

## ON THE INVARIANT REGION FOR COMPRESSIBLE EULER EQUATIONS WITH A GENERAL EQUATION OF STATE

*Dedicated to Professor Shuxing Chen on the occasion of his 80th birthday*

HAILIANG LIU\*

Iowa State University, Mathematics Department  
Ames, IA 50011, USA

FERDINAND THEIN

Otto-von-Guericke-Universität, Universitätsplatz 2  
Magdeburg, 39106, Germany

ABSTRACT. The state space for solutions of the compressible Euler equations with a general equation of state is examined. An arbitrary equation of state is allowed, subject only to the physical requirements of thermodynamics. An invariant region of the resulting Euler system is identified and the convexity property of this region is justified by using only very minimal thermodynamical assumptions. Finally, we show how an invariant-region-preserving (IRP) limiter can be constructed for use in high order finite-volume type schemes to solve the compressible Euler equations with a general constitutive relation.

**1. Introduction.** The dynamical evolution of a fluid is determined by the principles of conservation of mass, momentum, and energy. Material properties strongly influence the structure and dynamics of waves in any continuum-mechanical system, as shown by Menikoff and Plohr in their instructive and comprehensive work [14]. There the Riemann problem for the compressible Euler equations is discussed in detail. Constitutive relations are understood as the most fundamental level for fluids, where they are embodied in an equilibrium equation of state (EOS) and the physical principles of thermodynamics impose stringent constraints on the equation of state, cf. Landau and Lifshitz [13].

In this work we investigate the invariant region for solutions to the Euler equations with a general EOS. This system arises in fluid dynamics and is the paradigm for systems of hyperbolic conservation laws. It provides the motivation for many of the central ideas in the analysis of quasilinear hyperbolic partial differential equations. The Euler equations we consider here are given by

$$\partial_t \rho + \nabla_x \cdot (\rho \mathbf{u}) = 0, \tag{1.1a}$$

$$\partial_t (\rho \mathbf{u}) + \nabla_x \cdot (\rho \mathbf{u} \otimes \mathbf{u}) + \nabla_x P = 0, \tag{1.1b}$$

$$\partial_t E + \nabla_x \cdot ((E + P) \mathbf{u}) = 0, \tag{1.1c}$$

---

2020 *Mathematics Subject Classification.* Primary: 35L65, 76N15; Secondary: 65M08.

*Key words and phrases.* Euler equations, entropy, invariant region, equation of state, fundamental derivative.

Hailiang Liu was partially supported by the National Science Foundation under Grant DMS1812666.

\* Corresponding author: Hailiang Liu.

where  $\rho$  denotes the mass density,  $\mathbf{u}$  the fluid velocity,  $E$  the specific total energy, and  $P$  the pressure. The pressure must be specified through a constitutive relation, the EOS, which characterizes the fluid material. The qualitative character of the solutions of the conservation laws depends crucially on the thermodynamic properties of the material. For such a general system as we consider it here, a complete mathematical theory is still unavailable. A natural choice would be to consider solutions confined in a convex domain in state space, which is invariant either provable to hold in time along dynamic equations or known from physical considerations. An invariant region to (1.1) is an open set in phase space  $\mathbb{R}^d$  such that if initial data lie in this set, then the solution will remain in this set. It was proved by Hoff [7] that an invariant region for one dimensional hyperbolic conservation laws must be convex. The invariant region is important also for numerical algorithms to solve these equations. However, there are no rigorous convergence results to solutions for any numerical schemes approximating multi-dimensional systems of conservation laws. Nevertheless, if the numerical solution lies in some important invariant region, this would be an indication that the problem is well-posed or the solution is physically meaningful. Some first order methods such as the Lax-Friedrich method and the Godunov method are known to satisfy the invariant-region-preserving (IRP) property [3, 4, 20]. In the context of continuous finite elements, the IRP property has been studied by Guermond and Popov [5] using the first order approximation to solve general hyperbolic conservation law systems.

However, it is a rather difficult task to preserve an invariant region by a high order numerical method unless some nonlinear limiter is frequently imposed; cf. second-order limitation techniques are used in [12] to enforce the minimum entropy principle so that oscillations in numerical solutions may be damped. Indeed, recent efforts have been using limiting techniques to construct high order schemes, including maximum-principle-preserving schemes for scalar conservation laws (see [21]) and positivity-preserving schemes for hyperbolic systems including the compressible Euler equations (see e.g. [16, 22, 24]). The work by Zhang and Shu in [23] introduced a limiter to preserve the minimum-entropy-principle.

The emphasis in [9, 10, 11] is on the notion of invariant regions and how to numerically preserve an invariant region by high order discontinuous Galerkin schemes. We observe that the ideas for the high order IRP schemes studied in [10] can be extended to general hyperbolic conservation law systems as long as it features a convex invariant region. In this work we show how to identify a convex invariant region for the Euler system with a general EOS and present a simple IRP limiter which when coupled with any high order finite volume type methods can pull numerical solutions back to the invariant region. It is hoped that this information will be of service to the designers of numerical approximations of this important class of equations.

The remainder of the paper is organized as follows. In the next section, we review special models of the Euler equations and present the known invariant region for each of them. Further we discuss the basic structure for the full system of Euler equations, and present the main result to be proved in Section 3 and 4. In Section 3, we show a region defined by the entropy to be invariant. In Section 4, we provide a justification for a sub-level set of the internal energy. Section 5 is devoted to the IRP limiter and discussion on high order IRP schemes. Finally, a specific example of the general Euler system is given in Section 6.

