PULLBACK ATTRACTORS OF REACTION-DIFFUSION INCLUSIONS WITH SPACE-DEPENDENT DELAY

PETER E. KLOEDEN
School of Mathematics & Statistics
Huazhong University of Science & Technology
Wuhan 430074, China

THOMAS LORENZ
Applied Mathematics
RheinMain University of Applied Sciences
65197 Wiesbaden, Germany

Abstract. Inspired by biological phenomena with effects of switching off (maybe just for a while), we investigate non-autonomous reaction-diffusion inclusions whose multi-valued reaction term may depend on the essential supremum over a time interval in the recent past (but) pointwise in space. The focus is on sufficient conditions for the existence of pullback attractors. If the multi-valued reaction term satisfies a form of inclusion principle standard tools for non-autonomous dynamical systems in metric spaces can be applied and provide new results (even) for infinite time intervals of delay. More challenging is the case without assuming such a monotonicity assumption. Then we consider the parabolic differential inclusion with the time interval of delay depending on space and extend the approaches of norm-to-weak semigroups to a purely metric setting. This provides completely new tools for proving pullback attractors of non-autonomous dynamical systems in metric spaces.

1. Introduction. As a qualitative motivation from biology, consider inhomogeneous cell tissue under the influence of a chemical substance. Whenever the concentration $u(t, x)$ of that substance at a position $x$ exceeds a given threshold, the consequences for the cell there are significant, maybe even fatal (like apoptosis). Hence it is not just the present concentration $u(t, x)$ which influences the cells, but also its history – for a while, at least. Neighboring cells have a recovering effect on the tissue at position $x$, and the duration of such a process depends on the cell type, i.e., which position $x$ in the tissue we consider. From the mathematical point of view, the evolution depends on

$\text{ess sup}_{t - \hat{\Theta}(x) \leq s \leq t} u(s, x)$

2010 Mathematics Subject Classification. Primary: 35B41; Secondary: 35R70, 35K20, 35K57, 37L30.

Key words and phrases. Reaction-diffusion equation, parabolic differential inclusion with space-dependent delay, dead core, retarded functional evolution inclusion, existence of solutions, non-autonomous pullback attractors, set-valued dynamical system with norm-to-weak closed graph.

Partially supported by DFG grant KL 1203/7-1, the Spanish Ministerio de Ciencia e Innovación project MTM2011-22411, the Consejería de Innovación, Ciencia y Empresa (Junta de Andalucía) under the Ayuda 2009/FQM314 and the Proyecto de Excelencia P12-FQM-1492.
with a given non-negative function $\Theta(\cdot)$ of space. Strictly speaking, the threshold $c$ should also come into play; i.e., the special interest is on the question whether the threshold $c \geq 0$ has been exceeded in recent history or not. We will express the relevant non-local term by means of a cut-off function

$$\text{ess sup}_{t-\tilde{\Theta}(x) \leq s \leq t} [u(s,x)]_c$$

with $[\eta]_c := \begin{cases} c & \text{if } \eta > c, \\ \eta & \text{if } -c \leq \eta \leq c, \\ -c & \text{if } \eta < -c \end{cases}$ for $\eta \in \mathbb{R}$.

Considering any bounded open domain $\Omega \subset \mathbb{R}^d$ and time period $T \geq 0$, this non-local term is well-defined for any Lebesgue measurable function $u : (-\infty, T] \times \Omega \rightarrow \mathbb{R}$, but it leads to mathematically challenging questions since it clearly differs from standard forms of “delay” or “memory”.

Motivated by this model problem with the biological background, the focus of interest in this paper is the existence of a pullback attractor of the following non-autonomous reaction-diffusion inclusion on a bounded open domain $\Omega \subset \mathbb{R}^d$ with smooth boundary

$$\partial_t u - \Delta u \in G\left(t, x; u(t,x), \text{ess sup}_{t-\tilde{\Theta}(x) \leq s \leq t} [u(s,x)]_c, u(t,\cdot)\right)$$

$$u = u_0 \quad \text{a.e. in } (t_0, T) \times \Omega$$

$$u = 0 \quad \text{on } (t_0, T) \times \partial \Omega$$

In particular, we prefer a differential inclusion to its single-valued counterpart, i.e., a differential equation, in support of the possibility that the concept of relaxation is applied later whenever required in modeling, e.g., when terms are switched on or off.

Obviously, it is a parabolic differential inclusion subject to homogeneous Dirichlet boundary conditions. The “reaction term” on the right-hand side, however, does not depend just on the current state $u(t) \in L^2(\Omega)$ (in the pointwise or functional way) as usual. It is rather the “history” of the desired solution $u : (-\infty, T] \rightarrow L^2(\Omega)$ which comes into play. Hence, this problem can be regarded as a non-autonomous differential inclusion with delay (as, e.g., in \cite{7, 8, 30}) or with memory (as, e.g., in \cite{1, 29}) or as a retarded functional differential inclusion (similarly to, e.g., \cite{22, 40}). (From our point of view, the distinction between “delay” and “memory” is not handled uniformly in the literature about differential equations and inclusions – particularly if initial conditions are not specified.)

The remarkable aspect of problem (1) is the influence of the “history” depending on space, i.e., the duration of the time interval in the past depends on the spatial position $x \in \Omega$ and is assumed to be strictly positive. In more detail, the set-valued reaction term $G$ in inclusion (1) uses the representation

$$\tilde{\Theta}(x) := \theta + \Theta(x)$$

with a fixed constant $\theta > 0$ (indicating a positive lower bound) and a given non-negative function $\Theta : \Omega \rightarrow \mathbb{R}$.

The existence of strong solutions to problem (1) can be concluded directly from our earlier results in \cite{37}, essentially due to the positive lower bound $\theta$. (The details are explained in section 3 below). In this article, we focus on sufficient conditions for a pullback attractor in an appropriate sense of convergence.

This topic is analytically challenging. Indeed, standard tools for establishing the existence of attractors attractors start with the reformulation in terms of dynamical
systems. (In our case of a parabolic differential inclusion, we consider a multi-valued non-autonomous dynamical system as described in section 5 below.) In the framework of dynamical systems, the essential challenge is then to specify a norm or, more generally, a metric on the state space which serves two different purposes simultaneously: First, the dynamical system is appropriately (semi-) continuous w.r.t. initial states. Second, an additional compactness property enables us to draw conclusions about accumulation points whenever time is tending to $\pm \infty$ respectively. Pullback asymptotic compactness belongs to the most popular forms in this regard (see, e.g., [8, 9, 11, 38]).

In the special case of parabolic differential inclusion (1), we can specify a metric on the state space

$$C_\gamma := \left\{ u \in C^0\left((-\infty, 0], \mathbb{L}^2(\Omega)\right) \mid \lim_{s \to -\infty} (u(s) \cdot e^{\gamma s}) \text{ exists in } \mathbb{L}^2(\Omega) \right\}$$

and adapt arguments from [8] which guarantee the existence of a pullback attractor. This result is formulated in Theorem 4.2 below, but it is based on a technical assumption about the set-valued mapping $G : \mathbb{R} \times \Omega \times \mathbb{R} \times \mathbb{R} \times \mathbb{L}^2(\Omega) \rightrightarrows \mathbb{R}$ in the reaction term, which plays the role of a monotonicity condition w.r.t. the retarded argument:

$$z_1 \leq z_2 \implies G(t, x; y, z_1, w) \subset G(t, x; y, z_2, w). \quad (2)$$

(see hypothesis (G7) in section 2 for its complete formulation).

From our current point of view, there is no transparent way how to apply the “standard method” for pullback attractors of dynamical systems in metric spaces to parabolic inclusion problem (1) if we dispense with this special inclusion assumption (2). This is essentially due to the fact that the solution $u : (-\infty, T] \to \mathbb{L}^2(\Omega)$ enters the reaction term at time $t \in (-\infty, T]$ in form of

$$\Omega \to [-c, c], \quad x \mapsto \text{ess sup}_{t - \Theta(x) \leq s \leq t} [u(s, x)]_c.$$

Apparently, the latter function is bounded on the bounded domain $\Omega \subset \mathbb{R}^d$ and so, it also belongs to $\mathbb{L}^2(\Omega)$. But it is not obvious at all that its $\mathbb{L}^2(\Omega)$ norm depends continuously on one of the popular norms of $u(\cdot + T) \in C_\gamma \subset C^0\left((-\infty, 0], \mathbb{L}^2(\Omega)\right)$ (i.e., which are well established in the literature like the weighted supremum norm $\|\cdot\|_\gamma$ in [8] below, see, e.g., [8, 30]). Even under additional assumptions for $\Theta \in L^1(\Omega, \mathbb{R}^+)$, there can always be some positions $x \in \Omega$ in which the delay reaches infinitely long in the past. Hence on the one hand, the standard approaches to norms like exponentially decaying weights are to fail when verifying an appropriate form of (semi-) continuity of the dynamical system. On the other hand, we do not see how to draw conclusions about pullback asymptotic compactness if the metric of functions considers delay which is just integrably bounded in space.

Due to this observation, we consider problem (1) without monotonicity assumption (2) as a touchstone for a new approach to proving pullback attractors. This metric concept is the main contribution of our article to the general theory of non-autonomous dynamical systems.

Indeed, Zhong et al. have published interesting proofs for the existence of pullback attractors by supplying domain and value space of dynamical systems with different topologies. In [60], for example, so-called norm-to-weak continuous semigroups on a Banach space are used for investigating nonlinear parabolic differential...
equations. This idea has then been extended to non-autonomous dynamical systems or processes (see, e.g., [41, 52, 57, 55]). The joint key idea is assuming the condition on the dynamical system that for at each time instant, a norm converging sequence in the domain state space is always mapped to a (just) weakly convergent sequence of values in the same state space. This modification of topologies facilitates verifying the general conditions in concrete applications like reaction-diffusion equations.

We have borrowed this idea of different topologies in domain and value space, but (re-) interpret it in a purely metric setting. In particular, we do not use the norm or weak convergence in a Banach space in the standard meaning of functional analysis based on linear forms.

Indeed, the well-known consequence of Hahn-Banach theorem

$$\|\xi\|_Y = \sup \{ \ell(\xi) \mid \ell : Y \to \mathbb{R} \text{ linear and continuous, } \|\ell\|_{\text{Lin}(Y,\mathbb{R})} \leq 1 \}.$$ 

for any Banach space \((Y, \|\cdot\|_Y)\). Corollary 1.4 reveals a relationship between weak and norm convergence which we use as a starting point in a purely metric setting. Each continuous linear functional \(\ell : Y \to \mathbb{R}\) induces a “distance function” quantifying how “far” two vectors are away from each other:

$$Y \times Y \to \mathbb{R}, \quad (\xi, \zeta) \mapsto |\ell(\xi - \zeta)|$$

This is a pseudo-metric on \(Y\), i.e., it is reflexive, symmetric and satisfies the triangle inequality (see, e.g., [31]). But in general, it is not positive definite and so, it is not a metric. The kind of a “distance function”, however, leads to the metric induced by the norm

$$Y \times Y \to \mathbb{R}, \quad (\xi, \zeta) \mapsto \|\xi - \zeta\|_Y$$

if we take the pointwise supremum for all continuous linear functionals \(\ell : Y \to \mathbb{R}\) with \(\|\ell\|_{\text{Lin}(Y,\mathbb{R})} \leq 1\). Shortly speaking, these linear functionals \(\ell : Y \to \mathbb{R}\) with \(\|\ell\|_{\text{Lin}(Y,\mathbb{R})} \leq 1\) can be regarded as a parameter for the family of “distance functions”. Such a connection via pointwise supremum is now considered in a purely metric setting: Consider a nonempty set \(X\) and a family \((d_{X,j})_{j \in J}\) of functions \(X \times X \to [0, \infty)\) whose parameter \(j\) always belongs to a nonempty set \(J\). Each \(d_{X,j}\) \((j \in J)\) represents a form of quantifying distances between two elements of \(X\), but we do not assume that \(d_{X,j}\) satisfies all standard conditions on a metric. (This is what we just call a “distance function” in the following.) For subsequent conclusions about pullback attractors, however, the pointwise supremum of this family

$$d_X : X \times X \to [0, \infty), \quad (x_1, x_2) \mapsto \sup_{j \in J} d_{X,j}(x_1, x_2) =: d_X(x_1, x_2)$$

is supposed to be a metric on \(X\). For any sequence \((x_n)_{n \in \mathbb{N}}\) in \(X\) and an element \(x \in X\), the convergence w.r.t. the metric \(d_X\) is usually related to the equivalent conditions

$$\lim_{n \to \infty} d_X(x_n, x) = 0$$

$$\iff \forall \varepsilon > 0 \ \exists N = N(\varepsilon) \in \mathbb{N} \ \forall n \geq N : \ d_X(x_n, x) \leq \varepsilon$$

$$\iff \forall \varepsilon > 0 \ \exists N = N(\varepsilon) \in \mathbb{N} \ \forall n \geq N : \ \sup_{j \in J} d_{X,j}(x_n, x) \leq \varepsilon$$

$$\iff \forall \varepsilon > 0 \ \exists N = N(\varepsilon) \in \mathbb{N} \ \forall n \geq N : \ \forall j \in J : \ d_{X,j}(x_n, x) \leq \varepsilon.$$

This form of convergence can be regarded as “uniform with respect to the parameter \(j \in J\)”. Alternatively, we are also free to choose the criterion
\[ \forall j \in J : \lim_{n \to \infty} d_{X,j}(x_n, x) = 0 \]
\[ \iff \forall j \in J : \forall \varepsilon > 0 \ \exists N = N(j, \varepsilon) \in \mathbb{N} \ \forall n \geq N : d_{X,j}(x_n, x) \leq \varepsilon \]
and call it “pointwise with respect to the parameter \( j \)”. In general, these two criteria of convergence do not have to be equivalent, of course. Returning to the example of weak convergence in a Banach space \((Y, \| \cdot \|_Y)\) for a moment, the standard choice of parameters is

\[ J := \{ \ell : Y \to \mathbb{R} \mid \ell \text{ linear and continuous, } \|\ell\|_{\text{Lin}(Y, \mathbb{R})} \leq 1 \} \]

and, the related distance function is \( d_{Y,\ell}(:, :) = |\ell(x) - \ell(y)| \) for \( \ell \in J \). Then the convergence “pointwise in parameter” coincides with weak convergence (in the established sense of functional analysis). Moreover, \( d_Y \) is the metric induced by the norm of \( Y \) due to Hahn-Banach theorem and so, the convergence “uniform w.r.t. the parameter” coincides with norm convergence in \( Y \).

In connection with our parabolic problem \([1]\), the essential idea is to make a completely differential choice of parameters, which is not related to continuous linear functionals at all. (Hence, the convergence “pointwise w.r.t. the parameter” will not have anything to do with weak convergence in a Banach space.) Now the index \( j \in \mathbb{N} \) indicates the time interval \([-j, 0]\) how long the history is considered for the convergence criterion (see section 5.1 for more details), i.e., we will use the family

\[ C_\gamma \times C_\gamma \rightarrow [0, \infty), \]

\[ (u, v) \mapsto \sup_{s \leq 0} \left\{ e^{\gamma s} \|u(s) - v(s)\|_{L^2(\Omega)} \right\} + \left\| \text{ess sup}_{j \leq s \leq 0} \left[ u(s, \cdot) \right]_c - \left[ v(s, \cdot) \right]_c \right\|_{L^2(\Omega)} \]

\((j \in \mathbb{N})\) and its metric related to infinite delay

\[ C_\gamma \times C_\gamma \rightarrow [0, \infty), \]

\[ (u, v) \mapsto \sup_{s \leq 0} \left\{ e^{\gamma s} \|u(s) - v(s)\|_{L^2(\Omega)} \right\} + \left\| \text{ess sup}_{s \leq 0} \left[ u(s, \cdot) \right]_c - \left[ v(s, \cdot) \right]_c \right\|_{L^2(\Omega)} \]

A closer look reveals that their second terms can be written equivalently as

\[ \left\| \text{ess sup}_{j \leq s \leq 0} \left[ u(s, \cdot) \right]_c - \left[ v(s, \cdot) \right]_c \right\|_{L^2(\Omega)} = \left\| [u]_c - [v]_c \right\|_{L^2(\Omega; L^\infty([-j, 0]))} , \]

\[ \left\| \text{ess sup}_{s \leq 0} \left[ u(s, \cdot) \right]_c - \left[ v(s, \cdot) \right]_c \right\|_{L^2(\Omega)} = \left\| [u]_c - [v]_c \right\|_{L^2(\Omega; L^\infty(\mathbb{R}, [-\infty, 0]))} \]

and so, they can clearly differ from the standard norms in \( L^\infty([-j, 0]; L^2(\Omega)) \), \( L^\infty((-\infty, 0]; L^2(\Omega)) \) respectively. Our choice of the second term is motivated by the special structure of the fourth argument of \( G \) in parabolic inclusion \([1]\). The joint first term, i.e., the weighted supremum metric for all \( s \leq 0 \), induces a topology on \( C_\gamma \subset C^0((-\infty, 0]; L^2(\Omega)) \) which is finer than the compact-open topology, but there is no obvious relationship with the convergence in \( L^2(\Omega; L^\infty([-j, 0])) \), \( L^2(\Omega; L^\infty((-\infty, 0])) \) respectively.

Preferring such an alternative family \((d_{X,j})_{j \in J}\), however, implies that some standard tools from linear functional analysis require a counterpart in the new metric setting.

The Lemma of Mazur has proved to be a very useful implication from weak to norm convergence \([57] \ S. V.1 \text{ Theorem } 2\). Indeed, for any weakly converging sequence in a Banach space, there exists another sequence converging strongly to the same limit and, additionally, such a strongly converging sequence can be represented
in terms of the original sequence by means of convex combinations. In other words, we can obtain strong convergence to the same limit if we are willing to consider a possibly different sequence constructed by means of convex combinations. This gist can be extended to the purely metric setting of a nonempty set $X$ and the family $(d_{X,j})_{j \in J}$: We start with a sequence converging to an element $x \in X$ pointwise w.r.t. the parameter $j \in J$, but for conclusions about continuity later on, we would like to represent the same element $x \in X$ as the limit of a (possibly different) sequence converging uniformly w.r.t. the parameter $j \in J$. In the purely metric setting, however, we should feel free to rely on construction approaches different from convex combinations, which is established in vector spaces only. In connection with parabolic inclusion \ref{1}, for example, we will use a form of concatenation in the past in which the given limit function in $C_\gamma$ is involved explicitly (see Lemma 8.26 below for more details).

Strictly speaking, our focus of interest is on pullback attractors and so, we do not need a metric counterpart of Mazur’s Lemma for any sequence converging pointwise w.r.t. the parameter, but just for those sequences that occur in the construction of pullback $\omega$-limit points, i.e., they are related to a sequence of initial time instants tending to $-\infty$. This is an additional feature which can prove to be very useful for constructing alternative sequences whenever standard results from functional analysis cannot be applied to the respective convergence or if standard constructions such as convex combinations are no longer available in a purely metric setting.

Thus, we suggest a counterpart criterion for sequences tending to pullback $\omega$-limit points as a definition in the purely metric setting of $(d_{X,j})_{j \in J}$ and its supremum $d_X$ on a nonempty set $X$ (see Definition 6.7 below). This condition, however, has to be verified for each investigated metric example individually. Indeed, even if $X$ is a Banach space and if $d_X$ is induced by its norm, then we cannot just apply the Lemma of Mazur (in its standard form) unless the family $(d_{X,j})_{j \in J}$ metrizes the weak topology on $X$. But this will not be the case in our example related to parabolic problem \ref{1} (the key steps of verifying the metric counterpart are presented in Lemma 8.26 below, which is essentially based on concatenating solutions appropriately).

In this article, we develop an existence theory for pullback attractors in a purely metric setting using essentially the assumption that the underlying multi-valued non-autonomous dynamical system (MNDS) is sequentially continuous in (possibly) different senses of convergence in the domain and the value space respectively. This concept is then used for proving sufficient conditions on the parabolic differential inclusion \ref{1}, i.e., Theorems 4.1 and 4.2 below, cover the following examples:

Example 1.1.

\begin{enumerate}
  \item $\partial_t u - \Delta u \in \Lambda(t) \cdot \left( u(t, x) \cdot [-\varepsilon, \varepsilon] \right) + \left[ -\varepsilon, \varepsilon \right] \left( \frac{\text{ess sup}_{t-\theta - \Theta(x)} \leq s \leq t} \text{ess sup}_{t-\theta - \Theta(x)} \leq s \leq t} \left[ u(s, x) \right]_{[c]} \right), \varepsilon \right]$
  \item $\partial_t u - \Delta u \in e^{-\varepsilon^2} \frac{u(t, x)^2}{1 + u(t, x)^2} \cdot [-1, 1] + \frac{1}{\sqrt{1 + \varepsilon^2}} \text{ess sup}_{t-\theta - \Theta(x)} \leq s \leq t} \left[ u(s, x) \right]_{[c]} \cdot [-1, 0]$
  \item $\partial_t u - \Delta u \in e^{-|t|} \left( \text{ess sup}_{t-\theta - \Theta(x)} \leq s \leq t} \left[ u(s, x) \right]_{[c]} + \frac{1}{1 + \| u(t) \|_{L^2(\Omega)}} \int_{\Omega} \psi(y) u(t, y) dy \cdot [1, 2] \right)$
\end{enumerate}

with some $\psi \in C^0_\infty(\Omega), \Lambda \in L^2(\mathbb{R})$. 
Example 1.2.

a) $\partial_t u(t, x) - \Delta u(t, x) \in u(t, x) \cdot [-\varepsilon, \varepsilon] + \left[ -c - 1, \frac{\text{ess sup}_{s \leq t} [u(s, x)]_c}{2c + \text{ess sup}_{s \leq t} [u(s, x)]_c} \right].$

b) $\partial_t u(t, x) - \Delta u(t, x) \in \frac{u(t, x)^2}{1 + u(t, x)^2} \cdot [-1, 1] + \left[ -c, \text{ess sup}_{s \leq t} [u(s, x)]_c \right].$

These examples indicate the basic difference between the two theorems: If the set-valued mapping $G$ in the reaction term satisfies the inclusion principle (2) in addition, then we can even draw conclusions about the pullback attractor for examples with infinite delay (i.e., $\theta = \infty$ and so, $\Theta \geq 0$ arbitrary). Whenever not supposing such a monotonicity property of $G$, we need a form of decaying influence of the past or “vanishing” memory – for technical reasons in connection with asymptotic compactness. This is represented by the non-negative Lebesgue integrable function $\Theta$ of space.

The article has the following structure. In section 2, we list all assumptions used for differential inclusion (1) later. This concerns the set-valued mapping $G$ of the reaction term in particular. Section 3 summarizes the aspects concerning the existence of (both strong and mild) solutions to (1) and, we briefly sketch how they result from our earlier arguments in [37]. These solutions induce a multi-valued non-autonomous dynamical system (MNDS) on a function space and, here we aim at its pullback attractors. Section 4 provides two selections of assumptions from section 2 each of which is sufficient for the existence of a pullback attractor. In particular, Theorems 4.1 and 4.2 specify the sense of convergence to which their attraction refers. Section 5 suggests how to reformulate the parabolic differential problem (1) as an abstract evolution inclusion in a metric vector space. Its metric proves to be the pointwise supremum of a family of “distance functions”. This metric setting is the starting point for investigating pullback attractors in section 6. There we extend the concept of so-called norm-to-weak processes by Zhong et al. beyond Banach spaces.

For the sake of transparency, all proofs required in this article are collected in the last sections. We have aimed at a self-contained and detailed way of presentation. Section 7 provides all the proofs of results about attractors formulated in section 6. Finally section 8 consists of all the remaining proofs of statements in sections 3–5. In other words, we there verify the properties step by step required for applying our general metric theory (from section 6) to the original parabolic differential problem (1).

2. General hypotheses about parabolic differential inclusion. Let $\Omega \subset \mathbb{R}^d$ be a bounded open domain with smooth boundary. Suppose $c \geq 0$, $\theta \in (0, \infty]$ and $\Theta \in L^1(\Omega)$ with $\Theta \geq 0$. Define the auxiliary cut-off function $[\cdot]_c : \mathbb{R} \to \mathbb{R}$ by

$$\eta \mapsto [\eta]_c := \begin{cases} c & \text{if } \eta > c, \\ \eta & \text{if } -c \leq \eta \leq c, \\ -c & \text{if } \eta < -c. \end{cases}$$

We focus on the parabolic differential inclusion (1) with space-dependent delay and homogeneous Dirichlet boundary conditions, i.e.,
\[ \partial_t u - \Delta u \in G\left( t, x; u(t, x), \text{ess sup}_{t - \theta - \Theta(x) \leq s \leq t} [u(s, x)]_{c}, u(t, \cdot) \right) \]
\[ u = u_0 \quad \text{a.e. in } (t_0, T) \times \Omega \]
\[ u = 0 \quad \text{on } (t_0, T) \times \partial \Omega \]

where the set-valued reaction term \( G \) is assumed to fulfill some of the following assumptions:

**Hypotheses 2.1.** The set-valued mapping \( G : \mathbb{R} \times \Omega \times \mathbb{R} \times \mathbb{R} \times L^2(\Omega) \rightharbra \mathbb{R} \) satisfies:

(G1) For every \( t \in \mathbb{R} \), \( v_0 \in L^2(\Omega) \), the set-valued map \( G(t, \cdot; \cdot, \cdot, v_0) : \Omega \times \mathbb{R} \rightharbra \mathbb{R} \) has nonempty compact intervals as values and, its graph is a Lebesgue measurable subset of \( \Omega \times \mathbb{R}^2 \times \mathbb{R} \).

(G2) For every \( t \in \mathbb{R} \), \( x \in \Omega \) and \( v_0 \in L^2(\Omega) \), the set-valued map \( G(t, x; \cdot, \cdot, v_0) : \mathbb{R}^2 \rightharbra \mathbb{R} \) is Hausdorff upper semi-continuous, i.e., for every \( (y_0, z_0) \in \mathbb{R}^2 \) and \( \varepsilon > 0 \), there exists a radius \( \delta = \delta(t, x, v_0) > 0 \) with
\[
G(t, x; y, z, v_0) \subset B_\varepsilon(G(t, x; y_0, z_0, v_0)) \quad \text{def.} \quad \{ \sigma_1 + \sigma_2 \mid \sigma_1 \in G(t, x; y_0, z_0, v_0), \, \sigma_2 \in \mathbb{R}, \, |\sigma_2| < \varepsilon \}
\]
for all \( (y, z) \in B_\delta(y_0, z_0) \subset \mathbb{R}^2 \).

(G3) For every \( t \in \mathbb{R} \), there exists some \( \lambda_\bullet \in L^2(\Omega) \) with \( \lambda_\bullet \geq 0 \) such that the following uniform Lipschitz condition (w.r.t. the last argument) holds Lebesgue-almost everywhere in \( \Omega \):
\[
\| G(t, \cdot; y_0, z_0, v_0) \|_{L^2(\Omega)} \leq \lambda_\bullet(\cdot) \| v_0 - w_0 \|_{L^2(\Omega)} \cdot [-1, 1]
\]
for all \( y_0, z_0 \in \mathbb{R} \) and \( v_0, w_0 \in L^2(\Omega) \).

(G3') There exists some \( \Lambda \in L^2(\mathbb{R}) \) with \( 0 \leq \Lambda < \infty \) such that the following uniform Lipschitz condition holds Lebesgue-almost everywhere in \( \Omega \):
\[
G(t, \cdot; y_1, z_1, v_1) \subset G(t, \cdot; y_2, z_2, v_2) + \Lambda(t) \left( |y_1 - y_2| + |z_1 - z_2| + \| v_1 - v_2 \|_{L^2(\Omega)} \right) \cdot [-1, 1]
\]
for all \( t \in \mathbb{R} \), \( y_1, y_2, z_1, z_2 \in \mathbb{R} \) and \( v_1, v_2 \in L^2(\Omega) \).

(G4) For every \( (y_0, z_0, v_0) \in \mathbb{R} \times \mathbb{R} \times L^2(\Omega) \), the set-valued map \( G(\cdot; y_0, z_0, v_0) : \mathbb{R} \times \mathbb{R} \rightharbra \mathbb{R} \) has a Lebesgue measurable selection, i.e., there exists a Lebesgue measurable function \( g : \mathbb{R} \times \Omega \rightharbra \mathbb{R} \) with \( g(t, x) \in G(t, x; y_0, z_0, v_0) \) for all \( t \in \mathbb{R} \) and \( x \in \Omega \).

(G5) \( G \) satisfies the following growth condition: There are \( \tilde{\alpha}, \tilde{\beta} \geq 0 \) and \( \alpha \in L^\infty(\mathbb{R}) \) with \( 0 \leq \alpha \leq \tilde{\alpha} \) such that for all \( y, z \in \mathbb{R} \), \( w \in L^2(\Omega) \) and any \( t \in \mathbb{R} \), \( x \in \Omega \),
\[
\sup \left\{ |\sigma| \mid \sigma \in G(t, x; y, z, w) \right\} \leq \alpha(t) + \tilde{\beta} \cdot (|y| + |z|).
\]

(G5') \( G \) fulfills the stronger condition of linear growth: There exist constants \( \tilde{\alpha}, \tilde{\beta} \geq 0 \) and a function \( \alpha \in L^\infty(\mathbb{R}) \) with \( 0 \leq \alpha \leq \tilde{\alpha} \) such that for all \( y, z \in \mathbb{R} \), \( w \in L^2(\Omega) \) and \( t \in \mathbb{R} \), \( x \in \Omega \),
\[
\sup \left\{ |\sigma| \mid \sigma \in G(t, x; y, z, w) \right\} \leq \alpha(t) + \tilde{\beta} \cdot |y|.
\]

(G6) The parameter \( \tilde{\beta} > 0 \) in growth condition (G5), (G5') respectively is smaller than the smallest eigenvalue \( \lambda_1 > 0 \) of the negative Laplacian operator with homogeneous Dirichlet boundary conditions in \( \Omega \):
\[
0 < \tilde{\beta} < \lambda_1.
\]
(G7) \( G \) has the following inclusion property w.r.t. the fourth argument: For all \( t \in \mathbb{R} \), \( x \in \Omega \), \( y \in \mathbb{R} \) and \( w \in L^2(\Omega) \), the map \( G(t, x; y, \cdot, w) : \mathbb{R} \to \mathbb{R} \) satisfies \( G(t, x; y, z_1, w) \subset G(t, x; y, z_2, w) \) whenever \( z_1 \leq z_2 \).

This class of parabolic inclusion problem with memory clearly differs from what was investigated by Chepyzhov and collaborators, for example (see, e.g., [14, 15, 16] and references therein).

