Mathematical Theory of Nonlinear Single-Phase Poroelasticity

Cornelis Johannes van Duijn, Andro Mikelic

Mathematical Theory of Nonlinear Single-Phase Poroelasticity[⋆]

C.J. van Duijn[a], Andro Mikelić[b]

[a]*Darcy Center, Eindhoven University of Technology, Department of Mechanical Engineering and Utrecht University, Department of Earth Sciences, the Netherlands*
[b]*Univ Lyon, Université Claude Bernard Lyon 1, CNRS UMR 5208, Institut Camille Jordan, 43 blvd. du 11 novembre 1918, F-69622 Villeurbanne cedex, France*

ARTICLE INFO

ABSTRACT

In this paper we study the equations of nonlinear poroelasticity derived from mixture theory. They describe the quasi-static mechanical behaviour of a fluid saturated porous medium. The nonlinearity arises from the compressibility of the fluid and from the dependence of porosity and permeability on the divergence of the displacement. We point out analytical difficulties with the model. In our approach we discretize the quasi-static formulation in time and first consider the corresponding incremental problem. For this, we prove existence of a solution using Brézis' theory of pseudo-monotone operators. Generalizing Biot's free energy to the nonlinear setting we construct a Lyapunov functional for the model and prove stability. It allows constructing bounds that are uniform with respect to the time step. In the case when dissipative interface effects between the fluid and the solid are taken into account, we consider the continuous time case in the limit when the time step tends to zero. This yields existence of a weak free energy solution.

1. Introduction

The elastic quasi-static deformation of a fluid saturated porous medium received much attention in the civil engineering literature because of its relevance to many problems of practical interest. In the framework of consolidation in soil mechanics, these problems relate to the physical loading of soil layers or the effect of soil subsidence due to groundwater withdrawal for drinking water supply or industrial and agricultural purposes. Examples and underlying theories are given in the well-known works of Coussy [12], Lewis and Schrefler [18] and Verruijt [38]. They build on the classical theory of Terzaghi [36] and the pioneering approach of Biot [4],[37].

Recently, other examples of elastic deformation of porous media arise in the context of industrial and biomedical applications, such as paper printing [6], bone regeneration [13], [11], blood flow [29], [8] and car filters [20], [23].

In its simplest form, assuming both the fluid and the porous material (grains) to be incompressible and assuming the porous medium to be homogeneous and linear elastic with small strains , the mathematical formulation reads (see Bear [1], Verruijt [38] or van Duijn et al [14]) :

$$\text{div }\partial_t \mathbf{u} + \text{ div } \left(\frac{\mathbb{K}}{\eta_f}(\rho_f \mathbf{g} - \nabla p) \right) = q \tag{1}$$

and

$$- \text{div } \sigma = \mathbf{F}, \tag{2}$$

where

$$\sigma = \mathcal{G}e(\mathbf{u}) - \alpha p \mathbb{I}, \tag{3}$$

with

$$\mathcal{G}E = 2\mu E + \lambda \text{Tr}(E)\mathbb{I}, \quad \text{for symmetric matrices } E. \tag{4}$$

In these equations, $\mathbf{u}$ [m] denotes skeleton displacement, $\mathbb{K}$ [m^2] intrinsic permeability (a symmetric positive definite rank-2 tensor), η_f [Pa s] fluid viscosity, p [Pa] fluid pressure and q [1/s] sources/sinks. Further, σ [Pa] is the total stress, $\mathbf{F}$ a given body force (generally linked to gravitational effects), $\mathcal{G}$ the symmetric, positive-definite, rank-4 Gassmann tensor, $e(\mathbf{u})$ the linearized strain tensor and $\alpha \in (0, 1]$ Biot's effective stress parameter. Finally, μ [Pa] and λ [Pa] are Lamé's parameters. Using for $\mathcal{G}$ the specific form (4), i.e. Hooke's law, assumes that the skeleton is mechanically isotropic.

The linear quasi-static Biot system, as well as its dynamical analogue, was also derived by means of a multiscale approach, where the starting point is the linear fluid-structure interaction at the pore level. We refer to the monographs [32] and [21] for derivations using two-scale expansions and to [24] for a rigorous mathematical derivation by means of homogenization. The derivations using multiscale analysis confirm Biot's models in the linear setting. Hence from different points of view system (1)-(4) is well accepted.

In the engineering literature one writes $\alpha = 1 - K/K_g$, where K is the drained bulk modulus of the porous skeleton and K_g the bulk modulus of the grains. Since it is assumed that $K_g = +\infty$, we will set $\alpha = 1$ in (3).

From a mathematical perspective, equations (1)-(4) were studied by Ženišek [40], who was one of the first to demonstrate existence and uniqueness, and by Showalter [35] in the dynamic case. More recent studies include [28] and [20]. Later, Cao et al [9] considered a nonlinear extension of (1), by replacing the permeability tensor $\mathbb{K}$ by the product $\mathbb{K}k(\mathrm{div}\ \mathbf{u})$. The function $k(\cdot)$ is a relative permeability depending on the volumetric strain div $\mathbf{u}$. From (1) we notice that the overall mixture of two incompressible phases is not incompressible itself.

Though system (1)-(4) is linear, its mathematical complexity lies in the fact that it is of quasi-static nature. In particular (2)-(4) allow to control the size of the volumetric strain only through the size of the data. Some authors circumvent this by introducing a time dependence in (2)-(4) as well. For instance Bociu et al [5] replace $\mathbf{u}$ in (3) by $\mathbf{u} + \delta\partial_t\mathbf{u}$, $(\delta > 0)$, i.e. by introducing a viscoelastic effect. A different regularization was proposed by Murad and Cushman [27] who replaced (3) and (4) by

$$\sigma = 2\mu e(\mathbf{u}) + (\lambda\mathrm{div}\ \mathbf{u} + \lambda^*\mathrm{div}\ \partial_t\mathbf{u} - \alpha p)\mathbb{I}, \tag{5}$$

with $\lambda^* > 0$. This form arises in the non-equilibrium theory, where the fluid pressure and the solid pressure differ by λ^* div $\partial_t\mathbf{u}$.

In this paper we propose to study the quasi-static formulation in which we replace equation (1) by the nonlinear fluid phase mass balance based on the mixture theory of Bedford and Drumheller [2], [3], see e.g. Rutquist et al [31] and Lewis and Schrefler [18]:

$$n\partial_t\rho + \rho\mathrm{div}\ \partial_t\mathbf{u} + \mathrm{div}\ \mathbf{j} = Q, \tag{6}$$

where $\mathbf{j}$ denotes the Darcy mass flux

$$\mathbf{j} = \frac{\mathbb{K}k(n)\rho}{\eta_f}(\rho\mathbf{g} - \nabla p). \tag{7}$$

Here n denotes porosity, $\rho = \rho_f$ [kg/m^3] fluid density, k relative permeability and Q [kg/m^3 s] sources/sinks.

In equations (6)-(7), the porosity n is a given function of the volumetric strain: i.e.

$$n = n(\mathrm{div}\ \mathbf{u}). \tag{8}$$

An explicit expression for (8) is derived from the Lagrangian solid mass balance equation. This is shown in Section 2. Through (8), the relative permeability depends on div $\mathbf{u}$.

Since n is the volume fraction of voids in the porous medium, it should satisfy the natural bounds

$$0 < n < 1. \tag{9}$$

However, in Section 2 we show by means of a counter example that the porosity can attain negative - and thus physically unrealistic - values. Therefore, the bounds in (9) are a major concern in the mathematical model.

To close system (2)-(4), (6)-(7) we introduce a constitutive relation for the fluid density in terms of the pressure. Assuming weak compressibility we write

$$\rho = \rho(p) = \rho_0(1 + \beta(p - p_0)). \tag{10}$$

Further we propose an explicit expression for the relative permeability in terms of the porosity

$$k = k(n). \tag{11}$$

In (10) ρ_0 and p_0 are reference values for, respectively, density and pressure and β [Pa^{-1}] is the fluid compressibility coefficient. The relative permeability in (11) satisfies

$$\begin{cases} k \in C^1[0,1], \\ k(0) > 0 \quad \text{and} \quad k' > 0 \text{ in } ([0,1). \end{cases} \tag{12}$$

A well-known example is the Kozeny-Carman formula, see for instance [1],

$$k(n) = k_0 \frac{n^3}{(1-n)^2} \quad (k_0 > 0), \tag{13}$$

in a realistic porosity interval, bounded away from $n = 0$ and $n = 1$. Thus taking k such that (12) holds and

$$k(n) = \begin{cases} \geq \dfrac{k_0}{2} \dfrac{n_*^3}{(1-n_*)^2}, & \text{for } n \leq n_*, \\[2ex] k_0 \dfrac{n^3}{(1-n)^2}, & \text{for } 0 < n_* < n < n^* < 1, \\[2ex] \leq 2k_0 \dfrac{(n^*)^3}{(1-n^*)^2}, & \text{for } n \geq n^*, \end{cases} \tag{14}$$

for appropriately chosen $0 < n_* < n^* < 1$, gives a relative permeability satisfying (13) in the interval $[n_*, n^*]$.

We notice that equation (6), coupled with (2)) and (4), is nonlinear due to the relation $k = k(n)$ and the products involving time derivatives. Assuming constant fluid phase density in the poroelastic mixture is therefore an important simplification. This is studied in [9] and [5].

In studying system (2)-(4), (6)-(11) a crucial role is played by its free energy. The idea is to generalize Biot's free energy [4], which is quadratic in strain and fluid density, to the nonlinear poroelastic setting. This free energy serves as a Lyapunov functional. This approach is linked to general entropy methods for PDEs. For a detailed survey, covering various fields of applications we refer to [15] and to the recent book by Jüngel [16]. An interesting application of the entropy method is discussed in [22], [7], [10] and [25], where the authors consider dynamic capillary pressure effects in two-phase porous media flow.

This paper is organized as follows. In Section 2 we present details of the model formulation. Starting point is the mass balance for the fluid and the solid phase. The latter implies an explicit expression for (8). Introducing a lower bound for the porosity we modify the fluid mass balance so that a Lyapunov functional can be constructed for the modified system. This modification is such that the fluid equation reduces to its original form in the physical range of the fluid density ρ and solid volumetric strain $\mathcal{E}$. Section 2 is concluded by a weak formulation of the modified system.

In Section 3 we consider, for the relaxation parameter $\lambda^* \geq 0$, the incremental version of the modified system. Using Brézis' theory of pseudo monotone operators, existence is demonstrated. Applying the Lyapunov functional yields global (in time) estimates. Next, in Section 4, we use these estimates to solve the time continuous problem when $\lambda^* > 0$. In both Sections 3 and 4 we borrow ideas from Roubiček [30]. Finally, in Section 5 we present a discussion and conclusions.

2. Problem formulation

In a number of steps we construct in this section the equations that serve as starting point for the analysis. The general setting of the problem is as follows:

Let $\Omega \subset \mathbb{R}^m$ (m=2,3) denote a bounded domain, occupied by a linear elastic skeleton. The skeleton material (grains) is assumed incompressible: i.e. the bulk modulus of the grains is infinitely large. The voids in the porous structure are completely filled with a slightly compressible fluid, in the sense that the fluid pressure p and density ρ are related by (10).

2.1. Balance equations

For given $\xi \in \Omega$, let $\mathbf{x}(\xi, t)$ denote the location of a solid particle at time $t > 0$, that started at $\mathbf{x}(\xi, 0) = \xi$,. Then the skeleton velocity $\mathbf{v}_s$ is given by $\mathbf{v}_s = \partial_t \mathbf{x}|_\xi$.

Restricting themselves to small displacements $\mathbf{u}$ (within the elastic regime), Rutquist et al [31] and Lewis and Schrefler [18] argue that in the mass balance equation for the fluid and solid, the material derivative $\frac{D}{Dt} = \partial_t + \mathbf{v}_s \cdot \nabla$ can be replaced by the partial derivative ∂_t. This is made explicit by a scaling argument in van Duijn et al [14]. The resulting Lagrangian form of the mass balances reads:

$$n\partial_t \rho + \rho \operatorname{div} \mathbf{v}_s + \operatorname{div} \mathbf{j} = Q \quad \text{(fluid phase)} \tag{15}$$

and

$$\partial_t(1 - n) + (1 - n)\operatorname{div} \mathbf{v}_s = 0 \quad \text{(solid phase)}, \tag{16}$$

where $\mathbf{j}$ is mass flux (7).

Within the same approximation one may write

$$\operatorname{div} \mathbf{v}_s = \partial_t \operatorname{div} \mathbf{u}.$$

Using this in (15) and (16) gives

$$n\partial_t \rho + \rho \partial_t \operatorname{div} \mathbf{u} + \operatorname{div} \mathbf{j} = Q \tag{17}$$

and

$$\partial_t(1 - n) + (1 - n)\operatorname{div} \partial_t \mathbf{u} = 0. \tag{18}$$

Integrating (18) in time from $t = 0$, say, to $t > 0$, we have

$$1 - n = (1 - n_0)e^{-\operatorname{div}(\mathbf{u} - \mathbf{U}_0)} \quad \text{for } t > 0.. \tag{19}$$

Here $\mathbf{U}_0$ is the initial displacement and n_0 the initial porosity. With $n_0 \in (0, 1)$ in Ω, expression (19) ensures

$$n < 1 \quad \text{in} \quad \Omega \quad \text{for all } t > 0. \tag{20}$$

To avoid technical complications we restrict ourselves in this paper to the case when the initial porosity n_0 is constant in Ω.

For small displacements $\mathbf{u} - \mathbf{U}_0$, expression (19) is approximated by

$$n = n_0 + (1 - n_0)\operatorname{div}(\mathbf{u} - \mathbf{U}_0). \tag{21}$$

Remark 1. *Frequently, the linear form (21) is used for values of div $\mathbf{u}$ in a neighborhood of div $\mathbf{U}_0$: i.e. in practical circumstances (21) is applied when $\mathcal{E}_* < div(\mathbf{u} - \mathbf{U}_0) < \mathcal{E}^*$, where $\mathcal{E}_* < 0 < \mathcal{E}^*$ are appropriately chosen.*

Throughout the paper we redefine

$$\mathbf{u} := \mathbf{u} - \mathbf{U}_0, \tag{22}$$

where $\mathbf{U}_0 \in H_0^1(\Omega)^m \cap H^2(\Omega)^m$ is the initial displacement. Redefining accordingly

$$\mathbf{F} := \mathbf{F} + \operatorname{div}(\mathcal{G}e(\mathbf{U}_0)), \tag{23}$$

we obtain for the fluid pressure p and the skeleton displacement $\mathbf{u}$ the system:

$$n\partial_t \rho + \rho \operatorname{div} \partial_t \mathbf{u} + \operatorname{div}\left(\frac{\mathbb{K}k(n)\rho}{\eta_f}(\rho\mathbf{g} - \nabla p)\right) = Q, \tag{24}$$

$$- \operatorname{div}(\mathcal{G}e(\mathbf{u}) - p\mathbb{I}) = \mathbf{F}, \tag{25}$$

where

$$\rho = \rho(p) = \rho_0(1 + \beta(p - p_0)), \tag{26}$$

$$n = n(\operatorname{div} \mathbf{u}) = 1 - (1 - n_0)e^{-\operatorname{div} \mathbf{u}} \tag{27}$$

$$\approx n_0 + (1 - n_0)\operatorname{div} \mathbf{u} \quad \text{(small strains)}. \tag{28}$$
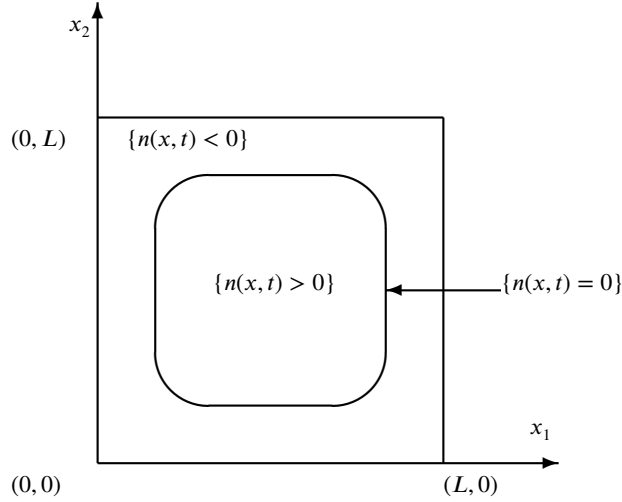
Figure 1: Sketch of level set of porosity n at some $t > 0$. The region where $n > 0$ shrinks with increasing time and disappears after a finite time $T_p > 0$.

