

Bachelorarbeit  
zur Erlangung des akademischen Grades  
Bachelor of Science

Zachary Schellin  
376930

December 6, 2020

# Contents

<b>Contents</b>	<b>2</b>
<b>1 Introduction</b>	<b>3</b>
1.1 State of the art . . . . .	3
1.2 Objective of this thesis . . . . .	4
1.3 Thesis outline . . . . .	4
<b>2 The BGK Equation</b>	<b>4</b>
<b>3 Reduced Order Algorithms</b>	<b>5</b>
3.1 Data Sampling . . . . .	5
3.2 POD . . . . .	5
3.3 Autoencoders . . . . .	6
3.4 Architectures . . . . .	6
3.4.1 Fully Connected . . . . .	6
3.4.2 Convolutional . . . . .	6
3.4.3 Training . . . . .	6
3.5 Reducd Order Model . . . . .	6
<b>4 Results and Latent Manifold Properties</b>	<b>6</b>
4.1 Evaluation Methods . . . . .	6
4.2 Results . . . . .	6
4.2.1 POD . . . . .	6
4.2.2 Fully-Connected Autoencoder . . . . .	7
4.2.3 Convolutional Autoncoder . . . . .	8
4.3 Discussion and Outlook . . . . .	8
<b>References</b>	<b>8</b>

# 1 Introduction

The Bhatnagar, Gross, Krook equation (BGK) is a kinetic collision model of ionized and neutral gases valid for rarefied as well as other pressure regimes [1]. Generating data of such a flow field is essential for various industry and scientific applications[**REF**]. With the intention to reduce time and cost during the data generating process, experiments were substituted with computational fluid dynamics (CFD) computations. Consequently reduced-order models (ROMs) coupled to aforementioned computations were introduced to further the reduction of time and cost. The thriving field of artificial intelligence operates in model order reduction for data visualization/analysis since the 80's (Quelle?) and has now surfaced in fluid mechanics. This thesis will cover the use of artificial intelligence for model order reduction in fluid mechanics.

## 1.1 State of the art

State of the art model reduction of dynamical systems is done via proper orthogonal decomposition (POD) which is an algorithm feeding on the idea of singular value decomposition (SVD)[2][3]. POD captures a low-rank representation on a linear manifold. So called POD modes, derived from SVD, describe the principle components of a problem which can be coupled within a Galerkin framework to produce an approximation of lower rank.

$$f(x) \approx \tilde{f}(x) \quad \text{with} \quad rk(f(x)) \gg rk(\tilde{f}(x)) \quad (1)$$

Bernard et al. use POD-Galerkin with an additional population of their snapshot database via optimal transport for the proposed BGK equation, bisecting computational run time (cost) in conjunction with an approximation error of 1 % [4]. Artificial intelligence in the form of autoencoders replacing the POD within a Galerkin framework is evaluated against the POD performance by Kookjin et al. for advection-dominated problems[5] resulting in sub 0.1% errors. An additional time inter- and extrapolation is evaluated. Using machine learning/ deep learning for reduced order modeling in CFD is a novel approach although "the idea of autoencoders has been part of the historical landscape of neural networks for decades"[6, p.493]. Autoencoders, or more precisely learning internal representations by the delta rule (backpropagation) and the use of hidden units in a feed forward neural network architecture, premiered by Rumelhart et al. (1986) [7]. Through so called hierarchical training Ballard et al.(1987) introduce a strategy to train auto autoassociative networks (nowadays referred to as autoencoders), in a reasonable time promoting further development despite computational limitations [8]. The so called bottleneck of autoencoders yields a non-smooth and entangled representation thus being uninterpretable by practitioners[9] leading to developments in this field. Rifai et al. introduce the contractive autoencoder (CAE) for classification tasks (2011), with the aim to extract robust features which are insensitive to input variations orthogonal to the low-dimensional non-linear manifold by adding a penalty on the frobenius norm of the intrinsic variables with respect to the input, surpassing other classification algorithms [9]. Subsequent development emerges with the manifold tangent classifier (MTC) [10]. A local chart for each datapoint is obtained hence characterizing the manifold which in turn improves classification performance. On that basis a generative process for the CAE is developed. Through movements along the manifold with directions defined by the Ja-

cobian of the bottleneck layer with respect to the input  $\vec{x}_m = JJ^T$ , sampling is realized [11].

## 1.2 Objective of this thesis

Due to the non-linearity of transport problems in particular shock fronts, the construction of a robust ROM for those cases poses several challenges. Proper orthogonal decomposition (POD) and it's numerous variants like shifted-POD[?], POD-Galerkin[?], POD+I [?] to name only a few of them, try to solve this problem by.....