3. Existence of solutions to the reaction-diffusion inclusions with delay. The existence of solutions has already been investigated in [37] for the special case of constant finite delay \( \theta \).

**Theorem 3.1.** Let \( \Omega \subset \mathbb{R}^d \) be a bounded open set with smooth boundary. In addition to \( T, \theta > 0 \) and \( c \geq 0 \) given, suppose that the set-valued map \( G : \mathbb{R} \times \Omega \times \mathbb{R} \times \mathbb{R} \times L^2(\Omega) \to \mathbb{R} \) satisfies the hypotheses (G1)–(G5).

Then, for every initial state \( u_0 \in C^0([-\theta, 0], L^2(\Omega)) \), there exists a continuous curve \( u : [-\theta, T] \to L^2(\Omega) \) such that \( u \) induces a strong solution to the following inclusion problem

\[
\begin{align*}
\partial_t u(t, x) - \Delta u(t, x) &\in G\left(t, x; u(t, x), \underset{t-\theta \leq s \leq t}{\text{ess sup}} [u(s, x)]_c, u(t, \cdot)\right) \\
u &\in u_0 \\
u &\equiv 0 \\
&\quad \text{a.e. in } [0, T] \times \Omega \\
&\quad \text{a.e. in } [-\theta, 0] \times \Omega \\
&\quad \text{on } (0, T) \times \partial \Omega.
\end{align*}
\]

Indeed, this theorem results from essentially the same proof as [37] Corollary 3.3. The only two differences are that first, the essential supremum there is taken over \( t - T \leq s \leq t \) (with the same time extent \( T \) as the interval of existence \([0, T]\)) and second, the initial state function \( u_0 \) there is constant w.r.t. time. These two modifications, however, do not have any significant influence on the proof in [37 § 3.1] due to the more general preparatory statements in [37 § 3.1] mainly based on [12].

Recall from [37 Corollary 2.7] and [37 Remark 2.12] that a strong solution to problem (1) is characterized equivalently in terms of integral solutions (in the sense of Bénilan [5]) and mild solutions \( u : [-\theta, T] \to L^2(\Omega) \) of the single-valued parabolic problem

\[
\begin{align*}
\partial_t u &= \Delta u + f(t) \\
u &\equiv u_0 \\
u &\equiv 0 \\
&\quad \text{in } (0, T) \times \Omega \\
&\quad \text{in } [-\theta, 0] \times \Omega \\
&\quad \text{on } (0, T) \times \partial \Omega
\end{align*}
\]

for a selector \( f \in L^2(0, T; L^2(\Omega)) \subset L^1(0, T; L^2(\Omega)) \) with

\[
f(t, x) \in G\left(t, x; u(t, x), \underset{t-\theta \leq s \leq t}{\text{ess sup}} [u(s, x)]_c, u(t, \cdot)\right)
\]

for Lebesgue-almost all \( t \in [0, T] \) and \( x \in \Omega \). Hence we have the additional representation

\[
u(t) = S(t) u_0(0) + \int_0^t S(t-s) f(s) \, ds, \quad t \in (0, T],
\]

where \( \{S(t)\}_{t \geq 0} \) denotes the heat semigroup on \( L^2(\Omega) \) with homogeneous Dirichlet boundary conditions. Parabolic differential equations are known to have some smoothing effect on their (weak) solutions and so, better results about the regularity in space are known (see, e.g., [4, 39, 42] for the general theory and [27, 37 § 4] for
Hence we just have to apply the preceding existence statement to the auxiliary map \( t \) by \( u \) Let \( \text{Lemma 3.3.} \)

the advantage that every \( u \) \( \in \mathbb{R}^d \) before is summarized in the following lemma whose proof is post-

Corollary 3.2. Let \( \Omega \subset \mathbb{R}^d \) be a bounded open set with smooth boundary, \( T \geq t_0, \theta > 0, c \geq 0 \) and \( \Theta \in L^1(\Omega, \mathbb{R}^+) \). Suppose that the set-valued map \( G : \mathbb{R} \times \Omega \times \mathbb{R} \times \mathbb{R} \times L^2(\Omega) \rightrightarrows \mathbb{R} \) satisfies the hypotheses (G1) – (G5).

Then, for every initial state \( u_0 \in C^0((-\infty, t_0], L^2(\Omega)) \), there exists a continuous curve \( u : (-\infty, T] \rightarrow L^2(\Omega) \) such that \( u \) induces a strong solution to the parabolic inclusion problem \( 1 \).

Proof. It results directly from Theorem 3.1 in a piecewise way. Without loss of generality we assume \( t_0 = 0 \). Indeed, we first focus on the wanted solution just in \([0, \frac{t}{2}]\) and then, the local result can be iterated until we have found a solution up to final time \( T \) after finitely many steps. This preliminary restriction to \([0, \frac{t}{2}]\) has the advantage that every \( u : (-\infty, \frac{t}{2}] \rightarrow L^2(\Omega) \) with \( u|_{(-\infty, 0]} = u_0 \) satisfies for each \( t \in [0, \frac{t}{2}] \) and \( x \in \Omega \)

\[
\text{ess sup}_{t - \theta - \Theta(x) \leq s \leq t} [u(s, x)]_c = \max \left\{ \text{ess sup}_{t - \theta - \Theta(x) \leq s \leq 0} [u_0(s, x)]_c, \text{ess sup}_{t - \theta \leq s \leq t} [u(s, x)]_c \right\}.
\]

Hence we just have to apply the preceding existence statement to the auxiliary map \( \hat{G} : \mathbb{R} \times \Omega \times \mathbb{R} \times \mathbb{R} \times L^2(\Omega) \rightrightarrows \mathbb{R} \) defined by

\[
\hat{G}(t, x, y, z, v) := G(t, x, y, \max \left\{ \text{ess sup}_{t - \theta - \Theta(x) \leq s \leq 0} [u_0(s, x)]_c, z \right\}, v),
\]

which obviously also satisfies conditions (G1) – (G5).

The essential tool for bridging the gap between our problem now and the one considered in \([37]\) before is summarized in the following lemma whose proof is post-

Lemma 3.3. Let \( \Omega \subset \mathbb{R}^d \) be a bounded open set with smooth boundary, \( 0 < \theta < \infty, c \geq 0 \) and \( \Theta \in L^1(\Omega, \mathbb{R}^+) \). Then the mapping \( C^0((-\infty, 0], L^2(\Omega)) \rightarrow L^2(\Omega) \) defined by \( u_0 \mapsto \text{ess sup}_{-\theta - \Theta(x) \leq s < 0} [u_0(s, x)]_c \) is well defined and continuous w.r.t. the metric \( d_{c, \gamma, \text{p.i.p.}} \) defined in \([14]\) below.

4. Main results. This paper focuses on the existence of pullback attractors of the non-autonomous dynamical system induced by parabolic differential inclusion \( 1 \).

For notational consistency with the literature a set-valued mapping \( \hat{A} : \mathbb{R} \rightrightarrows X \) for some space \( X \) will often be considered as a family \( \hat{A} = \{A(t) : t \in \mathbb{R}\} \) of subsets of \( X \) with the subsets of \( X \) denoted by \( A(t) \) instead of \( \hat{A}(t) \).

Theorem 4.1. Let \( \Omega \subset \mathbb{R}^d \) be bounded open with smooth boundary, \( c \geq 0, 0 < \theta < \infty \) and \( \Theta \in L^1(\Omega), \Theta \geq 0 \). Suppose that the set-valued map \( G : \mathbb{R} \times \Omega \times \mathbb{R} \times \mathbb{R} \times L^2(\Omega) \rightrightarrows \mathbb{R} \) satisfies the hypotheses (G1) – (G4), (G5'), (G6) and (G3'). Let \( \gamma > 0 \) be a constant larger than the smallest eigenvalue \( \lambda_1 > 0 \) of the negative Laplacian operator with homogeneous Dirichlet boundary conditions in \( \Omega \).

Then there exists a set-valued mapping \( \hat{A} : \mathbb{R} \rightrightarrows C^0((-\infty, 0], L^2(\Omega)) \), \( t \mapsto A(t) \) with following features:
1. For every $t \in \mathbb{R}$, the set $A(t) \subset C^0((-\infty,0], L^2(\Omega))$ is non-empty and compact with respect to locally uniform convergence in $(-\infty,0]$. For every function $\psi \in A(t)$, the limit of $e^{\gamma s} \psi(s)$ for $s \to -\infty$ exists in $L^2(\Omega)$.

2. $\hat{A}$ is negatively invariant, i.e., fixing $\tau \leq T$ arbitrarily, every $\psi \in A(T) \subset C^0((-\infty,0], L^2(\Omega))$ is related to $u : (-\infty, T] \times \Omega \to \mathbb{R}$ defined by $u(s,x) = \psi(s-T,x)$ which is a strong solution to the parabolic differential inclusion with space-dependent delay

$$\partial_t u - \Delta u \in G \left( t, x; u(t,x), \varphi \sup_{t-\theta \leq s \leq t} \left[ u(s,x) \right]_c, u(t,\cdot) \right)$$

on $\{ t \mid \tau \in (\tau, T) \times \Omega \}$.

3. $\hat{A}$ is positively invariant, i.e., for any real $\tau \leq T$ and $\psi \in A(\tau)$, every strong solution $u : (-\infty, T] \times \Omega \to \mathbb{R}$ to parabolic differential inclusion (7) with $u|_{(-\infty, \tau]} = \psi(\cdot - \tau)$ satisfies

$$u(\cdot + \tau)|_{(-\infty,0]} = \psi(\cdot + \tau - T)|_{(-\infty,0]} \in A(\tau).$$

4. $\hat{A}$ is pullback attracting in the following sense: Fix $T \in \mathbb{R}$, $q \in [0, \lambda_1 - \beta)$ and a sequence $(\tau_k)_{k \in \mathbb{N}}$ in $(-\infty, T)$ with $\tau_k \to -\infty$ arbitrarily. Choose any sequence $(\psi_k)_{k \in \mathbb{N}}$ in $C^0((-\infty,0], L^2(\Omega))$ such that $\sup_{k \in \mathbb{N}} \sup_{s \leq 0} \left( e^{\gamma s} \psi_k(s) \| L^2(\Omega) \right) < \infty$ and the limit of $e^{\gamma s} \psi_k(s)$ for $s \to -\infty$ exists for every $k \in \mathbb{N}$. Each $\psi_k$, $k \in \mathbb{N}$, initializes a strong solution $u_k : (-\infty, T] \times \Omega \to \mathbb{R}$ of parabolic differential inclusion (7) in $(\tau_k, T) \times \Omega$ with $u_k|_{(-\infty, \tau_k]} = \psi_k(\cdot - \tau_k)$ such that

$$\inf_{\xi \in A(T)} \sup_{\tau \leq s \leq T} \left\| u_k(s) - \xi(s - T) \right\|_{L^2(\Omega)} \to 0$$

and

$$\inf_{\xi \in A(T)} \left\| \varphi \sup_{\tau \leq s \leq T} \left[ u_k(s) \right]_c - \left[ \xi(s - T) \right]_c \right\|_{L^2(\Omega)} \to 0$$

for $k \to \infty$ at every time instant $\tau < T$.

A similar existence result can be shown for the problems with infinite delay (instead of space-dependent delay), i.e., depending on $\varphi \sup_{-\infty \leq s \leq t} u(s,x)_c$ in the fourth argument on the right-hand side. In particular, hypothesis (G7) specifies a monotonicity condition on the set-valued mapping $G$ with respect to its fourth argument, in which the essential supremum term usually occurs.

**Theorem 4.2.** Let $\Omega \subset \mathbb{R}^d$ be a bounded open set with smooth boundary, $c \geq 0$, $\theta := \infty$, $\gamma > \lambda_1$. Suppose that the set-valued map $G : \mathbb{R} \times \Omega \times \mathbb{R} \times L^2(\Omega) \to \mathbb{R}$ satisfies hypotheses (G1)–(G6), (G7).

Then there exists a set-valued mapping $\hat{A} : \mathbb{R} \to C^0((-\infty,0], L^2(\Omega))$, $t \mapsto A(t)$ satisfying the same conditions (1)–(4) as in Theorem 4.1.

From the conceptual point of view, we consider the reaction-diffusion inclusion with memory in the so-called history space setting, i.e., a past history variable is introduced as additional component and so, the underlying state space consists of (some) continuous functions $(-\infty,0) \to L^2(\Omega)$. This notion was proposed by Dafermos [19].
The first aspect of convergence in condition (4.) is usually called locally uniform convergence and, it is exactly the same topology as in many publications about so-called trajectory attractors (see, e.g., [14, 17, 18, 26, 34, 58] and references therein). Chepyzhov and Miranville suggested trajectory attractors for hyperbolic and parabolic differential equations with memory [15, 16]. Their underlying state space and the semigroup are similar to our considerations, but their criterion of attraction is based on convergence in forward time direction (see [16, Definition 0.1 (3)], in particular).

Our main results, however, focus on pullback attraction, i.e., convergence w.r.t. Hausdorff semi-distance for the initial time tending to $-\infty$ while the end time is fixed. To the best of our knowledge, the only statements combining trajectory attractors with pullback convergence so far were published by Zhao and Zhou [59]. They consider trajectories in forward time direction and so, their underlying state space consists of continuous functions on $[0, \infty) \rightarrow Y$ (with a Banach space $Y$).

In a word, main Theorems 4.1 and 4.2 specify sufficient conditions for a “pullback trajectory attractor” combining the modified gist of Chepyzhov and Miranville [16] with pullback attraction – but in a new way quite different from [59].

Our proofs are based on the concept of multi-valued non-autonomous dynamical systems in metric spaces (see §5.2 below for details). It is worth mentioning here that the general theory of their pullback attractors provides some more information about the convergence. Indeed, the semi-distance between sets used in criterion (17) below implies, for example, that the convergence in statement (4.) is uniform with respect to all strong solutions $u_k : (-\infty, T] \times \Omega \rightarrow \mathbb{R}$ initialized by the same $\psi_k$.

5. A way to a multi-valued non-autonomous dynamical system.

5.1. Reformulation as an abstract evolution inclusion in a metric vector space. The inclusion system (1) can be reformulated as an abstract evolution equation on the function space

$$C_\gamma := \left\{ u \in C^0\left((-\infty, 0], L^2(\Omega)\right) \bigg| \lim_{s \to -\infty} \left( u(s) \cdot e^{\gamma s} \right) \text{ exists in } L^2(\Omega) \right\}$$

where $\gamma > 0$ is a constant specified in a moment. $C_\gamma$ supplied with the weighted supremum norm

$$\|u\|_\gamma := \sup_{s \leq 0} \left( e^{\gamma s} \cdot \|u(s)\|_{L^2(\Omega)} \right) < \infty$$

is a separable Banach space [8] since the transformation argument on [30, page 15] can be extended to function values in a separable Banach space such as $L^2(\Omega)$ here. We reformulate system (1) as the abstract evolution inclusion

$$\frac{du}{dt} + Au \in \mathcal{G}(t, u_t),$$

where $A$ is defined in terms of the minus Laplace operator on $\Omega$ with homogeneous Dirichlet boundary condition, $u_t := u(\cdot + t) : (-\infty, 0] \rightarrow L^2(\Omega)$ and the set-valued map $\mathcal{G} : \mathbb{R} \times C_\gamma \rightharpoonup L^2(\Omega)$ is defined by

$$\mathcal{G}(t, \phi) := \left\{ \xi \in L^2(\Omega) \bigg| \xi(x) \in G\left(t, x; \phi(0, x), \esssup_{-\theta - \Theta(x) \leq s \leq 0} [\phi(s, x)]_c, \phi(0) \right) \right\}$$

for almost every $x \in \Omega$ (10)
The dependence of $G(t, \phi)$ on the essential supremum
$$\Omega \rightarrow \mathbb{R}, \quad x \mapsto \text{ess sup}_{-\theta - \Theta(x) \leq s \leq 0} [\phi(s, x)]_c$$
is rather difficult to handle by means of the weighted supremum norm $\| \cdot \|_{\gamma}$ on $C_\gamma$. It seems to be more recommendable to consider the metric $d_{c,\gamma} : C_\gamma \times C_\gamma \rightarrow [0, \infty)$ instead
$$d_{c,\gamma}(u, v) := \sup_{s \leq 0} (e^{\gamma s} \| u(s) - v(s) \|_{L^2(\Omega)}) + \text{ess sup}_{-\tau \leq s \leq 0} \| [u(s, \cdot)]_c - [v(s, \cdot)]_c \|_{L^2(\Omega)}.$$  \hfill (11)

In regard to pullback attractors, however, it has proved to be useful to model some “effect of vanishing influence of delay” or “vanishing memory” in the sense that the past becomes less and less relevant. Indeed, the first term in $d_{c,\gamma}(u, v)$ takes this aspect into consideration in form of the weighting factor $e^{\gamma s}$ (with $s \leq 0$ and fixed $\gamma > 0$). For implementing it also in the second term with essential supremum, we introduce additionally the following family of metrics $d_{c,\gamma,\tau} : C_\gamma \times C_\gamma \rightarrow [0, \infty)$, $\tau > 0$, defined by
$$d_{c,\gamma,\tau}(u, v) := \sup_{s \leq 0} (e^{\gamma s} \| u(s) - v(s) \|_{L^2(\Omega)}) + \text{ess sup}_{-\tau \leq s \leq 0} \| [u(s, \cdot)]_c - [v(s, \cdot)]_c \|_{L^2(\Omega)} + \| [u]_c - [v]_c \|_{L^2(\Omega; L^\infty([-\tau, 0]))}.$$  \hfill (12)

Levi’s theorem about monotone convergence provides directly a connection between these metrics:

**Lemma 5.1.** The distance $d_{c,\gamma}$ between any $u, v \in C_\gamma$ satisfies
$$d_{c,\gamma}(u, v) = \lim_{\tau \rightarrow \infty} d_{c,\gamma,\tau}(u, v) = \sup_{\tau > 0} d_{c,\gamma,\tau}(u, v).$$

The representation in (12) via supremum is analogous to a standard result about norms in any real Banach space $Y$, which is a consequence of Hahn-Banach theorem:
$$\| \xi \|_Y = \sup \{ \ell(\xi) \mid \ell : Y \rightarrow \mathbb{R} \text{ linear and continuous}, \| \ell \|_{\text{Lin}(Y, \mathbb{R})} \leq 1 \}. \hfill (13)$$

Indeed, one distance can be represented as supremum of a family of some other distance functions (each of which need not satisfy all typical conditions on metrics) and so, the related convergence is uniform w.r.t. the underlying parameter of the family. For avoiding misunderstandings with terms established in functional analysis, the convergence in $(C_\gamma, d_{c,\gamma})$ is called “uniform w.r.t. the parameter” here. If the convergence w.r.t. $d_{c,\gamma,\tau}$ holds for each $\tau > 0$ separately, we call it “pointwise in parameter” abbreviated as “p.i.p.”. (In [43] § 0.3.4 step (F)], the terms “strong” and “weak convergence” are used instead, but it is worth pointing out that this interpretation should be understood in the purely metric context, i.e., the linear structure of $C_\gamma$ is not relevant in this generalizing interpretation and so, the “weak convergence” a.k.a. “convergence pointwise in parameter $\tau$” here is not based on some dual space as usual in functional analysis.)

From the topological point of view, the condition on a sequence in $C_\gamma$ to converge w.r.t. each metric $d_{c,\gamma,\tau}$, $\tau > 0$, can be expressed by means of just a single metric on $C_\gamma$, e.g.,

$$d_{c,\gamma,\text{p.i.p.}} : (u, v) \mapsto \| u - v \|_{\gamma} + \sum_{\tau = 1}^{\infty} 2^{-\tau} \| [u]_c - [v]_c \|_{L^2(\Omega; L^\infty([-\tau, 0]))}. \hfill (14)$$
in which we have slightly modified the Fréchet product metric $\| \cdot \|_{\mathbb{R}^2}$. It is worth mentioning that $d_{c,\gamma,p.i.p.}(u,v)$ takes the norm of $L^2(\Omega; L^\infty([-\tau,0]))$ (for all $\tau \in \mathbb{N}$) into consideration – and not the standard norm of $L^\infty([-\tau,0]; L^2(\Omega))$. At first glance, this similarity does not imply that $d_{c,\gamma,p.i.p.}$ metrizes the compact-open topology for continuous functions $(-\infty,0] \to (L^2(\Omega), \| \cdot \|_{L^2(\Omega)})$.

**Lemma 5.2.** Each of the metric spaces $(C_\gamma, d_{c,\gamma}), (C_\gamma, d_{c,\gamma,\tau})$ for any $\tau > 0$ and $(C_\gamma, d_{c,\gamma,p.i.p.})$ is complete.

The proof of this lemma is postponed to §8.1.

Finally, we specify solutions to evolution inclusion (16) in $C_\gamma$: Fix initial state $\phi \in C_\gamma$ and initial time $t_0 \in \mathbb{R}$ and, choose final time $T > t_0$ arbitrarily. Under appropriate growth conditions on $G$, each solution $u : (-\infty, T] \to L^2(\Omega)$ to the parabolic differential inclusion

$$
\begin{align*}
\partial_t u - \Delta u &\in G(t, x; u(t, x), \sup_{t-\theta - \Theta(x) \leq s \leq t} [u(s, x)], u(t, \cdot)) \quad \text{a.e. in } (t_0, T) \times \Omega \\
u &= \phi(\cdot - t_0) \quad \text{a.e. in } (-\infty, t_0) \times \Omega \\
u &= 0 \quad \text{on } (t_0, T) \times \partial \Omega
\end{align*}
$$

is related to a selector $f \in L^2(t_0, T; L^2(\Omega))$ with

$$
f(t, x) \in G(t, x; u(t, x), \sup_{t-\theta - \Theta(x) \leq s \leq t} [u(s, x)], u(t, \cdot)).
$$

The equivalent mild characterization (16) leads to the following representation of the solution $y_\tau(\cdot) := y_\tau(\cdot; t_0, \phi) \in C^0((-\infty,0], L^2(\Omega))$ at time $t \geq t_0$

$$
y_\tau(s) = \begin{cases} 
S(t - t_0 + s) \phi(0) + \int_{t_0}^{t+s} S(t + s - \tau) f(\tau) \, d\tau & \text{if } s \in [-t + t_0), 0] \\
\phi(t - t_0 + s) & \text{if } s < -(t - t_0) 
\end{cases}
$$

with a selector $f : [t_0, t] \to L^2(\Omega)$ satisfying $f(\tau) \in G(\tau; y_\tau(\cdot + \tau - t)|_{(-\infty,0)})$ for a.e. $\tau \in [t_0, t]$.

By definition, $\phi \in C_\gamma$ implies that the limit of $e^{\gamma s} \phi(s) \in L^2(\Omega)$ for $s \to -\infty$ exists and so,

$$
\lim_{s \to -\infty} (e^{\gamma s} y_\tau(s)) = e^{-\gamma(t-t_0)} \lim_{s \to -\infty} (e^{\gamma(t-t_0+s)} \phi(t - t_0 + s))
$$

also exists in $L^2(\Omega)$, i.e., $y_\tau \in C_\gamma$.

**5.2. Multi-valued system generated by the evolution inclusion.** Define

$$
\mathbb{R}^2_\geq := \{(t,t_0) \in \mathbb{R}^2 \mid t \geq t_0\}
$$

and consider any metric space $(X,d_X)$. A set-valued map $U : \mathbb{R}^2_\geq \times X \rightrightarrows X$ is called a **multi-valued non-autonomous dynamical system** (MNDS) if it satisfies the following conditions:

1. each value $U(t,t_0,x_0)$ is a nonempty closed subset of $X$,
2. $U(t,t,x_0) = \{x_0\}$ for every $t \in \mathbb{R}$ and $x_0 \in X$,
3. $U(t_2,t_0,x_0) \subset U(t_2,t_1, U(t_1,t_0,x_0))$ whenever $t_0 \leq t_1 \leq t_2$ and $x_0 \in X$. 

In particular, \( U \) induces a set-valued mapping \( \mathbb{R}_+^2 \times \mathcal{P}(X) \to \mathcal{P}(X) \) (with \( \mathcal{P}(X) \) denoting the power set of \( X \)) according to \( U(t, t_0, M) := \bigcup_{x \in M} U(t, t_0, x) \subset X \) for any \( (t, t_0) \in \mathbb{R}_+^2 \) and \( M \subset X \). This notion is already underlying the right-hand side of the inclusion in condition (3.).

A multi-valued non-autonomous dynamical system \( U : \mathbb{R}_+^2 \times X \to X \) is said to be strict if even equality holds in condition (3.), i.e.,
\[
(4.) \quad U(t_2, t_0, x_0) = U(t_2, t_1, U(t_1, t_0, x_0)) \quad \text{for all} \quad t_0 \leq t_1 \leq t_2 \quad \text{and} \quad x_0 \in X.
\]

Further details about this setting are presented in [8, 9, 38, 45], for example.

Now the solutions to evolution inclusion (9) prove to induce a multi-valued non-autonomous dynamical system on \( C \) if \( \gamma > 0 \) is fixed sufficiently large (see § 8.2 for the details of the proof):

**Proposition 5.3.** Suppose \( G : \mathbb{R} \times \Omega \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}^2(\Omega) \to \mathbb{R} \) to satisfy the assumptions (G1) – (G5). Fix \( \theta > 0 \), \( \Theta \in L^1(\Omega) \) with \( \Theta \geq 0 \) and, let \( \gamma > 0 \) be a constant larger than the smallest eigenvalue \( \lambda_1 > 0 \) of the negative Laplacian operator with homogeneous Dirichlet boundary conditions in \( \Omega \).

For each \( (t, t_0) \in \mathbb{R}^2_+ \) and \( \phi \in C_\gamma \), let \( U(t, t_0, \phi) \subset C_\gamma \) consist of all solutions \( y_t(\cdot; t_0, \phi) \in C_\gamma \) of evolution inclusion (9) specified in § 5.1 (see, e.g., mild representation (10) with some selector \( f \)).

Then \( U : \mathbb{R}_+^2 \times C_\gamma \to C_\gamma \) is a strict multi-valued non-autonomous dynamical system. For every \( (t, t_0) \in \mathbb{R}_+^2 \), the set values of \( U(t, t_0, \cdot) : X \to X \) are closed w.r.t. both \( d_{c,\gamma} \) and \( d_{c,\gamma,\tau} \) for each \( \tau > t - t_0 \).

6. **The general theory of pullback attractors in spaces with (possibly) two metrics.**

6.1. **Pullback attractors of MNDS in a (standard) metric space \( (X, d_X) \): Definition and general existence via pullback asymptotic compactness.**

Now we specify the concept of pullback attractors with respect to a universe of sets and establish a condition sufficient for their existence. These considerations are continuing the lines of [9, 10], [38, Ch.9], for example. Existence Theorem 6.3 below slightly extends [8, Theorem 3.3] because we assume closed graph of \( U(t, t_0, \cdot) : C_\gamma \to C_\gamma \) instead of upper semi-continuity for each tuple \( (t, t_0) \in \mathbb{R}_+^2 \).

For the general setting consider any metric space \( (X, d_X) \) again and let \( U : \mathbb{R}_+^2 \times X \to X \) be a multi-valued non-autonomous dynamical system. A set-valued mapping \( \hat{D} : \mathbb{R} \to X \) (equivalently denoted by \( \hat{D} = \{ D(t) : t \in \mathbb{R} \} \)) is said to be negatively, strictly, or positively invariant (resp.) for the MNDS \( U \) if
\[
D(t) = \bigcap_{\tau \to +\infty} \text{dist}_X \left( U(t, t - \tau, D(t - \tau)) \right) \quad \text{for every} \quad (t, \tau) \in \mathbb{R}_+^2.
\]

Let \( D \) abbreviate the family of set-valued mappings \( \mathbb{R} \to X \) with nonempty closed values. A mapping \( \hat{K} \in \mathcal{D} \) is called pullback \( \mathcal{D} \)-attracting if every \( \hat{D} \in \mathcal{D} \) satisfies
\[
\lim_{\tau \to +\infty} \text{dist}_X \left( U(t, t - \tau, D(t - \tau)), K(t) \right) = 0 \quad \text{for each} \quad t \in \mathbb{R}.
\]

Combining pullback absorption and some form of compactness is to lay the basis for verifying pullback attraction. For this purpose we need an additional feature of the families \( \hat{D} \in \mathcal{D} \) for which the limit is required. Following a suggestion by
Let \( \hat{D} \subset D \) of multi-valued mappings \( \mathbb{R} \rightrightarrows X \) be a universe if it satisfies the following inclusion for every family \( \hat{D} = \{ D(t) : t \in \mathbb{R} \} \in \hat{D} \) in addition: Every further mapping \( \hat{D}' \in \hat{D} \) with \( D'(t) \subset D(t) \) for each \( t \in \mathbb{R} \) also belongs to \( \hat{D} \).

**Definition 6.1.** Let \( U : \mathbb{R}^2_+ \times X \rightrightarrows X \) be a MNDS in a metric space \((X, d_X)\) and \( D \) be a universe. A family \( \hat{A} = \{ A(t) : t \in \mathbb{R} \} \) is said to be a global pullback \( D \)-attractor for \( U \) if

(i) \( A(t) \subset X \) is compact for any \( t \in \mathbb{R} \),
(ii) \( \hat{A} \) is pullback \( D \)-attracting and,
(iii) \( \hat{A} \) is negatively invariant.

\( \hat{A} \) is said to be a strict global pullback \( D \)-attractor if the invariance property in (iii) is strict.

It is worth mentioning that the concept of pullback attractor does not imply any general equivalence to forward attractors or to so-called trajectory attractors, both of which have already been extended to non-autonomous dynamical systems (see, e.g., \[14, 16, 17, 20, 38, 58\]). As an obvious formal difference, pullback attractors consist of subsets of state space \( X \) corresponding to a set-valued mapping of time whereas (uniform) trajectory attractors are defined as a set of (usually continuous) functions \([0, \infty) \rightarrow X\) specified by means of the translation semigroup. There are examples showing that pullback and forward attractors do not have to coincide (see, e.g., \[36, 38\]) and, the relation between global forward and trajectory attractors (in the sense of Chepyzhov and Vishik) is investigated in \[34\], for example.

Following the standard approach to pullback attractors as presented in \[8, 38\], for example, the main tool to prove the existence of an attractor is the concept of pullback \( \omega \)-limit set. For a multi-valued mapping \( \hat{D} : \mathbb{R} \rightrightarrows X \), we define the pullback \( \omega \)-limit set as the \( t \)-dependent set \( \Lambda(\hat{D}, t, d_X) \subset X \) given by

\[
\Lambda(\hat{D}, t, d_X) := \bigcap_{\tau \geq 0} \bigcup_{s \geq \tau} U(t, t - s, D(t - s)).
\]

(19)

where the closure is considered w.r.t. \( d_X \). Obviously this set is closed in \((X, d_X)\), but it may be empty. It can be proved that \( y \in \Lambda(\hat{D}, t, d_X) \) if and only if there exist sequences \( t_n \nearrow +\infty \) in \( \mathbb{R} \) and \( y_n \in U(t, t - t_n, D(t - t_n)) \subset X \) with \( \lim_{n \to +\infty} y_n = y \).