Remark 2. *Concerning the initial displacement $\mathbf{U}_0$ we note that only div $\mathbf{U}_0$, the initial volumetric strain, is used. However, when discussing the free energy, one needs in addition that $\mathbf{U}_0$ is such that the corresponding elastic energy is finite. For simplicity we suppose $\mathbf{U}_0 \in H^2(\Omega)^m$.*

In the next sections we will develop the mathematical theory for system (24)-(28).

The issue of negative porosity in (27) (or, for that matter, a porosity exceeding one in approximation (28)), is discussed next.

2.2. Negative porosity

We consider a simplified version of the linear problem (1)-(4) and show that div $\mathbf{u}$ can attain values for which the porosity from (27)-(28) becomes negative.

For simplicity we give the construction in $\mathbb{R}^2$.

Let $\Omega = (0, L)^2$ for some $L > 0$. We suppose, as in the rest of this paper, that div $\mathbf{u}|_{t=0} = 0$. Further we set $\mathbf{F} = 0$ in (25). Using (4) in (25) gives

$$- \operatorname{div}\left(2\mu e(\mathbf{u}) + (\lambda \operatorname{div} \mathbf{u} - \alpha p)\mathbb{I}\right) = 0 \quad \text{in } \Omega., \tag{29}$$

Proceeding as in Verruijt [38], when he discusses the Mandel problem, we take the divergence of (29) to obtain

$$\Delta\left((2\mu + \lambda)\operatorname{div} \mathbf{u} - p\right) = 0 \quad \text{in } \Omega. \tag{30}$$

Hence the function

$$H = (2\mu + \lambda)\operatorname{div} \mathbf{u} - p$$

is harmonic in Ω.

The idea is to prescribe boundary conditions for equations (24) and (25) so that $H|_{\partial\Omega}$ is given. For instance, if we set along the four edges, see Figure 1,

$$\begin{cases} \{x_1 = 0\} : u_2 = 0, \ \sigma_{11} = \Sigma^{1,0} \text{ and } p = 0; \\ \{x_1 = L\} : u_2 = 0, \ \sigma_{11} = \Sigma^{1,L} \text{ and } p = 0; \\ \{x_2 = 0\} : u_1 = 0, \ \sigma_{22} = \Sigma^{2,0} \text{ and } p = 0; \\ \{x_2 = L\} : u_1 = 0, \ \sigma_{22} = \Sigma^{2,L} \text{ and } p = 0, \end{cases} \tag{31}$$

and use

$$\sigma_{11} = 2\mu \frac{\partial u_1}{\partial x_1} + \lambda \text{div } \mathbf{u} - p.$$

We have

$$\Sigma^{1,0} = (2\mu + \lambda)\frac{\partial u_1}{\partial x_1} \quad \text{at} \quad \{x_1 = 0\},$$

implying

$$H = \Sigma^{1,0} \quad \text{at} \quad \{x_1 = 0\}.$$

Repeating this along the other edges gives

$$H|_{\partial\Omega} = \Sigma^b,$$

where Σ^b denotes the given value of σ along the edges.

Then we have

Proposition 3. *Let $\mathcal{E} = \text{div } \mathbf{u}$ denote the volumetric stress and let $n(\mathcal{E})$ be given by (27). Suppose there exists a constant $\Sigma > 0$ such that $\Sigma^b \leq -\Sigma$. Then for Σ sufficiently large, there exists a $T_p = T_p(\Sigma) > 0$ such that*

$$n(\mathcal{E}(x,t)) < 0 \quad \text{for } t > T_p \text{ and } x \in \overline{\Omega}. \tag{32}$$

Proof. Note that the sign of Σ^b implies compression of the medium. Because Σ is constant, we have

$$(2\mu + \lambda)\text{div } \mathbf{u} - p = -\Sigma \quad \text{in} \quad \Omega \quad \text{for} \quad t > 0. \tag{33}$$

Restricting ourselves to the linear case (1) in a homogeneous and isotropic porous medium in which sources/sinks and gravity are absent, we have

$$\partial_t \text{div } \mathbf{u} - \frac{K}{\eta_f}\Delta p = 0 \quad \text{in} \quad \Omega, \ t > 0. \tag{34}$$

Combining (33)-(34) gives for $\mathcal{E}$ the problem

$$\begin{cases} \partial_t \mathcal{E} = \dfrac{(2\mu + \lambda)K}{\eta_f}\Delta\mathcal{E} \quad \text{in} \quad \Omega, \ t > 0; \\ \mathcal{E}|_{\partial\Omega} \leq -\dfrac{\Sigma}{2\mu + \lambda} \quad \text{for } t > 0; \\ \mathcal{E}|_{\{t=0\}} = 0 \quad \text{in} \quad \Omega. \end{cases} \tag{35}$$

By the strong maximum principle, $\mathcal{E} < \overline{\mathcal{E}}$ in Ω and for $t > 0$, where $\overline{\mathcal{E}}$ is the solution of problem (35 with $\overline{\mathcal{E}} = -\Sigma/(2\mu + \lambda)$ on $\partial\Omega$. Writing $\overline{\mathcal{E}}$ as a Fourier series, one observes that

$$\overline{\mathcal{E}}(x,t) \to -\frac{\Sigma}{2\mu + \lambda} \quad \text{as } t \to +\infty,$$

uniformly in $x \in \overline{\Omega}$.

Thus if

$$(1 - n_0)e^{\Sigma/(2\mu+\lambda)} > 1,$$

or

$$\Sigma > (2\mu + \lambda)\ln\frac{1}{1 - n_0},$$

the result is immediate. $\square$ $\hspace{10em}$ $\square$

This example shows that there is a problem with the model. A modification is needed to prevent the porosity (27, or (28), to become negative. Of course one could argue that this is outside the scope of the model or practical applications, since linear elasticity and small strains are supposed. However, since it is not clear how to ensure that indeed small displacements/strains are guaranteed, one needs to impose a porosity modification to prevent negative values.
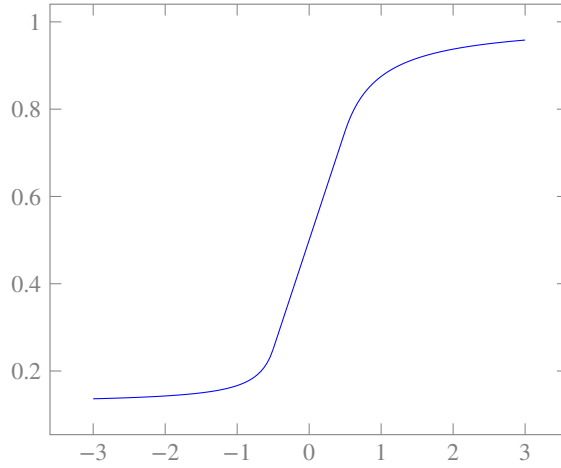
Nonlinear Poroelasticity

Figure 2: Sketch of porosity cut-off $\overline{n}(\mathcal{E})$, with $n_0 = 0.5$, $\mathcal{E}_* = -0.5$, $\mathcal{E}^* = 0.5$ and $\delta_0 = 0.125$.

2.3. Modification of balance equations

In a number of steps we modify equation (24) so that it becomes well-posed in a mathematical sense and reduces to its original form in the physical range of the unknowns.

First, to satisfy the natural bounds (9), we replace the porosity approximation (28) by a smooth increasing function $\overline{n} : \mathbb{R} \to \mathbb{R}$ such that

$$\overline{n}(\mathcal{E}) = \begin{cases} \lim_{\mathcal{E} \to +\infty} \overline{n}(\mathcal{E}) = 1, \\ n_0 + (1 - n_0)\mathcal{E}, \text{ for } \mathcal{E}_* \leq \mathcal{E} \leq \mathcal{E}^*; \\ \lim_{\mathcal{E} \to -\infty} \overline{n}(\mathcal{E}) = \delta_0 > 0. \end{cases} \tag{36}$$

Here $\mathcal{E}_*$ and $\mathcal{E}^*$ are practical values chosen such that $-n_0/(1 - n_0) < \mathcal{E}_* < 0 < \mathcal{E}^* < 1$ and $\delta_0 = n(\mathcal{E}_*)/2$, see Figure 2 for a sketch.

This construction ensures that the modified porosity $\overline{n}(\mathcal{E})$ remains in the physical range $(0, 1)$ and coincides with the linear approximation in the interval $(\mathcal{E}_*, \mathcal{E}^*)$. Realistic porosity measurements are always done away from the bounds $n = 0$ and $n = 1$, see e.g. Bear [1].

We choose to study equation (24) with the fluid density as primary unknown. Hence we need to express the pressure p in terms of ρ. Using (26) we have explicitly

$$p = p(\rho) := p_0 + \frac{\rho - \rho_0}{\beta \rho_0}. \tag{37}$$

When considering (24), one clearly has in mind that ρ takes values near the reference ρ_0. However the mathematical nature of the equations does not guarantee this behaviour. Hence a second modification is needed, now for ρ in the second and third term of the left-hand side of (24). Disregarding gravity, we replace (24) by the modified fluid mass balance equation

$$\overline{n}(\mathcal{E})\partial_t \rho + d(\rho)\partial_t \mathcal{E} - \operatorname{div}\left(k(\mathcal{E})\mathcal{D}(\rho)\mathbb{K}\nabla\rho\right) = Q, \tag{38}$$

where $\overline{n}(\mathcal{E})$ is given by (36) and $k(\mathcal{E}) = k(\overline{n}(\mathcal{E}))$. Further, $d, \mathcal{D} : \mathbb{R} \to \mathbb{R}$ are chosen such that

$$\left.\begin{array}{l} d(\rho) = \rho, \\ \mathcal{D}(\rho) = \dfrac{\rho}{\eta_f \beta \rho_0}, \end{array}\right\} \quad \text{for} \quad |\rho - \rho_0| \leq \rho_0 - \rho_*, \tag{39}$$

where $\rho_* \in (0, \rho_0)$ is a small constant. Outside this range we take for d and $\mathcal{D}$ extensions that suit the mathematical analysis. We clarify this at a later point in this section.

Remark 4. *The composite function $k(\mathcal{E}) = k(\bar{n}(\mathcal{E}))$ satisfies: $k \in C^1(\mathbb{R}) \cap L^\infty(\mathbb{R})$ and $k > k(\delta_0) > 0$, $k' > 0$ in $\mathbb{R}$.*

The balance of forces (25) is modified by adding the regularizing term $\lambda^* \partial_t \mathcal{E}$, as in the expression (5). This gives

$$-\text{div}\left(\mathcal{G}e(\mathbf{u}) + (\lambda^* \partial_t \mathcal{E} - \alpha p)\mathbb{I}\right) = \mathbf{F}, \tag{40}$$

where we have $\lambda^* \geq 0$.

We consider system (38), (40) in the set

$$Q_T = \{(x,t) : x \in \Omega, 0 < t < T\},$$

where $T > 0$ is arbitrarily chosen. To avoid technical complications we take $\partial\Omega \in C^1$ throughout the rest of this paper.

As initial conditions we have

$$\mathcal{E}|_{t=0} = 0 \quad \text{and} \quad \rho|_{t=0} = \rho^0 \quad \text{in} \quad \Omega, \tag{41}$$

where $\rho^0 : \Omega \to (0, +\infty)$ is taken near the reference value ρ_0. Along the boundary we prescribe

$$\mathbf{u}|_{\partial\Omega} = 0, \quad \nabla\rho \cdot \nu|_{\partial\Omega} = 0, \quad \text{for} \quad 0 < t \leq T. \tag{42}$$

where ν is the outward unit normal at $\partial\Omega$.

2.4. Lyapunov functional

In this section we derive an expression for the free energy which acts as a Lyapunov functional for system (38), (40). This a generalization of the free energy introduced originally by Biot [4].

Let $\{\mathbf{u}, \rho\}$ be a smooth solution of equations (38), (40) that satisfies conditions (41) and (42). Further, let $g : \mathbb{R} \to \mathbb{R}$ be a smooth, strictly increasing and globally Lipschitz function satisfying $g(\rho_0) = 0$.

We first multiply equation (40) by $\partial_t \mathbf{u}$ and integrate the result in Ω. This gives

$$\frac{1}{2}\frac{d}{dt}\int_\Omega \mathcal{G}e(\mathbf{u}) : e(\mathbf{u})\,dx + \lambda^*\int_\Omega (\partial_t\mathcal{E})^2\,dx - \frac{d}{dt}\int_\Omega \mathbf{F}\cdot\mathbf{u}\,dx - \int_\Omega p(\rho)\partial_t\mathcal{E}\,dx = -\int_\Omega \partial_t\mathbf{F}\cdot\mathbf{u}\,dx. \tag{43}$$

Next we multiply (38) by $g(\rho)$ and integrate the result in Ω. This results in

$$\int_\Omega \bar{n}(\mathcal{E})g(\rho)\partial_t\rho\,dx + \int_\Omega d(\rho)g(\rho)\partial_t\mathcal{E}\,dx - \int_\Omega k(\mathcal{E})D(\rho)g'(\rho)\mathbb{K}\nabla\rho\cdot\nabla\rho\,dx = \int_\Omega Qg(\rho)\,dx. \tag{44}$$

With

$$G(\rho) = \int_{\rho_0}^{\rho} g(z)\,dz, \tag{45}$$

the first term in (44) can be written as

$$\int_\Omega \bar{n}(\mathcal{E})\partial_t G(\rho)\,dx = \partial_t\int_\Omega \bar{n}(\mathcal{E})G(\rho)\,dx - \int_\Omega \bar{n}'(\mathcal{E})G(\rho)\partial_t\mathcal{E}\,dx. \tag{46}$$

Note that G is a nonnegative, convex function with $G(\rho_0) = 0$.

