbb{R}^d \times [0, \infty) \rightarrow [0, \infty)$ is an Ω -periodic (spatially periodic) solution to the BGK equation, that is,

$$\partial_t f + \mathbf{v} \cdot \nabla_{\mathbf{x}} f = \frac{1}{\tau} (M_f - f),$$

where $\tau > 0$ is constant. Then

$$\partial_t \langle f \ln(f) \rangle + \nabla_{\mathbf{x}} \cdot \langle \mathbf{v} f \ln(f) \rangle \leq 0.$$

and

$$\frac{d}{dt} \int_{\Omega} H[f] d\mathbf{x} \leq 0.$$

In other words, solutions of the BGK equation are entropy dissipative.

Proof. Similar to the last proof, but using Lemma 2.1.5, we have

$$\begin{aligned} \frac{\partial}{\partial t} H[f](t) &= \int_{\mathbb{R}^d} \frac{\partial f}{\partial t} (\ln(f) + 1) d\mathbf{v} \\ &= \int_{\mathbb{R}^d} \frac{\partial f}{\partial t} \ln(f) d\mathbf{v} + \frac{\partial n}{\partial t} \\ &= \frac{1}{\tau} \int_{\mathbb{R}^d} (M_f - f) \ln(f) d\mathbf{v} - \int_{\mathbb{R}^d} \mathbf{v} \cdot \nabla_x f \ln(f) d\mathbf{v} + \frac{\partial n}{\partial t} \\ &= \frac{1}{\tau} \int_{\mathbb{R}^d} (M_f - f) \ln(f) d\mathbf{v} - \nabla_x \cdot \int_{\mathbb{R}^d} \mathbf{v} f \ln(f) d\mathbf{v} + \int_{\mathbb{R}^d} \mathbf{v} f \cdot \nabla_x (\ln(f)) d\mathbf{v} + \frac{\partial n}{\partial t} \\ &= \frac{1}{\tau} \int_{\mathbb{R}^d} \ln(f) (M_f - f) d\mathbf{v} - \nabla_x \cdot \int_{\mathbb{R}^d} \mathbf{v} f \ln(f) d\mathbf{v} + \frac{\partial n}{\partial t} + \nabla_x \cdot (n\mathbf{u}) \\ &= \frac{1}{\tau} \int_{\mathbb{R}^d} \ln(f) (M_f - f) d\mathbf{v} - \nabla_x \cdot \int_{\mathbb{R}^d} \mathbf{v} f \ln(f) d\mathbf{v}. \end{aligned}$$

Therefore,

$$\partial_t \int_{\mathbb{R}^d} f \ln(f) d\mathbf{v} + \nabla_x \cdot \int_{\mathbb{R}^d} \mathbf{v} f \ln(f) d\mathbf{v} = -\frac{1}{\tau} \int_{\mathbb{R}^d} (f - M_f)(\ln(f) - \ln(M_f)) d\mathbf{v} \leq 0.$$

In other words,

$$\partial_t \langle f \ln(f) \rangle + \nabla_x \cdot \langle \mathbf{v} f \ln(f) \rangle \leq 0.$$

Using the Ω -periodicity of f and integrating over Ω , we have

$$d_t \int_{\Omega} \int_{\mathbb{R}^d} f \ln(f) d\mathbf{v} d\mathbf{x} = -\frac{1}{\tau} \int_{\Omega} \int_{\mathbb{R}^d} (f - M_f)(\ln(f) - \ln(M_f)) d\mathbf{v} d\mathbf{x} \leq 0,$$

and the proof is complete. ■

Finally, we have

Theorem 2.1.13. *Suppose that*

$$f_* : \Omega \times \mathbb{R}^d \rightarrow [0, \infty)$$

is a given Ω -periodic distribution. Let

$$\boldsymbol{\rho}_* = [n_*, n_* \mathbf{u}_*, E_*]^\top,$$

be the vector comprised of the first three moments of f_* , that is,

$$\boldsymbol{\rho}_*(\mathbf{x}) = \begin{bmatrix} n_*(\mathbf{x}) \\ n_*(\mathbf{x}) \mathbf{u}_*(\mathbf{x}) \\ E_*(\mathbf{x}) \end{bmatrix} = \left\langle f_*(\mathbf{x}, \mathbf{v}) \begin{bmatrix} 1 \\ \mathbf{v} \\ \frac{1}{2} |\mathbf{v}|^2 \end{bmatrix} \right\rangle, \quad \forall \mathbf{x} \in \Omega.$$

The Ω -periodic temperature, θ_* , is defined via

$$E_*(\mathbf{x}) = \frac{1}{2} n_*(\mathbf{x}) |\mathbf{u}_*(\mathbf{x})|^2 + \frac{d}{2} n_*(\mathbf{x}) \theta_*(\mathbf{x}).$$

The Maxwellian, as usual, is defined using the macroscopic densities as

$$M_*(\mathbf{x}, \mathbf{v}) := n_* \left(\frac{1}{2\pi\theta_*} \right)^{\frac{d}{2}} \exp \left(-\frac{|\mathbf{v} - \mathbf{u}_*|^2}{2\theta_*} \right).$$

Consider the admissible class of distribution functions defined as

$$\mathcal{A}_* := \left\{ g : \Omega \times \mathbb{R}^d \rightarrow \mathbb{R} \mid g \text{ is } \Omega\text{-periodic, } g \geq 0, \text{ and } \langle \mathbf{m}g \rangle = \boldsymbol{\rho}_* \right\},$$

where, recall,

$$\mathbf{m}(\mathbf{v}) := \left[1, \mathbf{v}, \frac{1}{2} |\mathbf{v}|^2 \right]^\top : \mathbb{R}^d \rightarrow \mathbb{R}^{d+2}.$$

Then, $M_*, f_* \in \mathcal{A}_*$. In other words, $\mathcal{A}_*$ is non-empty. For every $g \in \mathcal{A}_*$,

$$H[M_*](\mathbf{x}) \leq H[g](\mathbf{x}), \quad \forall \mathbf{x} \in \Omega.$$

In other words, the Maxwellian is the minimum entropy distribution out of the set of admissible distributions that gives rise to the macroscopic density $\boldsymbol{\rho}_*$.

Proof. Let us first show that $M_* \in \mathcal{A}_*$. Clearly M_* is nonnegative and Ω -periodic. The fact that $\langle \mathbf{m}M_* \rangle = \boldsymbol{\rho}_*$ follows from Lemma 2.1.2. Thus $M_* \in \mathcal{A}_*$. Next, we prove the inequality. Let $g \in \mathcal{A}_*$ be arbitrary. We may write $M_* = M_g$, since g gives rise to the same macroscopic fields as f_* . Define $h(s) := s \ln(s)$. This function h is convex on $(0, \infty)$, as is easy to show.

Its first two derivatives are

$$h'(s) = \ln(s) + 1, \quad h''(s) = \frac{1}{s} > 0, \quad \forall s > 0.$$

Thus,

$$h(s) \geq h(t) + h'(t)(s - t), \quad \forall s, t \in (0, \infty).$$

It follows that

$$\begin{aligned} h(g) &\geq h(M_g) + h'(M_g)(g - M_g) \\ &= h(M_g) + \ln(M_g)(g - M_g) + g - M_g. \end{aligned}$$

By Lemma 2.1.10,

$$\int_{\mathbb{R}^d} h(g) \, d\mathbf{v} \geq \int_{\mathbb{R}^d} h(M_g) \, d\mathbf{v} + \int_{\mathbb{R}^d} \ln(M_g)(g - M_g) \, d\mathbf{v} + \int_{\mathbb{R}^d} (g - M_g) \, d\mathbf{v} = \int_{\mathbb{R}^d} h(M_g) \, d\mathbf{v}.$$

This shows that, point-wise in Ω ,

$$H[g] \geq H[M_g] = H[M_*],$$

and the result is proven. ■

We should point out that workers in equilibrium thermodynamics and non-equilibrium thermodynamics generally prefer a definition of entropy that sees the entropy increasing as a function of time. But, for historical reasons, in the mathematical and numerical theory of the Boltzmann and the BGK equations, the prevailing definition of entropy is such that it is decreasing in time. In any case, this is simply a matter of a sign difference, and the mathematical dissipation property is an important marker for the design of numerical methods. In particular, a numerical approximation scheme should satisfy, if possible, some discrete form of entropy dissipation. However, designing fully discrete approximation schemes that *theoretically* satisfy discrete dissipation (as determined by a rigorous proof) is a challenging task. The papers [16, 17] address this issue for the single species BGK equation, but, this dissipation property comes at a rather high computational cost. Thus, it is not clear from the outset that it is practical to pursue this property from the theoretical point of view. On the other hand, checking the dissipation property numerically for benchmark simulations is certainly a worthwhile endeavor.

2.2 Discrete-Velocity BGK Models

Discrete velocity BGK models are constructed under the assumption that the velocity space is partitioned into discrete, allowable states. This discretized velocity space could be finite or countably infinite, in principle. But, typically the velocity space is constructed so that it is finite and approximates well the density of particles, f , for a given physical case, meaning that the velocity points might clustered around particular velocity values. To make the discussion simple, we will restrict our attention to the $d = 1$ case. The general case is constructed similarly.

2.2.1 Discrete-Velocity BGK Model in One Space Dimension

Let us make an analog model in one space dimension ($d = 1$) with discrete, rather than continuous, velocities, along the lines of the papers [16, 17]. Suppose that $h_v \in (0, 1)$ and $j \in \mathbb{N}$. Set

$$\mathcal{V}_{\mathcal{J}} := \left\{ v_j = h_v \cdot \left(j - \frac{1}{2} \right) \mid j \in \mathcal{J} \right\},$$

where

$$\mathcal{J} := \{-J + 1, \dots, J\}.$$

Note that the simple structure of $\mathcal{V}_{\mathcal{J}}$, with uniformly spaced points arranged symmetrically about 0, is used only for simplicity, and we may choose it in a more general fashion. In any case, set

$$\mathbf{m}_{\mathcal{J}}(j) := \left[1, v_j, \frac{1}{2}|v_j|^2 \right]^{\top}, \quad \forall j \in \mathcal{J}.$$

Define, for $\mathbf{g} : \mathcal{J} \rightarrow \mathbb{R}^d$

$$\langle \mathbf{g} \rangle_{\mathcal{J}} := \sum_{j \in \mathcal{J}} \mathbf{g}_j h_v.$$

Let us define the admissible class of solutions.

Definition 2.2.1. Suppose that $\Omega = [a, b] \subset \mathbb{R}$ and

$$f_{\mathcal{J},*} : \Omega \times \mathcal{J} \rightarrow [0, \infty)$$

is a given Ω -periodic discrete-velocity distribution. Let

$$\boldsymbol{\rho}_{\mathcal{J},*} = [n_{\mathcal{J},*}, n_{\mathcal{J},*} u_{\mathcal{J},*}, E_{\mathcal{J},*}]^{\top},$$

be the vector comprised of the first three discrete-velocity moments of $f_{\mathcal{J},*}$, that is,

$$\boldsymbol{\rho}_{\mathcal{J},*}(x) = \begin{bmatrix} n_{\mathcal{J},*}(x) \\ n_{\mathcal{J},*}(x)u_{\mathcal{J},*}(x) \\ E_{\mathcal{J},*}(x) \end{bmatrix} = \langle f_{\mathcal{J},*}(x, \cdot) \mathbf{m}_{\mathcal{J}} \rangle_{\mathcal{J}}, \quad \forall x \in \Omega. \quad (2.40)$$

The **admissible class of discrete-velocity distribution functions**, subordinate to $\boldsymbol{\rho}_{\mathcal{J},*}$ or $f_{\mathcal{J},*}$ is defined as the set

$$\mathcal{A}_{\mathcal{J},*} := \{g_{\mathcal{J}} : \Omega \times \mathcal{J} \rightarrow \mathbb{R} \mid g_{\mathcal{J}} \text{ is } \Omega\text{-periodic}, \quad g_{\mathcal{J}} \geq 0, \quad \text{and} \quad \langle \mathbf{m}_{\mathcal{J}} g_{\mathcal{J}} \rangle_{\mathcal{J}} = \boldsymbol{\rho}_{\mathcal{J},*}\}.$$

We need the following technical lemma.

Lemma 2.2.2. Suppose that the moment vector $\boldsymbol{\rho}_{\mathcal{J},*}$ is realized from an Ω -periodic discrete-velocity density as above. Fix $x \in \Omega = [a, b]$, and define the function $K : \mathbb{R}^3 \rightarrow \mathbb{R}$ via

$$K(\boldsymbol{\beta}) = K_x(\boldsymbol{\beta}) := \langle \exp(\boldsymbol{\beta} \cdot \mathbf{m}_{\mathcal{J}}) \rangle_{\mathcal{J}} - \boldsymbol{\beta} \cdot \boldsymbol{\rho}_{\mathcal{J},*}(x), \quad \forall \boldsymbol{\beta} \in \mathbb{R}^3.$$

$K(\cdot)$ is smooth, convex, and coercive and, therefore, has a unique minimizer, $\boldsymbol{\beta}_* \in \mathbb{R}^3$, which satisfies the nonlinear equation

$$\langle \exp(\boldsymbol{\beta}_* \cdot \mathbf{m}_{\mathcal{J}}) \mathbf{m}_{\mathcal{J}} \rangle_{\mathcal{J}} = \boldsymbol{\rho}_{\mathcal{J},*}(x).$$

Proof. Clearly, K is smooth. To see that it is convex, consider the first two functional derivatives. The first derivative is computed via

$$\begin{aligned} \frac{d}{ds} K(\boldsymbol{\beta} + s\boldsymbol{\eta}) &= \frac{d}{ds} \langle \exp(\boldsymbol{\beta} \cdot \mathbf{m}_{\mathcal{J}}) \exp(s\boldsymbol{\eta} \cdot \mathbf{m}_{\mathcal{J}}) \rangle_{\mathcal{J}} - \frac{d}{ds} (\boldsymbol{\beta} + s\boldsymbol{\eta}) \cdot \boldsymbol{\rho}_{\mathcal{J},*}(x) \\ &= \langle \exp(\boldsymbol{\beta} \cdot \mathbf{m}_{\mathcal{J}}) \boldsymbol{\eta} \cdot \mathbf{m}_{\mathcal{J}} \exp(s\boldsymbol{\eta} \cdot \mathbf{m}_{\mathcal{J}}) \rangle_{\mathcal{J}} - \boldsymbol{\eta} \cdot \boldsymbol{\rho}_{\mathcal{J},*}(x). \end{aligned}$$

Thus,

$$\begin{aligned} \left. \frac{d}{ds} K(\boldsymbol{\beta} + s\boldsymbol{\eta}) \right|_{s=0} &= \langle \exp(\boldsymbol{\beta} \cdot \mathbf{m}_{\mathcal{J}}) \boldsymbol{\eta} \cdot \mathbf{m}_{\mathcal{J}} \rangle_{\mathcal{J}} - \boldsymbol{\eta} \cdot \boldsymbol{\rho}_{\mathcal{J},*}(x) \\ &= \boldsymbol{\eta} \cdot \{ \langle \exp(\boldsymbol{\beta} \cdot \mathbf{m}_{\mathcal{J}}) \mathbf{m}_{\mathcal{J}} \rangle_{\mathcal{J}} - \boldsymbol{\rho}_{\mathcal{J},*}(x) \}, \end{aligned}$$

and we can write

$$\nabla K(\boldsymbol{\beta}) = \langle \exp(\boldsymbol{\beta} \cdot \mathbf{m}_{\mathcal{J}}) \mathbf{m}_{\mathcal{J}} \rangle_{\mathcal{J}} - \boldsymbol{\rho}_{\mathcal{J},*}(x).$$

Further,

$$\begin{aligned} \frac{d^2}{ds^2} K(\boldsymbol{\beta} + s\boldsymbol{\eta}) &= \frac{d}{ds} \langle \exp(\boldsymbol{\beta} \cdot \mathbf{m}_{\mathcal{J}}) \boldsymbol{\eta} \cdot \mathbf{m}_{\mathcal{J}} \exp(s\boldsymbol{\eta} \cdot \mathbf{m}_{\mathcal{J}}) \rangle_{\mathcal{J}} \\ &= \langle \exp(\boldsymbol{\beta} \cdot \mathbf{m}_{\mathcal{J}}) (\boldsymbol{\eta} \cdot \mathbf{m}_{\mathcal{J}})^2 \exp(s\boldsymbol{\eta} \cdot \mathbf{m}_{\mathcal{J}}) \rangle_{\mathcal{J}} \\ &= \boldsymbol{\eta}^T [\langle \mathbf{m}_{\mathcal{J}} \mathbf{m}_{\mathcal{J}}^T \exp(\boldsymbol{\beta} \cdot \mathbf{m}_{\mathcal{J}}) \exp(s\boldsymbol{\eta} \cdot \mathbf{m}_{\mathcal{J}}) \rangle_{\mathcal{J}}] \boldsymbol{\eta}. \end{aligned}$$

Thus,

$$\left. \frac{d^2}{ds^2} K(\boldsymbol{\beta} + s\boldsymbol{\eta}) \right|_{s=0} = \boldsymbol{\eta}^T [\langle \mathbf{m}_{\mathcal{J}} \mathbf{m}_{\mathcal{J}}^T \exp(\boldsymbol{\beta} \cdot \mathbf{m}_{\mathcal{J}}) \rangle_{\mathcal{J}}] \boldsymbol{\eta},$$

and we write

$$H_K(\boldsymbol{\beta}) = \langle \mathbf{m}_{\mathcal{J}} \mathbf{m}_{\mathcal{J}}^T \exp(\boldsymbol{\beta} \cdot \mathbf{m}_{\mathcal{J}}) \rangle_{\mathcal{J}},$$

which is the 3×3 Hessian matrix, of course. This matrix is positive definite, since

$$\boldsymbol{\eta}^T H_K(\boldsymbol{\beta}) \boldsymbol{\eta} = \langle (\boldsymbol{\eta} \cdot \mathbf{m}_{\mathcal{J}})^2 \exp(\boldsymbol{\beta} \cdot \mathbf{m}_{\mathcal{J}}) \rangle_{\mathcal{J}} > 0,$$

for every $\boldsymbol{\beta} \in \mathbb{R}^3$, for every $\boldsymbol{\eta} \in \mathbb{R}_*^3 := \mathbb{R}^3 \setminus \{\mathbf{0}\}$. It follows that K is convex.

Now, let us prove that K is coercive. Assume that $\boldsymbol{\beta} \in S^2 := \{\mathbf{x} \in \mathbb{R}^3 \mid \|\mathbf{x}\|_2 = 1\}$. To be precise, we will show the following: for every $M > 0$, there exists a number $R = R(M) > 0$, independent of $\boldsymbol{\beta}$, such that if s is any number satisfying $s > R$, then it follows that

$$K(s\boldsymbol{\beta}) > M,$$

where, of course,

$$K(s\boldsymbol{\beta}) := \langle \exp(s\boldsymbol{\beta} \cdot \mathbf{m}_{\mathcal{J}}) \rangle_{\mathcal{J}} - s\boldsymbol{\beta} \cdot \boldsymbol{\rho}_{\mathcal{J},*}(x).$$

Fix a vector $\boldsymbol{\beta} \in S^2$, and suppose that for this vector, for some $j_0 \in \mathcal{J}$,

$$\boldsymbol{\beta} \cdot \mathbf{m}_{\mathcal{J}}(j_0) > 0.$$

Then,

$$K(s\boldsymbol{\beta}) \geq h_v \exp(s\boldsymbol{\beta} \cdot \mathbf{m}_{\mathcal{J}}(j_0)) - s\boldsymbol{\beta} \cdot \boldsymbol{\rho}_{\mathcal{J},*}(x) \xrightarrow{s \rightarrow +\infty} +\infty.$$

In fact, by continuity, there is an $\varepsilon_{\boldsymbol{\beta}} > 0$ and a constant $C_1 > 0$, such that, if $\boldsymbol{\alpha} \in B(\boldsymbol{\beta}, \varepsilon_{\boldsymbol{\beta}})$ then,

$$\boldsymbol{\alpha} \cdot \mathbf{m}_{\mathcal{J}}(j_0) \geq C_1 > 0,$$

so that

$$K(s\alpha) \geq h_v \exp(sC_1) - s\alpha \cdot \rho_{\mathcal{J},*}(x) \xrightarrow{s \rightarrow +\infty} +\infty.$$

To sum up, given any $M > 0$, there is an $R_\beta = R_\beta(M)$, such that, if $s > R_\beta$, then

$$K(s\alpha) > M,$$

for all $\alpha \in B(\beta, \varepsilon_\beta)$.

Again, fix a vector $\beta \in S^2$, and suppose that for this vector, for all $j \in \mathcal{J}$,

$$\beta \cdot m_{\mathcal{J}}(j) \leq 0.$$

Now, we know that

$$\beta \cdot m_{\mathcal{J}}(j) = 0, \quad \forall j \in \mathcal{J} \quad \Longleftrightarrow \quad \beta = \mathbf{0},$$

but $\beta \neq \mathbf{0}$, since $\|\beta\|_2 = 1$. Therefore, for some $j_1 \in \mathcal{J}$,

$$\beta \cdot m_{\mathcal{J}}(j_1) < 0.$$

It follows from (2.40) that

$$\beta \cdot \rho_{\mathcal{J},*}(x) < 0.$$

Consequently,

$$K(s\beta) \geq h_v \exp(s\beta \cdot m(j_1)) - s\beta \cdot \rho_{\mathcal{J},*}(x) \xrightarrow{s \rightarrow +\infty} +\infty.$$

In fact, there is an $\varepsilon_\beta > 0$, such that, if $\alpha \in B(\beta, \varepsilon_\beta)$, then

$$\alpha \cdot m_{\mathcal{J}}(j_1) \leq C_2 < 0,$$

and

$$\alpha \cdot \rho_{\mathcal{J},*}(x) \leq C_3 < 0.$$

Hence,

$$K(s\alpha) \geq -sC_3 \xrightarrow{s \rightarrow +\infty} +\infty.$$

To sum up this second case, given any $M > 0$, there is an $R_\beta = R_\beta(M)$, such that, if $s > R_\beta$, then

$$K(s\alpha) > M,$$

for all $\alpha \in B(\beta, \varepsilon_\beta)$.

Observe that

$$S^2 \subset \bigcup_{\beta \in S^2} B(\beta, \varepsilon_\beta),$$

the object on the right being an open cover. Since S^2 is compact, there is a finite sub-cover. In other words, there are finite points $\beta_1, \dots, \beta_N \in S^2$ such that

$$S^2 \subset \bigcup_{i=1}^N B(\beta_i, \varepsilon_{\beta_i}).$$

Let $M > 0$ be given. Set

$$R(M) = \max_{i=1}^N R_{\beta_i}(M).$$

If $s > R(M)$, it follows that

$$K(s\beta) > M,$$

for all $\beta \in S^2$. Since $R(M)$ is independent of β , K is coercive.

Since K is smooth, strictly convex, and coercive, it has a unique minimizer, $\beta_* \in \mathbb{R}^3$ that satisfies

$$\nabla K(\beta_*) = \langle \exp(\beta_* \cdot m_{\mathcal{J}}) m_{\mathcal{J}} \rangle_{\mathcal{J}} - \rho_{\mathcal{J},*}(x) = \mathbf{0}.$$

The proof is complete. ■

Remark 2.2.3. *The compactness argument is required above because, without it, it is not clear that*

$$R(M) = \sup_{\beta \in S^2} R_\beta$$

is finite, which is needed for our definition of coercivity.

Lemma 2.2.4. *Suppose that $\Omega = [a, b] \subset \mathbb{R}$ and*

$$f_{\mathcal{J}} : \Omega \times \mathcal{J} \rightarrow [0, \infty)$$

is a given Ω -periodic discrete-velocity distribution. Let

$$\rho_{\mathcal{J}} = [n_{\mathcal{J}}, n_{\mathcal{J}} u_{\mathcal{J}}, E_{\mathcal{J}}]^\top,$$

be the vector comprised of the first three discrete-velocity moments of $f_{\mathcal{J}}$, that is,

$$\boldsymbol{\rho}_{\mathcal{J}}(x) = \begin{bmatrix} n_{\mathcal{J}}(x) \\ n_{\mathcal{J}}(x) \mathbf{u}_{\mathcal{J}}(x) \\ E_{\mathcal{J}}(x) \end{bmatrix} = \langle f_{\mathcal{J}}(x, \cdot) \mathbf{m}_{\mathcal{J}} \rangle_{\mathcal{J}}, \quad \forall x \in \Omega. \quad (2.41)$$

Fix $x \in [a, b]$, and define

$$\mathcal{M}_{\mathcal{J},f}(x, j) = \exp(\boldsymbol{\alpha}_{\mathcal{J},f}(x) \cdot \mathbf{m}_{\mathcal{J}}(j)), \quad \forall j \in \mathcal{J},$$

where $\boldsymbol{\alpha}_{\mathcal{J},f}(x) \in \mathbb{R}^3$ is the solution to

$$\langle \exp(\boldsymbol{\alpha}_{\mathcal{J},f}(x) \cdot \mathbf{m}_{\mathcal{J}}) \mathbf{m}_{\mathcal{J}} \rangle_{\mathcal{J}} = \boldsymbol{\rho}_{\mathcal{J}}(x),$$

which is guaranteed to exist, uniquely, by Lemma 2.2.2. Then

$$0 = \sum_{j \in \mathcal{J}} h_v (\mathcal{M}_{\mathcal{J},f}(x, j) - f_{\mathcal{J}}(x, j)), \quad (2.42)$$

$$0 = \sum_{j \in \mathcal{J}} h_v v_j (\mathcal{M}_{\mathcal{J},f}(x, j) - f_{\mathcal{J}}(x, j)), \quad (2.43)$$

$$0 = \sum_{j \in \mathcal{J}} h_v \frac{1}{2} |v_j|^2 (\mathcal{M}_{\mathcal{J},f}(x, j) - f_{\mathcal{J}}(x, j)). \quad (2.44)$$

This may be expressed succinctly as

$$\langle \mathbf{m}_{\mathcal{J}} (\mathcal{M}_{\mathcal{J},f}(x, \cdot) - f_{\mathcal{J}}(x, \cdot)) \rangle_{\mathcal{J}} = \mathbf{0}. \quad (2.45)$$

Furthermore, the function $\ln(\mathcal{M}_{\mathcal{J},f})$ is a discrete-velocity collision invariant, that is

$$\langle \ln(\mathcal{M}_{\mathcal{J},f}) (\mathcal{M}_{\mathcal{J},f} - f_{\mathcal{J}}) \rangle_{\mathcal{J}} = 0. \quad (2.46)$$

Proof. Identities (2.42) – (2.44) follow immediately from the definition. Since

$$\ln(\mathcal{M}_{\mathcal{J},f}(x, j)) = \ln(\exp(\boldsymbol{\alpha}_{\mathcal{J},f}(x) \cdot \mathbf{m}_{\mathcal{J}}(j))) = \boldsymbol{\alpha}_{\mathcal{J},f}(x) \cdot \mathbf{m}_{\mathcal{J}}(j),$$

identity (2.46) also follows. ■

Theorem 2.2.5. Define, for all $f : \Omega \times \mathcal{J} \rightarrow [0, \infty)$, the discrete-velocity entropy functional

$$H_{\mathcal{J}}[f](\cdot) := \langle f \ln(f) \rangle_{\mathcal{J}} : \Omega \rightarrow [0, \infty).$$

Let $\mathcal{A}_{\mathcal{J},*}$ be the admissible class of discrete-velocity distribution functions, subordinate to a given discrete-velocity distribution $f_{\mathcal{J},*} : \Omega \times \mathcal{J} \rightarrow [0, \infty)$, so that $\mathcal{A}_{\mathcal{J},*}$ is non-empty. Then, there is a unique function $\mathcal{M}_{\mathcal{J},*} \in \mathcal{A}_{\mathcal{J},*}$ such that

$$H_{\mathcal{J}}[\mathcal{M}_{\mathcal{J},*}](x) \leq H_{\mathcal{J}}[g](x), \quad (2.47)$$

for all $g \in \mathcal{A}_{\mathcal{J},*}$, for all $x \in \Omega$. Furthermore, at every point $x \in \Omega$, there is a unique vector $\alpha_{\mathcal{J},*}(x) \in \mathbb{R}^3$ such that

$$\mathcal{M}_{\mathcal{J},*}(x, j) = \exp(\alpha_{\mathcal{J},*}(x) \cdot m_{\mathcal{J}}(j)), \quad \forall j \in \mathcal{J}.$$

Proof. Fix $x \in [a, b]$. Define

$$\mathcal{M}_{\mathcal{J},*}(x, j) = \exp(\alpha_{\mathcal{J},*}(x) \cdot m_{\mathcal{J}}(j)), \quad \forall j \in \mathcal{J},$$

where $\alpha_{\mathcal{J},*}(x) \in \mathbb{R}^3$ is the solution to

$$\langle \exp(\alpha_{\mathcal{J},*}(x) \cdot m_{\mathcal{J}}) m_{\mathcal{J}} \rangle_{\mathcal{J}} = \rho_{\mathcal{J},*}(x),$$

which is guaranteed to exist, uniquely, by Lemma 2.2.2. This definition shows that the function $\mathcal{M}_{\mathcal{J},*} \in \mathcal{A}_{\mathcal{J},*}$. Thus, to conclude, we need only show that

$$H_{\mathcal{J}}[\mathcal{M}_{\mathcal{J},*}](x) \leq H_{\mathcal{J}}[g](x),$$

for all $g \in \mathcal{A}_{\mathcal{J},*}$. The proof follows that of the continuous-velocity case.

Let $g \in \mathcal{A}_{\mathcal{J},*}$ be arbitrary. We may write $\mathcal{M}_{\mathcal{J},*} = \mathcal{M}_{\mathcal{J},g}$, since g gives rise to the same macroscopic, discrete-velocity fields as $f_{\mathcal{J},*}$. Define $h(s) := s \ln(s)$. This function h is convex on $(0, \infty)$, and it follows that

$$\begin{aligned} h(g) &\geq h(\mathcal{M}_{\mathcal{J},g}) + h'(\mathcal{M}_{\mathcal{J},g})(g - \mathcal{M}_{\mathcal{J},g}) \\ &= h(\mathcal{M}_{\mathcal{J},g}) + \ln(\mathcal{M}_{\mathcal{J},g})(g - \mathcal{M}_{\mathcal{J},g}) + g - \mathcal{M}_{\mathcal{J},g}. \end{aligned}$$

By Lemma 2.2.4,

$$\langle h(g) \rangle_{\mathcal{J}} \geq \langle h(\mathcal{M}_{\mathcal{J},g}) \rangle_{\mathcal{J}} + \langle \ln(\mathcal{M}_{\mathcal{J},g})(g - \mathcal{M}_{\mathcal{J},g}) \rangle_{\mathcal{J}} + \langle g - \mathcal{M}_{\mathcal{J},g} \rangle_{\mathcal{J}} = \langle h(\mathcal{M}_{\mathcal{J},g}) \rangle_{\mathcal{J}}.$$

This shows that, point-wise in Ω ,

$$H_{\mathcal{J}}[g] \geq H_{\mathcal{J}}[\mathcal{M}_{\mathcal{J},g}] = H_{\mathcal{J}}[\mathcal{M}_{\mathcal{J},*}],$$

and the result is proven. ■

Definition 2.2.6. *The 1+1-dimensional discrete-velocity BGK model is defined as follows: given an $[a, b]$ -periodic function $f_0 : [a, b] \times \mathcal{J} \rightarrow (0, \infty)$, find $f : [a, b] \times \mathcal{J} \times [0, T] \rightarrow (0, \infty)$ such that $f(\cdot, \cdot, 0) = f_0$ and*

$$\partial_t f(x, j, t) + v_j \partial_x f(x, j, t) = \lambda (\mathcal{M}_{\mathcal{J},f}(x, j, t) - f(x, j, t)), \quad \forall j \in \mathcal{J}, \quad \forall x \in [a, b], \quad (2.48)$$

where

$$\mathcal{M}_{\mathcal{J},f}(x, j, t) = \exp(\alpha_{\mathcal{J},f}(x, t) \cdot \mathbf{m}_{\mathcal{J}}(j)), \quad \forall j \in \mathcal{J},$$

and $\alpha_{\mathcal{J},f}(x, t) \in \mathbb{R}^3$ is the solution to

$$\langle \exp(\alpha_{\mathcal{J},f}(x, t) \cdot \mathbf{m}_{\mathcal{J}}) \mathbf{m}_{\mathcal{J}} \rangle_{\mathcal{J}} = \langle f(x, \cdot, t) \mathbf{m}_{\mathcal{J}} \rangle_{\mathcal{J}}.$$

We have the following result concerning the solutions of the discrete-velocity model.