The next lemma extends \[8\] Lemma 3.2] to closed graphs of \( U(t, \tau, \cdot) \) and thus, it is a generalization of \[9\] Theorem 6 and Lemma 8] to the case of a general universe \( D \) (instead of just the bounded sets of \( X \)). It will be extended to a more general metric setting in Lemma 6.8 below and so, the proof is not presented in detail here.

**Lemma 6.2.** Let \( U : \mathbb{R}^2_+ \times X \rightrightarrows X \) be a multi-valued non-autonomous dynamical system in the metric space \((X, d_X)\). Assume for each \((t, \tau) \in \mathbb{R}^2_+\) that \( U(t, \tau, \cdot) : X \rightrightarrows X \) has closed graph. Let \( \hat{B} : \mathbb{R} \rightrightarrows X \) be a multi-valued mapping such that \( U \) is pullback asymptotically compact w.r.t. \( \hat{B} \), i.e., for any real sequence \( t_n \to +\infty \), every sequence \( y_n \in U(t, t - t_n, B(t - t_n)) \subset X \) with \( \lim_{n \to +\infty} y_n \) has a converging subsequence.

Then, for every \( t \in \mathbb{R} \), the pullback \( \omega \)-limit set \( \Lambda(\hat{B}, t, d_X) \subset X \) is non-empty, compact, and

\[
\lim_{\tau \to +\infty} \text{dist}_X \left( U(t, t - \tau, B(t - \tau)), \Lambda(\hat{B}, t, d_X) \right) = 0,
\]

(20)
\begin{align*}
\Lambda(\hat{B}, t, d_X) \subset U(t, \tau, \Lambda(\hat{B}, \tau, d_X)) & \quad \text{for all } (t, \tau) \in \mathbb{R}^2. \tag{21}
\end{align*}

We can now specify a condition sufficient for the existence of pullback attractors. It results from the preceding lemma for exactly the same reasons as [8, Theorem 3.3] as we show for a more general situation in section 7 below.

**Theorem 6.3.** In addition to the hypotheses in Lemma 6.2, suppose that \( \hat{B} \in \mathcal{D} \) is pullback \( \mathcal{D} \)-absorbing and has closed values. Then, the set-valued mapping \( \hat{A} : \mathbb{R} \rightharpoonup X, t \mapsto A(t) \) given by
\[
A(t) := \Lambda(\hat{B}, t, d_X)
\]
is a pullback \( \mathcal{D} \)-attractor. Moreover, \( \hat{A} \) is the unique element from \( \mathcal{D} \) with these properties. In addition, if \( U \) is a strict MNDS, then \( \hat{A} \) is strictly invariant.

6.2. Pullback flattening for MNDS in metric spaces (not necessarily Banach spaces). Obviously, some form of compactness has to be guaranteed for covering the asymptotic features of the dynamical systems independently of the respective initial time instants. In Lemma 6.2, we used the concept of pullback asymptotic compactness, which represents a form of sequential compactness: For any sequence \( t_n \to +\infty \) in \( \mathbb{R} \), every sequence \( y_n \in U(t, t - t_n, B(t - t_n)) \) has a converging subsequence (see, for example, [8, 9, 10, 38]).

An alternative criterion is based on the Kuratowski measure of non-compactness. Zhong et al. suggested the related “condition (C)” for attractors of autonomous dynamical systems in Banach spaces (e.g., [44, 53, 60]) and then extended it to pullback attractors of non-autonomous dynamical systems (e.g., [41, 51, 52, 54, 55, 56]). It has the basic notion in common with the flattening property, which was introduced for proving the existence of random attractors in [35] and which is presented in [11, §2.4] as well as [38, Ch. 12] for the general setting in Banach spaces. Here we adapt [11, Definition 2.24] to the multi-valued case in metric spaces, which do not have to be Banach spaces:

**Definition 6.4.** Let \((X, d_X)\) be a metric space, \( U : \mathbb{R}^2 \times X \rightharpoonup X \) a multi-valued non-autonomous dynamical system and \( \mathcal{D} \) a universe in \( X \).

\( U \) is said to be \( \mathcal{D} \)-pullback flattening w.r.t. \( d_X \) if for any \( t \in \mathbb{R} \) fixed arbitrarily, every \( \hat{B} = \{B(s) : s \in \mathbb{R}\} \in \mathcal{D} \) and \( \varepsilon > 0 \), there exist \( T_0 = T_0(t, \varepsilon, \hat{B}) > 0 \) and a nonempty subset \( M_\varepsilon \subset X \) such that

(i) \( M_\varepsilon \) is relatively compact,

(ii) \( \sup \{ \text{dist}_X(v, M_\varepsilon) \mid s \leq t - T_0, \ v \in U(t, s, B(s)) \} \leq \varepsilon, \)

For the sake of transparency, we formulate explicitly that whenever \( X \) is a Banach space, the established definition (as, e.g., in [11, 38, 41, 54, 55, 56, 60]) is just a special case. This implication is a rather obvious consequence of the compactness theorem of Heine-Borel and so, its proof is skipped.

**Lemma 6.5.** Let \( X \) be a real Banach space, \( U : \mathbb{R}^2 \times X \rightharpoonup X \) a MNDS and \( \mathcal{D} \) a universe in \( X \).

Then \( U \) is \( \mathcal{D} \)-pullback flattening (in the sense of Definition 6.4) if it satisfies the following condition: For any \( t \in \mathbb{R} \), every \( \hat{B} = \{B(s) : s \in \mathbb{R}\} \in \mathcal{D} \) and each \( \varepsilon > 0 \), there exist some \( T_0 = T_0(t, \varepsilon, \hat{B}) > 0 \), a finite-dimensional subspace \( X_\varepsilon \subset X \) and a bounded projector \( P_\varepsilon : X \to X_\varepsilon \) such that
and value space with different metrics, namely (continuity in Banach spaces (e.g., [41, 52, 56, 55, 60]) inspire us to supply domain induces a weaker condition on the considered system than the “strong” metric $d_j \in J \subset X$.

back attractors similar to Theorem 6.3. The main idea is to select the metrics the same arguments as in [57, page 27].

Definition 6.7. Let $(X, d_X)$ be a MNDS in a metric space $(X, d_X)$ and $D$ be a universe. Whenever $U$ is $D$-pullback flattening, then $\bar{U}$ is also pullback asymptotically compact w.r.t. any $\bar{B} \in D$ (in the sense of Lemma 6.2).

The lemma adapts [11] Theorem 2.25 to our metric setting. If $X$ is a uniformly convex Banach space, then the opposite inclusion can be verified by arguments similar to those used for [11] Theorem 2.25 or [38] Theorem 12.12 (the latter for skew-product flows). However, we are not going to use this opposite inclusion here.

6.3. Applying the gist of uniform and p.i.p. convergence in metric spaces to attractors. The concepts of strong and weak topologies can be extended to metric spaces which are possibly nonlinear – as sketched in § 5.1 (and § 3.4 step (F))). Indeed, consider a nonempty set $X$ with a metric $d_X$ and a family $(d_{X,j})_{j \in J}$ of distance functions $X \times X \to [0, \infty)$ satisfying $d_X = \sup_{j \in J} d_{X,j}$. Motivated by equation (13), which is well established in Banach spaces, we regard a sequence $(x_n)_{n \in \mathbb{N}}$ in $X$ as converging to some $x \in X$ “uniformly w.r.t. the parameter $j \in J$” if $d_X(x_n, x) \to 0 \ (n \to \infty)$ and $(x_n)_{n \in \mathbb{N}}$ is said to converge to $x$ “pointwise in parameter $j \in J$” (p.i.p.) if $d_{X,j}(x_n, x) \to 0 \ (n \to \infty)$ holds for every index $j \in J$.

If the index set $J$ is at most countable, we assume $J \subset \mathbb{N}$ without loss of generality and, the p.i.p. convergence can be expressed in terms of a single distance function $d_{X,p.i.p.} : X \times X \to [0, 1]$, e.g.,

$$d_{X,p.i.p.}(y, z) := \sum_{j \in J} 2^{-j} \frac{d_{X,j}(y, z)}{1 + d_{X,j}(y, z)} \quad (y, z \in X).$$

(22)

It is worth mentioning here that this interpretation does not require $d_{X,j}$ or $d_{X,p.i.p.}$ to satisfy all three typical conditions on metrics. Their essential purpose is just to quantify distances between any two points of $X$. If each $d_{X,j}$, $j \in J$, is a metric on $X$, however, then so is the proposed $d_{X,p.i.p.}$ (the triangle inequality results from the same arguments as in [57] page 27).

In this setting, we want to specify sufficient conditions for the existence of pullback attractors similar to Theorem 5.3. The main idea is to select the metrics $d_X$, $d_{X,p.i.p.}$ in a way which is more convenient to verify in the respective context. In regard to suitable (asymptotic) compactness, for example, the “p.i.p.” metric $d_{X,p.i.p.}$ induces a weaker condition on the considered system than the “strong” metric $d_X$, which is often used in Banach spaces (like the examples in [8, 41]). With respect to (semi-) continuity or closed graph, however, publications using norm-to-weak continuity in Banach spaces (e.g., [41, 52, 56, 55, 60]) inspire us to supply domain and value space with different metrics, namely $(X, d_X) \sim (X, d_{X,p.i.p.})$. To the best of our knowledge, this aspect is new in the metric setting.

The basic idea of using diverse metrics simultaneously leads to the question how to specify pullback $\omega$-limit sets.

Definition 6.7. Let $(X, d_X)$ be a metric space and $d_{X,j} : X \times X \to [0, \infty)$, $j \in J \subset \mathbb{N}$, a family of distance functions with $d_X = \sup_{j \in J} d_{X,j}$ and, define $d_{X,p.i.p.}$ by (22).
A multi-valued non-autonomous dynamical system $U : \mathbb{R}_+^2 \times X \rightrightarrows X$ is said to have the pullback $\omega$-Mazur property with respect to a mapping $\hat{B} : \mathbb{R} \rightrightarrows X$ and $d_X, d_{X,p.i.p.}$ if the related pullback $\omega$-limit sets satisfy the following condition for every tuple $(t, \tau) \in \mathbb{R}_+^2$: By definition (19), every point $y \in \Lambda(\hat{B}, \tau, d_{X,p.i.p.})$ is related to sequences $s_n \to +\infty$, $y_n \in U(\tau - s_n, x_n)$, $x_n \in B(\tau - s_n)$ with $d_{X,p.i.p.}(y_n, y) \to 0$. It is now required that some further sequences $\tilde{s}_k \to +\infty$, $\tilde{y}_k \in U(\tau - \tilde{s}_k, \tilde{x}_k)$, $\tilde{x}_k \in B(\tau - \tilde{s}_k)$ satisfy

(i) $d_X(\tilde{y}_k, y) \to 0$ for $k \to \infty$,

(ii) $\liminf_{n \to \infty} \inf_{k \geq n} \text{dist}_{X,p.i.p.}\left(U(t, \tau, y_n), U(t, \tau, \tilde{y}_k)\right) = 0$.

Obviously $d_{X,j} \leq d_X$ implies for each $j \in \mathcal{J}$ that the closure w.r.t. $d_X$ on the right-hand side of (19) is contained in its counterpart w.r.t. $d_{X,j}$ and so, $\Lambda(\hat{B}, t, d_X) \subset \Lambda(\hat{B}, t, d_{X,p.i.p.})$ is satisfied at every time $t \in \mathbb{R}$. Condition (i) specifying the pullback $\omega$-Mazur property ensures the equality of these pullback $\omega$-limit sets, i.e., each element of $\Lambda(\hat{B}, t, d_{X,p.i.p.})$ can be characterized as a limit w.r.t. $d_X$:

$$
\Lambda(\hat{B}, t, d_X) = \Lambda(\hat{B}, t, d_{X,p.i.p.}).
$$

Condition (ii) provides a connection between the respective "flows" of the related approximating sequences. It considers the Hausdorff semi-distance between sets w.r.t. $d_{X,p.i.p.}$ and so, it can be interpreted as some form of sequential upper semi-continuity. Here it is worth mentioning that this condition (ii) does not concern the set-valued mapping $U(t, \tau, \cdot) : X \rightrightarrows X$ in general, but merely those sequences which are related to pullback $\omega$-limit points and, they may depend on $t, \tau, y, (y_n)_{n \in \mathbb{N}}$. This is an essential difference from other standard assumptions like norm-to-weak upper semi-continuity.

The designation "$\omega$-Mazur property" is motivated by two aspects: First we focus essentially on the asymptotic features of sequences as the underlying initial time tends to $-\infty$. This is indicated in the letter $\omega$. Secondly, in any normed linear spaces, the well-known Lemma of Mazur implies an alternative characterization of weakly converging sequences in terms of strongly converging convex combinations (see, e.g., [57, Ch.V.1 Theorem 2 on page 120]). It is this type of relationship between weak and strong convergence which we extend beyond vector spaces here.

The next step is to extend Lemma 6.2 to this nonlinear setting with (possibly) several metrics. Obviously Lemma 6.2, which goes essentially back to [8], is the special case with $d_X = d_{X,j}$ for all indices $j \in \mathcal{J}$ since $d_X$ and $d_{X,p.i.p.}$ are then topologically equivalent and the pullback $\omega$-Mazur property is trivial (just setting $\tilde{y}_n := y_n, n \in \mathbb{N}$). This adapted lemma, however, leads to a further theorem about the existence of pullback attractors:

**Lemma 6.8.** Consider the metric space $(X, d_X)$, the family of metrics $(d_{X,j})_{j \in \mathcal{J}}$ and $d_{X,p.i.p.}$ as in Definition 6.7. Let $U : \mathbb{R}_+^2 \times X \rightrightarrows X$ be a multi-valued non-autonomous dynamical system. Suppose $\hat{B} : \mathbb{R} \rightrightarrows X$ to be a multi-valued mapping satisfying

1. $U$ is pullback asymptotically compact w.r.t. $\hat{B}$ and $d_{X,p.i.p.}$, i.e., for any real sequence $t_n \to +\infty$, every sequence $y_n \in U\left(t, t - t_n, B(t - t_n)\right)$ has a subsequence converging w.r.t. $d_{X,p.i.p.}$.
Then, for every $t \in \mathbb{R}$, the pullback $\omega$-limit set $\Lambda(\hat{B}, t, d_{X,p.i.p.}) \neq \emptyset$ is compact w.r.t. $d_{X,p.i.p.}$ and
\[
\lim_{\tau \to +\infty} \text{dist}_{X,p.i.p.} \left( U(t, t - \tau, B(t - \tau)), \Lambda(\hat{B}, t, d_{X,p.i.p.}) \right) = 0. \tag{23}
\]
Assume in addition that
(2.) for each $(t, \tau) \in \mathbb{R}^2_\tau$, the set-valued mapping $U(t, \tau, \cdot) : (X, d_X) \rightrightarrows (X, d_{X,p.i.p.})$ has closed graph,
(3.) $U$ has the pullback $\omega$-Mazur property w.r.t. $\hat{B}$, $d_X$, $d_{X,p.i.p.}$ (in the sense of Definition 6.14).

Then $\Lambda(\hat{B}, t, d_{X,p.i.p.}) = \Lambda(\hat{B}, t, d_X) \subset X$ is negatively invariant, i.e., for all $(t, \tau) \in \mathbb{R}^2_\tau$,
\[
\Lambda(\hat{B}, t, d_{X,p.i.p.}) \subset U(t, \tau, \Lambda(\hat{B}, \tau, d_{X,p.i.p.})). \tag{24}
\]

**Theorem 6.9.** In addition to all the assumptions of Lemma 6.8, let $\mathcal{D}$ be a universe in $X$. Suppose $\hat{B} \in \mathcal{D}$ to be pullback $\mathcal{D}$-absorbing and to have closed values w.r.t. $d_X$.

Then, the set-valued mapping $\hat{A} : \mathbb{R} \rightrightarrows X$, $t \mapsto A(t) := \Lambda(\hat{B}, t, d_{X,p.i.p.}) = \Lambda(\hat{B}, t, d_X)$ is a pullback $\mathcal{D}$-attractor w.r.t. $d_{X,p.i.p.}$. Furthermore, $\hat{A}$ is the unique element from $\mathcal{D}$ with these properties. In addition, if $U$ is a strict MNDS, then $\hat{A}$ is strictly invariant.

As an immediate consequence of Lemma 6.6, assumption (1.) about asymptotic pullback compactness is satisfied whenever $U$ is $\mathcal{D}$-pullback flattening and so, we obtain:

**Corollary 6.10.** Consider the metric space $(X, d_X)$, the family of metrics $(d_{X,j})_{j \in J}$ and $d_{X,p.i.p.}$, as in Definition 6.7. Suppose $\mathcal{D}$ to be a universe in $X$ and $\hat{B} : \mathbb{R} \rightrightarrows X$ to belong to $\mathcal{D}$. Let $U : \mathbb{R}^2_\tau \times X \rightrightarrows X$ be a multi-valued non-autonomous dynamical system satisfying hypotheses (2.), (3.) in Lemma 6.8. Furthermore assume $U$ to be $\mathcal{D}$-pullback flattening w.r.t. $d_{X,p.i.p.}$ (in the sense of Definition 6.4).

Then the conclusions of Theorem 6.9 hold, i.e., $\hat{A} : \mathbb{R} \rightrightarrows X$, $t \mapsto \Lambda(\hat{B}, t, d_{X,p.i.p.})$ is a pullback $\mathcal{D}$-attractor w.r.t. $d_{X,p.i.p.}$. Furthermore, $\hat{A}$ is the unique element from $\mathcal{D}$ with these properties. In addition, if $U$ is a strict MNDS, then $\hat{A}$ is strictly invariant.

7. Proofs about general existence of pullback attractors (stated in § 6).

**Proof of Lemma 6.9** Let $U$ be a MNDS on a metric space $(X, d_X)$ and $\mathcal{D}$ be a universe such that $U$ is $\mathcal{D}$-pullback flattening (in the sense of Definition 6.4).

Fix $t \in \mathbb{R}$ and $\hat{B} = \{B(s) : s \in \mathbb{R}\}$ in $\mathcal{D}$ arbitrarily. For each sequence $(t_n)_{n \in \mathbb{N}}$ in $[0, \infty)$ with $t_n \to \infty$, it remains to show that any sequence $(y_n)_{n \in \mathbb{N}}$ with $y_n \in U(t, t - t_n, B(t - t_n))$ has a convergent subsequence. Due to [77] Ch. 0.2 Theorem on page 13, it is sufficient to verify a subsequence $(y_{n_k})_{k \in \mathbb{N}}$ such that the set $\{y_{n_k} : k \in \mathbb{N}\} \subset X$ is totally bounded, i.e., for every $\rho > 0$, the set $\{y_{n_k} : k \in \mathbb{N}\}$ which can be covered by finitely many balls of radius $\rho$ w.r.t. $d_X$.

For each index $k \in \mathbb{N}$, there exist some $T_k > 0$ and a nonempty compact subset $M_k \subset X$ with
\[
\text{dist}_X(U(t, s, B(s)), M_k) \leq \frac{1}{k} \quad \text{for every } s \leq t - T_k.
\]
We can select a monotone sequence of indices \( n_k \to \infty \) with \( t_{n_k} \geq \max\{T_1, \ldots, T_k\} \) for every \( k \in \mathbb{N} \). Consider the subsequence \((y_{n_k})_{k \in \mathbb{N}}\).

Fixing \( \rho > 0 \) arbitrarily, there is some \( K = K(\rho) \in \mathbb{N} \) with \( \frac{1}{K} < \frac{\rho}{2} \). The compact subset \( M_K \subset X \) can be covered by finitely many open balls of radius \( \frac{\rho}{2} \) (w.r.t. \( d_X \)). Thus the union of open balls with the same respective centers and radius \( \rho \) covers \{\( y_{n_k} \mid k \geq K \}\} \subset \bigcup_{s \leq s_0} U(t, s, B(s)) \) with \( s_0 := t - \max\{T_1, \ldots, T_K\} \). Obviously, the set \{\( y_{n_k} \mid k \in \mathbb{N} \)\} can also be covered by finitely many open balls of radius \( \rho \).

**Proof of Lemma 6.2.** We follow essentially the arguments of [8, Lemma 3.2]. Consider any sequence \( y_n \in U(t, t - t_n, B(t - t_n)) \) with \( t_n \to +\infty \). As \( U \) is pullback asymptotically compact w.r.t. \( \hat{B} \) and \( d_{X,p.i.p.} \), there exists a subsequence converging w.r.t. \( d_{X,p.i.p.} \) and, its limit \( y \) belongs to \( \Lambda(\hat{B}, t, d_{X,p.i.p.}) \), i.e., \( \Lambda(\hat{B}, t, d_{X,p.i.p.}) \) is non-empty.

We now prove that \( \Lambda(\hat{B}, t, d_{X,p.i.p.}) \) is compact in \( (X, d_{X,p.i.p.}) \). For any sequence \((y_n)_{n \in \mathbb{N}} \in \Lambda(\hat{B}, t, d_{X,p.i.p.}) \), there exist sequences \( t_n \to +\infty \) and \( z_n \in U(t, t - t_n, B(t - t_n)) \) with \( d_{X,p.i.p.}(y_n, z_n) < \frac{\rho}{n} \) for each \( n \in \mathbb{N} \). Now the pullback asymptotic compactness of \( U \) again implies the existence of a subsequence \((z_{n_k})_{k \in \mathbb{N}} \) whose limit \( z \) w.r.t. \( d_{X,p.i.p.} \) is also contained in the closure \( \Lambda(\hat{B}, t, d_{X,p.i.p.}) \). Then, \( d_{X,p.i.p.}(y_{n_k}, z) \to 0 \) for \( k \to \infty \), so that \( \Lambda(\hat{B}, t, d_{X,p.i.p.}) \subset X \) is compact w.r.t. \( d_{X,p.i.p.} \).

The claimed limit in (23) is now proved by contradiction. If (23) does not hold, then there exist \( \varepsilon > 0 \) and sequences \( t_n \to +\infty \), \( y_n \in U(t, t - t_n, B(t - t_n))\) such that

\[
\text{dist}_{X,p.i.p.} \left( y_n, \Lambda(\hat{B}, t, d_{X,p.i.p.}) \right) \overset{\text{def}}{=} \inf \left\{ d_{X,p.i.p.}(y_n, \xi) \mid \xi \in \Lambda(\hat{B}, t, d_{X,p.i.p.}) \right\}
\]

is larger than \( \varepsilon \). As \( U \) is pullback asymptotically compact w.r.t. \( \hat{B} \) and \( d_{X,p.i.p.} \), there exists a subsequence \((y_{n_k})_{k \in \mathbb{N}} \) whose limit w.r.t. \( d_{X,p.i.p.} \) is also contained in the closed set \( \Lambda(\hat{B}, t, d_{X,p.i.p.}) \), but this contradicts the preceding lower distance bound \( \varepsilon \).

For proving the negative invariance (24) of \( \Lambda(\hat{B}, t, d_{X,p.i.p.}) \), we assume in addition:

1. For each \((t, \tau) \in \mathbb{R}^2_+ \), \( U(t, \tau, \cdot) : (X, d_X) \to (X, d_{X,p.i.p.}) \) has closed graph,
2. \( U \) has the pullback \( \omega \)-Mazur property w.r.t. \( \hat{B}, d_X, d_{X,p.i.p.} \) (in the sense of Definition 6.7).

Fix \((t, \tau) \in \mathbb{R}^2_+ \) and \( y \in \Lambda(\hat{B}, t, d_{X,p.i.p.}) \) arbitrarily. Then, by definition (19), there exist sequences \( t_n \to +\infty \), \( y_n \in U(t, t - t_n - \tau, x_n) \), \( x_n \in B(t - (t_n - \tau)) \) with \( d_{X,p.i.p.}(y_n, y) \to 0 \). For all \( t_n \geq t \), the standard process property implies

\[
U(t, t - t_n + \tau, x_n) \subset U(t, \tau, U(\tau, t - t_n + \tau, x_n)),
\]

and then \( y_n \in U(t, \tau, z_n) \) with some \( z_n \in U(\tau, t - t_n + \tau, x_n) \). As before, the pullback asymptotic compactness w.r.t. \( \hat{B} \) and \( d_{X,p.i.p.} \) ensures a subsequence (again denoted by) \((z_{n_k})_{k \in \mathbb{N}} \) which converges w.r.t. \( d_{X,p.i.p.} \). Its limit \( z \) is contained in \( \Lambda(\hat{B}, \tau, d_{X,p.i.p.}) \).

Assuming the pullback \( \omega \)-Mazur property w.r.t. \( \hat{B}, d_X \) and \( d_{X,p.i.p.} \), this element \( z \) is related to (possibly different) sequences \( \tilde{t}_k \to +\infty \), \( \tilde{x}_k \in U(\tau, t - (\tilde{t}_k - \tau), \tilde{x}_k) \).
and \( \tilde{x}_k \in B(t - (\tilde{t}_k - \tau)) \) with

\[
\begin{align*}
    d_X(\tilde{z}_k, z) &\to 0 \quad (k \to 0), \\
    \liminf_{n \to \infty} \inf_{k \geq n} \text{dist}_{\text{X,p.i.p.}}(U(t, \tau, z_n), U(t, \tau, \tilde{z}_k)) &= 0.
\end{align*}
\]

Considering appropriate subsequences (with the same notation) instead, we obtain index sequences \( n_j \nearrow \infty, k_j \nearrow \infty \) such that

\[
d_X(\tilde{z}_{k_j}, z) \to 0, \quad \text{dist}_{\text{X,p.i.p.}}(y_{n_j}, U(t, \tau, \tilde{z}_{k_j})) \to 0 \quad \text{for } j \to \infty.
\]

Finally the graph of \( U(t, \tau, \cdot) : (X, d_X) \sim (X, d_{\text{X,p.i.p.}}) \) is supposed to be closed and so, the respective limits for \( j \to \infty \) reveal

\[\begin{align*}
y \in U(t, \tau, z) \subset U(t, \tau, \Lambda(\hat{B}, \tau, d_{\text{X,p.i.p.}})).
\end{align*}\]

\[\square\]

**Proof of Theorem 6.9.** As in the proof of [8, Theorem 3.3], we first verify pullback \( \mathcal{D} \)-attraction, i.e., for every \( \hat{D} \in \mathcal{D}, \)

\[
\lim_{\tau \to +\infty} \text{dist}_{\text{X,p.i.p.}}(U(t, t - \tau, D(t - \tau)), A(t)) = 0.
\]

Indeed, thanks to the limit (23) in Lemma 6.8, for every \( \varepsilon > 0 \) and \( t \in \mathbb{R} \), there exists \( T_1(t, \varepsilon) > 0 \) such that for \( \tau \geq T_1(t, \varepsilon) \),

\[
\text{dist}_{\text{X,p.i.p.}}(U(t, t - \tau, B(t - \tau)), A(t)) < \varepsilon.
\]

\( \hat{B} \in \mathcal{D} \) is assumed to be pullback \( \mathcal{D} \)-absorbing and so, every \( \hat{D} \in \mathcal{D} \) has some \( T_2 = T_2(t - \tau, \hat{D}) > 0 \) satisfying

\[
U(t - \tau, t - \tau - T_2, D(t - \tau - T_2)) \subset B(t - \tau).
\]

Hence we have for all \( \tau > 0 \) sufficiently large

\[
\text{dist}_{\text{X,p.i.p.}}(U(t, t - \tau, D(t - \tau)), A(t)) < \varepsilon.
\]

Condition (iii) in Definition 6.1, i.e., negative invariance of pullback \( \mathcal{D} \)-attractors, follows from inclusion (24). Pullback \( \mathcal{D} \)-absorption of \( \hat{B} \in \mathcal{D} \) implies some \( T_3 = T_3(t, \hat{B}) > 0 \) with

\[
U(t, t - \tau, B(t - \tau)) \subset B(t) \quad \text{for all } \tau \geq T_3,
\]

Hence we have the inclusion \( A(t) \overset{\text{def.}}{=} \Lambda(\hat{B}, t, d_{\text{X,p.i.p.}}) = \Lambda(\hat{B}, t, d_X) \subset B(t) \) for each \( t \in \mathbb{R} \). In particular, \( \hat{A} \in \mathcal{D} \).

This observation leads to the conclusion that \( A \) is unique. Indeed, suppose that \( \hat{A}' \in \mathcal{D} \) is another pullback \( \mathcal{D} \)-attractor w.r.t. \( d_{\text{X,p.i.p.}} \), then as

\[
A'(t) \subset U(t, t - \tau, A'(t - \tau))
\]

and

\[
\lim_{\tau \to +\infty} \text{dist}_{\text{X,p.i.p.}}(U(t, t - \tau, A'(t - \tau)), A(t)) = 0,
\]

we have that \( A'(t) \subset A(t) \) for each \( t \in \mathbb{R} \). Exchanging \( \hat{A} \) and \( \hat{A}' \), it follows that \( \hat{A} = \hat{A}' \).
Finally, assume that the multi-valued non-autonomous dynamical system $U$ is strict (in addition). It remains to show $U(t, r, A(r)) \subset A(t)$ for any real $t \geq r$. Indeed, we have for every $\tau \geq 0$

$$U(t, r, A(r)) \subset U(t, r, U(r, r - \tau, A(r - \tau))) = U(t, r - \tau, A(r - \tau)).$$

As $\hat{A}$ pullback attracts itself w.r.t. $d_{X, p.i.p.}$, it follows from Lemma 6.8 that

$$\lim_{\tau \to +\infty} \text{dist}_{X, p.i.p.}(U(t, r - \tau, A(r - \tau)), A(t)) = 0.$$ Consequently, for $\varepsilon > 0$ fixed arbitrarily, there exists some $T_4(\varepsilon, t, r) > 0$ such that, for all $\tau \geq T_4$

$$\text{dist}_{X, p.i.p.}(U(t, r - \tau, A(r - \tau)), A(t)) < \varepsilon,$$

and as $U(t, r, A(r)) \subset U(t, r - \tau, A(r - \tau))$, we have

$$\text{dist}_{X, p.i.p.}(U(t, r, A(r)), A(t)) < \varepsilon.$$ Since $\varepsilon > 0$ was fixed arbitrarily, we conclude from the set $A(t) \subset X$ being closed w.r.t. $d_{X, p.i.p.}$ that $U(t, r, A(r)) \subset A(t)$, as required. \hfill \Box

8. Proofs about parabolic differential inclusions with delay (stated in §§3 – 5).

8.1. Topological preparations: Space-dependent delay and completeness of $C_\gamma$.

**Lemma 8.1** ([37, Corollary 2.14]). Let $\Omega \subset \mathbb{R}^d$ be bounded and open, and $T \in (0, \infty)$. Every function $v \in L^\infty([0, T]; L^\infty(\Omega))$ with $(t, x) \mapsto v(t, x)$ Lebesgue measurable satisfies

(a) $v \in L^\infty([0, T] \times \Omega)$,

(b) $x \mapsto \text{ess sup}_{0 \leq s \leq T} |v(s, x)|$ is essentially bounded on $\Omega$,

(c) $v(\cdot, x) \in L^\infty([0, T]) \subset \bigcap_{1 \leq p \leq \infty} L^p([0, T])$ for Lebesgue-almost every $x \in \Omega$,

(d) $\text{ess sup}_{0 \leq s \leq T} |v(s, x)| = \lim_{p \to \infty} \left(\frac{1}{T} \int_0^T |v(s, x)|^p \, ds\right)^{\frac{1}{p}}$ for a.e. $x \in \Omega$,

(e) $\text{ess sup}_{0 \leq s \leq T} |v(s, \cdot)| = \lim_{p \to \infty} \left(\frac{1}{T} \int_0^T |v(\cdot, s)|^p \, ds\right)^{\frac{1}{p}}$ in $L^q(\Omega)$ for all $q \geq 1$.