We substitute (46) back into (44). Adding the resulting expression and (43) yields

$$\frac{d}{dt}\int_\Omega\left(\frac{1}{2}\mathcal{G}e(\mathbf{u}) : e(\mathbf{u}) + \bar{n}(\mathcal{E})G(\rho) - \mathbf{F}\cdot\mathbf{u}\right)dx + \lambda^*\int_\Omega(\partial_t\mathcal{E})^2\,dx + \int_\Omega k(\mathcal{E})D(\rho)g'(\rho)\mathbb{K}\nabla\rho\cdot\nabla\rho\,dx+$$

$$\int_\Omega\left\{d(\rho)g(\rho) - \bar{n}'(\mathcal{E})G(\rho) - p(\rho)\right\}\partial_t\mathcal{E}\,dx = \int_\Omega Qg(\rho)\,dx - \int_\Omega \partial_t\mathbf{F}\cdot\mathbf{u}\,dx. \tag{47}$$

Before considering the general nonlinear case described by this expression, we first show its implication for the simplified linear setting. Then we use in (38) and (47)

$$\bar{n}(\mathcal{E}) = n_0, \quad d(\rho) = \rho_0, \quad k(\mathcal{E}) = 1 \quad \text{and} \quad D = \frac{1}{\eta_f\beta}.$$

For this choice, the term

$$\int_\Omega \left\{ d(\rho)g(\rho) - \overline{n}'(\mathcal{E})G(\rho) - p(\rho) \right\} \partial_t \mathcal{E} \, dx \tag{48}$$

in expression (47) simplifies to

$$\int_\Omega \left\{ \rho_0 g(\rho) - p(\rho) \right\} \partial_t \mathcal{E} \, dx. \tag{49}$$

Since

$$\int_\Omega \partial_t \mathcal{E} \, dx = 0,$$

expression (49) vanishes if $g(\rho)$ is chosen such that

$$\rho_0 g(\rho) - p(\rho) = \text{constant} = -p_0.$$

This gives

$$g(\rho) = \frac{\rho - \rho_0}{\beta \rho_0^2}$$

and

$$G(\rho) = \frac{(\rho - \rho_0)^2}{2\beta \rho_0^2}.$$

Using these expressions in (47) yields

$$\frac{d}{dt} \int_\Omega \left\{ \frac{1}{2} \mathcal{G}e(\mathbf{u}) : e(\mathbf{u}) + \frac{n_0}{\beta \rho_0^2}(\rho - \rho_0)^2 - \mathbf{F} \cdot \mathbf{u} \right\} dx + \lambda^* \int_\Omega (\partial_t \mathcal{E})^2 \, dx +$$

$$\int_\Omega \frac{1}{\eta_f \beta^2 \rho_0^2} \mathbb{K} \nabla \rho \cdot \nabla \rho \, dx = \int_\Omega Q g(\rho) \, dx - \int_\Omega \partial_t \mathbf{F} \cdot \mathbf{u} \, dx. \tag{50}$$

Hence

$$\mathcal{L}(\mathbf{u}, \rho) = \int_\Omega \left(\frac{1}{2} \mathcal{G}e(\mathbf{u}) : e(\mathbf{u}) + \frac{n_0}{2\beta \rho_0^2}(\rho - \rho_0)^2 - \mathbf{F} \cdot \mathbf{u} \right) dx \tag{51}$$

acts as a Lyapunov functional for the linear form of system (38), (40). The first term denotes the elastic energy of the skeleton, the second term the compression energy of the fluid and the third term the work done by the force $\mathbf{F}$.

Expression (51) coincides with Biot's original free energy expression from [4].

Next we return to the nonlinear case (47). As a first step we restrict ourselves to the physical range of the porosity. Then integral (48) becomes

$$\int_\Omega \left\{ d(\rho)g(\rho) - (1 - n_0)G(\rho) - p(\rho) \right\} \partial_t \mathcal{E} \, dx \tag{52}$$

This integral vanishes if $g(\rho)$ is chosen such that

$$d(\rho)g(\rho) - (1 - n_0)G(\rho) - p(\rho) = -p_0 \tag{53}$$

Differentiating the expression yields a first order equation for g. Thus for (52) to vanish, g should satisfy the initial value problem

$$\begin{cases} d(\rho)g'(\rho) + (d'(\rho) - (1 - n_0))g = \dfrac{1}{\rho_0 \beta}, & \text{for } \rho \in \mathbb{R}; \\ g(\rho_0) = 0. \end{cases} \tag{54}$$

We first consider this problem in the interval $|\rho - \rho_0| < \bar{\rho} := \rho_0 - \rho_*$ where $d(\rho) = \rho$. Then (54) reduces to

$$
\begin{cases}
\rho g' + n_0 g = \dfrac{1}{\beta \rho_0}, \\
g(\rho_0) = 0.
\end{cases}
\tag{55}
$$

Direct integration results in

$$
g(\rho) = \frac{1}{\beta n_0 \rho_0} \left(1 - (\frac{\rho_0}{\rho})^{n_0} \right).
\tag{56}
$$

A second integration yields

$$
G(\rho) = \int_{\rho_0}^{\rho} g(\xi)\, d\xi = \frac{1}{\beta n_0 (1 - n_0) \rho_0} \left((1 - n_0)\rho - \rho_0^{n_0} \rho^{1-n_0} + n_0 \rho_0 \right).
\tag{57}
$$

When $|\rho - \rho_0| > \bar{\rho}$, the function $d(\rho)$ has not yet been defined. We do this by first extending $g(\rho)$ for $|\rho - \rho_0| > \bar{\rho}$ and then by solving $d(\rho)$ from (53): i.e.

$$
d(\rho) = \frac{(1 - n_0)G(\rho) + p(\rho) - p_0}{g(\rho)}.
\tag{58}
$$

Clearly, (56) cannot be used for $\rho \leq 0$. Instead we extend (56) in a linear C^1−manner for $|\rho - \rho_0| > \bar{\rho}$. With $\tilde{\rho} = \rho_0 + \bar{\rho} = 2\rho_0 - \rho_*$, we set

$$
g(\rho) =
\begin{cases}
\dfrac{1}{\beta n_0 \rho_0} \left\{ 1 - (\dfrac{\rho_0}{\rho_*})^{n_0} + \dfrac{\rho - \rho_*}{\rho_*} (\dfrac{\rho_0}{\rho_*})^{n_0} \right\} & \text{for } \rho < \rho_*, \\[3mm]
\dfrac{1}{\beta n_0 \rho_0} \left\{ 1 - (\dfrac{\rho_0}{\tilde{\rho}})^{n_0} + \dfrac{\rho - \tilde{\rho}}{\tilde{\rho}} (\dfrac{\rho_0}{\tilde{\rho}})^{n_0} \right\} & \text{for } \rho > \tilde{\rho},
\end{cases}
\tag{59}
$$

yielding

$$
G(\rho) =
\begin{cases}
G(\rho_*) + \dfrac{\rho - \rho_*}{\beta n_0 \rho_0} \left(1 - (\dfrac{\rho_0}{\rho_*})^{n_0} + \dfrac{\rho - \rho_*}{2\rho_*} (\dfrac{\rho_0}{\rho_*})^{n_0} \right) & \text{for } \rho < \rho_*; \\[3mm]
G(\tilde{\rho}) + \dfrac{\rho - \tilde{\rho}}{\beta n_0 \rho_0} \left(1 - (\dfrac{\rho_0}{\tilde{\rho}})^{n_0} + \dfrac{\rho - \tilde{\rho}}{2\tilde{\rho}} (\dfrac{\rho_0}{\tilde{\rho}})^{n_0} \right) & \text{for } \rho > \tilde{\rho}.
\end{cases}
\tag{60}
$$

Substituting expressions (59) and (60) in (58), yields the desired extension for $d(\rho)$ when $|\rho - \rho_0| > \bar{\rho}$. Thus

$$
d(\rho) =
\begin{cases}
\rho & \text{for } |\rho - \rho_0| \leq \bar{\rho}, \\
\text{(58) with } g \text{ and } G \text{ given by (59) and (60)} & \text{for } |\rho - \rho_0| > \bar{\rho}.
\end{cases}
\tag{61}
$$

Hence the triple $\{g(\rho), G(\rho), d(\rho)\}$ constructed above satisfies (53). For this choice the integral (52) drops from expression (47).

Next we introduce a second modification to deal with a porosity satisfying (36). Starting point is (48). This integral vanishes if

$$
d(\rho)g(\rho) - \bar{n}'(\mathcal{E})G(\rho) = p(\rho) - p_0.
\tag{62}
$$

Keeping g as in (56), (59) and G as in (57), (60), we now modify $d(\rho)$, calling it $D(\rho, \mathcal{E})$, such that

$$
D(\rho, \mathcal{E}) = \frac{\bar{n}'(\mathcal{E})}{g(\rho)} G(\rho) + \frac{p(\rho) - p_0}{g(\rho)}.
\tag{63}
$$

Using (58) in this expression gives

$$
D(\rho, \mathcal{E}) = d(\rho) + (\bar{n}'(\mathcal{E}) - (1 - n_0)) \frac{G(\rho)}{g(\rho)}.
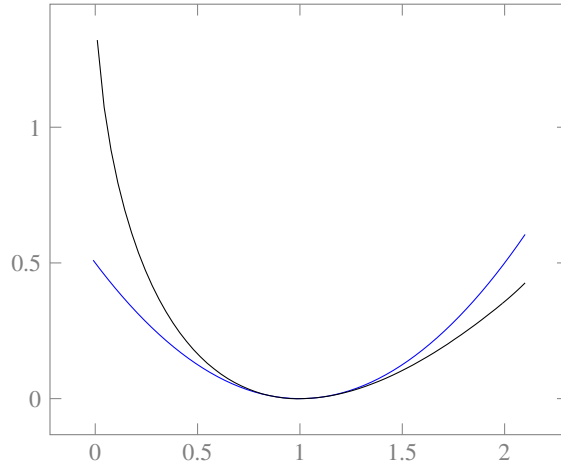\tag{64}
$$

Figure 3: Sketch of the free energy $\beta G(\rho/\rho_0)$. The linear case is in blue. The nonlinear case, see (57) and (60) with $n_0 = 1/3$ and $\rho_*/\rho_0 = 0.01$, is in black.

Clearly, for $|\rho - \rho_0| < \bar{\rho}$ and $\mathcal{E}_* < \mathcal{E} < \mathcal{E}^*$, this expression reduces to

$$D(\rho, \mathcal{E}) = \rho.$$

Finally we use in the Darcy term from equation (38)

$$D(\rho) = \frac{1}{\eta_f \rho_0 \beta} \begin{cases} \tilde{\rho}, & \text{for } \rho > \tilde{\rho}; \\ \rho, & \text{for } \rho_* < \rho < \tilde{\rho}; \\ \rho_*, & \text{for } \rho < \rho_*. \end{cases} \tag{65}$$

Thus in the end we consider the "second" modified fluid mass balance

$$\bar{n}(\text{div } \mathbf{u})\partial_t \rho + D(\rho, \text{div } \mathbf{u}) \text{ div } \partial_t \mathbf{u} = \text{ div } \left(k(\bar{n}(\text{div } \mathbf{u}))D(\rho)\mathbb{K}\nabla\rho \right) + Q. \tag{66}$$

System (40)-(66) serves as starting point of the analysis. The function $D(\rho, \mathcal{E})$ in (66) generalizes the fluid density. It is chosen so that

$$J(\mathbf{u}, \rho) = \frac{1}{2} \int_\Omega \mathcal{G}e(\mathbf{u}) : e(\mathbf{u}) \, dx + \int_\Omega \bar{n}(\text{div } \mathbf{u})G(\rho) \, dx - \int_\Omega \mathbf{F} \cdot \mathbf{u} \, dx \tag{67}$$

acts as a Lyapunov functional for the system. The function $G : \mathbb{R} \to \mathbb{R}$ satisfies $G(\rho_0) = 0$, $G(\rho) > 0$ if $\rho \neq \rho_0$ and G is strictly convex, with quadratic behavior for large values of $|\rho|$. It is explicitly given by (57) and (60).

2.5. Summary of equations and weak formulation

The problem describing the nonlinear poroelastic behavior of a fluid saturated porous medium is to find the displacement $\mathbf{u} : \overline{Q}_T \to \mathbb{R}^m$ and the fluid density $\rho : \overline{Q}_T \to \mathbb{R}$ satisfying

(i) the balance equations

$$\bar{n}(\mathcal{E})\partial_t \rho + D(\rho, \mathcal{E})\partial_t \mathcal{E} = \text{ div } \left(k(\mathcal{E})D(\rho)\mathbb{K}\nabla\rho \right) + Q, \tag{68}$$

$$- \text{ div } \left(\mathcal{G}e(\mathbf{u}) + \lambda^* \partial_t \mathcal{E}\mathbb{I} - p(\rho)\mathbb{I} \right) = \mathbf{F}, \tag{69}$$

in $Q_T = (0, T) \times \Omega$ and

(ii) the initial-boundary conditions (41)-(42).

The coefficients in equations (68)-(69) were introduced in this section. Specifically,
$\overline{n}(\mathcal{E})$ and $k(\mathcal{E})$ satisfy (36) and Remark 4,
$D(\rho, \mathcal{E})$, $\mathcal{D}(\rho)$ and $p(\rho)$ are given by (63), (65) and (37),
and $\lambda^* \geq 0$.

We recast this classical formulation in the following weak form.

Definition 5. *We call a triple* $(\mathbf{u}, \mathcal{E}, \rho) \in L^\infty(0, T; H^1(\Omega)^m) \times L^\infty(0, T; H^1_{loc}(\Omega)) \times \left(L^2(0, T; H^1(\Omega)) \cap L^\infty(0, T; L^2(\Omega)) \right)$,
$\partial_t \mathcal{E} \in L^2(Q_T) \cap L^\infty(0, T; H^1_{loc}(\Omega))$ *a* **weak free energy** *solution if*

(i)

$$-\int_0^T \int_\Omega \rho \overline{n}(\mathcal{E}) \partial_t \Phi \, dx dt - \int_\Omega n_0 \rho^0(x) \Phi(x, 0) \, dx + \int_0^T \int_\Omega \partial_t \mathcal{E} \left(D(\rho, \mathcal{E}) - \rho \overline{n}'(\mathcal{E}) \right) \Phi \, dx dt +$$

$$\int_0^T \int_\Omega k(\mathcal{E}) \mathcal{D}(\rho) \mathbb{K} \nabla \rho \cdot \nabla \Phi \, dx dt = \int_0^T \int_\Omega Q \Phi \, dx dt, \quad \forall \Phi \in H^1(Q_T), \ \Phi|_{t=T} = 0; \tag{70}$$

(ii)

$$\mathcal{E} = div \ \mathbf{u}; \tag{}$$

(iii)

$$\int_\Omega \mathcal{G} e(\mathbf{u}) : e(\xi) \, dx + \lambda^* \partial_t \int_\Omega \mathcal{E} \ div \ \xi \, dx - \int_\Omega p(\rho) div \ \xi \, dx =$$

$$\int_\Omega \mathbf{F} \cdot \xi \, dx, \quad \forall \xi \in H^1_0(\Omega)^3 \quad and \quad for \ almost \ all \quad t \in (0, T]; \tag{71}$$

(iv)

$$\mathcal{E}|_{t=0} = 0 \quad in \quad \Omega. \tag{72}$$

(v) For every $t_1, t_2 \in [0, T]$, $t_1 < t_2$,

$$\int_\Omega \left(\frac{1}{2} \mathcal{G} e(\mathbf{u}(t_2)) : e(\mathbf{u}(t_2)) + \overline{n}(\mathcal{E})(t_2)) G(\rho(t_2)) - \mathbf{F}(t_2) \cdot \mathbf{u}(t_2) \right) dx +$$

$$\int_{t_1}^{t_2} \int_\Omega \left(\lambda^* (\partial_t \mathcal{E})^2 + k(\mathcal{E}) \mathcal{D}(\rho) g'(\rho) \mathbb{K} \nabla \rho \cdot \nabla \rho - Q g(\rho) + \partial_t \mathbf{F} \cdot \mathbf{u} \right) dx dt \leq$$

$$\int_\Omega \left(\frac{1}{2} \mathcal{G} e(\mathbf{u}(t_1)) : e(\mathbf{u}(t_1)) + \overline{n}(\mathcal{E}(t_1)) G(\rho(t_1)) - \mathbf{F}(t_1) \cdot \mathbf{u}(t_1) \right) dx, \tag{73}$$

where $g(\rho)$ *and* $G(\rho)$ *are given, respectively, by (56), (59) and (57), (60).*

Here $\rho^0 \in L^2(\Omega)$, $Q \in C([0, T]; L^2(\Omega))$ *and* $\mathbf{F} \in H^1(0, T; L^2(\Omega)^m)$.

In Definition 5 we explicitly incorporate energy inequality (73). When dealing with classical solutions, equations (68)-(69) imply the energy balance (see (47), (48) and (67))

$$\partial_t J(\mathbf{u}, \rho) + \int_\Omega \lambda^* (\partial_t \mathcal{E})^2 \, dx + \int_\Omega k(\mathcal{E}) \mathcal{D}(\rho) g'(\rho) \mathbb{K} \nabla \rho \cdot \nabla \rho \, dx = \int_\Omega Q g(\rho) \, dx - \int_\Omega \partial_t \mathbf{F} \cdot \mathbf{u} \, dx. \tag{74}$$

However, in the weak formulation (70)-(71) we cannot use $\Phi = g(\rho)$ and $\xi = \partial_t \mathbf{u}$, due to lack of smoothness. Therefore (v) has to be added explicitly. Hence we consider only those weak solutions satisfying additionally (73). Therefore they are called **weak free energy** solutions.