Theorem 2.2.7. *Suppose that $\Omega = [a, b] \subset \mathbb{R}$ and*

$$f : \Omega \times \mathcal{J} \times [0, \infty) \rightarrow [0, \infty)$$

is a given Ω -periodic solution to the discrete-velocity BGK model. Define, for all $x \in [a, b]$ and $t \geq 0$, the discrete-velocity entropy density via

$$H_{\mathcal{J}}[f](x, t) := \langle f(x, \cdot, t) \ln(f(x, \cdot, t)) \rangle_{\mathcal{J}}.$$

Then, for all $x \in [a, b]$ and $t \geq 0$,

$$\frac{\partial}{\partial t} \langle \mathbf{m}_{\mathcal{J}} f(x, \cdot, t) \rangle_{\mathcal{J}} + \frac{\partial}{\partial x} \langle v(\cdot) \mathbf{m}_{\mathcal{J}} f(x, \cdot, t) \rangle_{\mathcal{J}} = \mathbf{0}; \quad (2.49)$$

and

$$\frac{\partial}{\partial t} \langle f(x, \cdot, t) \ln(f(x, \cdot, t)) \rangle_{\mathcal{J}} + \frac{\partial}{\partial x} \langle v(\cdot) f(x, \cdot, t) \ln(f(x, \cdot, t)) \rangle_{\mathcal{J}} \leq 0. \quad (2.50)$$

Consequently,

$$\frac{d}{dt} \int_{\Omega} H_{\mathcal{J}}[f](x, t) dx \leq 0. \quad (2.51)$$

Proof. The proofs are analogous to the continuous-velocity cases. Equation (2.49) follows directly from (2.45). To prove inequality (2.50), observe that

$$\begin{aligned} \frac{\partial}{\partial t} H_{\mathcal{J}}[f] &= \left\langle \frac{\partial f}{\partial t} (\ln(f) + 1) \right\rangle_{\mathcal{J}} \\ &= \left\langle \frac{\partial f}{\partial t} \ln(f) \right\rangle_{\mathcal{J}} + \frac{\partial}{\partial t} \langle f \rangle_{\mathcal{J}} \\ &= \frac{1}{\tau} \langle (\mathcal{M}_{\mathcal{J},f} - f) \ln(f) \rangle_{\mathcal{J}} - \langle v \partial_x f \ln(f) \rangle_{\mathcal{J}} - \partial_x \langle v f \rangle_{\mathcal{J}} \\ &= \frac{1}{\tau} \langle \ln(f) (\mathcal{M}_{\mathcal{J},f} - f) \rangle_{\mathcal{J}} - \partial_x \langle v f \ln(f) \rangle_{\mathcal{J}} \\ &= \frac{1}{\tau} \langle (\ln(f) - \ln(\mathcal{M}_{\mathcal{J},f})) (\mathcal{M}_{\mathcal{J},f} - f) \rangle_{\mathcal{J}} - \partial_x \langle v f \ln(f) \rangle_{\mathcal{J}} \\ &= -\frac{1}{\tau} \langle (\ln(\mathcal{M}_{\mathcal{J},f}) - \ln(f)) (\mathcal{M}_{\mathcal{J},f} - f) \rangle_{\mathcal{J}} - \partial_x \langle v f \ln(f) \rangle_{\mathcal{J}} \\ &\leq -\partial_x \langle v f \ln(f) \rangle_{\mathcal{J}}, \end{aligned}$$

where we used the fact that

$$(\ln(x) - \ln(y))(x - y) \geq 0, \quad \forall x, y \in (0, \infty).$$

Finally, using the Ω -periodicity of f and integrating over Ω , gives (2.51). ■

2.2.2 Space-Homogeneous Discrete-Velocity Model

Similar to the continuous velocity case, we can make the following definition:

Definition 2.2.8. *The **space-homogeneous, 1-dimensional discrete-velocity BGK model** is defined as follows: given the function $f_0 : \mathcal{J} \rightarrow (0, \infty)$, find $f : \mathcal{J} \times [0, T] \rightarrow (0, \infty)$ such that $f(\cdot, 0) = f_0$ and*

$$\partial_t f(j, t) = \lambda (\mathcal{M}_{\mathcal{J},f}(j, t) - f(j, t)), \quad \forall j \in \mathcal{J}, \quad (2.52)$$

where

$$\mathcal{M}_{\mathcal{J},f}(\mathbf{j}, t) = \exp(\boldsymbol{\alpha}_{\mathcal{J},f}(t) \cdot \mathbf{m}_{\mathcal{J}}(\mathbf{j})), \quad \forall \mathbf{j} \in \mathcal{J}, \quad (2.53)$$

and $\boldsymbol{\alpha}_{\mathcal{J},f}(t) \in \mathbb{R}^3$ is the solution to

$$\langle \exp(\boldsymbol{\alpha}_{\mathcal{J},f}(t) \cdot \mathbf{m}_{\mathcal{J}}) \mathbf{m}_{\mathcal{J}} \rangle_{\mathcal{J}} = \langle f(\cdot, t) \mathbf{m}_{\mathcal{J}} \rangle_{\mathcal{J}}. \quad (2.54)$$

Theorem 2.2.9. *Suppose that*

$$f : \mathcal{J} \times [0, \infty) \rightarrow (0, \infty)$$

is a given solution to the space-homogeneous discrete-velocity BGK model. Define, for all $t \geq 0$, the discrete-velocity entropy via

$$H_{\mathcal{J}}[f](t) := \langle f(\cdot, t) \ln(f(\cdot, t)) \rangle_{\mathcal{J}}.$$

Then, for all $t \geq 0$,

$$\frac{\partial}{\partial t} \langle \mathbf{m}_{\mathcal{J}} f(\cdot, t) \rangle_{\mathcal{J}} = \mathbf{0}; \quad (2.55)$$

and

$$\frac{\partial}{\partial t} \langle f(\cdot, t) \ln(f(\cdot, t)) \rangle_{\mathcal{J}} \leq 0. \quad (2.56)$$

Consequently,

$$\frac{\partial}{\partial t} \mathcal{M}_{\mathcal{J},f}(\mathbf{j}, t) = 0, \quad \forall \mathbf{j} \in \mathcal{J}. \quad (2.57)$$

Proof. The proofs are again analogous to the continuous-velocity cases. Equation (2.55) follows directly from (2.45). To prove inequality (2.56), observe that

$$\begin{aligned} \frac{\partial}{\partial t} H_{\mathcal{J}}[f] &= \left\langle \frac{\partial f}{\partial t} (\ln(f) + 1) \right\rangle_{\mathcal{J}} \\ &= \left\langle \frac{\partial f}{\partial t} \ln(f) \right\rangle_{\mathcal{J}} + \frac{\partial}{\partial t} \langle f \rangle_{\mathcal{J}} \\ &= \frac{1}{\tau} \langle (\mathcal{M}_{\mathcal{J},f} - f) \ln(f) \rangle_{\mathcal{J}} \\ &= \frac{1}{\tau} \langle \ln(f) (\mathcal{M}_{\mathcal{J},f} - f) \rangle_{\mathcal{J}} \\ &= \frac{1}{\tau} \langle (\ln(f) - \ln(\mathcal{M}_{\mathcal{J},f})) (\mathcal{M}_{\mathcal{J},f} - f) \rangle_{\mathcal{J}} \\ &= -\frac{1}{\tau} \langle (\ln(\mathcal{M}_{\mathcal{J},f}) - \ln(f)) (\mathcal{M}_{\mathcal{J},f} - f) \rangle_{\mathcal{J}} \end{aligned}$$

$$\leq 0,$$

where we used the fact that

$$(\ln(x) - \ln(y))(x - y) \geq 0, \quad \forall x, y \in (0, \infty).$$

Finally, (2.57) follows from (2.53) and (2.54), using (2.55). ■

Chapter 3

Numerical Approximations for the Single Species Case

In this chapter, we consider the numerical approximation of the BGK model for the $d = 1$ case. Therefore, phase space is two-dimensional, with one dimension for physical space and one for velocity space. It is not necessary to equate the dimensions of velocity and physical space, but this is a common practice, and, for the purpose of describing the numerical methods, one does not lose much generality using such simplifying assumptions. We call the present case the $1 \times 1v$ case. The methods that we describe in this section are scalable, meaning that, as more phase-space dimensions are added, the methods themselves change only slightly. The biggest obstacle for high-dimensional numerical simulation is the added number of degrees of freedom that accompany an increase in phase-space dimensions. Of course, this increase can be significant, since real-world phase space has 6 dimensions in the model.

3.1 Finite Volume Space and Velocity Discretization

Since the Vlasov-BGK equation is a nonlinear conservation-like law, shocks (discontinuities) and rarefaction waves can form and propagate in the solution. Thus, the integro-differential equation does not necessarily hold in the classical (strong) sense. In this case, finite volume methods, which are based on the integral form of the differential equation, are typically more appropriate and simpler to use.

First, we truncate the velocity space so that it is finite in size: $V = [-v_{\max}, v_{\max}]$, where $v_{\max} > 0$. This limits the range of velocities that may be approximated, but, as we will see, does not generally affect the accuracy of approximation as long as $v_{\max}$ is chosen sufficiently

large. We note that it is not necessary to make a symmetric truncation about zero velocity, and sometimes it is not advantageous to do so. This is only done for simplicity of presentation. In similar fashion, let us assume that $\Omega = [-L, L]$, with $L > 0$.

Let us define

$$h_x := \frac{2L}{N_x} \quad \text{and} \quad h_v := \frac{2v_{\max}}{N_v},$$

and then set

$$x_\ell := -L + (\ell - 1/2)h_x \quad \text{and} \quad v_\ell := -v_{\max} + (\ell - 1/2)h_v,$$

where ℓ can take integer and half-integer values. To discretize phase space, we break the rectangular $1 \times 1v$ domain, $\Omega \times V$, into a two dimensional grid of cells with finite volume, and approximate the cell average of the function. $\Omega \times V \subset \mathbb{R}^2$ can be written as the union of the cells:

$$\Omega \times V = \bigcup_{i,j=1}^{N_x, N_v} C_{i,j} = \bigcup_{i,j=1}^{N_x, N_v} C_i^x \times C_j^v = \bigcup_{i,j=1}^{N_x, N_v} \left[x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}} \right] \times \left[v_{j-\frac{1}{2}}, v_{j+\frac{1}{2}} \right], \quad (3.1)$$

with the point (x_i, v_j) at the center of cell $C_{i,j} = C_i^x \times C_j^v$. Next, we define

$$\bar{f}_{i,j}(t) := \bar{f}(x_i, v_j, t) := \frac{1}{|C_i^x \times C_j^v|} \int_{C_i^x \times C_j^v} f(x, v, t) dx dv. \quad (3.2)$$

The integral form of the Vlasov-BGK equation is

$$\partial_t \bar{f}_{i,j}(t) + \frac{1}{|C_{i,j}|} \int_{C_{i,j}} (\partial_x(vf) + \partial_v(a(x, t)f)) dx dv = \frac{\lambda}{|C_{i,j}|} \int_{C_{i,j}} M_f(x, v, t) dx dv - \lambda \bar{f}_{i,j}(t), \quad (3.3)$$

where we have assumed that the acceleration of particles due to the external field, a , is independent of velocity. Let us define the flux function, $\mathbf{F}$, via

$$\mathbf{F}(x, v, t) := [vf(x, v, t), a(x, t)f(x, v, t)]^\top := [F(x, v, t), G(x, v, t)]^\top,$$

where

$$F(x, v, t) := vf(x, v, t), \quad G(x, v, t) := a(x, t)f(x, v, t).$$

Define

$$\bar{M}_{i,j}(t) := \frac{1}{|C_{i,j}|} \int_{C_{i,j}} M_f(x, v, t) dx dv.$$

Applying the Divergence Theorem,

$$\begin{aligned} \partial_t \bar{f}_{i,j}(t) + \lambda \bar{f}_{i,j}(t) - \lambda \bar{M}_{i,j}(t) = & -\frac{1}{h_x h_v} \int_{C_j^v} (F(x_{i+1/2}, v, t) - F(x_{i-1/2}, v, t)) dv \\ & - \frac{1}{h_x h_v} \int_{C_i^x} (G(x, v_{j+1/2}, t) - G(x, v_{j-1/2}, t)) dx. \end{aligned} \quad (3.4)$$

Now, (3.4) is exact. To gain a practical numerical method, we must make some approximations. Let us first approximate the flux integrals:

$$\frac{1}{h_x h_v} \int_{C_j^v} (F(x_{i+1/2}, v, t) - F(x_{i-1/2}, v, t)) dv \approx \frac{F(x_{i+1/2}, v_j, t) - F(x_{i-1/2}, v_j, t)}{h_x}, \quad (3.5)$$

$$\frac{1}{h_x h_v} \int_{C_i^x} (G(x, v_{j+1/2}, t) - G(x, v_{j-1/2}, t)) dx \approx \frac{G(x_i, v_{j+1/2}, t) - G(x_i, v_{j-1/2}, t)}{h_v}. \quad (3.6)$$

To approximate the density function on the cell $C_{i,j}$, we use piecewise linear reconstructions:

$$p_{i,j}(x, v, t) = \bar{f}_{i,j}(t) + \sigma_{i,j}^x(x - x_i) + \sigma_{i,j}^v(v - v_j),$$

with the θ -minmod slope limiter,

$$\sigma_{i,j}^x = \text{minmod} \left(\left(\frac{\bar{f}_{i+1,j} - \bar{f}_{i-1,j}}{2h_x} \right), \theta^x \left(\frac{\bar{f}_{i,j} - \bar{f}_{i-1,j}}{h_x} \right), \theta^x \left(\frac{\bar{f}_{i+1,j} - \bar{f}_{i,j}}{h_x} \right) \right), \quad (3.7)$$

$$\sigma_{i,j}^v = \text{minmod} \left(\left(\frac{\bar{f}_{i,j+1} - \bar{f}_{i,j-1}}{2h_v} \right), \theta^v \left(\frac{\bar{f}_{i,j} - \bar{f}_{i,j-1}}{h_v} \right), \theta^v \left(\frac{\bar{f}_{i,j+1} - \bar{f}_{i,j}}{h_v} \right) \right), \quad (3.8)$$

where $\theta^x, \theta^v \in [1, 2]$. Note that for three *real* arguments, the minmod function is defined by

$$\text{minmod}(x, y, z) := \begin{cases} \min\{x, y, z\} & \text{if } x, y, z \geq 0 \\ \max\{x, y, z\} & \text{if } x, y, z \leq 0 \\ 0 & \text{otherwise} \end{cases}$$

The reconstructions make the density approximation multi-valued at the cell edges, a fact that we use to our advantage in the numerical approximation. Using the reconstructions, we define the density at the midpoints of each cell edge as follows:

$$f_{i+1/2,j}^-(t) := p_{i,j}(x_{i+1/2}, v_j, t), \quad f_{i+1/2,j}^+(t) := p_{i+1,j}(x_{i+1/2}, v_j, t),$$

and, likewise,

$$f_{i,j+1/2}^-(t) := p_{i,j}(x_i, v_{j+1/2}, t), \quad f_{i,j+1/2}^+(t) := p_{i,j+1}(x_i, v_{j+1/2}, t).$$

Next, the exact fluxes are replaced by numerical fluxes of the form

$$F(x_{i+1/2}, v_j, t) \approx \tilde{F}(f_{i+1/2,j}^+(t), f_{i+1/2,j}^-(t)) =: \hat{F}_{i+1/2,j}(t)$$

and

$$G(x_i, v_{j+1/2}, t) \approx \tilde{G}(f_{i,j+1/2}^+(t), f_{i,j+1/2}^-(t)) =: \hat{G}_{i,j+1/2}(t).$$

We use a simple upwind strategy to construct the numerical fluxes:

$$\hat{F}_{i+1/2,j}(t) = \tilde{F}(f_{i+1/2,j}^+(t), f_{i+1/2,j}^-(t)) = \begin{cases} v_j f_{i+1/2,j}^-(t) & \text{if } v_j \geq 0 \\ v_j f_{i+1/2,j}^+(t) & \text{if } v_j < 0 \end{cases}, \quad (3.9)$$

$$\hat{G}_{i,j+1/2}(t) = \tilde{G}(f_{i,j+1/2}^+(t), f_{i,j+1/2}^-(t)) = \begin{cases} a(x_i, t) f_{i,j+1/2}^-(t) & \text{if } a(x_i, t) \geq 0 \\ a(x_i, t) f_{i,j+1/2}^+(t) & \text{if } a(x_i, t) < 0 \end{cases}. \quad (3.10)$$

We can write the approximation scheme to this point as follows:

$$\begin{aligned} \partial_t \bar{f}_{i,j}(t) + \lambda \bar{f}_{i,j}(t) - \lambda \bar{M}_{i,j}(t) = & - \frac{\hat{F}_{i+1/2,j}(t) - \hat{F}_{i-1/2,j}(t)}{h_x} \\ & - \frac{\hat{G}_{i,j+1/2}(t) - \hat{G}_{i,j-1/2}(t)}{h_v} + \tilde{E}_{i,j}(t), \end{aligned} \quad (3.11)$$

where $\tilde{E}_{i,j}(t)$ is a local truncation (approximation) error.

Finally, to complete the spatial discretization of the Vlasov-BGK equation, we need to approximate the Maxwellian. For this we use the following:

$$n_i(t) = h_v \sum_{j=1}^{N_v} \bar{f}_{i,j}(t), \quad (3.12)$$

$$u_i(t) = \frac{h_v}{n_i(t)} \sum_{j=1}^{N_v} v_j \bar{f}_{i,j}(t), \quad (3.13)$$

$$\theta_i(t) = \frac{h_v}{n_i(t)} \sum_{j=1}^{N_v} |v_j|^2 \bar{f}_{i,j}(t) - |u_i(t)|^2, \quad (3.14)$$

$$\bar{M}_{i,j}(t) \approx \hat{M}_{i,j}(t) = \frac{n_i(t)}{(2\pi\theta_i(t))^{1/2}} \exp\left(-\frac{(v_j - u_i(t))^2}{2\theta_i(t)}\right). \quad (3.15)$$

Since we are replacing integrations by midpoint quadratures, we lose the collision invariances that we enjoyed at the continuum level. But, for the purpose of approximation, we will continue to assume that these invariances still exist at the discrete level.

We conclude this section by stating the spatially discrete approximation to the Vlasov-BGK equation:

$$\begin{aligned} d_t \bar{f}_{i,j}(t) + \lambda \bar{f}_{i,j}(t) - \lambda \hat{M}_{i,j}(t) = & -\frac{\hat{F}_{i+1/2,j}(t) - \hat{F}_{i-1/2,j}(t)}{h_x} \\ & - \frac{\hat{G}_{i,j+1/2}(t) - \hat{G}_{i,j-1/2}(t)}{h_v} + \hat{E}_{i,j}(t), \end{aligned} \quad (3.16)$$

where $\hat{E}_{i,j}(t)$ is a local truncation (approximation) error.

3.2 Implicit-Explicit Runge Kutta Time Stepping

In this section we introduce an implicit-explicit (IMEX) Runge Kutta (RK) method for integration in time that was proposed in [24]. The convection part is treated using an explicit method, while the collision part is solved using a diagonally implicit method. We will work with the phase-space continuous problem first, in order to take advantage of the collision invariants, before moving to the phase-space discrete problem. Let us rewrite the original Vlasov-BGK equation as

$$\frac{\partial}{\partial t} f(x, v, t) = T[f](x, v, t) + Q[f](x, v, t)$$

where T represents the transport term, and Q represents the collision term:

$$T[f](x, v, t) := -v \partial_x f(x, v, t) - a(x, t) \partial_v f(x, v, t), \quad (3.17)$$

$$Q[f](x, v, t) := \lambda(M_f(x, v, t) - f(x, v, t)). \quad (3.18)$$

The general ν -stage IMEX Runge Kutta scheme (diagonally implicit) is one of the form

$$y^{(s)} = y^n + \Delta t \sum_{r=1}^{s-1} \tilde{a}_{s,r} T[y^{(r)}] + \Delta t \sum_{r=1}^s a_{s,r} Q[y^{(r)}], \quad s \in \{1, \dots, \nu\}, \quad (3.19)$$

$$y^{n+1} = y^n + \Delta t \sum_{s=1}^{\nu} \tilde{b}_s T[y^{(s)}] + \Delta t \sum_{s=1}^{\nu} b_s Q[y^{(s)}], \quad (3.20)$$

where $\tilde{a}_{s,r}$, $\tilde{b}_s$, $a_{s,r}$, b_s are taken from the following Butcher tables, respectively:

$$\begin{array}{c|c} \tilde{c} & \tilde{A} \\ \hline & \tilde{b}^\top \end{array} \qquad \begin{array}{c|c} c & A \\ \hline & b^\top \end{array}. \quad (3.21)$$

The matrix $\tilde{A}$, for the explicit part, is strictly lower triangular, and A is lower triangular.

To run the algorithm, for each stage, we first calculate

$$\hat{y}^{(s-1)} := y^n + \Delta t \sum_{r=1}^{s-1} \tilde{a}_{s,r} T[y^{(r)}] + \Delta t \sum_{r=1}^{s-1} a_{s,r} Q[y^{(r)}]. \quad (3.22)$$

We then rearrange the terms in the final sum to find an expression for $y^{(s)}$:

$$\begin{aligned} y^{(r)} &= \hat{y}^{(s-1)} + \Delta t a_{s,s} Q[y^{(s)}] \\ &= \hat{y}^{(s-1)} + \Delta t a_{s,s} \lambda (M_{y^{(s)}} - y^{(s)}) \end{aligned} \quad (3.23)$$

$$\iff (1 + \Delta t a_{s,s} \lambda) y^{(s)} = \hat{y}^{(s-1)} + \Delta t a_{s,s} \lambda M_{y^{(s)}} \quad (3.24)$$

$$\iff y^{(s)} = \frac{1}{1 + \Delta t a_{s,s} \lambda} \hat{y}^{(s-1)} + \frac{\Delta t a_{s,s} \lambda}{1 + \Delta t a_{s,s} \lambda} M_{y^{(s)}}. \quad (3.25)$$

Note that we need to compute the Maxwellian $M_{y^{(s)}}$, which involves the current stage. This seems to present an issue. However, using the collision invariances properties in (2.1.2), we can circumvent this. In particular,

$$\begin{aligned} \int_{\mathbb{R}^d} \begin{pmatrix} 1 \\ \mathbf{v} \\ |\mathbf{v}|^2 \end{pmatrix} y^{(s)} d\mathbf{v} &= \int_{\mathbb{R}^d} \begin{pmatrix} 1 \\ \mathbf{v} \\ |\mathbf{v}|^2 \end{pmatrix} \hat{y}^{(s-1)} d\mathbf{v} + \Delta t a_{s,s} \lambda \int_{\mathbb{R}^d} \begin{pmatrix} 1 \\ \mathbf{v} \\ |\mathbf{v}|^2 \end{pmatrix} (M_{y^{(s)}} - y^{(s)}) d\mathbf{v} \\ &= \int_{\mathbb{R}^d} \begin{pmatrix} 1 \\ \mathbf{v} \\ |\mathbf{v}|^2 \end{pmatrix} \hat{y}^{(s-1)} d\mathbf{v}. \end{aligned} \quad (3.26)$$

Since the first, second, and third moments of $\hat{y}^{(s-1)}$ are equal to those of $y^{(s)}$, the two

Maxwellians at these stages are equal:

$$M_{y^{(s)}} = M_{\hat{y}^{(s-1)}}.$$

Therefore, we have the explicit update formula

$$y^{(s)} = \frac{1}{1 + \Delta t a_{s,s} \lambda} \hat{y}^{(s-1)} + \frac{\Delta t a_{s,s} \lambda}{1 + \Delta t a_{s,s} \lambda} M_{\hat{y}^{(s-1)}}. \quad (3.27)$$

With the above expression for $y^{(s)}$, a convex combination of $\hat{y}^{(s-1)}$ and $M_{\hat{y}^{(s-1)}}$, the IMEX-RK scheme can be completed without any complicated inversions.

3.3 Fully Discrete Scheme

When we discretize velocity space, we lose the collision invariances at the discrete level; this is because velocity integrals are replaced with midpoint rule quadrature, in addition to the fact that the velocity space is truncated. It follows that (3.27) breaks down. Nevertheless, we will assume that the phase-space discrete analog to (3.27) holds so that stages may be updated in an explicit fashion.

Suppose that $f_{i,j}^n$ is an approximation of the cell average of the density field over cell $C_{i,j}$ at time $t^n := t^{n-1} + \Delta t$. Similarly, we denote by $f_{i,j}^{(\ell)}$ the ℓ^{th} stage of the IMEX-RK scheme with respect to $f_{i,j}^n$. Define

$$T_{i,j}^\square := -\frac{\hat{F}_{i+1/2,j}^\square - \hat{F}_{i-1/2,j}^\square}{h_x} - \frac{\hat{G}_{i,j+1/2}^\square - \hat{G}_{i,j-1/2}^\square}{h_v}, \quad (3.28)$$

$$Q_{i,j}^\square := \lambda(M_{i,j}^\square - f_{i,j}^\square), \quad (3.29)$$

where the fluxes $\hat{F}_{i+1/2,j}^\square$ and $\hat{G}_{i,j+1/2}^\square$ are computed with respect to the cell-centered approximation $f_{i,j}^\square$ and $\square$ is a stage iteration or a time step index. The fully discrete Maxwellian is computed via

$$n_i^\square = h_v \sum_{j=1}^{N_v} f_{i,j}^\square, \quad (3.30)$$

$$u_i^\square = \frac{h_v}{n_i^\square} \sum_{j=1}^{N_v} v_j f_{i,j}^\square, \quad (3.31)$$

$$\theta_i^\square = \frac{h_\nu}{n_i^\square} \sum_{j=1}^{N_\nu} |v_j|^2 f_{i,j}^\square - |u_i^\square|^2, \quad (3.32)$$

$$M_{i,j}^\square = \frac{n_i^\square}{(2\pi\theta_i^\square)^{1/2}} \exp\left(-\frac{(v_j - u_i^\square)^2}{2\theta_i^\square}\right). \quad (3.33)$$

Then we compute

$$\widehat{f}_{i,j}^{(s-1)} := f_{i,j}^n + \Delta t \sum_{r=1}^{s-1} \widetilde{a}_{s,r} T_{i,j}^{(r)} + \Delta t \sum_{r=1}^{s-1} a_{s,r} Q_{i,j}^{(r)}, \quad (3.34)$$

$$f_{i,j}^{(s)} = \frac{1}{1 + \Delta t a_{s,s} \lambda} \widehat{f}_{i,j}^{(s-1)} + \frac{\Delta t a_{s,s} \lambda}{1 + \Delta t a_{s,s} \lambda} \widehat{M}_{i,j}^{(s-1)}, \quad (3.35)$$

$$f_{i,j} = f_{i,j}^n + \Delta t \sum_{s=1}^{\nu} \widetilde{b}_s T_{i,j}^{(s)} + \Delta t \sum_{s=1}^{\nu} b_s Q_{i,j}^{(s)}, \quad (3.36)$$

where $\widehat{M}_{i,j}^{(s-1)}$ is the discrete Maxwellian computed with respect to the fully discrete approximation $\widehat{f}_{i,j}^{(s-1)}$. This completes the description of the fully discrete scheme.

3.4 Poisson Solver for Vlasov-Poisson-BGK Equation

For the Vlasov-Poisson-BGK equation the acceleration a is determined by an electric field as follows:

$$a(x, t) = -\chi \partial_x \Phi,$$

where $\chi > 0$ is a physical constant and Φ is the electric potential determined via

$$-\partial_{xx} \Phi(x, t) = n(x, t),$$

subject to appropriate boundary conditions.

First let us consider *Dirichlet boundary conditions*:

$$\Phi(x_{1/2} = -L) = \alpha, \quad \Phi(x_{N_x+1/2} = L) = \beta. \quad (3.37)$$

Using the standard three-point stencil approximation, we have, for $i = 2, \dots, N_x - 1$,

$$-\frac{1}{h_x^2} [\Phi_{i+1} - 2\Phi_i + \Phi_{i-1}] = n_i. \quad (3.38)$$

The boundary conditions are specified at the peripheral edges of the domain, $-L = x_{1/2}$ and $L = x_{N_x+1/2}$. As is standard, we use the average of the ghost cell and neighboring interior cell. That is, we use the approximations

$$\frac{\Phi_0 + \Phi_1}{2} = \alpha, \quad \implies \Phi_0 = 2\alpha - \Phi_1, \quad (3.39)$$

and we cancel the term Φ_0 in the stencil approximation. The first equation in the approximation thus becomes

$$-\Phi_2 + 3\Phi_1 = h_x^2 n_1 + 2\alpha. \quad (3.40)$$

Using a similar procedure for right boundary condition, we have

$$\Phi_{N_x+1} = 2\beta - \Phi_{N_x}, \quad \implies -\Phi_{N_x-1} + 3\Phi_{N_x} = h_x^2 n_{N_x} + 2\beta. \quad (3.41)$$

Putting this all into a matrix system, we must solve

$$\begin{pmatrix} 3 & -1 & & & \\ -1 & 2 & -1 & & \\ & -1 & 2 & -1 & \\ & & & \ddots & \\ & & & -1 & 2 & -1 \\ & & & & -1 & 3 \end{pmatrix} \begin{pmatrix} \Phi_1 \\ \Phi_2 \\ \vdots \\ \Phi_{N_x-1} \\ \Phi_{N_x} \end{pmatrix} = \begin{pmatrix} h_x^2 n_1 + 2\alpha \\ h_x^2 n_2 \\ \vdots \\ h_x^2 n_{N_x-1} \\ h_x^2 n_{N_x} + 2\beta \end{pmatrix}. \quad (3.42)$$

This system is symmetric positive definite (SPD) and always has a unique solution.

Next, we turn our attention to the case of *Neumann boundary conditions*:

$$\frac{d\Phi}{dx}(-L) = \alpha, \quad \frac{d\Phi}{dx}(L) = \beta. \quad (3.43)$$

In this case, using a second-order approximation at the boundary, we must solve matrix system

$$\begin{pmatrix} 1 & -1 & & & \\ -1 & 2 & -1 & & \\ & -1 & 2 & -1 & \\ & & & \ddots & \\ & & & -1 & 2 & -1 \\ & & & & -1 & 1 \end{pmatrix} \begin{pmatrix} \Phi_1 \\ \Phi_2 \\ \vdots \\ \Phi_{N_x-1} \\ \Phi_{N_x} \end{pmatrix} = \begin{pmatrix} h_x^2 n_1 - h_x \alpha \\ h_x^2 n_2 \\ \vdots \\ h_x^2 n_{N_x-1} \\ h_x^2 n_{N_x} + h_x \beta \end{pmatrix}. \quad (3.44)$$

This is a symmetric positive semi-definite (SPSD) system. It has a unique mean-zero solution, that is, a solution satisfying

$$h_x \sum_{i=1}^{N_x} \Phi_i = 0, \quad \text{if and only if} \quad h_x \sum_{i=1}^{N_x} n_i = \alpha - \beta.$$

This is the discrete analog of the standard continuous compatibility conditions for unique solvability.

Given $f_{i,j}^\square$, we compute the accompanying macroscopic density,

$$n_i^\square = h_v \sum_{j=1}^{N_v} f_{i,j}^\square,$$

and then the associated electric potential $\Phi_i^\square$, as described above. Once the discrete potential is available, the acceleration may be approximated via

$$a_{i+1/2}^\square = -\chi \frac{\Phi_{i+1}^\square - \Phi_i^\square}{h_x}.$$

3.5 Sample Computations and Accuracy Tests

In this section, we report on several numerical tests showing accuracy of the numerical implementation.

