MNRAS **000**, 000–000 (0000) Preprint 25 April 2018 Compiled using MNRAS LATEX style file v3.0

Relativistic charged solitons created due to Nonlinear Landau Damping: A candidate for explaining coherent radio emission in pulsars

Taras Lakoba[1], Dipanjan Mitra[2,3], George Melikidze[3,4]

[1] *Department of Mathematics and Statistics, University of Vermont, Burlington VT 05401, USA*

[2] *National Centre for Radio Astrophysics, Tata Institute of Fundamental Research, Pune 411007, India*

[3] *Janusz Gil Institute of Astronomy, University of Zielona Góra, ul. Szafrana 2, 65-516 Zielona Góra, Poland*

[4] *Abastumani Astrophysical Observatory, Ilia State University, 3-5 Cholokashvili Ave., Tbilisi, 0160, Georgia*

25 April 2018

ABSTRACT

A potential resolution for the generation of coherent radio emission in pulsar plasma is the existence of relativistic charged solitons, which are solutions of nonlinear Schrödinger equation. In an earlier study, Melikidze et al. (2000) investigated the nature of these charged solitons; however, their analysis ignored the effect of nonlinear Landau damping, which is inherent in the derivation of the nonlinear Schödinger equation in the pulsar pair plasma. In this paper we include the effect of Landau damping and obtain solutions of the nonlinear Schödinger equation by applying a suitable numerical scheme. We find that for reasonable parameters for the nonlinearity and Landau damping, there exists soliton-like solutions that remain stable over a time during which they are capable of exciting the coherent curvature radiation in pulsars.

Key words: pulsars:general, MHD — plasmas — pulsars: general, radiation mechanism: nonthermal

1 INTRODUCTION

The region around a strongly magnetized ($\mathbf{B} \sim 10^{12}$ G) and fast-spinning neutron star generates enormous electric fields $\mathbf{E}$ and cannot be maintained as vacuum (Goldreich & Julian 1969). The problem of solving for the charge distribution around the magnetosphere is nontrivial and is a

matter of intense research (see e.g. Spitkovsky 2011; Pétri 2016). Most theories follow the idea that the region around the neutron star is a charge-separated magnetosphere which is force free, meaning that the electromagnetic energy is significantly larger than all other inertial, pressure and dissipative forces. The magnetosphere is initially charge-starved, and a supply of charged particles can come from the neutron star or due to pair creation in strong magnetic fields. To maintain co-rotation in the magnetosphere, the condition $\mathbf{E}.\mathbf{B} = 0$ should be satisfied, and this corresponds to a charge number density equal to the Goldreich–Julian density $n_{GJ} = \Omega.\mathbf{B}/2\pi ce$, where $\Omega = 2\pi/P$ and P is there rotational period of the pulsar, c is the velocity of light and e is the electron charge. In the presence of sufficient supply of charges the magnetosphere can have two distinct regions: the closed dipolar magnetic field line region where charges co-rotate with the neutron star and the open dipolar magnetic field line region where a outflowing relativistic pulsar wind can exist.

Radio emission from pulsar is thought to arise from the development of plasma instabilities in the electron–positron plasma streaming relativistically along open dipolar magnetic field lines in the pulsar magnetosphere. However, identifying the physical process that can explain the radio radiation properties in pulsars is a challenging problem in astrophysics. The key issues here are: (i) to explain the problem of coherency, which manifests itself as observed pulsar radio emission with unrealistically high brightness temperatures $\sim 10^{28} \ldots 10^{30}$K; and (ii) to explain the range of pulsar phenomena, such as micropulses, subpulse drift, nulling/moding, pulsar profile stability, polarization properties, etc.. Generally, the coherent pulsar radio emission can be generated by means of either a maser or a coherent curvature mechanism (e.g., Ginzburg & Zhelezniakov 1975; Ruderman & Sutherland 1975; Melikidze & Pataraya 1980; Melikidze & Pataraya 1984; Kazbegi et al. 1991; Melikidze et al. 2000) emitted in strongly magnetized electron—positron plasma well inside the light cylinder. However, as we will discuss in section 2 (see also Mitra (2017) for a recent review), a large body of observations appear to suggest that the pulsar radio emission is excited via a mechanism of coherent curvature radiation. This radiation emerges from regions of about 500 km above the neutron star surface. The high brightness temperature of this coherent radiation can be explained only if it is excited by charge bunches containing a very large number of charged particles rather than by a single charge. The physics of how these charge bunches are formed and how they emit coherent radio emission is still poorly understood. In this work we will focus on the problem of formation of charge bunches and their stability, and will also address the problem of coherency in pulsar radio emission.

There are several suggestion in the literature as to where and how the electron–positron pair production can happen in the pulsar magnetosphere, and amongst them the polar-cap models are

the ones where the pair production occurs just above the polar cap. These models can best explain the coherent radio emission properties of pulsars. Ruderman & Sutherland 1975 (hereafter RS75) were amongst the first to propose such a model, which attempted to explain the overall aspect of the pulsar emission, i.e., both coherency and radio pulsar observational phenomenology. In their model, there exists an inner acceleration region close to the polar cap, where a relativistic non-stationary flow of the electron–positron pair plasma can be established. To address the problem of coherent radio emission, RS75 suggested that charged bunches could be formed due to development of a two-stream instability that results from the fast-moving and slow-moving particles of the non-stationary plasma. This instability leads to the formation of linear electrostatic Langmuir waves with its frequency being the plasma frequency. As the wave propagates along the magnetic field, each type of particle is subject to the sinusoidal electric field, where for half of its period the field bunches charges of one sign, while for the next half-period it bunches charges of the opposite sign. RS75 proposed that these charge bunches can excite the coherent radio emission.

However, the explanation of coherent emission as occurring from such charge bunches has the following fundamental difficulty, as was pointed out by Lominadze et al. 1986 and Melikidze et al. 2000. On one hand, the spatial dimension Λ_b of an emitting bunch (along the magnetic field lines) should be smaller than the period of the coherently emitted wave λ_c:

$$\lambda_c > \Lambda_b. \tag{1a}$$

Indeed, if $\lambda_c < \Lambda_b$, then different regions of the bunch would emit independently and hence incoherently. As described above, the bunching is caused by linear Langmuir waves (having wavelength λ_l), and the size of a bunch is about half of the wave's period; i.e., $\Lambda_b \approx \lambda_l/2$. Since Langmuir waves have an approximately vacuum dispersion relation, $\omega = 2\pi c/\lambda$, the condition (1a) that the emission be coherent amounts to

$$\omega_c < 2\omega_l, \tag{1b}$$

where ω_c and ω_l the characteristic frequency of the emitted waves and the Langmuir waves, respectively. On the other hand, the temporal period of the emitted wave, i.e. $\mathcal{T}_c = 2\pi/\omega_c$, cannot exceed the time window over which the emitting bunch exists; this time window is half of the period of the Langmuir wave, i.e. $\mathcal{T}_b = \pi/\omega_l$. Indeed, if the condition

$$\mathcal{T}_c < \mathcal{T}_b \tag{2a}$$

does not hold, the charge bunch would disperse away before it has the chance to emit a radio wave. Equivalently to (2a), one must have

$$\omega_c > 2\omega_l, \tag{2b}$$

Clearly, the above two conditions: (1b) (coherency of the emission) and (2b) (non-dispersal of the charge bunch) are in contradiction with each other.

In the last few decades, significant refinement of the basic physical ideas that was postulated by RS75 has been achieved both theoretically and observationally (e.g., Melikidze et al. 2000 (hereafter MGP00); Gil et al. 2004; Mitra et al. 2009; Melikidze et al. 2014). To circumvent the fundamental difficulty described in the previous paragraph, MGP00 accounted for nonlinear effects due to sufficiently strong two-stream instability in the relativistic plasma. Their theory led to the nonlinear Schrödinger equation (NLSE) with a nonlinear Landau damping term, which describes propagation of the *slow envelope* of Langmuir waves. MGP00 neglected the Landau damping term, assuming it to be small, and showed that for reasonable pulsar parameters, the solution of the NLSE leads to formation of a nonlinear solitary wave, i.e., a soliton, and that these solitons carry an effective charge. Unlike the “half-period" charge bunches in the linear RS75 theory, the charged solitons can exist for times much longer than π/ω_l. Thus, since $\mathcal{T}_b$ is no longer related to π/ω_l, condition (2b) can no longer be deduced from condition (2a). (Let us note, in passing, that for solitons, condition (1b) also does not follow from condition (1a), because the soliton’s length is much greater than the period of the carrier Langmuir wave.) Hence, the bunch non-dispersal condition (1b) no longer contradicts the coherency condition (1a), and therefore charged solitons, at least in principle, can excite coherent radio emission in the plasma.

Yet, an explanation of the coherent emission relying on solitons of the “pure" nonlinear Schrödinger equation without a Landau damping term has a shortcoming of its own. A stably propagating soliton (or a few solitons) is known to emerge only from a certain class of initial conditions — a localized one. However, there is no reason to assume that such an initial state actually occurs in a magnetospheric plasma; rather, the initial condition there is likely to be a nonlocalized Langmuir wave with a randomly (but slowly) modulated envelope. A solution developing from such an initial condition is known to be a disordered ensemble of solitons and linear waves; in this disordered state, solitons continuously appear and disappear as a result of their interaction with one another and with linear waves. Consequently, such “flickering" solitons do not exist for times long enough that would let conditions (2a) and (1a) hold simultaneously. Thus, a mechanism that would preserve a soliton’s individuality for a sufficiently long time, is required for the MGP00 theory to become a strong contender in explaining the pulsar coherent radio emission.

In this paper we demonstrate that taking into account the effect of Landau damping in the MGP00 theory provides such a soliton-stabilizing mechanism. The main part of this paper is organized as follows. In section 2 we briefly describe the observational evidence from radio pulsars

that motivates invoking the charged soliton model. In section 3 we briefly outline the generation mechanism and features of the radio emitting plasma based on the polar-cap RS75 class of models. In section 4 we introduce the concept of the NLSE in pulsar plasma, and in section 5 we discuss the range of parameters which are reasonable to expect in charged bunches of plasma near a pulsar. In section 6 we present the main results: a numerical observation of an intense long-living electrostatic pulse with an internal structure, which is formed in the NLSE model in the presence of Landau damping. In section 7 we summarize the results. Appendix A contains a description of the numerical method, and Appendix B lists definitions of notations used in this work.

2 OBSERVATIONAL EVIDENCE OF COHERENT CURVATURE RADIATION OF PULSARS

Radio pulsar phenomenological studies performed over the years provide a sound basis for understanding some general properties of the pulsar radio emission (see, e.g., Mitra (2017)). Pulsars emit periodic signals ranging from about 1 msec to 8.5 sec, and the pulsed emission is restricted to an emission window which is typically 10% of the pulse period. In this study we will focus on properties of so-called normal-pulsars, whose periods, P, are longer than $\sim$50 msec and whose surface dipolar magnetic field is about $\sim 10^{12}$ G. In normal-pulsars the average pulse profile, which is obtained by averaging a large number of single pulses, is seen to be highly structured and can consists of one to several Gaussian-like components.

Pulsars are also highly linearly polarized, and the polarization position angle (PPA) across the pulsar profile shows a characteristic S-shaped swing. This has been interpreted by the rotating vector model (Radhakrishnan & Cooke 1969) as a signature of emission arising due to curvature radiation from charged bunches moving along the open dipolar magnetic field lines. The steepest gradient (SG) point of the PPA traverse corresponds to the fiducial magnetic plane which contains the rotation and magnetic axes.