In a number of steps we prove existence of weak solutions when $\lambda^* > 0$. We achieve this by first considering the incremental formulation. In this approximation, which is clearly relevant when treating the problem numerically, we obtain existence results which hold for all $\lambda^* \geq 0$.

3. Existence of a solution to the incremental problem

Because of the quasi-static nature of equation (69), in which the time dependence of the displacement enters mainly through the pressure, we study in this section the time discretized form of (68), (69).

In doing so we use the function $g = g(\rho)$, defined by (56) and (59) as the primary unknown. This is allowed since $g : \mathbb{R} \to \mathbb{R}$ is smooth and strictly increasing. The switch to g is done for mathematical convenience, because it allows us to obtain Lyapunov functional estimates in a straightforward way. We start with some definitions. Let

$$p(g) := p(\rho(g)) \quad \text{and} \quad D(g) := D(\rho(g))\rho'(g). \tag{75}$$

Further, since

$$G(\rho(z)) = \int_{\rho_0}^{\rho(z)} g(\xi)\, d\xi = \int_0^z \zeta \rho'(\zeta)\, d\zeta, \quad z \in \mathbb{R}, \tag{76}$$

let

$$\left.\begin{array}{l} G(g) := \int_0^g \zeta \rho'(\zeta)\, d\zeta \\[2mm] \text{and, from (63),} \\[2mm] D(g, \mathcal{E}) = \dfrac{\overline{n}'(\mathcal{E})}{g} G(g) + \dfrac{p(g) - p_0}{g}. \end{array}\right\} \tag{77}$$

Note that the first term in $D(g, \mathcal{E})$ is bounded with respect to $\mathcal{E}$ and grows linearly in g for large $|g|$. The second (pressure) term is bounded with respect to g since

$$p(g) - p_0 = p(\rho(g)) - p_0 = \frac{\rho(g) - \rho_0}{\beta \rho_0}.$$

Using these definitions in (68) and (69), we find in terms of g

$$\overline{n}(\mathcal{E})\partial_t \rho(g) + D(g, \mathcal{E})\partial_t \mathcal{E} = \text{div}\left(k(\mathcal{E})D(g)\mathbb{K}\nabla g\right) + Q, \tag{78}$$

$$-\text{div}\left(\mathcal{G}e(\mathbf{u}) + \lambda^* \partial_t \mathcal{E}\mathbb{I} - p(g)\mathbb{I}\right) = \mathbf{F}, \tag{79}$$

in Q_T.

Next we turn to the time discretized form of equations (78) and (79).

Let $\tau \in (0, 1)$ denote the time discretization step and $N \in \mathbb{N}$ a large integer such that $N\tau = T$. At each discrete time $t_j = j\tau$, with $j = 0, 1, \ldots, N$, we set

$$\mathbf{F}^j(x) = \mathbf{F}(x, j\tau), \; Q^j(x) = Q(x, j\tau), \quad x \in \Omega.$$

Let $\mathbf{u}^{j-1}$ and g^{j-1} denote, respectively, the displacement and transformed density at t_{j-1} for some $j \in \{1, 2, \ldots, N\}$: i.e.

$$\mathbf{u}^{j-1}(x) = \mathbf{u}(x, t_{j-1}), \; g^{j-1}(x) = g(x, t_{j-1}), \quad x \in \Omega.$$

Then $\mathbf{u}$ and g at time t_j are obtained as solutions of the incremental problem (writing $\mathbf{U} = \mathbf{u}^{j-1}$, $\Xi = g^{j-1}$ and $V = H_0^1(\Omega)^m \times H^1(\Omega)$:

Problem (**PD**): Given $(\mathbf{U}, \Xi) \in V$, find $(\mathbf{u}, g) \in V$ such that

$$\int_\Omega \frac{\overline{n}(\text{div }\mathbf{U})}{\tau}(\rho(g) - \rho(\Xi))\psi \, dx + \int_\Omega D_\tau(g, \text{div }\mathbf{u}, \text{div }\mathbf{U}) \; \text{div } \frac{\mathbf{u} - \mathbf{U}}{\tau} \psi \, dx$$

$$+ \int_\Omega k(\text{div }\mathbf{u})D(g)\mathbb{K}\nabla g \cdot \nabla \psi \, dx = \int_\Omega Q^j \psi \, dx, \quad \forall \psi \in H^1(\Omega); \tag{80}$$

$$\int_\Omega \mathcal{G}e(\mathbf{u}) : e(\xi) \, dx + \frac{\lambda^*}{\tau} \int_\Omega \text{div } (\mathbf{u} - \mathbf{U}) \, \text{div } \xi \, dx - \int_\Omega p(g) \text{div } \xi \, dx = \int_\Omega \mathbf{F}^j \cdot \xi \, dx, \quad \forall \xi \in H_0^1(\Omega)^m. \tag{81}$$

The coefficient D_τ in equation (80) is given by

$$D_\tau(g, \text{ div }\mathbf{u}, \text{ div }\mathbf{U}) = \frac{\overline{n}(\text{ div }\mathbf{u}) - \overline{n}(\text{ div }\mathbf{U})}{\text{div }\mathbf{u} - \text{ div }\mathbf{U}} \frac{G(g)}{g} + \frac{p(g) - p_0}{g}. \tag{82}$$

This expression results from $D(g, \mathcal{E})$ in (77), when the derivative $\overline{n}'(\mathcal{E})$ is replaced by the finite difference $\dfrac{\overline{n}(\text{ div }\mathbf{u}) - \overline{n}(\text{ div }\mathbf{U})}{\text{div }\mathbf{u} - \text{ div }\mathbf{U}}$. The specific choice of (82) appears convenient in the estimates concerning the time discrete Lyapunov functional.

Using the weak topology of the space $H_0^1(\Omega)^m \times H^1(\Omega)$, serious difficulties arise with the coefficients $\overline{n}$, D_τ and k depending on div $\mathbf{u}$. To remedy this, we introduce a Friedrichs mollifier Υ_ε, where ε is a small positive parameter (see e.g. [30], page 203), and replace div $\mathbf{u}$ in the nonlinearities by the convolution div $\mathbf{u} \star \Upsilon_\varepsilon = -\mathbf{u} \star \nabla\Upsilon_\varepsilon$. Using this substitution one can treat nonlinear coefficients containing div $\mathbf{u}$ as lower order terms in the equations. This allows us to use the theory of pseudo-monotone operators.

Applying this convolution, the regularized form of problem (**PD**) reads:

Problem (**PD**)$_\varepsilon$: Given $(\mathbf{U}, \Xi) \in V$, find $(\mathbf{u}_\varepsilon, g_\varepsilon) \in V$ such that, with $\mathcal{E}_\varepsilon = -\mathbf{u}_\varepsilon \star \nabla\Upsilon_\varepsilon$,

$$\int_\Omega \frac{\overline{n}(\text{div }\mathbf{U})}{\tau}(\rho(g_\varepsilon) - \rho(\Xi))\psi \, dx + \int_\Omega \left(\frac{\overline{n}(\mathcal{E}_\varepsilon) - \overline{n}(\text{div }\mathbf{U})}{\tau g_\varepsilon} G(g_\varepsilon) + \frac{p(g_\varepsilon) - p_0}{\tau g_\varepsilon} \right) \text{div }(\mathbf{u}_\varepsilon - \mathbf{U})\psi \, dx$$

$$+ \int_\Omega k(\mathcal{E}_\varepsilon)D(g_\varepsilon)\mathbb{K}\nabla g_\varepsilon \cdot \nabla\psi \, dx = \int_\Omega Q^j \psi \, dx, \quad \forall \psi \in H^1(\Omega), \tag{83}$$

$$\int_\Omega \mathcal{G}e(\mathbf{u}_\varepsilon) : e(\xi) \, dx + \frac{\lambda^*}{\tau} \int_\Omega \text{div }(\mathbf{u}_\varepsilon - \mathbf{U}) \text{ div } \xi \, dx - \int_\Omega p(g_\varepsilon)\text{div }\xi \, dx = \int_\Omega \mathbf{F}^j \cdot \xi \, dx, \quad \forall \xi \in H_0^1(\Omega)^m. \tag{84}$$

We have the following existence result

Proposition 6. *Let $\varepsilon > 0$ be a small positive constant. Under the assumptions of Definition 5, problem (**PD**)$_\varepsilon$ admits at least one solution $(\mathbf{u}_\varepsilon, g_\varepsilon) \in V$.*

Proof. We start by introducing a nonlinear operator $\mathcal{A}$, defined on V and with values in its dual V'. It results from adding (83) and (84), which we write as, with $(\mathbf{u}, g) \in V$,

$$\mathcal{A}(\mathbf{u}, g) = \mathbf{b}, \tag{85}$$

where

$$\langle \mathcal{A}(\mathbf{u}, g), (\xi, \psi) \rangle := \frac{1}{\tau} \int_\Omega \mathcal{G}e(\mathbf{u}) : e(\xi) \, dx + \frac{\lambda^*}{\tau^2} \int_\Omega \text{div }(\mathbf{u} - \mathbf{U}) \text{ div } \xi \, dx$$

$$- \int_\Omega \frac{p(g)}{\tau}\text{div }\xi \, dx + \int_\Omega k(-\mathbf{u} \star \nabla\Upsilon_\varepsilon)D(g)\mathbb{K}\nabla g \cdot \nabla\psi \, dx + \int_\Omega \frac{\overline{n}(\text{div }\mathbf{U})}{\tau}(\rho(g) - \rho(\Xi))\psi \, dx +$$

$$\int_\Omega \left(\frac{\overline{n}(-\mathbf{u} \star \nabla\Upsilon_\varepsilon) - \overline{n}(\text{div }\mathbf{U})}{\tau g} G(g) + \frac{p(g) - p_0}{\tau g} \right) \text{div }(\mathbf{u} - \mathbf{U})\psi \, dx, \quad \forall(\xi, \psi) \in V. \tag{86}$$

and

$$\langle b, (\xi, \psi) \rangle := \int_\Omega \mathbf{F}^j \cdot \xi \, dx + \int_\Omega Q^j \psi \, dx, \quad \forall(\xi, \psi) \in V. \tag{87}$$

The idea is to show that $\mathcal{A}$ is a perturbed monotone operator: i.e. $\mathcal{A}$ is monotone in its principal part containing derivatives of $\mathbf{u}$ and g. To be precise, we show that $\mathcal{A}$ is pseudomonotone and coercive. This allows to apply Brézis' theorem to (85) (see chapter 2 in monographs [19] and [30] or chapters 26 and 27 in [39]) to conclude existence for problem (**PD**)$_\varepsilon$.

For the comfort of the reader we recall that an operator $\mathcal{A} : V \to V'$ is pseudo-monotone if and only if $\mathcal{A}$ is bounded and

$$\left.\begin{array}{l} \{\mathbf{u}_r, g_r\} \rightharpoonup \{\mathbf{u}, g\} \text{ weakly in } V, \\ \limsup_{r\to+\infty}\langle \mathcal{A}(\mathbf{u}_r, g_r), (\mathbf{u}_r, g_r) - (\mathbf{u}, g)\rangle \le 0, \end{array}\right\} \Rightarrow \begin{array}{l} \forall (\mathbf{v}, h) \in V, \\ \langle \mathcal{A}(\mathbf{u}, g), (\mathbf{u}, g) - (\mathbf{v}, h)\rangle \le \\ \liminf_{r\to+\infty}\langle \mathcal{A}(\mathbf{u}_r, g_r), (\mathbf{u}_r, g_r) - (\mathbf{v}, h)\rangle, \end{array} \tag{88}$$

The boundedness of $\mathcal{A}$ is immediate. To show (88) we follow Chapter 2 from [30] or Chapter 17 from [34] and rewrite $\mathcal{A}$ in a form having a principal part containing partial derivatives of $\mathbf{u}$ (in $e(\mathbf{u})$ and div $\mathbf{u}$) and ∇g, and a lower order part containing $\mathbf{u}$ and g. Specifically, we introduce the operator $\mathcal{B} : V \times V \to V'$ by

$$\langle \mathcal{B}((\mathbf{w}, \ell), (\mathbf{u}, g)), (\xi, \psi)\rangle = \frac{1}{\tau}\int_\Omega \mathcal{G}e(\mathbf{u}) : e(\xi)\, dx + \frac{\lambda^*}{\tau^2}\int_\Omega \text{div }(\mathbf{u} - \mathbf{U})\text{ div }\xi\, dx -$$

$$\int_\Omega \frac{p(\ell)}{\tau}\text{div }\xi\, dx + \int_\Omega \left(\frac{\overline{n}(-\mathbf{w}\star\nabla\Upsilon_\varepsilon) - \overline{n}(\text{div }\mathbf{U})}{\tau\ell}G(\ell) + \frac{p(\ell) - p_0}{\tau\ell}\right)\text{div }(\mathbf{u} - \mathbf{U})\psi\, dx$$

$$+ \int_\Omega \frac{\overline{n}(\text{div }\mathbf{U})}{\tau}(\rho(\ell) - \rho(\Xi))\psi\, dx + \int_\Omega k(-\mathbf{w}\star\nabla\Upsilon_\varepsilon)D(\ell)\mathbb{K}\nabla g \cdot \nabla\psi\, dx, \ \forall(\xi, \psi) \in V. \tag{89}$$

We observe that $\mathcal{B}((\mathbf{u}, g), (\mathbf{u}, g)) = \mathcal{A}(\mathbf{u}, g)$. The introduction of $\mathcal{B}$ is useful because it reflects the monotonicity of the principal part of $\mathcal{A}(\mathbf{u}, g)$. This is a direct consequence of

$$\langle \mathcal{B}((\mathbf{w}, \ell), (\mathbf{u}_1, g_1)) - \mathcal{B}((\mathbf{w}, \ell), (\mathbf{u}_2, g_2)), (\mathbf{u}_1, g_1) - (\mathbf{u}_2, g_2)\rangle \ge 0, \tag{90}$$

with equality if and only if $\mathbf{u}_1 = \mathbf{u}_2$ and $g_1 = g_2$. Inequality (90) is checked by a short computation in (89).