Relaxation Test

This first test is designed to confirm that the BGK operator is calculated correctly for the single species case. We consider, in particular, the space homogeneous case: the IPDE becomes the following IODE:

$$\begin{cases} \frac{df}{dt} = \lambda(M_f - f) & (x, v, t) \in \Omega \times V \times [0, \infty) \\ f(x, v, 0) = f_0(x, v) & (x, v, t) \in \Omega \times V \times \{t = 0\} \end{cases}. \quad (3.45)$$

Recall that as shown in Section 2.1.2, the exact solution is

$$f(x, v, t) = e^{-\lambda t} f(x, v, 0) + (1 - e^{-\lambda t}) M_f(x, v, 0). \quad (3.46)$$

For the first test, we chose a function of the form $f_0(x, v) = b(v)g(x)$, where $b(v)$ is a

compactly supported C^∞ function, with nonzero values on $v \in (-2, 2)$. In particular, we use the function $g(x) = e^{-|x|}$. Consider the function

$$f(x, v, 0) = \begin{cases} e^{-|x|} \frac{5(v^2+4) \exp\left(\frac{5v}{v^2-4}\right)}{(v^2-4)^2 \left(1 + \exp\left(\frac{5v}{v^2-4}\right)\right)^2} & \text{if } v \in (-2, 2) \\ 0 & \text{otherwise} \end{cases}.$$

Observe that

$$n = \int_V f(x, v) dv = g(x) \int_V b(v) dv = g(x) = e^{-|x|}.$$

Note that $b(v)$ is an even function:

$$\begin{aligned} b(-v) &= \frac{5((-v)^2 + 4) \exp\left(\frac{5(-v)}{(-v)^2 - 4}\right)}{((-v)^2 - 4)^2 \left(1 + \exp\left(\frac{5(-v)}{(-v)^2 - 4}\right)\right)^2} \\ &= \frac{5(v^2 + 4) \exp\left(-\frac{5v}{v^2 - 4}\right)}{(v^2 - 4)^2 \left(1 + \exp\left(-\frac{5v}{v^2 - 4}\right)\right)^2} \cdot \frac{\exp\left(\frac{10v}{v^2 - 4}\right)}{\left[\exp\left(\frac{5v}{v^2 - 4}\right)\right]^2} \\ &= \frac{5(v^2 + 4) \exp\left(\frac{5v}{v^2 - 4}\right)}{(v^2 - 4)^2 \left(\exp\left(\frac{5v}{v^2 - 4}\right) + 1\right)^2} \\ &= b(v). \end{aligned}$$

Therefore, the function $h(v) = vb(v)$ is odd. Recall that when integrating an odd function over an interval that is symmetric about the origin, one obtains a zero integral. This means that $u = \frac{1}{n} \int_V v f dv \equiv 0$.

The energy density moment is not as easy to compute analytically, so we have found a numerical approximation:

$$E = \frac{1}{2} \int_V v^2 f dv \approx 0.3713094964845 e^{-|x|}.$$

The code is run with $\lambda = 1$ up to a final time $t_F = 1$. The error is calculated by taking the difference of the true solution minus the computed solution. The computational and true solutions are given in Figures 3.2 and 3.3, respectively. The errors are on the order of 10^{-5} when a mesh of size $N_x = 128$, $N_v = 130$ is used. The plot of the error is given in Figure 3.4.

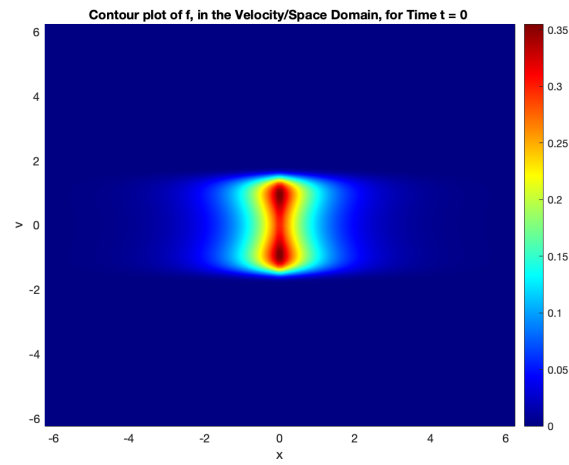


Figure 3.1: Initial condition function ("bump function").

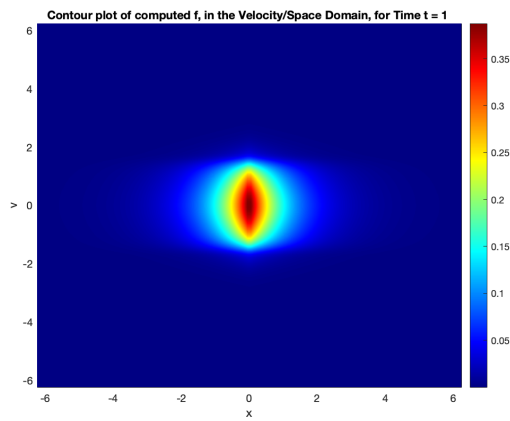


Figure 3.2: Numerical Solution

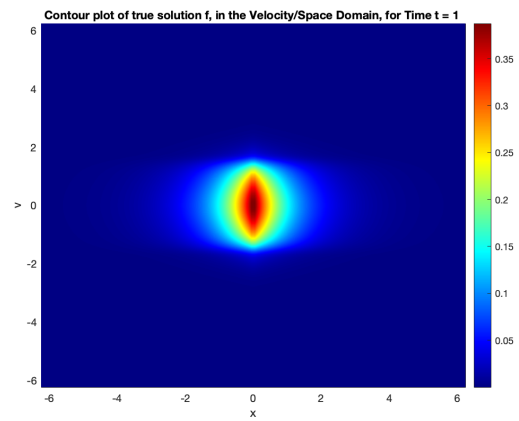


Figure 3.3: Theoretical Solution

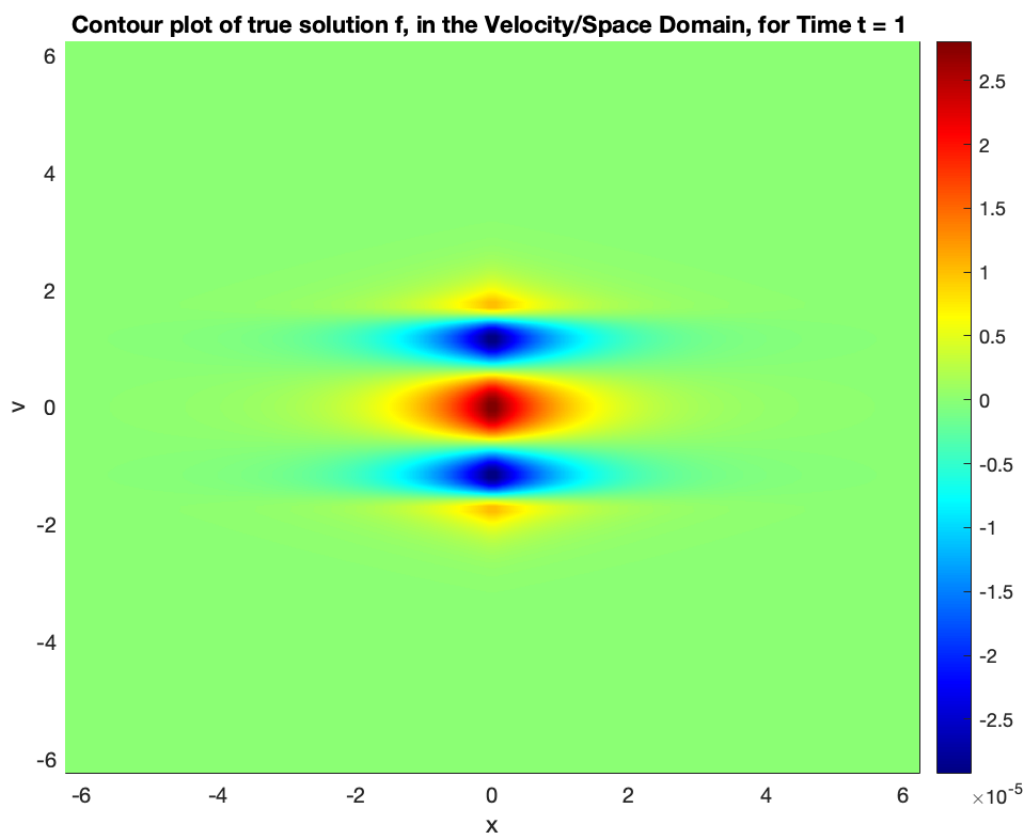


Figure 3.4: Error between theoretical solution and computed solution. This is the theoretical solution minus the code solution. ($N_x = 128$, $N_v = 130$.)

For the second test, we set $f_0(x, v)$ to be the sum of four Gaussians, centered at $(x, v) \in \{(\pm 2, \pm 2)\}$. See Figure 3.5. The trapezoidal rule is used for the initial moments. The code is run with $\lambda = 1$ up to time $t_F = 1$. The error is calculated by taking the difference of the true solution minus the computed solution. The computed and true solutions are given in Figures 3.6 and 3.7 (respectively). The errors are on the order of 10^{-3} , when $N_x = 64$, $N_v = 256$, and are shown in Figure 3.8.

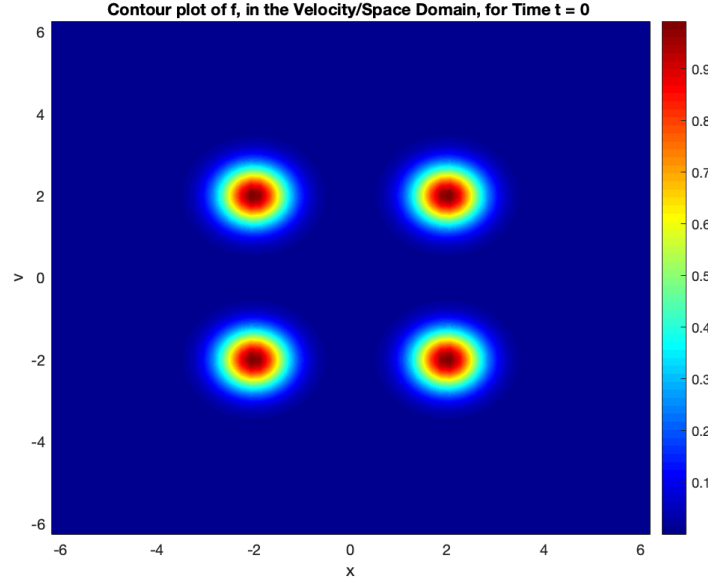


Figure 3.5: Initial condition function (sum of four Gaussians).

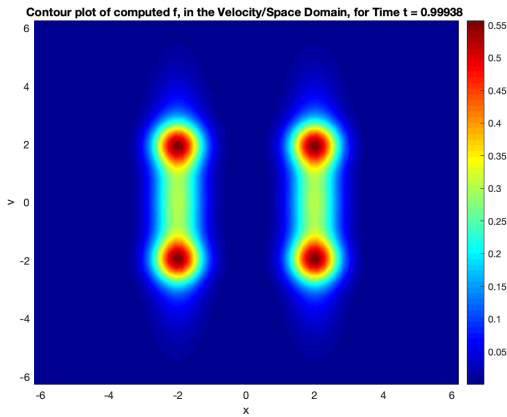


Figure 3.6: Numerical Solution

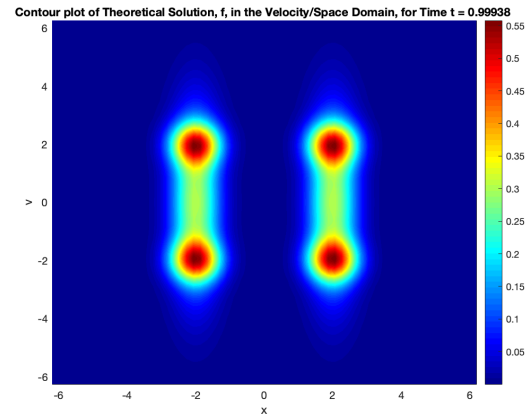


Figure 3.7: Theoretical Solution

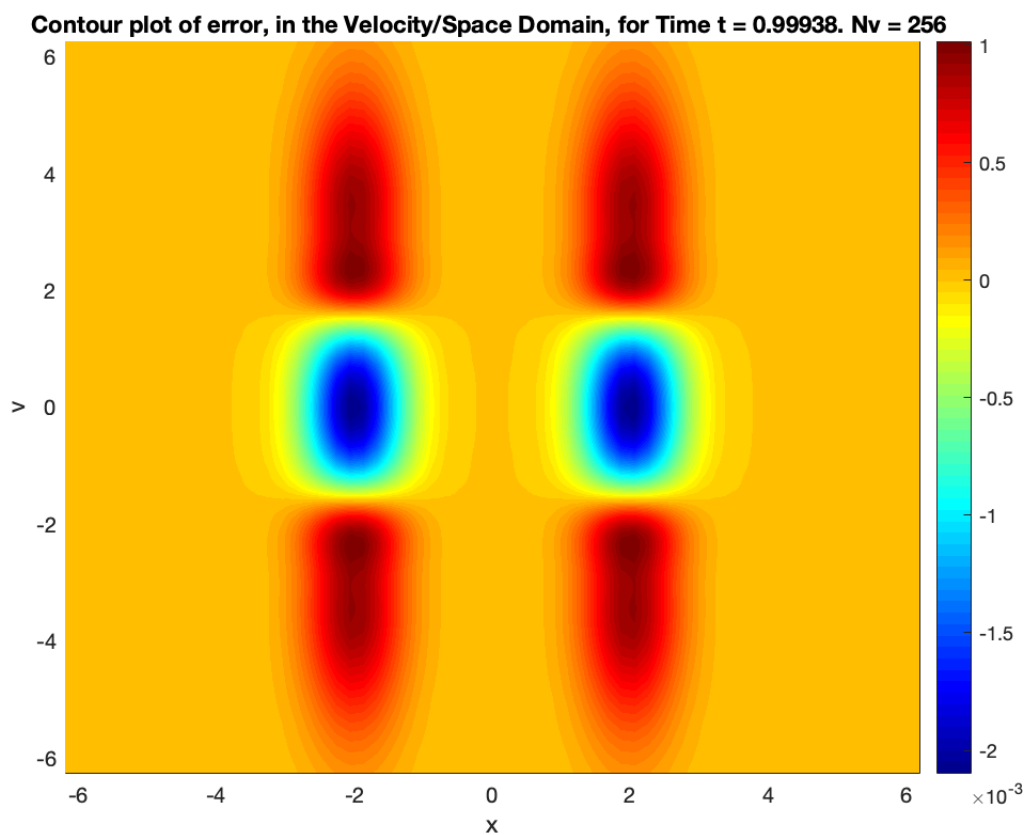


Figure 3.8: Error between theoretical solution and computed solution. This is the theoretical solution minus the code solution.

Sod Shock Tube (Euler Equation Limit of BGK)

The Sod shock tube test is a standard test. Formally, in the collision limit as $\lambda \rightarrow \infty$, the BGK equation is asymptotically equivalent to the Euler Equations. (That is, f converges to the Maxwellian M_f , and the moments n, u, θ follow the Euler equations.) Thus, we may test the ability of the code to solve the Sod shock tube problem, by letting λ be large (or equivalently, letting $\tau = \lambda^{-1}$ be small), and setting the initial conditions appropriately. For the test, we set $\lambda = 10^4$ ($\tau = 10^{-4}$).

The BGK equation in 1 dimension is

$$\frac{\partial f}{\partial t} + v \frac{\partial f}{\partial x} = \lambda(M_f - f). \quad (3.47)$$

To test our IMEX code, we compared the profiles of the number density ($n = \int f dv$), bulk velocity ($u = (\int v f dv)/n$), pressure ($P = n\theta$), and internal energy ($E = \frac{1}{2}\theta$) to the theoretical solution, worked out using the book by Toro [30]. We set the phase-space domain equal to $\Omega \times V = [-0.75, 0.75] \times [-10, 10]$. Since the moments are calculated over $\mathbb{R}^d$, then we must have a function that integrates to approximately the same value, when restricting $\mathbb{R}^d$ to the computational velocity domain (in this case, $[-10, 10] \subset \mathbb{R}$). That is, we must ensure that

$$\int_{\mathbb{R}^d} f(x, v, t) dv \approx \int_V f(x, v, t) dv,$$

where V is the truncated velocity domain. Using $[-10, 10]$ gives a reasonable approximation, as we show.

The setup for the Sod problem is a contact discontinuity separating gases of differing density and temperature, and zero velocity. Thus, for the Sod problem, the initial condition for the particle density is a piecewise Maxwellian with the following values:

$$\begin{pmatrix} n_L \\ u_L \\ \theta_L \end{pmatrix} = \begin{pmatrix} 1.0 \\ 0.0 \\ 1.0 \end{pmatrix}, \quad x \in (-0.75, 0), \quad \begin{pmatrix} n_R \\ u_R \\ \theta_R \end{pmatrix} = \begin{pmatrix} 0.125 \\ 0.0 \\ 0.8 \end{pmatrix}, \quad x \in (0, 0.75). \quad (3.48)$$

That is, the initial condition function (contour plot shown in Figure 3.9) is given by

$$f_0(x, v) = \begin{cases} \frac{1.0}{\sqrt{2\pi(1.0)}} \exp\left(-\frac{|v-0.0|^2}{2(1.0)}\right) & x \in [-0.75, 0], v \in [-10, 10] \\ \frac{0.125}{\sqrt{2\pi(0.8)}} \exp\left(-\frac{|v-0.0|^2}{2(0.8)}\right) & x \in (0, 0.75], v \in [-10, 10] \end{cases}. \quad (3.49)$$

Figure 3.10 shows the moments of the numerical/computed solution at time $t_F = 0.20$. The profiles of the solution seem to follow the correct values (as computed, using the Toro book, [30], for reference).

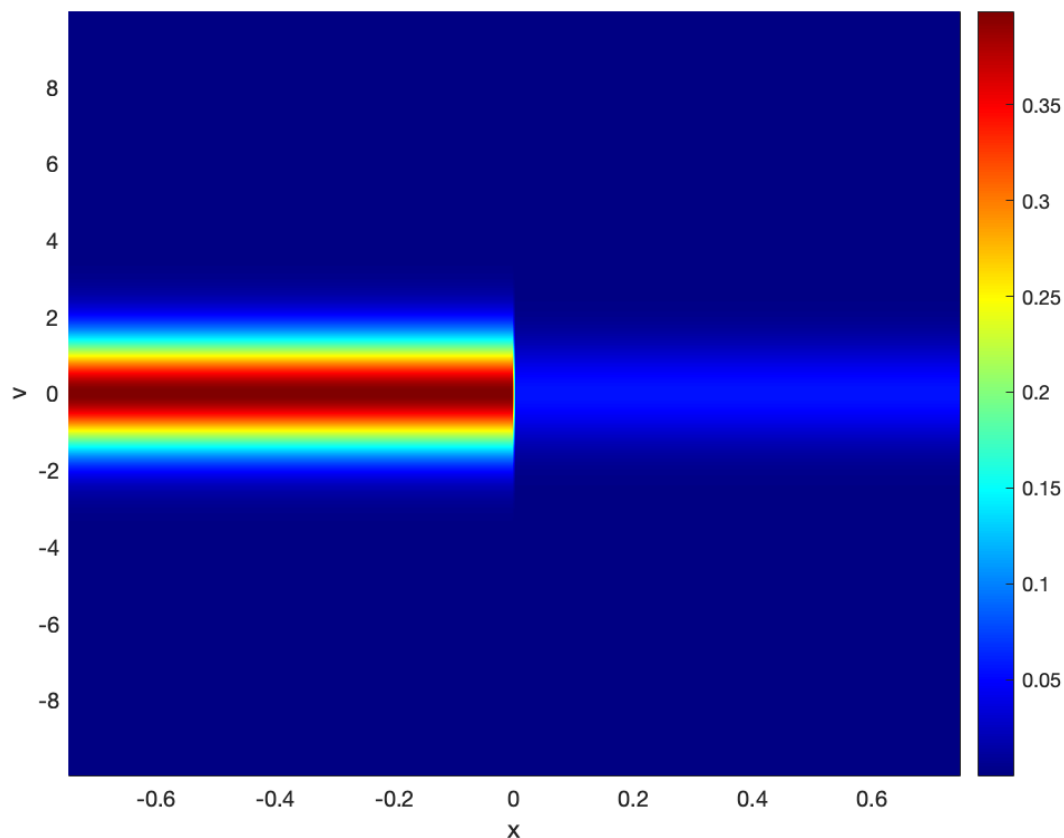


Figure 3.9: Contour plot of Sod initial condition function.

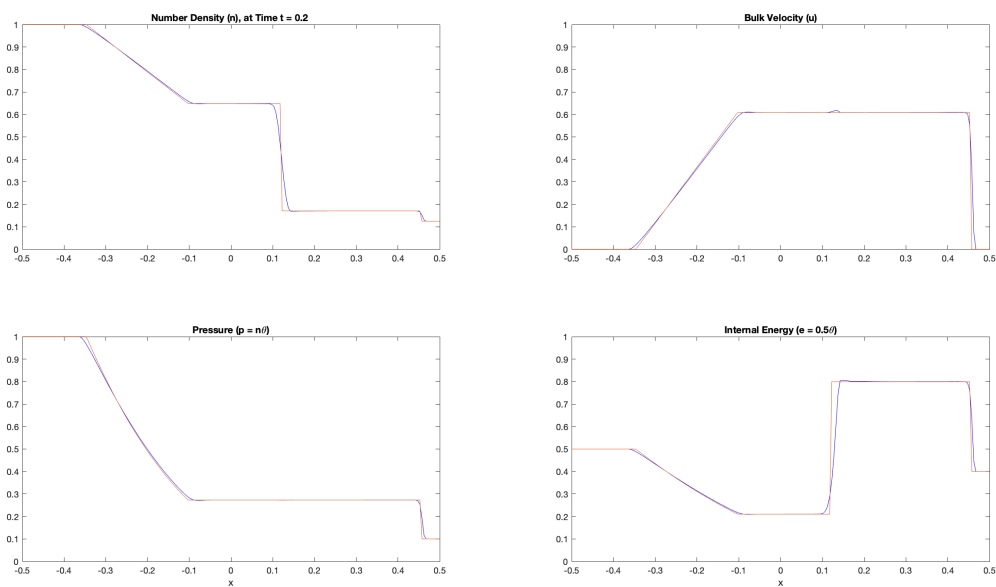


Figure 3.10: Numerical (Blue, BGK, $\lambda = 10^4$) and Theoretical (Orange, Euler) Solution for Sod Shock tube problem. ($\gamma = 3$) at final time $t_F = 0.20$.

Square Pulse Rotation

Consider the Vlasov Equation in 1 dimension:

$$\frac{\partial f}{\partial t} + v \frac{\partial f}{\partial x} + a \frac{\partial f}{\partial v} = 0 \quad (3.50)$$

Setting the acceleration term to $a = -x$ results in a system where an initial distribution is advected counterclockwise, along circular characteristics around the origin. For constructing the characteristics curves, we have the system of ODEs:

$$\frac{dt}{ds} = 1 \quad (3.51)$$

$$\frac{dx}{ds} = v \quad (3.52)$$

$$\frac{dv}{ds} = -x. \quad (3.53)$$

The first equation gives $t = s$, and the remaining two form a system of two coupled ODEs:

$$\dot{x} = v \quad (3.54)$$

$$\dot{v} = -x. \quad (3.55)$$

Taking the derivative of the first equation, and plugging in the second equation, we have

$$\ddot{x} = \dot{v} = -x \iff \ddot{x} + x = 0. \quad (3.56)$$

The solution of this equation is

$$x(t) = c_1 \cos(t) + c_2 \sin(t). \quad (3.57)$$

Thus,

$$\dot{v} = -x = -c_1 \cos(t) - c_2 \sin(t). \quad (3.58)$$

Integrating gives

$$v(t) = -c_1 \sin(t) + c_2 \cos(t). \quad (3.59)$$

Putting these together, we have

$$x(t) = c_1 \cos(t) + c_2 \sin(t) \quad (3.60)$$

$$v(t) = -c_1 \sin(t) + c_2 \cos(t). \quad (3.61)$$

This is a circle of radius $\sqrt{c_1^2 + c_2^2}$, traversed clockwise around the origin in the (x, v) -plane. Figure 3.11 shows an image of four square pulses rotated counterclockwise, until the final time $t_F = \frac{\pi}{2}$ is reached, representing a quarter rotation.

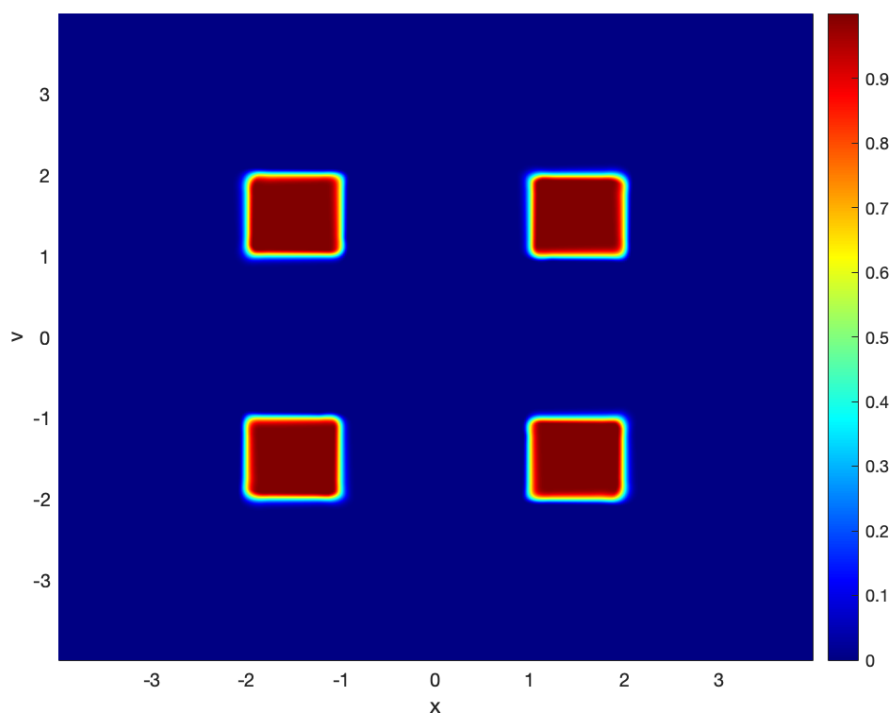


Figure 3.11: Rotation Problem: $\partial_t f + v\partial_x f - x\partial_v f = 0$. Grid = 400^2

Two Stream Instability (Vlasov-Poisson)

The next test that we performed is known as the *Two-Stream Instability Test*, and the present version is taken from Section 5.1.2 of [7]. In this test, two streams of electrons interact and create a highly filamented vortex. The test is designed to assess the code's ability to capture fine structure and examines only the advection piece of our equation (the Vlasov-Poisson equation):

$$\partial_t f + v \partial_x f + E \partial_v f = 0, \quad (x, v, t) \in [-2\pi, 2\pi] \times [-2\pi, 2\pi] \times [0, \infty) \quad (3.62)$$

where

$$-\partial_{xx}\Phi = \frac{e}{\epsilon_0}(n - \bar{n}), \quad E = -\partial_x\Phi, \quad n = \int f dv, \quad \bar{n} = \frac{1}{|V|} \int_V f dv. \quad (3.63)$$

For our test runs, we used $\frac{e}{\epsilon_0} = 1$. The initial condition function for the test is

$$f(x, v, 0) = \frac{v^2}{\sqrt{8\pi}} \left(2 - \cos\left(\frac{x}{2}\right) \right) e^{-\frac{v^2}{2}}. \quad (3.64)$$

The test uses periodic boundary conditions in x , and zero flow boundary conditions in v .

Figure 3.12, taken from [7] shows the plots of the solution at $t_F = 5$ and $t_F = 45$. Figures 3.13 and 3.14 show the results of running the code until $t_F = 5$ and $t_F = 45$, respectively. For these plots, we used 800^2 grid points; this is twice the number of grid points as in the paper (400^2). We find good agreement with the computed solutions.

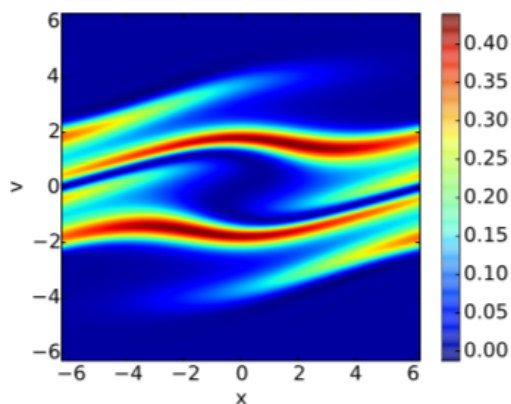
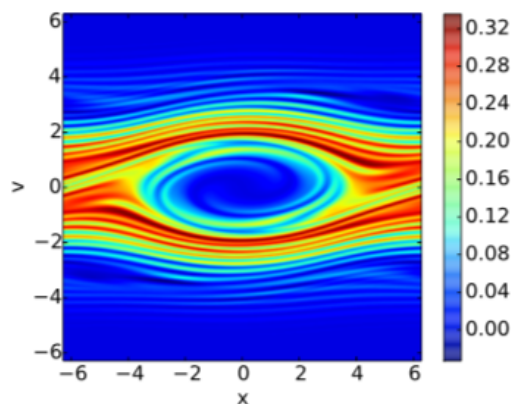
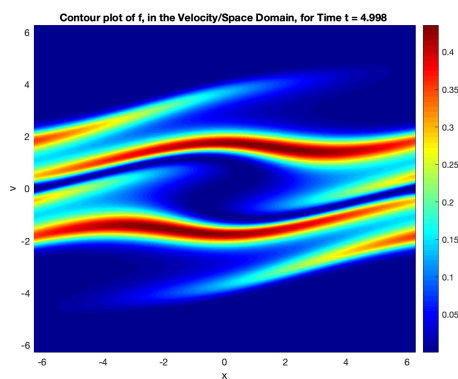
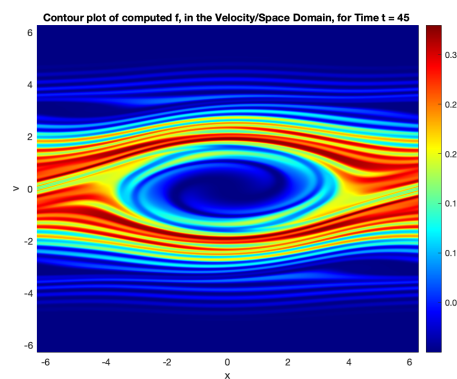
(b) $t = 5.0$ (c) $t = 45.0$

Figure 3.12: Two Stream Instability test case. Figure taken from Garrett & Hauck paper.

Figure 3.13: Code generated solution at $t = 5$.Figure 3.14: Code generated solution at $t = 45$.

Landau Damping (Vlasov-Poisson)

This test is from Section 5.1.3 of [7]. To perform the test, the Vlasov-Poisson equation is solved and the L^2 norm of the electric field is computed at each time step. The initial condition function is

$$f(x, v, 0) = \frac{1}{\sqrt{2\pi}} (1 + \alpha \cos(kx)) e^{-\frac{v^2}{2}}, \quad (3.65)$$

where $\alpha = 0.01$, $k = 0.5$. According to the paper, with these parameters, the L^2 norm of the electric field should decay exponentially at a rate of -0.1533 . Figure 3.15 shows the plots taken from the paper. Figure 3.16 shows the result from our code and indicates qualitative agreement.

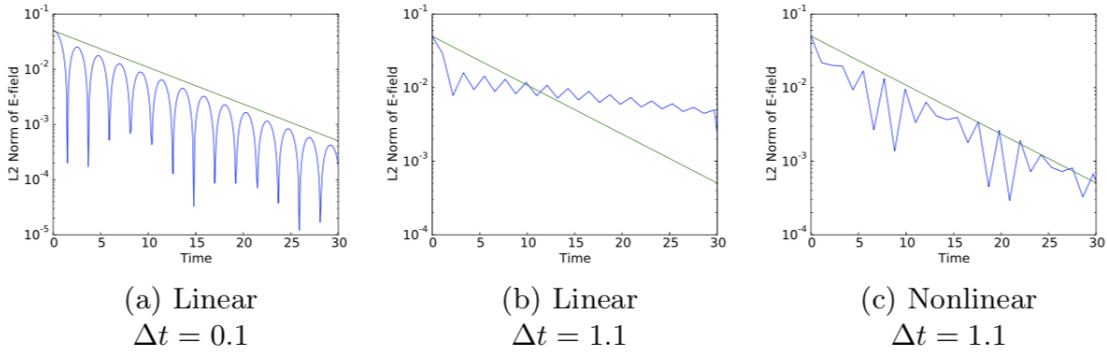


Figure 3.15: Landau Damping test case. Figure taken from Garrett and Hauck [7], with permission.