## 1.3 Thesis outline

# 2 The BGK Equation

The Knudsen number eq. (2) introduced by Danish physicist Martin Knudsen is a measure for the rarefaction of gases. In eq. (2)  $\lambda$  represents the mean free path and  $L$  the characteristic length [4] of a particle. The mean free path describes the average distance a particle may travel between successive impacts [WIKI]. For idealized gases the mean free path can be calculated via eq. (3). In eq. (3)  $k_b$  is the Boltzman constant,  $p$  is the total pressure,  $T$  is the thermodynamic temperature and  $d$  is the hard shell diameter [WIKI].

$$Kn = \frac{\lambda}{L} \quad (2)$$

$$\lambda = \frac{k_b T}{\sqrt{2} \pi d^2 p} \quad (3)$$

For Knudsen numbers  $Kn > 10^{-2}$  collisions are predominant in comparison to free transport, whereas for  $Kn < 10^{-2}$  free transport becomes the predominant behavior compared to collision[4]. The Boltzmann equation eq. (4) describes the dynamics of a gas flow, where  $f$  is the probability density distribution function for a particle at point  $\mathbf{x} \in \mathbb{R}^3$  with velocity  $\xi \in \mathbb{R}^3$  at time  $t \in \mathbb{R}$ .

$$\partial_t f(\mathbf{x}, \xi, t) + \xi \Delta_x f(\mathbf{x}, \xi, t) = Q(f, f) \quad (4)$$

Originally  $Q$  is often the binary Boltzmann collision term, which can be intractable in practice [1]. Binary, because Thus the BGK equation utilizes the BGK-Operator for  $Q(f, f)$  eq. (5)[4].

$$Q(f, f) = \frac{M_f(\mathbf{x}, \xi, t) - f(\mathbf{x}, \xi, t)}{\tau(\mathbf{x}, t)} \quad (5)$$

$$M_f(\mathbf{x}, \xi, t) = \frac{\rho(\mathbf{x}, t)}{(2\pi T(\mathbf{x}, t))^{\frac{3}{2}}} \exp\left(-\frac{\|\xi - U(\mathbf{x}, t)\|^2}{2T(\mathbf{x}, t)}\right) \quad (6)$$

$$\tau(\mathbf{x}, t) = \frac{Kn}{\rho(\mathbf{x}, t) T^{1-\nu}(\mathbf{x}, t)} \quad (7)$$

For many kinetic gas problems it is sufficient to replace the complexity of the Boltzmann collision term by a mean-free-path approach.  $\tau$  is referred to as the relaxation time, considering the fact, that collisions tend to relax the distribution function to an equilibrium

state  $f_0$ , where  $f_0 = M_f$  the equilibrium state is approximated by a Maxwellian distribution function. In eq. (6)  $U(\mathbf{x}, t) = (u(\mathbf{x}, t), v(\mathbf{x}, t), w(\mathbf{x}, t))^T$  is the macroscopic velocity,  $T(\mathbf{x}, t) \in \mathbb{R}$  and  $\rho(\mathbf{x}, t) \in \mathbb{R}$  is the temperature and density of the gas respectively. In eq. (7)  $\nu \in \mathbb{R}$  is the exponent of the viscosity law of the gas.

### 3 Reduced Order Algorithms

#### 3.1 Data Sampling

For the autencoder using fully connected layers, the input vectors  $y_o \in \mathbb{R}$  of size  $n_{input} = n_\xi = 40$  are arranged in the sampling matrix  $S_{AE} \in \mathbb{R}^{5000 \times 40}$  as seen in eq. (8) resulting in  $n_S = 5000$  available samples. Note that the POD uses the same matrix transposed  $S_{AE}^T$  as input.

$$S_{AE} = \begin{bmatrix} f(\xi_1, t_1, x_1) & \cdots & f(\xi_n, t_1, x_1) \\ f(\xi_1, t_1, x_2) & \cdots & f(\xi_n, t_1, x_2) \\ \vdots & \vdots & \vdots \\ f(\xi_1, t_1, x_n) & \cdots & f(\xi_n, t_1, x_n) \\ f(\xi_1, t_2, x_1) & \cdots & f(\xi_n, t_2, x_1) \\ \vdots & \vdots & \vdots \\ f(\xi_1, t_n, x_n) & \cdots & f(\xi_n, t_n, x_n) \end{bmatrix} \quad S_{Conv} = \begin{bmatrix} n_{Filters} & f(\xi_1, \mathbf{t}, \mathbf{x}) \\ n_{Filters} & f(\xi_2, \mathbf{t}, \mathbf{x}) \\ \vdots & \vdots \\ n_{Filters} & f(\xi_n, \mathbf{t}, \mathbf{x}) \end{bmatrix} \quad (9)$$

(8)

Convolutional autoencoders use a different sampling matrix  $S_{Conv}$  due to their two dimensional capability resulting in  $n_S = 40$  available samples eq. (9).  $n_{Filters}$  varies over the succeeding layers, growing with the shrinkage of  $(\mathbf{t}, \mathbf{x})$ .