To show (88) we consider a sequence $\{\mathbf{u}_r, g_r\} \subset V$ such that

$$(\mathbf{u}_r, g_r) \rightharpoonup (\mathbf{u}, g) \text{ weakly in } V \text{ and } \limsup_{r\to+\infty}\langle \mathcal{A}(\mathbf{u}_r, g_r), (\mathbf{u}_r, g_r) - (\mathbf{u}, g)\rangle \le 0. \tag{91}$$

As in [30] we set $(\mathbf{u}_\delta, g_\delta) = (1 - \delta)(\mathbf{u}, g) + \delta(\mathbf{v}, h)$, where $\delta \in [0, 1]$ and $(\mathbf{v}, h) \in V$. Using the monotonicity from (90), we obtain

$$\delta\langle \mathcal{A}(\mathbf{u}_r, g_r), (\mathbf{u}, g) - (\mathbf{v}, h)\rangle \ge -\langle \mathcal{A}(\mathbf{u}_r, g_r), (\mathbf{u}_r, g_r) - (\mathbf{u}, g)\rangle +$$

$$\langle \mathcal{B}((\mathbf{u}_r, g_r), (\mathbf{u}_\delta, g_\delta)), (\mathbf{u}_r, g_r) - (\mathbf{u}, g)\rangle + \delta\langle \mathcal{B}((\mathbf{u}_r, g_r), (\mathbf{u}_\delta, g_\delta)), (\mathbf{u}, g) - (\mathbf{v}, h)\rangle. \tag{92}$$

The sequence $(\mathbf{u}_r, g_r)$ is bounded in V and there exists a subsequence which strongly converges in $L^5(\Omega)^m$ and (a.e.) in Ω, to $(\mathbf{u}, g)$. Hence it suffices to pass to the limit along this subsequence. In (92) the terms containing the operator $\mathcal{B}$ are fixed with respect to the gradients. Hence

$$\lim_{r\to+\infty}\langle \mathcal{B}((\mathbf{u}_r, g_r), (\mathbf{u}_\delta, g_\delta)), (\mathbf{u}_r, g_r) - (\mathbf{u}, g)\rangle = 0, \tag{93}$$

and

$$\lim_{r\to+\infty}\langle \mathcal{B}((\mathbf{u}_r, g_r), (\mathbf{v}, h)), (\xi, \psi)\rangle = \langle \mathcal{B}((\mathbf{u}, g), (\mathbf{v}, h)), (\xi, \psi)\rangle, \tag{94}$$

for any $(\xi, \psi) \in V$. With these results, we are in position to pass to the limit $r \to +\infty$ in inequality (92). It yields

$$\delta\liminf_{r\to+\infty}\langle \mathcal{A}(\mathbf{u}_r, g_r), (\mathbf{u}, g) - (\mathbf{v}, h)\rangle \ge -\limsup_{r\to+\infty}\langle \mathcal{A}(\mathbf{u}_r, g_r), (\mathbf{u}_r, g_r) - (\mathbf{u}, g)\rangle +$$

$$\delta\langle \mathcal{B}((\mathbf{u}, g), (\mathbf{u}_\delta, g_\delta)), (\mathbf{u}, g) - (\mathbf{v}, h)\rangle. \tag{95}$$

By the pseudo monotonicity hypothesis (91), inequality (95) implies

$$\liminf_{r\to+\infty}\langle \mathcal{A}(\mathbf{u}_r, g_r), (\mathbf{u}, g) - (\mathbf{v}, h)\rangle \ge \langle \mathcal{B}((\mathbf{u}, g), (\mathbf{u}, g)), (\mathbf{u}, g) - (\mathbf{v}, h)\rangle = \langle \mathcal{A}(\mathbf{u}, g), (\mathbf{u}, g) - (\mathbf{v}, h)\rangle, \ \forall(\mathbf{v}, h) \in V. \tag{96}$$

We use this inequality to conclude

$$\liminf_{r\to+\infty}\langle \mathcal{A}(\mathbf{u}_r, g_r), (\mathbf{u}_r, g_r) - (\mathbf{v}, h)\rangle \ge \liminf_{r\to+\infty}\langle \mathcal{A}(\mathbf{u}_r, g_r), (\mathbf{u}_r, g_r) - (\mathbf{u}, g)\rangle + \liminf_{r\to+\infty}\langle \mathcal{A}(\mathbf{u}_r, g_r), (\mathbf{u}, g) - (\mathbf{v}, h)\rangle =$$

$$\liminf_{r\to+\infty} \underbrace{\langle \mathcal{B}\big((\mathbf{u}_r, g_r), (\mathbf{u}, g)\big), (\mathbf{u}_r, g_r) - (\mathbf{u}, g)\rangle}_{=0 \quad \text{by (93)}} + \liminf_{r\to+\infty} \underbrace{\langle \mathcal{B}\big((\mathbf{u}_r, g_r), (\mathbf{u}_r, g_r)\big) - \mathcal{B}\big((\mathbf{u}_r, g_r), (\mathbf{u}, g)\big), (\mathbf{u}_r, g_r) - (\mathbf{u}, g)\rangle}_{\geq 0 \quad \text{by (90)}}$$

$$+ \liminf_{r\to+\infty} \langle \mathcal{A}(\mathbf{u}_r, g_r), (\mathbf{u}, g) - (\mathbf{v}, h)\rangle \geq \liminf_{r\to+\infty} \langle \mathcal{A}(\mathbf{u}_r, g_r), (\mathbf{u}, g) - (\mathbf{v}, h)\rangle \geq \langle \mathcal{A}(\mathbf{u}, g), (\mathbf{u}, g) - (\mathbf{v}, h)\rangle, \ \forall \{\mathbf{v}, h\} \in V.$$

This completes the proof of the *pseudo monotonicity*.

It remains to prove *coercivity*. We evaluate directly the term

$$\langle \mathcal{A}(\mathbf{u}, g), (\mathbf{u} - \mathbf{U}, g)\rangle.$$

Taking $\xi = \mathbf{u} - \mathbf{U}$ and $\psi = g$ in (86), the cross terms involving the product $p(g)\,\mathrm{div}\,(\mathbf{u} - \mathbf{U})$ cancel and the term $p_0\,\mathrm{div}\,(\mathbf{u} - \mathbf{U})/\tau$ drops out after integration. What remains is

$$\langle \mathcal{A}(\mathbf{u}, g), (\mathbf{u} - \mathbf{U}, g)\rangle = \frac{1}{\tau}\int_\Omega \mathcal{G}e(\mathbf{u}) : e(\mathbf{u} - \mathbf{U})\,dx + \frac{\lambda^*}{\tau^2}\int_\Omega \mathrm{div}\,(\mathbf{u} - \mathbf{U})^2\,dx + \int_\Omega \frac{\overline{n}(\mathrm{div}\,\mathbf{U})}{\tau}(\rho(g) - \rho(\Xi))g\,dx +$$

$$\int_\Omega \frac{\overline{n}(-\mathbf{u} \star \nabla \Upsilon_\varepsilon) - \overline{n}(\mathrm{div}\,\mathbf{U})}{\tau} G(g)\,dx + \int_\Omega \mathbb{K}k(-\mathbf{u} \star \nabla \Upsilon_\varepsilon)\mathcal{D}(g)|\nabla g|^2\,dx. \tag{97}$$

The third and fourth term in the right hand side need special attention.

Since $\rho = \rho(g)$ is a C^1 monotonically increasing function, we have the elementary inequality

$$x(\rho(x) - \rho(y)) \geq \int_y^x \zeta \rho'(\zeta)\,d\zeta, \quad \forall x, y \in \mathbb{R}. \tag{98}$$

Using this inequality and the expression for G (see (77)) in these terms gives

$$\overline{n}(\mathrm{div}\,\mathbf{U})(\rho(g) - \rho(\Xi))g + (\overline{n}(-\mathbf{u} \star \nabla \Upsilon_\varepsilon) - \overline{n}(\mathrm{div}\,\mathbf{U}))\int_0^g \zeta \rho'(\zeta)\,d\zeta \geq$$

$$\overline{n}(-\mathbf{u} \star \nabla \Upsilon_\varepsilon)\int_0^g \zeta \rho'(\zeta)\,d\zeta - \overline{n}(\mathrm{div}\,\mathbf{U})\int_0^\Xi \zeta \rho'(\zeta)\,d\zeta. \tag{99}$$

Applying Korn's inequality, see Theorem 1.33 in [30], and inserting inequality (99) in equality (97) yields

$$\langle \mathcal{A}(\mathbf{u}, g), (\mathbf{u} - \mathbf{U}, g)\rangle \geq \frac{C_1}{\tau}||\mathbf{u}||^2_{H_0^1(\Omega)^m} + C_2||\nabla g||^2_{L^2(\Omega)^m} - \frac{C_3}{\tau} + \frac{C_4}{\tau}\int_\Omega \underbrace{\Big(\int_0^g \zeta \rho'(\zeta)\,d\zeta\Big)}_{\approx Cg^2 \text{ for large } |g|}\,dx, \tag{100}$$

where $C_i, i = 1, \dots, 4$ are positive constants. This proves the coercivity.

Having established pseudo monotonicity and coercivity of the operator $\mathcal{A}$, we are in position to apply Brézis' theorem. This concludes the assertion of the proposition. $\square$

Theorem 7. *Problem* (**PD**) *admits at least one solution* $(\mathbf{u}, g) \in V$.

Proof. For each $\varepsilon > 0$, let $(\mathbf{u}_\varepsilon, g_\varepsilon)$ be a solution of problem (**PD**$_\varepsilon$) as obtained in Proposition 6. From the coercivity part of the proof of Proposition 6 and equation (85), it follows that

$$||\mathbf{u}_\varepsilon||_{H_0^1(\Omega)^m} + ||g_\varepsilon||_{H^1(\Omega)} \leq C, \tag{101}$$

where C is independent of ε. Estimate (101) yields weak compactness in H^1. However this is not enough to prove that $-\mathbf{u}_\varepsilon \star \nabla \Upsilon_\varepsilon$ converges strongly in L^2 and (a.e.) on Ω as $\varepsilon \to 0$. The remedy is to consider the momentum equation (84), which gives us improved regularity through the elasticity term. Since $p(g_\varepsilon)$ is bounded in $H^1(\Omega)$, uniformly with respect to ε, we conclude that

$$||\mathbf{u}_\varepsilon||_{H^2(\Omega)^m} \leq C, \tag{102}$$

where C does not depend on ε. Using estimates (101)-(102), there is a subsequence $(\mathbf{u}_\varepsilon, g_\varepsilon)$, denoted by the same subscript, and a pair $(\mathbf{u}, g) \in (H_0^1(\Omega)^m \cap H^2(\Omega)^m) \times H^1(\Omega)$ such that

$$\mathbf{u}_\varepsilon \to \mathbf{u} \quad \text{strongly in} \quad H_0^1(\Omega)^m, \tag{103}$$

$$\text{div } \mathbf{u}_\varepsilon \to \text{div } \mathbf{u} \quad \text{strongly in} \quad L^2(\Omega) \text{ and (a.e) on } \Omega, \tag{104}$$

$$g_\varepsilon \rightharpoonup g \quad \text{weakly in} \quad H^1(\Omega), \tag{105}$$

$$g_\varepsilon \to g \quad \text{strongly in} \quad L^2(\Omega) \text{ and (a.e) on } \Omega, \tag{106}$$

as $\varepsilon \to 0$. The convergence properties allow to pass to the limit $\varepsilon \to 0$ in the system (83)-(84). Hence the pair $(\mathbf{u}, g)$ satisfies the equations of problem (**PD**), which proves the theorem. $\qquad\square$

To complete the study of the incremental problem, we need to estimate the behavior of solutions after at least $O(1/\tau)$ times steps. Here we use the discrete version of Lyapunov functional (67).

In problem (**PD**), where the discrete time step τ enters as parameter, one finds after one step $(\mathbf{u}^1, g^1)$ from the initial values $(\text{div } \mathbf{u}, \rho)|_{t=0} = (0, \rho^0)$. The idea is to repeat this procedure for an arbitrary number of steps. If $M \in \mathbb{N}$, $M \le N = T/\tau$, then $(\mathbf{u}^M, g^M)$ denotes the time discretized approximation of the original quasi-static equation, at $t = t_M = M\tau$.

The corresponding Lyapunov functional at $t = t_M$ reads

$$J^M = \int_\Omega \left(\frac{1}{2} \mathcal{G}e(\mathbf{u}^M) : e(\mathbf{u}^M) - \mathbf{F}^M \cdot \mathbf{u}^M + \bar{n}(\text{div } \mathbf{u}^M) G(g^M) \right) dx. \tag{107}$$

It satisfies

Theorem 8. *For each $M \in \mathbb{N}$, $M \le N = T/\tau$, we have*

$$J^M + \tau \sum_{j=1}^M \int_\Omega \left(\lambda^* \left(\frac{\text{div } (\mathbf{u}^j - \mathbf{u}^{j-1})}{\tau} \right)^2 + \frac{\mathbf{F}^j - \mathbf{F}^{j-1}}{\tau} \cdot \mathbf{u}^{j-1} + k(\text{div } \mathbf{u}^j) D(g^j) \mathbb{K} \nabla g^j \cdot \nabla g^j - Q^j g^j \right) dx \le J^0. \tag{108}$$

Here

$$J^0 = n_0 \int_\Omega G(g^0) \, dx, \quad g^0 = g(\rho^0).$$

Proof. At time $t = t_j$, with $j = 1, \ldots, N$, the equations in problem (**PD**) read

$$\int_\Omega \mathcal{G}e(\mathbf{u}^j) : e(\xi) \, dx + \frac{\lambda^*}{\tau} \int_\Omega \text{div } (\mathbf{u}^j - \mathbf{u}^{j-1}) \text{ div } \xi \, dx - \int_\Omega p(g^j) \text{div } \xi \, dx = \int_\Omega \mathbf{F}^j \cdot \xi \, dx, \quad \forall \xi \in H_0^1(\Omega)^m, \tag{109}$$

$$\int_\Omega \left(\frac{\bar{n}(\text{div } \mathbf{u}^{j-1})}{\tau} (\rho(g^j) - \rho(g^{j-1})) + \frac{\bar{n}(\text{div } \mathbf{u}^j) - \bar{n}(\text{div } \mathbf{u}^{j-1})}{\tau g^j} G(g^j) \right) \psi \, dx$$

$$+ \int_\Omega \frac{p(g^j) - p_0}{\tau g^j} \text{div } (\mathbf{u}^j - \mathbf{u}^{j-1}) \psi \, dx + \int_\Omega k(\text{div } \mathbf{u}^j) D(g^j) \mathbb{K} \nabla g^j \cdot \nabla \psi \, dx = \int_\Omega Q^j \psi \, dx, \quad \forall \psi \in H^1(\Omega). \tag{110}$$

Note that in equation (110) we have used explicitly the form of D_τ from (82). Next, we take $\xi = (\mathbf{u}^j - \mathbf{u}^{j-1})/\tau$ in (109) and $\psi = g^j$ in (110). The resulting two equalities are added and summed-up with respect to j up from $j = 1$ to $j = M$. Using the observations

(**i**) cross terms containing pressure cancel;

(**ii**)

$$\sum_{j=1}^M \mathcal{G}e(\mathbf{u}^j) : e(\mathbf{u}^j - \mathbf{u}^{j-1}) \ge \frac{1}{2} \left(\mathcal{G}e(\mathbf{u}^M) : e(\mathbf{u}^M) - \mathcal{G}e(\mathbf{u}^0) : e(\mathbf{u}^0) \right);$$

(iii)

$$\sum_{j=1}^{M} \left(\overline{n}(\text{div } \mathbf{u}^{j-1}) g^j (\rho(g^j) - \rho(g^{j-1})) + (\overline{n}(\text{div } \mathbf{u}^j) - \overline{n}(\text{div } \mathbf{u}^{j-1})) G(g^j) \right)$$

$$\geq \overline{n}(\text{div } \mathbf{u}^M) G(g^M) - \overline{n}(\text{div } \mathbf{u}^0) G(g^0),$$

where (98) is used;

(iv)

$$\sum_{j=1}^{M} \mathbf{F}^j \cdot (\mathbf{u}^j - \mathbf{u}^{j-1}) = \mathbf{F}^M \cdot \mathbf{u}^M - \mathbf{F}^0 \cdot \mathbf{u}^0 - \sum_{j=0}^{M-1} (\mathbf{F}^{j+1} - \mathbf{F}^j) \cdot \mathbf{u}^j,$$

one finds inequality (108). The reduced expression for J^0 results from $\mathbf{u}|_{t=0} = 0$.

$\square$

Having established existence for the discrete problem (**PD**) in Theorem 7 and a Lyapunov estimate in Theorem 8, we are now in a position to obtain estimates that are uniform in the time step τ.

Proposition 9. *There exists a constant $C > 0$ such that*

$$||\mathbf{u}^M||^2_{H^1(\Omega)^m} + ||g^M||^2_{L^2(\Omega)} \leq C, \tag{111}$$

and

$$\tau \sum_{j=1}^{M} \int_{\Omega} \left(\lambda^* \left(\frac{\text{div } (\mathbf{u}^j - \mathbf{u}^{j-1})}{\tau} \right)^2 + |\nabla g^j|^2 \right) dx \leq C, \tag{112}$$

for all M and τ such that $1 \leq M \leq N = T/\tau$, with τ sufficiently small.