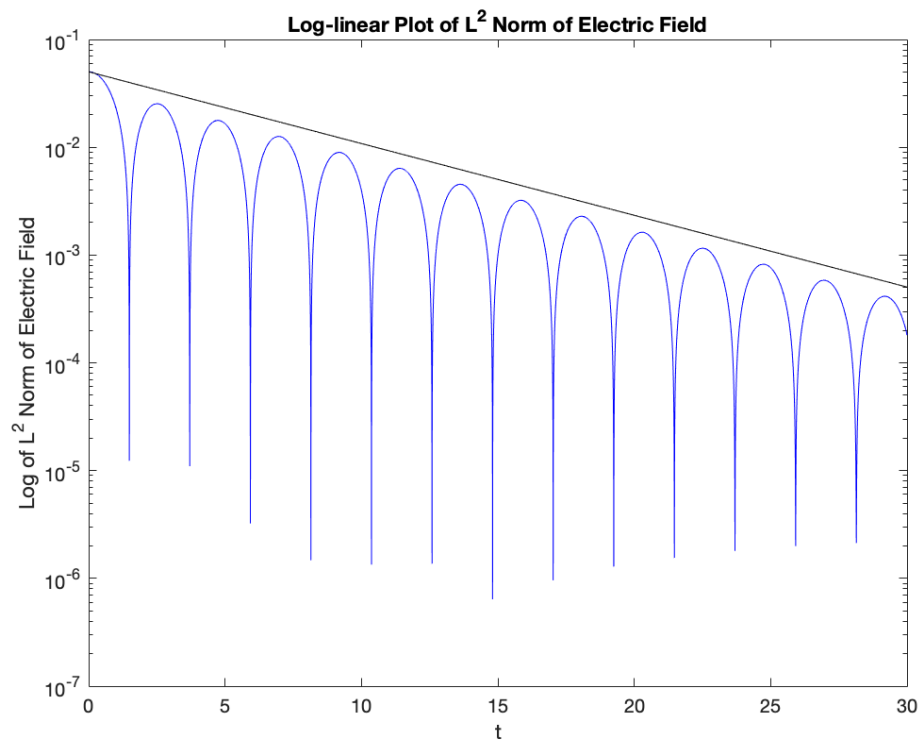


Figure 3.16: L^2 norm of the electric field.

Chapter 4

Discrete-Velocity Numerical Methods for the Single Species BGK Problem

Let us start with the standard discrete-velocity BGK problem first. After that, we will consider the space homogeneous case.

4.1 A Fully Explicit Scheme

Let us again consider the $d = 1$ case. In other words, for simplicity, we will assume that the dimension of physical space is one and the dimension of velocity space is one. We will work under the assumption that velocity space is discrete and the discrete-velocity Maxwellian is computed via optimization. In other words, our starting model is the following:

$$\partial_t f(x, j, t) + v_j \partial_x f(x, j, t) = \lambda (\mathcal{M}_{\mathcal{J}, f}(x, j, t) - f(x, j, t)), \quad \forall j \in \mathcal{J}, \quad \forall x \in [a, b], \quad (4.1)$$

where

$$\mathcal{M}_{\mathcal{J}, f}(x, j, t) = \exp(\boldsymbol{\alpha}_{\mathcal{J}, f}(x, t) \cdot \mathbf{m}_{\mathcal{J}}(j)), \quad \forall j \in \mathcal{J},$$

and $\boldsymbol{\alpha}_{\mathcal{J}, f}(x, t) \in \mathbb{R}^3$ is the solution to

$$\langle \exp(\boldsymbol{\alpha}_{\mathcal{J}, f}(x, t) \cdot \mathbf{m}_{\mathcal{J}}) \mathbf{m}_{\mathcal{J}} \rangle_{\mathcal{J}} = \langle f(x, \cdot, t) \mathbf{m}_{\mathcal{J}} \rangle_{\mathcal{J}}.$$

For simplicity, we assume that the discrete velocities satisfy

$$v_j = h_v \cdot \left(j - \frac{1}{2} \right), \quad j \in \mathcal{J}, \quad \mathcal{J} := \{-J + 1, \dots, J\}.$$

Definition 4.1.1. *The first-order explicit discrete-velocity scheme is defined as follows:*

$$f_{i,j}^{n+1} = f_{i,j}^n - \frac{\Delta t}{h_x} \left(\widehat{F}(x_{i+1/2}, v_j, t^n) - \widehat{F}(x_{i-1/2}, v_j, t^n) \right) + \frac{\Delta t}{\tau} (\mathcal{M}_{i,j}^n - f_{i,j}^n)$$

where

$$\mathcal{M}_{i,j}^n = \exp(\alpha_i^n \cdot \mathbf{m}_{\mathcal{J}}(j)),$$

and $\alpha_i^n \in \mathbb{R}^3$ is the solution to

$$\langle \exp(\alpha_i^n \cdot \mathbf{m}_{\mathcal{J}}) \mathbf{m}_{\mathcal{J}} \rangle_{\mathcal{J}} = \langle f_{i,\cdot}^n \mathbf{m}_{\mathcal{J}} \rangle_{\mathcal{J}}.$$

The first-order numerical flux is obtained via a simple upwind methodology:

$$\widehat{F}(x_{i+1/2}, v_j, t^n) = \begin{cases} v_j f_{i,j}^n & \text{if } v_j \geq 0, \\ v_j f_{i,j+1}^n & \text{if } v_j < 0. \end{cases}$$

Based on our construction, we have the following result:

Theorem 4.1.2. *Suppose that $f_{i,j}^0 > 0$, for all $j \in \mathcal{J}$, and all $1 \leq i \leq N_x$. Then, for all $\mathbf{n} \in \mathbb{N}$,*

$$h_x \sum_{i=1}^{N_x} \langle \mathbf{m}_{\mathcal{J}}(\cdot) f_{i,\cdot}^n \rangle_{\mathcal{J}} = h_x \sum_{i=1}^{N_x} \langle \mathbf{m}_{\mathcal{J}}(\cdot) f_{i,\cdot}^0 \rangle_{\mathcal{J}}. \quad (4.2)$$

Additionally, provided that the CFL condition

$$\Delta t \left(\frac{1}{\tau} + \max_{j \in \mathcal{J}} \frac{|v_j|}{h_x} \right) \leq 1$$

holds, it follows that, for all $\mathbf{n} \in \mathbb{N}$, $j \in \mathcal{J}$, and $i \in \{1, 2, \dots, N_x\}$, $f_{i,j}^n > 0$. Furthermore, for all $\mathbf{n} \in \mathbb{N}$,

$$h_x \sum_{i=1}^{N_x} \langle \ln(f_{i,\cdot}^n) f_{i,\cdot}^n \rangle_{\mathcal{J}} \leq h_x \sum_{i=1}^{N_x} \langle \ln(f_{i,\cdot}^{n-1}) f_{i,\cdot}^{n-1} \rangle_{\mathcal{J}}. \quad (4.3)$$

In other words, the scheme is entropy-stable.

Proof. For $v_j > 0$, the scheme can be written as

$$\begin{aligned} f_{i,j}^{n+1} &= f_{i,j}^n - \frac{\Delta t v_j}{h_x} (f_{i,j}^n - f_{i,j-1}^n) + \frac{\Delta t}{\tau} (\mathcal{M}_{i,j}^n - f_{i,j}^n) \\ &= \left(1 - \frac{\Delta t v_j}{h_x} - \frac{\Delta t}{\tau} \right) f_{i,j}^n + \frac{\Delta t v_j}{h_x} f_{i,j-1}^n + \frac{\Delta t}{\tau} \mathcal{M}_{i,j}^n \end{aligned}$$

$$= \left(1 - \frac{\Delta t |v_j|}{h_x} - \frac{\Delta t}{\tau}\right) f_{i,j}^n + \frac{\Delta t |v_j|}{h_x} f_{i,j-1}^n + \frac{\Delta t}{\tau} \mathcal{M}_{i,j}^n.$$

For $v_j < 0$, the scheme is

$$\begin{aligned} f_{i,j}^{n+1} &= f_{i,j}^n - \frac{\Delta t v_j}{h_x} (f_{i,j+1}^n - f_{i,j}^n) + \frac{\Delta t}{\tau} (\mathcal{M}_{i,j}^n - f_{i,j}^n) \\ &= \left(1 + \frac{\Delta t v_j}{h_x} - \frac{\Delta t}{\tau}\right) f_{i,j}^n - \frac{\Delta t v_j}{h_x} f_{i,j+1}^n + \frac{\Delta t}{\tau} \mathcal{M}_{i,j}^n \\ &= \left(1 - \frac{\Delta t |v_j|}{h_x} - \frac{\Delta t}{\tau}\right) f_{i,j}^n + \frac{\Delta t |v_j|}{h_x} f_{i,j+1}^n + \frac{\Delta t}{\tau} \mathcal{M}_{i,j}^n. \end{aligned}$$

The composite scheme can be written as

$$f_{i,j}^{n+1} = \left(1 - \frac{\Delta t |v_j|}{h_x} - \frac{\Delta t}{\tau}\right) f_{i,j}^n + \frac{\Delta t (|v_j| - v_j)}{2h_x} f_{i,j+1}^n + \frac{\Delta t (|v_j| + v_j)}{2h_x} f_{i,j-1}^n + \frac{\Delta t}{\tau} \mathcal{M}_{i,j}^n,$$

which is a convex combination of $f_{i,j}^n$, $f_{i,j-1}^n$, $f_{i,j+1}^n$, and $\mathcal{M}_{i,j}^n$. If all of these values are positive, then $f_{i,j}^{n+1}$ is positive. Since the function $x \ln(x)$ is convex, it follows that, if $\alpha \in (0, 1)$ and $x \neq y$, $x, y \in (0, \infty)$, then

$$(\alpha x + (1 - \alpha)y) \ln(\alpha x + (1 - \alpha)y) \leq \alpha x \ln(x) + (1 - \alpha)y \ln(y).$$

Thus,

$$\begin{aligned} f_{i,j}^{n+1} \ln(f_{i,j}^{n+1}) &\leq \left(1 - \frac{\Delta t |v_j|}{h_x} - \frac{\Delta t}{\tau}\right) f_{i,j}^n \ln(f_{i,j}^n) + \frac{\Delta t (|v_j| - v_j)}{2h_x} f_{i,j+1}^n \ln(f_{i,j+1}^n) \\ &\quad + \frac{\Delta t (|v_j| + v_j)}{2h_x} f_{i,j-1}^n \ln(f_{i,j-1}^n) + \frac{\Delta t}{\tau} \mathcal{M}_{i,j}^n \ln(\mathcal{M}_{i,j}^n). \end{aligned}$$

Summing on k , we have

$$\begin{aligned} \langle f_{i,\cdot}^{n+1} \ln(f_{i,\cdot}^{n+1}) \rangle_{\mathcal{J}} &\leq \left\langle \left(1 - \frac{\Delta t |v_{\cdot}|}{h_x} - \frac{\Delta t}{\tau}\right) f_{i,\cdot}^n \ln(f_{i,\cdot}^n) \right\rangle_{\mathcal{J}} \\ &\quad + \left\langle \frac{\Delta t (|v_{\cdot}| - v_{\cdot})}{2h_x} f_{i+1,\cdot}^n \ln(f_{i+1,\cdot}^n) \right\rangle_{\mathcal{J}} \\ &\quad + \left\langle \frac{\Delta t (|v_{\cdot}| + v_{\cdot})}{2h_x} f_{i-1,\cdot}^n \ln(f_{i-1,\cdot}^n) \right\rangle_{\mathcal{J}} + \left\langle \frac{\Delta t}{\tau} \mathcal{M}_{i,\cdot}^n \ln(\mathcal{M}_{i,\cdot}^n) \right\rangle_{\mathcal{J}}, \end{aligned}$$

We next invoke the minimization property in (2.47): for all $\mathbf{i}$,

$$\begin{aligned}
\langle f_{\mathbf{i},\cdot}^{n+1} \ln(f_{\mathbf{i},\cdot}^{n+1}) \rangle_{\mathcal{J}} &\leq \left\langle \left(1 - \frac{\Delta t |\mathbf{v}_\cdot|}{h_x} - \frac{\Delta t}{\tau} \right) f_{\mathbf{i},\cdot}^n \ln(f_{\mathbf{i},\cdot}^n) \right\rangle_{\mathcal{J}} \\
&\quad + \left\langle \frac{\Delta t (|\mathbf{v}_\cdot| - v_\cdot)}{2h_x} f_{\mathbf{i}+1,\cdot}^n \ln(f_{\mathbf{i}+1,\cdot}^n) \right\rangle_{\mathcal{J}} \\
&\quad + \left\langle \frac{\Delta t (|\mathbf{v}_\cdot| + v_\cdot)}{2h_x} f_{\mathbf{i}-1,\cdot}^n \ln(f_{\mathbf{i}-1,\cdot}^n) \right\rangle_{\mathcal{J}} + \left\langle \frac{\Delta t}{\tau} f_{\mathbf{i},\cdot}^n \ln(f_{\mathbf{i},\cdot}^n) \right\rangle_{\mathcal{J}} \\
&= \left\langle \left(1 - \frac{\Delta t |\mathbf{v}_\cdot|}{h_x} \right) f_{\mathbf{i},\cdot}^n \ln(f_{\mathbf{i},\cdot}^n) \right\rangle_{\mathcal{J}} \\
&\quad + \left\langle \frac{\Delta t (|\mathbf{v}_\cdot| - v_\cdot)}{2h_x} f_{\mathbf{i}+1,\cdot}^n \ln(f_{\mathbf{i}+1,\cdot}^n) \right\rangle_{\mathcal{J}} \\
&\quad + \left\langle \frac{\Delta t (|\mathbf{v}_\cdot| + v_\cdot)}{2h_x} f_{\mathbf{i}-1,\cdot}^n \ln(f_{\mathbf{i}-1,\cdot}^n) \right\rangle_{\mathcal{J}},
\end{aligned}$$

Summing on $\mathbf{i}$ and using the spatial periodicity, completes the proof of (4.3). The proof of (4.2) is an easy exercise. $\blacksquare$

4.2 An Explicit Scheme for the Space-Homogeneous Case

In this section, we will explore an explicit scheme for the space-homogeneous, discrete-velocity BGK problem. Recall that the one-dimensional, space homogeneous, discrete-velocity model is as follows:

$$\partial_t f(\mathbf{j}, t) = \lambda (\mathcal{M}_{\mathcal{J},f}(\mathbf{j}, t) - f(\mathbf{j}, t)), \quad \forall \mathbf{j} \in \mathcal{J}, \quad (4.4)$$

where

$$\mathcal{M}_{\mathcal{J},f}(\mathbf{j}, t) = \exp(\boldsymbol{\alpha}_{\mathcal{J},f}(t) \cdot \mathbf{m}_{\mathcal{J}}(\mathbf{j})), \quad \forall \mathbf{j} \in \mathcal{J},$$

and $\boldsymbol{\alpha}_{\mathcal{J},f}(t) \in \mathbb{R}^3$ is the solution to

$$\langle \exp(\boldsymbol{\alpha}_{\mathcal{J},f}(t) \cdot \mathbf{m}_{\mathcal{J}}) \mathbf{m}_{\mathcal{J}} \rangle_{\mathcal{J}} = \langle f(\cdot, t) \mathbf{m}_{\mathcal{J}} \rangle_{\mathcal{J}}.$$

As before, it may be useful to suppose that the discrete velocities satisfy

$$v_j = h_v \cdot \left(j - \frac{1}{2} \right), \quad j \in \mathcal{J}, \quad \mathcal{J} := \{-J+1, \dots, J\},$$

though this can be generalized without any difficulty. The explicit scheme for the space-homogeneous case is given by the following:

Definition 4.2.1. *The first-order explicit discrete-velocity scheme is defined as follows:*

$$f_j^{n+1} = f_j^n + \frac{\Delta t}{\tau} (\mathcal{M}_j^n - f_j^n),$$

where

$$\mathcal{M}_j^n = \exp(\boldsymbol{\alpha}^n \cdot \mathbf{m}_{\mathcal{J}}(j)), \quad (4.5)$$

and $\boldsymbol{\alpha}^n \in \mathbb{R}^3$ is the solution to

$$\langle \exp(\boldsymbol{\alpha}^n \cdot \mathbf{m}_{\mathcal{J}}) \mathbf{m}_{\mathcal{J}} \rangle_{\mathcal{J}} = \langle f^n \mathbf{m}_{\mathcal{J}} \rangle_{\mathcal{J}}. \quad (4.6)$$

We have the following result, which shows that an explicit scheme will always have a time step restriction owing to the stiffness of the BGK problem:

Theorem 4.2.2. *Suppose that $f_j^0 > 0$, for all $j \in \mathcal{J}$. Then, for all $n \in \mathbb{N}$,*

$$\langle \mathbf{m}_{\mathcal{J}}(\cdot) f^n \rangle_{\mathcal{J}} = \langle \mathbf{m}_{\mathcal{J}}(\cdot) f^0 \rangle_{\mathcal{J}}. \quad (4.7)$$

Additionally, provided that the condition

$$\frac{\Delta t}{\tau} \leq 1$$

holds, it follows that, for all $n \in \mathbb{N}$ and $j \in \mathcal{J}$, $f_j^n > 0$. Furthermore, for all $n \in \mathbb{N}$,

$$\langle \ln(f^n) f^n \rangle_{\mathcal{J}} \leq \langle \ln(f^{n-1}) f^{n-1} \rangle_{\mathcal{J}}. \quad (4.8)$$

In other words, the scheme is entropy-stable. Finally,

$$\mathcal{M}_j^n = \mathcal{M}_j^0, \quad \forall j \in \mathcal{J}, \quad (4.9)$$

for all $n \in \mathbb{N}$.

Proof. The scheme can be written as

$$f_j^{n+1} = \left(1 - \frac{\Delta t}{\tau}\right) f_j^n + \frac{\Delta t}{\tau} \mathcal{M}_j^n,$$

which is a convex combination of f_j^n and $\mathcal{M}_j^n$. Clearly, f_j^{n+1} is positive. Since the function $x \ln(x)$ is convex, it follows that

$$f_j^{n+1} \ln(f_j^{n+1}) \leq \left(1 - \frac{\Delta t}{\tau}\right) f_j^n \ln(f_j^n) + \frac{\Delta t}{\tau} \mathcal{M}_j^n \ln(\mathcal{M}_j^n).$$

Summing on j , we have

$$\langle f^{n+1} \ln(f^{n+1}) \rangle_{\mathcal{J}} \leq \left\langle \left(1 - \frac{\Delta t}{\tau}\right) f^n \ln(f^n) \right\rangle_{\mathcal{J}} + \left\langle \frac{\Delta t}{\tau} \mathcal{M}^n \ln(\mathcal{M}^n) \right\rangle_{\mathcal{J}},$$

Using the minimization property in (2.47), we find that,

$$\begin{aligned} \langle f^{n+1} \ln(f^{n+1}) \rangle_{\mathcal{J}} &\leq \left\langle \left(1 - \frac{\Delta t}{\tau}\right) f^n \ln(f^n) \right\rangle_{\mathcal{J}} + \left\langle \frac{\Delta t}{\tau} \mathcal{M}^n \ln(\mathcal{M}^n) \right\rangle_{\mathcal{J}} \\ &\leq \left\langle \left(1 - \frac{\Delta t}{\tau}\right) f^n \ln(f^n) \right\rangle_{\mathcal{J}} + \left\langle \frac{\Delta t}{\tau} f^n \ln(f^n) \right\rangle_{\mathcal{J}} \\ &= \langle f^n \ln(f^n) \rangle_{\mathcal{J}}. \end{aligned}$$

The proof of (4.7) is an easy exercise and is skipped for brevity. The proof of (4.9) follows from (4.7), (4.5), and (4.6). ■

4.3 An Implicit Scheme for the Space Homogeneous Case

In the last section, we learned that an explicit scheme for the space homogeneous case still requires a potentially restrictive time step constraint for stability. But, this can be easily overcome in the space homogeneous case.

Definition 4.3.1. *The first-order implicit discrete-velocity scheme is defined as follows:*

$$f_j^{n+1} = f_j^n + \frac{\Delta t}{\tau} (\mathcal{M}_j^n - f_j^{n+1}),$$

where

$$\mathcal{M}_j^n = \exp(\boldsymbol{\alpha}^n \cdot \mathbf{m}_{\mathcal{J}}(j)), \tag{4.10}$$

and $\boldsymbol{\alpha}^n \in \mathbb{R}^3$ is the solution to

$$\langle \exp(\boldsymbol{\alpha}^n \cdot \mathbf{m}_{\mathcal{J}}) \mathbf{m}_{\mathcal{J}} \rangle_{\mathcal{J}} = \langle f^n \mathbf{m}_{\mathcal{J}} \rangle_{\mathcal{J}}. \tag{4.11}$$

Technically, the last scheme is an IMEX scheme, but we will stick to our definition, since we can show that, actually, $\mathcal{M}_j^n = \mathcal{M}_j^{n+1}$. Specifically, we have the following result:

Theorem 4.3.2. *Suppose that $f_j^0 > 0$, for all $j \in \mathcal{J}$. Then, for all $n \in \mathbb{N}$,*

$$\langle m_{\mathcal{J}}(\cdot) f^n \rangle_{\mathcal{J}} = \langle m_{\mathcal{J}}(\cdot) f^0 \rangle_{\mathcal{J}}. \quad (4.12)$$

Additionally, for all $n \in \mathbb{N}$ and $j \in \mathcal{J}$, $f_k^n > 0$, and for all $n \in \mathbb{N}$,

$$\langle \ln(f^n) f^n \rangle_{\mathcal{J}} \leq \langle \ln(f^{n-1}) f^{n-1} \rangle_{\mathcal{J}}. \quad (4.13)$$

In other words, the scheme is entropy-stable. Finally,

$$\mathcal{M}_j^n = \mathcal{M}_j^0, \quad \forall j \in \mathcal{J}, \quad (4.14)$$

for all $n \in \mathbb{N}$. All of these properties hold unconditionally.

Proof. The scheme can be written as

$$\left(1 + \frac{\Delta t}{\tau}\right) f_j^{n+1} = f_j^n + \frac{\Delta t}{\tau} \mathcal{M}_j^n,$$

or, equivalently, as

$$f_j^{n+1} = \frac{1}{\left(1 + \frac{\Delta t}{\tau}\right)} f_j^n + \frac{\frac{\Delta t}{\tau}}{\left(1 + \frac{\Delta t}{\tau}\right)} \mathcal{M}_j^n, \quad (4.15)$$

which is again a convex combination of f_j^n and $\mathcal{M}_j^n$. Clearly, f_j^{n+1} is positive, provided f_j^n and $\mathcal{M}_j^n$ are positive. Let us write, for simplicity,

$$\alpha := \frac{1}{1 + \frac{\Delta t}{\tau}} \quad \text{and} \quad \beta := \frac{\frac{\Delta t}{\tau}}{1 + \frac{\Delta t}{\tau}}.$$

Since the function $x \ln(x)$ is convex, it follows that

$$f_j^{n+1} \ln(f_j^{n+1}) \leq \alpha f_j^n \ln(f_j^n) + \beta \mathcal{M}_j^n \ln(\mathcal{M}_j^n).$$

Summing on j , we have

$$\langle f^{n+1} \ln(f^{n+1}) \rangle_{\mathcal{J}} \leq \langle \alpha f^n \ln(f^n) \rangle_{\mathcal{J}} + \langle \beta \mathcal{M}^n \ln(\mathcal{M}^n) \rangle_{\mathcal{J}},$$

Using the minimization property in (2.47), we find that,

$$\begin{aligned} \langle f^{n+1} \ln(f^{n+1}) \rangle_{\mathcal{J}} &\leq \langle \alpha f^n \ln(f^n) \rangle_{\mathcal{J}} + \langle \beta \mathcal{M}^n \ln(\mathcal{M}^n) \rangle_{\mathcal{J}} \\ &\leq \langle \alpha f^n \ln(f^n) \rangle_{\mathcal{J}} + \langle \beta f^n \ln(f^n) \rangle_{\mathcal{J}} \\ &= \langle f^n \ln(f^n) \rangle_{\mathcal{J}}. \end{aligned}$$

Next, summing $m_{\mathcal{J}}(\cdot)$ times (4.15) on j , we have

$$\begin{aligned} \langle m_{\mathcal{J}}(\cdot) f^{n+1} \rangle_{\mathcal{J}} &= \langle \alpha m_{\mathcal{J}}(\cdot) f^n \rangle_{\mathcal{J}} + \langle \beta m_{\mathcal{J}}(\cdot) \mathcal{M}^n \rangle_{\mathcal{J}} \\ &= \langle \alpha m_{\mathcal{J}}(\cdot) f^n \rangle_{\mathcal{J}} + \langle \beta m_{\mathcal{J}}(\cdot) f^n \rangle_{\mathcal{J}} \\ &= \langle m_{\mathcal{J}}(\cdot) f^n \rangle_{\mathcal{J}}, \end{aligned}$$

which proves (4.12), and, finally, (4.14) follows directly from (4.12). ■

Chapter 5

Multispecies BGK Equations

In the present section, we describe a relatively recent BGK-type model for the multi-species setting [9]. This model, which generalizes the one-species case, satisfies a multi-species analog of Boltzmann's H-Theorem [9]. In the multispecies setting, the locally conserved quantities are the species number $m_i \langle f_i \rangle$, the total momentum $\sum_i m_i \langle \mathbf{v} f_i \rangle$, and the total energy $\sum_i m_i \langle |\mathbf{v}|^2 f_i \rangle$; and the dissipated entropy (Lyapunov functional) is the total (mathematical) entropy $\sum_i \eta(f_i)$, where $\eta(f) := f \ln(f) - f$ is the kinetic entropy density. In the zero-relaxation-time limit, hydrodynamic equations can be derived along the lines of the single species BGK equation [9]. Other consistent multi-species models can be found in [11, 14].

5.1 Formulation and Properties

Suppose that there are $N_s \in \mathbb{N}$ ($N_s \geq 2$) species of particles in a mixture of gases. For each species, $i \in \{1, \dots, N_s\}$, denote the kinetic distribution of particles with mass m_i , by $f_i(\mathbf{x}, \mathbf{v}, t)$, so that $f_i(\mathbf{x}, \mathbf{v}, t)$ describes the density of species- i particles at $\mathbf{x} \in \Omega \subset \mathbb{R}^d$, with microscopic velocity $\mathbf{v} \in \mathbb{R}^d$, at time $t \geq 0$, with respect to the measure $d\mathbf{v} d\mathbf{x}$. Define the *number density* n_i , *bulk velocity* $\mathbf{u}_i$, and *temperature* T_i , via

$$n_i = \int_{\mathbb{R}^d} f_i d\mathbf{v}, \quad \mathbf{u}_i = \frac{1}{n_i} \int_{\mathbb{R}^d} f_i \mathbf{v} d\mathbf{v}, \quad T_i = \frac{m_i}{n_i d} \int_{\mathbb{R}^d} f_i |\mathbf{u}_i - \mathbf{v}|^2 d\mathbf{v}, \quad (5.1)$$

and the *mass density* ρ_i , *momentum density* $\mathbf{k}_i$, and *energy density* E_i by

$$\rho_i = m_i n_i \quad \mathbf{k}_i = \rho_i \mathbf{u}_i \quad E_i = \frac{1}{2} \rho_i |\mathbf{u}_i|^2 + \frac{d}{2} n_i T_i. \quad (5.2)$$

The *multispecies Vlasov-BGK equation* models the evolution of the distribution fields $f_i(\mathbf{x}, \mathbf{v}, t)$, $i = 1, \dots, N_s$, via the system

$$\frac{\partial f_i}{\partial t} + \mathbf{v} \cdot \nabla_{\mathbf{x}} f_i + \mathbf{a}_i \cdot \nabla_{\mathbf{v}} f_i = \sum_{j=1}^{N_s} \lambda_{i,j} (M_{i,j} - f_i), \quad i = 1, \dots, N_s, \quad (5.3)$$

where $\mathbf{a}_i$ describes the acceleration of particles of species i , $\lambda_{i,j}$ is the collision frequency between species i and j , and the Maxwellian distribution $M_{i,j}$ is defined as

$$M_{i,j} = M_{i,j}[f_i, f_j](\mathbf{x}, \mathbf{v}, t) = n_i \left(\frac{m_i}{2\pi T_{i,j}} \right)^{\frac{d}{2}} \exp \left(-\frac{m_i |\mathbf{v} - \mathbf{u}_{i,j}|^2}{2T_{i,j}} \right), \quad (5.4)$$

using the *mixture velocities* and *mixture temperatures*,

$$\mathbf{u}_{i,j} = \frac{\lambda_{i,j} \rho_i \mathbf{u}_i + \lambda_{j,i} \rho_j \mathbf{u}_j}{\lambda_{i,j} \rho_i + \lambda_{j,i} \rho_j} \quad (5.5a)$$

$$T_{i,j} = \frac{\lambda_{i,j} n_i T_i + \lambda_{j,i} n_j T_j}{\lambda_{i,j} n_i + \lambda_{j,i} n_j} + \frac{\lambda_{i,j} \rho_i (|\mathbf{u}_i|^2 - |\mathbf{u}_{i,j}|^2) + \lambda_{j,i} \rho_j (|\mathbf{u}_j|^2 - |\mathbf{u}_{i,j}|^2)}{d(\lambda_{i,j} n_i + \lambda_{j,i} n_j)}. \quad (5.5b)$$

The mixture velocities and temperatures, $\mathbf{u}_{i,j}$ and $T_{i,j}$, are chosen so that certain collision invariances hold, as we show momentarily. The existence and uniqueness of nonnegative mild solutions to the multispecies BGK equation was proved by Klingenberg & Pirner (2018) [13], for periodic physical space and under certain restrictions on the collision frequencies.

It is straightforward to check that the proposed system satisfies the usual conservation properties and an entropy dissipation property via an H-Theorem-like result. See [9] for details. In this next computation, we show that the system, as defined above, satisfies certain collision invariances.