Pulsar profile along with linear polarization information is used in a statistical sense to infer that the pulsar radio emission beam is composed of a central core emission surrounded by nested conal emission. The components in the single pulses are more dynamic in their location inside the pulse window, which leads to such phenomena as: (i) subpulse drifting, where in subsequent single pulses the emission components are seen to systematically move across the pulse window; (ii) the small-scale quasiperiodic temporal structures seen in components of single pulses called "micro-structures"; and (iii) nulling and moding, where the average or radio emission either

switches off completely or changes its pattern for a certain duration, and then returns back to its original state. All these phenomena can be considered as non-stationary effects in the pulsar magnetospheric plasma.

In the following three subsections we will briefly summarize the basic outcome from pulsar radio observations and point out the constraints they provide in formulation of the theory of coherent radio emission from pulsar.

2.1 Emission height

There are three different techniques that can be used to determine the location where the radio emission detaches from the pulsar magnetosphere. Two of these techniques, namely the geometrical method and the aberration and retardation (A/R) method, rely on the fact that pulsar emission arises in the region of open dipolar diverging magnetic field lines; merits, drawbacks, and usage of these height estimation methods can be found in Mitra & Li (2004) and Dyks (2008). Between these two methods, the A/R method, proposed by Blaskiewicz et al. (1991), is known to give more robust estimates for radio emission locations in normal pulsars. [1] Moreover, the A/R method revealed that emission heights can be estimated independently of pulsar's geometry (see Dyks et al. 2004). The A/R effect is seen as a shift between the center of the total intensity profile and the fiducial plane containing the magnetic and spin axes, which is often identified as the steepest gradient point of the PPA traverse or the peak of the core emission. The A/R methods suggest that the core and conal emission, i.e. the overall emission across the pulsar beam, arises from approximately the same height (Mitra et al. 2016). A few notable studies dedicated to finding emission heights using the A/R method are: Blaskiewicz et al. (1991), von Hoensbroech & Xilouris (1997), Mitra & Li (2004), Mitra & Rankin (2011), Weltevrede & Johnston (2008). These studies suggest that the radio emission arises from about $\sim$500 km above the neutron star's surface (see also Fig. 3 of Mitra 2017) The third method for finding emission heights is based on using pulsar scintillation. In this method, one uses the fact that the emission from the compact emission region of the pulsar passes through the interstellar medium which can act as a varying lens, thus modulating the pulsar signal. The nature of this modulation depends on the spatial transverse extent of the source, which can be recovered by performing extremely high spatial resolution interferometry. The method has

[1] The geometrical method involves estimation of emission heights by solving for the the geometry of the pulsar beam, which in turn involves fitting the rotating vector model to the PPA traverse to estimate the angle between the rotation axis and magnetic axis and the angle between the magnetic axis and the observer's line of sight. These estimated parameters turn out to be highly correlated (see e.g. Everett & Weisberg (2001)), and hence robust estimates of actual height using this method are not possible.

been applied successfully on a few pulsars, and accurate results are only available for the Vela pulsar which imply that the spatial transverse extent of the emission source is about 4 km and the corresponding radio emission altitude is estimated to be about 340 km, in agreement with the other methods (Johnson et al. 2012).

The pulsar radio emission height $R_{em} \sim 500$ km is a very significant input to the pulsar radio emission mechanism problem. The only plasma instability that can develop at these heights (where the magnetic field is very strong and the plasma is constrained to move along the magnetic field lines) is the two–stream instability. Hence, resonance-type instabilities like the cyclotron maser instability (which can develop only near the light cylinder, where the magnetic field is weak), can be ruled out.

2.2 Evidence for curvature radiation

The estimated emission heights R_{em} is the location where the emission detaches from the pulsar magnetosphere. It is quite possible that the pulsar emission is generated in the emitting plasma at a certain height $R_g < R_{em}$, and then emerges out of the plasma at R_{em}. There is, however, no direct way to probe this effect, and one has to resort to wave propagation properties in electron–positron plasma at such strong magnetic fields. Once the radiation is generated in the plasma, it naturally splits as ordinary, or O-mode (polarized in the plane of the wave vector $\mathbf{k}$ and the magnetic field $\mathbf{B}$ plane) and the extraordinary, X-mode (polarized perpendicularly to the $\mathbf{k}$ and $\mathbf{B}$ plane). The O-mode can strongly interact with the plasma and can be damped or ducted away, while the X-mode can escape the plasma at $R_g \sim R_{em}$ as if it were in vacuum (see Mitra et al. 2009 and Melikidze et al. 2014 for details).

It turns out that there is multiple observational evidence that allows determination of the orientation of the emerging polarization direction with respect to the dipolar magnetic field planes. The most direct evidence comes from the x-ray image of the Vela pulsar wind nebula and fiducial or the SG point of the absolute PPA, which can be used to establish that the electric vector emanating out of the pulsar is orthogonal to the magnetic field planes, and hence represents the extraordinary (X) mode. Lai et al. (2001) also showed that the proper motion direction (PM) of the pulsar is aligned with the rotation axis. Johnston et al. (2005) and Rankin (2007) produced a distribution of the quantity $|\mathrm{PM} - \text{absolute PPA}|$ for a few pulsars and found a bimodal distribution around zero and 90°. Assuming that the pulsar's PMs are parallel to the rotation axis, the bimodality could be explained as occurring due to the emerging radiation being either parallel or perpendicular to

the magnetic field planes, since pulsars are known to have orthogonal polarization modes. Alternatively, PMs of pulsars can also be parallel or perpendicular to the rotation axis. While both the above explanations are possible, it is clear that the electric vectors of the waves which detach from the pulsar magnetosphere to reach the observer follow the magnetic field planes.

These observations can hence be interpreted as suggesting that the observed emission is associated with curvature radiation mechanism, since this is the only known emission mechanism that can distinguish the magnetic field planes. Further evidence of curvature radiation can also be obtained from single pulse polarization, where Mitra et al. (2009) demonstrated that the instantaneous polarization of components of single pulses closely follow the average PPA.

2.3 Evidence for non-dipolar surface magnetic field

At a distance of R_{em} where the pulsar radio emission originates, the magnetic field is significantly dipolar. However, the magnetic field at the surface of the neutron star needs to be significantly non-dipolar, so that a sufficient amount of the electron–positron pair plasma can be generated to explain the observed pulsar radiation. Pulsars are known to slow down at a certain rate $\dot{P}$, and this slow-down can be used to estimate only the surface dipolar magnetic field component to be $B_d \sim 6 \times 10^{19}\sqrt{P\dot{P}}$ G; here P is the pulsar rotation period (in seconds) and $\dot{P}$ is non-dimensional. There is, however, observational evidence that suggests the presence of a surface non-dipolar magnetic field. The strongest piece of such evidence, from which the existence of non-dipolar magnetic field can be inferred, comes from the discovery of a long-period ($P = 8.5$ sec) pulsar PSR J2144$-$3933 (Young et al. 1999). Gil & Mitra (2001) argued that significant acceleration of pair plasma in this pulsar, which is essential for producing the radio emission, can only happen if the radius of curvature of the surface magnetic field is $\rho \sim 10^5$ cm, which is about an order of magnitude smaller than ρ values in normal pulsars with $P \sim 1$. The smaller value of ρ_c in PSR J2144$-$3933 implies that the magnitude of the surface non-dipolar magnetic field there is about 10^{14} G, which is about 100 times higher than the dipolar magnetic field. Furthermore, in some radio pulsars, soft x-ray blackbody radiation is seen from hot polar caps, and the estimated area of the polar cap is often found to be smaller than the dipolar area, suggesting the presence of a strong non-dipolar field on the neutron star surface (see, e.g., Table 1 of Geppert (2017) and references therein).

Thus, in summary, the basic input to the pulsar emission models from observations is that coherent radio emission is excited in a non-stationary plasma flowing away from a pulsar and

detaches from it at a height of about a few hundred km above the pulsar's surface, in a region of open dipolar magnetic field lines. The magnetic field on the neutron star surface, however, is significantly non-dipolar.

3 PLASMA CONDITION IN THE MAGNETOSPHERE AND MECHANISM FOR RADIO EMISSION IN PULSARS

As we have described in the Introduction, the RS75 class of the polar cap models provide a framework whereby the observed coherent curvature radiation is attributed to the emission of radio waves by charge bunches in the plasma. In this section we will summarize subsequent stages of creation of this radio emission. We will refer to a a pulsar, i.e., a neutron star, having the following parameters: radius R_s, pulsar period P (measured in seconds), pulsar slow-down rate $\dot{P}$, surface magnetic field B_s, and dipolar magnetic field B_d (see the beginning of section 2.3). For future use we also introduce the ratio $b = B_s/B_d$, a non-dimensional parameter $\dot{P}_{-15} = \dot{P}/(10^{-15})$, which is assumed to be of order one, and the vector of angular velocity of the rotating star, whose magnitude is $\Omega = 2\pi/P$.

3.1 Gap formation

RS75 suggested that if the condition $\mathbf{\Omega}.\mathbf{B_s} < 0$ holds above the pulsar polar cap, then the polar cap is positively charged. They envisaged a situation where initially there is only a limited supply of stray positive charges above the polar cap, which is relativistically flowing away from the pulsar along open magnetic field lines as a pulsar wind. Consequently, if the binding energy of the ions in the neutron star surface are sufficiently strong, then the region above the polar cap will be deficient in positive charges, and a vacuum gap can be created, where an enormously high electric field exists. Photons of energy $> 2m_ec^2$, where m_e is the mass of electron, are split inside the vacuum gap into electron–positron pairs, and the electric field in the gap separates these two types of charges. They are then further accelerated along the curved magnetic field lines (hence the term 'curvature radiation') and can annihilate again, generating a secondary pair of high-energy photons, which, in their turn, after traveling some mean free path can produce another electron–positron pair. In terms of pulsar parameters, the potential drop ΔV across the gap and the gap's height h can be expressed as

$$\Delta V \sim 2 \times 10^{12} b^{-1/7} P^{-3/14} \dot{P}_{-15}^{-1/14} \rho_6^{4/7} \ \mathrm{V}, \tag{3}$$

$$h \sim 5 \times 10^3 b^{-4/7} P^{1/7} \dot{P}_{-15}^{-2/7} \rho_6^{2/7} \ \mathrm{cm} \gtrsim 10^4 \ \mathrm{cm}\,. \tag{4}$$

Here $\rho_6 \equiv \rho\,(\mathrm{cm}) \times 10^{-6}$, where ρ is the curvature radius of magnetic field lines in the gap region. In obtaining the numeric estimate in (4), we used the values $\rho_6 \sim P \sim \dot{P}_{-15} \sim 1$ and $b \sim 10$.

3.2 Spark formation

A number of such localized discharges can form in the gap, and each such a discharge undergoes a pair-creation cascade. The electric field in the gap accelerates the electrons towards the stellar surface, while the positrons are accelerated away from the surface. At the top of the gap these positrons acquire Lorentz factors of γ_b such that

$$\gamma_b \approx e\Delta V/m_e c^2 \ \sim \ 2 \times 10^6. \tag{5}$$

As these particles move away from the gap to the region where $\mathbf{E}.\mathbf{B} = 0$, they continue to create high energy photons, which further create pairs, and this cascade leads to the generation of a cloud of secondary electron–positron plasma, which has a significantly lower Lorentz factor with a mean value of γ_p. If there are n_b pairs in the primary beam, then the number of pairs in the secondary plasma can be estimated as $n_p \sim (0.5\gamma_b/\gamma_p)n_b$, and thus the density of the secondary plasma increases by a factor $\kappa = n_p/n_b$. In this work we will use the value $\kappa \sim 10^4$ Sturrock (1971). The burst of pair production process increases the charge density along the gap discharge stream and screens the potential in the gap. This process occurs exponentially, and after a certain from the star, which is estimated to be $\sim 30 - 40h \sim 500$ m (RS75), the charge density becomes close to n_{GJ}, and the particle acceleration process stops. During this time the discharge spreads in the lateral direction thus acquiring a width of $\sim h$. We will call this fully formed discharge a spark, and each spark is associated with a secondary plasma cloud. According to the above description, such a cloud has the shape of a column with a longitudinal dimension of ~ 500 m and a diameter of $\gtrsim 10$ m near the pulsar's surface. As the cloud moves away from the star to a distance R from its center, the cloud's diameter grows proportionally to $(R/R_s)^{1.5}$ due to the divergence of dipolar magnetic field lines.