Proof. Combining expression (107) for J^M and inequality (108), yields for any $1 \leq M \leq N$

$$\frac{1}{2} \int_{\Omega} \mathcal{G}e(\mathbf{u}^M) : e(\mathbf{u}^M) \, dx + \int_{\Omega} \overline{n}(\text{div } \mathbf{u}^M) G(g^M) \, dx \leq \int_{\Omega} \mathbf{F}^M \cdot \mathbf{u}^M \, dx + J^0 + \tau \sum_{j=1}^{M} Q^j g^j \, dx +$$

$$\tau \sum_{j=1}^{M} \int_{\Omega} \frac{\mathbf{F}^j - \mathbf{F}^{j-1}}{\tau} \cdot \mathbf{u}^{j-1} \, dx \leq \frac{\delta}{2} ||\mathbf{u}^M||^2_{L^2(\Omega)^m} + \frac{1}{2\delta} ||\mathbf{F}^M||^2_{L^2(\Omega)^m} + J^0 + \frac{\tau}{2} \sum_{j=1}^{M} ||g^j||^2_{L^2(\Omega)} +$$

$$\frac{\tau}{2} \sum_{j=1}^{M} ||\mathbf{u}^{j-1}||^2_{L^2(\Omega)^m} + \frac{\tau}{2} \sum_{j=1}^{M} ||Q^j||^2_{L^2(\Omega)} + \frac{\tau}{2} \sum_{j=1}^{M} ||\frac{\mathbf{F}^j - \mathbf{F}^{j-1}}{\tau}||^2_{L^2(\Omega)^m}.$$

By the assumptions on Q and $\mathbf{F}$, the last two terms are uniformly bounded with respect to τ and M. We estimate the left hand side from below by applying Korn's inequality to the first term and the quadratic growth of G to the second term. Then for δ and τ sufficiently small, we obtain for the combination

$$\mathcal{U}_j = ||\mathbf{u}^j||^2_{H^1(\Omega)^m} + ||g^j||^2_{L^2(\Omega)}, \quad j = 0, \dots, M,$$

the inequality

$$\mathcal{U}_M \leq C_1 + C_2 \tau \sum_{j=0}^{M-1} \mathcal{U}_j,$$

where C_1 and C_2 do not dependent on τ and M. Next we apply the *discrete Gronwall inequality*[1], see footnote, to find

$$\mathcal{U}_M \leq C_1 e^{C_2(M-1)\tau} < C_1 e^{C_2 T} \quad \text{for all } 1 \leq M \leq N.$$

The second estimate follows directly from Theorem 8.

$\square$

---

[1] *Discrete version of Gronwall's lemma:* Let $\{\mathcal{U}_n\}$ and $\{w_n\}$ be nonnegative sequences satisfying $\mathcal{U}_n \leq A + \sum_{j=0}^{M-1} \mathcal{U}_j w_j$. Then for all n, $\mathcal{U}_n \leq A \exp\{\sum_{j=0}^{M-1} w_j\}$.

However, to pass to the limit $\tau \to 0$ in the nonlinearities, one needs more information on the behavior of the ratios $\{ \operatorname{div} (u^j - u^{j-1})/\tau \}$ and $\{(g^j - g^{j-1})/\tau\}$. In fact, we must establish relative compactness of the sequences $\{\operatorname{div} \mathbf{u}^j\}$ and $\{g^j\}$.

We start with a local H^1-estimate for $\mathcal{E}^j = \operatorname{div} \mathbf{u}^j$.

Lemma 10. *Let $\varphi \in C_0^\infty(\Omega)$ and $\tau > 0$ sufficiently small. Then there exists a constant $C = C(\varphi)$ such that*

$$\tau \sum_{j=1}^N ||\varphi\mathcal{E}^j||_{H^1(\Omega)}^2 + \frac{\lambda^*}{2\mu + \lambda} \max_{1 \le M \le N} ||\varphi\mathcal{E}^M||_{H^1(\Omega)}^2 \le C. \tag{113}$$

Proof. Let

$$L^j = (2\mu + \lambda)\mathcal{E}^j - p(g^j) + \lambda^*\frac{\operatorname{div} (\mathbf{u}^j - \mathbf{u}^{j-1})}{\tau}, \quad j = 1, \dots M. \tag{114}$$

Inequality (111) implies

$$\tau \sum_{j=1}^M \left(||\mathbf{u}^j||_{H^1(\Omega)^m}^2 + ||g^j||_{L^2(\Omega)}^2 \right) \le \tau M C \le T C.$$

Combined with (112) this gives for L^j

$$\tau \sum_{j=1}^M ||L^j||_{L^2(\Omega)}^2 \le C. \tag{115}$$

As in the counterexample for negative porosity, we take the divergence of the time discrete momentum equation. This yields

$$-\Delta L^j = \operatorname{div} \mathbf{F}^j \quad \text{in} \quad \Omega. \tag{116}$$

In general, however, there are no boundary conditions for L^j available. Here we must rely on local estimates to obtain (113). Let us first write the equation for $\varphi L^j \in H^2(\Omega) \cap H_0^1(\Omega)$:

$$\Delta(\varphi L^j) = -\varphi \operatorname{div} \mathbf{F}^j + 2\nabla\varphi \cdot \nabla L^j + L^j \Delta\varphi.$$

Its weak form reads

$$\int_\Omega \nabla(\varphi L^j)\zeta \, dx = \int_\Omega \operatorname{div} \mathbf{F}^j \varphi\zeta \, dx + \int_\Omega L^j(2\nabla\varphi \cdot \nabla\zeta + \zeta\Delta\varphi) \, dx, \quad \forall \zeta \in H_0^1(\Omega). \tag{117}$$

Taking $\zeta = \varphi L^j$ results in

$$\int_\Omega |\nabla(\varphi L^j)|^2 \, dx = -\int_\Omega \mathbf{F}^j \cdot \nabla(\varphi L^j)\varphi \, dx - \int_\Omega \mathbf{F}^j \cdot \nabla\varphi\varphi L^j \, dx + \int_\Omega (L^j)^2 \varphi\Delta\varphi \, dx - \int_\Omega 2L^j \nabla\varphi \cdot \nabla(\varphi L^j) \, dx.$$

With $C = C(\varphi)$ denoting a generic constant depending on φ, we have

$$||\varphi L^j||_{H^1(\Omega)}^2 \le C\left(||\mathbf{F}^j||_{L^2(\Omega)}^2 + ||L^j||_{L^2(\Omega)}^2 \right), \tag{118}$$

for $1 \le j \le M \le N$. Combing this inequality with (115) gives

$$\tau \sum_{j=1}^M ||\varphi L^j||_{H^1(\Omega)}^2 \le C(\tau \sum_{j=1}^M ||\mathbf{F}^j||_{L^2(\Omega)^m}^2 + 1) \le C. \tag{119}$$

Next we multiply expression (114) by $\tau\varphi$ and write it as

$$\tau(2\mu + \lambda)\varphi\mathcal{E}^j + \lambda^*\varphi(\mathcal{E}^j - \mathcal{E}^{j-1}) = \tau L^j\varphi + \tau p(g^j)\varphi \in H^1(\Omega).$$

Taking the H^1-inner product of this expression with $\varphi \mathcal{E}^j$ gives

$$(2\mu + \lambda)\tau ||\varphi\mathcal{E}^j||^2_{H^1(\Omega)} + \lambda^*(\varphi(\mathcal{E}^j - \mathcal{E}^{j-1}), \varphi\mathcal{E}^j)_{H^1(\Omega)} = \tau(\varphi(L^j + p(g^j)), \varphi\mathcal{E}^j)_{H^1(\Omega)}$$

or

$$(2\mu + \lambda)\frac{\tau}{2}||\varphi\mathcal{E}^j||^2_{H^1(\Omega)} + \lambda^*(\varphi(\mathcal{E}^j - \mathcal{E}^{j-1}), \varphi\mathcal{E}^j)_{H^1(\Omega)} \leq \frac{\tau}{2(2\mu + \lambda)}||\varphi(L^j + p(g^j))||^2_{H^1(\Omega)}. \tag{120}$$

Using the identity

$$\sum_{j=1}^{M} a^j(a^j - a^{j-1}) = \frac{(a^M)^2}{2} - \frac{(a^0)^2}{2} + \frac{1}{2}\sum_{j=1}^{M}(a^j - a^{j-1})^2,$$

when summing-up (120) gives

$$\tau\sum_{j=1}^{M}||\varphi\mathcal{E}^j||^2_{H^1(\Omega)} + \frac{\lambda^*}{2\mu + \lambda}||\varphi\mathcal{E}^M||^2_{H^1(\Omega)} \leq +\frac{\tau}{(2\mu + \lambda)^2}\sum_{j=1}^{M}||\varphi(L^j + p(g^j))||^2_{H^1(\Omega)}.$$

Combining this inequality with (112) and (119), results in the estimate of the lemma. □

We conclude this section with an estimate for $(\rho(g^j) - \rho(g_{j-1}))/\tau$. However, since in equations (68) or (70) the (discrete) time derivative is multiplied by $\bar{n}(\mathcal{E})$, we look for an estimate for

$$\mathcal{N}^j = \bar{n}(\mathcal{E}^j)\rho(g^j). \tag{121}$$

With the results of Proposition 9 and Lemma 10, the space-time compactness of $\mathcal{N}$ will imply the same property of g.

We summarize our findings in the next proposition

Proposition 11. *For given $\tau > 0$ and $j = 1, \ldots, N$,, let $(\mathbf{u}_\tau(t_j), g_\tau(t_j)) \in V$ denote a solution of problem (**PD**). Then we have*

$$\max_{1 \leq j \leq N} \left(||\mathbf{u}_\tau(t_j)||_{H^1(\Omega)^m} + ||g_\tau(t_j)||_{L^2(\Omega)}\right) \leq C, \tag{122}$$

$$\tau\sum_{j=1}^{N}\int_\Omega \left(\lambda^*\left(\frac{div\,(\mathbf{u}_\tau(t_j) - \mathbf{u}_\tau(t_{j-1}))}{\tau}\right)^2 + |\nabla g_\tau(t_j)|^2\right) dx \leq C, \tag{123}$$

$$\tau\sum_{j=1}^{N}||\varphi\,div\,\mathbf{u}_\tau(t_j)||^2_{H^1(\Omega)} + \lambda^*\max_{1 \leq j \leq N}||\varphi div\,\mathbf{u}_\tau(t_j)||^2_{H^1(\Omega)} \leq C, \tag{124}$$

$$\tau\sum_{j=1}^{N}\left(||\frac{\mathcal{N}^j - \mathcal{N}^{j-1}}{\tau}||^2_{H^{-2}(\Omega)} + ||\varphi\mathcal{N}^j||^2_{H^1(\Omega)}\right) \leq C, \tag{125}$$

where

$$\mathcal{N}^j = \bar{n}(div\,\mathbf{u}_\tau(t_j))\rho(g_\tau(t_j))$$

and where $\varphi \in C_0^\infty(\Omega)$.

Proof. We only need to prove estimate (125). Rewriting equation (80) we have

$$\int_\Omega \frac{\mathcal{N}^j - \mathcal{N}^{j-1}}{\tau}\psi\,dx = \int_\Omega \frac{\bar{n}(\mathcal{E}^{j-1})(\rho(g^j) - \rho(g^{j-1}))}{\tau}\psi\,dx + \int_\Omega \frac{(\bar{n}(\mathcal{E}^j) - \bar{n}(\mathcal{E}^{j-1}))\rho(g^j)}{\tau}\psi\,dx =$$

$$\int_\Omega \frac{\bar{n}(\mathcal{E}^j) - \bar{n}(\mathcal{E}^{j-1})}{\tau}\left(\rho(g^j) - G(g^j)\right)\psi\,dx - \int_\Omega \frac{(\mathcal{E}^j - \mathcal{E}^{j-1})(\rho(g^j) - p_0)}{\tau g^j}\psi\,dx + \int_\Omega Q^j\psi\,dx$$

$$- \int_\Omega k(\mathcal{E}^j)D(g^j)\mathbb{K}\nabla g^j\nabla\psi\,dx, \quad \text{for } \psi \in H_0^2(\Omega).$$

Recalling that for $m \leq 3$, $H^2(\Omega) \subset L^\infty(\Omega)$, we have

$$||\frac{\mathcal{N}^j - \mathcal{N}^{j-1}}{\tau}||^2_{H^{-2}(\Omega)} \leq C\left(||\frac{\mathcal{E}^j - \mathcal{E}^{j-1}}{\tau}||^2_{L^2(\Omega)}||g^j||^2_{L^2(\Omega)} + ||Q^j||^2_{L^2(\Omega)} + ||\nabla g^j||^2_{L^2(\Omega)^m}\right)$$

and the full estimate reads

$$\tau \sum_{j=1}^N ||\frac{\mathcal{N}^j - \mathcal{N}^{j-1}}{\tau}||^2_{H^{-2}(\Omega)} \leq$$

$$C\left(\max_{1 \leq j \leq N} ||g^j||^2_{L^2(\Omega)} \tau \sum_{j=1}^N ||\frac{\mathcal{E}^j - \mathcal{E}^{j-1}}{\tau}||^2_{L^2(\Omega)} + 1 + \tau \sum_{j=1}^N ||\nabla g^j||^2_{L^2(\Omega)^m}\right) \leq C \qquad (126)$$

The local estimate for the space derivatives is given by

$$\tau \sum_{j=1}^N ||\nabla(\varphi \mathcal{N}^j)||^2_{L^{3/2}(\Omega)} \leq C\left((\max_{1 \leq j \leq N} ||\nabla(\varphi \mathcal{E}^j)||^2_{L^2(\Omega)}) \tau \sum_{j=1}^N ||g^j||^2_{L^6(\Omega)} + C + \tau \sum_{j=1}^N ||\nabla g^j||^2_{L^2(\Omega)^m}\right) \leq C. \qquad (127)$$

This results in estimate (125). $\qquad \square$

4. Existence for continuous time problem with $\lambda^* > 0$

In Proposition 11, where the time step τ enters as a parameter, one finds $\{(\mathbf{u}_\tau(t_j), g_\tau(t_j))\}_{j=1,\dots,N}$ from the "initial value" div $\mathbf{u}(0) = 0$ and $g(0) = g^0$. Here $N = O(1/\tau)$ and $g^0 = g(\rho^0)$. This procedure yields a time discretized approximation of the original quasi-static equations.

In this section we investigate the limit $\tau \searrow 0$. Here a crucial role is played by the parameter λ^*, which is needed to control the behaviour in time of $\mathcal{E} = $ div $\mathbf{u}$.

Using the discrete solution $(\mathbf{u}_\tau(t_j), g_\tau(t_j))$, we construct two approximations that hold for all $0 \leq t \leq T$. The first is the piecewise constant approximation

$$(\overline{\mathbf{u}}_\tau(t), \overline{g}_\tau(t)) = (\mathbf{u}_\tau(t_j), g_\tau(t_j)) \quad \text{for} \quad j\tau \leq t < (j+1)\tau. \qquad (128)$$

The second is the Rothe interpolant, which is the piecewise linear time-continuous approximation

$$(\tilde{\mathbf{u}}_\tau(t), \tilde{g}_\tau(t)) = \left(j + 1 - \frac{t}{\tau}\right)(\mathbf{u}_\tau(t_j), g_\tau(t_j)) + \left(\frac{t}{\tau} - j\right)(\mathbf{u}_\tau(t_{j+1}), g_\tau(t_{j+1})),$$

$$\text{for} \quad j\tau \leq t \leq (j+1)\tau. \qquad (129)$$

In (128) and (129) the index j runs from $j = 0$ to $j = N - 1$.