Lemma 5.1.1. *If the mixture velocities $\mathbf{u}_{i,j}$ and the mixture temperatures $T_{i,j}$ are as given in (5.5), then the multispecies BGK collision operators*

$$Q_{i,j}[f_i, f_j] = \lambda_{i,j} (M_{i,j} - f_i) \quad (5.6)$$

satisfy the following conservation properties, which correspond to the conservation of mass, total momentum, and total energy: for any $i, j \in \{1, \dots, N_s\}$,

$$\int_{\mathbb{R}^d} \lambda_{i,j} (M_{i,j} - f_i) d\mathbf{v} = 0, \quad (5.7a)$$

$$\int_{\mathbb{R}^d} \lambda_{i,j}(M_{i,j} - f_i)m_i \mathbf{v} d\mathbf{v} + \int_{\mathbb{R}^d} \lambda_{j,i}(M_{j,i} - f_j)m_j \mathbf{v} d\mathbf{v} = \mathbf{0}, \quad (5.7b)$$

$$\int_{\mathbb{R}^d} \lambda_{i,j}(M_{i,j} - f_i)m_i |\mathbf{v}|^2 d\mathbf{v} + \int_{\mathbb{R}^d} \lambda_{j,i}(M_{j,i} - f_j)m_j |\mathbf{v}|^2 d\mathbf{v} = 0. \quad (5.7c)$$

Proof. Note that using the substitution

$$\mathbf{s} = \frac{\mathbf{v} - \mathbf{u}_{i,j}}{\left(\frac{2T_{i,j}}{m_i}\right)^{\frac{1}{2}}} \implies d\mathbf{s} = \left(\frac{m_i}{2T_{i,j}}\right)^{\frac{d}{2}} d\mathbf{v}, \quad (5.8)$$

and the fact that

$$\int_{\mathbb{R}^d} \exp(-|\mathbf{s}|^2) d\mathbf{s} = \pi^{\frac{d}{2}}, \quad \int_{\mathbb{R}^d} \mathbf{s} \exp(-|\mathbf{s}|^2) d\mathbf{s} = \mathbf{0}, \quad \int_{\mathbb{R}^d} |\mathbf{s}|^2 \exp(-|\mathbf{s}|^2) d\mathbf{s} = \frac{d}{2} \pi^{\frac{d}{2}}, \quad (5.9)$$

straightforward computations show that the moments of the Maxwellian $M_{i,j}$ satisfy

$$\int_{\mathbb{R}^d} M_{i,j} d\mathbf{v} = n_i, \quad \int_{\mathbb{R}^d} \mathbf{v} M_{i,j} d\mathbf{v} = n_i \mathbf{u}_{i,j}, \quad \int_{\mathbb{R}^d} |\mathbf{v}|^2 M_{i,j} d\mathbf{v} = n_i |\mathbf{u}_{i,j}|^2 + dn_i \frac{T_{i,j}}{m_i}. \quad (5.10)$$

Using these identities, the collision invariance properties of (5.7) follow. First, note that

$$I_1 := \int_{\mathbb{R}^d} Q_{i,j}^{BGK} d\mathbf{v} = \lambda_{i,j} \int_{\mathbb{R}^d} M_{i,j} d\mathbf{v} - \lambda_{i,j} \int_{\mathbb{R}^d} f_i d\mathbf{v} = \lambda_{i,j} n_i - \lambda_{i,j} n_i = 0, \quad (5.11)$$

which establishes (5.7a). To prove (5.7b), note that

$$\begin{aligned} I_2 &:= \int_{\mathbb{R}^d} Q_{i,j}^{BGK} m_i \mathbf{v} d\mathbf{v} + \int_{\mathbb{R}^d} Q_{j,i}^{BGK} m_j \mathbf{v} d\mathbf{v} \\ &= m_i \lambda_{i,j} \int_{\mathbb{R}^d} M_{i,j} \mathbf{v} d\mathbf{v} - m_i \lambda_{i,j} \int_{\mathbb{R}^d} f_i \mathbf{v} d\mathbf{v} + m_j \lambda_{j,i} \int_{\mathbb{R}^d} M_{j,i} \mathbf{v} d\mathbf{v} - m_j \lambda_{j,i} \int_{\mathbb{R}^d} f_j \mathbf{v} d\mathbf{v} \\ &= \rho_i \lambda_{i,j} \mathbf{u}_{i,j} - \lambda_{i,j} \rho_i \mathbf{u}_i + \rho_j \lambda_{j,i} \mathbf{u}_{j,i} - \rho_j \lambda_{j,i} \mathbf{u}_j. \end{aligned} \quad (5.12)$$

Rearranging this, and using the symmetry assumption, $\mathbf{u}_{i,j} = \mathbf{u}_{j,i}$ the result follows:

$$\mathbf{u}_{i,j} = \frac{\lambda_{i,j} \rho_i \mathbf{u}_i + \lambda_{j,i} \rho_j \mathbf{u}_{j,i}}{\lambda_{i,j} \rho_i + \lambda_{j,i} \rho_j} \implies (\lambda_{i,j} \rho_i + \lambda_{j,i} \rho_j) \mathbf{u}_{i,j} - \lambda_{i,j} \rho_i \mathbf{u}_i - \lambda_{j,i} \rho_j \mathbf{u}_j = I_2 = \mathbf{0}. \quad (5.13)$$

Finally, for (5.7c), note that

$$\begin{aligned}
I_3 &= \int_{\mathbb{R}^d} Q_{i,j}^{BGK} m_i |\mathbf{v}|^2 d\mathbf{v} + \int_{\mathbb{R}^d} Q_{j,i}^{BGK} m_j |\mathbf{v}|^2 d\mathbf{v} \\
&= m_i \lambda_{i,j} \int_{\mathbb{R}^d} M_{i,j} |\mathbf{v}|^2 d\mathbf{v} - m_i \lambda_{i,j} \int_{\mathbb{R}^d} f_i |\mathbf{v}|^2 d\mathbf{v} + m_j \lambda_{j,i} \int_{\mathbb{R}^d} M_{j,i} |\mathbf{v}|^2 d\mathbf{v} - m_j \lambda_{j,i} \int_{\mathbb{R}^d} f_j |\mathbf{v}|^2 d\mathbf{v} \\
&= m_i \lambda_{i,j} \left(n_i |\mathbf{u}_{i,j}|^2 + d n_i \frac{T_{i,j}}{m_i} \right) - m_i \lambda_{i,j} \left(n_i |\mathbf{u}_i|^2 + d n_i \frac{T_i}{m_i} \right) \\
&\quad + m_j \lambda_{j,i} \left(n_j |\mathbf{u}_{j,i}|^2 + d n_j \frac{T_{j,i}}{m_j} \right) - m_j \lambda_{j,i} \left(n_j |\mathbf{u}_j|^2 + d n_j \frac{T_j}{m_j} \right) \\
&= \rho_i \lambda_{i,j} [|\mathbf{u}_{i,j}|^2 - |\mathbf{u}_i|^2] + \rho_j \lambda_{j,i} [|\mathbf{u}_{j,i}|^2 - |\mathbf{u}_j|^2] + d n_i \lambda_{i,j} [T_{i,j} - T_i] + d n_j \lambda_{j,i} [T_{j,i} - T_j] \quad (5.14)
\end{aligned}$$

Using symmetry assumptions $\mathbf{u}_{i,j} = \mathbf{u}_{j,i}$ and $T_{i,j} = T_{j,i}$, and gathering terms as in (5.13),

$$\begin{aligned}
T_{i,j} &= \frac{n_i \lambda_{i,j} T_i + n_j \lambda_{j,i} T_j}{n_i \lambda_{i,j} + n_j \lambda_{j,i}} + \frac{\rho_i \lambda_{i,j} (|\mathbf{u}_i|^2 - |\mathbf{u}_{i,j}|^2) + \rho_j \lambda_{j,i} (|\mathbf{u}_j|^2 - |\mathbf{u}_{j,i}|^2)}{d(n_i \lambda_{i,j} + d n_j \lambda_{j,i})} \\
\Rightarrow T_{i,j} d[n_i \lambda_{i,j} + n_j \lambda_{j,i}] &= d[n_i \lambda_{i,j} T_i + n_j \lambda_{j,i} T_j] + \rho_i \lambda_{i,j} [|\mathbf{u}_i|^2 - |\mathbf{u}_{i,j}|^2] + \rho_j \lambda_{j,i} [|\mathbf{u}_j|^2 - |\mathbf{u}_{j,i}|^2] \\
\Rightarrow I_3 &= 0. \quad (5.15)
\end{aligned}$$

The proof is complete. ■

The converse of the last result is also true, as is easy to show. We omit the proof for the sake of brevity.

Lemma 5.1.2. *Suppose that, for any $i, j \in \{1, \dots, N_s\}$, the following collision invariances hold:*

$$\int_{\mathbb{R}^d} \lambda_{i,j} (M_{i,j} - f_i) d\mathbf{v} = 0, \quad (5.16)$$

$$\int_{\mathbb{R}^d} \lambda_{i,j} (M_{i,j} - f_i) m_i \mathbf{v} d\mathbf{v} + \int_{\mathbb{R}^d} \lambda_{j,i} (M_{j,i} - f_j) m_j \mathbf{v} d\mathbf{v} = \mathbf{0}, \quad (5.17)$$

$$\int_{\mathbb{R}^d} \lambda_{i,j} (M_{i,j} - f_i) m_i |\mathbf{v}|^2 d\mathbf{v} + \int_{\mathbb{R}^d} \lambda_{j,i} (M_{j,i} - f_j) m_j |\mathbf{v}|^2 d\mathbf{v} = 0. \quad (5.18)$$

Consequently, it must be that the mixture velocities and temperatures satisfy

$$\mathbf{u}_{i,j} = \frac{\lambda_{i,j} \rho_i \mathbf{u}_i + \lambda_{j,i} \rho_j \mathbf{u}_j}{\lambda_{i,j} \rho_i + \lambda_{j,i} \rho_j} \quad (5.19)$$

$$T_{i,j} = \frac{\lambda_{i,j} n_i T_i + \lambda_{j,i} n_j T_j}{\lambda_{i,j} n_i + \lambda_{j,i} n_j} + \frac{\lambda_{i,j} \rho_i (|\mathbf{u}_i|^2 - |\mathbf{u}_{i,j}|^2) + \lambda_{j,i} \rho_j (|\mathbf{u}_j|^2 - |\mathbf{u}_{j,i}|^2)}{d(\lambda_{i,j} n_i + \lambda_{j,i} n_j)}. \quad (5.20)$$

5.2 Numerical Approximation

For certain parameter regimes, the problem may become numerically stiff, requiring prohibitively small time steps to resolve the dynamics of the particle interactions. To enable taking larger time steps, we aim to develop implicit and semi-implicit methods for the multispecies problem, wherein the collision term is computed implicitly. Thus, to see what difficulties lie ahead for the numerical analysis of the multi-species case, let us first examine the space homogeneous problem,

$$\frac{\partial f_i}{\partial t} = \sum_{j=1}^{N_s} \lambda_{ij}(M_{ij} - f_i), \quad \text{for } i \in \{1, \dots, N_s\}. \quad (5.21)$$

As a prototype for the implicit treatment of the collision term, we discretize using the Backward Euler method:

$$\frac{f_i^{n+1} - f_i^n}{\Delta t} = \sum_{j=1}^{N_s} \lambda_{ij}(M_{ij}^{n+1} - f_i^{n+1}) \quad (5.22)$$

$$\iff f_i^{n+1} = f_i^n + \Delta t \sum_{j=1}^{N_s} \lambda_{ij} M_{ij}^{n+1} - \Delta t f_i^{n+1} \sum_{j=1}^{N_s} \lambda_{ij} \quad (5.23)$$

$$\iff f_i^{n+1} \left(1 + \Delta t \sum_{j=1}^{N_s} \lambda_{ij} \right) = f_i^n + \Delta t \sum_{j=1}^{N_s} \lambda_{ij} M_{ij}^{n+1} \quad (5.24)$$

$$\iff f_i^{n+1} = \frac{f_i^n + \Delta t \sum_{j=1}^{N_s} \lambda_{ij} M_{ij}^{n+1}}{1 + \Delta t \sum_{j=1}^{N_s} \lambda_{ij}}. \quad (5.25)$$

Note that (5.25) has the terms M_{ij}^{n+1} . Recall that for the space homogeneous, single species model, the Maxwellian is time invariant (see (2.35)), so that $M^n = M^{n+1}$. This temporal invariance does not hold in the multispecies case, and this represents a significant numerical challenge. Care must be taken to give a proper implicit update for the collision operator, and this is the subject of current work in the project.

As with fluid equations, spatially adaptive meshes are a requirement for highly efficient simulations of flows with fine-scale structures in phase space. However, kinetic equations also require adaptivity, not only for the resolution of fine scale structures, but also to address the fact that the effective support of the kinetic distribution may vary dramatically in phase space. Typically the size of the domain is based on the temperature of the distribution; for multispecies problems, each species may have its own temperature, which adds a complication

not found in the single species setting. In other words, the Maxwellians may require different phase space resolutions. One way that this can be addressed is by defining the Maxwellians on different compatible grids that adequately resolve their individual supports.

Chapter 6

Asymptotic Relaxation of Multi-Species Moment Equations

In the current chapter, which is based on the preprint [10], we study the space-homogeneous, multi-species BGK equations from [9], which can be understood as a special, illustrative case of the equations in [3] or [15]. We do so in the context of collision frequencies that are independent of the phase-space velocity, but depend on both number density and temperature. In particular, we analyze the ODE system for the momentum and energy moments (the number density evolution is trivial), which is important for asymptotic analysis and for the development of numerical methods.

Indeed, many numerical simulations of the full BGK system (including phase-space advection) rely on implicit-explicit (IMEX) methods that treat the space-homogeneous component implicitly. In this setting, the result of solving the associated moments equations first is to effectively linearize and diagonalize (with respect to the phase-space velocity) the required implicit solve, yielding a significant reduction in computational cost. (See [5, 23] for applications to the single species case and [25, Section 4] for an extension to the multi-species setting when the collision frequencies depend only on the number densities.)

In this chapter, we first establish existence and uniqueness of the momentum-energy ODE system, which follows from standard theory once a lower positive bound can be established on the species temperatures. We show monotonicity of the minimum temperature envelope as well as upper and lower bounds on the bulk velocity. We then prove exponential decay of the species momenta and energies to their steady state values, thereby generalizing some of the results in [6] to a system with an arbitrary number of species.

6.1 Model Equations

Here we recap the multi-species BGK equation, specifically for the case that the collision frequencies are dependent upon species temperatures. We denote by f_i the kinetic distribution of particles of species $i \in \{1, \dots, N\}$ having mass m_i . More specifically, $f_i(\mathbf{x}, \mathbf{v}, t)$ is the density of species i particles at the point $\mathbf{x} \in \Omega \subset \mathbb{R}^d$, with microscopic velocity $\mathbf{v} \in \mathbb{R}^d$, at time $t \geq 0$, with respect to the measure $d\mathbf{v} d\mathbf{x}$. Associated to each f_i are the species number density n_i , mass density ρ_i , bulk velocity $\mathbf{u}_i$, and temperature T_i , defined by

$$\rho_i = m_i n_i = m_i \int_{\mathbb{R}^d} f_i d\mathbf{v}, \quad \mathbf{u}_i = \frac{1}{\rho_i} \int_{\mathbb{R}^d} m_i \mathbf{v} f_i d\mathbf{v}, \quad T_i = \frac{m_i}{n_i d} \int_{\mathbb{R}^d} |\mathbf{v} - \mathbf{u}_i|^2 f_i d\mathbf{v}. \quad (6.1)$$

The species momentum densities $\mathbf{k}_i$ and species energy densities E_i are given by

$$\mathbf{k}_i = \rho_i \mathbf{u}_i \quad \text{and} \quad E_i = \frac{1}{2} \rho_i |\mathbf{u}_i|^2 + \frac{d}{2} n_i T_i. \quad (6.2)$$

BGK models are expressed in terms of Maxwellian distributions that depend on these moments.

Definition 6.1.1. Given $n > 0$, $\mathbf{u} \in \mathbb{R}^d$, $\theta > 0$, a Maxwellian $M_{n,\mathbf{u},\theta}$ is a function of the form

$$M_{n,\mathbf{u},\theta}(\mathbf{v}) = \frac{n}{(2\pi\theta)^{\frac{d}{2}}} \exp\left(-\frac{|\mathbf{v} - \mathbf{u}|^2}{2\theta}\right). \quad (6.3)$$

Straightforward computations show that the moments of a Maxwellian satisfy

$$\int_{\mathbb{R}^d} M_{n,\mathbf{u},\theta} d\mathbf{v} = n, \quad \int_{\mathbb{R}^d} \mathbf{v} M_{n,\mathbf{u},\theta} d\mathbf{v} = n\mathbf{u}, \quad \int_{\mathbb{R}^d} |\mathbf{v}|^2 M_{n,\mathbf{u},\theta} d\mathbf{v} = n|\mathbf{u}|^2 + dn\theta. \quad (6.4)$$

We consider in this paper the multi-species BGK equation from [9], given by

$$\frac{\partial f_i}{\partial t} + \mathbf{v} \cdot \nabla_{\mathbf{x}} f_i = \frac{1}{\varepsilon} \sum_j \lambda_{i,j} (M_{i,j} - f_i), \quad \forall i \in \{1, \dots, N\}, \quad (6.5)$$

where $\varepsilon > 0$ is the Knudsen number, $\lambda_{i,j} > 0$ is the (nondimensional) frequency of collisions between species i and j (independent of $\mathbf{v}$), and $M_{i,j}(\mathbf{v}) = M_{n_i, \mathbf{u}_{i,j}, T_{i,j}/m_i}(\mathbf{v})$ is a Maxwellian defined by (6.3), using the mixture velocities and temperatures

$$\mathbf{u}_{i,j} = \frac{\rho_i \lambda_{i,j} \mathbf{u}_i + \rho_j \lambda_{j,i} \mathbf{u}_j}{\rho_i \lambda_{i,j} + \rho_j \lambda_{j,i}}, \quad T_{i,j} = \frac{n_i \lambda_{i,j} T_i + n_j \lambda_{j,i} T_j}{n_i \lambda_{i,j} + n_j \lambda_{j,i}} + \frac{1}{d} \frac{\rho_i \rho_j \lambda_{i,j} \lambda_{j,i}}{\rho_i \lambda_{i,j} + \rho_j \lambda_{j,i}} \frac{|\mathbf{u}_i - \mathbf{u}_j|^2}{n_i \lambda_{i,j} + n_j \lambda_{j,i}}. \quad (6.6)$$

The existence and uniqueness of nonnegative mild solutions to (a more general) multi-species BGK equation was proved in [13] for periodic spatial domains and the collision frequencies that depend on the number densities.

6.1.1 Model Properties

The BGK collision operators in (6.5) satisfy the following invariance properties: For any i, j ,

$$\int_{\mathbb{R}^d} \lambda_{i,j}(M_{i,j} - f_i) d\mathbf{v} = 0 \quad (6.7)$$

$$\int_{\mathbb{R}^d} \lambda_{i,j}(M_{i,j} - f_i) m_i \mathbf{v} d\mathbf{v} + \int_{\mathbb{R}^d} \lambda_{j,i}(M_{j,i} - f_j) m_j \mathbf{v} d\mathbf{v} = \mathbf{0} \quad (6.8)$$

$$\int_{\mathbb{R}^d} \lambda_{i,j}(M_{i,j} - f_i) m_i |\mathbf{v}|^2 d\mathbf{v} + \int_{\mathbb{R}^d} \lambda_{j,i}(M_{j,i} - f_j) m_j |\mathbf{v}|^2 d\mathbf{v} = 0, \quad (6.9)$$

corresponding to conservation of species mass, total momentum, and total energy, respectively. Furthermore, an entropy dissipation condition is satisfied: For a spatially homogeneous mixture, the entropy function $\mathcal{H} = \sum_i \int_{\mathbb{R}^d} f_i \log f_i d\mathbf{v}$ satisfies $\frac{d}{dt} \mathcal{H} \leq 0$. Moreover, $\frac{d}{dt} \mathcal{H} = 0$ if and only if $f_i = M_{n_i, \mathbf{u}_{\text{eq}}, T_{i,j}/m_i}(\mathbf{v})$ for all $i \in \{1, \dots, N\}$, where $\mathbf{u}_{\text{eq}}$ and T_{eq} are the equilibrium bulk velocity and temperature, respectively, which are common to all species [9].

The collision frequencies $\lambda_{i,j}$ are typically expressed as functions of the species moments given in (6.1). To define these collision frequencies, [9] presents a recipe based on matching either momentum or energy relaxation rates of the Boltzmann equation, given either a differential cross section or a momentum transfer cross section. In this paper, we consider the case of hard sphere collisions, for which the momentum transfer cross section is independent of microscopic velocity and given by [2]

$$\sigma_{\text{MT}} = \pi \left(\frac{d_i + d_j}{2} \right)^2 = \pi d_{i,j}^2, \quad (6.10)$$

where d_i and d_j are reference diameters for the particles of species i and j ; reference diameters for several species are given in [2] and [4]. Following the recipe for matching energy relaxation rates from Sec. 4.3 of [9], we obtain the following collision frequencies for hard sphere interactions:

$$\lambda_{i,j}^{\text{HS}} = \frac{32\pi^2}{3(2\pi)^{3/2}} \frac{m_i m_j}{(m_i + m_j)^2} (d_i + d_j)^2 n_j \sqrt{\frac{T_i}{m_i} + \frac{T_j}{m_j}}. \quad (6.11)$$

While the formula in (6.11) is specific to the physically relevant case $d = 3$, collision frequen-

cies can be derived in arbitrary dimensions in a similar fashion.

The dependence of the collision frequencies on the temperature makes the BGK model considered here more complicated than many models considered in the literature which assume that $\lambda_{i,j}$ is a constant or depends only on n_j . This is largely due to the non-Lipschitz nature of the square root function near zero and the fact that the frequencies are time-dependent functions. While we focus on the hard spheres model in this paper for simplicity, the analysis provided here is readily applied to other models in the literature [2].

6.1.2 Moment Equations

The focus of the present work is the space-homogeneous multi-species BGK equation, obtained by removing the advection terms from (6.5). In this case $f_i = f_i(\mathbf{v}, t)$ satisfies

$$\frac{\partial f_i}{\partial t} = \frac{1}{\varepsilon} \sum_j \lambda_{i,j} (M_{i,j} - f_i), \quad i \in \{1, \dots, N\}. \quad (6.12)$$

Multiplication of (6.12) by m_i , $m_i \mathbf{v}$, and $m_i |\mathbf{v}|^2$, followed by integration with respect to $\mathbf{v}$, gives equations for the mass density, momentum density, and energy density of each species

$$\frac{d\rho_i}{dt} = 0, \quad (6.13a)$$

$$\frac{d(\rho_i \mathbf{u}_i)}{dt} = \frac{1}{\varepsilon} \sum_j A_{i,j} (\mathbf{u}_j - \mathbf{u}_i), \quad (6.13b)$$

$$\frac{dE_i}{dt} = \frac{1}{\varepsilon} \sum_j B_{i,j} \left(\frac{E_j}{n_j} - \frac{E_i}{n_i} \right) + \frac{1}{2\varepsilon} \sum_j B_{i,j} S_{i,j} (m_i - m_j), \quad (6.13c)$$

where the equations for $\rho_i \mathbf{u}_i$ and E_i are derived using the definitions for $\mathbf{u}_{i,j}$ and $T_{i,j}$ from (6.6) and the $N \times N$ matrices A , B , C , D , F , G , and S are given by

$$[A]_{i,j} = \frac{\rho_i \rho_j \lambda_{i,j} \lambda_{j,i}}{\rho_i \lambda_{i,j} + \rho_j \lambda_{j,i}}, \quad [B]_{i,j} = \frac{n_i n_j \lambda_{i,j} \lambda_{j,i}}{n_i \lambda_{i,j} + n_j \lambda_{j,i}}, \quad [S]_{i,j} = |\mathbf{u}_{i,j}|^2, \quad [C]_{i,j} = B_{i,j} S_{i,j}, \quad (6.14a)$$

$$[D]_{i,j} = \left(\sum_k A_{i,k} \right) \delta_{i,j}, \quad [F]_{i,j} = \left(\sum_k B_{i,k} \right) \delta_{i,j}, \quad [G]_{i,j} = \left(\sum_k C_{i,k} \right) \delta_{i,j}. \quad (6.14b)$$

According to (6.13a), each individual species mass density ρ_i (or, equivalently, number density n_i) is conserved (i.e., constant in time). However, only the *total* momentum and energy are

conserved. Indeed, summing (6.13b) and (6.13c) over all species yields:

$$\sum_{i=1}^N \frac{d(\rho_i \mathbf{u}_i)}{dt} = \mathbf{0} \quad \text{and} \quad \sum_{i=1}^N \frac{dE_i}{dt} = 0, \quad (6.15)$$

respectively.

In terms of the vectorized quantities

$$\mathbf{U} = [\mathbf{u}_1, \dots, \mathbf{u}_N]^\top \in \mathbb{R}^{N \times d} \quad \text{and} \quad \mathbf{E} = [E_1, \dots, E_N]^\top \in \mathbb{R}^N, \quad (6.16)$$

the equations in (6.13) take the form

$$P \frac{d\mathbf{U}}{dt} = -\frac{1}{\varepsilon} (D - A) \mathbf{U}, \quad (6.17a)$$

$$\frac{d\mathbf{E}}{dt} = -\frac{1}{\varepsilon} (F - B) Q^{-1} \mathbf{E} + \frac{1}{2\varepsilon} (G - C) M \mathbf{1}, \quad (6.17b)$$

where $P = \text{diag}\{\rho_k\} \in \mathbb{R}^{N \times N}$, $Q = \text{diag}\{n_k\} \in \mathbb{R}^{N \times N}$, and $M = \text{diag}\{m_k\} \in \mathbb{R}^{N \times N}$. To simplify the analysis later, it will be convenient to introduce variables

$$\mathbf{W} = P^{\frac{1}{2}} \mathbf{U} \in \mathbb{R}^{N \times d} \quad \text{and} \quad \boldsymbol{\xi} = Q^{-\frac{1}{2}} \mathbf{E} \in \mathbb{R}^N. \quad (6.18)$$

In terms of $\mathbf{W}$ and $\boldsymbol{\xi}$, (6.17) takes the form

$$\frac{d\mathbf{W}}{dt} = -\frac{1}{\varepsilon} Z \mathbf{W}, \quad (6.19a)$$

$$\frac{d\boldsymbol{\xi}}{dt} = -\frac{1}{\varepsilon} \widehat{Z} \boldsymbol{\xi} + \frac{1}{2\varepsilon} Q^{-\frac{1}{2}} (G - C) M \mathbf{1}, \quad (6.19b)$$

where

$$Z = P^{-\frac{1}{2}} (D - A) P^{-\frac{1}{2}} \in \mathbb{R}^{N \times N} \quad \text{and} \quad \widehat{Z} = Q^{-\frac{1}{2}} (F - B) Q^{-\frac{1}{2}} \in \mathbb{R}^{N \times N}. \quad (6.20)$$

6.2 Properties of the ODE system

The main result of this section is to show that there is a unique solution to the ODE system (6.17) for all physically meaningful initial conditions. When the collision frequencies are constant in time, such a result follows directly from the standard ODE theory for systems with Lipschitz dynamics. However, because the collision frequencies in (6.11) depend on $\sqrt{T_i}$,

only a local Lipschitz condition can be obtained for the right-hand side of (6.17). Thus the key step is to show that the species temperatures are bounded below away from zero. This bound will also be important for establishing exponential decay to steady-state in Section 6.3. We also establish upper and lower bounds on the bulk velocity components that will be used in the decay estimates.

6.2.1 Positivity of the Temperature

To make the following proofs easier to manage, we introduce the parameters

$$\alpha_{i,j} = \frac{\rho_i \lambda_{i,j}}{\rho_i \lambda_{i,j} + \rho_j \lambda_{j,i}} \quad \text{and} \quad \beta_{i,j} = \frac{n_i \lambda_{i,j}}{n_i \lambda_{i,j} + n_j \lambda_{j,i}}, \quad (6.21)$$

which satisfy the conditions $\alpha_{i,j} + \alpha_{j,i} = 1 = \beta_{i,j} + \beta_{j,i}$. With these parameters, the mixture values in (6.6) can be expressed as

$$\mathbf{u}_{i,j} = \alpha_{i,j} \mathbf{u}_i + \alpha_{j,i} \mathbf{u}_j \quad \text{and} \quad T_{i,j} = \beta_{i,j} T_i + \beta_{j,i} T_j + \frac{1}{d} m_i \alpha_{j,i} \beta_{i,j} |\mathbf{u}_i - \mathbf{u}_j|^2. \quad (6.22)$$

Lemma 6.2.1. *The temperatures T_i , $i \in \{1, \dots, N\}$, satisfy the ODE*

$$\frac{dT_i}{dt} = \frac{1}{\varepsilon} \sum_j \lambda_{i,j} \beta_{j,i} (T_j - T_i) + \frac{1}{\varepsilon d} \sum_j \lambda_{i,j} m_i \alpha_{j,i} (\alpha_{j,i} + \beta_{i,j}) |\mathbf{u}_i - \mathbf{u}_j|^2. \quad (6.23)$$

Proof. The proof is a direct calculation. From (6.13b) and the fact that $A_{i,j} = \alpha_{j,i} \rho_i \lambda_{i,j}$, it follows that

$$\rho_i \frac{d}{dt} |\mathbf{u}_i|^2 = \frac{1}{\varepsilon} \sum_j \rho_i \lambda_{i,j} (2\alpha_{j,i} \langle \mathbf{u}_i, \mathbf{u}_j \rangle - 2\alpha_{j,i} |\mathbf{u}_i|^2), \quad (6.24)$$

where $\langle \cdot, \cdot \rangle$ denotes the standard Euclidean inner product on $\mathbb{R}^d$. The equation in (6.23) can be obtained by (i) using (6.2) to express (6.13c) in terms of the temperatures and velocities; (ii) invoking (6.24) to eliminate the time derivative of $|\mathbf{u}_i|^2$; (iii) using the formula in (6.22) to replace $\mathbf{u}_{i,j}$ wherever it appears; and (iv) applying the elementary relations $\alpha_{j,i}^2 = \alpha_{j,i} - \alpha_{i,j} \alpha_{j,i} = 2\alpha_{j,i} + \alpha_{i,j}^2 - 1$. ■

Theorem 6.2.2. *Suppose that for some $t_f > 0$, there exists a local solution $\mathbf{U} \in C^1([0, t_f]; \mathbb{R}^{N \times d})$ and $\mathbf{E} \in C^1([0, t_f]; \mathbb{R}^N)$ to the system given by (6.17), with initial conditions $\mathbf{U}(0)$ and $\mathbf{E}(0)$ such that $T_i(0) > 0$ for all $i \in \{1, \dots, N\}$. Then $T_i(t) \geq T_{\min} := \min_k \{T_k(0)\}$ for all $t \in [0, t_f]$ and all $i \in \{1, \dots, N\}$.*

Proof. The proof relies on an integrating factor technique that yields a Grönwall-type estimate. Define a lower temperature envelope $T_*(t) = \min_k \{T_k(t)\}$. From (6.23) it follows that, for $i \in \{1, \dots, N\}$,

$$\frac{dT_i}{dt} + c_i(t)T_i(t) \geq c_i(t)T_*(t), \quad \text{where} \quad c_i(t) = \frac{1}{\varepsilon} \sum_j \lambda_{i,j}(t)\beta_{j,i}(t). \quad (6.25)$$

Let $a_i(t) = \int_0^t c_i(s) ds$. Then $a_i'(t) = c_i(t)$ and

$$\frac{d}{dt} (e^{a_i(t)} T_i(t)) = e^{a_i(t)} \left[\frac{d}{dt} T_i(t) + c_i(t) T_i(t) \right] \stackrel{(6.25)}{\geq} e^{a_i(t)} c_i(t) T_*(t) = \frac{d}{dt} (e^{a_i(t)} T_*(t)). \quad (6.26)$$

For each $\gamma > 0$, define the set

$$S_\gamma = \left\{ t \in [0, t_f] \mid T_*(t) < \frac{1}{1+\gamma} T_*(0) \right\}. \quad (6.27)$$

We now argue by contradiction. Assume that S_γ is nonempty and let $t_* = \inf S_\gamma$. Then by definition,

$$T_*(t_*) \leq \frac{1}{1+\gamma} T_*(0) < T_*(0). \quad (6.28)$$

However, because T_i is continuous on $[0, t_f]$, it follows that $T_*(t) \geq T_*(t_*)$ for every $t \in [0, t_*]$ which, along with (6.26), implies that

$$\frac{d}{dt} (e^{a_i(t)} T_i(t)) \geq \frac{d}{dt} (e^{a_i(t)} T_*(t_*)), \quad \forall t \in [0, t_*]. \quad (6.29)$$

Integrating both sides of (6.29) from $[0, t_*]$ and then multiplying the result by $e^{-a_i(t_*)} \geq 0$ gives

$$T_i(t_*) \geq T_i(0)e^{-a_i(t_*)} + T_*(t_*) (1 - e^{-a_i(t_*)}) \geq T_*(0)e^{-a_i(t_*)} + T_*(t_*) (1 - e^{-a_i(t_*)}). \quad (6.30)$$

Define i_* such that $T_{i_*}(t_*) = T_*(t_*)$. If $i = i_*$, then (6.30) becomes

$$T_*(t_*) \equiv T_{i_*}(t_*) \geq T_*(0)e^{-a_{i_*}(t_*)} + T_*(t_*) (1 - e^{-a_{i_*}(t_*)}), \quad (6.31)$$

which, after some simple algebra, implies that $T_*(t_*) \geq T_*(0)$. However this results contradicts (6.28) which means that no such t_* exists and the set S_γ must be empty. As a

consequence, $T_*(t) \geq \frac{1}{1+\gamma} T_*(0)$ for all $t \in [0, t_f]$. Since $\gamma > 0$ is arbitrary, the limit $\gamma \rightarrow 0$ yields $T_*(t) \geq T_*(0) = T_{\min}$ for all $t \in [0, t_f]$. The proof is complete. ■

6.2.2 Velocity Bounds

For each component $j \in \{1, \dots, d\}$, define the lower and upper velocity envelopes

$$\underline{u}_k(t) = \min_j \{\mathbf{U}_{j,k}(t)\}, \quad \bar{u}_k(t) = \max_j \{\mathbf{U}_{j,k}(t)\}, \quad (6.32)$$

and the components of the vector $\mathbf{u}_{\max} \in \mathbb{R}^d$

$$[\mathbf{u}_{\max}(t)]_k = \max \{|\underline{u}_k(t)|, |\bar{u}_k(t)|\}. \quad (6.33)$$

Theorem 6.2.3. *Suppose that the assumptions of Theorem 6.2.2 hold. Then for each $j \in \{1, \dots, d\}$, the velocity components follow the inequalities*

$$\underline{u}_k(0) \leq \underline{u}_k(t) \leq \mathbf{U}_{i,k}(t) \leq \bar{u}_k(t) \leq \bar{u}_k(0) \quad (6.34)$$

for all $t \in [0, t_f]$ and all $i \in \{1, \dots, N\}$. Further, for any $t \in [0, t_f]$, and all $i \in \{1, \dots, N\}$,

$$\|\mathbf{u}_i(t)\|_2 \leq \|\mathbf{u}_{\max}(t)\|_2 \leq \|\mathbf{u}_{\max}(0)\|_2 =: u_{\max}. \quad (6.35)$$

Proof. The proof of the inequalities in (6.34) are similar to the proof of the temperature lower bound given in Theorem 6.2.2. From the velocity equation, (6.13b),

$$\frac{d}{dt} \mathbf{U}_{i,k} + \left(\frac{1}{\varepsilon} \sum_j \lambda_{i,j} \alpha_{j,i} \right) \mathbf{U}_{i,k} = \frac{1}{\varepsilon} \sum_j \lambda_{i,j} \alpha_{j,i} \mathbf{U}_{j,k}, \quad (6.36)$$

we obtain, using the definitions $\mathbf{U}_{j,k}(t) \leq \bar{u}_k(t)$, and $c_i = \frac{1}{\varepsilon} \sum_j \lambda_{i,j} \alpha_{j,i}$,

$$\frac{d}{dt} \mathbf{U}_{i,k} + c_i(t) \mathbf{U}_{i,k} \leq c_i(t) \bar{u}_k(t). \quad (6.37)$$

From this point the proof of the upper bound in (6.34) follows that of Theorem 6.2.2, with (6.37) being the analog of (6.25). The proof of the lower bound is even more similar.