**2. Equation of state.** To obtain a closed system it is necessary to prescribe an equation of state, which relates the pressure  $P$  to the density and the total energy. It is provided by thermodynamics. The following three distinct situations are important:

**2.1. Pressureless gases.** The pressure  $P$  vanishes and so the total energy reduces to just the kinetic energy:  $E = \frac{1}{2}\rho|\mathbf{u}|^2$ . The equations (1.1) take the form

$$\begin{aligned}\partial_t \rho + \nabla_x \cdot (\rho \mathbf{u}) &= 0, \\ \partial_t (\rho \mathbf{u}) + \nabla_x \cdot (\rho \mathbf{u} \otimes \mathbf{u}) &= 0,\end{aligned}$$

and the energy equation in (1.1) follows formally from the continuity and momentum equations. The pressureless Euler system has been proposed as a simple model describing the formation of galaxies in the early stage of the universe. There is a maximum principle for the velocity field  $u$ , though its gradient can become discontinuous – shock formation. If fluid elements meet at the same location, then they stick together to form larger compounds and so  $\rho(x, t)$  can have singular parts (in particular, Dirac measures). Consequently the system must be understood in the sense of distributions, and admits an invariant region of the form

$$\Sigma = \{(\rho, \mathbf{m}) \mid \rho > 0, \quad -\alpha_i \rho \leq m_i \leq \alpha_i \rho, i = 1, \dots, d\}$$

for some  $\alpha_i > 0$  with  $\alpha_i = \sup_x |u_i(x, 0)|$ , and  $\mathbf{m} = \rho \mathbf{u}$ .

**2.2. Isentropic gases.** In this regime, the thermodynamical entropy of the fluid is assumed to be constant in space and time. Consequently, the pressure is a function of the density only. For polytropic gases,

$$P(\rho) = U'(\rho)\rho - U(\rho), \quad \rho \geq 0,$$

so that  $\rho e = U(\rho)$  is convex in  $\rho$  and  $E = \frac{1}{2}\rho|\mathbf{u}|^2 + U(\rho)$ . As in the pressureless case, the energy equation in (1.1) follows formally from the continuity and the momentum equation. For solutions with shocks, the continuity and the momentum equation must be considered in the sense of distributions, and the energy equation does no longer follow automatically. A physically reasonable relaxation is to assume that no energy can be created by the fluid: the energy equality in (1.1) must be replaced by the inequality

$$\partial_t \left( \frac{1}{2}\rho|\mathbf{u}|^2 + U(\rho) \right) + \nabla \cdot \left( \left( \frac{1}{2}\rho|\mathbf{u}|^2 + U'(\rho)\rho \right) \mathbf{u} \right) \leq 0$$

in distributional sense. If the above inequality is strict, this means physically that mechanical energy is transformed into heat. A form of energy that is not accounted for by the model. For one dimensional flows, we find Riemann invariants of the form

$$R_{\pm} = \frac{m}{\rho} \pm \int^{\rho} \sqrt{\frac{U''(r)}{r}} dr,$$

so that for  $\lambda_{\pm} = u \pm \sqrt{U''(\rho)\rho}$ ,

$$\begin{aligned}\partial_t R_- + \lambda_- \partial_x R_- &= 0, \\ \partial_t R_+ + \lambda_+ \partial_x R_+ &= 0.\end{aligned}$$

The corresponding invariant region is given by

$$\Sigma = \{(\rho, m)^{\top} \mid R_+ \leq c_2, \quad R_- \geq c_1\},$$

where  $c_i$  is determined by the given initial data.

**2.3. Full Euler equations.** We consider a polytropic gas with the adiabatic coefficient  $\gamma_0 > 1$ . Then the pressure is given in terms of the conserved quantities  $\mathbf{w} := (\rho, \mathbf{m}, E)^\top$  by the formula

$$P = (\gamma_0 - 1) \left( E - \frac{|\mathbf{m}|^2}{2\rho} \right) = (\gamma_0 - 1) \left( E - \frac{1}{2}\rho|\mathbf{u}|^2 \right) = (\gamma_0 - 1)\rho e.$$

Density and pressure define the specific thermodynamical entropy, given as

$$s := \log \left( \frac{P}{c\rho^{\gamma_0}} \right), \quad c = k(\gamma_0 - 1) > 0$$

in the case of polytropic gases. The internal energy density has the form

$$e = k \exp(s) \rho^{\gamma_0 - 1}.$$

The specific entropy  $s$  must be constant along characteristics:

$$\partial_t s + \mathbf{u} \cdot \nabla_x s = 0.$$

But since the solutions to the compressible Euler equations may become discontinuous in finite time, the physically reasonable relaxation is that the specific entropy should be non-decreasing forward in time, which expresses the second law of thermodynamics. It follows that

$$\inf s(x, t) \geq \inf s(x, 0) =: s_0,$$

where  $s(x, 0)$  is the initial specific entropy. Hence the region

$$\Sigma = \{ \mathbf{w} \mid \rho > 0, P > 0, s \geq s_0 \}$$

is invariant and convex (see e.g., [10]).