The charge number density of the primary plasma beam, $n_b = n_{GJ}$, can be expressed in terms of pulsar parameters as

$$n_{GJ} \sim 6 \times 10^5 \left(\dot{P}_{-15}/P\right)^{0.5} R_{50}^{-3} \mathrm{cm}^{-3}, \tag{6}$$

where $R_{50} = R/(50R_s)$; and a value of $R_s = 10$ km will be used for subsequent calculations in

this paper. Here and below we normalize the distance to $50R_s$ since the coherent radio emission occurs around that altitude (see next subsection); hence $R_{50} \sim 1$. The charge number density of the secondary plasma is $n_p = \kappa n_{GJ}$, and hence the mean Lorentz factor of the secondary plasma can be estimated to be

$$\gamma_p \approx \gamma_b/(2\kappa) \sim 100\,. \tag{7}$$

The plasma frequency ω_p is given by

$$\omega_p \sim 4 \times 10^9 R_{50}^{-1.5} \kappa_4^{0.5} (\dot{P}_{-15}/P)^{0.25}\ \mathrm{Hz}, \tag{8}$$

where $\kappa_4 = \kappa/10^4 \sim 1$.

For later use in the next subsection, we will also mention that once the spark-associated plasma column leaves the polar cap and the secondary plasma cloud is screened from the gap's potential, the next discharge can be initiated. Thus, the distance between two consecutive plasma clouds is estimated to be $\sim h$.

3.3 Development of linear two-stream instability in secondary plasma

The overall sparking process leads to a non-stationary flow of successive plasma clouds flowing along a bundle of magnetic dipolar field lines. Indeed, since the magnetic field is strong and the ratio of the plasma frequency to cyclotron frequency $\omega_p/\omega_B \ll 1$ at the radio emission heights, the charges are confined to move tightly along the magnetic field lines. Therefore, we will hence restrict ourselves to discussing the plasma properties in a one-dimensional flow.

Within each plasma cloud, there is a spread of velocities of charged particles and hence a spread of their Lorentz factors γ_p. RS75 showed that due to this spread of velocities and a spatial overlap between the electron and positron species within the same cloud, a two-stream instability can develop. However, later, Asseo & Melikidze (1998) showed that a much stronger two-stream instability develops due to an overlap of the slow- and fast- moving particles of two *successive* clouds. Asseo & Melikidze 1998 showed that this two-stream instability provides a sufficiently strong Langmuir turbulence in the plasma, with the frequency of Langmuir waves ω_l being

$$\omega_l \approx \gamma_p \omega_p \sim 4 \times 10^{11} R_{50}^{-1.5}\,. \tag{9}$$

Moreover, disturbances of the envelope of these slowly modulated Langmuir waves, obeying the NLSE (MGP00), propagate with the group velocity that can coincide with the velocity of some portion of the charged particles. This occurs because of the aforementioned spread of particle velocities, both within one species and well as between the two species. This synchronism is behind the mechanism of the nonlinear Landau damping, first derived for this situation by MGP00.

To quantify that effect, as well as the size of nonlinear and dispersive terms in the NLSE in the next section, we will need to refer to the velocity distribution function of electrons and positrons in the secondary plasma. They can be approximated by a Gaussian function centered around momentum p_α and having a spread of p_T:

$$f_\alpha \propto \exp[-((p-p_\alpha)/p_T)^2] \sim \exp[-((\gamma-\gamma_{s,\alpha})/\gamma_T)^2]. \tag{10}$$

Here $\alpha = +$ or $-$, corresponding to positron or electron species of the plasma, and $\gamma_{s,\alpha}$ and γ_T are the Lorentz factors corresponding to the momenta p_α and p_T. It is important to notice that $f_+ \neq f_-$. Namely, it was shown by Cheng & Ruderman (1979) that due to the flow of plasma along the curved magnetic field lines, there is a change in the distribution function, and one can show (Asseo & Melikidze 1998) that:

$$|\gamma_{s,+} - \gamma_{s,-}|/\gamma_p \equiv \Delta\gamma/\gamma_p \approx \Delta\sigma\gamma_p^3/\gamma_b\,, \tag{11}$$

where the value of parameter $\Delta\sigma$ depends on whether the surface magnetic field at polar cap region is assumed to be strictly dipolar or to have a multipolar structure (MGP00); see section 2.3. Importantly for the estimate of the Landau damping parameter is section 5 below, it was shown in MGP00 that $\Delta\gamma/\gamma_p$ is in the range $0.5\ldots 2$.

Finally, an important parameter when solving the NLSE in subsequent sections is the distance from the pulsar's surface where the two-stream instability develops, thereby leading to strong Lamgmuir turbulence and hence strong radio emission. Simple kinematic estimates performed by MGP00 show that the distance at which this instability can set in is $\sim 2\gamma_p^2 h \gtrsim 200$ km. In terms of pulsar parameters, this can be written as,

$$R_{\rm onset} \sim 20R_s \times \gamma_2^2 \rho_6^{2/7} b^{-4/7} B_{12}^{-4/7} P^{3/7}\,, \tag{12a}$$

where $B_{12} = B_d/10^{12}$ G and $\gamma_2 = \gamma_p/10^2$. The longitudinal dimension of the instability region is limited by the divergence of the magnetic field lines, which leads to a decrease of the plasma charge density in proportion to $(R/R_s)^{1.5}$. As suggested in Asseo & Melikidze (1998), the dimension of this region is

$$\Delta R \sim 500 \ \ {\rm km}. \tag{12b}$$

Therefore, we can estimate that most of the radiation comes from the middle (or the second half) of this region, where, on one hand, the Langmuir turbulence is *already* strong because the two-stream instability has had sufficient time to develop, and, on the other hand, is *still* strong enough, having not been weakened by the divergence of the magnetic field lines. Thus, the location of the

radiating region can be estimated to be:

$$R_{em} \sim (R_{\rm onset} + \Delta R/2) \sim 500 \ \ {\rm km} \tag{12c}$$

from the star. Equivalently, one has $R_{50} \sim 1$ in this region. This is the value we used in (8) to obtain the estimate (9). Let us note that estimate (12c) is in good agreement with observations, as discussed in section 2.1.

In the next two sections we will discuss the nonlinear equation satisfied by the envelope of the Langmuir waves and the range of parameters for which it should be solved. Then, in section 6, we will demonstrate that the solution of that equation is capable of explaining features of the observed coherent radiation from pulsars.

4 NONLINEAR PROCESS AND CREATION OF RELATIVISTIC CHARGED SOLITON

In the Introduction we described a difficulty which the (linear) RS75 model, presented above, has explaining the origin of the coherent radio emission by plasma clouds. To address that difficulty, MGP00 proposed a nonlinear model for pulsar radio emission, whereby they included three new effects. First, they argued that due to a (relatively small) spread of particle velocities in the plasma, frequencies of the Langmuir waves have a corresponding spread $\Delta\omega \ll \omega_l$. Since the dispersion relation of Langmuir waves is not exactly linear, the group velocity dispersion needed to be included. Second, the nonlinearity of plasma's electric susceptibility was also included. As is well-known from the theory of nonlinear waves, an interplay between the nonlinearity and dispersion leads to modulational instability of the slowly changing envelope of the Langmuir waves, which in turn may lead to formation of localized pulses (solitons) out of suitable, localized initial conditions. Third, a resonant interaction between plasma particles propagating with speed that coincides with the (mean) group velocity of Langmuir wavepackets, results in the nonlinear Landau damping. Thus, the MGP00 model is described by the non linear Schrödinger equation (NLSE) with nonlinear Landau damping. Ichikawa & Taniuti (1973) derived the NLSE taking into account the effect of Landau damping for the non-relativistic case and applied it to an electron–ion plasma. Pataraya & Melikidze (1982) derived the NLSE for the relativistic case, and MGP00 applied it to the pulsar system in detail.

Let us now present the NLSE derived by MGP00:

$$i\frac{\partial E_\parallel}{\partial \tau} + G\frac{\partial^2 E_\parallel}{\partial \chi^2} + qE_\parallel\left(\left|E_\parallel\right|^2 + \frac{s}{\pi q}\fint \frac{|E_\parallel(\chi',\tau)|^2\, d\chi'}{\chi-\chi'}\right) = 0. \tag{13}$$

In this equation, $E_\parallel$ is the amplitude of the electric field which is parallel to the external magnetic

field, and χ and τ are the dimensional space and time coordinates in the moving frame of reference (MFR). They are related to the coordinates $\chi_\circ$ and $\tau_\circ$ in the observer frame of reference (OFR) by: $\chi = \gamma_\circ(\chi_\circ - v_g\tau_\circ)$ and $\tau = \gamma_\circ(\tau_\circ - (v_g/c^2)\chi_\circ)$, where v_g is the group velocity of Langmuir waves and $\gamma_\circ \approx \gamma_p$ is the corresponding Lorentz factor. In particular, any length and time intervals in the OFR are related to such intervals in the MFR by: $\delta\chi_\circ = \delta\chi/\gamma_p$ and $\delta\tau_\circ = \delta\tau/\gamma_p$. The coefficients G, q and s in Eq. (13) correspond to the dispersion, nonlinearity and Landau damping terms, respectively. These coefficients depend on the distribution functions $f_\alpha(p)$ (see section 3.3) and other parameters of the plasma and are given by Eq. (A20) in MGP00:

$$G = \frac{1}{4}\frac{\gamma_p^2 c^2}{\omega_p} G_d\,, \tag{14}$$

$$q = \left(\frac{e^2}{m_e c}\right)^2 \frac{1}{\gamma_p^2 \omega_p} q_d\,, \tag{15}$$

$$s = \left(\frac{e^2}{m_e c}\right)^2 \frac{1}{\gamma_p^2 \omega_p} s_d\,, \tag{16}$$

where G_d, q_d and s_d depend only on the distribution functions of the plasma. MGP00 solved the NLSE ignoring the effect of Landau damping. This, however, is justified only when the Landau damping term, $\propto s$, is so small compared to the nonlinear term, $\propto q$, that it does not have the chance to affect the evolution of the Langmuir wave's envelope over typical times considered. However, we will show below for reasonable parameters of pulsar plasma, this assumption does not necessarily hold; hence the effect of Landau damping over such sufficiently long times cannot be neglected. With this motivation, we will now proceed to solving the NLSE (13) *with* nonlinear Landau damping numerically.

5 NONDIMENSIONALIZATION OF EQ. (13)

We begin by nondimensionalizing the space variable in the MFR:

$$\chi = l\,\theta\,x, \tag{17a}$$

where x is the nondimensional space variable, factor θ is to be defined shortly, and the characteristic scale l is defined as:

$$l = c/\omega_p = \lambda_p/(2\pi)\,, \tag{17b}$$

with λ_p being the period of Langmuir waves in the MFR, and the plasma frequency ω_p is given by Eq. (8). The parameter $2\pi\,l$ is of the same order of magnitude as the Langmuir spatial period λ_p (in the MFR) at some intermediate location of the plasma cloud (see section 3). Therefore, factor $\theta/(2\pi)$ in (17a) characterizes the ratio of scales of the Langmuir wave envelope to the Langmuir

period λ_p.[2] In what follows we will consider this ratio to be on the order of $10^2 \ldots 10^3$, whence $\theta = 10^3 \ldots 10^4$.