Finally, the bound in (6.35) follows from (6.34). Specifically, if a , b , and c are real-valued

scalars such that $a \leq b \leq c$, then $|b| \leq \max\{|a|, |c|\}$. Hence

$$\|\mathbf{u}_i\|_2^2 = \sum_{k=1}^d \mathbf{u}_{i,k}^2 \leq \sum_{k=1}^d \max\{|\underline{u}_k(0)|, |\bar{u}_k(0)|\}^2 = \|\mathbf{u}_{\max}(0)\|_2^2 = u_{\max}^2. \quad (6.38)$$

■

6.2.3 Existence and Uniqueness

According to (6.13a), the number densities n_i and mass densities ρ_i are non-negative constants in time:

$$n_i(0) = n_i(t), \quad \rho_i(0) = \rho_i(t) = m_i n_i(t), \quad \forall i \in \{1, \dots, N\}, \quad t \in [0, t_f]. \quad (6.39)$$

To avoid any degenerate cases, we assume that $n_{\min} := \min_j n_j(0) > 0$ and $\rho_{\min} := \min_j \rho_j(0) > 0$. Then, given any velocity-energy pair $(\mathbf{u}_i, E_i)$, the associated temperature T_i is given by (cf. (6.1))

$$T_i = \vartheta_i(\mathbf{u}_i, E_i) := \frac{2}{dn_i} E_i - \frac{m_i}{d} |\mathbf{u}_i|^2, \quad (6.40)$$

and the set of all realizable velocity and energy states is given by

$$\mathcal{R} = \left\{ (\mathbf{U}, \mathbf{E}) \in \mathbb{R}^{N \times d} \times \mathbb{R}^N \mid \vartheta_i(\mathbf{u}_i, E_i) \geq 0, \forall i \in \{1, \dots, N\} \right\}, \quad (6.41)$$

where ϑ_i is defined in (6.40). Given positive scalars $T_{\min} \in (0, \infty)$ and $E_{\text{tot}} \in (0, \infty)$, let

$$\mathcal{D}(T_{\min}, E_{\text{tot}}) = \left\{ (\mathbf{U}, \mathbf{E}) \in \mathcal{R} \mid \sum_{i=1}^N E_i = E_{\text{tot}} \quad \text{and} \quad \vartheta_i(\mathbf{u}_i, E_i) \geq T_{\min}, \forall i \in \{1, \dots, N\} \right\}. \quad (6.42)$$

Recall that by Theorem 6.2.3, $|\mathbf{u}_i| \leq u_{\max} =: \|\mathbf{u}_{\max}(0)\|_2 \leq \sqrt{\frac{2E_{\text{tot}}}{\rho_{\min}}}$ for all $(\mathbf{U}, \mathbf{E}) \in \mathcal{D}(T_{\min}, E_{\text{tot}})$. Thus $\mathcal{D}(T_{\min}, E_{\text{tot}})$ is a closed and bounded subset of $\mathbb{R}^{N \times d} \times \mathbb{R}^N$.

Lemma 6.2.4. *Let $T_{\min} \in (0, \infty)$ and $E_{\text{tot}} \in (0, \infty)$ be given. Suppose that for some $t_f > 0$, there exists a local solution $\mathbf{U} \in C^1([0, t_f]; \mathbb{R}^{N \times d})$ and $\mathbf{E} \in C^1([0, t_f]; \mathbb{R}^N)$ of the system in (6.17) with initial condition $(\mathbf{U}^0, \mathbf{E}^0) \in \mathcal{D}(T_{\min}, E_{\text{tot}})$. Then $(\mathbf{U}(t), \mathbf{E}(t)) \in \mathcal{D}(T_{\min}, E_{\text{tot}})$ for all $t \in [0, t_f]$. In particular, $T_{\min} \leq T_i(t) \leq T_{\max} := \frac{2}{dn_{\min}} E_{\text{tot}}$ for all $t \in [0, t_f]$.*

Proof. To conclude that $(\mathbf{U}(t), \mathbf{E}(t)) \in \mathcal{D}$ for all $t \in [0, t_f]$, two conditions must be satisfied.

The first condition:

$$\sum_{i=1}^N E_i(t) = \sum_{i=1}^N E_i(0) = E_{\text{tot}}, \quad \forall t \in [0, t_f], \quad (6.43)$$

follows immediately from (6.15), and the upper bound $T_i(t) \leq T_{\max}$ follows:

$$T_i(t) = \frac{2}{dn_i} E_i \leq \frac{2}{dn_{\min}} E_{\text{tot}} = T_{\max}. \quad (6.44)$$

The second condition:

$$T_{\min} \leq T_i(t) = \vartheta(\mathbf{u}_i(t), E_i(t)), \quad \forall t \in [0, t_f], \quad (6.45)$$

is a direct consequence of Theorem 6.2.2. ■

Theorem 6.2.5. *Suppose that $m_i > 0$ and $n_i > 0$ for all $i \in \{1, \dots, N\}$. Then for any $(\mathbf{U}^0, \mathbf{E}^0) \in \text{int } \mathcal{R}$, there exists a global, unique solution $(\mathbf{U}, \mathbf{E}) \in C^1([0, \infty); \mathcal{R})$ of the system (6.17) with the initial conditions $(\mathbf{U}(0), \mathbf{E}(0)) = (\mathbf{U}^0, \mathbf{E}^0)$. Moreover, the associated temperatures $T_i(t) = \vartheta_i(\mathbf{u}_i(t), E_i(t))$ are bounded below by their initial values; that is*

$$\min_i T_i(t) \geq \min_i T_i(0), \quad \forall t > 0. \quad (6.46)$$

Proof. The system (6.17) can be written as

$$\frac{d}{dt} \begin{pmatrix} \mathbf{U} \\ \mathbf{E} \end{pmatrix} = \mathbf{f}(\mathbf{U}, \mathbf{E}) \quad \text{where} \quad \mathbf{f}(\mathbf{U}, \mathbf{E}) := \begin{pmatrix} -\frac{1}{\epsilon} P^{-1}(D - A)\mathbf{U} \\ -\frac{1}{\epsilon}(F - B)Q^{-1}\mathbf{E} + \frac{1}{2\epsilon}(G - C)M\mathbf{1} \end{pmatrix}. \quad (6.47)$$

Let $T_{\min} = \min_i T_i(0)$ and $E_{\text{tot}} = \sum_i E_i(0)$. Since $(\mathbf{U}, \mathbf{E}) \in \text{int } \mathcal{R}$, it follows that $T_{\min} > 0$ and $E_{\text{tot}} > 0$. Moreover, there exists an $\epsilon > 0$ such that the closed ϵ -neighborhood of $\mathcal{D}(T_{\min}, E_{\text{tot}})$, denoted by $\mathcal{D}_\epsilon(T_{\min}, E_{\text{tot}})$, is contained in $\text{int } \mathcal{R}$, i.e., $\mathcal{D}_\epsilon(T_{\min}, E_{\text{tot}}) \subset \text{int } \mathcal{R}$. In particular, $\min_i T_i = \min_i \vartheta(\mathbf{u}_i, E_i)$ is bounded below on $\mathcal{D}_\epsilon(T_{\min}, E_{\text{tot}})$. Hence, $\mathbf{f}$ is Lipschitz on $\mathcal{D}_\epsilon(T_{\min}, E_{\text{tot}})$ with Lipschitz constant $L > 0$ and bound $\|\mathbf{f}(\mathbf{U}, \mathbf{E})\|_2 \leq \mathcal{M} < \infty$. Since $(\mathbf{U}^0, \mathbf{E}^0) \in \mathcal{D}$, there is a number $\beta = \beta(\epsilon) > 0$, independent of the point $(\mathbf{U}^0, \mathbf{E}^0)$, such that the closed ball $\overline{B}((\mathbf{U}, \mathbf{E}), \beta) \subset \mathcal{D}_\epsilon$. Appealing to the Picard-Lindelöf Theorem B.0.1, there exists a unique solution on the interval $[0, \delta]$, where $\delta = \min(\frac{1}{2L}, \frac{\beta}{\mathcal{M}})$. By Lemma 6.2.4, $(\mathbf{U}(t), \mathbf{E}(t)) \in \mathcal{D}$, for all $t \in [0, \delta]$. Since $(\mathbf{U}(\delta), \mathbf{E}(\delta)) \in \mathcal{D}$, we can apply the Picard-Lindelöf Theorem to (6.17) again, but with the initial condition $= (\mathbf{U}(\delta), \mathbf{E}(\delta))$, and thereby extend

the unique solution to the interval $[0, 2\delta]$. Again, $(\mathbf{U}(2\delta), \mathbf{E}(2\delta)) \in \mathcal{D}$, and the process can be continued indefinitely, with, crucially, no degradation of $\delta > 0$, since $\mathcal{M}$ and L cannot grow larger and β does not shrink. The result follows. $\blacksquare$

6.3 Long-time Behavior

Having proven the existence of unique, global, physically realizable solutions for (6.17), we now seek to characterize the asymptotic behavior of those solutions. In particular, we want to show that solutions converge to

$$\mathbf{U}^\infty(t) = \mathbf{1}(\mathbf{u}^\infty)^\top \quad \text{and} \quad \mathbf{E}^\infty(t) = \frac{|\mathbf{u}^\infty(t)|^2}{2} P\mathbf{1} + \frac{dT^\infty(t)}{2} Q\mathbf{1}, \quad (6.48)$$

where

$$\mathbf{u}^\infty(t) = \frac{\sum_i \rho_i \mathbf{u}_i(t)}{\sum_i \rho_i} \quad \text{and} \quad T^\infty(t) = \frac{\sum_i n_i T_i(t)}{\sum_i n_i} + \frac{\sum_i \rho_i (|\mathbf{u}_i(t)|^2 - |\mathbf{u}^\infty(t)|^2)}{d \sum_i n_i}. \quad (6.49)$$

are constant functions of time.

Proposition 6.3.1. *Suppose that $(\mathbf{U}, \mathbf{E}) \in C^1([0, \infty); \mathcal{R})$ is a unique global-in-time solution to the system (6.17a)-(6.17b) with assumptions as given in Theorem 6.2.5. Further let $\mathbf{T} \in C^1([0, \infty); \mathbb{R}^N)$ be the associated temperature vector, whose components satisfy (6.23). Then $\mathbf{u}^\infty$, T^∞ , and $\mathbf{E}^\infty$ are time invariant quantities, and $T^\infty > 0$.*

Proof. By conservation of mass, (6.13a), and total momentum, (6.15), $\sum_i \rho_i$ and $\sum_i \rho_i \mathbf{u}_i$ are constant. Thus $\mathbf{u}^\infty$ is time invariant. To show that T^∞ is constant in time, multiply (6.49) by the expression $\frac{d}{2} \sum_i n_i$, which is constant by (6.13a), to obtain

$$\frac{d}{2} \left(\sum_i n_i \right) T^\infty = \sum_i \left[\frac{d}{2} n_i T_i + \frac{1}{2} \rho_i |\mathbf{u}_i|^2 \right] - \frac{1}{2} \left(\sum_i \rho_i \right) |\mathbf{u}^\infty|^2 = \sum_i E_i - \frac{1}{2} \left(\sum_i \rho_i \right) |\mathbf{u}^\infty|^2. \quad (6.50)$$

The first term on the right-hand side of (6.50) is constant by conservation of total energy, (6.15); the second term is constant by conservation of species mass, (6.13a), and the fact that $\mathbf{u}^\infty$ is constant. Thus T^∞ is time invariant. It follows immediately from (6.48) that $\mathbf{E}^\infty$ is also time invariant.

Since $(\mathbf{U}(t), \mathbf{E}(t)) \in \mathcal{R}$ for all $t > 0$, it follows that $T_i(t) = \vartheta(\mathbf{u}_i(t), E_i(t)) \geq 0$ for all

$t > 0$. Thus to show that $T^\infty > 0$, it is sufficient to show that $\sum_i \rho_i (|\mathbf{u}_i|^2 - |\mathbf{u}^\infty|^2) \geq 0$. Let $r_i = \frac{\rho_i}{\sum_k \rho_k} > 0$. Then $\sum_i r_i = 1$ and $|\mathbf{u}^\infty|^2 = (\sum_i r_i \mathbf{u}_i)^\top (\sum_j r_j \mathbf{u}_j) = \sum_i \sum_j r_i r_j \mathbf{u}_i^\top \mathbf{u}_j$. Therefore

$$\sum_i \rho_i (|\mathbf{u}_i|^2 - |\mathbf{u}^\infty|^2) = \sum_k \rho_k \left[\sum_{i,j} r_i r_j |\mathbf{u}_i|^2 - \sum_{i,j} r_i r_j \mathbf{u}_i^\top \mathbf{u}_j \right] = \frac{1}{2} \sum_k \rho_k \sum_{i,j} r_i r_j |\mathbf{u}_i - \mathbf{u}_j|^2 \geq 0. \quad (6.51)$$

■

Remark 6.3.2. The quantities $\mathbf{u}^\infty$ and T^∞ were referred to as the mixture mean velocity and mixture temperature, respectively, in [8]. However, in this paper, we reserve these terms for the quantities $\mathbf{u}_{i,j}$ and $T_{i,j}$, respectively.

The convergence of $\mathbf{u}_i$, T_i , and $\mathbf{E}$ as $t \rightarrow \infty$ is established by the following result, the proof of which is the focus of the rest of this section.

Theorem 6.3.3. Under the assumptions of Proposition 6.3.1, for all $i \in \{1, \dots, N\}$,

$$\lim_{t \rightarrow \infty} \mathbf{u}_i(t) = \mathbf{u}^\infty, \quad \lim_{t \rightarrow \infty} T_i(t) = T^\infty, \quad \text{and} \quad \lim_{t \rightarrow \infty} \mathbf{E}(t) = \mathbf{E}^\infty, \quad (6.52)$$

and bounds for the decay rates of the bulk velocities and energies are given by

$$\|\mathbf{u}_i(t) - \mathbf{u}^\infty\|_2 \leq C_U \exp\left(-\frac{Z_{\min}}{\varepsilon} t\right), \quad i \in \{1, \dots, N\} \quad (6.53)$$

$$\|\mathbf{E}(t) - \mathbf{E}^\infty\|_2 \leq C_1 e^{-\frac{\hat{Z}_{\min}}{\varepsilon} t} + C_2 \frac{e^{-\frac{Z_{\min}}{\varepsilon} t} - e^{-\frac{\hat{Z}_{\min}}{\varepsilon} t}}{\hat{Z}_{\min} - Z_{\min}}, \quad (6.54)$$

where $Z_{\min} > 0$ and $\hat{Z}_{\min} > 0$ are lower bounds on the positive eigenvalues of Z and $\hat{Z}$, respectively, and $C_U > 0$, $C_1 > 0$, and $C_2 > 0$ are constants depending only on the initial conditions of the system.

6.3.1 Null Spaces of $D - A$, Z , $F - B$, and $\hat{Z}$

To prove Theorem 6.3.3, we first characterize the null spaces of the matrices $D - A$, Z , $F - B$, and $\hat{Z}$, which are defined in (6.14), and (6.20).

Lemma 6.3.4. The matrices $D - A$ and Z are symmetric positive semi-definite (SPSD), each with a one dimensional null space. In particular, $\mathcal{N}(D - A) = \text{span}(\{\mathbf{1}\})$ and $\mathcal{N}(Z) =$

$\text{span}\left(\left\{P^{\frac{1}{2}}\mathbf{1}\right\}\right)$. Moreover, the null space and range space of Z are invariant with respect to time.

Proof. Clearly $D - A$ and Z are symmetric by inspection. For any $\mathbf{y} \in \mathbb{R}^N$,

$$\mathbf{y}^\top (D - A) \mathbf{y} = \frac{1}{2} \sum_{i,j} A_{i,j} (y_i - y_j)^2 \geq 0. \quad (6.55)$$

Moreover, since $A_{i,j} > 0$, $\frac{1}{2} \sum_{i,j} A_{i,j} (y_i - y_j)^2 = 0$ if and only if $y_i = y_j$ for all i and j , in which case $\mathbf{y} = c\mathbf{1}$ for some $c \in \mathbb{R}$. Thus $D - A$ is SPSPD, with a one-dimensional null space spanned by the eigenvector $\mathbf{1}$. Similarly $\mathbf{y}^\top Z \mathbf{y} = (P^{-\frac{1}{2}}\mathbf{y})^\top (D - A) (P^{-\frac{1}{2}}\mathbf{y}) \geq 0$, and $Z\mathbf{y} = \mathbf{0} \iff P^{-\frac{1}{2}}\mathbf{y} = c\mathbf{1} \iff \mathbf{y} = cP^{\frac{1}{2}}\mathbf{1}$, for some $c \in \mathbb{R}$. Thus Z is SPSPD, with a one-dimensional null space spanned by the eigenvector $P^{\frac{1}{2}}\mathbf{1}$. By conservation of mass, (6.13a), $P^{\frac{1}{2}}\mathbf{1}$ is independent of time. Thus the null space of Z is invariant with respect to time. The symmetry of Z implies that $\mathcal{R}(Z) = \mathcal{N}(Z)^\perp$ is also invariant. ■

Lemma 6.3.5. *The matrices $F - B$ and $\hat{Z}$ are symmetric positive semi-definite (SPSPD). In particular $\mathcal{N}(F - B) = \text{span}(\{\mathbf{1}\})$ and $\mathcal{N}(\hat{Z}) = \text{span}\left(\left\{Q^{\frac{1}{2}}\mathbf{1}\right\}\right)$. Moreover, the null space and range space of $\hat{Z}$ are invariant with respect to time.*

Proof. The proof follows closely that of Lemma 6.3.4. Details are left to the reader. ■

6.3.2 Velocity Relaxation Proof

Theorem 6.3.6. *Under the assumptions of Proposition 6.3.1, $\lim_{t \rightarrow \infty} \mathbf{u}_i(t) = \mathbf{u}^\infty$ and*

$$\|\mathbf{u}_i(t) - \mathbf{u}^\infty\|_2 \leq C_U \exp\left(-\frac{z_{\min}}{\varepsilon} t\right), \quad i = 1, \dots, N, \quad (6.56)$$

where $C_U > 0$ is a constant that depends on the initial conditions of the system and $z_{\min} > 0$ is a lower bound on the positive eigenvalues of Z .

Proof. The proof is based on (6.19a) which is equivalent to (6.17a). Let $\mathbf{W}^\infty = P^{\frac{1}{2}}\mathbf{U}^\infty$, where $\mathbf{U}^\infty = \mathbf{1}(\mathbf{u}^\infty)^\top$ is given in (6.48). Then Lemma 6.3.4 implies that $Z\mathbf{W}^\infty = \mathbf{0} \in \mathbb{R}^{N \times d}$. Moreover, since P and $\mathbf{U}^\infty$ are independent of time, so too is $\mathbf{W}^\infty$. Thus in terms of $\widetilde{\mathbf{W}} := \mathbf{W} - \mathbf{W}^\infty$, (6.19a) takes the form

$$\frac{d\widetilde{\mathbf{W}}}{dt} = -\frac{1}{\varepsilon} Z \widetilde{\mathbf{W}}. \quad (6.57)$$

Write the element $\widetilde{\mathbf{W}}$ as a sum of the null and range space components with respect to Z : $\widetilde{\mathbf{W}} = \widetilde{\mathbf{W}}_{\mathcal{R}} + \widetilde{\mathbf{W}}_{\mathcal{N}}$. As shown in Appendix D, $\widetilde{\mathbf{W}}_{\mathcal{N}} = 0$. Hence $\widetilde{\mathbf{W}} = V\mathbf{A}$, where the columns of $V = [\mathbf{v}_1, \dots, \mathbf{v}_{N-1}] \in \mathbb{R}^{N \times (N-1)}$ are a set of orthonormal eigenvectors of Z that span $\mathcal{R}(Z)$ and $\mathbf{A} = [\mathbf{a}_1, \dots, \mathbf{a}_d] \in \mathbb{R}^{(N-1) \times d}$. Using this formulation, the ODE system (6.57) becomes

$$\frac{d\mathbf{A}}{dt} = -\frac{1}{\varepsilon} V^\top Z V \mathbf{A}. \quad (6.58)$$

The eigenvalues of $V^\top Z V$ are strictly positive and bounded below by $z_{\min} := N \max\{\rho_k\}^{-1} \min_{i,j} A_{i,j}$ (see Appendix C). Thus $(\mathbf{A}, V^\top Z V \mathbf{A})_{\text{F}} \geq z_{\min} \|\mathbf{A}\|_{\text{F}}^2$, where $\|\cdot\|_{\text{F}}$ is the Frobenius norm. Let $s(t) = \|\mathbf{A}\|_{\text{F}}^2$. Then, using (6.58),

$$\frac{1}{2} \frac{ds}{dt} = \left(\mathbf{A}, \frac{d\mathbf{A}}{dt} \right)_{\text{F}} = \left(\mathbf{A}, -\frac{1}{\varepsilon} V^\top Z V \mathbf{A} \right)_{\text{F}} \leq -\frac{z_{\min}}{\varepsilon} \|\mathbf{A}\|_{\text{F}}^2 = -\frac{z_{\min}}{\varepsilon} s \quad (6.59)$$

which implies that

$$\|\mathbf{A}(t)\|_{\text{F}}^2 \leq \|\mathbf{A}(0)\|_{\text{F}}^2 e^{-\frac{2z_{\min}}{\varepsilon} t}. \quad (6.60)$$

Since $V^\top V = I_{N-1}$, it follows then $\|\widetilde{\mathbf{W}}\|_{\text{F}}^2 = \|V\mathbf{A}\|_{\text{F}}^2 = \|\mathbf{A}\|_{\text{F}}^2$. Thus, (6.60) becomes

$$\|\mathbf{W}(t) - \mathbf{W}^\infty\|_{\text{F}}^2 \leq \|\mathbf{W}(0) - \mathbf{W}^\infty\|_{\text{F}}^2 e^{-\frac{2z_{\min}}{\varepsilon} t}. \quad (6.61)$$

Since $\mathbf{W} = P^{\frac{1}{2}} \mathbf{U}$, it follows that

$$\|\mathbf{W} - \mathbf{W}^\infty\|_{\text{F}}^2 = \sum_{k=1}^d \sum_{j=1}^N [\rho_j ((\mathbf{u}_j)_k - \mathbf{u}_k^\infty)^2] = \sum_{j=1}^N \rho_j \|\mathbf{u}_j - \mathbf{u}^\infty\|_2^2 \geq \min\{\rho_k\} \|\mathbf{u}_i - \mathbf{u}^\infty\|_2^2, \quad (6.62)$$

for any $i \in \{1, \dots, N\}$. Thus, (6.61) gives the decay bound

$$\|\mathbf{u}_i(t) - \mathbf{u}^\infty\|_2^2 \leq \frac{1}{\min\{\rho_k\}} \|\mathbf{W}(t) - \mathbf{W}^\infty\|_{\text{F}}^2 \leq \frac{1}{\min\{\rho_k\}} \|\mathbf{W}^0 - \mathbf{W}^\infty\|_{\text{F}}^2 e^{-\frac{2z_{\min}}{\varepsilon} t} =: C_{\mathbf{U}}^2 e^{-\frac{2z_{\min}}{\varepsilon} t}, \quad (6.63)$$

and $\lim_{t \rightarrow \infty} \mathbf{u}_i = \mathbf{u}^\infty$, as desired. ■

6.3.3 Energy Relaxation Proof

Theorem 6.3.7. *With the same assumptions as in Proposition 6.3.1, $\lim_{t \rightarrow \infty} \mathbf{E}(t) = \mathbf{E}^\infty$ and*

$$\|\mathbf{E}(t) - \mathbf{E}^\infty\|_2 \leq C_1 e^{-\frac{\hat{Z}_{\min}}{\varepsilon} t} + C_2 \frac{e^{-\frac{Z_{\min}}{\varepsilon} t} - e^{-\frac{\hat{Z}_{\min}}{\varepsilon} t}}{\hat{Z}_{\min} - Z_{\min}}, \quad (6.64)$$

where $C_1 > 0$ and $C_2 > 0$ are constants that depend on the initial conditions of the system (6.17) and $\hat{Z}_{\min}$ is a lower bound on the positive eigenvalues of $\hat{Z}$.

Proof. The proof is based on (6.19), which is equivalent to (6.17). Set $\boldsymbol{\xi}^\infty := Q^{-\frac{1}{2}} \mathbf{E}^\infty$, and note that

$$\hat{Z} \boldsymbol{\xi}^\infty = \frac{|\mathbf{u}^\infty|^2}{2} Q^{-\frac{1}{2}} (F - B) M \mathbf{1}. \quad (6.65)$$

Since Q and $\mathbf{E}^\infty$ are time invariant, so too is $\boldsymbol{\xi}^\infty$. Hence $\tilde{\boldsymbol{\xi}} := \boldsymbol{\xi} - \boldsymbol{\xi}^\infty$ satisfies the ODE

$$\frac{d\tilde{\boldsymbol{\xi}}}{dt} = -\frac{1}{\varepsilon} \hat{Z} \tilde{\boldsymbol{\xi}} + \frac{1}{2\varepsilon} Q^{-\frac{1}{2}} [(G - C) - |\mathbf{u}^\infty|^2 (F - B)] M \mathbf{1}. \quad (6.66)$$

Consider the null and range space components of $\tilde{\boldsymbol{\xi}}$ with respect to $\hat{Z}$: $\tilde{\boldsymbol{\xi}} = \tilde{\boldsymbol{\xi}}_{\mathcal{R}} + \tilde{\boldsymbol{\xi}}_{\mathcal{N}}$. It can be shown that $\tilde{\boldsymbol{\xi}}_{\mathcal{N}} \equiv \mathbf{0}$. (See Appendix D.) Therefore $\tilde{\boldsymbol{\xi}} = \tilde{\boldsymbol{\xi}}_{\mathcal{R}} = \hat{V} \mathbf{b}$, where $\hat{V} = [\hat{\mathbf{v}}_1, \dots, \hat{\mathbf{v}}_{N-1}] \in \mathbb{R}^{N \times (N-1)}$ is a matrix whose columns form an orthonormal basis for the range of $\hat{Z}$ and $\mathbf{b} = \hat{V}^\top \tilde{\boldsymbol{\xi}} \in \mathbb{R}^{N-1}$. Multiplication of (6.66) by $\mathbf{b}^\top \hat{V}^\top$ gives

$$\frac{1}{2} \frac{d}{dt} \|\mathbf{b}\|_2^2 = -\frac{1}{\varepsilon} \mathbf{b}^\top \hat{V}^\top \hat{Z} \hat{V} \mathbf{b} + \frac{1}{2\varepsilon} \mathbf{b}^\top \hat{V}^\top Q^{-\frac{1}{2}} [(G - C) - |\mathbf{u}^\infty|^2 (F - B)] M \mathbf{1}. \quad (6.67)$$

The eigenvalues of $\hat{V}^\top \hat{Z} \hat{V}$ are strictly positive and bounded below by $\hat{Z}_{\min} := N \max\{n_k\}^{-1} \min_{i,j} B_{i,j}$ (see Appendix C), which gives the bound $\mathbf{b}^\top \hat{V}^\top \hat{Z} \hat{V} \mathbf{b} \geq \hat{Z}_{\min} \|\mathbf{b}\|_2^2$ for any $\mathbf{b} \in \mathbb{R}^{N-1}$. With this bound and an application of the Cauchy-Schwarz inequality, (6.67) becomes

$$\frac{d}{dt} \|\mathbf{b}\|_2^2 \leq -\frac{2\hat{Z}_{\min}}{\varepsilon} \|\mathbf{b}\|_2^2 + \frac{1}{\varepsilon} \|\mathbf{b}\|_2 \frac{1}{\min\{n_k\}^{\frac{1}{2}}} \left\| [(G - C) - |\mathbf{u}^\infty|^2 (F - B)] M \mathbf{1} \right\|_2. \quad (6.68)$$

The next step is to bound the source term in (6.68). Standard norm inequalities and the definitions in (6.14), and Appendix C, give

$$\left\| [(G - C) - |\mathbf{u}^\infty|^2 (F - B)] M \mathbf{1} \right\|_2 \leq 2 \max\{m_k\} \sum_{i=1}^N \sum_{\substack{j=1 \\ j \neq i}}^N B_{\max} \left| |\mathbf{u}_{i,j}|^2 - |\mathbf{u}^\infty|^2 \right|. \quad (6.69)$$

Meanwhile, to bound $|S_{ij} - |\mathbf{u}^\infty|^2|$, we use (i) the decay bound in (6.56), (ii) the definitions of $\mathbf{u}_{i,j}$ and $\mathbf{u}^\infty$, (iii) the triangle inequality and Cauchy Schwarz inequality, (iv) the fact that $\alpha_{i,j} + \alpha_{j,i} = 1$, and (v) the fact that $|\mathbf{u}_{i,j}| \leq \max_{i,j}\{|\mathbf{u}_i|, |\mathbf{u}_j|\} \leq u_{\max}$ to obtain

$$\begin{aligned} ||\mathbf{u}_{i,j}|^2 - |\mathbf{u}^\infty|^2| &= |(\mathbf{u}_{i,j} + \mathbf{u}^\infty)^\top (\mathbf{u}_{i,j} - \mathbf{u}^\infty)| \leq 2u_{\max} \|\mathbf{u}_{i,j} - \mathbf{u}^\infty\|_2 \\ &\leq 2u_{\max} (\alpha_{i,j} \|\mathbf{u}_i - \mathbf{u}^\infty\|_2 + \alpha_{j,i} \|\mathbf{u}_j - \mathbf{u}^\infty\|_2) \leq 2C_U u_{\max} \exp\left(-\frac{z_{\min}}{\varepsilon} t\right). \end{aligned} \quad (6.70)$$

Together (6.69) and (6.70) give

$$\|[(G - C) - |\mathbf{u}^\infty|^2(F - B)] M\mathbf{1}\|_2 \leq 4N(N-1)C_U u_{\max} B_{\max} \max\{m_k\} e^{-\frac{z_{\min}}{\varepsilon} t}, \quad (6.71)$$

and (6.68) becomes, after dividing through by $\|\mathbf{b}\|_2$,

$$\frac{d}{dt} \|\mathbf{b}\|_2 \leq -\frac{\hat{z}_{\min}}{\varepsilon} \|\mathbf{b}\|_2 + \frac{1}{\varepsilon} C_0 e^{-\frac{z_{\min}}{\varepsilon} t}, \quad (6.72)$$

where $C_0 = \frac{1}{2} \frac{4N(N-1)C_U u_{\max} B_{\max} \max\{m_k\}}{\min\{n_k\}^{\frac{1}{2}}}$, or equivalently,

$$\frac{d}{dt} \left(\|\mathbf{b}\|_2 e^{\frac{\hat{z}_{\min}}{\varepsilon} t} \right) \leq \frac{C_0}{\varepsilon} e^{\frac{\hat{z}_{\min} - z_{\min}}{\varepsilon} t}. \quad (6.73)$$

Integrating both sides above in t gives

$$\|\mathbf{b}\|_2 \leq \|\mathbf{b}^0\|_2 e^{-\frac{\hat{z}_{\min}}{\varepsilon} t} + \frac{C_0}{\varepsilon} \frac{\varepsilon}{\hat{z}_{\min} - z_{\min}} \left(e^{-\frac{z_{\min}}{\varepsilon} t} - e^{-\frac{\hat{z}_{\min}}{\varepsilon} t} \right). \quad (6.74)$$

Since $\hat{V}^\top \hat{V} = I_{N-1}$, it follows that $\|\tilde{\boldsymbol{\xi}}\|_2^2 = \|\hat{V}\mathbf{b}\|_2^2 = \|\mathbf{b}\|_2^2$. Thus, (6.74) becomes

$$\|\boldsymbol{\xi} - \boldsymbol{\xi}^\infty\|_2 \leq \|\boldsymbol{\xi}^0 - \boldsymbol{\xi}^\infty\|_2 e^{-\frac{\hat{z}_{\min}}{\varepsilon} t} + C_0 \frac{e^{-\frac{z_{\min}}{\varepsilon} t} - e^{-\frac{\hat{z}_{\min}}{\varepsilon} t}}{\hat{z}_{\min} - z_{\min}}. \quad (6.75)$$

Further, since

$$\|\boldsymbol{\xi} - \boldsymbol{\xi}^\infty\|_2^2 = \|Q^{-\frac{1}{2}}(\mathbf{E} - \mathbf{E}^\infty)\|_2^2 = \sum_{i=1}^N n_i^{-1} |E_i - E_i^\infty|^2 \geq \frac{1}{\max\{n_k\}} \|\mathbf{E} - \mathbf{E}^\infty\|_2^2, \quad (6.76)$$

it follows that

$$\|\mathbf{E}(t) - \mathbf{E}^\infty\|_2 \leq \max\{n_k\}^{\frac{1}{2}} \left(\|\boldsymbol{\xi}^0 - \boldsymbol{\xi}^\infty\|_2 e^{-\frac{\hat{z}_{\min}}{\varepsilon} t} + C_0 \frac{e^{-\frac{z_{\min}}{\varepsilon} t} - e^{-\frac{\hat{z}_{\min}}{\varepsilon} t}}{\hat{z}_{\min} - z_{\min}} \right), \quad (6.77)$$

which recovers (6.64) with $C_1 := \max\{n_k\}^{\frac{1}{2}} \|\xi^0 - \xi^\infty\|_2$ and $C_2 := C_0 \max\{n_k\}^{\frac{1}{2}}$. Thus $\lim_{t \rightarrow \infty} E(t) = E^\infty$, as desired. ■

6.3.4 Temperature Relaxation

Corollary 6.3.8. *With the same assumptions as in Proposition 6.3.1, for each species $i \in \{1, \dots, N\}$, $\lim_{t \rightarrow \infty} T_i(t) = T^\infty$, and*

$$|T_i - T^\infty| \leq \frac{2}{d \min\{n_k\}} \left[C_1 e^{-\frac{\hat{z}_{\min}}{\varepsilon} t} + C_2 \frac{e^{-\frac{z_{\min}}{\varepsilon} t} - e^{-\frac{\hat{z}_{\min}}{\varepsilon} t}}{\hat{z}_{\min} - z_{\min}} \right] + \frac{\max\{m_k\}}{d} 2u_{\max} C_U e^{-\frac{z_{\min}}{\varepsilon} t}. \quad (6.78)$$

Proof. For each i ,

$$T_i = \frac{2}{dn_i} E_i - \frac{m_i}{d} |\mathbf{u}_i|^2 \quad \text{and} \quad T^\infty = \frac{2}{dn_i} E_i^\infty - \frac{m_i}{d} |\mathbf{u}^\infty|^2. \quad (6.79)$$

Therefore

$$|T_i - T^\infty| \leq \frac{2}{d \min\{n_k\}} |E_i - E_i^\infty| + \frac{\max\{m_k\}}{d} ||\mathbf{u}^\infty|^2 - |\mathbf{u}_i|^2|. \quad (6.80)$$

The velocity decay bound (6.56), gives

$$||\mathbf{u}^\infty|^2 - |\mathbf{u}_i|^2| \leq 2u_{\max} \|\mathbf{u}^\infty - \mathbf{u}_i\|_2 \leq 2u_{\max} C_U e^{-\frac{z_{\min}}{\varepsilon} t}. \quad (6.81)$$

Using this and the energy decay bound (6.64), the result follows. ■

Remark 6.3.9. *The result in Corollary 6.3.8 can also be verified directly, using the formulation of the temperature equation (6.23) and an approach similar to the proof of Theorem 6.3.7. The difference in the proofs comes from the fact that the nullspace component analogous to $\tilde{\xi}_N$ is no longer zero when using (6.23).*

6.4 Numerical Demonstration

In this section, we compute temperature and velocity profiles using a fully implicit (backward) Euler time stepping scheme.¹ We assume a slab geometry in the x_1 direction. Thus while

¹The Backward Euler method preserves the monotonicity properties of the temperature and bulk velocity established in Section 6.2. This fact will be proved in future work.