**2.4. General EOS.** From now on, we consider a general constitutive relation derived from thermodynamics. By the second law of thermodynamics we know that there exists a function  $s$ , called entropy, which is a twice differential function of form

$$s = G(e, v),$$

where  $e$  is the specific internal energy and  $v = 1/\rho > 0$  is the specific volume. This formula means

$$s(x, t) = G(e(x, t), v(x, t)),$$

for  $t > 0$  and  $x$  in the fluid region. The thermodynamic properties of a material are embodied in this relation. This fundamental equation contains all the information we need to close the Euler system. As usual, we assume that  $s$  has the following properties:

- (I)  $G(e, v)$  is concave in  $(e, v)$ ,
- (II)  $\partial_e G(e, v) > 0$ .

As a result,  $s = G(e, v)$  can be reformulated as

$$e = F(s, v),$$

which for  $G$  strict concave can be shown to have the following properties:

- (I)  $F(s, v)$  is convex in  $(s, v)$ ,
- (II)  $\partial_s F(s, v) = \frac{1}{G_e} > 0$ .

We can define the pressure  $P$  and the temperature  $\theta$  by

$$P = -F_v, \quad \theta = F_s,$$

in accordance with the fundamental thermodynamic identity

$$de = -Pdv + \theta ds,$$

i.e., the first law of thermodynamics (see e.g., [2]). Imposed by the Clausius-Duhem inequality, the entropy inequality and the energy equation take the form

$$0 \leq \rho(\theta - F_s)(\partial_t s + \mathbf{u} \cdot \nabla_x s) + (-P - F_v)\nabla_x \mathbf{u},$$

which must hold for all admissible thermodynamic processes. The above choice is a valid candidate for dynamic equations and also consistent with classical formulas for a simple fluid. Rewriting the above we also have

$$ds = \frac{1}{\theta} de + \frac{P}{\theta} dv.$$

Hence  $\theta = 1/G_e$  and  $P = \theta G_v$  and thus we have

$$P = \frac{G_v}{G_e}. \quad (2.1)$$

This formula for the pressure enables us to close the Euler system of equations. For our purposes we further follow Menikoff and Plohr and define dimensionless quantities as in [14]. In the situation under consideration it is possible to describe the fundamental properties of the studied system using these three quantities.

**Definition 2.1** (Dimensionless quantities). Given a thermodynamic system described by the fundamental equation  $e = F(s, v)$  the following dimensionless quantities can be defined:

- (i) The *adiabatic exponent* or synonymous the *dimensionless sound speed*

$$\gamma = \frac{v}{P} \frac{\partial^2 F}{\partial v^2}. \quad (2.2)$$

- (ii) The *Grüneisen coefficient*

$$\Gamma = -\frac{v}{\theta} \frac{\partial^2 F}{\partial s \partial v}. \quad (2.3)$$

- (iii) The *dimensionless specific heat*

$$g = \frac{Pv}{\theta^2} \frac{\partial^2 F}{\partial s^2}. \quad (2.4)$$

- (iv) The *fundamental derivative*

$$\mathcal{G} = -\frac{1}{2}v \frac{\frac{\partial^3 F}{\partial v^3}}{\frac{\partial^2 F}{\partial v^2}} = -\frac{1}{2}v \frac{\frac{\partial^2 P}{\partial v^2}}{\frac{\partial P}{\partial v}}. \quad (2.5)$$

These dimensionless quantities are very helpful in discussing the Riemann problem for the Euler equations. Further they have several useful interpretations which we will not present here in detail. We therefore again highly recommend [14] and just make some remarks.

**Remark 1.**

- (i) Systems with a convex fundamental equation  $F(s, v)$  are called thermodynamically stable.

(ii) The convexity of  $F(s, v)$  can be expressed in terms of  $g, \gamma$  and  $\Gamma$ , i.e.

$$F(s, v) \text{ is convex} \quad \Leftrightarrow \quad g \geq 0, \gamma \geq 0 \text{ and } g\gamma \geq \Gamma^2. \quad (2.6)$$

- (iii) The thermodynamic laws do not confine the sign of  $\mathcal{G}$ .
- (iv) If  $\gamma > 0$  holds, the system of Euler equations is strictly hyperbolic.
- (v) The characteristic families corresponding to the largest and smallest eigenvalue of the full Euler system are genuine nonlinear iff  $\mathcal{G} \neq 0$  holds. In particular shock waves are compressive iff  $\mathcal{G} > 0$ .
- (vi) For the polytropic gas presented above we obtain (cf. [14])

$$\gamma = \gamma_0, \quad g = \gamma_0 - 1, \quad \Gamma = \gamma_0 - 1 \text{ and } \mathcal{G} = \frac{\gamma_0 + 1}{2}.$$

Thermodynamic stability requires  $\gamma_0 \geq 1$  and thus using this EOS the system of Euler equations is strictly hyperbolic.

For further useful remarks and insights we refer to [14]. Assuming  $\mathcal{G} > 0$  and convexity of  $F(s, v)$  it is possible to prove the following useful inequality

$$\frac{Pv}{e} < 2\gamma.$$

The proof relies on the convexity of the isentropes in the  $P - v$ -plane, see [14] (p. 95, Lem. 4.3). In view of equation (2.1) this gives

$$\frac{vG_v}{eG_e} < 2\gamma. \quad (2.7)$$

From now on we assume the convexity of the fundamental equation, see (2.6) and additionally we demand  $\mathcal{G} > 0$ . For consequences when  $\mathcal{G} \leq 0$  we refer to Müller and Voss [15] and again to [14].