**Proof** of Lemma 8.1. For any given $u_0 \in L^\infty(-\infty, 0; L^2(\Omega))$ and $c > 0$, we first verify that

$$v_0 := \text{ess sup}_{-\Theta(x) \leq s \leq 0} [u_0(s, x)]_c : \Omega \to \mathbb{R}$$

belongs to $L^2(\Omega)$. Indeed, for every $p \in [1, \infty)$ and Lebesgue-almost every $x \in \Omega$, the Lebesgue integral

$$\frac{1}{\Theta(x)} \int_{-\Theta(x)}^0 |[u_0(s, x)]_c|^p \, ds$$

is well defined and has some value in $[0, c^p]$. For the same reasons as in Lemma 8.1, the function

$$\Omega \to \mathbb{R}, \quad x \mapsto \left(\frac{1}{\Theta(x)} \int_{-\Theta(x)}^0 |[u_0(s, x)]_c|^p \, ds\right)^{\frac{1}{p}}$$
converges to \( v_0 \) Lebesgue-almost everywhere in \( \Omega \) for \( p \to \infty \). Due to the joint \( L^\infty \) bound, we conclude from Lebesgue’s theorem of dominated convergence that \( v_0 \in L^\infty(\Omega) \subset L^2(\Omega) \) and the convergence even holds w.r.t. \( L^2(\Omega) \).

The continuity of \((C_\gamma, d_{c,\gamma,p.i.p.}) \to L^2(\Omega)\) remains to be verified. Choose any \( u_0 \in C_\gamma \) and a sequence \((u_k)_{k \in \mathbb{N}} \) in \( C_\gamma \) converging to \( u_0\) w.r.t. \( d_{c,\gamma,p.i.p.}. \) Set
\[
\begin{align*}
v_0 &:= \text{ess sup}_{-\Theta(x) \leq s \leq 0} [u_0(s,x)]_c, \\
v_k &:= \text{ess sup}_{-\Theta(x) \leq s \leq 0} [u_k(s,x)]_c \quad (k \in \mathbb{N}).
\end{align*}
\]

For each \( \varepsilon > 0 \), there exists some \( \bar{T} = \bar{T}(\Theta, \varepsilon) > 0 \) such that \( \Omega_{\bar{T}} := \{ x \in \Omega \mid \Theta(x) \leq \bar{T} \} \) satisfies \( 2c \cdot \mathcal{L}^d(\Omega \setminus \Omega_{\bar{T}}) < \frac{\varepsilon}{2} \) due to \( \Theta \in L^1(\Omega) \). (This next step reveals the key technical advantage of space-dependent delay.) Moreover, there exists some \( K = K(\varepsilon) \in \mathbb{N} \) with \( d_{c,\gamma,0+\bar{T}}(u_k, u) < \frac{\varepsilon}{2} \) for all \( k \geq K \). Choosing any integer \( \bar{\tau} = \bar{\tau}(\varepsilon) \) with \( \bar{\tau} > \Theta + \bar{T} \), we obtain for all indices \( k \geq K \)
\[
\begin{align*}
\|v_k - v\|_{L^2(\Omega)} &= \| \text{ess sup}_{-\Theta(x) \leq s \leq 0} [u_k(s,x)]_c - \text{ess sup}_{-\Theta(x) \leq s \leq 0} [u(s,x)]_c \|_{L^2(\Omega)} \\
&\leq \| \text{ess sup}_{-\Theta(x) \leq s \leq 0} [u_k(s,x)]_c - [u(s,x)]_c \|_{L^2(\Omega)} \\
&\leq \| \text{ess sup}_{\Theta(x) \leq s \leq 0} [u_k(s,x)]_c - [u(s,x)]_c \|_{L^2(\Omega)} + 2c \cdot \mathcal{L}^d(\Omega \setminus \Omega_{\bar{T}}) + \varepsilon \\
&\leq d_{c,\gamma,\bar{\tau}}(u_k, u) + \frac{\varepsilon}{2}.
\end{align*}
\]

Finally the assumption \( d_{c,\gamma,p.i.p.}(u_k, u) \to 0 \) \((k \to \infty)\) implies \( \|v_k - v\|_{L^2(\Omega)} \leq d_{c,\gamma,\bar{\tau}}(u_k, u) + \frac{\varepsilon}{2} \leq \varepsilon \) for all \( k \in \mathbb{N} \) sufficiently large. \( \square \)

**Proof of Lemma 5.4** For any \( \tau > 0 \), let \( (u_k)_{k \in \mathbb{N}} \) be a Cauchy sequence in \( C_\gamma \) w.r.t. the metric \( d_{c,\gamma,\tau} \). Obviously \( (u_k)_{k \in \mathbb{N}} \) is also a Cauchy sequence in the normed vector space \((C_\gamma, \| \cdot \|_s)\). Restricting our considerations to any bounded interval \([-T, 0] \) for a moment, the sequence of restrictions \((u_k)_{[-T,0]} \) is a Cauchy sequence in \((C^0([-T, 0], L^2(\Omega)), \| \cdot \|_{L^2(\Omega)}) \) and so, we can find a unique continuous function \( u : (-\infty, 0] \to (L^2(\Omega), \| \cdot \|_{L^2(\Omega)}) \) such that for every \( T > 0 \),
\[
\sup_{-T \leq s \leq 0} (e^{\gamma s} \cdot \| u_k(s) - u(s) \|_{L^2(\Omega)}) \to 0 \quad (k \to \infty).
\]

Moreover, \( u \in C_\gamma \) and \( \| u_k - u \|_\gamma \to 0 \) \((k \to \infty)\). Indeed, for each index \( k \in \mathbb{N} \), the limit of \( e^{\gamma s} u_k(s) \) for \( s \to -\infty \) exists in \( L^2(\Omega) \) and the Cauchy property of \((u_j)_{j \in \mathbb{N}} \) w.r.t. \( \| \cdot \|_\gamma \) ensures that it does not depend on \( k \in \mathbb{N} \), i.e., there exists \( v \in L^2(\Omega) \) with
\[
\| e^{\gamma s} u_k(s) - v \|_{L^2(\Omega)} \to 0 \quad (s \to -\infty)
\]
for all \( k \in \mathbb{N} \). Fix any \( \varepsilon > 0 \). There exist \( K = K(\varepsilon) \in \mathbb{N}, \sigma = \sigma(\varepsilon, K) < 0 \) such that for all \( j \geq K \),
\[
\begin{align*}
\| u_j - u_K \|_\gamma &\overset{\text{Def}}{=} \sup_{s \leq 0} (e^{\gamma s} \| u_j(s) - u_K(s) \|_{L^2(\Omega)}) < \varepsilon, \\
\sup_{s \leq \sigma} \| v(s) - e^{\gamma s} u_K(s) \|_{L^2(\Omega)} &< \varepsilon
\end{align*}
\]
and so, we obtain for all \( s \leq \sigma, j \geq K \)
\[
\|v(s) - e^{\gamma s} u(s)\|_{L^2(\Omega)} \\
\leq \|v(s) - e^{\gamma s} u_k(s)\|_{L^2(\Omega)} + \|e^{\gamma s} u_k(s) - e^{\gamma s} u_j(s)\|_{L^2(\Omega)} + \|e^{\gamma s} u_j(s) - e^{\gamma s} u(s)\|_{L^2(\Omega)} \\
\leq \varepsilon + \varepsilon + \|e^{\gamma s} u_j(s) - e^{\gamma s} u(s)\|_{L^2(\Omega)}.
\]

The limit for \( j \to \infty \) reveals \( \|v(s) - e^{\gamma s} u(s)\|_{L^2(\Omega)} \leq 2\varepsilon \), for each \( s \leq \sigma \), i.e., \( v \) is also the limit of \( e^{\gamma s} u(s) \) for \( s \to -\infty \). Hence, \( u \in C_{\gamma} \). Since the index \( K \) depends only on \( \varepsilon \), we have \( \|u_k - u\|_{\gamma} \to 0 \).

It is worth mentioning that this unique limit function \( u : (-\infty, 0] \to L^2(\Omega) \) does not depend on the parameter \( \tau \) and so, limits of Cauchy sequences never depend on \( \tau \). In particular, the conclusions about \( d_{c,\gamma,\tau} \) with any \( \tau > 0 \) also hold for the joint metric \( d_{c,\gamma,p,i.p.} \) as well as the “strong” metric \( d_{c,\gamma} \).

It remains to prove for \( \tau > 0 \) fixed initially that \( (u_k)_{k\in\mathbb{N}} \) converges to \( u \) w.r.t. \( d_{c,\gamma,\tau} \). \( ([u_k])_{k\in\mathbb{N}} \) is a Cauchy sequence in the Lebesgue-Bochner space \( L^2(\Omega, L^\infty([-\tau, 0])) \) which is complete \( 23 \) § II.2.

Thus there is a unique limit \( v \in L^2(\Omega, L^\infty([-\tau, 0])) \) and, we will show
\[
[u]_c = v \quad \text{Lebesgue-almost everywhere in } [-\tau, 0] \times \Omega.
\]

Indeed, the embeddings \( L^2(\Omega, L^\infty([-\tau, 0])) \hookrightarrow L^2([-\tau, 0] \times \Omega), C^0([-\tau, 0], L^2(\Omega)) \hookrightarrow L^2([-\tau, 0] \times \Omega) \) are continuous since \( \Omega \subset \mathbb{R}^n \) is bounded. As a first consequence, there is a subsequence \( ([u_k])_{k\in\mathbb{N}} \) which converges to \( v \) Lebesgue-almost everywhere in \([0, 0] \times \Omega\). Secondly we can select a further subsequence (again denoted by) \( ([u_k])_{k\in\mathbb{N}} \) converging to \( u \) Lebesgue-almost everywhere in \([0, 0] \times \Omega\). Then \(([u_k])_{k\in\mathbb{N}} \) converges to \( [u]_c \) Lebesgue-almost everywhere in \([-\tau, 0] \times \Omega \) and so, \( [u]_c = v \) holds Lebesgue-almost everywhere in \([-\tau, 0] \times \Omega \).

8.2. Proof of Proposition 5.3 about the MNDS \( U \) related to evolution inclusion \( 9 \). According to Corollary 3.2, every state \( \phi \in C_{\gamma} \) at time \( t_0 \in \mathbb{R} \) initializes at least one strong solution to parabolic differential inclusion \( 1 \), with space-dependent delay. Hence, the set-valued mapping \( U \) specified in Proposition 5.3 satisfies \( U(t, t_0, \phi) \neq \emptyset \).

Lemma 8.2. The multi-valued mapping \( U : \mathbb{R}_+^2 \times C_{\gamma} \rightharpoonup C_{\gamma} \) from Proposition 5.3 satisfies the 2-parameter semigroup property, i.e., for any real \( t_0 \leq t_1 \leq t_2 \) and \( \phi \in C_{\gamma} \),
\[
U(t_2, t_0, \phi) = U(t_2, t_1, U(t_1, t_0, \phi)).
\]

Proof. First we consider the inclusion \( U(t, t_0, \phi) \subset U(t, s, U(s, t_0, \phi)) \) for any \( t_0 \leq s \leq t \) and \( \phi \in C_{\gamma} \), for each \( \chi \in U(t, t_0, \phi) \), there is a solution \( y_t \) related to a selector \( f \) such that \( \chi = y_t(\cdot; t_0, \phi) \) with
\[
\chi(\theta) = y_t(\theta; t_0, \phi) = S(t - t_0 + \theta) \phi(0) + \int_{t_0}^{t+\theta} S(t + \theta - \sigma) f(\sigma) \, d\sigma \\
= S(t - s + \theta) S(s - t_0) \phi(0) + \int_{t_0}^{s} S(t - s + \theta) S(s - \sigma) f(\sigma) \, d\sigma.
\]
\[
\int_s^{t+\theta} S(t+\theta-\sigma) f(\sigma) \, d\sigma
\]
\[
= S(t-s+\theta) \left\{ S(s-t_0) \phi(0) + \int_{t_0}^s S(s-\sigma) f(\sigma) \, d\sigma \right\} + \\
\int_s^{t+\theta} S(t+\theta-\sigma) f(\sigma) \, d\sigma
\]
\[
= S(t-s+\theta) \, y_s(0; t_0, \phi) + \int_s^{t+\theta} S(t+\theta-\sigma) f(\sigma) \, d\sigma
\]

for \( \theta \in [-t-s, 0] \), i.e., \( \chi(\theta) = y_t(\theta; s, y_s(\cdot; t_0, \phi)) \) holds for every \( \theta \in [-t-s, 0] \).

On the other hand, if \( \theta \in (-(\infty, -(t-s)) \), then \( y_s(t+\theta-s; t_0, \phi) = y_t(\theta; t_0, \phi) \)
due to the joint selector \( f \) and so, \( \chi = y_t(\cdot; s, y_s(\cdot; t_0, \phi)) \). This implies \( \chi \in U(t, s, y_s(\cdot; t_0, \phi)) \subset U(t, s, U(s, t_0, \phi)) \) and, consequently,
\[
U(t, t_0, \phi) \subset U(t, s, U(s, t_0, \phi)).
\]

For the opposite inclusion consider any \( \chi \in U(t, s, U(s, t_0, \phi)) \). Then, there exist \( \chi_1 \in U(s, t_0, \phi) \) with \( \chi_1 = y_t^1(\cdot; t_0, \phi) \) for a solution \( y_t^1(\cdot; t_0, \phi) \) corresponding to a selector \( f_1 \) and a solution \( y_s^2(\cdot; s, \chi_1) \) corresponding to a selector \( f_2 \) such that \( \chi = y_t^2(\cdot; s, \chi_1) = y_t^2(\cdot; s, y_s^1(\cdot; t_0, \phi)) \). Concatenating these solutions, gives
\[
u_\sigma = \begin{cases} 
y_1^1(\cdot; t_0, \phi) & \text{if } \sigma < s, \\
y_2^2(\cdot; s, y_s^1(\cdot; t_0, \phi)) & \text{if } s \leq \sigma,
\end{cases}
\]
which is the mild solution of \( F \) corresponding to the selector
\[
f(\sigma) = \begin{cases} 
f_1(\sigma) \in \mathcal{G}(\sigma, y_1^1) = \mathcal{G}(\sigma, u_\sigma) & \text{if } \sigma < s \\
f_2(\sigma) \in \mathcal{G}(\sigma, y_2^2) = \mathcal{G}(\sigma, u_\sigma) & \text{if } s \leq \sigma.
\end{cases}
\]
Hence \( \chi = u_\epsilon = y_t(\cdot; t_0, \phi) \), which means that \( \chi \in U(t, t_0, \phi) \).

In regard to conclusions about closed values of \( U \), we require the Nemyskii operator adapted to the arguments on the right-hand side of evolution problem \( [8] 
- similarly to \( [32] \) \( \S 2.4 \).

**Proposition 8.3** (Set-valued Nemyskii operator \([32] \text{ Ch. 2 Theorem 7.26}\)). Let \((\tilde{\Omega}, \Sigma, \mu)\) be a \( \sigma \)-finite measure space and \( E_1, E_2 \) separable Banach spaces. For \( 1 \leq p, q < \infty \) and the set-valued map \( F : \tilde{\Omega} \times E_1 \rightrightarrows E_2 \) suppose:

(i) All values of \( F \) are nonempty closed subsets of \( E_2 \).

(ii) The graph of \( F \) is a measurable subset of \( \tilde{\Omega} \times E_1 \times E_2 \). (In particular, this is satisfied if the set-valued map \( F \) is measurable \([32] \text{ Ch. 2 Proposition 1.7}\).)

(iii) For every \( \omega \in \tilde{\Omega} \), the set-valued map \( F(\omega, \cdot) : E_1 \rightrightarrows E_2 \) is Hausdorff upper semi-continuous.

(iv) There exist \( \alpha \in L^q(\tilde{\Omega}) \) and a constant \( c > 0 \) such that \( \|F(\omega, z)\|_{E_2} \leq \alpha(\omega) + c \cdot \|z\|^{q^*}_{E_2} \) holds for \( \mu \)-almost all \( \omega \in \tilde{\Omega} \) and every \( z \in E_1 \).

Then, the set-valued map \( \mathbb{F} : L^p(\tilde{\Omega}, E_1) \rightrightarrows L^q(\tilde{\Omega}, E_2) \) defined by
\[
\mathbb{F}(u) := \{ v \in L^q(\tilde{\Omega}, E_2) \mid v(\omega) \in F(\omega, u(\omega)) \subset E_2 \text{ for } \mu\text{-a.e. } \omega \in \tilde{\Omega} \}
\]
is Hausdorff upper semi-continuous, i.e., for every \( u_0 \in L^p(\tilde{\Omega}, E_1) \) and \( \varepsilon > 0 \), there exists a radius \( \rho > 0 \) such that every \( u \in L^p(\tilde{\Omega}, E_1) \) with \( \|u - u_0\|_{L^p(\tilde{\Omega}, E_1)} < \rho \) satisfies \( \mathbb{F}(u) \subset \mathbb{B}_\varepsilon(\mathbb{F}(u_0)) \subset L^q(\tilde{\Omega}, E_2) \).
This general statement has already laid the basis for the following conclusions in [37]:

**Corollary 8.4 ([37 Corollary 2.22]).** Let \( \Omega \subset \mathbb{R}^d \) be a bounded open set with smooth boundary. Suppose that hypotheses (G1) - (G5) hold for the set-valued mapping \( G : \mathbb{R} \times \Omega \times \mathbb{R} \times \mathbb{R} \times L^2(\Omega) \rightrightarrows \mathbb{R} \). Consider the set-valued mapping \( \mathcal{G} : \mathbb{R} \times (L^2(\Omega))^3 \rightrightarrows L^2(\Omega) \) defined by

\[
(t; u_0, v_0, w_0) \mapsto \{ \xi \in L^2(\Omega) \mid \xi(x) \in G(t; x; u_0(x), v_0(x), w_0(\cdot)) \subset \mathbb{R} \}
\]

for Lebesgue-almost every \( x \in \Omega \). (26)

Then, the set-valued mapping \( \mathcal{G} \) has the following properties:

(G1) Each value of \( \mathcal{G} \) is a nonempty bounded closed convex subset of \( H := L^2(\Omega) \).

(G2) \( \mathcal{G}(t; \cdot, \cdot, \cdot) : H^3 \rightrightarrows H \) is Hausdorff upper semi-continuous for every \( t \in [0, T] \).

(G3) \( \mathcal{G}(\cdot; u_0, v_0, w_0) : [0, T] \rightrightarrows H \) has a Lebesgue measurable selection for every \( (u_0, v_0, w_0) \in H^3 \).

(G4) \( \mathcal{G} \) satisfies a linear growth condition: There are \( \alpha, \beta \geq 0 \) such that for all \( u, v, w \in H := L^2(\Omega) \) and Lebesgue-almost all \( t \in [0, T] \),

\[
\sup \{ \| \eta \|_{L^2(\Omega)} \mid \eta \in \mathcal{G}(t; u, v, w) \} \leq \alpha \cdot L^d(\Omega)^{1/2} + \beta \cdot (\| u \|_{L^2(\Omega)} + \| v \|_{L^2(\Omega)}).
\]

**Lemma 8.5.** Let \( \Omega \subset \mathbb{R}^d \) be a bounded open set with smooth boundary. Suppose hypotheses (G1) - (G5) for the set-valued mapping \( G : \mathbb{R} \times \Omega \times \mathbb{R} \times \mathbb{R} \times L^2(\Omega) \rightrightarrows \mathbb{R} \). Fix \( \theta > 0 \), \( \Theta \in L^1(\Omega) \) with \( \Theta \geq 0 \).

Then for each \( t_0 \in \mathbb{R} \), the mapping \( \mathcal{G}(t_0; \cdot) : (C_7, d_{c,\gamma}) \rightrightarrows L^2(\Omega) \) specified in (10) is Hausdorff upper semi-continuous with nonempty convex closed values.

**Proof.** Fix \( t_0 \in \mathbb{R} \) arbitrarily. The claimed Hausdorff upper semi-continuity of \( \mathcal{G}(t_0; \cdot) : (C_7, d_{c,\gamma}) \rightrightarrows L^2(\Omega) \) results directly from Corollary 8.4 since the continuity mentioned in Lemma [33] implies the continuity of the single-valued function

\[
(C_7, d_{c,\gamma}) \rightarrow L^2(\Omega)^3, \quad \varphi \mapsto (\varphi(0), \text{ess sup}_{-\Theta \leq s \leq 0} [\varphi(s, \cdot)]_{c,\varphi(0)}).
\]

Lemma 8.6 (Standard convergence theorem for differential inclusions, e.g., [13 Theorem VI-4]). Let \( X \) be a Hausdorff locally convex space, \( Y \) be a separable Banach space and \( T > 0 \). Suppose for the set-valued map \( \mathcal{F} : [0, T] \times X \rightrightarrows Y \) and the sequences \( (x_j(\cdot))_{j \in \mathbb{N}}, (y_j(\cdot))_{j \in \mathbb{N}} \) in \( L^1(0, T; Y) \):

(i) Each value of \( \mathcal{F} \) is a nonempty closed convex subset of \( Y \).

(ii) For Lebesgue-a.e. \( t \in [0, T] \), \( \mathcal{F}(t, \cdot) : X \rightrightarrows Y \) is Hausdorff upper semi-continuous.

(iii) \( (x_j(\cdot))_{j \in \mathbb{N}} \) converges to some \( x(\cdot) : [0, T] \rightarrow X \) Lebesgue-almost everywhere.

(iv) \( (y_j(\cdot))_{j \in \mathbb{N}} \) converges to some \( y(\cdot) \) weakly in \( L^1(0, T; Y) \).

(v) \( \text{dist}(y_j(t), \mathcal{F}(t, x_j(t))) \rightarrow 0 \) as \( j \rightarrow \infty \) for Lebesgue-almost every \( t \in [0, T] \),

Then, \( y(t) \in \mathcal{F}(t, x(t)) \) holds for Lebesgue-almost every \( t \in [0, T] \).

**Lemma 8.7.** For every \( (t, t_0) \in \mathbb{R}^2_+ \), the set values of \( U(t, t_0, \cdot) : X \rightrightarrows X \) (specified in Proposition 5.3 in terms of parabolic differential inclusion [13]) are closed w.r.t. both \( d_{c,\gamma} \) and \( d_{c,\gamma,T} \) for each \( \tau > t - t_0 \).
Proof. Fix $\phi \in C_\gamma$, $(t, t_0) \in \mathbb{R}^2_+$ and $\tau > t - t_0$ arbitrarily and consider a sequence $(u^{(k)})_{k \in \mathbb{N}}$ in $C_\gamma$ with $u^{(j)} \in U(t, t_0, \phi)$ for each index $j \in \mathbb{N}$ and 

$$d_{c, \gamma, \tau}(u^{(k)}, u) \to 0 \quad (k \to \infty).$$

Due to (16), each $u^{(k)} \in C_\gamma$ is represented by means of a selector $f_k : [t_0, t] \to L^2(\Omega)$ with $f_k(s) \in \mathcal{G}(s, u^{(k)}(\cdot + s - t)|_{(-\infty, 0]})$ for a.e. $s \in [t_0, t]$:

$$u^{(k)}(s) = \begin{cases} S(t - t_0 + s) \phi(0) + \int_{t_0}^{t+s} S(t + s - \tau) f_k(\tau) \, d\tau & \text{if } s \in [-t - t_0, 0] \\ \phi(t - t_0 + s) & \text{if } s < -t - t_0. \end{cases}$$

In particular, we obtain $u(s) = \phi(t - t_0 + s)$ for every $s < -(t - t_0)$ and so, the assumption $\tau > t - t_0$ leads to the convergence w.r.t. $d_{c, \gamma}$, i.e., $d_{c, \gamma}(u^{(k)}, u) \to 0$ $(k \to \infty)$. Assumption (G5) about the linear growth of $G$ implies

$$\|f_k(s)\|_{L^2(\Omega)} \leq \left(\alpha + \beta c \right) \cdot L^n(\Omega)^{1/2} + \beta \cdot \|u^{(k)}(- (t - s))\|_{L^2(\Omega)}$$

for Lebesgue-almost every $s \in [t_0, t]$ and so, $(f_k)_{k \in \mathbb{N}}$ is bounded in $L^\infty(t_0, t; L^2(\Omega))$ due to

$$\sup_{t_0 \leq s \leq t} \|u^{(k)}(-(t-s)) - u(-(t-s))\|_{L^2(\Omega)} = \sup_{-(t-t_0) \leq \sigma \leq 0} \|u^{(k)}(\sigma) - u(\sigma)\|_{L^2(\Omega)} \leq e^{\gamma(t-t_0)} \cdot d_{c, \gamma, \tau}(u^{(k)}, u) \leq \text{const.}$$

As a consequence of Alaoglu’s theorem (see, e.g., [57, Ch.V.2, Theorem 1]), there exists a sequence of indices $k_j \nearrow \infty$ such that $(f_{k_j})_{j \in \mathbb{N}}$ converges to some $f \in L^2(t_0, t; L^2(\Omega))$ weakly in $L^\infty(t_0, t; L^2(\Omega))$. Lemmas 8.5 and 8.6 lead to $f(s) \in \mathcal{G}(s, u(\cdot + s - t)|_{(-\infty, 0]})$ for Lebesgue-almost every $s \in [t_0, t]$.

It remains to show $u \in U(t, t_0, \phi)$ or, in other terms,

$$u(s) = S(t - t_0 + s) \phi(0) + \int_{t_0}^{t+s} S(t + s - \tau) f(\tau) \, d\tau$$

for every $s \in [-t - t_0, 0]$. In fact we have for each $w \in L^2(\Omega)^* = L^2(\Omega)$ by means of the adjoint $C_0$ semigroup [24 § I.5.14] (see [3] supplementarily)

$$\langle u^{(k_j)}(s), w \rangle_{L^2(\Omega)} = \langle S(t - t_0 + s) \phi(0), w \rangle_{L^2(\Omega)} + \int_{t_0}^{t+s} \langle S(t + s - \tau) f_{k_j}(\tau), w \rangle_{L^2(\Omega)} \, d\tau$$

$$= \langle S(t - t_0 + s) \phi(0), w \rangle_{L^2(\Omega)} + \int_{t_0}^{t+s} \langle f_{k_j}(\tau), S(t + s - \tau)^* w \rangle_{L^2(\Omega)} \, d\tau$$

$$\longrightarrow \langle S(t - t_0 + s) \phi(0), w \rangle_{L^2(\Omega)} + \int_{t_0}^{t+s} \langle f(\tau), S(t + s - \tau)^* w \rangle_{L^2(\Omega)} \, d\tau$$

as $j \to \infty$ due to $f_{k_j} \rightharpoonup f$ in $L^2(t_0, t; L^2(\Omega))$. Hence $U(t, t_0, \phi) \subset C_\gamma$ is closed w.r.t. $d_{c, \gamma, \tau}$ for every $\tau > t - t_0$. This implies closedness w.r.t. both $d_{c, \gamma}$ and $d_{c, \gamma, \text{p.i.p.}}$. $
$}

8.3. Closed graphs of $U(t, t_0, \cdot)$ with respect to various metrics on $C_\gamma$.

In this and the following subsections, $U : \mathbb{R}^2_+ \times C_\gamma \rightharpoonup \gamma$ always denotes the strict multi-valued non-autonomous dynamical system induced by the parabolic differential inclusion [15] as specified in the mild representation [16]. The coefficient
mapping \( G : \mathbb{R} \times \Omega \times \mathbb{R} \times \mathbb{R} \times L^2(\Omega) \mapsto \mathbb{R} \) is supposed to fulfill hypotheses (G1) – (G5). Fix \( \theta > 0 \) and \( \Theta \in L^1(\Omega) \) with \( \Theta \geq 0 \). Let \( \gamma > 0 \) be larger than the smallest eigenvalue \( \lambda_1 > 0 \) of the negative Laplacian operator with homogeneous Dirichlet boundary conditions in \( \Omega \) – as in Proposition \ref{5.3} The metrics \( d_{c,\gamma} \), \( d_{c,\gamma,p.i.p.} \) and \( d_{c,\gamma,p.i.p.} \) on \( C_\gamma \) are defined in \( \S\ref{5.1} \).

From now on, our essential goal is to verify in several steps that \( U \) satisfies the assumptions of Theorem \ref{6.9} This implies the main Theorem \ref{4.1}. Finally, the main Theorem \ref{4.5} (about globally infinite delay) can be concluded from (standard) Theorem \ref{6.9}. This implies the main Theorem \ref{4.1}. Finally, the main

**Lemma 8.8.** For each \((t, t_0) \in \mathbb{R}^2\), the graph of the set-valued mapping \( U(t, t_0, \cdot) : (C_\gamma, d_{c,\gamma}) \mapsto (C_\gamma, d_{c,\gamma,p.i.p.}) \) is closed.