$d = 3$, the quantities of interest depend only on x_1 . Moreover the bulk velocities in the x_2 and x_3 direction are identically zero, i.e., $U_{i,j} = 0$ for all $j \in \{2, 3\}$.

We consider systems with three different gas species chosen from a collection of four possible elements: Helium (He), Argon (Ar), Krypton (Kr), and Xenon (Xe). The masses and diameters of these elements are taken from [4] and are given below in SI units:

$$\begin{pmatrix} m_{\text{He}} \\ m_{\text{Ar}} \\ m_{\text{Kr}} \\ m_{\text{Xe}} \end{pmatrix} = \begin{pmatrix} 6.6464731 \\ 66.335209 \\ 139.14984 \\ 218.01714 \end{pmatrix} \times 10^{-27} \text{ kg}, \quad \begin{pmatrix} d_{\text{He}} \\ d_{\text{Ar}} \\ d_{\text{Kr}} \\ d_{\text{Xe}} \end{pmatrix} = \begin{pmatrix} 2.193 \\ 3.659 \\ 4.199 \\ 4.939 \end{pmatrix} \times 10^{-10} \text{ m}. \quad (6.82)$$

In all of the examples below, the monotonicity of the minimum temperature (Theorem 6.2.2) and the upper and lower bounds on the bulk velocities (Theorem 6.2.3) are respected.

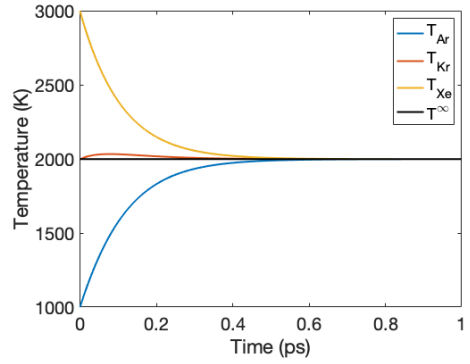
Example 1: Temperature Decay, Ar-Kr-Xe Mixture. The purpose of this example is to demonstrate temperature relaxation when the bulk velocities are zero. The initial number densities, velocities, and temperatures are given by

$$\begin{pmatrix} n_{\text{Ar}}^0 \\ n_{\text{Kr}}^0 \\ n_{\text{Xe}}^0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \times 10^{28} \text{ m}^{-3}, \quad \mathbf{U}^0 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \text{ m/s}, \quad \mathbf{T}^0 = \begin{pmatrix} 1000 \\ 2000 \\ 3000 \end{pmatrix} \text{ K}. \quad (6.83)$$

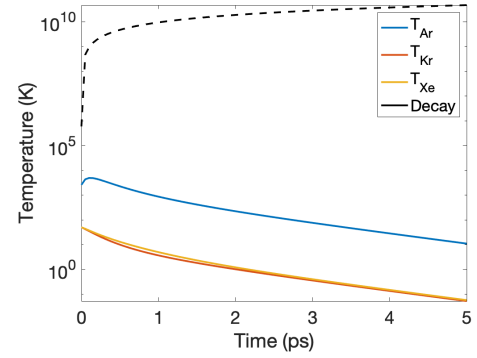
Figure 6.1 shows the relaxation of the temperatures to their steady state value. In this case the decay rate established in (6.78) is fairly sharp, even though the temperatures may not approach steady state monotonically.

Example 2: Velocity Decay, Ar-Kr-Xe Mixture. The purpose of this example is to demonstrate the velocity relaxation, with minimal effects from the temperature. A positive velocity is given to one particle type, and each temperature is set to the same constant. The initial number densities, velocities, and temperatures are given by

$$\begin{pmatrix} n_{\text{Ar}}^0 \\ n_{\text{Kr}}^0 \\ n_{\text{Xe}}^0 \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix} \times 10^{28} \text{ m}^{-3}, \quad \mathbf{U}^0 = \begin{pmatrix} 100 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \text{ m/s}, \quad \mathbf{T}^0 = \begin{pmatrix} 1000 \\ 1000 \\ 1000 \end{pmatrix} \text{ K}. \quad (6.84)$$



(a) Species and steady state temperatures.



(b) Deviation from steady state temperature and analytical decay estimate from (6.78).

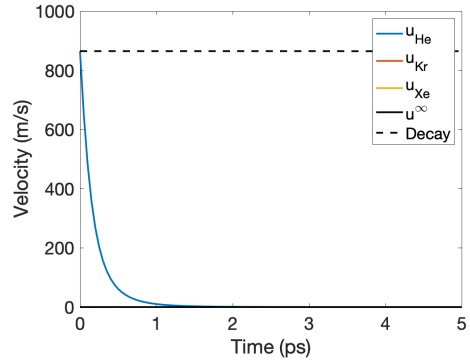
Figure 6.1: Temperature Decay of Ar-Kr-Xe Mixture in Example 1.

Figure 6.2 shows the relaxation of the velocities and temperatures to the steady state values. The decay rate established for the velocities and temperatures in (6.56) and (6.78) underestimate the true decay rates in the example. The results also demonstrate that, unlike the two-species case, particle velocities may not converge to the steady-state value monotonically.

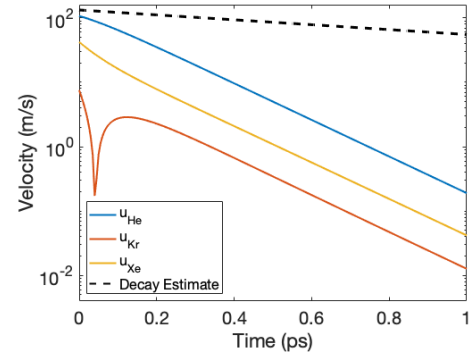
Example 3: Velocity-Temperature Relaxation, He-Kr-Xe Mixture This test case exercises the model due to the large kinetic energy differences between the species. The initial number densities, velocities, and temperatures are given by

$$\begin{pmatrix} n_{\text{He}}^0 \\ n_{\text{Kr}}^0 \\ n_{\text{Xe}}^0 \end{pmatrix} = \begin{pmatrix} 0.01 \\ 1 \\ 1 \end{pmatrix} \times 10^{28} \text{ m}^{-3}, \quad \mathbf{U}^0 = \begin{pmatrix} 864.8 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \text{ m/s}, \quad \mathbf{T}^0 = \begin{pmatrix} 3000 \\ 300 \\ 300 \end{pmatrix} \text{ K.} \quad (6.85)$$

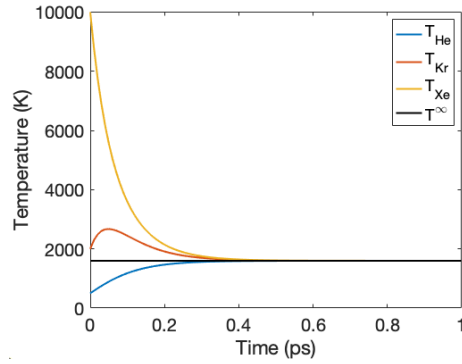
The momentum equilibration has a corresponding kinetic energy redistribution which causes non-monotonic changes in the temperature profile, as seen in Figure 6.3c. Because of the large differences in the particle masses and number densities, the estimates for the decay rates established in (6.56) and (6.78) are very weak.



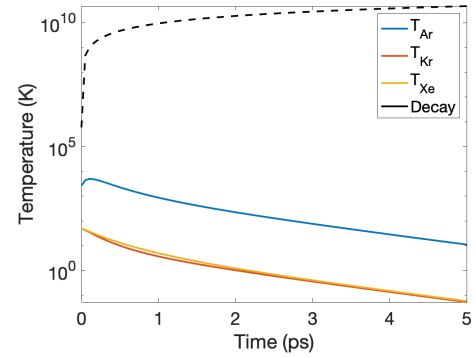
(a) Species and steady state velocities.



(b) Deviation from steady state velocity and analytical decay estimate from (6.56).



(c) Species and steady state temperatures.



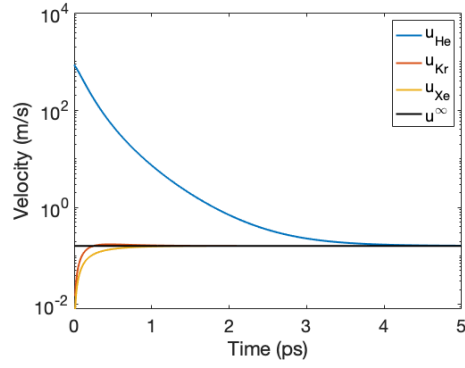
(d) Deviation from steady state velocity and analytical decay estimate from (6.78).

Figure 6.2: Velocity and Temperature Decay of Ar-Kr-Xe Mixture in Example 2.

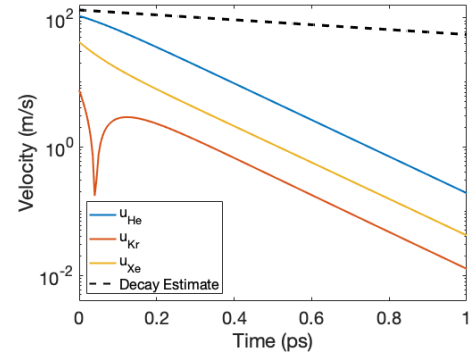
6.5 Conclusions

In this chapter, we have studied the moment equations associated to a recently developed multi-species BGK kinetic model for rarefied gas dynamics in the spatially homogeneous setting. The model includes collision frequencies that, unlike most models in the literature, depend on the species temperatures. This fact complicates the analysis of the moment equations and subsequent numerical tools. We have proven that all species temperatures are bounded below by a positive, non-decreasing temperature envelope, thus establishing that species temperatures always remain positive. We have also established upper and lower bounds on the components of the bulk velocity that are in terms of the initial data.

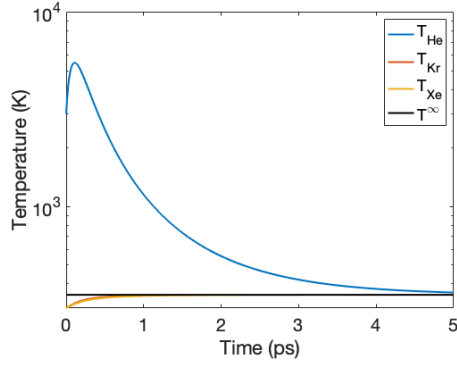
Using the lower bound on the species temperatures, we have shown that the moments



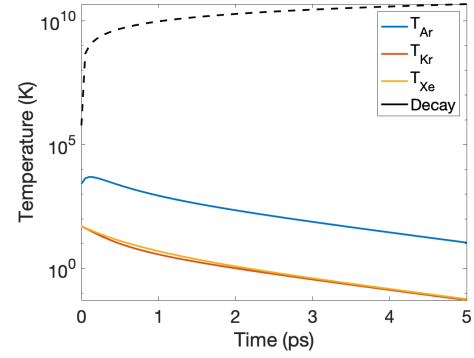
(a) Species and steady state velocities.



(b) Deviation from steady state velocity and analytical decay estimate from (6.56).



(c) Species and steady state temperatures.



(d) Deviation from steady state velocity and analytical decay estimate from (6.78).

Figure 6.3: Velocity and Temperature Decay of He-Kr-Xe Mixture in Example 3.

always stay within a bounded, time-invariant set of physically realizable states, which, in turn, leads to the existence of global unique solutions. Finally, we have proved that unique equilibria exist for the moment equations, and solutions converge to these equilibria exponentially in time. In addition, we have established bounds on the convergence rates to equilibria. We concluded the examination of the space homogeneous problem with some basic numerical simulations that demonstrate some of the established theoretical behavior.

The numerical results in this chapter are computed with a backward Euler method, using an iterative scheme to solve the relevant non-linear algebraic equations at each time-step. In future work, we will describe the scheme more fully and prove convergence under a suitable time step restriction that does not require resolution of the parameter ε . We will use the backward Euler method as a component in an IMEX scheme for simulating multi-species BGK

models that include phase-space advection.

Chapter 7

Progress Summary and Next Steps

7.1 Summary

This Year-2 progress report presents a review of basic background on the theory and numerical solution of BGK approximations for Boltzmann-type kinetic equations. The BGK equations, which simplify their Boltzmann equation counterparts, are highly nonlinear, nonlocal, and high-dimensional models of particle kinetics in rarified gases and plasmas. We discuss both single species models and self-consistent extensions to multi-species BGK models. Preliminary work toward the efficient numerical simulation of BGK-type kinetic equations is presented via several benchmark problems.

Theoretical aspects of single species BGK kinetic models have been presented in order to motivate and describe the numerical methods used. In particular, conservation and entropy dissipation properties were presented, along with an analysis of the space homogeneous (no advection) problem. Numerical methods must be sophisticated enough to respect these conservation and dissipation properties at the fully discrete level. A stable finite volume numerical approximation framework has been outlined to capture potential discontinuities in the approximate solutions. Since the problem can be numerically stiff for some parameter regimes, implicit and semi-implicit time integration schemes are of particular importance for stability. Particular focus is given to implicit-explicit Runge-Kutta (IMEX-RK) time stepping methods, as they give a reasonable balance between accuracy, efficiency, and stability. However, in some cases, fully implicit integration strategies are demanded, and the work here is a stepping stone toward developing such algorithms and codes.

Several numerical benchmark problems and tests of the prototype MATLAB codes are presented, including the Sod shock tube and two-stream instability test cases. The tests are

presented for $1 \times 1 \nu$ phase spaces, but code is currently under development for more realistic higher-dimensional cases. In particular, $1 \times 3 \nu$ and $2 \times 3 \nu$ codes for “slab geometries” will be designed and benchmarked in the near future. The numerical methods that are presented are inherently scalable, and, thus, the only impediment for efficient and stable numerical simulation is the increased number of degrees of freedom. An example $1 \times 1 \nu$ code, used for numerical solutions of the Sod shock tube problem, is given in Appendix E. This work is done to address the goal of developing practical stable and efficient semi-implicit and fully implicit time integrator strategies for the problem.

An introduction to a particular self-consistent BGK-type model for multi-species particle kinetics is given. The combinatorial complexity of the model grows with the addition of distinct chemical constituents, making such multi-species models even more challenging for numerical solution. We have discussed the moment equations for the multispecies case, as the moments can be used to update the Maxwellians in an IMEX approach. We have studied the asymptotic relaxation behavior of the moment equations and have proven that the temperatures always remain positive, which guarantees the existence and uniqueness of the moment equations.

7.2 Next Steps

There are several outstanding issues that this work will address in the future. In particular, as in the single species case, an implicit (or, at least, semi-implicit) approach is desired for the computation of the stiff collision operators in the multi-species case to achieve numerical stability. However, the procedures used in the single species case do not work directly in the multi-species case. Solving the moment equation system is a fundamental bridge toward an efficient IMEX scheme for the multi-species case. A sophisticated fixed point iterative scheme is currently in development to address this issue in computing the implicit update for the moment equations. The scheme is designed to relax the implicit updates of the velocity and temperature to appropriate values, depending on collisional frequencies that themselves depend on the implicit moments. This allows for the desired implicit update of the BGK collision operators. Beyond the multispecies BGK model presented, we will examine a multispecies ellipsoidal statistical BGK model (ES-BGK), that allows for simulation with the correct Prandtl number, and captures anisotropic effects in the flow behavior. We will examine the moment system for this model, and develop iterative methods akin to the ones currently in development, to update the moment system.

In the longer term, in addition to semi- and fully-implicit solvers, this work will focus on the design of fast adaptive phase-space methods and block-structured adaptive mesh refinement (AMR) that will efficiently accommodate disparate scales that are inherent in multi-species problems, owing, for example, to disparate particle sizes and temperatures. Incorporating implicit solver technology with AMR, especially in the context of such highly nonlinear and nonlocal models is expected to be challenging. We will also develop efficient discrete velocity methods for multi-species BGK models, extending the framework that we outlined herein for the single species case.

Chapter 8

Acknowledgements

This work is supported by the ASINA Program at Oak Ridge National Laboratory under subcontract UTB-CW24420, “BGK Kinetic Equations.” The authors wish to thank Cory Hauck (ORNL) and Ryan Glasby (ORNL) for several useful discussions regarding numerical methods for BGK-type equations as well as their permission to use the results of some joint work in this report.

Appendix A

An Important Integral

We need to compute the following generalized Gaussian integral in several places in the report.

Lemma A.0.1.

$$\int_{\mathbb{R}^d} |\mathbf{s}|^2 e^{-|\mathbf{s}|^2} d\mathbf{s} = \frac{d}{2} \pi^{\frac{d}{2}}. \quad (\text{A.1})$$

Proof.

$$\begin{aligned} \int_{\mathbb{R}^d} |\mathbf{s}|^2 e^{-|\mathbf{s}|^2} d\mathbf{s} &= \int_{\mathbb{R}^d} (s_1^2 + \dots + s_d^2) e^{-(s_1^2 + \dots + s_d^2)} d\mathbf{s} \\ &= \int_{\mathbb{R}^d} \left(\sum_{j=1}^d s_j^2 \right) \prod_{k=1}^d e^{-s_k^2} d\mathbf{s} \\ &= \int_{\mathbb{R}^d} s_1^2 \prod_{k=1}^d e^{-s_k^2} d\mathbf{s} + \dots + \int_{\mathbb{R}^d} s_d^2 \prod_{k=1}^d e^{-s_k^2} d\mathbf{s} \\ &= \left(\prod_{\substack{k=1 \\ k \neq 1}}^d \int_{\mathbb{R}} e^{-s_k^2} ds_k \right) \int_{\mathbb{R}} s_1^2 e^{-s_1^2} ds_1 + \dots + \left(\prod_{\substack{k=1 \\ k \neq d}}^d \int_{\mathbb{R}} e^{-s_k^2} ds_k \right) \int_{\mathbb{R}} s_d^2 e^{-s_d^2} ds_d \\ &= \left(\pi^{\frac{1}{2}} \right)^{d-1} \int_{\mathbb{R}} s_1^2 e^{-s_1^2} ds_1 + \dots + \left(\pi^{\frac{1}{2}} \right)^{d-1} \int_{\mathbb{R}} s_d^2 e^{-s_d^2} ds_d \\ &= \pi^{\frac{d-1}{2}} \sum_{i=1}^d \int_{\mathbb{R}} s_i^2 e^{-s_i^2} ds_i, \end{aligned} \quad (\text{A.2})$$

where we have used many elementary integration techniques, and the fact that $\int_{\mathbb{R}} e^{-x^2} dx = \pi^{\frac{1}{2}}$. It remains to determine the value of the terms of the form $\int_{\mathbb{R}} s^2 e^{-s^2} ds$. To that end,

utilizing integration by parts, with

$$u = s_i \implies du = ds_i \quad (\text{A.3})$$

$$dv = 2s_i e^{-s_i^2} ds_i \implies v = -e^{-s_i^2}, \quad (\text{A.4})$$

we have

$$\begin{aligned} \frac{1}{2} \int_{\mathbb{R}} 2s_i^2 e^{-s_i^2} ds_i &= \frac{1}{2} \left[\left[-s_i e^{-s_i^2} \right]_{-\infty}^{\infty} - \int_{\mathbb{R}} -e^{-s_i^2} ds_i \right] \\ &= \frac{1}{2} \int_{\mathbb{R}} e^{-s_i^2} ds_i \\ &= \frac{1}{2} \pi^{\frac{1}{2}}. \end{aligned} \quad (\text{A.5})$$

Thus, Equation (A.2) gives us

$$\int_{\mathbb{R}^d} |\mathbf{s}|^2 e^{-|\mathbf{s}|^2} d\mathbf{s} = \pi^{\frac{d-1}{2}} \sum_{i=1}^d \frac{1}{2} \pi^{\frac{1}{2}} = \pi^{\frac{d-1}{2}} \left(\frac{d}{2} \pi^{\frac{1}{2}} \right) = \frac{d}{2} \pi^{\frac{d}{2}}, \quad (\text{A.6})$$

as desired. ■

Appendix B

Picard-Lindelöf Theorem

We will make use of the following version of the Picard-Lindelöf Theorem for autonomous systems:

Theorem B.0.1. *Suppose that $\mathbf{u}_0 \in \mathbb{R}^d$ and $\mathbf{f} \in C(\overline{B}(\mathbf{u}_0, \beta); \mathbb{R}^d)$, where $B(\mathbf{u}_0, \beta)$ ($\overline{B}(\mathbf{u}_0, \beta)$) denotes the open (closed) Euclidean ball of radius β centered at $\mathbf{u}_0$. Suppose further that $\|\mathbf{f}(\mathbf{v})\|_2 \leq M$, for all $\mathbf{v} \in \overline{B}(\mathbf{u}_0, \beta)$, and $\mathbf{f}$ is Lipschitz continuous, with constant $L > 0$, on $\overline{B}(\mathbf{u}_0, \beta)$. Set $\delta = \min(\frac{1}{2L}, \frac{\beta}{M})$. Then the IVP $\mathbf{u}'(t) = \mathbf{f}(\mathbf{u}(t))$, $\mathbf{u}(t_0) = \mathbf{u}_0$ has a unique solution on the interval $[t_0 - \delta, t_0 + \delta]$.*

Appendix C

Bounding the Eigenvalues of Z and $\hat{Z}$

The matrices Z and $\hat{Z}$, defined in (6.20), play a key role in the dynamics of (6.17) and the equivalent system (6.19). In this appendix, we prove bounds on the eigenvalues of these matrices. With the collision frequencies defined in (6.11), the matrices A and B , defined in (6.14a), take the form

$$A_{i,j} = \frac{16}{3} \sqrt{\frac{\pi}{2}} \frac{m_i m_j (d_i + d_j)^2}{(m_i + m_j)^3} \rho_i \rho_j \sqrt{\frac{T_i}{m_i} + \frac{T_j}{m_j}}, \quad (\text{C.1})$$

$$B_{i,j} = \frac{8}{3} \sqrt{\frac{\pi}{2}} \frac{(d_i + d_j)^2}{(m_i + m_j)^2} \rho_i \rho_j \sqrt{\frac{T_i}{m_i} + \frac{T_j}{m_j}}. \quad (\text{C.2})$$

Since each value used to define the matrix elements is positive and bounded above and below (importantly, the temperature), then their values can be bounded as follows:

$$A_{\min}(t) := \min_{i,j} A_{i,j}(t) \leq A_{i,j}(t) \leq \max_{i,j} A_{i,j}(t) =: A_{\max}(t) \quad (\text{C.3})$$

$$B_{\min}(t) := \min_{i,j} B_{i,j}(t) \leq B_{i,j}(t) \leq \max_{i,j} B_{i,j}(t) =: B_{\max}(t) \quad (\text{C.4})$$

Theorem C.0.1 (Eigenvalue Bounds). *Let $A, B, D, F \in \mathbb{R}^{N \times N}$ be the matrices defined in (6.14), and let $A_{\min} = \min_{i,j} A_{i,j}$, $B_{\min} = \min_{i,j} B_{i,j}$, $A_{\max} = \max_{i,j} A_{i,j}$, and $B_{\max} = \max_{i,j} B_{i,j}$. Then the eigenvalues of the matrices $Z = P^{-\frac{1}{2}}(D - A)P^{-\frac{1}{2}}$ and $\hat{Z} = Q^{-\frac{1}{2}}(F - B)Q^{-\frac{1}{2}}$ satisfy*

$$0 = z_0 < z_{\min} \leq z_1 \leq \cdots \leq z_{N-1} \leq z_{\max}, \quad (\text{C.5})$$

$$0 = \hat{z}_0 < \hat{z}_{\min} \leq \hat{z}_1 \leq \cdots \leq \hat{z}_{N-1} \leq \hat{z}_{\max}, \quad (\text{C.6})$$

where

$$z_{\min} = \frac{A_{\min} N}{\max\{\rho_k\}}, \quad z_{\max} = \frac{A_{\max}(N-1)}{\min\{\rho_k\}}, \quad (\text{C.7})$$

$$\hat{z}_{\min} = \frac{B_{\min} N}{\max\{n_k\}}, \quad \hat{z}_{\max} = \frac{B_{\max}(N-1)}{\min\{n_k\}}. \quad (\text{C.8})$$

Proof. The proofs of (C.5) and (C.6) are similar; thus we prove (C.5) and leave the remaining details to the reader. The (single) zero eigenvalue of $(D - A)$ corresponds to the eigenvector, $\mathbf{1}$, spanning the null space of $(D - A)$. To find bounds on the other eigenvalues, let $\mathbf{y} \in \mathcal{R}(D - A) = \mathcal{N}(D - A)^\perp$, so that $\mathbf{y}^\top \mathbf{1} = \sum_i y_i = 0$. Recall that $\mathbf{y}^\top (D - A) \mathbf{y} = \frac{1}{2} \sum_i \sum_j A_{i,j} (y_i - y_j)^2$. Using $\sum_i y_i = 0$,

$$\mathbf{y}^\top (D - A) \mathbf{y} \leq \frac{A_{\max}}{2} \sum_i \sum_j (y_i - y_j)^2 = \frac{A_{\max}}{2} \sum_i \sum_{j \neq i} (y_i^2 + y_j^2) = A_{\max}(N-1) \|\mathbf{y}\|_2^2. \quad (\text{C.9})$$

Thus if $\mathbf{y} = P^{-\frac{1}{2}} \mathbf{z}$, then

$$\mathbf{z}^\top Z \mathbf{z} \leq (N-1) A_{\max} \mathbf{z}^\top P^{-1} \mathbf{z} = (N-1) A_{\max} \sum_i \frac{1}{\rho_i} z_i^2 \leq \frac{(N-1) A_{\max}}{\min\{\rho_k\}} \|\mathbf{z}\|_2^2 =: z_{\max} \|\mathbf{z}\|_2^2, \quad (\text{C.10})$$

which gives the upper bound in (C.5). For the lower bound,

$$\mathbf{y}^\top (D - A) \mathbf{y} \geq \frac{A_{\min}}{2} \sum_{i=1}^N \sum_{j=1}^N (y_i^2 + y_j^2) = A_{\min} N \|\mathbf{y}\|_2^2. \quad (\text{C.11})$$

Thus if $\mathbf{y} = P^{-\frac{1}{2}} \mathbf{z}$, then

$$\mathbf{z}^\top Z \mathbf{z} \geq N A_{\min} \mathbf{z}^\top P^{-1} \mathbf{z} \geq \frac{N A_{\min}}{\max\{\rho_k\}} \|\mathbf{z}\|_2^2 =: z_{\min} \|\mathbf{z}\|_2^2. \quad (\text{C.12})$$

■

Appendix D

Null Space Components of $\widetilde{W}$ and $\widetilde{\xi}$

In this appendix, we show that $\widetilde{W}_{\mathcal{N}}$ (the projection of $\widetilde{W}$ onto $\mathcal{N}(Z)$) and $\widetilde{\xi}_{\mathcal{N}}$ (the projection of $\widetilde{\xi}$ onto $\mathcal{N}(\widehat{Z})$) are both zero. These results are used in the proofs of Theorems 6.3.6 and 6.3.7, respectively.

Lemma D.0.1. $\widetilde{W}_{\mathcal{N}} \equiv \mathbf{0}$.

Proof. Since the null space of Z is spanned by $P^{\frac{1}{2}}\mathbf{1}$, it is sufficient to show that $\widetilde{W}^{\top} P^{\frac{1}{2}}\mathbf{1} = \mathbf{0}$. Write $\mathbf{u}^{\infty} = \frac{\mathbf{U}^{\top} P \mathbf{1}}{\mathbf{1}^{\top} P \mathbf{1}}$ and $(\mathbf{U}^{\infty})^{\top} = \mathbf{u}^{\infty} \mathbf{1}^{\top}$, so that

$$\widetilde{W}^{\top} P^{\frac{1}{2}}\mathbf{1} = (\mathbf{W} - \mathbf{W}^{\infty})^{\top} P^{\frac{1}{2}}\mathbf{1} = \mathbf{U}^{\top} P \mathbf{1} - (\mathbf{U}^{\infty})^{\top} P \mathbf{1} = \mathbf{U}^{\top} P \mathbf{1} - \frac{\mathbf{U}^{\top} P \mathbf{1}}{\mathbf{1}^{\top} P \mathbf{1}} \mathbf{1}^{\top} P \mathbf{1} = \mathbf{0}. \quad (\text{D.1})$$

■

Lemma D.0.2. $\widetilde{\xi}_{\mathcal{N}} \equiv \mathbf{0}$.

Proof. Since the null space of $\widehat{Z}$ is spanned by $Q^{\frac{1}{2}}\mathbf{1}$, it is sufficient to show that $\widetilde{\xi}^{\top} Q^{\frac{1}{2}}\mathbf{1} = 0$. Recall that $\widetilde{\xi} = Q^{-\frac{1}{2}}(\mathbf{E} - \mathbf{E}^{\infty})$; hence

$$\widetilde{\xi}^{\top} Q^{\frac{1}{2}}\mathbf{1} = (\mathbf{E} - \mathbf{E}^{\infty})^{\top} \mathbf{1} = \sum_{i=1}^N E_i - \sum_{i=1}^N E_i^{\infty} = 0, \quad (\text{D.2})$$

where the fact that $\sum_i E_i = \sum_i E_i^{\infty}$ is a consequence of (6.50) in the proof of Proposition 6.3.1. ■

Appendix E

Code

The MATLAB code listed in this appendix is that used to generate the results of Section 3.5 for the Sod shock tube benchmark problem. This code is developed for prototyping and demonstration purposes only and is not meant to represent production-quality software.