We would like to stress that the above assumption about the range for θ does not impact the main conclusion of our study, for two reasons. First, as we will show below, a value of θ results merely in a ball-park estimate of the maximum nondimensional simulation time, while the freedom to adjust the nonlinearity coefficient in the nondimensional equation will still allow us to observe the important changes in the electric field's evolution. Second, it is realistic to assume that there exists a wider range of θ values than what we assumed above; this simply corresponds to the envelope of the initial electric field having a wider range of spatial scales. Then, for a given amplitude of the electric field's envelope (i.e., for a given size of nonlinear terms in Eq. (13)), only those of its fluctuations whose wavelength fall into a narrow(er) range of values will exhibit the phenomenon of pulse formation described below. In other words, the fluctuations with the spatial scale of interest to us will be selected by the governing Eq. (13) and not by our assumption of the range of θ (as long as that range is sufficiently broad).

In addition to an uncertainty in the value range of our nondimensional parameters that occur due to an uncertainty of the scaling parameter θ, there is also a (much smaller) uncertainty due to these parameters' dependence on the height R above the pulsar surface where the coherent radiation is emitted. In (12c) we estimated that this occurs around 500 km above the surface. Correspondingly, $R_{50} \sim 1$, as defined after Eq. (6), and this value is to be used to estimate the plasma frequency ω_p in (8) and hence the parameter l in (17b). For other values of R_{50}, one has $l \propto R_{50}^{1.5}$. We will use this fact in section 6.1 below.

Next, we normalize $E_\parallel$ to a typical magnitude of the electric field, E_0, at a location where the two-stream instability sets in in the cloud of secondary plasma (see (12a)):

$$E_\parallel = E_0\, u; \tag{18}$$

thus u is the nondimensional electric field. Note that by this definition, $u = O(1)$ initially.

Finally, we introduce the nondimensional time, t, via

$$\tau = \frac{(l\theta)^2}{G}\, t = \frac{4\theta^2}{\gamma_p^2 \omega_p\, G_d}\, t, \tag{19}$$

where we have used formula (17b) and the relation (14). Then, upon dividing through by $G/(l\theta)^2$, Eq. (13) attains the form:

$$i\frac{\partial u}{\partial t} + \frac{\partial^2 u}{\partial x^2} + Q\, u \left(|u|^2 + \frac{s}{\pi q} \fint \frac{|u(x',t)|^2\, dx'}{x - x'} \right) = 0. \tag{20}$$

[2] In other words, $x = O(1)$ corresponds to the dimensional scale of $O\big(\theta/(2\pi)\big)$ of Langmuir periods in the MFR.

where using Eqs. (14) and (15) we have: $Q = qE_0^2(l\,\theta)^2/G = \left(2E_0\theta\, e^2/(\gamma_p^2 m_e c\omega_p)\right)^2 q_d/G_d$. The initial condition to this equation has, by design, the nondimensional magnitude and spatial scale of order one: $|u| \sim 1$ and $|u|/|u_x| \sim 1$.

Let us now estimate the maximum nondimensional simulation time. First, the dimensional prototype of that time, $\tau_{\max}$, is that needed for a Langmuir wave packet to travel, within the cloud of secondary plasma, a length ΔR, where the Langmuir turbulence is sufficiently strong (i.e. the nonlinear and dispersive terms in (13) significantly affect the wave packet's evolution). As estimated in Eq. (12b), $\Delta R \sim 500$ km. Since that length is referenced in the OFR while τ is measured in the MFR, the Lorentz factor $\gamma_p \sim 10^2$ (see Eq (7)) needs to be accounted for; thus

$$\tau_{\max} \approx \frac{\Delta R}{c\,\gamma_p} \sim 10^{-5}\mathrm{s}. \tag{21a}$$

Using now relation (19), where $G_d \sim 1$ (see Fig. 1), and an estimate $\omega_p \sim 4\times 10^9$ Hz for a typical pulsar with $P \sim \dot{P}_{-15} \sim \kappa_4 \sim R_{50} \sim 1$ (see Eq. (8))), one obtains that:

$$t_{\max} \sim 10^8/\theta^2 \;\sim\; 1\ldots 10^2. \tag{21b}$$

Here the last estimate follows from our earlier assumption $\theta = 10^3 \ldots 10^4$.

Now, as we will demonstrate in the next section, the hallmark of the evolution of a Langmuir wave packet governed by Eq. (13) is the formation, out of an initially disordered state, of an intense pulse with an internal structure. In light of this, we should consider such values of nonlinear parameters Q and (s/q) in the nondimensional Eq. (20) that result in such formation over the times estimated in (21b). As for parameter (s/q), which characterizes the strength of Landau damping relative to the purely cubic nonlinearity (and does not depend on details of nondimensionalization; see (15) and (16)), its size was estimated following the procedure described in MGP00 (see MGP00 Appendix A). Correcting an arithmetic error in that paper, we re-calculated the parameters G_d and s_d/q_d for $\Delta\gamma/\gamma_p$ = 0.5 and 1; the result is plotted as a function of the thermal spread of the plasma, γ_T/γ_p, in Fig.1. For moderately large values of the thermal spread, $\gamma_T/\gamma_p \sim 1\ldots 2$, one finds the size of the Landau damping term relative to the size of the purely cubic nonlinear term fall in the range

$$s/q \;=\; s_d/q_d \;=\; 0.05\ldots 0.2. \tag{22}$$

On the other hand, it is not possible to estimate the correct order of magnitude of the nondimensional parameter Q in (20) based on physical grounds, because one does not have an estimate for the unknown electric field E_0. Therefore, we simply had to use trial and error to find a range

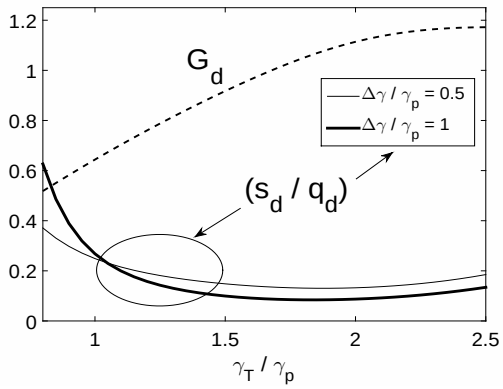

Figure 1. Parameters G_d and s_d/q_d in (15) and (16) as functions of the thermal spread of the plasma γ_T/γ_p, plotted for two representative values of the relative shift $\Delta\gamma/\gamma_p = 0.5,\ 1$ between the centers of the electron and positron distributions; see (11). The ellipse indicates grouping of the two s_d/q_d-curves, to distinguish them from the G_d-curve.

of values of Q such that, for (s/q) being in the range (22), the times of intense pulse formation are in the range (21b). As we will demonstrate in the next section, such Q have the magnitude $O(1)$.

6 NUMERICAL SIMULATIONS OF EQ. (20)

In order to solve Eq. (20), we used a numerical method recently proposed in Lakoba (2016). This method combines the leap-frog (LF) solver with the idea of the integrating factor (IF) method and hence will be referred to as IF-LF. For the reader's convenience we present its details in Appendix A. The non-physical parameters of the simulations, such as time step Δt and mesh size Δx, are also listed there.

The initial condition $u(x,0)$ for the envelope of the Langmuir wave is taken in the form of a mixture of a constant and a random field. In the simulations we vary both the ratio r of these two parts in the mixture and the correlation length $l_{\rm corr}$ of the random part, since the actual range of these physical parameters in the plasma is not known. Thus:

$$u(x,0) = (1-r) + r \int \frac{\exp[-0.5(k/k_{\rm corr})^2 - ikx]}{\sqrt{\sqrt{\pi}\, k_{\rm corr}}}\, \widehat{w}(k)\, dk; \tag{23a}$$

where $k_{\rm corr} = 2\pi/l_{\rm corr}$ and $\widehat{w}$ is a white noise in Fourier space:

$$\langle \hat{w}^*(k_1)\hat{w}(k_2)\rangle = 2\boldsymbol{\delta}(k_1 - k_2), \qquad \langle \hat{w}(k_1)\hat{w}(k_2)\rangle = 0; \tag{23b}$$

here $\langle\cdots\rangle$ stand for the ensemble averaging and $\boldsymbol{\delta}$, for the delta-function.

As we have announced in the previous section, we will be interested in an evolution that leads to the emergence of an intense pulse from an initially disordered state (23). Due to Landau damping, spectral components of the evolving electric field's envelope shift off-center during the evolution, which causes the forming pulse to move (with non-zero acceleration) in the reference frame of Eq. (20). Modeling realistically such a moving pulse would require a very large spatial com-

putational domain, which is out of reach for our computational resources. We therefore resorted to a common numerical trick: We impose periodic boundary conditions on a finite-length domain, thereby modeling repeated passing of a pulse (or some disordered field) through this domain. We restricted our consideration to the nondimensional value $L = 40\pi \approx 126$.

Thus, in our simulations we have four physical parameters that we varied: ratio r of the random and constant components in the initial state of the field (23); correlation length $l_{\rm corr}$ of the random component there; Landau damping coefficient (s/q), and the nonlinearity coefficient Q in (20).

6.1 Main results

6.1.1 Formation of an intense pulse due to Landau damping

We will now present the first of the two key findings of this study: the formation, due to Landau damping, of an intense pulse from a disordered initial field. A typical evolution of the field and its Fourier spectrum, representative of such a process, is illustrated in Fig. 2. At first, there is a transitional time interval (approximately until $t = 40$ in the case shown in Fig. 2), during which the field and its spectrum remain statistically similar to the initial ones. Some moderately intense field fluctuations routinely occur during that time, but they keep on quickly dissolving. This is a typical evolution of a disordered field in the NLSE: see, e.g., Solli et al. 2007; Fedele et al. 2010; Lakoba 2015. Then, a *new phenomenon* appears due to Landau damping: Within a short period (somewhere in $40 \lesssim t < 45$ in the case of Fig. 2 (b)), an intense pulse begins to form and, most importantly, *no longer dissolves back into a disordered state*. As the pulse keeps on becoming taller and narrower, its spectrum develops a secondary peak (circled in Fig. 2), which begins to shift exceedingly fast away from the original central wavenumber of the field. This shift of the field's spectral components occurs due to Landau damping, whereby energy from Fourier harmonics with $k > 0$ is transferred to those with $k < 0$ (for $(s/q) > 0$ in Eq. (20)). This stage, where the pulse "matures", takes a relatively short time (approximately corresponding to $45 < t < 55$ in the case of Fig. 2), after which the growth and narrowing of the pulse in physical space slow down and then cease. Then, the only result of its evolution remain the accelerated moving (in the reference frame of Eq. (20)) in the physical space and the moving of the secondary peak in the Fourier space. For the parameters used in this simulation, one is able to reliably observe this "mature" stage of the pulse evolution only for a relatively short time. The reason is that by $t \sim 65$, the secondary peak has already moved too close to the left edge of the computational spectral domain, so that spectral components of the solution near $k \approx -k_{\rm max}$ have increased considerably above the initial