**Proof.** We use essentially the same arguments as in the proof of Lemma \ref{8.7} but now take a converging sequence for the argument of \( U(t, t_0, \cdot) \) into additional consideration. Choose any sequences \((\phi(k))_{k \in \mathbb{N}}, (u(k))_{k \in \mathbb{N}} \) in \( C_\gamma \) with \( u(j) \in U(t, t_0, \phi(j)) \) for each index \( j \in \mathbb{N} \) and

\[
d_{c,\gamma}(\phi(k), \phi) + d_{c,\gamma,p.i.p.}(u(k), u) \rightarrow 0 \quad (k \rightarrow \infty).
\]

Each \( u(k) \in C_\gamma \) is related to a selector \( f_k : [t_0, t] \mapsto L^2(\Omega) \) with

\[
f_k(s) \in \mathbb{G}(s, u(k)(\cdot + s - t)|_{(-\infty, 0]}
\]

for a.e. \( s \in [t_0, t] \) and

\[
u(k)(s) = \begin{cases} 
S(t-t_0+s) \phi(k)(0) + \int_{t_0}^{t+s} S(t+s-\tau) f_k(\tau) \, d\tau & \text{if } s \in [-t-t_0, 0] \\
\phi(k)(t-t_0 + s) & \text{if } s < -(t-t_0).
\end{cases}
\]

Hypothesis (G5) implies

\[
\|f_k(s)\|_{L^2(\Omega)} \leq \text{const}(\alpha, \beta, c, \Omega) \cdot \left(1 + \|u(k)(\cdot - (t - s))\|_{L^2(\Omega)}\right)
\]

for Lebesgue-almost every \( s \in [t_0, t] \) and so, \((f_k)_{k \in \mathbb{N}} \) is bounded in \( L^\infty(t_0, t; L^2(\Omega)) \) due to

\[
\sup_{t_0 \leq s \leq t} \left\| u(k)(\cdot - (t - s)) - u(\cdot - (t - s))\right\|_{L^2(\Omega)} \leq e^{\gamma(t-t_0)} \cdot d_{c,\gamma,p.i.p.}(u(k), u) \leq \text{const}.
\]

Due to Alaoglu’s theorem, there exists a sequence of indices \( k_j \uparrow \infty \) such that \((f_{k_j})_{j \in \mathbb{N}} \) converges to some \( f \in L^2(t_0, t; L^2(\Omega)) \) weakly in \( L^2(t_0, t; L^2(\Omega)) \). We conclude from Lemmas \ref{8.5} and \ref{8.6} that \( f(s) \in \mathbb{G}(s, u(\cdot + s - t)|_{(-\infty, 0]}) \) holds for Lebesgue-almost every \( s \in [t_0, t] \).

In regard to the claim \( u \in U(t, t_0, \phi) \), it remains to verify

\[
u(s) = \begin{cases} 
S(t-t_0+s) \phi(0) + \int_{t_0}^{t+s} S(t+s-\tau) f(\tau) \, d\tau & \text{if } s \in [-t-t_0, 0] \\
\phi(t-t_0 + s) & \text{if } s < -(t-t_0).
\end{cases}
\]

The statement for \( s < -(t-t_0) \) is rather obvious since the convergence of \((\phi(k))_{k \in \mathbb{N}} \) w.r.t. \( d_{c,\gamma} \) implies the convergence in \( \| \cdot \|_\gamma \) and hence pointwise convergence in time.
Fixing any \( s \in [-t_0, 0] \), the adjoint \( C_0 \) semigroup \( \{S(\tau)\} \) is the essential tool to obtain for each \( w \in L^2(\Omega) \)
\[
\langle u^{(k)}(s), w \rangle_{L^2(\Omega)}
= \langle S(t - t_0 + s) \phi^{(k)}(0), w \rangle_{L^2(\Omega)} + \int_{t_0}^{t+s} \langle S(t + s - \tau) f_{k}(\tau), w \rangle_{L^2(\Omega)} \, d\tau
= \langle S(t - t_0 + s) \phi^{(k)}(0), w \rangle_{L^2(\Omega)} + \int_{t_0}^{t+s} \langle f_{k}(\tau), S(t + s - \tau) w \rangle_{L^2(\Omega)} \, d\tau
\rightarrow \langle S(t - t_0 + s) \phi(0), w \rangle_{L^2(\Omega)} + \int_{t_0}^{t+s} \langle f(\tau), S(t + s - \tau) w \rangle_{L^2(\Omega)} \, d\tau
\]
as \( j \to \infty \) due to \( f_{k} \to f \) in \( L^2(t_0, t; L^2(\Omega)) \).

For preparing the proof of main Theorem \( \text{[4.2]} \) we briefly formulate the weaker closed graph property w.r.t. \( d_{c,\gamma,p.i.p.} \) for the case of infinite delay, i.e., \( \theta := \infty \), but under the additional monotonicity assumption \( (G7) \).

**Lemma 8.9.** Let \( G : \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times L^2(\Omega) \to \mathbb{R} \) satisfy condition \( (G7) \) — in addition to \( (G1) - (G5) \). Set \( \theta := \infty \). Then for any \( (t, t_0) \in \mathbb{R}_+^2 \), the graph of \( U(t, t_0, \cdot) : C_\gamma \to C_\theta \) is sequentially closed w.r.t. \( d_{c,\gamma,p.i.p.} \) or, equivalently, with respect to the simultaneous convergence in \( d_{c,\gamma,\tau} \) for every \( \tau > 0 \).

Its proof is based on essentially the same arguments. We just replace preceding Lemma \( \text{[8.5]} \) by the following auxiliary result. Indeed, the additional assumption \( (G7) \) about the inclusion property of \( G \) (w.r.t. its fourth argument) now provides the essential tool for extending this upper semi-continuity to the set-valued map \( \mathcal{G} \) on the right-hand side of evolution inclusion \( [4] \). It helps us to overcome the obstacle that \( C_\gamma \to L^2(\Omega), \ u \mapsto \text{ess sup}_{s \leq t} \ [u(s, \cdot)]_c \) is not continuous w.r.t. \( d_{c,\gamma,p.i.p.} \):

**Lemma 8.10.** Let \( \Omega \subset \mathbb{R}^d \) be a bounded open set with smooth boundary, \( \theta := \infty \). Suppose hypotheses \( (G1) - (G5) \) and \( (G7) \) for the set-valued map \( G : \mathbb{R} \times \Omega \times \mathbb{R} \times \mathbb{R} \times L^2(\Omega) \to \mathbb{R} \).

Then for each \( t_0 \in \mathbb{R} \), the mapping \( \mathcal{G}(t_0, \cdot) : (C_\gamma, d_{c,\gamma,p.i.p.}) \to L^2(\Omega) \) specified in \( [10] \) is Hausdorff upper semi-continuous with nonempty convex closed values.

**Proof.** Fix \( t_0 \in \mathbb{R} \), \( \varphi \in C_\gamma \) and \( \varepsilon > 0 \) arbitrarily. Due to Corollary \( \text{[8.4]} \) there is a radius \( \rho > 0 \) such that all \( u, v, w \in L^2(\Omega) \) with
\[
\|u - \varphi(0)\|_{L^2(\Omega)} + \|v - \text{ess sup}_{s \leq 0} [\varphi(s)]_c\|_{L^2(\Omega)} + \|w - \varphi(0)\|_{L^2(\Omega)} < \rho
\]
satisfy
\[
\mathcal{G}(t_0; u, v, w) \subset \mathbb{B}_\varepsilon \left( \mathcal{G}(t_0; \varphi(0), \text{ess sup}_{s \leq 0} [\varphi(s)]_c, \varphi(0)) \right) \quad \mathbb{B}_\varepsilon \left( \mathcal{G}(t_0, \varphi) \right) \subset L^2(\Omega).
\]
Levi’s theorem of monotone convergence and Hölder’s inequality guarantee some \( \tau = \tau(\varphi, \rho) \in \mathbb{N} \) sufficiently large such that
\[
\left\| \text{ess sup}_{-\infty < s \leq 0} [\varphi(s)]_c - \text{ess sup}_{-\tau \leq s \leq 0} [\varphi(s)]_c \right\|_{L^2(\Omega)} < \rho/4.
\]
Now set the radius \( r := \frac{\rho}{2\sqrt{\tau}} \in (0, \frac{\rho}{4}] \) and then, all \( \psi \in C_\gamma \) with \( d_{c,\gamma,p.i.p.}(\varphi, \psi) < r \) satisfy
\[
\left\| \text{ess sup}_{-\tau \leq s \leq 0} [\psi(s)]_c - \text{ess sup}_{-\tau \leq s \leq 0} [\varphi(s)]_c \right\|_{L^2(\Omega)} \leq \left\| [\psi]_c - [\varphi]_c \right\|_{L^2(\Omega; L^\infty([-\tau, 0]))} < \frac{\rho}{4}
\]
and
\[ \|\psi(0) - \varphi(0)\|_{L^2(\Omega)} \leq d_{c,\gamma,p.i.p.}(\varphi, \psi) < r \leq \frac{\rho}{4}, \]
so
\[ \mathcal{G}(t_0, \psi(0), \text{ess sup}_{-\tau \leq s \leq 0} [\psi(s)]_c, \psi(0)) \subset \mathcal{B}_\varepsilon(\mathcal{G}(t_0, \varphi)) \subset L^2(\Omega). \]
By (G7) this implies
\[ \mathcal{G}(t_0, \psi(0), \text{ess sup}_{-\infty < s \leq 0} [\psi(s)]_c, \psi(0)) \subset \mathcal{B}_\varepsilon(\mathcal{G}(t_0, \varphi)) \subset L^2(\Omega), \]
which is equivalent to
\[ \mathcal{G}(t_0, \psi) \subset \mathcal{B}_\varepsilon(\mathcal{G}(t_0, \varphi)) \subset L^2(\Omega). \]
The remaining claims about nonempty convex closed values result from Corollary 8.4 about \( \mathcal{G} \) directly. \( \Box \)

8.4. A priori estimates for solutions to parabolic differential inclusion \([15]\).

Lemma 8.11. Under the assumptions (G1) - (G6), the solution \( y_t(\cdot; t_0, \phi) \in C_\gamma \) of \([16]\) satisfies these estimates for every \( (t, t_0) \in \mathbb{R}_+^2 \) and \( \phi \in C_\gamma \) with \( C \) denoting the constant \( C = (\bar{\alpha} + \bar{\beta} \cdot c) \cdot \mathcal{L}^n(\Omega)^{1/2} \)
\[ \|y_t(\cdot; t_0, \phi)\|_\gamma \leq e^{-(\lambda_1 - \bar{\beta})(t-t_0)} \cdot \|\phi\|_\gamma + \frac{c}{\lambda_1 - \bar{\beta}} \cdot (1 - e^{-(\lambda_1 - \bar{\beta})(t-t_0)}), \]
\[ d_{c,\gamma}(0, y_t(\cdot; t_0, \phi)) \leq e^{-(\lambda_1 - \bar{\beta})(t-t_0)} \cdot d_{c,\gamma}(0, \phi) + \frac{c}{\lambda_1 - \bar{\beta}} \cdot (1 - e^{-(\lambda_1 - \bar{\beta})(t-t_0)} + \frac{c}{\lambda_1 - \bar{\beta}} \cdot \mathcal{L}^n(\Omega)^{1/2}. \]

Proof. By definition \([16]\), each solution \( y_t(\cdot; t_0, \phi) \in C_\gamma \) is related to a selector \( f : [t_0, t] \rightarrow L^2(\Omega) \) satisfying \( f(\tau) \in \mathcal{G}(\tau, y_t(\cdot + \tau - t)|_{(-\infty,0)} \) for a.e. \( \tau \in [t_0, t]: \)
\[ y_t(s; t_0, \phi) = \begin{cases} S(t-t_0+s) \phi(0) + \int_{t_0}^{t+s} S(t+s-\tau) f(\tau) \, d\tau & \text{if } s \in [-t-t_0, 0] \\ \phi(t-t_0+s) & \text{if } s < -(t-t_0) \end{cases} \]
and so
\[ \|y_t(s; t_0, \phi)\|_{L^2(\Omega)} \leq \begin{cases} e^{-\lambda_1(\cdot - t-t_0+s)} \|\phi(0)\|_{L^2} + \int_{t_0}^{t+s} e^{-\lambda_1(\cdot - t-t_0+s-\tau)} \|f(\tau)\|_{L^2} \, d\tau & \text{if } s \geq -(t-t_0) \\ \|\phi(t-t_0+s)\|_{L^2(\Omega)} & \text{if } s < -(t-t_0) \end{cases} \]
hold at every time instant \( s \leq 0 \) because every absolutely continuous curve \( u : [t_0, t] \rightarrow L^2(\Omega) \) solving \( \partial_t u - \Delta u = f \) in the weak sense satisfies the upper estimate
\[ \frac{1}{2} \cdot \frac{d}{dt} \|u(t)\|_{L^2(\Omega)}^2 + \|\nabla u(t)\|_{L^2(\Omega)}^2 \leq \|u(t)\|_{L^2(\Omega)} \cdot \|f(t)\|_{L^2(\Omega)} \]
and in combination with the Cauchy-Schwarz and Poincaré inequalities, it leads to
\[ \frac{d}{dt} \|u(t)\|_{L^2(\Omega)} + \lambda_1 \cdot \|u(t)\|_{L^2(\Omega)} \leq \|f(t)\|_{L^2(\Omega)} \quad (28) \]
to which we finally apply the Gronwall inequality for integrating. Here \( \lambda_1 > 0 \) still denotes the smallest eigenvalue of the negative Laplacian operator with homogeneous Dirichlet boundary conditions in \( \Omega \). We generally assume \( \gamma > \lambda_1 \). As a
The definition of $d$

Assume hypotheses (G1) – (G4), (G5′) and (G6) for $G : \mathbb{R} \times \Omega \times \mathbb{R}^2 \times L^2(\Omega) \rightarrow \mathbb{R}$. Then for each $\phi \in C_\gamma$ and $t_0 \in \mathbb{R}$, the solution $y_t(\cdot ; t_0, \phi) \in C_\gamma$.
of (10) satisfies
\[ \|y_t(t; t_0, \phi)\|_\gamma \leq \|\phi\|_\gamma \cdot e^{-\left(\lambda_1 - \tilde{\beta}\right) \cdot (t-t_0)} + \mathcal{L}^n(\Omega)^{1/2} \cdot \int_{t_0}^t \alpha(s) \cdot e^{-\left(\lambda_1 - \tilde{\beta}\right) \cdot (t-s)} ds. \]

Proof. It is based on essentially the same arguments as for recent Lemma 8.11. As key difference, assumption (G5’) now implies for Lebesgue-almost every \( s \in [t_0, t] \) and \( x \in \Omega \)
\[ |f(s, x)| \leq \alpha(s) + \tilde{\beta} \cdot |y_t(- (t-s); t_0, \phi)(x)|, \]
which in turn implies that
\[ \|f(s)\|_{L^2(\Omega)} \leq \alpha(s) \cdot \mathcal{L}^n(\Omega)^{1/2} + \tilde{\beta} \cdot \|y_t(- (t-s); t_0, \phi)\|_{L^2(\Omega)} \]
due to the Minkowski inequality. This upper bound is now playing the role of estimate (27). Hence we conclude from the same arguments as in the recent proof of Lemma 5.11
\[ \|y_t(t; t_0, \phi)\|_\gamma \leq e^{-\lambda_1 \cdot (t-t_0)} \cdot \|\phi\|_\gamma + \int_{t_0}^t e^{-\lambda_1 \cdot (t-\tau)} \left( \alpha(\tau) \cdot \mathcal{L}^n(\Omega)^{1/2} + \tilde{\beta} \cdot \|y_t(\cdot; t_0, \phi)\|_\gamma \right) d\tau \]
and so,
\[ e^{\lambda_1 \cdot t} \cdot \|y_t(t; t_0, \phi)\|_\gamma \leq e^{\lambda_1 \cdot t} \cdot \|\phi\|_\gamma + \int_{t_0}^t \left( \alpha(\tau) \cdot \mathcal{L}^n(\Omega)^{1/2} \cdot e^{\lambda_1 \cdot \tau} + \tilde{\beta} \cdot e^{\lambda_1 \cdot \tau} \cdot \|y_t(\cdot; t_0, \phi)\|_\gamma \right) d\tau. \]
The Gronwall inequality in its integral form implies for every \( t \geq t_0 \)
\[ e^{\lambda_1 \cdot t} \cdot \|y_t(\cdot; t_0, \phi)\|_\gamma \]
\[ \leq e^{\lambda_1 \cdot t_0} \cdot \|\phi\|_\gamma + \mathcal{L}^n(\Omega)^{1/2} \cdot \int_{t_0}^t \alpha(\tau) \cdot e^{\lambda_1 \cdot \tau} d\tau \]
\[ + \int_{t_0}^t e^{\tilde{\beta} \cdot (t-\tau)} \cdot \tilde{\beta} \cdot \left( e^{\lambda_1 \cdot t_0} \cdot \|\phi\|_\gamma + \mathcal{L}^n(\Omega)^{1/2} \cdot \int_{t_0}^\tau \alpha(s) \cdot e^{\lambda_1 \cdot s} ds \right) d\tau \]
\[ = e^{\lambda_1 \cdot t_0} \cdot \|\phi\|_\gamma \cdot e^{\tilde{\beta} \cdot (t-t_0)} + \mathcal{L}^n(\Omega)^{1/2} \cdot \int_{t_0}^t \alpha(\tau) \cdot e^{\lambda_1 \cdot \tau} d\tau \]
\[ + \mathcal{L}^n(\Omega)^{1/2} \cdot \int_{t_0}^t \alpha(s) \cdot e^{\lambda_1 \cdot s + \tilde{\beta} \cdot (t-s)} d(s, \tau) \]
\[ \leq e^{\lambda_1 \cdot t_0} \cdot \|\phi\|_\gamma \cdot e^{\tilde{\beta} \cdot (t-t_0)} + \mathcal{L}^n(\Omega)^{1/2} \cdot \int_{t_0}^t \alpha(\tau) \cdot e^{\lambda_1 \cdot \tau} d\tau \]
\[ + \mathcal{L}^n(\Omega)^{1/2} \cdot \int_{t_0}^t \alpha(s) \cdot e^{\lambda_1 \cdot s + \tilde{\beta} \cdot t} \left( e^{\tilde{\beta} \cdot s} - e^{\tilde{\beta} \cdot t} \right) ds \]
which implies that
\[ \|y_t(t; t_0, \phi)\|_\gamma \leq \|\phi\|_\gamma \cdot e^{-\left(\lambda_1 - \tilde{\beta}\right) \cdot (t-t_0)} + \mathcal{L}^n(\Omega)^{1/2} \cdot \int_{t_0}^t \alpha(s) \cdot e^{\left(\lambda_1 - \tilde{\beta}\right) \cdot s} ds \cdot e^{-\left(\lambda_1 - \tilde{\beta}\right) \cdot t}. \]
8.5. Conclusions about pullback absorbing and positively invariant balls w.r.t. $\|\cdot\|$. As an obvious consequence of the a priori estimates in Lemma 8.11 balls in $C_{\gamma}$ of radius $\rho_0 = \frac{(\gamma + \beta \epsilon) \mathcal{L}^{\alpha}(\Omega)^{1/2}}{1 - \beta}$ are pullback absorbing for all families of subsets $\tilde{D} = \{D(t) : t \in \mathbb{R}\}$ which do not grow “too fast” as $t_0 \to -\infty$. In more detail, we even conclude:

**Corollary 8.13.** Suppose assumptions (G1) – (G6) as in Lemma 8.11 and consider the MNDS $U : \mathbb{R}^2 \times C_{\gamma} \rightarrow C_{\gamma}$ specified in Proposition 8.5 by means of inclusion (15).

Set $\rho_0 = \frac{(\gamma + \beta \epsilon) \mathcal{L}^{\alpha}(\Omega)^{1/2}}{1 - \beta} > 0$.

Then the set-valued mapping $\tilde{B} = \{B(t) : t \in \mathbb{R}\}$ of “balls” $B(t) := \{\phi \in C_{\gamma} | \|\phi\|_{\gamma} \leq \rho(t)\}$ with radii

$$\rho(t) := \rho_0 \cdot \left(1 + e^{-(\lambda_1 - \tilde{\beta})t}\right) \in (\rho_0, \infty)$$

has the following properties:

1. $\tilde{B}$ is positively invariant w.r.t. $U$, i.e., $U(t, \tau, B(\tau)) \subset B(t)$ for any $(t, \tau) \in \mathbb{R}^2_{\geq}$.

2. Fixing $q \in [0, \lambda_1 - \tilde{\beta})$ arbitrarily, let $\mathcal{D}$ consist of families $\tilde{D} = \{D(t) : t \in \mathbb{R}\}$ of nonempty closed bounded sets in $C_{\gamma}$ satisfying $\sup \{e^{qt} \|D(t)||_{\gamma} | t < 0\} < \infty$.

Then $\tilde{B}$ is pullback $\mathcal{D}$-absorbing with respect to $U$.

The radius function $\rho : \mathbb{R} \rightarrow \mathbb{R}$ here is chosen as the unique solution to the initial value problem

$$\begin{align*}
\rho(t) &:= \rho_0 \cdot \left(1 + e^{-(\lambda_1 - \tilde{\beta})t}\right) \in (\rho_0, \infty) \\
\rho(0) &= 2 \rho_0.
\end{align*}$$

In statement (2.), we consider universes which may grow even exponentially (in backward time direction), but their exponential growth rate is bounded from above by the given parameters of $U$.

**Proof of Corollary 8.13**

1. For any $(t, \tau) \in \mathbb{R}^2_{\geq}$, every $y_t(\cdot; \tau, \phi) \in C_{\gamma}$ with $\phi \in B(\tau) \subset C_{\gamma}$ satisfies due to Lemma 8.11:

$$\begin{align*}
\|y_t(\cdot; \tau, \phi)\|_{\gamma} &\leq e^{-(\lambda_1 - \tilde{\beta})(t-\tau)} \cdot \|\phi\|_{\gamma} + \rho_0 \cdot \left(1 - e^{-(\lambda_1 - \tilde{\beta})(t-\tau)}\right) \\
&\leq e^{-(\lambda_1 - \tilde{\beta})(t-\tau)} \cdot \rho(\tau) + \rho_0 \cdot \left(1 - e^{-(\lambda_1 - \tilde{\beta})(t-\tau)}\right) \\
&= \rho_0 \cdot \left(e^{-(\lambda_1 - \tilde{\beta})(t-\tau)} \cdot (1 + e^{-(\lambda_1 - \tilde{\beta})\tau}) + 1 - e^{-(\lambda_1 - \tilde{\beta})(t-\tau)}\right) \\
&= \rho_0 \cdot \left(e^{-(\lambda_1 - \tilde{\beta})t} + 1\right) = \rho(t),
\end{align*}$$

e.g., $y_t(\cdot; \tau, \phi) \in B(t)$.

2. For $q \in [0, \lambda_1 - \tilde{\beta})$ fixed arbitrarily, choose any family $\tilde{D} = \{D(t) : t \in \mathbb{R}\}$ of nonempty closed bounded sets in $C_{\gamma}$ with $C_{\tilde{D}} := \sup \{e^{qt} \|D(t)||_{\gamma} | t < 0\} < \infty$.

Then we conclude from Lemma 8.11 that for any $(t, \tau) \in \mathbb{R}^2_{\geq}$ and $\phi \in D(\tau)$,

$$\begin{align*}
\|y_t(\cdot; \tau, \phi)\|_{\gamma} &\leq e^{-(\lambda_1 - \tilde{\beta})(t-\tau)} \cdot \|\phi\|_{\gamma} + \rho_0 \cdot \left(1 - e^{-(\lambda_1 - \tilde{\beta})(t-\tau)}\right) \\
&\leq e^{-(\lambda_1 - \tilde{\beta})(t-\tau)} \cdot C_{\tilde{D}} \cdot e^{-q\tau} + \rho_0 \cdot \left(1 - e^{-(\lambda_1 - \tilde{\beta})(t-\tau)}\right) \\
&= C_{\tilde{D}} e^{-(\lambda_1 - \tilde{\beta})t} \cdot e^{(\lambda_1 - \tilde{\beta} - q)\tau} + \rho_0 \cdot \left(1 - e^{-(\lambda_1 - \tilde{\beta})(t-\tau)}\right).
\end{align*}$$
Whenever we keep $t \in \mathbb{R}$ fixed and consider $\tau \to -\infty$, we obtain due to $\lambda_1 - \beta - q > 0$

$$\limsup_{\tau \to -\infty} \sup_{\phi \in D(\tau)} \|u_{\tau}(\cdot; \tau, \phi)\|_\gamma \leq \rho_0 < \rho(t).$$

As a consequence, $\hat{B}$ is pullback absorbing w.r.t. $U$. \hfill $\square$

8.6. **Standard results about a priori estimates for solutions to parabolic differential equations.**

**Lemma 8.14** ([4] Theorem IV.2.1 & Lemma IV.2.1, [33] Lemma 6.15). Let $\Omega \subset \mathbb{R}^n$ be a nonempty open set with smooth boundary. For the parabolic problem

$$\begin{aligned}
\partial_t u &= \Delta u + f(t) \quad \text{in } (0, T) \times \Omega \\
u(0) &= u_0 \quad \text{on } \Omega \\
u &= 0 \quad \text{on } (0, T) \times \partial \Omega
\end{aligned}$$

with a function $f \in L^2(0, T; L^2(\Omega))$ and any initial state $u_0 \in L^2(\Omega)$, there exists a unique strong solution $u \in C^0([0, T], L^2(\Omega))$ with the following properties:

- the map $[0, T] \ni t \mapsto \sqrt{t} \cdot u'(t)$ is in $L^2(0, T; L^2(\Omega))$,
- the map $(0, T] \ni t \mapsto \int_\Omega |\nabla u(t)|^2 \, dx$ defines $\|\nabla u(t)\|_{L^2(\Omega)}^2$ is continuous and integrable,
- $\|u'(t)\|_{L^2(\Omega)}^2 + \frac{d}{dt} \|\nabla u\|_{L^2(\Omega)}^2 = (f(t), u'(t))_{L^2(\Omega)}$ holds for Lebesgue-almost every $t$.

Moreover, if $u_0 \in W^{1,2}_0(\Omega) \cap W^{2,2}(\Omega)$, then $u' \in L^2(0, T; L^2(\Omega))$ and the map $t \mapsto \|\nabla u(t)\|_{L^2(\Omega)}^2$ is absolutely continuous.

An a priori estimate of the time derivative for weak solutions to parabolic equations will be used. The following bounds are special cases of inequalities (6.3), (6.6), respectively, in [39] Ch. III § 6] and the existence is stated in [39] Ch. III Theorem 6.1:

**Lemma 8.15.** Consider the parabolic problem

$$\begin{aligned}
\partial_t u - \sum_{i,j=1}^d \partial_{x_i} \left( a_{ij}(t, x) \cdot \partial_{x_j} u \right) &= h(t, x) \quad \text{in } [0, T] \times \Omega, \\
u &= 0 \quad \text{on } (0, T) \times \partial \Omega, \\
u(0, \cdot) &= u_0 \quad \text{on } \Omega
\end{aligned}$$

with $h \in L^2([0, T] \times \Omega)$ and Lebesgue measurable coefficients $a_{ij} : [0, T] \times \Omega \to \mathbb{R}$ satisfying the condition of uniform parabolicity and

$$\mu(T) := \int_0^T \operatorname{ess sup}_{x \in \Omega} |\partial_t a_{ij}(t, x)| \, dt < \infty.$$

Then, for every $u_0 \in W^{1,2}_0(\Omega)$, there exists a unique weak solution $u : [0, T] \times \Omega \to \mathbb{R}$ in $W^{1,2}([0, T] \times \Omega)$ and it satisfies the estimates for every $t \in (0, T]$

$$\begin{aligned}
\|u(t, \cdot)\|_{L^2(\Omega)}^2 + \|\nabla_x u \|_{L^2([0, t] \times \Omega)}^2 &\leq C(\Omega, t) \cdot \left( \|u_0\|_{L^2(\Omega)}^2 + \|h\|_{L^2([0, t] \times \Omega)}^2 \right), \\
\|\nabla_x u(t, \cdot)\|_{L^2(\Omega)}^2 + \|\partial_t u \|_{L^2([0, t] \times \Omega)}^2 &\leq C(\Omega, \mu(t)) \cdot \left( \|\nabla_x u_0\|_{L^2(\Omega)}^2 + \|h\|_{L^2([0, t] \times \Omega)}^2 \right).
\end{aligned}$$
Uniqueness implies that the several types of solutions coincide as mentioned in \[ \text{[37] \S 2.1}. \] Hence, for integral solutions to the non-homogeneous heat equation, we obtain:

**Corollary 8.16.** Let \( \Omega \subset \mathbb{R}^d \) be a bounded open set. Consider an integral solution \( u : [0, T] \to L^2(\Omega) \) to non-autonomous problem (29) with \( f \in L^2(0, T; L^2(\Omega)) \subset L^1(0, T; L^2(\Omega)) \) and \( u_0 \in W^{1,2}_0(\Omega). \)

Then, \( u \) is both a strong and a weak solution. In particular, \( u \in W^{1,2}(0, T; \Omega) \) and, it satisfies the following a priori estimate for every \( t \in [0, T] \):

\[
\|u(t)\|_{W^{1,2}(\Omega)}^2 + \|u\|_{W^{1,2}(0, T; \Omega)}^2 \leq C(\Omega, T) \cdot \left( \|u_0\|_{W^{1,2}(\Omega)}^2 + \int_0^t \|f(s)\|_{L^2(\Omega)}^2 \,ds \right).
\]

**8.7. Supplementary \( L^2(\Omega; L^\infty([t_0, T])) \) estimates due to assumption \( (G^3') \).** Assumption \((G^3')\) about Lipschitz continuity of \( G \) w.r.t. the third and fourth argument implies condition \((G^2)\) about Hausdorff upper semi-continuity directly. Now we formulate some further conclusions from \((G^3')\) which will be used for proving main Theorem 4.1 later.

**Lemma 8.17.** Let \( G : \mathbb{R} \times \Omega \times \mathbb{R} \times \mathbb{R} \times L^2(\Omega) \rightarrow \mathbb{R} \) satisfy conditions to \((G1) - (G5), (G^3)\). Consider any \( \theta > 0 \) and \( \Theta \in L^1(\Omega) \) with \( \Theta \geq 0 \). Fixing \((T, t_0) \in \mathbb{R}^2_+ \) and \( \phi \in C_\gamma \) with \( \phi(0) \in W^{1,2}_0(\Omega) \) arbitrarily, suppose \( u : (-\infty, T) \rightarrow L^2(\Omega) \) to be a mild solution to parabolic differential inclusion \([15]\).