E.1 Main Driver: vlasovPoissonBGKMain.m

```
% Script to solve Vlasov-Poisson-BGK Equation:
%
% 1X1V
%
%  $f_t + v * f_x + a * f_v = 1/\tau * (M-f)$ 
%
% This set of code is designed to solve the Sod Shock tube problem.
% In the limit as ( $\tau \rightarrow 0$ ) or ( $\lambda \rightarrow \infty$ ), BGK  $\rightarrow$  Euler.
% So, for the current test, set  $\tau=10^{-N}$ , set  $a = 0$ .
%
% Scheme:
% IMEX RK: Explicit Advection, Implicit Collision.
%
clear;
clc;
tic
%
% Number of ghost cells:
```

```

del = 1;
%
% Collision time:
tau = 10^(-4);
%
% Theta values for minmod:
thetaX = 2;
thetaV = 2;
%
% Spatial domain:
xL = -0.5;
xR = 0.5;
%
% Number of cells:
Nx = 256;
%
% Cell edge points:
x = linspace(xL,xR,Nx+1);
%
% Cell center points:
xC = x(1:end-1) + 0.5 * (x(2:end) - x(1:end-1));
xCenter = [xC(1:del) , xC , xC(end-del+1:end)];
hX = xC(2) - xC(1);
%
% Velocity domain:
vMin = -10;
vMax = 10;
%
% Number of cells:
Nv = 258;
%
% Cell edge points:
v = linspace(vMin,vMax,Nv+1);
%
% Cell center points:
vC = v(1:end-1) + 0.5 * (v(2:end) - v(1:end-1));
vCenter = [vC(1:del) , vC , vC(end-del+1:end)];
hV = vC(2) - vC(1);
%
% Define time levels:

```

```

tInit = 0;
tFin = 0.2;
%
% Vectors for ghost cells
gVx = [1,length(xC)]+del;
gVv = [1,length(vC)]+del;
%
tableNum = 8;
[Ae,be,ce,Ai,bi,ci] = butcherTable(tableNum);
butcher.Ae = Ae; butcher.Ai = Ai;
butcher.be = be; butcher.bi = bi;
butcher.ce = ce; butcher.ci = ci;
%
% Stiffness matrix for the Poisson solve...zero Dirichlet BC
% (as per the test problems). This is not needed for the Sod problem,
% but I kept it so I didn't have to change all my function
% dependencies.
%
temp = zeros(1,Nx);
temp(1)=2; temp(2)=-1;
A = toeplitz(temp);
A(1,1) = 3; A(end,end) = 3;
%
grid.del = del;
grid.gVx = gVx; grid.Nx = Nx; grid.xC = xC; grid.xCenter = xCenter;
grid.gVv = gVv; grid.Nv = Nv; grid.vC = vC; grid.vCenter = vCenter;
grid.hX = hX;
grid.hV = hV;
grid.xL = xL ; grid.xR = xR ;
grid.vMin = vMin; grid.vMax = vMax;
grid.tInit = tInit; grid.tFin = tFin;
grid.thetaX = thetaX;
grid.thetaV = thetaV;
%
testNum = 1;
[f] = initialCondition(testNum,grid);
%
[tVec,y,MAX] = vlasovPoissonBGKSolver(f,butcher,grid,tau,A,f);
%
figure(107)

```

```

pcolor(xC,vC,y);
title(['Contour plot of computed f, in the Velocity/Space', ...
      ' Domain, for Time t = ', num2str(tFin), ' '])
xlabel('x')
ylabel('v')
shading interp;
colormap(jet);
colorbar;
%
toc
%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%
% Embedded functions below:
%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%
function [f] = initialCondition(testNum,grid)
%
xC = grid.xC; xCenter = grid.xCenter;
vC = grid.vC; vCenter = grid.vCenter;
%
% Initialize solution array:
f = zeros(length(vC),length(xC),1);
%
switch testNum
    case 1 % tFin = 0.250
        nL = 1.00000; uL = 0.00000; thetaL = 0001.000;
        nR = 0.12500; uR = 0.00000; thetaR = 0000.800;
    case 2 % tFin = 0.150
        nL = 1.00000; uL = -2.00000; thetaL = 0000.400;
        nR = 1.00000; uR = 2.00000; thetaR = 0000.400;
    case 3 % tFin = 0.012
        nL = 1.00000; uL = 0.00000; thetaL = 1000.000;
        nR = 1.00000; uR = 0.00000; thetaR = 0000.010;
    case 4 % tFin = 0.035
        nL = 1.00000; uL = 0.00000; thetaL = 0000.010;
        nR = 1.00000; uR = 0.00000; thetaR = 0100.000;
    case 5 % tFin = 0.035
        nL = 5.99924; uL = 19.59750; thetaL = 0460.894;

```

```

    nR = 5.99242; uR = -6.19633; thetaR = 046.0950;
end
%
% Initial condition (make a separate function?):
%
for j = 1:length(vC)
    for i = 1:length(xC)
        if xCenter(i) <= 0.0
            f(j,i) = nL / sqrt(2*pi*thetaL) * exp(-(vC(j)-uL)^2 ...
                / (2*thetaL));
        else
            f(j,i) = nR / sqrt(2*pi*thetaR) * exp(-(vC(j)-uR)^2 ...
                / (2*thetaR));
        end
    end
end
end
%
% %Plot the initial conditions....
% figure(110)
% % pcolor(x_c,v_c,f);
% surf(x_c,v_c,f); view(-30,50); zlim([0,2.1]);
% % title(['Contour plot of f, in the Velocity/Space Domain,' ...
% ' for Time t = ',num2str(0),' '])
% xlabel('x')
% ylabel('v')
% shading interp;
% colormap(jet);
% colorbar;
%
end
%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%
function [Ae,be,ce,Ai,bi,ci] = butcherTable(table)
%
% Butcher Tableaux for the IMEX-RK scheme. Tables 2-6 come from
% the Pareschi & Russo paper. Table 7 is Backward Euler. Table 8
% (needs reference, from Cory).
%
switch table

```

```

case 2
    Ae = [0,0 ; 1,0];
    be = [0.5 ; 0.5];
    ce = [0 ; 1];

%
    gam = 1-1/sqrt(2);
    Ai = [gam,0 ; 1-2*gam,gam];
    bi = [0.5 ; 0.5];
    ci = [gam ; 1-gam];

%
case 3
    Ae = [0,0,0 ; 0,0,0 ; 0,1,0];
    be = [0 ; 0.5 ; 0.5];
    ce = [0 ; 0 ; 1];

%
    Ai = [0.5,0,0 ; -0.5,0.5,0 ; 0,0.5,0.5];
    bi = [0 ; 0.5 ; 0.5];
    ci = [0.5 ; 0 ; 1];

%
case 4
    Ae = [0,0,0 ; 0.5,0,0 ; 0.5,0.5,0];
    be = [1/3 ; 1/3 ; 1/3];
    ce = [0 ; 0.5 ; 1];

%
    Ai = [0.25,0,0 ; 0,0.25,0 ; 1/3,1/3,1/3];
    bi = [1/3 ; 1/3 ; 1/3];
    ci = [0.25 ; 0.25 ; 1];

%
case 5
    Ae = [0,0,0 ; 1,0,0 ; 0.25,0.25,0];
    be = [1/6 ; 1/6 ; 2/3];
    ce = [0 ; 1 ; 0.5];

%
    gam = 1-1/sqrt(2);
    Ai = [gam,0,0 ; 1-2*gam,gam,0 ; 0.5-gam,0,gam];
    bi = [1/6 ; 1/6 ; 2/3];
    ci = [gam ; 1-gam ; 0.5];

%
case 6
    Ae = [0,0,0,0 ; 0,0,0,0 ; 0,1,0,0 ; 0,0.25,0.25,0];

```

```

be = [0 ; 1/6 ; 1/6 ; 2/3];
ce = [0 ; 0 ; 1 ; 0.5];

%
alpha = 0.24169426078821;
beta  = 0.06042356519705;
eta   = 0.12915286960590;
Ai = [alpha,0,0,0; -alpha,alpha,0,0; 0,1-alpha,alpha,0; ...
      beta,eta,0.5-beta-eta-alpha,alpha];
bi = [0 ; 1/6 ; 1/6 ; 2/3];
ci = [alpha ; 0 ; 1 ; 0.5];

%
case 7

%
% Forward Euler:
Ae = 0;
be = 1;
ce = 0;

%
% Backward Euler
Ai = 1;
bi = 1;
ci = 1;

%
case 8
gam = 1 - 1/sqrt(2);
delt = 1 - 1/(2*gam);

%
Ae = [0,0,0 ; gam,0,0 ; delt,1-delt,0];
be = [delt;1-delt;0];
ce = [0;gam;1];

%
Ai = [0,0,0 ; 0,gam,0 ; 0,1-gam,gam];
bi = [0;1-gam;gam];
ci = [0;gam;1];
end
end

```

Listing E.1: Main driver: vlasovPoissonBGKMain.m.

E.2 vlasovPoissonBGKSolver.m

```

function [t,Z,M] = vlasovPoissonBGKSolver(y0,butcher,grid,tau,A,f)
%
% y0 = IC grid function (j,i) = (velocity,space)
% Ae, be, ce: Butcher Tableau for explicit solve
% Ai, bi, ci: Butcher tableau for implicit solve
%
Ai = butcher.Ai;
bi = butcher.bi;
ci = butcher.ci;
Ae = butcher.Ae;
be = butcher.be;
ce = butcher.ce;
%
gVx = grid.gVx; Nx = grid.Nx; hX = grid.hX;
gVv = grid.gVv; Nv = grid.Nv; hV = grid.hV;
%
% A is the stiffness matrix for the Poisson solve.
%
% Number of ghost cells
del = grid.del;
%
% Spatial domain:
xL = grid.xL; xR = grid.xR;
%
% Velocity domain:
vMin = grid.vMin; vMax = grid.vMax;
%
% Define CFL(s)
CFLMaxX = min(hX) / max(abs(vMax),abs(vMin));
CFLMaxV = min(hV) / max(abs(xL),abs(xR));
%
% Define initial/final time levels:
tCurrent = grid.tInit;
tFinal = grid.tFin;
%
% size of grid function array y including ghost cells:
d = size(y0)+2*del;

```

```

%
% Initialize array of solutions:
y = zeros([d,2]);
%
% Input initial conditions into solution vector
y(gVv(1):gVv(2),gVx(1):gVx(2),1) = f;
%
% Store appropriate initial values in ghost cells
for i = 1:del
    y(:,gVx(1)-i,1) = y(:,gVx(1),1);
    y(:,gVx(2)+i,1) = y(:,gVx(2),1);
end
for j = 1:del
    y(gVv(1)-j,:,1) = y(gVv(1),,:,1);
    y(gVv(2)+j,:,1) = y(gVv(2),,:,1);
end
%
% Number of stages of RK scheme:
s = length(ce);
%
% Initial vectors for RK scheme:
uo = zeros([d,s]);
fI = zeros([d,s]);
fE = zeros([d,s]);
M = zeros([d,s]);
%
redo = 0;
EMax = 1;
t = 1;
%
while tCurrent < tFinal
%
    dt = min(0.24 * min(CFLMaxX,CFLMaxV),tFinal-tCurrent) ...
        / (1 + (redo > 0));
%
% This flag indicates that the electric field gets bigger than E_max.
% If so, the step needs to recalculate.
    redo = 0;
%
    uo(:, :, :) = 0;

```

```

uo(:, :, 1) = y(:, :, 1);
[~, M(:, :, 1)] = BGKCollision(tCurrent, y(:, :, 1), grid, tau, tCurrent);
u(:, :, 1) = (tau*uo(:, :, 1) + dt*Ai(1, 1)*M(:, :, 1)) ...
    / (tau + dt*Ai(1, 1));
fI(:, :, 1) = BGKCollision(tCurrent+ci(1)*dt, u(:, :, 1), grid, ...
    tau, tCurrent);
%
[fE(:, :, 1), re] = divFlux(tCurrent+ce(1)*dt, u(:, :, 1), grid, EMax, A);
redo = redo+re;

for r = 2:s
    uo(:, :, r) = y(:, :, 1) ...
        + dt * reshape(Ae(r, :) * reshape(permute(fE, [3, 1, 2]), s, ...
            (Nv+2*del)*(Nx+2*del)), Nv+2*del, Nx+2*del) ...
        + dt * reshape(Ai(r, :) * reshape(permute(fI, [3, 1, 2]), s, ...
            (Nv+2*del)*(Nx+2*del)), Nv+2*del, Nx+2*del);
%
    [~, M(:, :, r)] = BGKCollision(tCurrent+ci(r)*dt, uo(:, :, r), ...
        grid, tau, tCurrent);
%
    u(:, :, r) = (tau*uo(:, :, r) + dt*Ai(r, r)*M(:, :, r)) ...
        / (tau + dt*Ai(r, r));

    [fE(:, :, r), re] = divFlux(tCurrent+ce(r)*dt, u(:, :, r), grid, EMax, A);
    redo = redo+re;

    fI(:, :, r) = BGKCollision(tCurrent+ci(r)*dt, u(:, :, r), ...
        grid, tau, tCurrent);
end
%
y(:, :, 2) = y(:, :, 1) ...
    + (dt * reshape(be' * reshape(permute(fE, [3, 1, 2]), s, ...
        (Nv+2*del)*(Nx+2*del)), Nv+2*del, Nx+2*del) ...
    + dt * reshape(bi' * reshape(permute(fI, [3, 1, 2]), s, ...
        (Nv+2*del)*(Nx+2*del)), Nv+2*del, Nx+2*del) ...
    ) * (1 - (redo > 0));
%
y(:, :, 1) = y(:, :, 2);
tCurrent = tCurrent + dt*(1 - (redo>0));
%
```

```
% This computes the moments and Maxwellian at the final time step.
if tCurrent == tFinal
    BGKCollision(tCurrent+ci(r)*dt,y(:, :, 1),grid,tau,tCurrent);
end
end
%
% Return solution at final time
Z = y(gVv(1):gVv(2),gVx(1):gVx(2),1);
%
end
```

Listing E.2: vlasovPoissonBGKSolver.m.

E.3 BGKCollision.m

```

function [z,M] = BGKCollision(t,y,grid,tau,tCurrent)
%
del = grid.del;
hx = grid.hX; gVx = grid.gVx; x = grid.xC;
hv = grid.hV; gVv = grid.gVv; v = grid.vC;
%
% Grid function array must have ghost cells already.
M = zeros(length(x)+2*del,length(v)+2*del);
nDens = zeros(1,length(x));
mDens = zeros(1,length(x));
EDens = zeros(1,length(x));

uDens = zeros(1,length(x));
theta = zeros(1,length(x));
%
% Moment integrals: n, n*u, and E.
nDens(1:length(x)) = hv * sum(y(gVv(1):gVv(2),1+del:length(x)+del));
mDens(1:length(x)) = hv * sum(v(:).*y(gVv(1):gVv(2), ...
    1+del:length(x)+del));
EDens(1:length(x)) = 0.5 * hv * sum(v(:).^2 .* y(gVv(1):gVv(2), ...
    1+del:length(x)+del));

uDens(:) = mDens(:)./nDens(:);

theta(:) = 2*EDens(:)./nDens(:)-uDens(:).^2;

M(gVx(1):gVx(2),gVv(1):gVv(2)) = nDens(:) ./ sqrt(2*pi*theta(:)) ...
    .* exp(-0.5 * abs(v - uDens(:)).^2 ./ theta(:));
%
% Ghost cells: zero flow boundaries.
for i = 1:del
    M(:,gVv(1)-i) = M(:,gVv(1));
    M(:,gVv(2)+i) = M(:,gVv(2));
end
for j = 1:del
    M(gVx(1)-j,:) = M(gVx(1),:);
    M(gVx(2)+j,:) = M(gVx(2),:);

```

```

end
%
if tCurrent == 0.2
%
% initial shock happens here.
x0 = 0;
gam = 3;
uL = [1,0,1]; uR = [0.125,0,0.1];
%
pDens = nDens .* theta;
%
[nSod,uSod,pSod,eSod] = sodSoln(x,tCurrent,uL,uR,gam,x0);
%
figure(1)
plot(x,nDens,'b','LineWidth',3)
hold on
plot(x,nSod,'LineWidth',3)
xlim([-0.5,0.5])
ylim([0,1])
set(gca,'fontsize',24)
legend('BGK','Euler')
hold off
%
figure(2)
plot(x,uDens,'b','LineWidth',3)
hold on
plot(x,uSod,'LineWidth',3)
xlim([-0.5,0.5])
ylim([0,1])
set(gca,'fontsize',24)
legend('BGK','Euler')
hold off
%
figure(3)
plot(x,theta,'b','LineWidth',3)
hold on
plot(x,2*eSod,'LineWidth',3)
xlim([-0.5,0.5])
ylim([0,2]) % may need to shift this
set(gca,'fontsize',24)

```

```
    legend('BGK','Euler')
    hold off
end
%
z = (M'-y)/tau;
%
% M is a [Nv]X[Nx] array after transpose
M = M';
%
end
```

Listing E.3: BGKCollision.m

E.4 divFlux.m

```

function [z,redo] = divFlux(t,y,grid,EMax,A)
%
% y = grid function of f_{ij}^k values for fixed k
%
del = grid.del;
hx = grid.hX; gVx = grid.gVx; x = grid.xCenter; thetaX = grid.thetaX;
hv = grid.hV; gVv = grid.gVv; v = grid.vCenter; thetaV = grid.thetaV;
%
fP = zeros(size(y));
fM = zeros(size(y));
gP = zeros(size(y));
gM = zeros(size(y));
%
% construct slopes
[sX(:, :), sV(:, :)] = slopeReconstruction(y,grid);
%
% % Spatial fluxes
fP(:,gVx(1):gVx(2)) = (y(:,gVx(1):gVx(2)) ...
    + 0.5*hx*sX(:,gVx(1):gVx(2))) .* v(:).*(v>0)' ...
    +(y(:,gVx(1)+1:gVx(2)+1) - 0.5*hx*sX(:,gVx(1)+1:gVx(2)+1)) ...
    .* v(:).*(v<=0)' ;
fM(:,gVx(1):gVx(2)) = (y(:,gVx(1)-1:gVx(2)-1) ...
    + 0.5*hx*sX(:,gVx(1)-1:gVx(2)-1)) .* v(:).*(v>0)' ...
    +(y(:,gVx(1):gVx(2)) - 0.5*hx*sX(:,gVx(1):gVx(2))) ...
    .* v(:).*(v<=0)' ;
%
% The electric field is unnecessary for this test, so we set it to
% zero here.
E = zeros(1,length(x));
redo = 0;

% % Velocity fluxes
gP(gVv(1):gVv(2), :) = (y(gVv(1):gVv(2), :) ...
    + 0.5*hv*sV(gVv(1):gVv(2), :)) .* E.*(E>0) ...
    + (y(gVv(1)+1:gVv(2)+1, :) - 0.5*hv*sV(gVv(1)+1:gVv(2)+1, :)) ...
    .* E.*(E<=0) ;
gM(gVv(1):gVv(2), :) = (y(gVv(1)-1:gVv(2)-1, :) ...

```

```

+ 0.5*hv*sV(gVv(1)-1:gVv(2)-1,:)) .* E.*(E>0) ...
+ (y(gVv(1):gVv(2),:) - 0.5*hv*sV(gVv(1):gVv(2),,:)) ...
.* E.*(E<=0) ;

z = - (fP - fM)/hx - (gP - gM)/hv;

end

```

Listing E.4: divFlux.m

E.5 slopeReconstruction.m

```

function [sX,sV] = slopeReconstruction(y,grid)
%
sX = zeros(size(y));
sV = zeros(size(y));
%
del = grid.del;
hx = grid.hX; gVx = grid.gVx; thetaX = grid.thetaX;
hv = grid.hV; gVv = grid.gVv; thetaV = grid.thetaV;
%
SV(:, :, 1) = (y(gVv(1)+1:gVv(2)+1,gVx(1):gVx(2)) ...
    - y(gVv(1)-1:gVv(2)-1,gVx(1):gVx(2))) / 2 ;
SV(:, :, 2) = thetaV * (y(gVx(1):gVv(2),gVx(1):gVx(2)) ...
    - y(gVv(1)-1:gVv(2)-1,gVx(1):gVx(2))) ;
SV(:, :, 3) = thetaV * (y(gVx(1)+1:gVv(2)+1,gVx(1):gVx(2)) ...
    - y(gVv(1):gVv(2),gVx(1):gVx(2))) ;
%
sV(gVv(1):gVv(2),gVx(1):gVx(2)) = minMod(SV)/hv;
%
SX(:, :, 1) = (y(gVv(1):gVv(2),gVx(1)+1:gVx(2)+1) ...
    - y(gVv(1):gVv(2),gVx(1)-1:gVx(2)-1)) / 2 ;
SX(:, :, 2) = thetaX * (y(gVx(1):gVv(2),gVx(1):gVx(2)) ...
    - y(gVv(1):gVv(2),gVx(1)-1:gVx(2)-1)) ;
SX(:, :, 3) = thetaX * (y(gVx(1):gVv(2),gVx(1)+1:gVx(2)+1) ...
    - y(gVv(1):gVv(2),gVx(1):gVx(2))) ;
%
sX(gVv(1):gVv(2),gVx(1):gVx(2)) = minMod(SX)/hx;
%
for i = 1:del
    sX(:,gVx(1)-i) = sX(:,gVx(1));
    sX(:,gVx(2)+i) = sX(:,gVx(2));
end
for j = 1:del
    sV(gVv(1)-j,:) = sV(gVv(2),:);
    sV(gVv(2)+j,:) = sV(gVv(1),:);
end
%
```

end

Listing E.5: `slopeReconstruction.m`

E.6 minMod.m

```
function M = minMod(x)
%
M(:, :) = 0.25*abs(sign(x(:, :, 1))+sign(x(:, :, 2))) ...
    .* (sign(x(:, :, 1))+sign(x(:, :, 3))) .* min(abs(x), [], 3);
%
end
```

Listing E.6: minMod.m

E.7 sodSoln.m

```

function [n,u,p,e] = sodSoln(x,tCurrent,phiL,phiR,gam,x0)
%
% Script to compute theoretical sod solution at given time with
% given gamma.
%
%
% phiL and phiR are ordered as (n,u,p)
nL = phiL(1); uL = phiL(2); pL = phiL(3);
nR = phiR(1); uR = phiR(2); pR = phiR(3);
%
% This function does an iterative solve to compute the value of
% pStar.
pStar = pStarSolve(10^(-8));
thL = pL / nL;
thR = pR / nR;

aL = sqrt(3*thL);
aStarL = aL * (pStar/pL)^((gam-1)/(2*gam));
aR = sqrt(3*thR);

nFan = nL * (2/(gam+1) + (gam-1)*(uL-x/tCurrent) ...
    / (gam+1)/aL).^ (2/(gam-1));

nStarL = nL * (pStar/pL)^(1/gam);
nStarR = nR * ( (pStar/pR + (gam-1)/(gam+1)) ...
    / ((gam-1)*pStar/(gam+1)/pR + 1) );

uFan = 2/(gam+1) * (aL + (gam-1)*uL/2 + x/tCurrent);

uStar = uL - 2*aL / (gam-1) * ((pStar/pL)^((gam-1)/(2*gam)) - 1);

pFan = pL * (2/(gam+1) + (gam-1)*(uL-x/tCurrent) ...
    / (gam+1) / aL).^ (2*gam/(gam-1));

eL = 0.5*pL/nL;
eStarL = 0.5*pStar/nStarL;
eStarR = 0.5*pStar/nStarR;

```

```

eR = 0.5*pR/nR;

sHL = uL - aL;
sTL = uStar - aStarL;
%
% speed of the contact wave.
% lambda_2(U_star_L) = S_2 = lambda_2(U_star_R)
% Toro book: Equation 2.134, and page 96
sC = uStar; sR = uR + aR*( (gam+1)*pStar / (2*gam*pR) ...
+ 0.5*(gam-1)/gam)^(0.5);

nFunc = @(x) nL .* ((x<x0 + sHL*tCurrent)) ...
+ nFan .* ((x0 + sHL*tCurrent <= x) & (x < x0 + sTL*tCurrent)) ...
+ nStarL .* ((x0 + sTL*tCurrent <= x) & (x < x0 + sC*tCurrent)) ...
+ nStarR .* ((x0 + sC*tCurrent <= x) & (x < x0 + sR*tCurrent)) ...
+ nR .* ((x0 + sR*tCurrent <= x));

uFunc = @(x) uL .* ((x<x0 + sHL*tCurrent)) ...
+ uFan .* ((x0 + sHL*tCurrent <= x) & (x < x0 + sTL*tCurrent)) ...
+ uStar .* ((x0 + sTL*tCurrent <= x) & (x < x0 + sR*tCurrent)) ...
+ uR .* ((x0 + sR*tCurrent <= x));

pFunc = @(x) pL .* ((x<x0 + sHL*tCurrent)) ...
+ pFan .* ((x0 + sHL*tCurrent <= x) & (x < x0 + sTL*tCurrent)) ...
+ pStar .* ((x0 + sTL*tCurrent <= x) & (x < x0 + sR*tCurrent)) ...
+ pR .* ((x0 + sR*tCurrent <= x));

eFunc = @(x) eL .* ((x<x0 + sHL*tCurrent)) ...
+ (0.5*pFan./nFan) .* ((x0 + sHL*tCurrent <= x) ...
& (x < x0 + sTL*tCurrent)) ...
+ eStarL .* ((x0 + sTL*tCurrent <= x) & (x < x0 + sC*tCurrent)) ...
+ eStarR .* ((x0 + sC*tCurrent <= x) & (x < x0 + sR*tCurrent)) ...
+ eR .* ((x0 + sR*tCurrent <= x));

n = nFunc(x);
u = uFunc(x);
p = pFunc(x);
e = eFunc(x);

```

end

Listing E.7: `sodSoln.m`

E.8 pStarSolve.m

```
function pStar = pStarSolve(TOL)
%
% Newton Raphson to compute the pressure in Sod problem.
%
ERR = 1;
%
% Initial guess.
pOld = 0.9;
k = 0;
%
while ERR > TOL
    pNew = pOld - (sqrt(3)*pOld^(1/3) - sqrt(3) + ...
        2*(pOld-0.1)/sqrt(pOld+0.05)) / (sqrt(3)*pOld^(-2/3)/3 ...
        + 2*(pOld+0.05)^(-0.5) - (pOld-0.1)*(pOld+0.05)^(-1.5));
    ERR = abs(pOld-pNew) / (0.5*(pOld+pNew));
    pOld = pNew;
    k = k+1;
end
%
pStar = pNew;
%
```

Listing E.8: pStarSolve.m

Appendix F

Bibliography

- [1] P.L. Bhatnagar, E.P. Gross, and M. Krook. Model for collision processes in gases I: Small amplitude processes in charged and neutral one-component systems. *Phys. Rev.*, 94:511, 1954.
- [2] G. Bird. *Molecular Gas Dynamics and the Direct Simulation of Gas Flows, Volume 1*. Clarendon Press, 1994.
- [3] Alexander V Bobylev, Marzia Bisi, Maria Groppi, Giampiero Spiga, and Irina F Potapenko. A general consistent BGK model for gas mixtures. *Kinetic & Related Models*, 11(6), 2018.
- [4] S. Chapman and T.G. Cowling. *The Mathematical Theory of Non-uniform Gases: An Account of the Kinetic Theory of Viscosity, Thermal Conduction and Diffusion in Gases*. Cambridge Mathematical Library. Cambridge University Press, 1990.
- [5] F. Coron and B. Perthame. Numerical passage from kinetic to fluid equations. *SIAM J. Numer. Anal.*, 28:26–42, 1991.
- [6] Anaïs Crestetto, Christian Klingenberg, and Marlies Pirner. Kinetic/fluid micro-macro numerical scheme for a two component gas mixture. *Multiscale Modeling & Simulation*, 18(2):970–998, 2020.
- [7] C. Kristopher Garrett and Cory D. Hauck. A fast solver for implicit integration of the Vlasov-Poisson system in the Eulerian framework. *SIAM J. Sci. Comput.*, 40, 2018.

- [8] Jeffrey Haack, Cory D Hauck, Christian F Klingenberg, Marlies Pirner, and Sandra Warnecke. Numerical schemes for a multi-species BGK model with velocity-dependent collision frequency. *Journal of Computational Physics*, 473:111729, Jan 2023.
- [9] J.R. Haack, C.D. Hauck, and M.S. Murillo. A conservative, entropic multispecies BGK model. *J. Stat. Phys.*, 168:826–856, 2017.
- [10] Evan Habbershaw, Ryan S. Glasby, Jeffrey R. Haack, Cory D. Hauck, and Steven M. Wise. Asymptotic relaxation of moment equations for a multi-species, homogeneous BGK model. *arXiv preprint 2310.12885*, 2023.
- [11] B.B. Hamel. Kinetic model for binary gas mixtures. *Physics of Fluids*, 8:418–425, 1965.
- [12] H. Holway. New statistical models for kinetic theory: Methods of construction. *Phys. Fluids*, 9:1658, 1966.
- [13] Christian Klingenberg and Marlies Pirner. Existence, uniqueness and positivity of solutions for BGK models for mixtures. *Journal of Differential Equations*, 264, 09 2017.
- [14] Christian Klingenberg, Marlies Pirner, and Gabriella Puppo. A consistent kinetic model for a two-component mixture with an application to plasma. *Kinetic and Related Models*, 10(2):445–465, 2017.
- [15] Christian Klingenberg, Marlies Pirner, and Gabriella Puppo. A consistent kinetic model for a two-component mixture with an application to plasma. *Kinetic & Related Models*, 10(2), 2017.
- [16] L. Mieussens. Discrete velocity model and implicit scheme for the BGK equation of rarefied gas dynamics. *Math. Models and Methods Appl. Sci.*, 10:1121–1149, 2000.
- [17] L. Mieussens. Schemes for Boltzmann-BGK equation in plane and axisymmetric geometries. *J. Comput. Phys.*, 162:429–466, 2000.
- [18] L. Mieussens and H. Struchtrup. Numerical comparison of Bhatnagar-Gross-Krook models with proper Prandtl number. *Phys. Fluids*, 16:2797–2813, 2004.
- [19] Stéphane Mischler. Uniqueness for the BGK-equation in $\mathbb{R}^N$ and rate of convergence for a semi-discrete scheme. *Differential and Integral Equations*, 9(5):1119 – 1138, 1996.

- [20] Lorenzo Pareschi and Giovanni Russo. Implicit–explicit Runge–Kutta schemes and applications to hyperbolic systems with relaxation. *J. Sci. Comp.*, 25(1):129–155, 2005.
- [21] B. Perthame and M. Pulvirenti. Weighted L^∞ bounds and uniqueness for the Boltzmann BGK model. *Archive Rat. Mech. Anal.*, 125(3):289–295, 1993.
- [22] Benoît Perthame. Global existence to the BGK model of Boltzmann equation. *J. Diff. Eq.*, 82:191–205, 1989.
- [23] Sandra Pieraccini and Gabriella Puppo. Implicit–explicit schemes for BGK kinetic equations. *Journal of Scientific Computing*, 32:1–28, 2007.
- [24] Sandra Pieraccini and Gabriella Puppo. Implicit-explicit schemes for BGK kinetic equations. *J. Sci. Comput.*, 32:1–28, 2007.
- [25] Gabriella Puppo. Kinetic models of BGK type and their numerical integration. *arXiv preprint*, 2019.
- [26] E. Ringeissen. *Thesis*. PhD thesis, University of Paris VII, 1991.
- [27] L. Saint-Raymond. From the BGK model to the Navier-Stokes equations. *Ann. Sci. Ecole Norm. Sup.*, 36:271–317, 2003.
- [28] E.M. Shakhov. On the generalization of the Krook kinetic equation. *Izv. Russ. Acad. Sci. Fluid Dyn.*, 5:142–145, 1968.
- [29] H. Struchtrup. *Macroscopic Transport Equations for Rarefied Gas Flows: Approximation Methods in Kinetic Theory*. Springer Verlag, Berlin, Germany, 2005.
- [30] Eleuterio Toro. *Riemann Solvers and Numerical Methods for Fluid Dynamics: A Practical Introduction*. Springer, Germany, 2nd edition, 1999.
- [31] W.G. Vincenti and C.H. Kruger. *Introduction to Physical Gas Dynamics*. Krieger, Malabar, Florida, 1986.
- [32] K. Xu. A gas-kinetic BGK scheme for the Navier-Stokes equations and its connection with artificial dissipation and Godunov methods. *J. Comput. Phys.*, 171:289–335, 2001.