We recall that for  $2 \times 2$  systems such as the isentropic Euler system, an invariant region can be described by two Riemann invariants [18]. For the Euler equations (1.1) with a general EOS, we shall identify a region which is both convex and invariant. The convexity may be determined by the structure conditions of the EOS, and the invariance property depends on the dynamic equations. The main result can be stated as follows.

**Theorem 2.2.** *Consider system (1.1) and let the specific internal energy be given by  $e = F(s, v)$  with the properties (I) and (II) given above. The quantities  $\gamma$ ,  $g$ ,  $\Gamma$  and  $\mathcal{G}$  are given as in Definition 2.1. Furthermore the following constraint holds*

$$\mathcal{G} > 0.$$

*Then there exists a convex invariant region for the system of Euler equations:*

$$\Sigma = \{\mathbf{w} = (\rho, \mathbf{m}, E) \mid \rho > 0, R > 0, q < 0\}, \quad (2.8)$$

where

$$R = E - \frac{|\mathbf{m}|^2}{2\rho}, \quad q = \rho(\inf s_0 - s).$$

Throughout the following two sections we will prove this theorem.

**3. Entropy minimization.** Using  $e_s = \theta$  and  $e_\rho = P/\rho^2$ , and the energy equation, one can derive

$$0 = \partial_t E + \nabla_x \cdot ((E + P)\mathbf{u}) = \theta(\partial_t(\rho s) + \nabla_x \cdot (\rho s \mathbf{u})),$$

leading to the transport equation

$$\partial_t s + \mathbf{u} \cdot \nabla_x s = 0,$$

at least for solution without shocks. For weak solutions, we need to recall the celebrated Clausius-Duhem inequality

$$\partial_t(\rho s) + \nabla_x \cdot (\rho s \mathbf{u}) \geq \frac{r\rho}{\theta} - \nabla_x \cdot \left( \frac{f}{\theta} \right),$$

which, when both the heat supply  $r$  and heat flux  $f$  vanish, leads to

$$\partial_t s + \mathbf{u} \cdot \nabla_x s \geq 0.$$

In either case, we have the following entropy minimization principle:

$$\inf s(x, t) = \inf s(x, 0) =: s_0, \quad s(x, t) - s_0 \geq 0.$$

This ensures that the region  $\{\mathbf{w} \mid \rho > 0, q \leq 0\}$  with  $q = \rho(s_0 - s)$  is invariant. Moreover, we have the following.

**Proposition 1.** *Let  $q = \rho(s_0 - s)$ , then  $q$  is convex in terms of  $\mathbf{w} := (\rho, m, E)$  with  $m = |\mathbf{m}|$ .*

*Proof.* For simplicity of computation, we take  $s_0 = 0$  so that  $q = -\rho G(e, v)$ . A direct calculation gives

$$\begin{aligned} q_\rho &= -G + vG_e\xi + vG_v, \quad \xi = E - vm^2 \\ q_m &= mvG_e, \\ q_E &= -G_e \end{aligned}$$

and second order derivatives

$$\begin{aligned} q_{\rho\rho} &= -v^3(G_{vv} + G_{ee}\xi^2 + 2\xi G_{ev} - m^2 G_e), \\ q_{\rho m} &= -mv^2 G_e - G_{ee}v^3 m\xi - mv^3 G_{ev}, \\ q_{E\rho} &= v^2 \xi G_{ee} + v^2 G_{ev}, \\ q_{mm} &= vG_e - v^3 m^2 G_{ee}, \\ q_{mE} &= v^2 m G_{ee}, \\ q_{EE} &= -vG_{ee}, \\ q_{E\rho} &= v^2 \xi G_{ee} + v^2 G_{ev}. \end{aligned}$$

Note that from concavity of  $s = G(e, v)$  we have

$$G_{vv} < 0, \quad G_{ee} < 0, \quad G_{vv}G_{ee} > G_{ev}^2. \quad (3.1)$$

We shall only use this basic assumption to show  $q$  is convex in  $\mathbf{w}$ . In order to do so, we only need to show

$$q_{\rho\rho} > 0, \quad A = q_{\rho\rho}q_{mm} - q_{m\rho}^2 > 0 \text{ and } B = \det(D^2 q) > 0.$$

First from  $q_{\rho\rho}$  we have

$$\begin{aligned} q_{\rho\rho} &\geq -v^3 \left( G_{vv} + G_{ee}\xi^2 + 2|\xi| \sqrt{G_{vv}G_{ee}} - m^2 G_e \right) \\ &= v^3 m^2 G_e + v^3 (\sqrt{|G_{vv}|} + |\xi| \sqrt{|G_{ee}|})^2 > 0. \end{aligned}$$

Next we estimate  $A$  as follows.

$$\begin{aligned} A &= q_{\rho\rho}q_{mm} - q_{\rho m}^2 \\ &= -v^4(G_e - v^2m^2G_{ee})(G_{vv} + G_{ee}\xi^2 + 2\xi G_{ev} - m^2G_e) - m^2v^4(G_e + v\xi G_{ee} + vG_{ev})^2. \end{aligned}$$