For each \( \psi \in C_\gamma \) with \( \psi(0) \in W^{1,2}_0(\Omega) \) and

\[
\rho_0 := \|\psi(0) - \phi(0)\|_{W^{1,2}(\Omega)} + \left\| \text{ess sup}_{t - \theta \leq s \leq t} \left| [\psi(s)]_c - [\phi(s)]_c \right| \right\|_{L^2(\Omega)},
\]

there exists a (both strong and mild) solution \( v : (-\infty, T) \rightarrow L^2(\Omega) \) of

\[
\begin{align*}
\partial_t v - \Delta v &\in G\left(t, x; v(t, x), \text{ess sup}_{t - \theta \leq s \leq t} \left[ v(s, x)_c, v(t, \cdot) \right] \right) \\
v &= \psi(\cdot - t_0) \quad \text{a.e. in } (t_0, T) \times \Omega \quad \text{a.e. in } (-\infty, t_0) \times \Omega \\
v &= 0 \quad \text{on } (t_0, T) \times \partial\Omega
\end{align*}
\]

with the supplementary property for every \( t \in [t_0, T] \):

\[
\|u(t) - v(t)\|_{L^2(\Omega)} + \left\| \text{ess sup}_{t_0 \leq s \leq t} \left[ [u(s, \cdot)]_c - [v(s, \cdot)]_c \right] \right\|_{L^2(\Omega)} \leq \text{const} \cdot (\Omega, T - t_0) \cdot \rho_0 \cdot e^{\text{const} \cdot (\Omega, T - t_0) \cdot \|\Lambda\|_{L^2([t_0, T])}}.
\]

**Corollary 8.18.** As in Lemma 8.17, let \( G : \mathbb{R} \times \Omega \times \mathbb{R} \times \mathbb{R} \times L^2(\Omega) \rightarrow \mathbb{R} \) satisfy the conditions \((G1) - (G5), (G^3)\). For any \( \phi, \psi \in C_\gamma \) with \( \phi(0), \psi(0) \in W^{1,2}_0(\Omega) \) and \((t, t_0) \in \mathbb{R}^2_+ \), each solution \( y_t(\cdot; t_0, \phi) \in C_\gamma \) of \([16]\) satisfies

\[
\inf \left\{ d_{c, \gamma}(y_t(\cdot; t_0, \phi), z) \mid z \in U(t, t_0, \psi) \right\} \leq C \cdot (d_{c, \gamma}(\phi, \psi) + \|\psi(0) - \phi(0)\|_{W^{1,2}(\Omega)})
\]

and (similarly, but related to space-dependent delay)

\[
\inf_{z \in U(t, t_0, \psi)} \left( \|y_t(\cdot; t_0, \phi) - z\|_{\gamma} + \left\| \text{ess sup}_{t - \theta \leq s \leq t} \left[ [y_t(s, \cdot)]_c - [z(s)]_c \right] \right\|_{L^2(\Omega)} \right) \leq C \cdot \|\phi(0) - \psi(0)\|_{W^{1,2}(\Omega)} + \|\phi - \psi\|_{\gamma} + \left\| \text{ess sup}_{t - \theta \leq s \leq t} \left[ [\phi(s)]_c - [\psi(s)]_c \right] \right\|_{L^2(\Omega)}.
\]
with a constant $C = C(\Omega, \|A\|_{L^2(\mathbb{R})}, t_0, t) \geq 1$.

For proving Lemma 8.17 we use several standard tools in set-valued analysis and evolution equations for imitating conclusions which are usually formulated as Filippov’s theorem about differential inclusions (e.g., [11 § 2.4, Theorem 1], [43 Theorem A.6]):

**Proposition 8.19** ([2 Theorem 9.7.2], [17]). Consider a metric space $X$ and a set-valued mapping $\mathcal{F} : [a, b] \times X \sim \mathbb{R}^m$ satisfying the following conditions

(i) the values of $\mathcal{F}$ are nonempty compact convex subsets of $\mathbb{R}^m$,
(ii) for every $x \in X$, the mapping $\mathcal{F}(\cdot, x) : [a, b] \sim \mathbb{R}^m$ is Lebesgue measurable,
(iii) there exists some $\lambda \in L^1([a, b])$ with $\lambda \geq 0$ such that for every almost every $t \in [a, b]$, the mapping $\mathcal{F}(t, \cdot) : X \sim \mathbb{R}^m$ is $\lambda(t)$-Lipschitz continuous.

Then there exists a measurable/Lipschitz parametrization $f$ of $\mathcal{F}$, i.e., a single-valued function $f : [a, b] \times X \times \mathbb{R} \to \mathbb{R}^m$ (with $\mathbb{R}$ denoting the closed unit ball in $\mathbb{R}^m$) and a constant $c > 0$ (independent of $\mathcal{F}$) such that

1. $\mathcal{F}(t, x) = \{ f(t, x, u) \mid u \in \mathbb{R} \}$ holds for every $t \in [a, b]$ and $x \in X$,
2. for a.e. $t \in [a, b]$ and any $u \in \mathbb{R}$, the function $f(t, \cdot, u) : X \to \mathbb{R}^m$ is $c \lambda(t)$-Lipschitz continuous,
3. for a.e. $t \in [a, b]$ and any $x \in X$, the function $f(t, x, \cdot) : \mathbb{R} \to \mathbb{R}^m$ is Lipschitz continuous.

**Lemma 8.20** (Filippov, [2 Theorem 8.2.10]). Consider a complete $\sigma$-finite measure space $(\mathcal{O}, \mathcal{A}, \mu)$, complete separable metric spaces $X, Y$ and a measurable set-valued mapping $\mathcal{U} : \mathcal{O} \sim X$ with nonempty closed values. Let $f : \mathcal{O} \times X \to Y$ be a Carathéodory map.

Then for every measurable map $h : \mathcal{O} \to Y$ satisfying

$$h(\omega) \in f(\omega, \mathcal{U}(\omega)) \quad \text{for } \mu\text{-almost all } \omega \in \mathcal{O}$$

there exists a measurable selector $u : \mathcal{O} \to X$ of $\mathcal{U}$ such that

$$\begin{cases}
  u(\omega) \in \mathcal{U}(\omega) \\
  h(\omega) = f(\omega, u(\omega))
\end{cases} \quad \text{for } \mu\text{-almost all } \omega \in \mathcal{O}.$$

**Lemma 8.21** ([20 § 1.3], [49 Appendix B]). Let $\Omega$ be an arbitrary domain in $\mathbb{R}^d$ and let $A$ be an elliptic differential operator on $L^2(\Omega)$ subject to homogeneous Dirichlet boundary conditions.

Then, the solutions of $\partial_t u = Au$ form a strongly continuous semigroup $(S(t))_{t \geq 0}$ of bounded linear operators on $L^2(\Omega)$. For every $t \geq 0$, the intersection $L^1(\Omega) \cap L^\infty(\Omega)$ is invariant w.r.t. $S(t)$, i.e.,

$$S(t) \left( L^1(\Omega) \cap L^\infty(\Omega) \right) \subset L^1(\Omega) \cap L^\infty(\Omega) \subset L^2(\Omega).$$

The restriction $(S(t)|_{L^1(\Omega) \cap L^\infty(\Omega)})_{t \geq 0}$ can be extended to a unique positive non-expansive semigroup on $L^p(\Omega)$ for every $p \in [1, \infty]$, i.e.,

$$\|S(t)v\|_{L^p(\Omega)} \leq \|v\|_{L^p(\Omega)} \quad \text{for every } p \in [1, \infty], t \geq 0, v \in L^p(\Omega).$$

If the domain $\Omega \subset \mathbb{R}^d$ is bounded and $A$ the Laplace operator, then the heat semigroup $(S(t))_{t \geq 0}$ with homogeneous Dirichlet boundary conditions satisfies for every $p \in [1, \infty]$, $t \geq 0$ and $v \in L^p(\Omega)$

$$\|S(t)v\|_{L^p(\Omega)} \leq \|v\|_{L^p(\Omega)}, \quad \|S(t)v\|_{L^p(\Omega)} \leq C(\Omega) e^{-\lambda_t} \|v\|_{L^p(\Omega)}$$
and for every \( t > 0 \), \( v \in L^p(\Omega), 1 \leq p < \infty \) with \( \frac{1}{p} = \frac{1}{r} - \frac{1}{q} \)
\[
\|S(t)v\|_{L^q(\Omega)} \leq (4\pi t)^{-\frac{d}{2q}} \|v\|_{L^r(\Omega)}.
\]

Proof of Lemma 8.17 Fix \((T, t_0) \in \mathbb{R}^2_+\) and \( \phi \in C_\gamma \) with \( \phi(0) \in W^{1,2}_0(\Omega) \) arbitrarily and let \( u : (-\infty, T) \rightarrow L^2(\Omega) \) be any mild solution to parabolic differential inclusion (15), i.e., there exists some \( f \in L^2(t_0, T; L^2(\Omega)) \) with
\[
f(t, x) \in G\left(t, x; u(t, x), \operatorname{ess sup}_{t - \theta - \Theta(x) \leq s \leq t} [u(s, x)]_c, u(t, \cdot)\right)
\]
for Lebesgue-almost all \( t \in [t_0, T] \) and \( x \in \Omega \) such that we have the representation
\[
u(t) = S(t - t_0) \phi(0) + \int_{t_0}^t S(t - s) f(s) \, ds, \quad t \in (t_0, T].
\]

In particular, \( u \) is also a weak solution to a non-homogeneous heat equation with homogeneous Dirichlet condition. Hence in combination with linear growth assumption (G5) and Gronwall inequality, Corollary 8.16 implies \( u \in W^{1,2}([t_0, T] \times \Omega) \).

Due to additional assumption \((G3')\), Proposition 8.19 about measurable/Lipschitz parametrizations leads to a constant \( \tilde{c} \geq 0 \) (independent of \( G \)) and a function \( g : \mathbb{R} \times \Omega \times \mathbb{R} \times \mathbb{R} \times L^2(\Omega) \times [-1, 1] \rightarrow \mathbb{R} \) with the following properties
\begin{enumerate}[(1.)]
\item \( G(t, y; z, v, \eta) = \{g(t, y, z, v, \eta) \mid \eta \in [-1, 1]\} \) for any \( t \in [t_0, T] \), \( x \in \Omega \), \( y, z \in \mathbb{R} \) and \( v \in L^2(\Omega) \),
\item for \( a.e. \) \( t \in [t_0, T] \) and any \( \eta \in [-1, 1] \), the function \( g(t, \cdot; \cdot; \cdot; \cdot; \eta) : \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times L^2(\Omega), \|\cdot\|_{L^2} \rightarrow \mathbb{R} \) is \( \tilde{c} \Lambda(t) \)-Lipschitz continuous,
\item for \( a.e. \) \( t \in [t_0, T] \) and any \( x \in \Omega \), \( y, z \in \mathbb{R} \), \( v \in L^2(\Omega) \), the function \( g(t, x; y, z, v; \cdot) : [-1, 1] \rightarrow \mathbb{R} \) is Lipschitz continuous.
\end{enumerate}

Filippov’s Lemma 8.20 provides a Lebesgue measurable function \( \eta : [t_0, T] \times \Omega \rightarrow [-1, 1] \) such that
\[
f(t, x) = g\left(t, x; u(t, x), \operatorname{ess sup}_{t - \theta - \Theta(x) \leq s \leq t} [u(s, x)]_c, u(t, \cdot); \eta(t, x)\right)
\]
holds for Lebesgue-almost all \( t \in [t_0, T] \) and \( x \in \Omega \). Now Corollary 3.2 is applied to a single-valued mapping on the right-hand side of (14) and so, for each \( \psi \in C_\gamma \), we obtain the existence of a (both strong and mild) solution \( v : (-\infty, T) \rightarrow L^2(\Omega) \) of
\[
\begin{cases}
\partial_t v - \Delta v = g\left(t, x; v(t, x), \operatorname{ess sup}_{t - \theta - \Theta(x) \leq s \leq t} [v(s, x)]_c, v(t, \cdot); \eta(t, x)\right) & \text{a.e. in } (t_0, T) \times \Omega \\
v = \psi(\cdot - t_0) & \text{a.e. in } (-\infty, t_0] \times \Omega \\
v = 0 & \text{on } (t_0, T) \times \partial \Omega
\end{cases}
\]
Similarly to \( u \), we conclude \( v \in W^{1,2}([t_0, T] \times \Omega) \) from Corollary 8.16 and linear growth assumption (G5). The difference \( w := u - v : [t_0, T] \rightarrow \mathbb{R} \) also belongs to \( W^{1,2}([t_0, T] \times \Omega) \). Thus, \( t \mapsto w(t, x) \) is absolutely continuous in \([t_0, T]\) for Lebesgue-almost every \( x \in \Omega \) (see, e.g., [46] Theorem 3.1.2 (g) & Lemma 3.1.1, [61] Theorem 2.1.4 cited in [37] Lemma 4.4). In particular, we obtain for Lebesgue-almost every \( x \in \Omega \) and every \( t \in [t_0, T] \)
\[
|w(t, x)| \leq |w(t_0, x)| + \int_{t_0}^t |\partial_s w(s, x)| \, ds,
\]
which implies that
\[
\ell_s \leq \int_{t_0}^t |\partial_s w(s, x)| \, ds.
\]
Hence
\[
\left( \ell_s \right)^2 \leq 2 \left| w(t_0, x) \right|^2 + 2 \cdot \left( \int_{t_0}^t |\partial_s w(s, x)| \, ds \right)^2
\]
\[
\leq 2 \left| w(t_0, x) \right|^2 + 2 \cdot (t - t_0) \cdot \int_{t_0}^t |\partial_s w(s, x)|^2 \, ds.
\]
Thus, the function \( x \mapsto \ell_s \) belongs to \( L^2(\Omega) \) for each \( t \in [t_0, T] \) and, it satisfies
\[
\left\| \ell_s \right\|_{L^2(\Omega)} \leq 2 \left\| w(t_0) \right\|_{L^2(\Omega)}^2 + 2 \cdot (t - t_0) \cdot \|\partial_s w\|_{L^2([t_0, t] \times \Omega)}^2. \tag{34}
\]
In regard to an explicit upper \( L^2(\Omega) \) bound of this essential supremum, Lemma \( 8.15 \) will lead to an a priori estimate of \( \|\partial_s w\|_{L^2([t_0, t] \times \Omega)}^2 \) in a moment since \( w = u - v \) solves an initial-boundary value problem of the non-homogeneous heat equation. In more details, Lipschitz hypothesis \((G3')\) and its consequences for \( g \) imply for Lebesgue-almost every \( t \in [t_0, T], x \in \Omega \)
\[
\begin{align*}
|f(t, x) - g(t, x; v(t, x), \ell_s)\|_{L^2(\Omega)} &\leq \hat{C}(t) \cdot \left( \left\| u(t, x) - v(t, x) \right\|_{L^2(\Omega)} + \left\| u(t) - v(t) \right\|_{L^2(\Omega)} + \left\| \partial_s w \right\|_{L^2([t_0, t] \times \Omega)}^2 \right) \\
&\leq \hat{C}(t) \cdot \left( \left\| u(t, x) - v(t, x) \right\|_{L^2(\Omega)} + \left\| u(t) - v(t) \right\|_{L^2(\Omega)} + \left\| \partial_s w \right\|_{L^2([t_0, t] \times \Omega)}^2 \right) + \left\| \partial_s \right\|_{L^2([t_0, t] \times \Omega)}^2.
\end{align*}
\]
Due to the general inequality \( \left( \sum_{j=1}^{4} \alpha_j \right)^2 \leq 2^{3} \cdot \sum_{j=1}^{4} \alpha_j^2 \) for all \( \alpha_1, \ldots, \alpha_4 \in \mathbb{R} \), the integration w.r.t. \( x \in \Omega \) reveals with some constant \( C = C(\Omega, \hat{C}) > 0 \)
\[
\begin{align*}
\left\| f(t, \cdot) - g(t, \cdot; v(t, \cdot), \ell_s)\right\|_{L^2(\Omega)}^2 &\leq C \hat{C}^2 \left( \left\| u(t) - v(t) \right\|_{L^2(\Omega)}^2 + \left\| \ell_s \right\|_{L^2(\Omega)}^2 \right) \\
&\leq C \hat{C}^2 \left( \left\| u(t) - v(t) \right\|_{L^2(\Omega)}^2 + \left\| \ell_s \right\|_{L^2(\Omega)}^2 \right) + \left\| \partial_s \right\|_{L^2([t_0, t] \times \Omega)}^2.
\end{align*}
\]
Due to equation (34), the monotone increasing auxiliary function \( \delta \) satisfies the following implicit inequality with a constant \( C \) for every \( t \in [t_0, T] \):

\[
\| u(t) - v(t) \|_{L^2(\Omega)} \leq \left\| \text{ess sup}_{t_0 \leq s \leq t} \left| u(s) - v(s) \right| \right\|_{L^2(\Omega)}
\]

for every \( t \in [t_0, T] \), i.e., with a modified constant (again denoted by) \( C = C(\Omega, \tilde{c}) \):

\[
\left\| f(t, \cdot) - g(t, \cdot), \text{ess sup}_{t-\Theta(t) \leq s \leq t} \left| v(s, \cdot) \right|, \text{ess sup}_{t-\Theta(t) \leq s \leq t} \left| v(s, \cdot) \right|, \text{ess sup}_{t-\Theta(t) \leq s \leq t} \left| v(s, \cdot) \right| \right\|_{L^2(\Omega)}
\]

\[
\leq C \cdot \Lambda(t)^2 \cdot \left( \rho_0^2 + \left\| \text{ess sup}_{t_0 \leq s \leq t} \left| u(s) - v(s) \right| \right\|_{L^2(\Omega)} \right)
\]

Now Lemma 8.15 provides an a priori estimate of \( \| \partial_s w \|_{L^2([t_0, t] \times \Omega)} \) for every \( t \in [t_0, T] \):

\[
\| \partial_s w \|_{L^2([t_0, t] \times \Omega)} \leq C \cdot \left( \| \nabla_x w(t_0) \|_{L^2(\Omega)} + \right)
\]

\[
\int_{t_0}^{t} \Lambda(s)^2 \cdot \left( \rho_0^2 + \left\| \text{ess sup}_{t_0 \leq \sigma \leq s} \left| w(\sigma) \right| \right\|_{L^2(\Omega)} \) \right) \ ds.
\]

Due to equation (34), the monotone increasing auxiliary function \( \delta : [t_0, T] \to \mathbb{R} \) defined by

\[
\delta(t) := \left\| \text{ess sup}_{t_0 \leq s \leq t} \left| u(s, \cdot) - v(s, \cdot) \right| \right\|_{L^2(\Omega)}
\]

satisfies the following implicit inequality with a constant \( C = C(\Omega, \tilde{c}) \):

\[
\delta(t) \leq C \cdot (1 + t - t_0) \cdot \| \phi(0) - \psi(0) \|_{W^{1,2}(\Omega)} + C \cdot (t - t_0) \cdot \rho_0^2 \cdot \int_{t_0}^{t} \Lambda(s)^2 \ ds
\]

\[
+ C \cdot (t - t_0) \cdot \int_{t_0}^{t} \Lambda(s)^2 \cdot \delta(s) \ ds.
\]

The Gronwall inequality for monotone increasing (and not necessarily continuous) functions (see, e.g., [33 Proposition A.1]) bridges the gap to an explicit estimate and leads to the following upper bound of \( \delta(t) \):

\[
C \cdot \left( (1 + t - t_0) \cdot \| \phi(0) - \psi(0) \|_{W^{1,2}(\Omega)} + (t - t_0) \cdot \rho_0^2 \cdot \| \Lambda \|_{L^2([t_0, t])}^2 \right) e^{C \cdot (t - t_0) \cdot \| \Lambda \|_{L^2([t_0, t])}^2}
\]

\[
\leq C \cdot \left( e^{t - t_0} \cdot \| \phi(0) - \psi(0) \|_{W^{1,2}(\Omega)} + e^{t - t_0} \cdot \rho_0^2 \cdot \| \Lambda \|_{L^2([t_0, t])}^2 \right) e^{C \cdot (t - t_0) \cdot \| \Lambda \|_{L^2([t_0, t])}^2}
\]

\[
\leq C \cdot \rho_0^2 \left( 1 + \| \Lambda \|_{L^2([t_0, t])}^2 \right) e^{C \cdot (t - t_0) \cdot \| \Lambda \|_{L^2([t_0, t])}^2}
\]

\[
\leq C \cdot \rho_0^2 \cdot e^{C \cdot \| \Lambda \|_{L^2([t_0, t])}^2}
\]

for each \( t \in [t_0, T] \) with adapted constants (always denoted by) \( C = C(\Omega, \tilde{c}, T - t_0) \) \( \geq 1 \). As a direct consequence of inequality (35), the proof of estimate (31) in Lemma 8.17 is completed.
Proof of Corollary 8.18. Choose any \( \phi, \psi \in C_\gamma \) with \( \phi(0), \psi(0) \in W^{1,2}(\Omega) \) and \( (t, t_0) \in \mathbb{R}_+^2 \) as well as any \( y_t(\cdot; t_0, \phi) \in U(t, t_0, \phi) \subset C_\gamma \). In particular, \( u : (-\infty, t) \rightarrow L^2(\Omega) \) defined by \( u(s) := y_t(-t - s); t_0, \phi) \) is a solution to parabolic inclusion (15) (restricted to \( [0, t_0] \) instead of \([0, T]\)).

Now Lemma 8.17 guarantees the existence of a (both strong and mild) solution \( v : (-\infty, t) \rightarrow L^2(\Omega) \) to the corresponding parabolic inclusion (30) related to initial function \( \psi \) (instead of \( \phi \)) and satisfying

\[
\|u(s) - v(s)\|_{L^2(\Omega)} + \left\| \text{ess sup}_{t_0 \leq \sigma \leq s} |u(\sigma, \cdot)| - |v(\sigma, \cdot)|\right\|_{L^2(\Omega)} \leq C \cdot \rho_0 \cdot e^{C \cdot \|\Lambda\|_{L^2(\Omega, t_0, t)}}
\]

for every \( s \in [t_0, t] \) with

\[
\rho_0 := \|\psi(0) - \phi(0)\|_{W^{1,2}(\Omega)} + \left\| \text{ess sup}_{-\theta - \Theta(\cdot) \leq s \leq 0} |\psi(s)| - |\phi(s)|\right\|_{L^2(\Omega)}
\]

and a constant \( C = C(\Omega, t - t_0) \geq 1 \). In particular, \( z := v(\cdot + t) : (-\infty, 0) \rightarrow L^2(\Omega) \) belongs to \( U(t, t_0, \psi) \subset C_\gamma \) as specified in Proposition 5.3 and, we obtain

\[
d_{c,\gamma}(y_t(\cdot; t_0, \phi), z)
\]

\[
\sup_{\sigma \leq 0} \left( e^{\gamma \sigma} \cdot \|y_t(\sigma; t_0, \phi) - z(\sigma)\|_{L^2(\Omega)} \right)
\]

\[
+ \left\| \text{ess sup}_{\sigma \leq 0} \|[y_t(\sigma; t_0, \phi)]_c - [z(\sigma)]_c\right\|_{L^2(\Omega)}
\]

\[
= \sup_{s \leq t} \left( e^{\gamma(s-t)} \cdot \|u(s) - v(s)\|_{L^2(\Omega)} \right) + \left\| \text{ess sup}_{s \leq t} |u(s)|_c - |v(s)|_c\right\|_{L^2(\Omega)}
\]

\[
\leq e^{-\gamma t} \cdot \max \left\{ e^{\gamma t_0} \|\phi - \psi\|_{L^2(\Omega)} + \left( e^{\gamma s} \cdot \|u(s) - v(s)\|_{L^2(\Omega)} \right) \right\}
\]

\[
+ \left\| \text{ess sup}_{\sigma \leq 0} \|\phi(\sigma)|_c - |\psi(\sigma)|_c\right\|_{L^2(\Omega)} + \left\| \text{ess sup}_{s \leq t} |u(s)|_c - |v(s)|_c\right\|_{L^2(\Omega)}
\]

\[
\leq d_{c,\gamma}(\phi, \psi) + e^{-\gamma t} \cdot \sup_{s \leq t} \left( e^{\gamma s} \cdot \|u(s) - v(s)\|_{L^2(\Omega)} \right)
\]

\[
+ \left\| \text{ess sup}_{t_0 \leq s \leq t} |u(s)|_c - |v(s)|_c\right\|_{L^2(\Omega)}
\]

\[
\leq d_{c,\gamma}(\phi, \psi) + C \cdot \rho_0 \cdot e^{C \cdot \|\Lambda\|_{L^2(\Omega, t_0, t)}}
\]

with a constant \( C = C(\Omega, t - t_0) > 0 \) and

\[
\rho_0 \overset{\text{Def.}}{=} \|\psi(0) - \phi(0)\|_{W^{1,2}(\Omega)} + \left\| \text{ess sup}_{-\theta - \Theta(\cdot) \leq s \leq 0} |\psi(s)|_c - |\phi(s)|_c\right\|_{L^2(\Omega)}
\]

i.e., the claimed estimate (32) holds:

\[
d_{c,\gamma}(y_t(\cdot; t_0, \phi), z) \leq \text{const}(\Omega, \|\Lambda\|_{L^2(\Omega, t_0, t)} \cdot (d_{c,\gamma}(\phi, \psi) + \|\psi(0) - \phi(0)\|_{W^{1,2}(\Omega)})).
\]

Finally inequality (33) results from essentially the same arguments (just considering the essential suprema w.r.t. space-dependent intervals in time instead of \((-\infty, 0])\).
Lemma 8.22. As in Corollary 8.13, suppose assumptions\( (G1) - (G6) \) and consider the MNDS \( U \subset C_\gamma \supseteq C_\gamma \) specified in Proposition 5.3 by means of parabolic differential inclusion \( (15) \). Define \( \tilde{B} = \{ B(t) : t \in \mathbb{R} \} \) consisting of “closed balls” \( B(t) := \{ \phi \in C_\gamma \mid \| \phi \|_{\gamma} \leq \rho(t) \} \) with radii \( \rho(t) := \rho_0 \cdot \left( 1 + e^{-\left(\lambda_1 - \beta\right)t} \right) \in (\rho_0, \infty) \), \( \rho_0 := \frac{(\bar{\alpha} + \bar{\beta} c) \cdot L^\alpha(\Omega)^{1/2}}{\lambda_1 - \beta} > 0 \). (36)

Then \( U \) is pullback asymptotically compact w.r.t. \( \tilde{B} \) and the norm \( \| \cdot \|_{\gamma} \), i.e., for any real sequence \( t_n \to +\infty \), every sequence \( y_n \in U(t, t - t_n, B(t - t_n)) \) has a subsequence converging w.r.t. \( \| \cdot \|_{\gamma} \).

Proof. Fix \( t \in \mathbb{R} \) arbitrarily. Consider any sequences \( t_n \to +\infty \) and \( (\phi_n)_{n \in \mathbb{N}} \), \( (y_n)_{n \in \mathbb{N}} \) in \( C_\gamma \) with \( y_n \in U(t, t - t_n, \phi_n) \) and \( \phi_n \in B(t - t_n) \) for each \( n \in \mathbb{N} \), i.e., in particular, \( \| \phi_n \|_{\gamma} \leq \rho(t - t_n) \) and

\[
y_n(s) = \begin{cases} S(t_n + s) \phi_n(0) + \int_{t - t_n}^{t + s} S(t + s - \tau) f_n(\tau) \, d\tau & \text{if } s \in [-t_n, 0] \\ \phi_n(s + t_n) & \text{if } s < -t_n\end{cases}
\]

with a selector \( f_n : [t - t_n, t] \to L^2(\Omega) \) satisfying \( f_n(\tau) \in \mathbb{G}(\tau, y_n(\cdot + \tau - t)|_{(-\infty, 0]} \) for a.e. \( \tau \). Corollary 8.13 (1.) states the positive invariance of \( \tilde{B} \) w.r.t. \( U \) and so, \( y_n(\cdot + s - t)|_{(-\infty, 0]} \in U(s, t - t_n, \phi_n) \subset U(s, t - t_n, B(t - t_n)) \subset B(s) \) (37) holds for every \( s \in [t - t_n, t] \).