Applying Proposition 11, yields for both approximations, with $\natural \in \{^-, ^\sim\}$,

$$\max_{0 \leq t \leq T} \left(||\mathbf{u}^\natural_\tau(t)||^2_{H^1(\Omega)^m} + ||g^\natural_\tau(t)||^2_{L^2(\Omega)}\right) dt \leq C, \qquad (130)$$

$$\int_0^T \int_\Omega |\nabla g^\natural_\tau(t)|^2 \, dxdt \leq C, \qquad (131)$$

$$\int_0^T ||\varphi \, \mathcal{E}^\natural_\tau(t)||^2_{H^1(\Omega)} \, dt \leq C, \qquad (132)$$

$$\lambda^* \max_{0 \leq t \leq T} ||\varphi \mathcal{E}^\natural_\tau(t)||^2_{H^1(\Omega)} \leq C, \qquad (133)$$

$$\int_0^T ||\varphi \mathcal{N}^\natural_\tau(t)||^2_{W^{1,3/2}(\Omega)}) \leq C, \qquad (134)$$

where $\mathcal{E}^\natural_\tau = $ div $\mathbf{u}^\natural_\tau$, $\overline{\mathcal{N}}_\tau = \overline{n}(\overline{\mathcal{E}}_\tau)\rho(\overline{g}_\tau)$ and $\tilde{\mathcal{N}}_\tau(t) = (j + 1 - t/\tau)\mathcal{N}^j + (t/\tau - j)\mathcal{N}^{j+1}$.

Further we have

$$\partial_t \tilde{\mathcal{N}}_\tau = \frac{\mathcal{N}^{j+1} - \mathcal{N}^j}{\tau} \quad \text{and} \quad \partial_t \tilde{\mathcal{E}}_\tau = \frac{\mathcal{E}^{j+1} - \mathcal{E}^j}{\tau},$$
$$\text{for} \quad t_j \leq t \leq t_{j+1} \quad \text{and} \quad j = 0, \dots, N-1.$$

Hence, by (123)

$$\int_0^T \int_\Omega \lambda^* |\partial_t \tilde{\mathcal{E}}_\tau(t)|^2 \, dxdt \leq C \tag{135}$$

and

$$\int_0^T ||\partial_t \tilde{\mathcal{N}}_\tau(t)||^2_{H^{-2}(\Omega)} \, dt \leq C. \tag{136}$$

In what follows we rely heavily on the material and theory collected in [[30], chapters 7 and 8]. Since the piecewise constant approximation $(\overline{\mathbf{u}}_\tau(t), \overline{g}_\tau(t))$ is discontinuous in time, its time derivative is only a measure. To deal with this we introduce the space $\mathcal{M}(0, T; L^2(\Omega))$ of regular Borel measures in $[0, T]$ with values in $L^2(\Omega)$, which is the dual space of $C([0, T]; L^2(\Omega))$. With $\delta(t_j)$ denoting the Dirac measure concentrated in t_j, we have

$$||\partial_t \overline{\mathcal{E}}_\tau||_{\mathcal{M}(0,T;L^2(\Omega))} = || \sum_{j=1}^N (\mathcal{E}^j - \mathcal{E}^{j-1}) \delta(t_j) ||_{\mathcal{M}(0,T;L^2(\Omega))}$$

$$= \tau \sum_{j=1}^N || \frac{\mathcal{E}^j - \mathcal{E}^{j-1}}{\tau} ||_{L^2(\Omega)} = ||\partial_t \tilde{\mathcal{E}}_\tau||_{L^1(0,T;L^2(\Omega))} \leq$$

$$\sqrt{T} ||\partial_t \tilde{\mathcal{E}}_\tau||_{L^2(0,T;L^2(\Omega))} \leq C. \tag{137}$$

Analogously

$$||\partial_t \overline{\mathcal{N}}_\tau||_{\mathcal{M}(0,T;H^{-2}(\Omega))} \leq C, \tag{138}$$

where $\mathcal{M}(0, T; H^{-2}(\Omega))$ is the dual space of $C([0, T]; H_0^2(\Omega))$.

For the convergence of the time continuous approximation (129) we use estimates (130)-(136) and the well-known weak and weak* compactness theorems. The result is that there exists a quadruple $\{\tilde{\mathbf{u}}, \tilde{g}, \tilde{\mathcal{E}}, \tilde{\mathcal{N}}\}$ such that along a subsequence $\tau \searrow 0$ we have

$$\tilde{\mathbf{u}}_\tau \to \tilde{\mathbf{u}} \quad \text{weak* in} \quad L^\infty(0, T; H_0^1(\Omega)^m), \tag{139}$$

$$\tilde{g}_\tau \rightharpoonup \tilde{g} \quad \text{weakly in} \quad L^2(0, T; H^1(\Omega)), \tag{140}$$

$$\tilde{\mathcal{E}}_\tau \rightharpoonup \tilde{\mathcal{E}} \quad \text{weakly in} \quad L^2(0, T; H^1(\omega)), \tag{141}$$

$$\partial_t \tilde{\mathcal{E}}_\tau \rightharpoonup \partial_t \tilde{\mathcal{E}} \quad \text{weakly in} \quad L^2(0, T; L^2(\Omega)), \tag{142}$$

$$\tilde{\mathcal{N}}_\tau \rightharpoonup \tilde{\mathcal{N}} \quad \text{weakly in} \quad L^2(0, T; W^{1,3/2}(\omega)), \tag{143}$$

$$\partial_t \tilde{\mathcal{N}}_\tau \rightharpoonup \partial_t \tilde{\mathcal{N}} \quad \text{weakly in} \quad L^2(0, T; H^{-2}(\Omega)). \tag{144}$$

Concerning the convergence of $(\overline{\mathbf{u}}_\tau, \overline{g}_\tau)$, we use estimates (130)-(134), now combined with (137)-(138). Moreover, applying [([30]), Corollary 7.9], we use that the spaces

$$W^{1,2,\mathcal{M}}(0, T; H^1(\omega), L^2(\omega)) = \{z \in L^2(0, T; H^1(\omega)) \mid \frac{dz}{dt} \in \mathcal{M}(0, T; L^2(\omega))\}$$

and $W^{1,2,\mathcal{M}}(0, T; W^{1,3/2}(\omega), H^{-2}(\omega))$ are compactly embedded in $L^2(0, T; L^2(\omega))$, for any smooth bounded subset ω of Ω. The result is that there exists $(\overline{\mathbf{u}}, \overline{g}, \overline{\mathcal{E}}, \overline{\mathcal{N}})$ such that along a subsequence $\tau \searrow 0$ one has the same convergence as in (139)-(141) and (143). The convergence in (142) and (144) is now replaced by weak$-*$ convergence in $\mathcal{M}(0, T; L^2(\Omega))$ for $\partial_t \overline{\mathcal{E}}_\tau$ and in $\mathcal{M}(0, T; H^{-2}(\Omega))$ for $\partial_t \overline{\mathcal{N}}_\tau$.

Furthermore, the estimates allow us to conclude

$$\overline{\mathcal{E}}_\tau \to \overline{\mathcal{E}} \quad \text{strongly in } L^2((0,T) \times \omega) \text{ and (a.e) on } (0,T) \times \omega, \tag{145}$$

$$\overline{\mathcal{N}}_\tau \to \overline{\mathcal{N}} \quad \text{strongly in } L^2((0,T) \times \omega) \text{ and (a.e) on } (0,T) \times \omega. \tag{146}$$

As a consequence

$$\rho(\overline{g}_\tau) = \frac{\overline{\mathcal{N}}_\tau}{\overline{n}(\overline{\mathcal{E}}_\tau)} \to \frac{\overline{\mathcal{N}}}{\overline{n}(\overline{\mathcal{E}})} \tag{147}$$

and

$$\overline{g}_\tau = \rho^{-1}\Big(\frac{\overline{\mathcal{N}}_\tau}{\overline{n}(\overline{\mathcal{E}}_\tau)}\Big) \to \rho^{-1}\Big(\frac{\overline{\mathcal{N}}}{\overline{n}(\overline{\mathcal{E}})}\Big) = \overline{g}. \tag{148}$$

strongly in $L^2((0,T) \times \omega)$ and a.e. on $(0,T) \times \omega$. This in turn implies

$$\begin{cases} \rho(\overline{g}_\tau) \to \rho(\overline{g}) \quad \text{strongly in } L^2((0,T) \times \omega) \text{ and (a.e) on } (0,T) \times \omega; \\ \mathcal{D}(\overline{g}_\tau) \to \mathcal{D}(\overline{g}) \quad \text{strongly in } L^2((0,T) \times \omega) \text{ and (a.e) on } (0,T) \times \omega. \end{cases} \tag{149}$$

Inherited from $\overline{\mathcal{E}}_\tau = \operatorname{div} \overline{\mathbf{u}}_\tau$, the convergence properties imply

$$\overline{\mathcal{E}} = \operatorname{div} \overline{\mathbf{u}} \quad \text{a.e. in} \quad (0,T) \times \Omega. \tag{150}$$

As in [[30], pages 224-226] one shows that $\tilde{\mathbf{u}} = \overline{\mathbf{u}}$ and $\tilde{g} = \overline{g}$. Then (150) implies that $\overline{\mathcal{E}} = \tilde{\mathcal{E}}$. Alternatively, this follows from estimate (135) which gives

$$\int_0^T ||\overline{\mathcal{E}}_\tau(t) - \tilde{\mathcal{E}}_\tau(t)||^2_{L^2(\Omega)} \, dt = \frac{\tau^3}{2} \sum_{j=1}^{N-1} ||\mathcal{E}^j - \mathcal{E}^{j+1}||^2_{L^2(\Omega)} = C\tau^2 ||\partial_t \tilde{\mathcal{E}}_\tau||^2_{L^2(0,T;L^2(\Omega))} = C\tau^2. \tag{151}$$

Similarly,

$$\int_0^T ||\overline{\mathcal{N}}_\tau(t) - \tilde{\mathcal{N}}_\tau(t)||^2_{H^{-2}(\Omega)} \, dt = C\tau^2, \tag{152}$$

which yields $\overline{\mathcal{N}} = \tilde{\mathcal{N}}$.

From this point on we denote the limit, as $\tau \searrow 0$, by the quadruple $(\mathbf{u}, g, \mathcal{E}, \mathcal{N})$, where

$$\mathcal{E} = \operatorname{div} \mathbf{u} \quad \text{and} \quad \mathcal{N} = \overline{n}(\mathcal{E})\rho(g).$$

We are now in a position to prove the main existence result for a weak solution of the time continuous case.

Theorem 12. *Let $\lambda^* > 0$. Then there exists at least one* **weak free energy solution** *$(\mathbf{u}, \mathcal{E}, \rho)$ satisfying Definition 5.*

Proof. In the proof we use approximations (128) and (129), and their convergence properties.

Let $\tau > 0$, sufficiently small, and let $t \in (\tau, T)$. Then $t_j \le t < t_{j+1}$ for some $j \in \{1, \dots, N-1\}$ and $\overline{\mathbf{u}}_\tau(t) = \mathbf{u}^j$ and $\overline{g}_\tau(t) = g^j$.

We first consider the momentum balance equation (81).

Starting point is problem **(PD)**. Using equation (109) we have for any $\xi \in H_0^1(\Omega)^m$

$$\int_\Omega \mathcal{G}e(\overline{\mathbf{u}}_\tau) : e(\xi) \, dx = \int_\Omega \mathcal{G}e(\mathbf{u}^j) : e(\xi) \, dx = -\frac{\lambda^*}{\tau} \int_\Omega (\mathcal{E}^j - \mathcal{E}^{j-1}) \operatorname{div} \xi \, dx =$$

$$\int_\Omega p(g^j) \operatorname{div} \xi \, dx + \int_\Omega \mathbf{F}^j \cdot \xi \, dx = -\lambda^* \int_\Omega \partial_t \tilde{\mathcal{E}}_\tau(t - \tau) \operatorname{div} \xi \, dx + \int_\Omega p(\overline{g}_\tau) \operatorname{div} \xi \, dx + \int_\Omega \mathbf{F}_\tau \cdot \xi \, dx \tag{153}$$

Here we introduced
$$\mathbf{F}_\tau(t) = \mathbf{F}(t_j) = \mathbf{F}^j \quad \text{for} \quad t_j \le t < t_{j+1} \quad \text{and} \quad j = 0, \dots, N-1.$$

Multiplying equation (153) by $\alpha \in C_0^\infty(0, T)$ and integrating the result over (τ, T), yields

$$\int_\tau^T \{\int_\Omega \mathcal{G}e(\overline{\mathbf{u}}_\tau) : e(\xi)\,dx\}\alpha(t)\,dt + \lambda^* \int_\tau^T \{\int_\Omega \partial_t \tilde{\mathcal{E}}_\tau(t-\tau)\,\text{div}\,\xi\,dx\}\alpha(t)\,dt +$$

$$\int_\tau^T \{\int_\Omega p(\overline{g}_\tau)\text{div}\,\xi\,dx\}\alpha(t)\,dt = \int_\tau^T \{\int_\Omega \mathbf{F}_\tau \cdot \xi\,dx\}\alpha(t)\,dt. \tag{154}$$

Next we send $\tau \searrow 0$ along the appropriate subsequence to have convergence of the terms containing $\mathbf{u}$, $\mathcal{E}$ and $\mathbf{F}$. What remains is the pressure term. We recall that $p(g)$ is the composite function $(p \circ \rho)(g)$, where $p(\rho)$ is given by (37) and $\rho(g)$ is defined through (54) and (56). Since $\overline{g}_\tau \to g$ strongly in $L^2((0, T) \times \omega)$, see (148), we have similarly

$$p(\overline{g}_\tau) = (p \circ \rho)(\overline{g}_\tau) \to (p \circ \rho)(\overline{g}) = (p \circ \rho)(g) = p(g)$$

strongly in $\quad L^2((0, T) \times \omega) \quad$ and a.e. in $\quad (0, T) \times \omega$.

This concludes the first part of the proof.