Regrouping terms so that

$$\begin{aligned} -\rho^4 A &= G_e(G_{vv} + G_{ee}\xi^2 + 2\xi G_{ev} - m^2G_e + m^2G_e + 2vm^2(\xi G_{ee} + G_{ev})) \\ &\quad - m^2v^2G_{ee}(G_{vv} + G_{ee}\xi^2 + 2\xi G_{ev} - m^2G_e) \\ &\quad + v^2m^2(\xi^2G_{ee}^2 + G_{ev}^2 + 2\xi G_{ee}G_{ev}) \\ &= G_e(E^2G_{ee} + 2EG_{ev} + G_{vv}) + v^2m^2(G_{ev}^2 - G_{ee}G_{vv}) \\ &= -G_e(E\sqrt{|G_{ee}|} + \sqrt{|G_{vv}|})^2 - v^2m^2(G_{ee}G_{vv} - G_{ev}^2) < 0. \end{aligned}$$

Next we turn to estimate  $B$ . Note that

$$\begin{aligned} B - q_{EE}A &= 2q_{\rho E}q_{\rho m}q_{mE} - q_{mm}q_{\rho E}^2 - q_{\rho\rho}q_{mE}^2 \\ &= -2m^2v^6(G_{ev} + \xi G_{ee}) \cdot G_{ee} \cdot (G_e + \xi v G_{ee} + v G_{ev}) \\ &\quad - v^6(\rho G_e - vm^2G_{ee}) \cdot (G_{ev} + \xi G_{ee})^2 \\ &\quad + m^2v^7G_{ee}^2(G_{vv} + \xi^2G_{ee} + 2\xi G_{ev} - m^2G_e) \\ &= -v^7G_e [2\rho m^2(G_{ev} + \xi G_{ee}) + \rho^2(G_{ev} + \xi G_{ee})^2 + m^4G_{ee}^2] \\ &\quad + m^2v^7G_{ee}(G_{ee}G_{vv} - G_{ev}^2) \\ &= -v^7G_e (m^2G_{ee} + \rho(G_{ev} + \xi G_{ee}))^2 + m^2v^7G_{ee}(G_{ee}G_{vv} - G_{ev}^2). \end{aligned}$$

This together with  $q_{EE} = -vG_{ee}$  gives

$$\begin{aligned} B &= q_{EE}A - v^7G_e (m^2G_{ee} + \rho(G_{ev} + \xi G_{ee}))^2 + m^2v^7G_{ee}(G_{ee}G_{vv} - G_{ev}^2) \\ &= v^5G_eG_{ee}(E^2G_{ee} + 2EG_{ev} + G_{vv}) - v^5G_e(EG_{ee} + G_{ev})^2 \\ &= v^5G_e(G_{vv}G_{ee} - G_{ev}^2) > 0. \end{aligned}$$

We thus conclude that  $q$  is convex.  $\square$

**4. Internal energy.** Recall that from the Boltzmann kinetic equation with  $f(t, x, \xi)$  as the probability density of gas particles at position  $x$  and moving with velocity  $\xi$ , the corresponding energy takes the form

$$E = \frac{1}{2} \int |\xi|^2 f(t, x, \xi) d\xi,$$

with

$$\rho = \int f(t, x, \xi) d\xi, \quad \rho \mathbf{u} = \int \xi f(t, x, \xi) d\xi.$$

This relation between kinetic description and the averaged fluid variables allows one to show  $E \geq \frac{1}{2}\rho|\mathbf{u}|^2$ , since

$$\frac{1}{2}\rho|\mathbf{u}|^2 = \frac{1}{2} \frac{|\int \xi f(t, x, \xi) d\xi|^2}{\int f(t, x, \xi) d\xi} \leq \frac{1}{2} \int |\xi|^2 f(t, x, \xi) d\xi = E.$$

Motivated by this fact, we introduce the following quantity

$$R = E - \frac{|\mathbf{m}|^2}{2\rho} = \rho e.$$

While one can verify that this function is concave in  $\mathbf{w}$ , the question here is of course to decide whether  $\{\mathbf{w} | R > 0, \rho > 0\}$  is invariant by the Euler dynamics.

Now we make use of the inequality (2.7)

$$\frac{vG_v}{eG_e} < 2\gamma. \quad (4.1)$$

With this inequality we are able to show that  $e > 0$ , hence  $R > 0$ , is invariant. More precisely we must demand  $\gamma$  to be bounded. But this is true for many EOS and away from vacuum. Formally from both the energy equation and the momentum equation we deduce

$$\partial_t(\rho e) + \nabla_x \cdot (\rho e \mathbf{u}) + P \nabla_x \cdot \mathbf{u} = 0.$$

For smooth solutions we have

$$\partial_t e + \mathbf{u} \cdot \nabla_x e = \left( -\nabla_x \cdot \mathbf{u} \frac{P}{\rho e} \right) e,$$

where the ratio on the right hand side can be written as

$$\frac{P}{\rho e} = \frac{vG_v}{eG_e} < 2\gamma.$$

Thus  $e > 0$  is invariant.