As a consequence of Cantor’s diagonal method, it suffices to show for every \( \varepsilon > 0 \) and infinite index set \( J \subset \mathbb{N} \) that there exist an infinite subset \( J_\varepsilon \subset J \) and some \( \zeta \in C_\gamma \) with \( \| y_j - \zeta \|_{\gamma} < \varepsilon \) for all \( j \in J_\varepsilon \). Obviously, it holds for every \( n \in \mathbb{N} \)

\[
\sup_{s < -t_n} \left( e^{\gamma s} \cdot \| \phi_n(s + t_n) \|_{L^2(\Omega)} \right) \leq \sup_{s < -t_n} \left( e^{\gamma s} \cdot \| \phi_n \|_{\gamma} \right) e^{-\gamma(t + t_n)} \leq \rho_0 \cdot (1 + e^{-\left(\lambda_1 - \beta\right)(t + t_n)}),
\]

Furthermore, we obtain for every \( n \in \mathbb{N} \) and \( s \in [-t_n, 0] \)

\[
\| S(t_n + s) \, \phi_n(0) \|_{L^2(\Omega)} \leq e^{-\lambda_1 (t + t_n + s)} \| \phi_n(0) \|_{L^2(\Omega)} \leq e^{-\lambda_1 (t + t_n + s)} \| \phi_n \|_{\gamma} \leq e^{-\lambda_1 (t + t_n + s)} \rho_0 \cdot \left( 1 + e^{-\left(\lambda_1 - \beta\right)(t - t_n)} \right),
\]

so

\[
\sup_{s \in [-t_n, 0]} \left( e^{\gamma s} \cdot \| S(t_n + s) \, \phi_n(0) \|_{L^2(\Omega)} \right) \leq \rho_0 \cdot (1 + e^{-\left(\lambda_1 - \beta\right)(t + t_n)}) \cdot \sup_{s \in [-t_n, 0]} e^{\gamma s - \lambda_1 (t + t_n + s)} \leq \rho_0 \cdot \left( 1 + e^{-\left(\lambda_1 - \beta\right)(t - t_n)} \right) \cdot e^{-\lambda_1 t_n} = \rho_0 \cdot \left( e^{-\lambda_1 t_n} + e^{-\left(\lambda_1 - \beta\right)t_n} \right).\]
In addition, the following estimate is satisfied for all $n \in \mathbb{N}$ and $s \in [-t_n, 0]$

$$\left\| \int_{t-t_n}^{t+s} S(t+s-\tau) f_n(\tau) \, d\tau \right\|_{L^2(\Omega)} \leq \int_{t-t_n}^{t+s} e^{-\lambda_1 (t+s-\tau)} \|f_n(\tau)\|_{L^2(\Omega)} \, d\tau$$

Due to the assumption $\gamma > \lambda_1$ and the prior estimate

$$\|f_n(\tau)\|_{L^2(\Omega)} \leq C + \bar{\beta} \cdot \rho_0 \left(1 + e^{-\lambda_1 \tau} \right),$$

we can choose $N_\varepsilon \in \mathbb{N}$, $T_\varepsilon > 1$ and $\delta_\varepsilon \in [0, 1]$ successively such that for all $n \geq N_\varepsilon$ and $-t_n < -T_\varepsilon - 2$,

$$\begin{align*}
\sup_{s < -t_n} \left( e^{\gamma s} \cdot \|\phi_n(s + t_n)\|_{L^2(\Omega)} \right) &< \frac{\varepsilon}{8}, \\
\sup_{s \in [-t_n, 0]} \left( e^{\gamma s} \cdot \|S(t_n + s) \phi_n(0)\|_{L^2(\Omega)} \right) &< \frac{\varepsilon}{8}, \\
\sup_{s \in [-t_n, -T_\varepsilon]} \left( e^{\gamma s} \cdot \left\| \int_{t-t_n}^{t+s} S(t+s-\tau) f_n(\tau) \, d\tau \right\|_{L^2(\Omega)} \right) &< \frac{\varepsilon}{8}, \\
\sup_{s \in [-T_\varepsilon, 0]} \left( e^{\gamma s} \cdot \left\| \int_{t+s-\delta_\varepsilon}^{t+s} S(t+s-\tau) f_n(\tau) \, d\tau \right\|_{L^2(\Omega)} \right) &< \frac{\varepsilon}{8}.
\end{align*}$$

This implies

$$\sup_{s \leq -T_\varepsilon} \left( e^{\gamma s} \cdot \|y_n(s)\|_{L^2(\Omega)} \right) < \frac{\varepsilon}{4} \quad \text{for every } n \geq N_\varepsilon. \quad (40)$$

For further conclusions about compactness in $L^2(\Omega)$, we consider the following representation for every $n \geq N_\varepsilon$ and $s \in [-T_\varepsilon, 0]$

\begin{align*}
y_n(s) &= S(t_n + s) \phi_n(0) + \int_{t-t_n}^{t+s} S(t+s-\tau) f_n(\tau) \, d\tau \\
&= S(t_n + s) \phi_n(0) + \int_{t-t_n}^{t-T_\varepsilon-1} S(t+s-\tau) f_n(\tau) \, d\tau \\
&\quad + \int_{t-T_\varepsilon}^{t+s-\delta_\varepsilon} S(t+s-\tau) f_n(\tau) \, d\tau + \int_{t+s-\delta_\varepsilon}^{t+T_\varepsilon} S(t+s-\tau) f_n(\tau) \, d\tau \\
&= S(s+T_\varepsilon+1) \left( S(t_n-T_\varepsilon+1) \phi_n(0) + \int_{t-T_\varepsilon}^{t+s-\delta_\varepsilon} S(t+s-\tau) f_n(\tau) \, d\tau \right) \\
&\quad + \int_{t-T_\varepsilon-1}^{t+s-\delta_\varepsilon} S(t+s-\tau) f_n(\tau) \, d\tau + \int_{t+s-\delta_\varepsilon}^{t+T_\varepsilon} S(t+s-\tau) f_n(\tau) \, d\tau \\
&= S(s+T_\varepsilon+1) \ y_n(-T_\varepsilon) + \int_{t-T_\varepsilon-1}^{t+s-\delta_\varepsilon} S(t+s-\tau) f_n(\tau) \, d\tau
\end{align*}
is uniformly continuous. This observation implies first that the image set \( \zeta \) is relatively compact. Hence the theorem of Arzelà-Ascoli (as in, e.g., [28]) ensures that the sequence of functions from \([-T_\varepsilon, 0]\) is also relatively compact in \( L^2(\Omega) \) for all \( n \geq N_\varepsilon \).

Furthermore, the sequence of functions

\[
[-T_\varepsilon, 0] \ni s \mapsto S(s+T_\varepsilon+1) y_n(-T_\varepsilon-1) \in L^2(\Omega), \quad n \geq N_\varepsilon,
\]

is relatively compact in \( C^0([-T_\varepsilon, 0], L^2(\Omega)) \). Indeed, \( y_n(-T_\varepsilon-1) \in B(-T_\varepsilon-1) \) for all \( n \in \mathbb{N} \) implies that \( \{y_n(-T_\varepsilon-1) \mid n \geq N_\varepsilon\} \) is bounded in \( L^2(\Omega) \) and so, the auxiliary set

\[
M_1 := \{ S(1) y_n(-T_\varepsilon-1) \mid n \geq N_\varepsilon\} \subset L^2(\Omega)
\]

is relatively compact in \( L^2(\Omega) \) since the heat semigroup \( (S(t))_{t \geq 0} \) on \( L^2(\Omega) \) is immediately compact. Due to [24, Lemma I.5.2], the continuity of \( L \) is relatively compact in \( N \) for \( t \geq 0 \) because if this sequence proves to be relatively compact w.r.t. supremum norm (due to the theorem of Arzelà-Ascoli) then so is the sequence of respective restrictions to \( \{s = \sigma\} \subset [-T_\varepsilon, 0]^2 \). In regard to the image sets, estimate [39] and Lemma [21] guarantee that

\[
\left\{ \int_{t-T_\varepsilon-1}^{t+\delta_\varepsilon} S(t+\sigma-\delta_\varepsilon-\tau) f_n(\tau) \ d\tau \mid \sigma, s \in [-T_\varepsilon, 0], \sigma \geq s, n \geq N_\varepsilon \right\} \subset L^2(\Omega)
\]

is bounded and so, the union of the image sets

\[
\left\{ S(\delta_\varepsilon) \int_{t-T_\varepsilon-1}^{t+\delta_\varepsilon} S(t+\sigma-\delta_\varepsilon-\tau) f_n(\tau) \ d\tau \mid \sigma, s \in [-T_\varepsilon, 0], \sigma \geq s, n \geq N_\varepsilon \right\} \subset L^2(\Omega)
\]
is relatively compact since \( S(\delta_z) : L^2(\Omega) \to L^2(\Omega) \) is a compact linear operator. Moreover, the same conclusions from \([27]\) and \( \| y_n \|_\gamma \leq \rho(t) \) as on our way to inequality \([39]\) reveal that

\[
\Lambda_\varepsilon := \sup_{n \geq N_\varepsilon} \sup_{\sigma \in [-T_\varepsilon, 0]} \sup_{\tau \in [t-T_\varepsilon, t+\sigma]} \| S(t+\sigma-\tau) f_n(\tau) \|_{L^2(\Omega)} < \infty
\]

and so, the family of integral functions

\[
F_n(\sigma, \cdot)\mid_{[-T_\varepsilon, \sigma]} : [-T_\varepsilon, \sigma] \to L^2(\Omega)
\]

for all \( n \geq N_\varepsilon, \sigma \in [-T_\varepsilon, 0] \) is \( \Lambda_\varepsilon \)-Lipschitz continuous. It is worth mentioning here that this Lipschitz constant is uniform w.r.t. both \( n \geq N_\varepsilon \) and \( \sigma \in [-T_\varepsilon, 0] \).

Next, we verify equicontinuity of the family \( \{ F_n(\cdot, s) \mid_{[s, 0]} \mid n \geq N_\varepsilon, s \in [-T_\varepsilon, 0] \} \) because then the triangle inequality leads to the equicontinuity of \( (F_n)_{n \geq N_\varepsilon} \) on their joint domain \([-T_\varepsilon, 0]^2 \cap \mathbb{R}_2^\varepsilon\). The representation

\[
F_n(\sigma, s) = \int_{t-T_\varepsilon-1}^{t+s-\delta_\varepsilon} S(t+\sigma-\tau) f_n(\tau) \, d\tau
\]

lays the basis for the same arguments as before: The auxiliary set

\[
M_2 := \{ S(\delta_z) f_n(\tau) \mid n \geq N_\varepsilon, \tau \in [-T_\varepsilon-1, 0] \} \subset L^2(\Omega)
\]

results from a bounded set in \( L^2(\Omega) \) being mapped by the compact linear operator \( S(\delta_z) \) and so, it is relatively compact in \( L^2(\Omega) \). Due to \([24]\) Lemma I.5.2] (again), the function

\[
[0, T_\varepsilon+1] \times (M_2, \| \cdot \|_{L^2(\Omega)}) \ni (\tau', \zeta) \mapsto S(\tau') \zeta \in (L^2(\Omega), \| \cdot \|_{L^2(\Omega)})
\]

is uniformly continuous. This implies equicontinuity of all the functions

\[
F_n(\cdot, s) \mid_{[s, 0]} : [s, 0] \ni \sigma \mapsto \int_{t-T_\varepsilon-1}^{t+s-\delta_\varepsilon} S(t+\sigma-\delta_z-\tau) S(\delta_z) f_n(\tau) \, d\tau
\]

for \( n \geq N_\varepsilon, s \in [-T_\varepsilon, 0] \) by means of the simple estimate for any \( \sigma_1, \sigma_2 \in [s, 0] \)

\[
\| F_n(\sigma_1, s) - F_n(\sigma_2, s) \|_{L^2(\Omega)} \leq \int_{t-T_\varepsilon-1}^{t+s-\delta_\varepsilon} \| (S(t+\sigma_1-\delta_z-\tau) - S(t+\sigma_2-\delta_z-\tau)) S(\delta_z) f_n(\tau) \|_{L^2(\Omega)} \, d\tau.
\]

Together with the \( \Lambda_\varepsilon \)-Lipschitz continuity of all the functions \( F_n(\sigma, \cdot)\mid_{[-T_\varepsilon, \sigma]} \) \( n \geq N_\varepsilon, \sigma \in [-T_\varepsilon, 0] \), we conclude from the triangle inequality (similarly to the proof of \([37]\) Lemma 2.21]) that

\[
F_n : [-T_\varepsilon, 0]^2 \cap \mathbb{R}_2^\varepsilon \ni (\sigma, s) \mapsto \int_{t-T_\varepsilon-1}^{t+s-\delta_\varepsilon} S(t+\sigma-\tau) f_n(\tau) \, d\tau \in L^2(\Omega), \ n \geq N_\varepsilon,
\]

are equicontinuous (as functions of two variables though). The relative compactness of the image set in \( L^2(\Omega) \) has already been verified and so, the theorem of Arzelà-Ascoli guarantees that \( (F_n)_{n \geq N_\varepsilon} \) is relatively compact in \( C^0([-T_\varepsilon, 0]^2 \cap \mathbb{R}_2^\varepsilon, L^2(\Omega)) \).
Thus, the respective restrictions to the “diagonal”, i.e.,
\[ [-T_\varepsilon, 0] \ni s \mapsto F_n(s, s) = \int_{t-T_\varepsilon - 1}^{t+s-\delta_\varepsilon} S(t+s-\tau) f_n(\tau) \, d\tau \in L^2(\Omega), \quad n \geq N_\varepsilon, \]
are also relatively compact in \( C^0([-T_\varepsilon, 0], L^2(\Omega)) \).

Finally we again consider the complete representation
\[
y_n(s) = S(s+T_\varepsilon + 1) y_n(-T_\varepsilon - 1) + \int_{t-T_\varepsilon - 1}^{t+s-\delta_\varepsilon} S(t+s-\tau) f_n(\tau) \, d\tau
\]
and conclude from these two aspects of relative compactness in \( C^0([-T_\varepsilon, 0], L^2(\Omega)) \) that for every infinite index set \( J \subset \mathbb{N} \), there is a sequence \( n_k \not\to \infty \) in \( J \) with the following three features:

- \( n_k \geq N_\varepsilon \) for all \( k \in \mathbb{N} \),
- \( (s \mapsto S(s+T_\varepsilon + 1) y_{n_k}(-T_\varepsilon - 1))_{k \in \mathbb{N}} \) converges in \( C^0([-T_\varepsilon, 0], L^2(\Omega)) \),
- \( (s \mapsto \int_{t-T_\varepsilon - 1}^{t+s-\delta_\varepsilon} S(t+s-\tau) f_{n_k}(\tau) \, d\tau)_{k \in \mathbb{N}} \) converges in \( C^0([-T_\varepsilon, 0], L^2(\Omega)) \),
- \( \sup_{k \in \mathbb{N}} \sup_{s \in [-T_\varepsilon, 0]} \left( e^{\gamma s} \left\| \int_{t+s-\delta_\varepsilon}^{t+s} S(t+s-\tau) f_{n_k}(\tau) \, d\tau \right\|_{L^2(\Omega)} \right) < \frac{\varepsilon}{8} \).

Due to the underlying Cauchy property w.r.t. the supremum norm on \([T_\varepsilon, 0]\), there is an index \( \kappa = \kappa(\varepsilon, J) \in \mathbb{N} \) sufficiently large such that for all \( k \geq \kappa \),

\[
\sup_{s \in [-T_\varepsilon, 0]} \left( e^{\gamma s} \left\| y_{n_k}(s) - y_{n_\kappa}(s) \right\|_{L^2(\Omega)} \right) < \frac{\varepsilon}{2}.
\]

Together with the preceding estimate \( 40 \) about \( s \leq -T_\varepsilon \), we obtain for all indices \( k \geq \kappa \)

\[
\left\| y_{n_k} - y_{n_\kappa} \right\|_{L^2(\Omega)} = \sup_{s \leq 0} \left( e^{\gamma s} \left\| y_{n_k}(s) - y_{n_\kappa}(s) \right\|_{L^2(\Omega)} \right) < \varepsilon.
\]

\( \square \)

8.9. Preparing the pullback asymptotic compactness (w.r.t. \( d_{c, \gamma, p.i.p.} \)) for solutions to \( 15 \).

**Lemma 8.23.** Let \( \Omega \subset \mathbb{R}^d \) be bounded open with smooth boundary, and suppose hypotheses \((G1) - (G5)\). Fix real \( t_0 < t \), \( 0 < \tau < t - t_0 \), and consider a sequence \((\phi_k)_{k \in \mathbb{N}} \) in \( C_0 \) bounded w.r.t. \( \| \cdot \|_{\gamma} \).

Then every sequence \((y_\varepsilon(\cdot; t_0, \phi_k))_{k \in \mathbb{N}} \) induced by respective solutions to parabolic differential inclusion \( 15 \) is relatively compact with respect to the norm of \( L^2(\Omega); L^\infty([-\tau, 0]) \).

The basic idea for proving Lemma 8.23 is to find a candidate for a converging subsequence by means of standard compactness theorems for \( L^2([-\tau, 0] \times \Omega) \) first. Then we will use former results from \( 37 \) to conclude that this subsequence converges to the same limit function even w.r.t. the norm of \( L^2(\Omega); L^\infty([-\tau, 0]) \). In more detail, we will apply the following direct consequence of \( 37 \) Lemmas 4.6, 4.7 which is stated here without proof:

**Lemma 8.24.** Let \( \Omega \subset \mathbb{R}^d \) be a bounded open set with smooth boundary and fix \( \tau > 0 \). Suppose for the sequence \((v_k)_{k \in \mathbb{N}} \) and the function \( v \) in \( W^{1,2}([-\tau, 0] \times \Omega) \) that
Then, the sequence of functions \( x \mapsto \text{ess sup}_{-\tau \leq s \leq 0} v_k(s, x) \), \( k \in \mathbb{N} \), defined on \( \Omega \) converges to \( x \mapsto \text{ess sup}_{-\tau \leq s \leq 0} v(s, x) \) in \( L^2(\Omega) \).

Proof of Lemma 8.23. Consider any sequence \( (u^{(k)})_{k \in \mathbb{N}} \) with \( u^{(k)} = y_t(\cdot; t_0, \phi_k) \in U(t, t_0, \phi_k) \subset C_0 \) for \( k \in \mathbb{N} \) and the sequence \( (f_k)_{k \in \mathbb{N}} \) of respective selectors, i.e., \( f_k(s) \in G(s, u^{(k)}(\cdot + s - t))_{(-\infty,0]} \) for a.e. \( s \in [t_0, t] \). According to growth assumption (G5),

\[
|f_k(s, x)| \leq \alpha(s) + \beta \cdot \left(|u^{(k)}((t-s), x)| + c\right)
\]

hence

\[
\|f_k(s)\|_{L^2(\Omega)} \leq (\alpha + \beta c) \cdot \|u^{(k)}((t-s))\|_{L^2(\Omega)}.
\]

Estimate (28) for any weak solution and the Gronwall inequality lead to an a priori bound \( \rho > 0 \) with

\[
\sup \left\{ \|u^{(k)}((t-s))\|_{L^2(\Omega)}, \|f_k(s)\|_{L^2(\Omega)} \right\} \leq \rho.
\]

As a further consequence, there is a constant \( C = C(\rho, \Omega, t-t_0) > 0 \) such that

\[
\|\nabla_x u^{(k)}((t-s))\|_{L^2(\Omega)} \leq C \frac{\sqrt{s-t}}{\sqrt{t}}
\]

holds for all \( k \in \mathbb{N} \) and \( s \in (t_0, t] \). Indeed, the heat semigroup \( \{S(t)\}_{t \geq 0} \) on \( L^2(\Omega) \) subject to homogeneous Dirichlet boundary conditions is known to be analytic since \( \Omega \) is assumed to be a bounded open domain with smooth boundary (see, e.g., [85 § 7 Theorem 2.7]). Hence \( \|S(t)\|_{\text{Lip}(\Omega)} \) guarantees a constant \( c = c(\Omega) > 0 \) such that for any \( T > 0 \) and \( v \in L^2(\Omega) \), the function \( S(T)v \in L^2(\Omega) \) even belongs to \( W^{1,2}(\Omega) \cap W^{2,2}(\Omega) \) and satisfies

\[
\|D_x S(\tau) v\|_{L^2(\Omega)} \leq c \frac{\sqrt{T}}{\tau} \|v\|_{L^2(\Omega)}.
\]

Now we conclude from an interpolation inequality (as, e.g., in [8 § 9, page 313])

\[
\|\nabla_x S(\tau) v\|_{L^2(\Omega)} \leq \frac{\sqrt{T}}{\tau} \|v\|_{L^2(\Omega)}
\]

for all \( \tau \in (0, t-t_0] \) and \( v \in L^2(\Omega) \) with a (possibly different) constant \( c = c(\Omega, t-t_0) > 0 \). Finally the (piecewise) mild representation (16) leads to

\[
\|\nabla_x u^{(k)}((t-s))\|_{L^2(\Omega)}
\]

\[
\leq \|\nabla_x S(t-t_0) \phi_k(0)\|_{L^2(\Omega)} + \int_{t_0}^{s} \|\nabla_x S(s-\tau) f_k(\tau)\|_{L^2(\Omega)} d\tau
\]

\[
\leq \frac{c}{\sqrt{s-t_0}} \|\phi_k(0)\|_{L^2(\Omega)} + \int_{t_0}^{s} \frac{c}{\sqrt{s-\tau}} \rho \ d\tau \leq \frac{C}{\sqrt{s-t_0}}
\]

for all \( s \in (t_0, t] \) with a suitable constant \( C = C(\rho, \Omega, t-t_0) > 0 \).
for each \( s \in [-\tau, 0] \). Furthermore \( (\partial_t u^{(k)})_{k \in \mathbb{N}} \) is bounded in \( L^2([-\tau, 0] \times \Omega) \) as a consequence of Lemma 8.15.

The compactness theorems of Kakutani and Rellich-Kondrachov (e.g., [6] Theorems 3.17, 9.16) lead to a subsequence (again denoted by) \( (u^{(k)})_{k \in \mathbb{N}} \) and a function \( u \in W^{1,2}([-\tau, 0] \times \Omega) \) such that

\[
\begin{align*}
    u^{(k)} \rightharpoonup u & \quad \text{in } W^{1,2}([-\tau, 0] \times \Omega), \\
    u^{(k)} \rightarrow u & \quad \text{in } L^2([-\tau, 0] \times \Omega), \\
    u^{(k)}(s) \rightarrow u(s) & \quad \text{in } L^2(\Omega)
\end{align*}
\]

for \( k \rightarrow \infty \) and Lebesgue-almost every \( s \in [-\tau, 0] \), because the convergence in \( L^2([-\tau, 0] \times \Omega) \) implies the convergence in \( L^2(-\tau, 0; L^2(\Omega)) \) and hence the \( L^2(\Omega) \)-convergence of a subsequence Lebesgue-almost everywhere in \([-\tau, 0]\).

Finally we will apply Lemma 8.24 to the sequence \( \{[u^{(k)}]_c - [u]_c\}_{k \in \mathbb{N}} \) and the limit function \( \equiv 0 \) for obtaining the claimed convergence

\[
\text{ess sup}_{-\tau \leq s \leq 0} |[u^{(k)}(s, \cdot)]_c - [u(s, \cdot)]_c| \rightarrow 0 \quad \text{in } L^2(\Omega) \quad (k \rightarrow \infty).
\]

Indeed, for any Sobolev function \( v \in W^{1,2}([-\tau, 0] \times \Omega) \), it is known that both \( \max\{v, 0\} \) and \( \min\{v, 0\} \) also belong to the Sobolev space \( W^{1,2}([-\tau, 0] \times \Omega) \) with the weak derivatives

\[
D \max\{v, 0\} = \begin{cases} 
    Dv & \text{if } v > 0, \\
    0 & \text{if } v \leq 0
\end{cases}
\]

\[
D \min\{v, 0\} = \begin{cases} 
    Dv & \text{if } v < 0, \\
    0 & \text{if } v \geq 0
\end{cases}
\]

(see, e.g., [61] Corollary 2.1.8]). As a consequence, the functions \( v_k := [u^{(k)}]_c - [u]_c \), \( k \in \mathbb{N} \), and \( v = 0 \) are contained in \( W^{1,2}([-\tau, 0] \times \Omega) \) and satisfy the assumptions (i), (iii) of Lemma 8.24. The compactness theorem of Kakutani and the lower semi-continuity of norms w.r.t. weak convergence imply for Lebesgue-almost every \( s \in [-\tau, 0] \)

\[
\|u(s)\|_{W^{1,2}(\Omega)} \leq \liminf_{k \rightarrow \infty} \|u^{(k)}(s)\|_{W^{1,2}(\Omega)} \leq \text{const}(\rho, \Omega, t - t_0)
\]

and so, assumption (ii) of Lemma 8.24 is also fulfilled.

\[\square\]

8.10. **Pullback asymptotic compactness of \( U \) w.r.t. the metric \( d_{c,\gamma,p,i.p.} \).**

According to Lemma 8.22 the MNDS \( U \) is pullback asymptotically compact with respect to \( \tilde{B} \) (consisting of balls with time-dependent radii) and the norm \( \| \cdot \|_\gamma \). Now we extend this long-term feature of \( U \) and \( \tilde{B} \) to the metric \( d_{c,\gamma,p,i.p.} \).

**Lemma 8.25.** As in Corollary 8.13 and Lemma 8.22, suppose assumptions (G1) – (G6) and consider the MNDS \( U : \mathbb{R}^2_+ \times C_\gamma \rightharpoonup C_\gamma \) specified in Proposition 5.3 by means of differential inclusion [15]. Define \( \tilde{B} = \{B(t) : t \in \mathbb{R}\} \) consisting of \( B(t) := \{\phi \in C_\gamma \mid \|\phi\|_\gamma \leq \rho(t)\} \) with radii as in (56).

Then \( U \) is pullback asymptotically compact (even) w.r.t. \( \tilde{B} \) and the metric \( d_{c,\gamma,p,i.p.} \), i.e., for any real sequence \( t_n \rightarrow +\infty \), every sequence \( y_n \in U(t, t - t_n, B(t - t_n)) \) has a subsequence converging w.r.t. \( d_{c,\gamma,p,i.p.} \).

**Proof.** Consider any real sequence \( t_n \rightarrow +\infty \) and any sequence \( (y_n)_{n \in \mathbb{N}} \) in \( C_\gamma \) with \( y_n \in U(t, t - t_n, B(t - t_n)) \). According to Lemma 8.22 there exists a subsequence (again denoted by) \( (y_n)_{n \in \mathbb{N}} \) which converges to some \( y \in C_\gamma \) with respect to the norm \( \| \cdot \|_\gamma \).
Now fix \( \tau > 0 \) arbitrarily. It remains to prove for a subsequence \((y_n)_{n \in \mathbb{N}}\) of \((y_n)_{n \in \mathbb{N}}\)

\[
\| \text{ess sup}_{-\tau \leq s \leq 0} \left[ |y_n(s, \cdot)|_c - |y(s, \cdot)|_c \right] \|_{L^2(\Omega)} = \| |y_n|_c - |y|_c \|_{L^2(\Omega; L^\infty([-\tau,0]))} \to 0
\]

for \( j \to \infty \) because then Cantor’s diagonal method (in regard to \( \tau \to \infty \)) provides a further subsequence of \((y_n)_{n \in \mathbb{N}}\) converging to \( y \) w.r.t. \( d_{C,\gamma} \), as claimed.

Set \( t_0 := t - \tau - 1 \) and choose the index \( n_0 \in \mathbb{N} \) sufficiently large with \( t - t_n \leq t_0 \) for every \( n \geq n_0 \). Then for each \( n \geq n_0 \), we can select some \( \phi_n \in B(t_0) \subset C_\gamma \) with \( y_n = y_l(\cdot; t_0, \phi_n) \in U(t, t_0, \phi_n) \) in the sense \([16]\) of solutions to \([15]\) because \( \hat{B} \) is positively invariant w.r.t. the two-parameter semigroup \( U \) according to Lemma \([8,2\) and Corollary \([8,13,1]\).

Finally Lemma \([8,23]\) states that \((|y_n|_c)_{n \in \mathbb{N}}\) is relatively compact w.r.t. the norm of \( L^2(\Omega; L^\infty([-\tau,0])) \). Hence there exist a subsequence \((y_n)_{n \in \mathbb{N}}\) and a function \( z \in L^2(\Omega; L^\infty([-\tau,0])) \) with

\[
\| |y_n|_c - z \|_{L^2(\Omega; L^\infty([-\tau,0]))} \to 0 \quad (j \to \infty).
\]

It implies the convergence of a further subsequence to \( z \) Lebesgue-almost everywhere in \([-\tau,0] \times \Omega \) and so, we conclude from the convergence to \( y \) w.r.t. \( \| \cdot \| \), that \( z = |y|_c \) Lebesgue-almost everywhere in \([-\tau,0] \times \Omega \). \( \square \)

8.11. The proof of main Theorem \([4,2]\) Now we have all the tools for proving main Theorem \([4,2]\) (about infinite delay) under the additional hypothesis \((G7)\):

Proof of Theorem \([4,2]\). In short, it results from Theorem \([6,3]\) i.e., from the general existence theorem for pullback attractors of multi-valued non-autonomous dynamical systems in metric spaces. Indeed, we supply \( C_\gamma \) with the metric \( d_{C,\gamma,p.i.p.} \) defined in \([14]\). \((C_\gamma, d_{C,\gamma,p.i.p.})\) is complete according to Lemma \([5,2]\). Proposition \([5,3]\) specifies the strict MNDS \( U : \mathbb{R}_+^2 \times C_\gamma \to C_\gamma, (t_1, t_0, \phi) \mapsto u(\cdot - t_1) \) by means of the parabolic differential inclusion (with \( \theta := \infty \))

\[
\begin{cases}
\partial_t u - \Delta u \in G(t, x; u(t, x), \text{ess sup}_{-\infty < s \leq t} |u(s, x)|_c, u(t, \cdot)) \text{ a.e. in } (t_0, t_1) \times \Omega \\
\quad u = \phi(\cdot - t_0) \text{ a.e. in } (-\infty, t_0] \times \Omega \\
\quad u = 0 \text{ on } (t_0, t_1) \times \partial \Omega.
\end{cases}
\]

For each tuple \((t, \tau) \in \mathbb{R}_+^2\), the multi-valued mapping \( U((t, \tau), \cdot) : (C_\gamma, d_{C,\gamma,p.i.p.}) \to (C_\gamma, d_{C,\gamma,p.i.p.}) \) has a closed graph according to Lemma \([8,9]\).

Furthermore consider the set-valued mapping \( \hat{B} = \{ B(t) : t \in \mathbb{R} \} \) of “balls” \( B(t) := \{ \phi \in C_\gamma \ | \| \phi \|_\gamma \leq \rho(t) \} \) specified in Corollary \([8,13]\). Then \( U \) is pullback asymptotically compact w.r.t. \( \hat{B} \) and \( d_{C,\gamma,p.i.p.} \) due to recent Lemma \([8,25]\). Hence, all the assumptions of Lemma \([6,2]\) are satisfied – as required for Theorem \([6,3]\).

Finally for \( q \in [0, \lambda_1 - \tilde{\beta}] \) fixed arbitrarily, we consider the universe \( \mathcal{D} \) consisting of all families \( \hat{D} = \{ D(t) : t \in \mathbb{R} \} \) of nonempty closed bounded sets in \( C_\gamma \) satisfying \( \sup \{ e^{q t} \| D(t) \|_\gamma \ | t < 0 \} < \infty \).

According to Corollary \([8,13,2]\), \( \hat{B} \) is pullback \( \mathcal{D} \)-absorbing w.r.t. \( U \). Hence, Theorem \([6,3]\) guarantees that \( U \) has a pullback \( \mathcal{D} \)-attractor with respect to \( d_{C,\gamma,p.i.p.} \), which is even strictly invariant. \( \square \)
8.12. The pullback ω-Mazur property of $U$ w.r.t. $\hat{B}$, $d_{c,\gamma}$ and $d_{c,\gamma,p.i.p.}$.

Lemma 8.26. As in Corollary 8.19 and Lemma 8.24, suppose assumptions (G1) – (G6), consider the MNDS $U : \mathbb{R}_+ \times C_\gamma \rightarrow C_\gamma$ specified in Proposition 5.3 and define $\hat{B} = \{ B(t) : t \in \mathbb{R} \}$ by means of closed balls with radii as in (36).

Then $U$ has the pullback ω-Mazur property w.r.t. $\hat{B}$, $d_{c,\gamma}$ and $d_{c,\gamma,p.i.p.}$, i.e., according to Definition 6.7: Fix $(t, \tau) \in \mathbb{R}_+^2$ arbitrarily and, suppose for $y \in C_\gamma$, $(s_n)_{n \in \mathbb{N}}$ in $\mathbb{R}$, $(y_n)_{n \in \mathbb{N}}$, $(\phi_n)_{n \in \mathbb{N}}$ in $C_\gamma$

\[
\begin{cases}
  s_n \rightarrow +\infty, & d_{c,\gamma,p.i.p.}(y_n, y) \rightarrow 0 \quad \text{for } n \rightarrow \infty, \\
  y_n \in U(\tau, \tau - s_n, \phi_n), & \phi_n \in \hat{B}(\tau - s_n) \quad \text{for each } n \in \mathbb{N}.
\end{cases}
\]

Then there exist some further sequences $\tilde{s}_k \rightarrow +\infty$, $\tilde{y}_k \in U(\tau, \tau - \tilde{s}_k, \tilde{\varphi}_k)$, $\tilde{\varphi}_k \in \hat{B}(\tau - \tilde{s}_k)$ satisfying

(i) $d_{c,\gamma}(\tilde{y}_k, y) \rightarrow 0$ for $k \rightarrow \infty$,

(ii) $\lim_{n \rightarrow \infty} \inf_{k \geq n} \text{dist}_{c,\gamma,p.i.p.}(U(t, \tau, y_n), U(t, \tau, \tilde{y}_k)) = 0$.