#### 3.2 POD

The singular value decomposition of the input  $X$  [REF to Section 1] gives the optimal low-rank approximation  $\tilde{X}$  of  $X$  eq. (11)[Eckard-Young]. Figure 1 shows the singular values (left) and the cumulative energy (right) derived from eq. (10):

$$S_N = \sum_{k=1}^N a_k \quad \text{with a sequence} \quad \{a_k\}_{k=1}^n \quad (10)$$

$$\underset{\tilde{X}, s.t. rank(\tilde{X})=r}{\operatorname{argmin}} \quad ||X - \tilde{X}||_F = \tilde{U} \tilde{\Sigma} \tilde{V}^* \quad (11)$$

### 3.3 Autoencoders

### 3.4 Architectures

#### 3.4.1 Fully Connected

The autoencoder architecture 1.0 is a composition of five fully connected layers, eq. (12). The subscript F denotes fully connected layers opposed to the subscript C which denotes convolutional layers in section 3.4.2.

$$y_p = f_F^5(f_F^4(f_F^3(f_F^2(f_F^1(y_o)))))) \quad (12)$$

Two input/output layers of size  $n_{input} = 40$ , two hidden layers of size  $n_{hidden} = 20$  and one "bottleneck" layer of size  $n_{code} = 3$ . The trainingloss is the mean-squared error (MSE) between the input and output of the autoencoder, where  $y_0$  is the original inpput vector and  $y_p$  is the reconstructed output vector eq. (13).

$$MSE = \frac{(y_p - y_o)^2}{n_{batch}} \quad (13)$$

#### 3.4.2 Convolutional

The convolutional autoencoder architecture 1.1 is a composition of six convolutional and three fully conected layers, eq. (14).

$$y_p = f_C^9(f_C^8(f_C^7(f_F^6(f_F^5(f_F^4(f_C^3(f_C^2(f_C^1(y_o)))))))))) \quad (14)$$

#### 3.4.3 Training

During training every 1000 epochs a sample against its prediction was printed in order to link the value of the L1-Loss to a prediction. Using this method a first verification of the model was achieved. Continuing the search for any possible shortage of the models performance, that this method could not cover, eg. samples lying between every 1000 sample, that the model was not able to reconstruct correctly, a second verification process is conducted.

### 3.5 Reducd Order Model

## 4 Results and Latent Manifold Properties

### 4.1 Evaluation Methods

### 4.2 Results

#### 4.2.1 POD

The first five singular values give an accurate approximation  $\tilde{X}$  of  $X$ . As a means to evaluate the low-rank approximation of  $X$  we will compare the density derived from ??, computed from  $X$  and  $\tilde{X}$ .

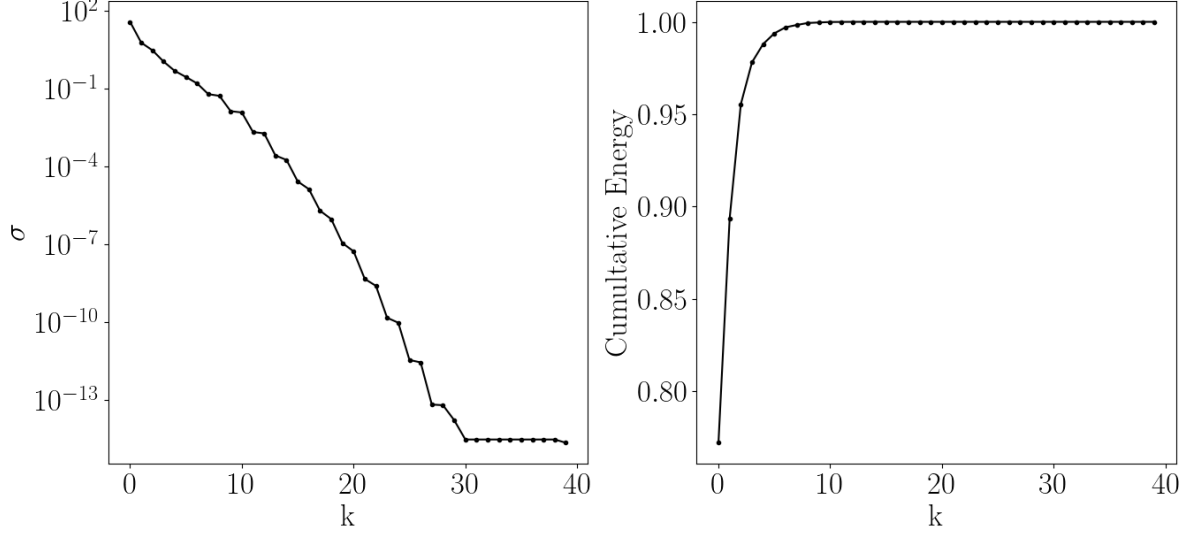


Figure 1: Singular Values (left) and cumulative enrgy (right) over the number of singular values