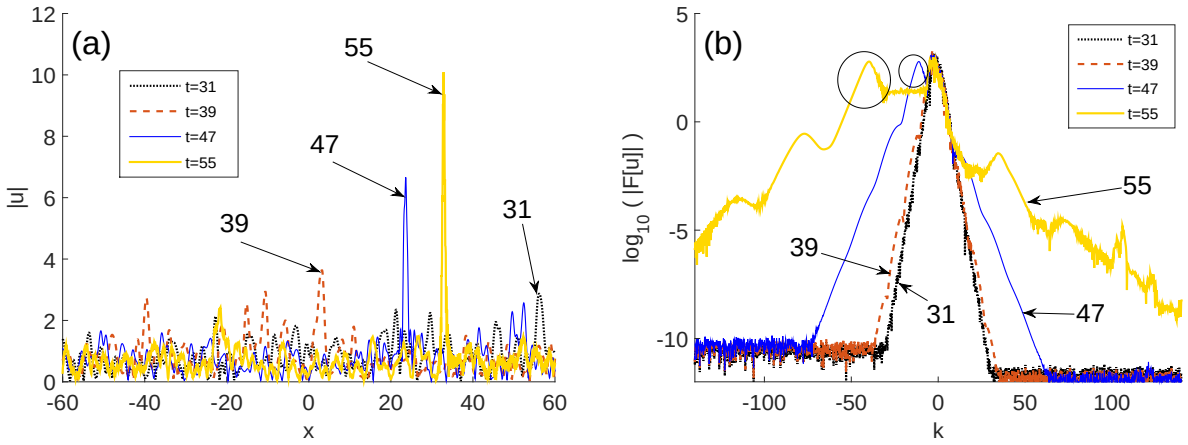

Figure 2. (Color online) Field (a) and spectrum (b) of the solution of Eq. (20) with the initial condition (23) for $Q = 0.25$, $r = 1$, $l_{\rm corr} = \pi$, and $s/q = 0.05$. (Only part of the spectral domain is shown for better visibility of details.) The arrows mark the times at which the solution is plotted. The circles in panel (b) mark the "solitonic" part of the pulse, corresponding to the intense peak, as discussed in the text.

noise level. To prevent those components from invalidating the numerical solution, we stopped simulations when the Fourier amplitude of those components would reach (an arbitrarily chosen) value 10^{-4}. A detailed observation and examination of the "mature" stage of the pulse evolution is possible either for a wider computational spectral domain (see below) or for smaller values of Q or (s/q). However, the latter would increase the time needed for the pulse formation to several hundred units, which is beyond the physically relevant range (21b), and therefore is not presented here.

The unbounded widening of the solution's spectrum presents an issue for the validity of that solution not only from the numerical, but also from a physical perspective. Indeed, one of the key assumptions under which the governing equation (13) is valid is that the characteristic scale of the initial perturbation in the plasma must be much greater than the Langmuir wavelength. After Eqs. (17), we assumed that the ratio of these two spatial scales, denoted there as $\theta/(2\pi)$, is on the order of $10^2 \ldots 10^3$. When the solution's spectrum widens M times, its characteristic spatial scale decreases by the same factor. Therefore, the governing model remains valid only as long as $\theta/(2\pi M) \gg 1$, whence one must require that the spectrum widening factor be limited by $M < 10^2$. If the initial (nondimensional) spectral half-width of the solution is $k \sim 2$, as in Fig. 2, then the solution will remain *physically* valid as long as the separation between the secondary and primary spectral peaks does not exceed ~ 100 units.

A closer examination of the solution's Fourier spectrum reveals that while the remnants of the original spectral peak remain "noisy" (i.e., jagged in Fig. 2), the secondary, "breakaway" peak, corresponding to the intense pulse, is smooth and has clearly seen exponentially decaying tails. This leads us to hypothesize that the created pulse is a soliton of the perturbed nonlinear Schrödinger

equation (20). Creation of a *long-living* intense pulse out of a disordered initial condition has been reported before (Jordan & Josserand 2001; Genty et al. 2010) for the generalized NLSE. However, in those cases, the size of the perturbation, interpreted as the difference between the given NLSE and the purely cubic one, was not small. Namely, either the nonlinearity was *considerably* different from cubic (Jordan & Josserand (2001) and references therein) or a higher-order dispersion term in the cubic NLSE was of order one (Genty et al. 2010). Moreover, there are two differences in our observation of the pulse's emergence compared to such observations in those earlier studies. First, our Eq. (20) contains only a *small* perturbation to the NLSE: $(s/q) \lesssim 0.1$.[3] This makes the second difference even more surprising: in strongly non-cubic generalized NLSE considered earlier by Jordan & Josserand (2001), the time that it took an intense pulse to emerge was about two to three orders of magnitude *greater* than in the *slightly* perturbed NLSE (20). A theoretical explanation of these differences, as well as of the very fact that a small Landau damping causes the emergence of an intense pulse from a disordered state, remains an open problem.

Thus, to summarize our first key finding: A small Landau damping leads to two qualitative changes of a disordered initial field. First, unexpectedly, it causes formation of an intense pulse that exists over a long time (at least as long as the model (13) remains physically and numerically valid). Second, expectedly, it leads to an (accelerated) shift of that pulse in the spectral domain towards lower wavenumbers (for $(s/q) > 0$).

6.1.2 Internal structure of the pulse, and features of radiation

Our second key finding concerns the *wavelength* of the radiation emitted by the plasma where an intense pulse has formed due to the mechanism described above. The intense pulse creates a ponderomotive force which prevents the charged bunch from collapsing. The ponderomotive force is $\propto \nabla|E_\parallel|^2$, which, when used in conjunction with the Poisson equation and restricted to the one-dimensional motion along magnetic field lines (see section 1), gives the charge density across the pulse to be proportional to $\partial^2|E_\parallel|^2/\partial x^2$ (see MGP00). The charge density, in turn, is a coherent structure bounded by the width of the intense pulse, which moves along curved magnetic field lines to produce coherent curvature radiation. To illustrate that finding, in Figs. 3 and 4 we compare the quantity $\partial^2|u|^2/\partial x^2$ (recall u is the nondimensional electric field given by Eq.(18)), in the initial state (Fig. 3) and in a state where a pulse has formed (Fig. 4). The simulation was run for the same

[3] We observed pulse formation even for values of $(s/q) < 0.01$, but it occurs over proportionally longer times, which are outside the physically relevant range (21b).

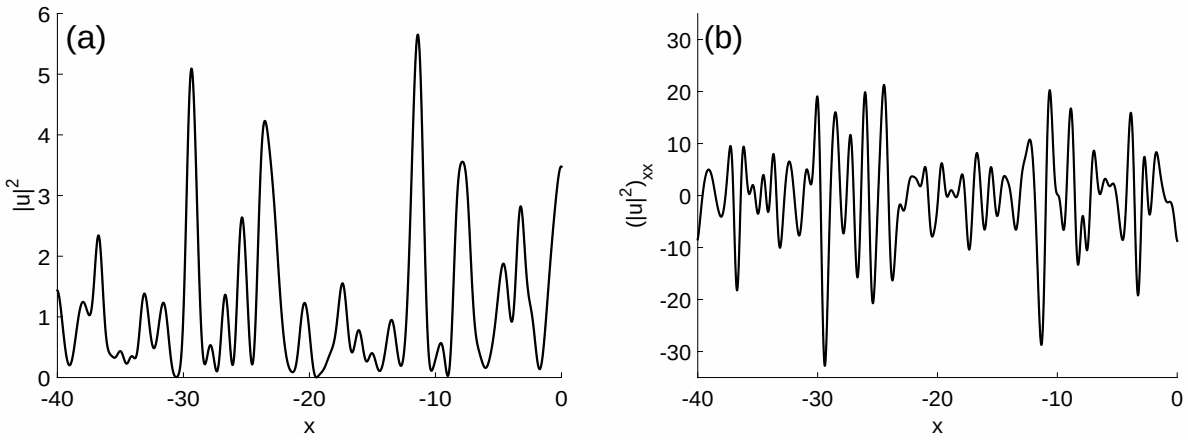

Figure 3. Quantities proportional to the intensities of the electric field (a) and radiation (b) in the initial state ($t = 0$). Simulation parameters are the same as in Fig. 2, except that the computational spectrum is three times broader. Only part of the spatial computational domain is shown for clarity.

parameters as in Fig. 2, except that, in order to resolve a high-wavenumber ripple in Fig. 4, we used a three-time wider spectral window and a correspondingly smaller time step (see Appendix A). Since the time of formation of an intense pulse is sensitive to the initial condition (and hence the computational spectrum), in Figs. 3 and 4 it is different from that in Fig. 2.

A comparison of Fig. 3(b) and Fig. 4(c) shows a more than three-order of magnitude increase of the radiation's intensity. It is important to note that this increase occurs due to two separate reasons: first, formation of a pulse whose amplitude, i.e. $\max|u|$, is an order of magnitude greater than that of the initial field (compare Fig. 3(a) to Fig. 4(a)), and second, existence of a highly oscillatory ripple "on top" of the pulse (see Fig. 4(b)). The spatial period of this ripple is about an order of magnitude smaller than unity, which adds approximately two orders of magnitude to the size of the second derivative of $|u|^2$ (see Fig. 4(c)).

The wavelength of the ripple $\delta x_{\rm ripple}$ "on top" of the intense pulse is explained by the spectrum of the field: see Fig. 4(d). Specifically, the "bulk" of the pulse corresponds to the "solitonic" spectral peak located at $k \approx -80$ in that figure. On the other hand, "remnants" of the initial field have $k \approx 0$. Thus, the intensity of the superposition of these two parts of the field, i.e.

$$|u|^2 \approx \left|A_1 e^{-i80x} + A_2 e^{i0x}\right|^2 = (|A_1|^2 + |A_2|^2) + 2|A_1||A_2|\cos(80x + \phi), \qquad (24)$$

has the approximate wavelength of the ripple as, $\delta x_{\rm ripple} = (2\pi)/80 \approx 0.07$. (In (24), A_1, A_2, and ϕ are some constants.) This wavelength is the smallest scale of a coherent structure inside the intense solitonic pulse and is clearly visible in Fig. 4(b) and 4(c). In section 7 we will demonstrate that while this structure is "short-lived" compared to the pulse itself, it still can be a source of coherent emission.

Thus, we have found two types of structures that emerge from an initially disordered state of