Next we tackle the mass balance equation (70).
We first put equation (78) in the form

$$\partial_t \mathcal{N} - \partial_t \overline{n}(\mathcal{E})\rho(g) + D(g, \mathcal{E})\partial_t \mathcal{E} - \text{div}\left(k(\mathcal{E})D(g)\mathbb{K}\nabla g\right) = Q$$

and apply the discretization of problem **(PD)**. Similar to (154) this gives for any $\psi \in C_0^\infty(\Omega)$ and $\alpha \in C^\infty[0, T]$

$$\int_\tau^T \int_\Omega \left(\partial_t \tilde{\mathcal{N}}_\tau(t-\tau) - \partial_t \tilde{v}_\tau(t-\tau)\left(\rho(\overline{g}_\tau) - \frac{G(\overline{g}_\tau)}{\overline{g}_\tau}\right) + \partial_t \tilde{\mathcal{E}}_\tau(t-\tau)\frac{p(\overline{g}_\tau) - p_0}{\overline{g}_\tau}\right)\psi(x)\alpha(t)\,dxdt +$$

$$\int_\tau^T \int_\Omega k(\overline{\mathcal{E}}_\tau)D(\overline{g}_\tau)\mathbb{K}\nabla\overline{g}_\tau \cdot \nabla\psi(x)\alpha(t)\,dxdt = \int_\tau^T \int_\Omega Q_\tau\psi(x)\alpha(t)\,dxdt, \tag{155}$$

where

$$\tilde{v}_\tau(t) = \left(j + 1 - \frac{t}{\tau}\right)\overline{n}(\mathcal{E}^j) + \left(\frac{t}{\tau} - j\right)\overline{n}(\mathcal{E}_{j+1}),$$

and

$$Q_\tau(t) = Q(t_j) = Q^j$$

for $j\tau \le t < (j+1)\tau$. \hfill (156)

The boundedness of $\overline{n}'$ implies

$$||\partial_t \tilde{v}_\tau||_{L^2((0,T) \times \Omega)} \le C \tag{157}$$

and inherited from (132)

$$||\tilde{v}_\tau||_{L^2(0,T;H^1(\omega)} \le C. \tag{158}$$

Hence

$$\partial_t \tilde{v}_\tau \rightharpoonup \partial_t \overline{n}(\mathcal{E}) \quad \text{weakly in} \quad L^2((0, T) \times \Omega) \tag{159}$$

and

$$\tilde{v}_\tau \to \overline{n}(\mathcal{E}) \quad \text{strongly in} \quad L^2((0, T) \times \omega) \text{ and a.e. in} \quad (0, T) \times \omega. \tag{160}$$

We are now in position to pass to the limit $\tau \searrow 0$ in (155) and obtain

$$\int_0^T \langle \partial_t(\overline{n}(\mathcal{E})\rho(g)), \psi\rangle_{H^{-2}(\Omega), H_0^2(\Omega)}\alpha(t)\, dt - \int_0^T \int_\Omega \left(\partial_t\overline{n}(\mathcal{E})\big(\rho(g) - \frac{G(g)}{g}\big) + \partial_t\mathcal{E}\frac{p(g) - p_0}{g}\right)\psi(x)\alpha(t)\, dxdt+$$

$$\int_0^T \int_\Omega k(\mathcal{E})D(g)\mathbb{K}\nabla\overline{g} \cdot \nabla\psi(x)\alpha(t)\, dxdt = \int_0^T \int_\Omega Q\psi(x)\alpha(t)\, dxdt, \tag{161}$$

or

$$\partial_t\left(\overline{n}(\mathcal{E})\rho(g)\right) - \partial_t\overline{n}(\mathcal{E})\rho(g) + D(\rho, \mathcal{E})\partial_t\mathcal{E} - \operatorname{div}\left(k(\mathcal{E})D(g)\mathbb{K}\nabla g\right) = Q \text{ in } \mathcal{D}'((0,T) \times \Omega), \tag{162}$$

It remains to check the initial and boundary conditions and the energy inequality (73).
First we notice that (141)-(142) imply

$$\tilde{\mathcal{E}}_\tau \rightharpoonup \mathcal{E} \quad \text{weakly in} \quad W^{1,2,2}(0,T; H^1(\omega), L^2(\omega)), \tag{163}$$

where $W^{1,2,2}(0,T; H^1(\omega), L^2(\omega)) = \{z \in L^2(0,T; H^1(\omega)) \mid \partial_t z \in L^2(0,T; L^2(\omega))\}$. In this space the trace in time $\mathcal{E} \to \mathcal{E}(0)$ is a weakly continuous map from $W^{1,2,2}(0,T; H^1(\omega), L^2(\omega))$ to $L^2(\omega)$. Hence

$$\tilde{\mathcal{E}}_\tau(0) \rightharpoonup \mathcal{E}(0) \quad \text{weakly in} \quad L^2(\omega), \tag{164}$$

where $\mathcal{E}(0) = \operatorname{div} \mathbf{u}^0 = 0$.
    Next, using (143), (144) and (146), we conclude that

$$\tilde{\mathcal{N}}_\tau(0) \rightharpoonup \mathcal{N}(0) \text{ weakly in } H^{-2}(\omega), \tag{165}$$

which justifies the initial condition for $\mathcal{N}$. Since $\mathcal{N} = \overline{n}(\mathcal{E})\rho(g)$, we have simultaneously the initial conditions for the density ρ and for g.
    We still miss the flux boundary condition for the mass balance equation (68). Starting point is again equation (155), now with $\psi \in H^1(\Omega)$ and $\alpha(T) = 0$. Since

$$\int_\tau^T \int_\Omega \partial_t\tilde{\mathcal{N}}_\tau(x, t - \tau)\psi(x)\alpha(t)\, dxdt = -\int_\tau^T \int_\Omega \tilde{\mathcal{N}}_\tau(x, t - \tau)\psi(x)\frac{d}{dt}\alpha(t)\, dxdt-$$

$$\int_\Omega \tilde{\mathcal{N}}_\tau(x, 0)\psi(x)\alpha(\tau)\, dx \to -\int_0^T \int_\Omega \tilde{\mathcal{N}}(x, t)\psi(x)\frac{d}{dt}\alpha(t)\, dxdt - \alpha(0)\int_\Omega \overline{n}(0)\rho^0(x)\psi(x)\, dx$$

and since the strong convergence (145)-(146), together with the weak convergence (141) and (143), implies the same for Ω, we may pass to the limit $\tau \searrow 0$ and conclude that

$$-\int_0^T \int_\Omega \rho(g)\overline{n}(\mathcal{E})\partial_t\Phi(x, t)\, dxdt - \int_\Omega \overline{n}(0)\rho^0(x)\Phi(x, 0)\, dx + \int_0^T \int_\Omega \partial_t\mathcal{E}\Big(D(\rho(g), \mathcal{E})-$$

$$\rho(g)\overline{n}'(\mathcal{E})\Big)\Phi(x, t)\, dxdt + \int_0^T \int_\Omega k(\mathcal{E})D(g)\mathbb{K}\nabla g \cdot \nabla_x\Phi(x, t)\, dxdt$$

$$= \int_0^T \int_\Omega Q\Phi(x, t)\, dxdt, \quad \forall\Phi \in H^1(\Omega \times (0, T)), \text{ with } \Phi|_{t=T} = 0.$$

Inequality (73) is a direct consequence of Theorem 8 and the weak lower semi-continuity of the gradient terms in $L^2(Q_T)$. $\qquad \square$

5. Discussion and conclusion

    In this paper we study a model that describes the quasi-static mechanical behaviour of a fluid saturated porous medium. In it simplest (linear) form it is described by equations (1)-(4), where (1) results from the fluid phase mass balance in the case that the fluid is incompressible.

We follow Rutquist et al [31] and Lewis and Schrefler [18] and propose a fluid mass balance that is based on the mixture theory of Bedford and Drumheller [2], [3]. This yields equation (6) and the resulting nonlinear system is given by (2)-(4) and (6). Note that the time derivative of the fluid density ρ appears in (6), since the fluid is assumed weakly compressible, see expression (10). Models where the fluid density is constant (see [5] and [9]) do not contain this source term. Moreover the porosity n and the deformation of the medium are related through (8). An expression for this relation is derived from the solid phase mass balance. It is given by (19) or, when the deformation is small, by approximation (21).

It is shown by means of a counterexample that the porosity may admit non-physical, i.e. negative, values. This is made precise in Proposition 3. To obtain a well-posed mathematical problem the porosity is modified according to cut-off (36). This cut-off is chosen such that it reduces to the correct expression in the physical range. Outside this range it remains positive. Likewise a cut-off for the density is introduced through expressions (61) and (65).

The momentum balance equation (2)-(4) is modified as well. Following Murad and Cushman [27] we add the term

$$\lambda^* \text{div} \, \partial_t \mathbf{u} \quad (\lambda^* \geq 0) \tag{166}$$

to the expression for the total stress. This result in expression (5). Murad and Cushman give a thermodynamically based derivation of the equation in which (166) appears as the difference between the fluid and solid pressures. Having $\lambda^* > 0$, (166) acts as a time regularization of the volumetric stress for our quasi-static problem.

An important role in the analysis of the equations is played by the free energy of the system. This free energy acts as a Lyapunov functional. It is given by (67), which generalizes Biot's original expression developed for the linear case [4]. In the case that the deformation and fluid density are in the physical range, the free energy simplifies to, see also (57),

$$J(\mathbf{u}, \rho) = \int_\Omega \left(\frac{1}{2} \mathcal{G}e(\mathbf{u}) : e(\mathbf{u}) - \mathbf{F} \cdot \mathbf{u} + \frac{n(\text{div } \mathbf{u})}{\beta_0 n_0 (1-n_0)\rho_0} \left((1-n_0)\rho - \rho_0^{n_0}\rho^{1-n_0} + n_0\rho_0 \right) \right) dx. \tag{167}$$

We introduce a weak formulation and prove existence of a solution in a number of steps. Discretizing in time, we first consider the incremental equations. Using Brézis' fundamental theorem for pseudo monotone operators, see for instance Lions [19] and Roubiček [30], we obtain existence for the corresponding incremental problem. The result holds for any $\lambda^* \geq 0$. Moreover, using the free energy, estimates that are global in time are derived. These (stability) estimates are crucial when considering the time continuous, quasi-static, formulation for which we prove existence at the expense of having $\lambda^* > 0$. The free energy implies global stability of the solution.

We note that only in the proof of the local $H^1(\Omega)-$ estimates for div $\mathbf{u}$, we use the fact that the Gassmann tensor has the specific form of Hooke's law (4). In the incremental problem we could have replaced $\mathcal{G}$ by a general rank-4, symmetric, positive-definite Gassmann tensor.

Some particular cases of system (24)-(25) were studied before. An interesting example is the consolidation with an irrotational composite flow rate, when the system reduces to a scalar pseudo-parabolic PDE. For details see [17].

We notice also that the model studied in this paper was extensively used by Schrefler et al, see [33] and [18] and references therein. It is broadly accepted in the computational poromechanics community. A review of different numerical methods and software is given in [31] and [26].

References

[1] Bear J, Bachmat Y. *Introduction to Modeling of Transport Phenomena in Porous Media*, Springer Science and Business Media, 1990.

[2] Bedford A, Drumheller DS. A variational theory of immiscible mixtures. *Arch. Ration. Mech. Anal.*, 1978; 68: 37-51.

[3] Bedford A, Drumheller DS. Theories of immiscible and structured mixtures. *Internat. J Engrg Sci.*, 1983; 21: 863-960.

[4] Biot MA. Mechanics of Deformation and Acoustic Propagation in Porous Media. *J. Appl. Phys.*, 1962; 33: 1482.

[5] Bociu L, Guidoboni G, Sacco R, Webster JT. Analysis of nonlinear poro-elastic and poro-visco-elastic models. *Arch. Ration. Mech. Anal.*, 2016; 222: 1445–1519.

[6] Bosco E, Peerlings RH, Geers MG. Predicting hygro-elastic properties of paper sheets based on an idealized model of the underlying fibrous network. *Int. J. Solids Struct.*, 2015; 56: 43-52.

[7] Cances C, Choquet C, Fan Y, Pop IS. Existence of weak solutions to a degenerate pseudo-parabolic equation modeling two-phase flow in porous media. *CASA Report*, 2010; 10-75.

[8] Čanić S, Hartley CJ, Rosenstrauch D, Tambača J, Guidoboni G, A. Mikelić. Blood Flow in Compliant Arteries: An Effective Viscoelastic Reduced Model, Numerics and Experimental Validation. *Ann. Biomed. Eng.*, 2006; 34: 575 - 592

[9] Cao Y, Chen S, Meir AJ. Analysis and numerical approximations of equations of nonlinear poroelasticity. *Discrete Contin. Dyn. Syst. Ser. B,* 2013; 18: 1253–1273.

[10] Cao X, Pop IS . Degenerate two-phase porous media flow model with dynamic capillarity. *J. Differ. Equations*, 2016; 260: 2418-2456.

[11] Cardoso L, Fritton SP, Gailani G, Benalla M, Cowin SC. Advances in assessment of bone porosity, permeability, and interstitial fluid flow. *J. Biomech.*, 2013; 46: 253–265.

[12] Coussy O. *Poromechanics*, John Wiley and Sons, 2004.

[13] Cowin SC. Bone Poroelasticity. *J. Biomech.*, 1999; 32: 218-238.

[14] van Duijn CJ, Mikelić A, Wick T. A monolithic phase-field model of a fluid-driven fracture in a nonlinear poroelastic medium, *Math. Mech. Solids*, 2019; 24: 1530–1555.

[15] Evans LC. A survey of entropy methods for partial differential equations. *B. Am. Math. Soc.*, 2004; 41:409–438.

[16] Jüngel A. *Entropy methods for diffusive partial differential equations*. BCAM Springer Briefs: Springer, 2016.

[17] Holland E, Showalter RE. Poro-Visco-Elastic Compaction in Sedimentary Basins. *SIAM J. Math. Anal.*, 2018; 50: 2295-2316.

[18] Lewis R W, Schrefler B A. *The Finite Element Method in the Static and Dynamic Deformation and Consolidation of Porous Media*. John Wiley and Sons, 1998.

[19] Lions JL. *Quelques méthodes de résolution des problémes aux limites non linéaires*. Dunod, Paris, 1969.

[20] Marciniak-Czochra A, Mikelić A. A Rigorous Derivation of the Equations for the Clamped Biot-Kirchhoff-Love Poroelastic Plate. *Arch. Rational Mech. Anal.*, 2015; 215: 1035–1062.

[21] Mei CC, Vernescu B. *Homogenization methods for multiscale mechanics*. World scientific, 2010.

[22] Mikelić A. A global existence result for the equations describing unsaturated flow in porous media with dynamic capillary pressure. *J. Differ. Equations*, 2010; 248: 1561-1577.

[23] Mikelić A, Tambača J. Derivation of a poroelastic flexural shell model, *Multiscale Model. Simul.,* 2016; 14: 364-397.

[24] Mikelić A, Wheeler MF. On the interface law between a deformable porous medium containing a viscous fluid and an elastic body. *Math. Models Methods Appl. Sci.* 2012; 22: 1250031.

[25] Milišić JP. The unsaturated flow in porous media with dynamic capillary pressure. *J. Differ. Equations*, 2018; 264: 5629-5658.

[26] Minkoff SE, Stone CM, Bryant S, Peszynska M, Wheeler MF. Coupled fluid flow and geomechanical deformation modeling. *J. Petrol. Sci. Eng.*, 2003; 38: 37–56.

[27] Murad MA, Cushman JH. Multiscale flow and deformation in hydrophilic swelling porous media. *Internat. J Engrg Sci.*, 1996; 34: 313–338.

[28] Owczarek S. A Galerkin method for Biot consolidation model. *Math. Mech. Solids*, 2010; 15: 42–56.

[29] Prosi M, Zunino P, Perktold K, Quarteroni A. Mathematical and numerical models for transfer of low-density lipoproteins through the arterial walls: a new methodology for the model set up with applications to the study of disturbed lumenal flow. *J. Biomech.*, 2005; 38: 903-917.

[30] Roubiček T, *Nonlinear Partial Differential Equations with Applications*. Springer Science & Business Media, 2005.

[31] Rutqvist J, Börgesson L, Chijimatsu M, Kobayashi A, Jing L, Nguyen TS, Noorishad J, Tsang CF. Thermohydromechanics of partially saturated geological media: governing equations and formulation of four finite element models. *Int. J. Rock Mech. Min. Sci.*, 2001; 38: 105–127.

[32] Sanchez-Palencia E. *Non-Homogeneous Media and Vibration Theory*, Springer Lecture Notes in Physics **129**, 1980.

[33] Schrefler BA, Simoni L, Xikui L, Zienkiewicz OC. Mechanics of partially saturated porous media, in "Numerical Methods and Constitutive Modelling in Geomechanics", D. S. Desai, G. Gioda (Eds.), Courses and Lectures CISM no. 311, Springer New York 1990, 169-211.

[34] Schweizer B, *Partielle Differentialgleichungen*, Springer-Verlag Berlin 2018.

[35] Showalter RE. Diffusion in poro-elastic media. *J. Math. Anal. Appl.*, 2000; 251: 310–340.

[36] Terzaghi K. *Theoretical soil mechanics*. Chapman And Hall, Limited, London, 1951.

[37] Tolstoy I., ed. *Acoustics, elasticity, and thermodynamics of porous media. Twenty-one papers by M.A. Biot*. Acoustical Society of America, New York, 1992.

[38] Verruijt A, Van Baars S. *Soil mechanics*. Delft, the Netherlands: VSSD, 2007.

[39] Zeidler E. *Nonlinear functional analysis and its applications II/B Nonlinear monotone operators*, Springer-Verlag, New York, 1990.

[40] Ženíšek A. The existence and uniqueness theorem in Biot's consolidation theory, *Appl. Math.*, 1984; 29: 194–211.