#### 4.2.2 Fully-Connected Autoencoder

The search for a reduced model of the BGK equation yields a first reduction and analysis of the provided data with a low Knudsen number of  $Kn = 0.00001$ . For this flow field the Navier Stokes equations are still valid. Analysing the batch size for the architecture 1.0 the test errors in table 1 can be produced. For the evaluation of the prediction the

Architecture	1.0
Batch Size	L2-Error
64	0.008
32	0.0049
16	0.0038
8	0.0037
4	0.0026
2	0.0021

Table 1: L2-Error over Batch-Size

Autoencoder produces normalized conservative quantities are analysed. The conservative quantities are the total energy  $E$ , the density  $\rho$  and the impulse  $\rho u$ . Normalization is done over the temporal mean as seen in eq. (15a) to eq. (15c).

$$\frac{\frac{d}{dt} \int E dx}{\bar{E}} = 0 \quad \text{with} \quad \bar{E} = \frac{\iint E dt dx}{\Delta t} \quad (15a)$$

$$\frac{\frac{d}{dt} \int \rho dx}{\bar{\rho}} = 0 \quad \text{with} \quad \bar{\rho} = \frac{\iint \rho dt dx}{\Delta t} \quad (15b)$$

$$\frac{\frac{d}{dt} \int \rho u dx}{\bar{\rho} u} = 0 \quad \text{with} \quad \bar{\rho} u = \frac{\iint \rho u dt dx}{\Delta t} \quad (15c)$$

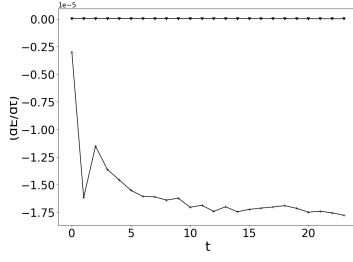


Figure 2: Normalized total change in Energy  $\hat{E}$  over time

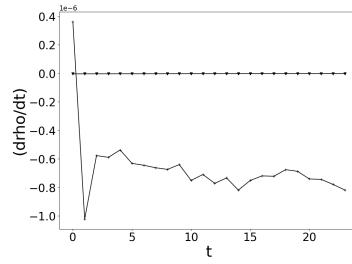


Figure 3: Normalized total change in Energy  $\hat{\rho}$  over time

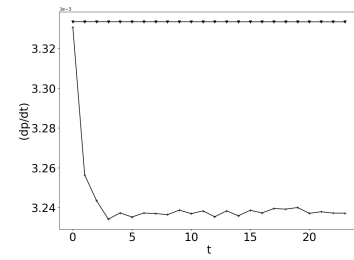


Figure 4: Normalized total change in Energy  $\hat{p}$  over time

#### 4.2.3 Convolutional Autocoder

### 4.3 Discussion and Outlook

## References

- [1] Bhatnagar, Gross, and Krook. A model for collision processes in gases. 1954.
- [2] Thomas Franz. *Reduced-order modeling of steady transonic flows via manifold learning*. 2016.
- [3] Steve L. Brunton and J. Nathan Kutz. *Data driven science and engineering*. 2019.
- [4] Florian Bernard, Angelo Iollo, and Sebastian Riffaud. Reduced-order model for the bgk equation based on pod and optimal transport. 2018.
- [5] Kookjin Lee and Kevin T. Carlberg. Model reduction of dynamical systems on nonlinear manifolds using deep convolutional autoencoders. 2019.
- [6] Ian J. Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. MIT Press, Cambridge, MA, USA, 2016. <http://www.deeplearningbook.org>.
- [7] D.E. Rumelhart, G.E. Hinton, and R.J. Williams. Learning internal representations by error propagation. 1986.
- [8] Dana H. Ballard. Modular learning in neural networks. 1987.
- [9] Salah Rifai, Pascal Vincent, Xavier Muller, Xavier Glorot, and Yoshua Bengio. Contractive auto-encoders: Explicit invariance during feature extraction. 2011.
- [10] Salah Rifai, Yann N Dauphin, Pascal Vincent, Yoshua Bengio, and Xavier Muller. *The Manifold Tangent Classifier*. Curran Associates, Inc., 2011.
- [11] Salah Rifai, Yoshua Bengio, Yann Dauphin, and Pascal Vincent. A generative process for sampling contractive auto-encoders, 2012.