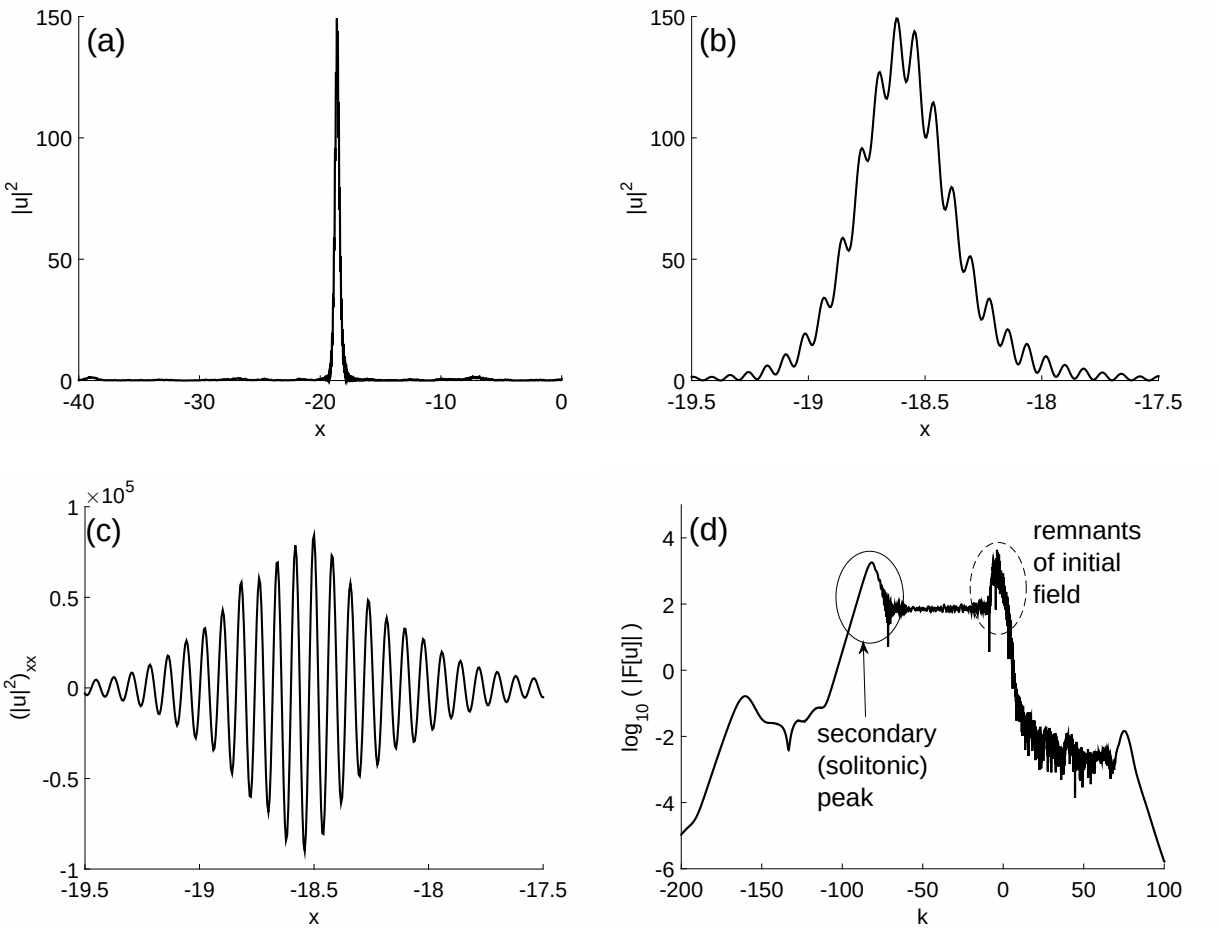

Figure 4. Same simulation as in Fig. 3, but at $t = 66$, after an intense pulse has formed and "matured". Panel (a) is a counterpart of that in Fig. 3. Panel (b) is a close-up of (a) focusing on the vicinity of the pulse. Panel (c) is a counterpart of Fig. 3(b). Panel (d) shows part of the Fourier spectrum of the numerical solution.

electric field in the pulsar plasma: (i) an intense pulse of the envelope of Langmuir waves and (ii) a "ripple" on top of this pulse. We will conclude this subsection with an estimation of their characteristic spatial scales and the corresponding frequencies that they can emit. We will begin with the intense pulse and use the parameters reported for, and shown in, Fig. 4. In dimensional units, the width of the pulse, $\delta\chi_{\mathrm{pulse},\circ}$, in the OFR can be estimated as follows. First, since the nondimensional time that it takes the intense pulse to form is $t \approx 60 \sim 100$, then (21b) implies that $\theta \sim 10^3$. Using this value and $\omega_p \sim 4 \times 10^9$ Hz in Eqs. (17) and a value $\gamma_p \sim 10^2$ to convert between the OFR and MFR variables, one obtains:

$$\delta\chi_{\mathrm{pulse},\circ} = \delta\chi/\gamma_p = l\theta\delta x_{\mathrm{pulse}}/\gamma_p \sim 3\times 10^8 \ \mathrm{m\cdot s^{-1}}/(4\times 10^9 \ \mathrm{s^{-1}})\cdot 10^3 \cdot 1/10^2 \sim 70 \ \mathrm{cm}. \quad (25a)$$

This corresponds to the radiation's frequency of about $c/\delta\chi_{\mathrm{pulse},\circ} \sim 400$ MHz and thus falls in the mid-range of the spectrum of observed pulsar coherent radio emission, which extends from several tens of MHz to a few GHz. Let us note that this estimate can go down by a factor of two or so if one allows for the possibility that the radiation is emitted from altitudes R_{em} that are lower

(s/q) \ Q	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0	1.1	1.2	1.3	1.4	1.5
0.05	40	30	20	12	6.1	5.8	4.0	3.0	3.4	2.9	2.0	1.8	1.5
0.10	21	11	7.5	5.5	4.4	3.5	3.0	3.5	2.8	2.1	2.0	1.6	1.5

Table 1. Evolution times $t_{4\times}$ that it takes the pulse's amplitude to exceed four times the average amplitude of the initial state. The initial condition (23) has parameters $r = 0.9$ and $l_{\rm corr} = \pi$. All simulations use the same seed of the random number generator, leading to the same initial pulse profile. Since details of the pulse evolution depend on the (randomly chosen) initial profile, the times $t_{4\times}$ are listed only to two significant figures, which suffices to illustrate the general trend.

than the value of 500 km assumed in (12c); see the paragraph before Eq. (18). Similarly, since the nondimensional wavelength $\delta x_{\rm ripple}$ of the ripple on top of the pulse is seen to be about an order of magnitude smaller, then

$$\delta\chi_{{\rm ripple},\circ} \sim 7 \ \ {\rm cm}, \tag{25b}$$

and the corresponding frequency is about 4 GHz. In the next subsection we will show that both frequencies following from estimates (25) may go down by about an order of magnitude if one assumes a higher value of the nonlinearity coefficient. In section 7 we will further discuss the relevance of our numerical observations to the problem of coherent emission by change bunches in plasma.

Having presented our main findings, we now describe how the formation of an intense pulse is affected by the physical parameters of the governing equation (20).

6.2 Dependence of pulse formation on Q

Predictably, as one increases Q without changing other parameters in (20), the field evolution due to nonlinear terms (both pure cubic and Landau damping ones) occurs faster, and the time required for the pulse formation decreases. This is illustrated in Table 1, where we list the times $t_{4\times}$ that it takes the pulse's amplitude, $\max_x |u|$, to exceed an arbitrarily set threshold of 4, which is four times the average amplitude of the initial state (23). These times are seen to decrease with the increase of Q, as expected. At the same time, we found that the spatial scale of the intense pulse is not significantly affected by Q. This fact will play a role in the forthcoming estimate of the dimensional wavelength and frequency of the coherent radio emission that such a pulse can generate.

Namely, we can accept that the distance ΔR that it takes for the instability in the secondary plasma cloud to lead to formation of an intense pulse is given by (12b); then the *dimensional* time of pulse formation continues to be given by (21a). In such a case, the decrease in the *nondimensional* time, seen in Table 1 as Q increases, implies that θ takes on values from the upper part of its range, e.g., $\theta \sim 10^4$: see (21b). Now, for a greater θ, a given nondimensional spatial scale

corresponds to a greater dimensional scale: see (17). Thus, as Q increases, the dimensional scale of both the solitonic pulse and the fine structure on top of it can go up by an order of magnitude compared to (25). Consequently, the frequencies of the coherent emission can be found in the range from several tens to several hundreds of MHz for the pulse and the "ripple", respectively.

We should be careful to note that this is only one interpretation of the observed decrease of $t_{4\times}$ with Q; other interpretations may be possible. For example, one can assume that the increase of Q implies that the Langmuir turbulence in the plasma cloud is so strong that the formation of an intense pulse occurs not over 500 km, as in (12b), but much sooner, say, over 100 km. In this case, *both* θ in (21b) and R_{50} in (8) would change compared to their values in (25). However, a more detailed analysis of such a possibility is outside the scope of this study.

Coming back to Table 1, we note that for the two different values of the Landau damping coefficient (s/q), the most pronounced decrease of $t_{4\times}$ occurs for a higher range of Q values for the smaller (s/q): for $Q \in (0.4,\, 1.0)$ for $(s/q) = 0.05$, and for $Q \in (0.3,\, 0.5)$ for $(s/q) = 0.10$. For both values of (s/q), the decrease of $t_{4\times}$ significantly slows down for Q values above those respective ranges.

In the next subsection we will discuss other changes in the pulse evolution that occur with changing the Landau damping coefficient (s/q).

6.3 Dependence of pulse formation on (s/q)

The expected effect of varying the Landau damping coefficient is that the formation time of an intense pulse decreases as (s/q) increases. However, and perhaps less expectedly, an increase of Landau damping beyond a certain point begins to have less effect on the speed of the pulse formation. This can be seen from the data reported for larger values of (s/q) in Table 2. Moreover, at least four other changes occur with the increase of Landau damping. First, the shape of the pulse becomes visibly asymmetric, with a "tail" forming behind the pulse; see Fig. 5(a). Second, this change in the shape is accompanied by a decrease of the amplitude of the "matured" pulse; compare Fig. 5(a) to Figs. 4(a,b). Third, the intensity of the radiation emitted by the pulse decreases, whereas its wavelength increases: compare Fig. 5(b) to Fig. 4(c) and note that respective horizontal and vertical scales are different. These effects are manifested in the Fourier space as follows: (i) the spectral peak corresponding to the intense pulse is now much less prominent over a spectral "plateau" that is observed immediately on its right side, and (ii) "remnants" of the initial field with spectral components near $k = 0$ have been considerably reduced for the larger value of (s/q).

$(Q;\ l_{\rm corr})$ \ (s/q)	0.01	0.02	0.04	0.06	0.08	0.10	0.13	0.16	0.20
$(1;\ \pi)$	14	8.0	4.5	4.1	3.5	3.4	3.1	3.0	2.6
$(1;\ 1)$	> 100	65	32	27	16	12	12	9.2	9.4
$(2;\ 1)$	30	14	11	7.0	6.1	4.6	3.5	2.5	4.4

Table 2. Dependence of the times $t_{4\times}$ on the Landau damping coefficient. The initial condition for each simulation is the same and has $r = 0.9$.

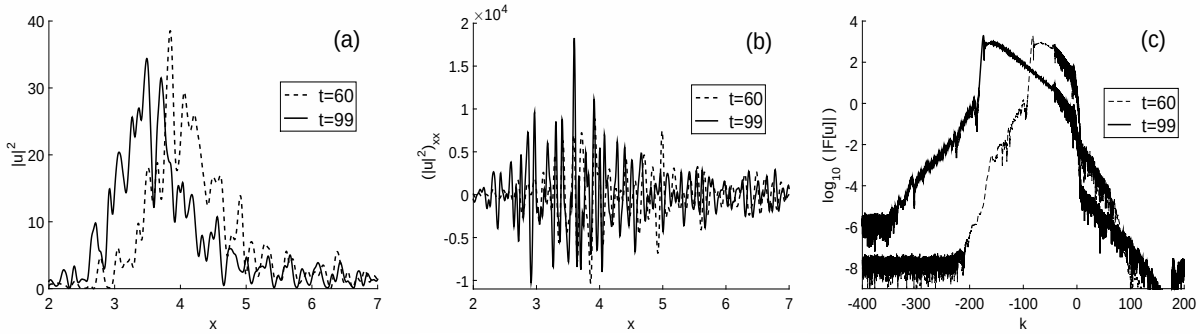

Figure 5. Same simulation parameters as for Fig. 4, but for a larger value of Landau damping $(s/q) = 0.15$. Solution at two times are shown to illustrate slow evolution of the field's spectrum for larger values of (s/q). The two times are chosen so that the locations of the intense pulse would nearly coincide, to facilitate visual comparison of the two pulses. Only part of spatial and spectral domains is shown for better visibility.

Thus, the spectrum of the field is considerably narrower for the larger values of Landau damping. Finally, and perhaps unexpectedly, the shift of the spectrum occurs much slower for larger values of (s/q). Specifically, the pulse shown in Fig. 5 for $(s/q) = 0.15$ forms around $t = 25$, whereas that shown in Fig. 4 for $(s/q) = 0.05$ forms around $t = 50$. The spectrum of the latter pulse approaches the left edge of the computational domain (i.e., $-k_{\max} \approx -600$ in this case) already for $t = 75$, by which time the numerical solution becomes invalid (see the beginning of section 6). On the contrary, the spectrum of the pulse for $(s/q) = 0.15$ is seen not to reach even one half of the computational spectral window by $t = 100$.

For completeness, we also verified that as (s/q) increases to become of order one, a pulse no longer forms. Instead, a "step" with an oscillatory "tail" is formed. The spectrum of this solution is approximately flat at the top, with the top's width increasing with time.