**5. IRP limiter.** For the Euler equations with a general EOS, we have identified a useful invariant region (2.8), i.e.,

$$\Sigma = \{\mathbf{w} \mid \rho > 0, R > 0, q < 0\},$$

where

$$R = E - \frac{|\mathbf{m}|^2}{2\rho}, \quad q = \rho(\inf s_0 - s).$$

In this section we discuss techniques of how to limit numerical solutions obtained from a high order finite-volume type schemes back to  $\Sigma$ , whenever they are out of  $\Sigma$  at some solution points.

We first recall the general explicit limiter introduced in [10]. Assume the multi-dimensional system of conservation laws admits an invariant region  $\Sigma$ , characterized by

$$\Sigma = \{\mathbf{w} \mid U(\mathbf{w}) \leq 0\},$$

where  $U$  is convex. Denote the interior of  $\Sigma$  by  $\Sigma_0$ . A key fact we have used is that for any bounded domain  $K$ , the averaging defined by

$$\bar{\mathbf{w}} = \frac{1}{|K|} \int_K \mathbf{w}(x) dx$$

is a contraction operator.

**Lemma 5.1** ([10]). *Let  $\mathbf{w}(x)$  be non-trivial piecewise continuous vector functions. If  $\mathbf{w}(x) \in \Sigma$  for all  $x \in K \subset \mathbb{R}^d$  and  $U$  is strictly convex, then  $\bar{\mathbf{w}} \in \Sigma_0$  for any bounded domain  $K$ .*

This lemma sets the foundation for using the domain average as a reference to limit the obtained solution polynomials, through a linear convex combination as in [21, 23]. If we consider a system, the question of particular interest is whether the limiting approximation is still high order accurate.

Let  $\mathbf{w}_h(x)$  be a sequence of vector polynomials over  $K$ , which is a high order accurate approximation to the function  $\mathbf{w}(x) \in \Sigma$ . Assume  $\bar{\mathbf{w}}_h \in \Sigma_0$ , but  $\mathbf{w}_h(x)$  is not entirely located in  $\Sigma$ . We construct

$$\tilde{\mathbf{w}}_h(x) = \theta \mathbf{w}_h(x) + (1 - \theta) \bar{\mathbf{w}}_h,$$

where  $\theta \in (0, 1]$  is defined by  $\theta = \min\{1, \theta_1\}$ , where

$$\theta_1 = \frac{U(\bar{\mathbf{w}}_h)}{U(\bar{\mathbf{w}}_h) - U_h^{\max}}, \quad U_h^{\max} = \max_{x \in K} U(\mathbf{w}_h(x)) > 0.$$

If  $\Sigma = \bigcap_{i=1}^M \{\mathbf{w} \mid U_i(\mathbf{w}) \leq 0\}$ , then the limiter parameter needs to be modified as

$$\theta = \min\{1, \theta_1, \dots, \theta_M\}.$$

This reconstruction has been shown to satisfy three desired properties.

**Theorem 5.2.** [10] *The reconstructed polynomial  $\tilde{\mathbf{w}}_h(x)$  satisfies the following three properties:*

- (i) *the average is preserved, i.e.,  $\bar{\mathbf{w}}_h = \bar{\tilde{\mathbf{w}}}_h$ ;*
- (ii)  *$\tilde{\mathbf{w}}_h(x)$  lies entirely within invariant region  $\Sigma, \forall x \in K$ ;*
- (iii) *order of accuracy is maintained, i.e., if  $\|\mathbf{w}_h - \mathbf{w}\|_\infty \leq 1$ , then*

$$\|\tilde{\mathbf{w}}_h - \mathbf{w}\|_\infty \leq \frac{C}{|U(\bar{\mathbf{w}}_h)|} \|\mathbf{w}_h - \mathbf{w}\|_\infty,$$

where  $C > 0$  depends on  $\mathbf{w}$  and  $\Sigma$ .

Let  $\mathbf{w}_h^n$  be the numerical solution at the  $n$ -th time step generated from a high order finite-volume-type scheme of an abstract form

$$\mathbf{w}_h^{n+1} = \mathcal{L}(\mathbf{w}_h^n), \quad \mathbf{w}_h^n = \mathbf{w}_h^n(x) \in V_h.$$

Provided that the scheme has the following property: there exists  $\lambda_0$  and a test set  $S$  such that if

$$\lambda := \frac{\Delta t}{\Delta x} \leq \lambda_0 \quad \text{and} \quad \mathbf{w}_h^n(x) \in \Sigma \text{ for } x \in S,$$

then

$$\bar{\mathbf{w}}_h^{n+1} \in \Sigma_0;$$

the limiter can then be applied with  $K$  replaced by  $S_K : S \cap K$ , i.e.,

$$U_h^{\max} = \max_{x \in S_K} U(\mathbf{w}_h(x)),$$

through the following algorithm:

**Step 1. Initialization:** take the piecewise  $L^2$  projection of  $\mathbf{w}_0$  onto  $V_h$ , such that

$$\langle \mathbf{w}_h^0 - \mathbf{w}_0, \phi \rangle = 0, \quad \forall \phi \in V_h.$$

**Step 2. Limiting:** Impose the modified limiter on  $\mathbf{w}_h^n$  for  $n = 0, 1, \dots$  to obtain  $\tilde{\mathbf{w}}_h^n$ .

**Step 3. Update** by the scheme:

$$\mathbf{w}_h^{n+1} = \mathcal{L}(\tilde{\mathbf{w}}_h^n).$$

Return to Step 2.