Proof. Fix $(t, \tau) \in \mathbb{R}_+^2$ arbitrarily and, make for $y \in C_\gamma$, $(s_n)_{n \in \mathbb{N}}$ in $\mathbb{R}$, $(y_n)_{n \in \mathbb{N}}$, $(\phi_n)_{n \in \mathbb{N}}$ in $C_\gamma$ the assumptions of convergence specified in the lemma. For each index $k \in \mathbb{N}$, we are going to select an index $n_k \in \mathbb{N}$ and construct some $\tilde{y}_k, \tilde{\varphi}_k \in C_\gamma$ such that

(i') $d_{c,\gamma}(\tilde{y}_k, y) \leq \frac{1}{k}$,

(ii') $\lim_{n \rightarrow \infty} \inf_{k \geq n} \text{dist}_{c,\gamma,p.i.p.}(U(t, \tau, y_n), U(t, \tau, \tilde{y}_k)) \leq \frac{1}{k}$,

(iii') $\tilde{y}_k \in U(\tau, \tau - k, \tilde{\varphi}_k)$,

(iv') $\tilde{\varphi}_k \in \hat{B}(\tau - k)$, i.e., $\| \tilde{\varphi}_k \|_\gamma \leq \rho(\tau - k)$.

Then a final step of relabeling the sequence $(\tilde{y}_k)_{k \in \mathbb{N}}$ (if required) leads to all the claimed features.

Choose any $k \in \mathbb{N}$. $\hat{C}_k = \text{const}(\Omega, \| \Lambda \|_{L^2(\mathbb{R})}, \tau - k, \tau) \geq 1$ and $\hat{C}_k = \text{const}(\Omega, \| \Lambda \|_{L^2(\mathbb{R})}, \tau - k, t) \geq 1$ denote the constant specified in Corollary 8.18 for the time tuples $(\tau, \tau - k), (t, \tau - k) \in \mathbb{R}_+^2$ respectively. As mentioned in the proof of Lemma 3.3 (in section 8.1), there exists some $\hat{T} = \hat{T}(\Theta, \tau, k) \geq 1$ such that $\Omega_{\hat{T}} := \{ x \in \Omega \mid \Theta(x) \leq \hat{T} \}$ satisfies

\[
2c \cdot L^d(\Omega \setminus \Omega_{\hat{T}})^{1/2} < \frac{1}{12k \hat{C}_k \hat{C}_k} \quad \text{and} \quad 2c \cdot \sum_{r \in \mathbb{N} : r \geq \hat{T}} 2^{-r} < \frac{1}{2k} \quad (43)
\]
due to $\Theta \in L^1(\Omega)$ and the converging geometric series. Set $\sigma_k := k + \hat{T} + \theta + 2 > 0$ as abbreviation.

Due to $s_n \rightarrow \infty (n \rightarrow \infty)$, there is an index $N = N(\Theta, \theta, \tau, k) \in \mathbb{N}$ with $s_n \geq \sigma_k$ for all $n \geq N$. The assumption $d_{c,\gamma,p.i.p.}(y_n, y) \rightarrow 0 (n \rightarrow \infty)$ and Definition 14 of $d_{c,\gamma,p.i.p.}$ imply an index $n_k \geq N$ sufficiently large with

\[
e^{\gamma k} \| y_{n_k} - y \|_\gamma + \sup_{-\sigma_k \leq s \leq 0} \| y_{n_k}(s) - y(s) \|_{L^2(\Omega)} < \frac{1}{12k \hat{C}_k \hat{C}_k} \quad \text{and} \quad \sup_{-\sigma_k \leq s \leq 0} \| y_{n_k}(s) - y(s) \|_{L^2(\Omega)} < \frac{1}{12k \hat{C}_k \hat{C}_k} \quad (44)
\]
Now define $\tilde{z}_k$, $\tilde{\varphi}_k$, $\varphi_k \in C^0\left(\left(-\infty, 0\right], L^2(\Omega)\right)$ as

$$\tilde{z}_k(s) := \begin{cases} y_{n_k}(s) & \text{if } -\sigma_k + 1 \leq s \leq 0, \\ (s + \sigma_k) \cdot y_{n_k}(s) - (s + \sigma_k - 1) \cdot y(s) & \text{if } -\sigma_k \leq s < -\sigma_k + 1, \\ y(s) & \text{if } s < -\sigma_k, \end{cases}$$

$$\tilde{\varphi}_k(s) := \tilde{z}_k(s - k) = \begin{cases} y_{n_k}(s - k) & \text{if } -\theta - \tilde{T} - 1 \leq s \leq 0, \\ (s + \theta + \tilde{T} + 2) \cdot y_{n_k}(s - k) - (s + \theta + \tilde{T} + 1) \cdot y(s - k) & \text{if } -\theta - \tilde{T} - 2 \leq s < -\theta - \tilde{T} - 1, \\ y(s - k) & \text{if } s < -\theta - \tilde{T} - 2, \end{cases}$$

$$\varphi_k(s) := y_{n_k}(s - k).$$

The first observation is $\tilde{\varphi}_k \in B(t - k)$. Indeed, the set-valued mapping $\hat{B} = \{B(t) : t \in \mathbb{R}\}$ is positively invariant w.r.t. $U$ due to Corollary 8.13 (1). Hence, the special construction of $U$ implies for every index $n \in \mathbb{N}$ and time parameter $s \in [-s_0, 0]$ $y_{n}(-s) \in U(\tau + s, \tau - s_n, \phi_n) \subset U(\tau + s, \tau - s_n, B(\tau - s_n)) \subset B(\tau + s)$, i.e., $\|y_{n}(-s)\| \leq \rho(\tau + s)$ for each $s \in [-s_0, 0]$ with the radius function $\rho(\cdot)$ defined in [30]. The assumption $d_{c,\gamma\cdot p.i.p.}(y_n, y) \rightarrow 0$ ($n \rightarrow \infty$) has the consequence

$$\|y_{n}(-s) - y(-s)\| \leq e^{-\gamma s} \cdot \|y_n - y\| \leq e^{-\gamma s} \cdot d_{c,\gamma\cdot p.i.p.}(y_n, y) \rightarrow 0$$

($n \rightarrow \infty$) for every $s \in [-\sigma_k, 0]$ and so, $\|y(-s)\| \leq \rho(\tau + s)$. In particular,

$$\max\left\{\|y_{n_k}(-k)\|, \|y(-k)\|\right\} \leq \rho(\tau - k)$$

and, we conclude in a piecewise way that $\|\tilde{\varphi}_k\| \leq \rho(\tau - k)$, i.e., feature (iv') is verified.

Next, aiming at property (iv') by means of estimate [33] in Corollary 8.18 we start with the distance

$$\rho_k := \|\tilde{\varphi}_k - \varphi_k\| + \|\text{ess sup}_{-\theta - \tilde{T} - 2 \leq s \leq 0} \|\tilde{\varphi}_k(s) - [\varphi_k(s)]_c\|\|_L^2(\Omega)$$

and conclude from the choice of $\Omega_{\tilde{T}} \subset \Omega$ with $\tilde{T} = \tilde{T}(\Theta, k) \geq 1$

$$\rho_k = \sup_{s \leq 0} \left\{ e^{\gamma s} \cdot \|\tilde{\varphi}_k(s) - \varphi_k(s)\|_L^2 \right\} + \|\text{ess sup}_{-\theta - \tilde{T} - 2 \leq s \leq 0} \|\tilde{\varphi}_k(s) - [\varphi_k(s)]_c\|\|_L^2 \leq \sup_{s \leq -\theta - \tilde{T} - 1} \left\{ e^{\gamma s} \cdot \|y_{n_k}(s - k) - y_{n_k}(s - k)\|_L^2 \right\} + \|\text{ess sup}_{[-\theta - \tilde{T} - 2, 0]} \|\tilde{\varphi}_k - [\varphi_k]_c\|\|_L^2 \leq e^{\gamma k} \cdot \|y - y_{n_k}\| + \|\text{ess sup}_{-\theta - \tilde{T} - 2 \leq s \leq 0} \|y_{n_k}(s - k) - [y(s - k)]_c\|\|_L^2(\Omega) \leq e^{\gamma k} \cdot \|y - y_{n_k}\| + \|\text{ess sup}_{-\sigma_k \leq s \leq 0} \|y_{n_k}(s) - [y(s)]_c\|\|_L^2(\Omega) \leq \frac{1}{4 k \tilde{C}_k} \cdot \frac{\tilde{C}_k}{\gamma \kappa \leq 1} = \frac{1}{4 k \tilde{C}_k} \cdot \frac{1}{\tilde{C}_k}.$
By construction (in Proposition 5.3), \( y_{nk} \in U(\tau, \tau - k, \varphi_k) \) and so, the arguments of Corollary 8.18 provide some \( \tilde{y}_k \in U(\tau, \tau - k, \tilde{\varphi}_k) \) with

\[
\| y_{nk} - \tilde{y}_k \|_\gamma + \left\| \text{ess sup}_{\theta = -\tilde{T} - k - 2 \leq s \leq 0} \| [y_{nk}(s)]_c - [\tilde{y}_k(s)]_c \|_{L^2(\Omega)} \right. \\
\leq \tilde{C}_k \cdot \left( \| \varphi_k(0) - \tilde{\varphi}_k(0) \|_{W^{1,2}(\Omega)} + \| \varphi_k - \tilde{\varphi}_k \|_\gamma \\
+ \left\| \text{ess sup}_{\theta = -\tilde{T} - 2 \leq s \leq 0} \| [\varphi_k(s)_c - [\tilde{\varphi}_k(s)]_c \|_{L^2(\Omega)} \right. \\
= \tilde{C}_k \cdot \left( \| y_{nk}(-k) - y_{nk}(-k) \|_{W^{1,2}(\Omega)} + \rho_k \right) \leq \frac{1}{4k}
\]

since \( y_{nk} \in U(\tau, \tau-s_{nk}, \phi_{nk}) \), \( s_{nk} > \sigma_k > k \) and the “smoothing effect” of parabolic equations (in the sense of Corollary 8.16) imply \( y_{nk}(-k) = \varphi_k(0) \in W^{1,2}(\Omega) \).

The additional advantage of this construction is \( \tilde{y}_k(s) = y(s) \) for every \( s \leq -\sigma_k \) and \( \theta - \tilde{T} - 2 \). In regard to the aimed estimate (i′), we obtain

\[
d_{c,\gamma}(\tilde{y}_k, y) = \| \tilde{y}_k - y \|_\gamma + \left\| \text{ess sup}_{s \leq 0} \| [\tilde{y}_k(s)]_c - [y(s)]_c \|_{L^2(\Omega)} \right. \\
\leq \| \tilde{y}_k - y_nk \|_\gamma + \left\| \text{ess sup}_{-\sigma_k \leq s \leq 0} \| [\tilde{y}_k(s)]_c - [y_{nk}(s)]_c \|_{L^2(\Omega)} \right. \\
+ \| y_{nk} - y \|_\gamma + \left\| \text{ess sup}_{-\sigma_k \leq s \leq 0} \| [y_{nk}(s)]_c - [y(s)]_c \|_{L^2(\Omega)} \right. \\
\leq \frac{1}{4k} + \left\| \text{ess sup}_{-\sigma_k \leq s \leq 0} \| [\tilde{y}_k(s)]_c - [y_{nk}(s)]_c \|_{L^2(\Omega)} \right. \\
+ \frac{1}{12k} + \frac{1}{12k} \\
\leq \frac{5}{12k} + \frac{1}{4k} \leq \frac{1}{k}.
\]

Finally, property (ii′) remains to be verified, i.e.,

\[
dist_{c,\gamma,p.i.p.}(U(t, \tau, y_{nk}), U(t, \tau, \tilde{y}_k)) \leq \frac{1}{k}.
\]

Indeed, every \( \xi \in U(t, \tau, y_{nk}) \) has the form \( \xi = yึก(\cdot, \tau - k, \varphi_k) \) since \( U \) is a strict MNDS (Proposition 5.3) and \( \varphi_k = y_{nk}(\cdot - k) \) on \( (-\infty, 0] \). Lemma 8.17 leads to a (both strong and mild) solution \( v: (-\infty, t) \to L^2(\Omega) \) of

\[
\begin{cases}
\partial_t v - \Delta v \in G(t, x; v(t, x), \text{ess sup}_{t - \theta - \Theta(x) \leq s \leq t} [v(s, x)]_c, v(t, \cdot)) \quad \text{a.e. in } (\tau, t) \times \Omega \\
v = \tilde{y}_k(\cdot - \tau) \quad \text{a.e. in } (-\infty, \tau) \times \Omega \\
v = 0 \quad \text{on } (\tau, t) \times \partial\Omega
\end{cases}
\]

with the supplementary estimate for every \( s \in [\tau, t] \)

\[
\| \xi(s - t) - v(s) \|_{L^2(\Omega)} + \left\| \text{ess sup}_{\sigma \leq s \leq \tau} \| [\xi(\sigma - t, \cdot)]_c - [v(\sigma, \cdot)]_c \|_{L^2(\Omega)} \right. \\
\]
Due to the original choice of $\bar{y}_k \in U(t, \tau - k, \bar{\varphi}_k)$ satisfying an additional estimate (related to Corollary 8.18), the notion of concatenating solutions reveals for $\bar{\xi} \in C_\gamma$ defined as $\bar{\xi}(s) := \nu(s + t)$ for each $s \in (-\infty, 0]$.

1. $\bar{\xi} \in U(t, \tau, \bar{y}_k) \subset U(t, \tau - k, \bar{\varphi}_k)$

2. $\|\xi - \bar{\xi}\|_\gamma + \left\|\operatorname{ess sup}_{-\theta - \Theta(t) - (t - \gamma + k) \leq \sigma \leq 0} \left[\|\xi(s)\|_c - \|\bar{\xi}(s)\|_c\right]\right\|_{L^2(\Omega)}$

Due to the remaining property (ii‘):

$$\operatorname{dist}_c,\gamma,p.i.p.\left(\xi, U(t, \tau, \bar{y}_k)\right) \leq d_{c,\gamma,p.i.p.}(\xi, \bar{\xi})$$

This leads to the following estimate:

$$\operatorname{dist}_c,\gamma,p.i.p.\left(\xi, U(t, \tau, \bar{y}_k)\right) \leq \bar{\mathcal{C}}_k \cdot \left(\|\varphi_k - \bar{\varphi}_k\|_\gamma + \left\|\operatorname{ess sup}_{-\hat{T} \leq s \leq 0} \left[\|\varphi_k(s)\|_c - \|\bar{\varphi}_k(s)\|_c\right]\right\|_{L^2(\Omega)} + \frac{1}{2T_\gamma}\right) + \frac{1}{2T_\gamma}$$

Hence,

$$\operatorname{dist}_c,\gamma,p.i.p.\left(\xi, U(t, \tau, \bar{y}_k)\right) \leq \bar{\mathcal{C}}_k \cdot \left(\|\varphi_k - \bar{\varphi}_k\|_\gamma + \left\|\operatorname{ess sup}_{-\hat{T} \leq s \leq 0} \left[\|\varphi_k(s)\|_c - \|\bar{\varphi}_k(s)\|_c\right]\right\|_{L^2(\Omega)} + \frac{1}{2T_\gamma}\right) + \frac{1}{2T_\gamma}$$

i.e., $\operatorname{dist}_c,\gamma,p.i.p.\left(U(t, \tau, y_{n_k}), U(t, \tau, \bar{y}_k)\right) \leq \frac{1}{4k} + \frac{1}{12k} + \frac{1}{2k} < \frac{1}{k}$ for each $k \in \mathbb{N}$. \qed

8.13. **The proof of main Theorem 4.1.** Briefly speaking, it is an immediate consequence of Theorem 5.9, which specifies sufficient conditions for the existence of pullback attractors of multi-valued non-autonomous dynamical systems in terms of two metrics.
Indeed, we again supply $C_\gamma$ with the metric $d_{c,\gamma,\text{p.i.p.}}$ defined in \cite{14}. Due to Lemma \ref{5.2} the metric space $(C_\gamma, d_{c,\gamma,\text{p.i.p.}})$ is complete. Proposition \ref{5.3} specifies the strict MNDS $U : \mathbb{R}_+^2 \times C_\gamma \sim C_\gamma$, $(t_1, t_0, \phi) \mapsto u(-t_1)$ by means of the parabolic differential inclusion (with $0 < \theta < \infty$ and $\Theta \in L^1(\Omega, [0, \infty[)$ given)
\[
\begin{aligned}
\partial_t u - \Delta u &\in G\left(t, x; u(t, x), \sup_{t-\theta \leq s \leq t} |u(s, x)|_C, u(t, \cdot)\right) \\
&\quad \text{a.e. in } (t_0, t_1) \times \Omega \\
u &= \phi(\cdot - t_0) \\
u &= 0 \quad \text{on } (t_0, t_1) \times \partial \Omega.
\end{aligned}
\]
For each $(t, \tau) \in \mathbb{R}_+^2$ the set-valued mapping $U(t, \tau, \cdot) : (C_\gamma, d_{c,\gamma}) \sim (C_\gamma, d_{c,\gamma,\text{p.i.p.}})$ has a closed graph according to Lemma \ref{8.8}.

Now consider the set-valued mapping $\mathcal{B} = \{B(t) : t \in \mathbb{R}\}$ of “balls” $B(t) := \{} \phi \in C_\gamma \mid \|\phi\|_C \leq \rho(t)\}$ specified in Corollary \ref{8.13} Then $U$ is pullback asymptotically compact w.r.t. $\mathcal{B}$ and $d_{c,\gamma,\text{p.i.p.}}$ due to Lemma \ref{8.25} The recently proved Lemma \ref{8.26} states that $U$ has the pullback $\omega$-Mazur property w.r.t. $\mathcal{B}$, $d_\gamma$, $d_{X,\text{p.i.p.}}$ (in the sense of Definition \ref{6.7}). Thus, all the assumptions of Lemma \ref{6.8} are fulfilled – as required for Theorem \ref{6.9}.

Finally for $q \in (0, \lambda_1 - \lambda_1)$ fixed arbitrarily, we (again) consider the universe $\mathcal{D}$ of all families $\tilde{D} = \{D(t) : t \in \mathbb{R}\}$ of nonempty closed bounded sets in $C_\gamma$ satisfying $\sup \{e^{q \xi} \|D(t)\|_C \mid t < 0\} < \infty$. Due to Corollary \ref{8.13}(2.), $\mathcal{B}$ is pullback $\mathcal{D}$-absorbing with respect to $U$.

Hence, Theorem \ref{6.9} ensures that $U$ has a pullback $\mathcal{D}$-attractor with respect to $d_{c,\gamma,\text{p.i.p.}}$, which is even strictly invariant – as claimed. 

\begin{thebibliography}{99}
\setlength{\itemsep}{0pt}
\bibitem[1]{1} J.-P. Aubin and A. Cellina, \textit{Differential Inclusions}, volume 264 of Grundlehren der mathematischen Wissenschaften, Springer-Verlag, Berlin, 1984. Set-valued maps and viability theory.
\bibitem[2]{2} J.-P. Aubin and H. Frankowska, \textit{Set-valued Analysis}, volume 2 of Systems & Control: Foundations & Applications, Birkhäuser Boston Inc., Boston, MA, 1990.
\bibitem[3]{3} J. M. Ball, \textit{Strongly continuous semigroups, weak solutions, and the variation of constants formula}, \textit{Proc. Amer. Math. Soc.}, \textbf{63} (1977), 370–373.
\bibitem[4]{4} V. Barbu, \textit{Nonlinear Semigroups and Differential Equations in Banach Spaces}, Editura Academiei Republicii Socialiste România, Bucharest; Noordhoff International Publishing, Leiden, 1976. Translated from the Romanian.
\bibitem[5]{5} P. Bénilan, \textit{Solutions intégrales d'équations d'évolution dans un espace de Banach}, \textit{C. R. Acad. Sci. Paris Sér. A-B}, \textbf{274} (1972), A47–A50.
\bibitem[6]{6} H. Brezis, \textit{Functional Analysis, Sobolev Spaces and Partial Differential Equations}, Universitext. Springer, New York, 2011.
\bibitem[7]{7} T. A. Burton, \textit{Volterra Integral and Differential Equations}, volume 202 of Mathematics in Science and Engineering, Elsevier B. V., Amsterdam, second edition, 2005.
\bibitem[8]{8} T. Caraballo and P. E. Kloeden, \textit{Non-autonomous attractor for integro-differential evolution equations}, \textit{Discrete Contin. Dyn. Syst. Ser. S}, \textbf{2} (2009), 17–36.
\bibitem[9]{9} T. Caraballo, J. A. Langa, V. S. Melnik, and J. Valero, \textit{Pullback attractors of nonautonomous and stochastic multivalued dynamical systems}, \textit{Set-Valued Anal.}, \textbf{11} (2003), 153–201.
\bibitem[10]{10} T. Caraballo, G. Łukaszewicz, and J. Real, \textit{Pullback attractors for asymptotically compact non-autonomous dynamical systems}, \textit{Nonlinear Anal.}, \textbf{64} (2006), 484–498.
\bibitem[11]{11} A. N. Carvalho, J. A. Langa, and J. C. Robinson, \textit{Attractors for Infinite-Dimensional Non-Autonomous Dynamical Systems}, volume 182 of Applied Mathematical Sciences, Springer, New York, 2013.
\bibitem[12]{12} C. Castaing, L. A. Faik, and A. Salvadori, \textit{Evolution equations governed by m-accretive and subdifferential operators with delay}, \textit{Int. J. Appl. Math.}, \textbf{2} (2000), 1005–1026.
\end{thebibliography}
PULLBACK ATTRACTORS OF REACTION-DIFFUSION INCLUSIONS

[12] C. Castaing and M. Valadier, *Convex Analysis and Measurable Multifunctions*, volume 580 of Lecture Notes in Mathematics, Springer-Verlag, Berlin-New York, 1977.

[13] V. V. Chepyzhov, S. Gatti, M. Grasselli, A. Miranville, and V. Pata, Trajectory and global attractors for evolution equations with memory, *Appl. Math. Lett.*, 19 (2006), 87–96.

[14] V. V. Chepyzhov and A. Miranville, Trajectory and global attractors of dissipative hyperbolic equations with memory, *Commun. Pure Appl. Anal.*, 4 (2005), 115–142.

[15] V. V. Chepyzhov and A. Miranville, On trajectory and global attractors for semilinear heat equations with fading memory, *Indiana Univ. Math. J.*, 55 (2006), 119–167.

[16] V. V. Chepyzhov and M. I. Vishik, Evolution equations and their trajectory attractors, *J. Math. Pures Appl. (9)*, 76 (1997), 913–964.

[17] V. V. Chepyzhov and M. I. Vishik, *Attractors for Equations of Mathematical Physics*, volume 49 of American Mathematical Society Colloquium Publications, American Mathematical Society, Providence, RI, 2002.

[18] C. M. Dafermos, Asymptotic stability in viscoelasticity, *Arch. Rational Mech. Anal.*, 37 (1970), 297–308.

[19] E. B. Davies, *Heat Kernels and Spectral Theory*, volume 92 of Cambridge Tracts in Mathematics, Cambridge University Press, Cambridge, 1989.

[20] M. M. Deza and E. Deza, *Encyclopedia of Distances*, Springer, Heidelberg, third edition, 2014.

[21] O. Diekmann, S. A. van Gils, S. M. Verduyn Lunel, and H.-O. Walther, *Delay Equations*, volume 110 of Applied Mathematical Sciences, Springer-Verlag, New York, 1995. Functional, complex, and nonlinear analysis.

[22] J. Diestel and J. J. Uhl, Jr, *Vector Measures*, volume 15 of Mathematical Surveys and Monographs, American Mathematical Society, Providence, R.I., 1977.

[23] K.-J. Engel and R. Nagel, *One-parameter Semigroups for Linear Evolution Equations*, volume 194 of Graduate Texts in Mathematics, Springer-Verlag, New York, 2000.

[24] J. R. Giles, *Introduction to the Analysis of Metric Spaces*, volume 3 of Australian Mathematical Society Lecture Series, Cambridge University Press, Cambridge, 1987.

[25] N. V. Gorban, O. V. Kapustyan and P. O. Kasyanov, Uniform trajectory attractor for non-autonomous reaction-diffusion equations with Carathéodory’s nonlinearity, *Nonlinear Anal.*, 98 (2014), 13–26.

[26] B. Goullet and O. V. Kapustyan and P. O. Kasyanov, On regularity of all weak solutions and their attractors for reaction-diffusion inclusion in unbounded domain, In *Continuous and distributed systems*, volume 211 of Solid Mech. Appl., pages 205–220. Springer, Cham, 2013.

[27] J. W. Green and F. A. Valentine, On the Arzelà-Ascoli theorem, *Math. Mag.*, 34 (1960/61), 199–202.

[28] P. E. Kloeden and J. A. Langa, Flattening, squeezing and the existence of random attractors, *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.*, 463 (2007), 163–181.

[29] P. E. Kloeden and T. Lorenz, Construction of nonautonomous forward attractors, *Proc. Amer. Math. Soc.*, 144 (2016), 259–268.

[30] P. E. Kloeden, T. Lorenz and M. Yang, Reaction-diffusion equations with a switched-off reaction zone, *Commun. Pure Appl. Anal.*, 13 (2014), 1907–1933.

[31] P. E. Kloeden and M. Rasmussen, *Nonautonomous Dynamical Systems*, volume 176 of Mathematical Surveys and Monographs, American Mathematical Society, Providence, RI, 2011.
O. A. Ladyženskaja, V. A. Solonnikov and N. N. Ural’ceva, *Linear and Quasilinear Equations of Parabolic Type*, Translated from the Russian by S. Smith. Translations of Mathematical Monographs, Vol. 23. American Mathematical Society, Providence, R.I., 1968.

V. Lakshmikantham, L. Z. Wen and B. G. Zhang, *Theory of Differential Equations with Unbounded Delay* volume 298 of Mathematics and its Applications, Kluwer Academic Publishers Group, Dordrecht, 1994.

Y. Li and C. Zhong, *Pullback attractors for the norm-to-weak continuous process and application to the nonautonomous reaction-diffusion equations* *Appl. Math. Comput.*, 190 (2007), 1020–1029.

G. M. Lieberman, *Second Order Parabolic Differential Equations* World Scientific Publishing Co., Inc., River Edge, NJ, 1996.

T. Lorenz, *Mutational Analysis*, volume 1996 of Lecture Notes in Mathematics. Springer-Verlag, Berlin, 2010. A joint framework for Cauchy problems in and beyond vector spaces.

Q. Ma, S. Wang and C. Zhong, *Necessary and sufficient conditions for the existence of global attractors for semigroups and applications* *Indag. Math. N.S.*, 11 (2000), 53–62.

V. S. Melnik and J. Valero, *On attractors of multivalued semi-flows and differential inclusions* *Set-Valued Anal.*, 6 (1998), 83–111.

C. B. Morrey, Jr, *Multiple Integrals in the Calculus of Variations*, Die Grundlehren der mathematischen Wissenschaften, Band 130. Springer-Verlag New York, Inc., New York, 1966.

A. Ornelas, *Parametrization of Carathéodory multifunctions*, *Rend. Sem. Mat. Univ. Padova*, 83 (1990), 33–44.

A. Pazy, *Semigroups of Linear Operators and Applications to Partial Differential Equations* volume 44 of Applied Mathematical Sciences, Springer-Verlag, New York, 1983.

P. Quittner and P. Souplet, *Superlinear Parabolic Problems*, Birkhäuser Advanced Texts: Basler Lehrbücher. [Birkhäuser Advanced Texts: Basel Textbooks]. Birkhäuser Verlag, Basel, 2007. Blow-up, global existence and steady states.

B. Schmalfuss, *Attractors for the non-autonomous dynamical systems* In *International Conference on Differential Equations, Vol. 1, 2 (Berlin, 1999)*, pages 684–689. World Sci. Publ., River Edge, NJ, 2000.

H. Song, *Pullback attractors of non-autonomous reaction-diffusion equations in H10* *J. Differential Equations*, 249 (2010), 2357–2376.

H. Song and H. Wu, *Pullback attractors of nonautonomous reaction-diffusion equations* *J. Math. Anal. Appl.*, 325 (2007), 1200–1215.

C.-Y. Sun and C.-K. Zhong, *Attractors for the semilinear reaction-diffusion equation with distribution derivatives in unbounded domains* *Nonlinear Anal.*, 63 (2005), 49–65.

Y. Wang, C. Zhong, and S. Zhou, *Pullback attractors of nonautonomous dynamical systems* *Discrete Contin. Dyn. Syst.*, 16 (2006), 587–614.

Y. Wang and S. Zhou, *Kernel sections and uniform attractors of multi-valued semiprocesses* *J. Differential Equations*, 232 (2007), 573–622.

Y. Wang and S. Zhou, *Kernel sections on multi-valued processes with application to the nonlinear reaction-diffusion equations in unbounded domains* *Quart. Appl. Math.*, 67 (2009), 343–378.

K. Yosida, *Functional Analysis*, volume 123 of Grundlehren der mathematischen Wissenschaften, Springer-Verlag, Berlin-New York, sixth edition, 1980.

M. Z. Zgurovsky and P. O. Kasyanov, *Evolution inclusions in nonsmooth systems with applications for earth data processing: Uniform trajectory attractors for nonautonomous evolution inclusions solutions with pointwise pseudomonotone mappings* In *Advances in global optimization*, volume 95 of Springer Proc. Math. Stat., pages 283–294. Springer, Cham, 2015.

C. Zhao and S. Zhou, *Pullback trajectory attractors for evolution equations and application to 3D incompressible non-Newtonian fluid* *Nonlinearity*, 21 (2008), 1691–1717.

C.-K. Zhong, M.-H. Yang, and C.-Y. Sun, *The existence of global attractors for the norm-to-weak continuous semigroup and application to the nonlinear reaction-diffusion equations* *J. Differential Equations*, 223 (2006), 367–399.

W. P. Ziemer, *Weakly Differentiable Functions*, volume 120 of Graduate Texts in Mathematics, Springer-Verlag, New York, 1989.

Received March 2015; revised April 2016.

E-mail address: kloeden@math.uni-frankfurt.de
E-mail address: thomas.lorenz@hs-rm.de