6.4 Dependence of pulse formation on $l_{\rm corr}$

The effect of the spatial scale of the initial condition on the field evolution is predictable, at least withing some range. Namely, as $l_{\rm corr}$ decreases, the effect of the dispersive term u_{xx} in Eq. (20) increases compared to that of the nonlinear term. Thus, having a smaller $l_{\rm corr}$ in the initial condition is, essentially, similar to having a smaller Q: it delays pulse formation. This is confirmed by comparing the first and second lines in Table 2, which show the effect of decreasing $l_{\rm corr}$. In comparison, the third and second lines of the same Table show that a similar effect of pulse formation delay occurs when Q is decreased, as has already been noted in section 6.2.

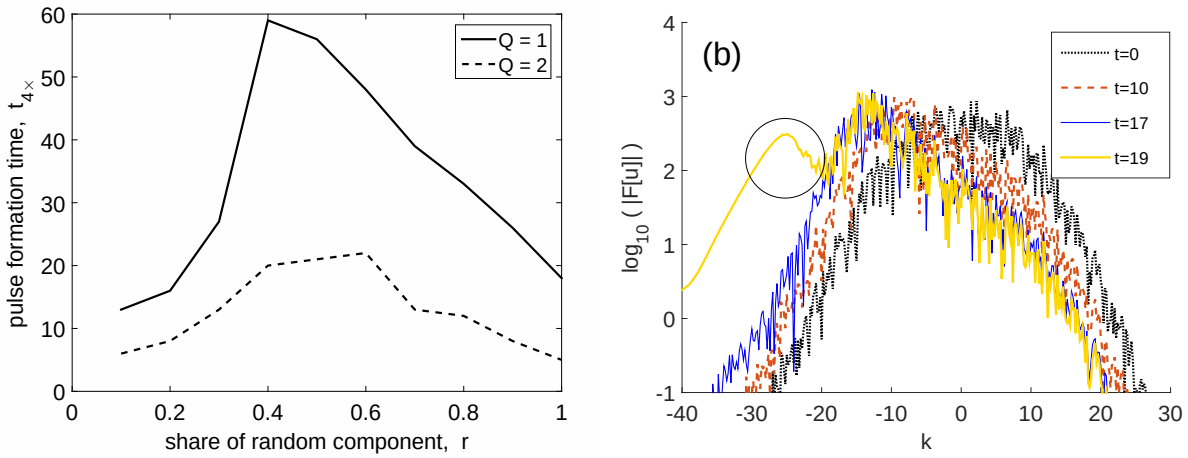

Figure 6. (a) Typical dependence of the pulse formation time on the share of the random component in the initial field. Parameters are: $l_{\rm corr} = 1$ and $(s/q) = 0.05$. To generate the random part of the initial condition, we used the same seed of the random number generator for all r. (b) (Color online) Fourier spectra illustrating the evolution of an initial random state ($r = 1.0$) into an intense pulse, as explained in the text; $Q = 1$. Only part of the spectral domain is shown for better visibility of details.

6.5 Dependence of pulse formation on r

Finally, we investigated how the “degree of randomness" of the initial state affects the pulse formation. Initially, we expected that the formation times would increase with the share of the random component in the initial state. However, in numerical experiments we observed that while those times indeed initially increase with r, they reach a maximum around $r = 0.5$ and then begin to decrease. Representative results are shown in Fig. 6(a). We observed qualitatively the same results for several other values of parameters Q and (s/q) than reported in Fig. 6, as well as for a different initial random state (controlled by the seed of the random number generator in the numerical code).

In Fig. 6(b) we show different stages of the field evolution, which were found to be similar for all $r \gtrsim 0.4$. Initially, Landau damping leads mostly to a (rather slow) shift of the field’s Fourier spectrum, with only minor changes of the spectrum’s shape; compare the curves for $t = 0$ and $t = 10$. In physical space, the field appears disordered during that stage. Then, the spectrum begins to become noticeably narrower and asymmetric; it is shown at $t = 17$, which is shortly before the formation of an intense pulse. (Incidentally, the spectral narrowing corresponds to the increase of the correlation length of the field, and, according to the previous subsection, this facilitates the pulse formation.) Finally, an intense pulse forms within a relatively short time interval; [4] see the curve for $t = 19$ in Fig. 6(b). In the spectrum, the formation of a pulse corresponds to the emergence of a secondary peak around $k \approx -30$, marked with a circle; compare it with Figs. 2(b) and

[4] In other simulations we observed that the duration of this time interval scales inversely proportional to Q, but seems to be rather insensitive to (s/q) for sufficiently small values of that parameter.

4(d). Once the pulse is formed, its height and width remain almost unchanged, until the numerical solution loses its validity due to a significant part of the spectrum shifting near the left edge of the spectral window (see section 6.1.1).

7 SUMMARY AND DISCUSSION

We have addressed the open problem of explaining a mechanism of coherent curvature radio emission by the electron–positron plasma in pulsar magnetosphere. As the mathematical model of this phenomenon we considered the generalized nonlinear Schrödinger equation (NLSE) proposed by Melikidze et al. (2000) (MGP00), which includes effects of group velocity dispersion, nonlinearity of electric susceptibility, and resonant interaction between Langmuir waves and plasma particles (Landau damping). In the absence of Landau damping, the (pure) NLSE can, in principle, support solitons, which in the plasma would be manifested as charge bunches that propagate stably and therefore are capable of emitting coherent radiation. However, formation of solitons in the pure NLSE requires that initially, the Langmuir wave have the envelope that is either localized or has several well-separated localized “bumps". It is only then that charge solitons could form and maintain their shape for a sufficiently long time to radiate coherently. There is no reason to expect that such a special initial condition of Langmuir waves would exist in a disordered pulsar plasma. Then, it is known that evolution of a *disordered initial state* in the pure NLSE leads to an ensemble of strongly interacting pulses, which constantly appear, disappear, and change their shape due to the interaction. Such a disordered, in both time and space, ensemble of pulses cannot be expected to emit coherently.

Motivated by this failure of the pure NLSE to identify a candidate mechanism of coherent emission, we numerically solved the NLSE *with the nonlinear Landau damping* term, as derived by MGP00. We found that for a range of realistic values of pulsar parameters, the presence of Landau damping leads to the formation of an intense, soliton-like pulse out of an initially disordered Langmuir wave. Such a stable pulse can emit coherently and thus is a reasonable candidate as a source of coherent radio emission.

Moreover, two important notes about this pulse formation are in order. First, the Landau damping coefficient has to fall in a certain range (namely, the lower part of (22)). If it is too high, then the intensity of the emerging pulse is lower, or a stable pulse may even not form at all; see section 6.3. On the other hand, if the Landau damping is too low, the pulse may not have the time to form

during the stage when the charge density in the plasma cloud is sufficiently large to produce strong radiation; see section 5.

Second, the intense pulse, formed for appropriate values of the Landau damping coefficient, has an internal structure whose spatial scale can be about an order of magnitude smaller than the spatial extent of the pulse itself; see Fig. 4(b). In this work we did not undertake an actual *calculation* of the coherent emission by charged stable pulses; this clearly requires a separate study. Without such a calculation, one cannot tell to what extent each of these structures: the "bulk" solitonic pulse itself and the finer "ripple" on top of it, contribute to the coherent emission. It appears intuitively plausible that frequencies in the lower end of the observed spectrum (tens to hundreds of MHz) are generated by the pulse as a whole, while frequencies from the higher end (up to several GHz) are generated by the "ripple". This is because the spatial scale of the pulse is about an order of magnitude greater than that of the "ripple"; see sections 6.1.2 and 6.2. However, a *calculation* of the spectrum emitted by such a *two-scale* structure of charges remains an open problem.

Let us now demonstrate that while the "ripple" on top of the solitonic pulse keeps changing its shape on a time scale that is small compared to the time scale where such a pulse exists, those changes are still "slow enough" to allow the "ripple" to emit coherently in the range of frequencies estimated in section 6.1.2 (several GHz), and even at lower frequencies. To that end, note that in order for the "ripple" to be a source of coherent radiation, it must exist long enough to guarantee condition (2a). Namely, the time $\mathcal{T}_b$ over which the shape of this "ripple" remains mostly unchanged must be much greater than the period $\mathcal{T}_c$ of the coherent radio emission. Let us demonstrate, using the illustrating example of Fig. 4, that this is indeed the case. In Fig. 7 we show that the profiles of both the electric field's intensity $|u|^2$ and the ponderomotive force $|u|^2_{xx}$ are mostly preserved over $t \approx 0.01$. Now, if $t \sim 100$ corresponds to 500 km (see section 5), then $\mathcal{T}_b \sim 0.01$ corresponds to about 50 m. Then, condition (2a) implies that the lower limit of frequencies ω_c is about $c/(50\ \mathrm{m}) \sim 10$ MHz. This is consistent (within a two-order of magnitude margin) with the value of several GHz mentioned after estimate (25b).

Note that since the solitonic pulse itself stably propagates over $t > O(10)$ nondimensional units, there is, for practical purposes, no lower limit from condition (2a) on the frequencies that it, as a whole, can emit coherently.

Finally, let us note that since we had to use periodic boundary conditions in our numerical simulations (see the preamble to section 6), we always observed that only one intense pulse forms as a result of many collisions with smaller pulses. In an actual plasma cloud, where the pulse

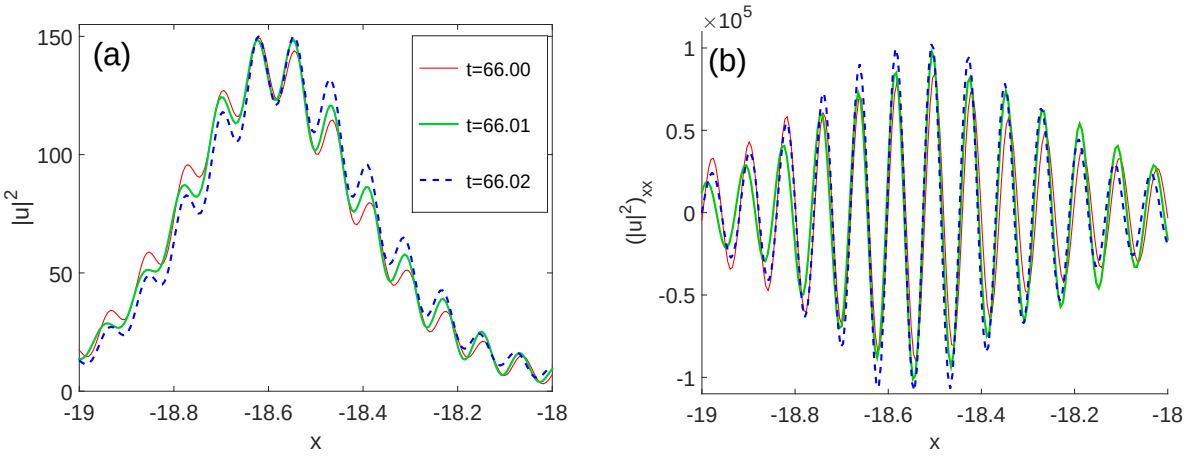

Figure 7. (Color online) Same quantities as in Fig. 4(b),(c), respectively, but for $t = 66.00,\ 66.01,\ 66.02$. Line colors and styles are as shown in the legend in panel (a). The x-window is smaller than in Fig. 4(b),(c) in order to make the details appear more clearly. In both panels, the centers of the pulses at the different times are manually superimposed in order to clearly show the changes of the profile. (If a pulse moves without changing its shape, its does not affect its ability to emit coherently.) We also observed that at $t = 66.10$ the profile of the "ripple" has changed completely relatively to that at $t = 66.00$; the corresponding curves are not shown in order not to clutter the picture.

passes through it only once rather than repeatedly, many well-separated and long-living solitonic pulses may form. Then, taking into account emission by this ensemble of stable charge bunches, as opposed to by a single charge bunch, is yet another open problem.