For multi-dimensional Euler equations (1.1), we use the invariant region defined in (2.8) to formulate the IRP limiter which serves as a basis for constructing high order IRP numerical schemes for (1.1).

A limiter as such was first reported in [9] for one-dimensional Euler equations, and in [11] for the isentropic gas dynamics. The limiter in [9] is explicit and simultaneously preserves the positivity of density and pressure and also a minimum principle for the specific entropy [20].

**6. An example EOS.** It was shown previously that the polytropic gas EOS admits an invariant region and it easily fulfills the conditions formulated above. Here we want to present another EOS which fulfills the requirements. To this end we will consider the so-called Tait EOS, cf. [1]. Let the specific internal energy be given as follows

$$e = F(s, v) = A(v - v_r) + B\Phi(v) + \frac{1}{2C} [(s - s_r) - D(v - v_r)]^2 + \theta_r [(s - s_r) - D(v - v_r) + C\theta_r] + e_r, \quad (6.1)$$

$$\Phi(v) = \begin{cases} \ln \frac{v_r}{v}, & \nu = 1 \\ \frac{1}{1 - \nu} \left( \frac{1}{v_r^{\nu-1}} - \frac{1}{v^{\nu-1}} \right), & \nu > 1 \end{cases}. \quad (6.2)$$

The constants  $A, B, C, D$  will be specified later. The quantities  $e_r, v_r, s_r$  and  $\theta_r$  denote reference constants of the quantities which may be chosen at a given reference state in the phase space of the material at hand. The exponent  $\nu \geq 1$  can also be chosen according to the specific material under consideration. Now we calculate the pressure and the temperature according to the formulas given above, i.e.,

$$P = -\frac{\partial F}{\partial v}(s, v) = -\left( A + B\Phi'(v) - \frac{D}{C} [(s - s_r) - D(v - v_r)] - D\theta_r \right), \quad (6.3)$$

$$\theta = \frac{\partial F}{\partial s}(s, v) = \frac{1}{C} [(s - s_r) - D(v - v_r)] + \theta_r. \quad (6.4)$$

Using equation (6.4) for the temperature we can express the pressure as a function of the volume and the temperature, i.e.,

$$P = f(v, \theta) = D(\theta - \theta_r) - A - B\Phi'(v). \quad (6.5)$$

Such a relation is called thermal EOS. We now choose

$$A := K_r - p_r, \quad B := K_r v_r^\nu, \quad \text{and} \quad \bar{p}(\theta) := D(\theta - \theta_r) + p_r.$$

The constant quantity  $K_r > 0$  denotes the modulus of compression at a given reference state. The constants  $K_r, v_r$  may be chosen as the saturation values of the material under consideration for a given reference temperature  $\theta_r$ , cf. [19]. Accordingly  $\bar{p}(\theta)$  can be obtained as the linearization of the saturation curve near a given temperature. A similar approach is used in [6]. Usually the saturation curve is given as the saturation pressure being a function of the temperature

$$p_0 = f(\theta),$$

e.g. [13, 19], and thus we would obtain

$$\bar{p}(\theta) = f'(\theta_r)(\theta - \theta_r) + p_r.$$

With these assumptions we can rewrite (6.3) into a more well known form

$$P = \bar{p}(\theta) + K_r \left[ \left( \frac{v_r}{v} \right)^\nu - 1 \right]. \quad (6.6)$$

The requirement  $P > 0$  places a restriction on the allowed region of the state space. The non-linear Tait EOS is for example used in [8, 17] and may serve as an EOS

for liquids. In order to verify the convexity of the internal energy we calculate the second order derivatives, i.e.,

$$\frac{\partial^2 F}{\partial v^2}(s, v) = B\Phi''(V) + \frac{D^2}{C}, \quad \frac{\partial^2 F}{\partial s^2}(s, v) = \frac{1}{C}, \quad \frac{\partial^2 F}{\partial s \partial v}(s, v) = -\frac{D}{C}.$$

It is easily verified that  $F$  is strictly convex if  $C > 0$ . Indeed  $C$  is related to the specific isochoric heat capacity

$$c_v = \frac{\partial F}{\partial \theta}(s(v, \theta), v) = C\theta$$

and can be chosen to be  $C = c_{v,r}/\theta_r$ . We can also calculate the speed of sound which is given by

$$a = \sqrt{v^2 \frac{\partial^2 F}{\partial v^2}} = \sqrt{v^2 \left( B\Phi''(V) + \frac{D^2}{C} \right)}.$$

For the dimensionless quantities we yield

$$\begin{aligned} g &= \frac{Pv}{\theta^2} \frac{\partial^2 F}{\partial s^2} = \frac{Pv}{\theta c_v}, \quad \Gamma = -\frac{v}{\theta} \frac{\partial^2 F}{\partial s \partial v} = \frac{Dv}{c_v}, \\ \gamma &= \frac{v}{P} \frac{\partial^2 F}{\partial v^2} = \frac{v}{P} \left( B\Phi''(V) + \frac{D^2}{C} \right) = \nu \frac{K_r v_r^\nu}{P v^\nu} + \frac{\Gamma^2}{g}, \\ \mathcal{G} &= -\frac{1}{2} v \frac{\frac{\partial^3 F}{\partial v^3}}{\frac{\partial^2 F}{\partial v^2}} = -\frac{1}{2} v^2 \frac{B\Phi'''(V)}{\gamma P} = \frac{\nu + 1}{2} \frac{\nu K_r v_r^\nu}{\gamma P v^\nu}. \end{aligned}$$

Clearly we have  $\mathcal{G} > 0$ . Thus the isentropes are convex in the  $P - v$ -plane. Hence the presented EOS meets all requirements needed above.

**Acknowledgments.** Parts of this research were done during a research stay of Hailiang Liu in Magdeburg. The authors want to thank Gerald Warnecke for this possibility and his hospitality. Moreover the authors want to express their gratitude to Marshall Slemrod for helpful discussions.

## REFERENCES

- [1] J. H. Dymond and R. Malhotra, The Tait equation: 100 years on, *Int. J. Thermophys.*, **9** (1988), 941–951.
- [2] L. C. Evans, *Entropy and Partial Differential Equations*, Lecture notes, 2010.
- [3] H. Frid, Invariant regions under Lax-Friedrichs scheme for multidimensional systems of conservation laws, *Discrete Contin. Dyn. Syst.*, **1** (1995), 585–593.
- [4] H. Frid, Maps of Convex Sets and Invariant Regions for Finite-Difference Systems of Conservation Laws, *Arch. Ration. Mech. Anal.*, **160** (2001), 245–269.
- [5] J. L. Guermond and B. Popov, Invariant domains and first-order continuous finite element approximation for hyperbolic systems, *SIAM J. Numer. Anal.*, **54** (2016), 2466–2489.
- [6] M. Hantke and F. Thein, Why condensation by compression in pure water vapor cannot occur in an approach based on Euler equations, *Quart. Appl. Math.*, **73** (2015), 575–591.
- [7] D. Hoff, Invariant regions for systems of conservation laws, *Trans. Amer. Math. Soc.*, **289** (1985), 591–610.
- [8] M. J. Iivings, D. M. Causon and E. F. Toro, On Riemann solvers for compressible liquids, *Int. J. Numer. Methods Fluids*, **28** (1998), 395–418.
- [9] Y. Jiang and H. Liu, *An Invariant-region-preserving (IRP) Limiter to DG Methods for Compressible Euler Equations*, Springer, Cham, 2018.
- [10] Y. Jiang and H. Liu, Invariant-region-preserving DG methods for multi-dimensional hyperbolic conservation law systems, with an application to compressible Euler equations, *J. Comput. Phys.*, **373** (2018), 385–409.

- [11] Y. Jiang and H. Liu, [An invariant region preserving limiter for DG schemes to isentropic Euler equations](#), *Numer. method. PDEs*, **35** (2019), 5–33.
- [12] B. Khobalatte and B. Perthame, [Maximum principle on the entropy and second-order kinetic schemes](#), *Math. Comput.*, **62** (1994), 119–131.
- [13] L. D. Landau and E. M. Lifšic, *Lehrbuch der Theoretischen Physik, Bd. V Statistische Physik*, Akad.-Verl., Berlin, 1987.
- [14] R. Menikoff and B. J. Plohr, [The Riemann problem for fluid flow of real materials](#), *Rev. Mod. Phys.*, **61** (1989), 75–130.
- [15] S. Müller and A. Voss, [The Riemann Problem for the Euler Equations with Nonconvex and Nonsmooth Equation of State: Construction of Wave Curves](#), *SIAM J. Sci. Comput.*, **28** (2006), 651–681.
- [16] B. Perthame and C. W. Shu, [On positivity preserving finite volume schemes for Euler equations](#), *Numer. Math.*, **73** (1996), 119–130.
- [17] R. Saurel, P. Cocchi and P. Butler, [Numerical Study of Cavitation in the Wake of a Hypervelocity Underwater Projectile](#), *J. Propul. Power*, **15** (1999), 513–522.
- [18] J. Smoller, *Shock Waves and Reaction-Diffusion Equations*, Grundlehren der Mathematischen Wissenschaften, Springer-Verlag, New York-Berlin, 1983.
- [19] F. Thein, *Results for Two Phase Flows with Phase Transition*, Dissertation, Otto-von-Guericke University of Magdeburg, 2018.
- [20] E. Tadmor, [A minimum entropy principle in the gas dynamics equations](#), *Appl. Numer. Math.*, **2** (1986), 211–219.
- [21] X. Zhang and C. W. Shu, [On maximum-principle-satisfying high order schemes for scalar conservation laws](#), *J. Comput. Phys.*, **229** (2010), 3091–3120.
- [22] X. Zhang and C. W. Shu, [On positivity preserving high order discontinuous Galerkin schemes for compressible Euler equations on rectangular meshes](#), *J. Comput. Phys.*, **229** (2010), 8918–8934.
- [23] X. Zhang and C. W. Shu, [A minimum entropy principle of high order schemes for gas dynamics equations](#), *Numer. Math.*, **121** (2012), 545–563.
- [24] X. Zhang, Y. Xia and C. W. Shu, [Maximum-principle-satisfying and positivity-preserving high order discontinuous Galerkin schemes for conservation laws on triangular meshes](#), *J. Sci. Comput.*, **50** (2012), 29–62.

Received February 2021; revised April 2021.

*E-mail address:* [hliu@iastate.edu](mailto:hliu@iastate.edu)

*E-mail address:* [ferdinand.thein@ovgu.de](mailto:ferdinand.thein@ovgu.de)