ACKNOWLEDGMENTS

The work of T.I.L. was supported in part by the NSF grant DMS-1217006. D.M. and G.M. would like to thank the Physics Department of the University of Vermont, where this project was started when D.M. held a visiting faculty position there working in collaboration with Prof J. Rankin. D.M. also thanks Prof. V.N. Kotov for useful discussions and encouragement during his work on this project.

REFERENCES

Asseo, E., & Melikidze, G. I. 1998, MNRAS, 301, 59

Blaskiewicz, M., Cordes, J. M., & Wasserman, I. 1991, ApJ, 370, 643

Briggs, W., Newell, A. C., & Sarie, T. 1983, J. of Computational Physics, 51, 83

Cheng, A. F., & Ruderman, M. A. 1979, ApJ, 229, 348

Dyks, J. 2008, MNRAS, 391, 859

Dyks, J., Rudak, B., & Harding, A. K. 2004, ApJ, 607, 939

Everett, J. E., & Weisberg, J. M. 2001, ApJ, 553, 341

Fedele, F., Cherneva, Z., Tayfun, M. A., & Guedes Soares, C. 2010, Physics of Fluids, 22, 036601

Genty, G., de Sterke, C. M., Bang, O., et al. 2010, Physics Letters A, 374, 989

Geppert, U. 2017, Journal of Astrophysics and Astronomy, 38, 46

Gil, J., Lyubarsky, Y., & Melikidze, G. I. 2004, ApJ, 600, 872

Gil, J., & Mitra, D. 2001, ApJ, 550, 383

Ginzburg, V. L., & Zhelezniakov, V. V. 1975, ARA&A, 13, 511

Goldreich, P., & Julian, W. H. 1969, ApJ, 157, 869

Johnson, M. D., Gwinn, C. R., & Demorest, P. 2012, ApJ, 758, 8

Johnston, S., Hobbs, G., Vigeland, S., et al. 2005, MNRAS, 364, 1397

Jordan, R., & Josserand, C. 2001, Math. Comp. Simul., 55, 443

Kazbegi, A. Z., Machabeli, G. Z., & Melikidze, G. I. 1991, MNRAS, 253, 377

Lai, D., Chernoff, D. F., & Cordes, J. M. 2001, ApJ, 549, 1111

Lakoba, T. I. 2015, Physics Letters A, 379, 1821

—. 2016, J. Sci. Comp., 72, 14

Lominadze, D. G., Machabeli, G. Z., Melikidze, G. I., & Pataraia, A. D. 1986, Fizika Plazmy, 12, 1233

Melikidze, G. I., Gil, J. A., & Pataraya, A. D. 2000, ApJ, 544, 1081

Melikidze, G. I., Mitra, D., & Gil, J. 2014, ApJ, 794, 105

Melikidze, G. I., & Pataraya, A. D. 1980, Astrophysics, 16, 104

—. 1984, Astrophysics, 20, 100

Mitra, D. 2017, Journal of Astrophysics and Astronomy, 38, 52

Mitra, D., Gil, J., & Melikidze, G. I. 2009, ApJ, 696, L141

Mitra, D., & Li, X. H. 2004, A&A, 421, 215

Mitra, D., Rankin, J., & Arjunwadkar, M. 2016, MNRAS, 460, 3063

Mitra, D., & Rankin, J. M. 2011, ApJ, 727, 92

Pétri, J. 2016, Journal of Plasma Physics, 82, 635820502

Radhakrishnan, V., & Cooke, D. J. 1969, Astrophys. Lett., 3, 225

Rankin, J. M. 2007, ApJ, 664, 443

Ruderman, M. A., & Sutherland, P. G. 1975, ApJ, 196, 51

Sloan, D. M., & Mitchell, A. R. 1986, J. of Computational Physics, 67, 372–395

Solli, D. R., Ropers, C., Koonath, P., & Jalali, B. 2007, Nature, 450, 1054

Spitkovsky, A. 2011, Astrophysics and Space Science Proceedings, 21, 139

Sturrock, P. A. 1971, ApJ, 164, 529

von Hoensbroech, A., & Xilouris, K. M. 1997, A&A, 324, 981

Weltevrede, P., & Johnston, S. 2008, MNRAS, 391, 1210

Young, M. D., Manchester, R. N., & Johnston, S. 1999, Nature, 400, 848

APPENDIX A: NUMERICAL METHOD FOR SOLVING EQ. (20)

We will present this method for equations of the form

$$iu_t + \mathcal{L}u + \mathcal{N} = 0, \tag{26}$$

where $\mathcal{L}$ is a linear operator with spatially constant coefficients and $\mathcal{N}$ includes all other terms. The generalized NLS (20) is a special case of (26), with $\mathcal{L} = \partial^2/\partial x^2$ and $\mathcal{N}$ being the entire nonlinear term. Below we will label the Fourier transform of any quantity with an over-hat:

$$\widehat{u}(k,t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} u(x,t)\, e^{-ikx} dx, \qquad u(x,t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \widehat{u}(k,t)\, e^{ikx} dx.$$

Similarly, $\widehat{\mathcal{L}} \equiv \widehat{\mathcal{L}}(k)$ and $\widehat{\mathcal{N}} \equiv \widehat{\mathcal{N}}(k)$ will denote the Fourier symbols of operators $\mathcal{L}$ and $\mathcal{N}$, respectively. For example, in (20), $\widehat{\mathcal{L}} = -k^2$. Taking the Fourier tranform of (26) yields

$$i\widehat{u}_t + \widehat{\mathcal{L}}\widehat{u} + \widehat{\mathcal{N}} = 0. \tag{27}$$

Solving this from t_1 to t_2 as a linear inhomogeneous equation yields:

$$e^{-i\widehat{\mathcal{L}}t_2}\,\widehat{u}(t_2) - e^{-i\widehat{\mathcal{L}}t_1}\,\widehat{u}(t_1) = \int_{t_1}^{t_2} e^{-i\widehat{\mathcal{L}}t'}\, i\widehat{\mathcal{N}}(t')\,dt', \tag{28}$$

where we have suppressed the obvious k-dependence of variables. Thus, the linear term in (26) is accounted for exactly by (28).

In Lakoba (2016) it was proposed to use the leap-frog scheme to discretize the integral on the r.h.s. of Eq. (28) and thereby turn that equation into a numerical method. The leap-frog scheme is well-known to quasi-preserve[5] the L_2-norm of the numerical solution. Choosing it is, therefore, appropriate for the wide class of equations (26) where this norm is conserved; the generalized NLS (20) with Landau damping belongs to that class. It should also be noted that the leap-frog scheme is explicit and hence easy to implement.

Within the leap-frog scheme, two versions of the discretization of the integral in (28) are still possible. As discussed in Lakoba (2016), one of them significantly distorts the solution's spectrum at the edges of the spectral computational window, while the other does not. For our purposes of modeling Landau damping, which constitutes a transfer of energy from higher- to lower-k Fourier harmonics (for $(s/q) > 0$), it is essential to have the spectrum undistorted at the edges. Therefore, in this study we used the latter method, which was called IF-LF (integrating factor–leap-frog) in Lakoba (2016).[6] The form of the IF-LF method is Lakoba (2016):

$$e^{-i\widehat{\mathcal{L}}t_{n+1}}\,\widehat{u}(t_{n+1}) - e^{-i\widehat{\mathcal{L}}t_{n-1}}\,\widehat{u}(t_{n-1}) = 2\Delta t\, e^{-i\widehat{\mathcal{L}}t_n}\, i\widehat{\mathcal{N}}(t_n)\,, \tag{29a}$$

where Δt is the time discretization step. For the generalized NLS (20) this simplifies to:

$$e^{ik^2\Delta t}\,\widehat{u}(t_{n+1}) - e^{-ik^2\Delta t}\,\widehat{u}(t_{n-1}) = 2i\Delta t\,\widehat{\mathcal{N}}(t_n)\,. \tag{29b}$$

This method has accuracy $O(\Delta t^2)$ in time; the discrete Fourier transform yields an exponential accuracy in space, provided that the solution with all its derivatives is continuous.

An extra step is now required to turn method (29) into a useful tool. Namely, if implemented just as above, the method will become numerically unstable over a time t_{inst}, which is on the order of a hundred time units for $Q = O(1)$. This numerical instability occurs for low-k Fourier harmonics and is caused by a so-called "parasitic" solution, which is well-known to be engendered

[5] i.e., possibly allow to fluctuate about the initial value, but not shift systematically

[6] Let us note, in passing, that the other method, which distorts the spectrum at the edges, has advantages over the IF-LF in numerical stability.

by the leap-frog scheme. As discussed in Lakoba (2016), for $Q < 0$ this instability is essentially a linear, modulational-type instability. On the other hand, for $Q > 0$, the instability is nonlinear and, to our knowledge, was first analyzed in Briggs et al. (1983); Sloan & Mitchell (1986). A method to suppress the instability caused by the "parasitic" solution was demonstrated and extensively tested in Lakoba (2016). It consists of averaging the solution every $t_{\rm stab}$ time units, with $\Delta t \ll t_{\rm stab} \ll t_{\rm instab}$, in a way that distorts the solution by a negligible amount. In this study, we used $t_{\rm stab} = 1$. We will now describe the aforementioned averaging procedure which we used to stabilize the numerical solution.

Denote $\widehat{v}_n = \widehat{u}_n e^{ik^2 t_n} \equiv \widehat{u}_n e^{ik^2 n\Delta t}$. Note that (29b) is then rewritten as:

$$\widehat{v}_{n+1} - \widehat{v}_{n-1} = 2i\Delta t\, e^{ik^2 t_n}\, \widehat{\mathcal{N}}(t_n)\,, \tag{30}$$

where $\widehat{\mathcal{N}}(t_n)$ depends on $\widehat{v}_n$. Suppose one has computed the solution up to time t_{n+3} inclusively. Using the solution computed in the last eight time steps, do the following. First, find the average at $t = t_n$:

$$\widehat{\overline{v}}_n = \frac{11}{64}\widehat{v}_n + \frac{15}{64}(\widehat{v}_{n-1} + \widehat{v}_{n+1}) - \frac{3}{32}(\widehat{v}_{n-2} + \widehat{v}_{n+2}) + \frac{1}{64}(\widehat{v}_{n-3} + \widehat{v}_{n+3}). \tag{31a}$$

One can verify that $|\widehat{v}_n - \widehat{\overline{v}}_n| = O(\Delta t^6)$. Next, repeat this for $t = t_{n-1}$:

$$\widehat{\overline{v}}_{n-1} = \frac{11}{64}\widehat{v}_{n-1} + \frac{15}{64}(\widehat{v}_{n-2} + \widehat{v}_n) - \frac{3}{32}(\widehat{v}_{n-3} + \widehat{v}_{n+1}) + \frac{1}{64}(\widehat{v}_{n-4} + \widehat{v}_{n+2}). \tag{31b}$$

Finally, use $\widehat{\overline{v}}_n$ and $\widehat{\overline{v}}_{n-1}$ to restart the IF-LF method (30) by replacing, on the l.h.s., $\widehat{v}_{n-1}$ with $\widehat{\overline{v}}_{n-1}$ and, on the r.h.s., $\widehat{v}_n$ with $\widehat{\overline{v}}_n$. As demonstrated in Lakoba (2016), this procedure suppresses the numerical instability while introducing only negligible dissipation to the solution. A logical flow-chart for implementing this stabilizing averaging in Matlab is found at:
`http://www.cems.uvm.edu/~tlakoba/recent_publications/stabilization_step_E_logical_flowchart.txt`.

The IF-LF method with a stabilization step based on (31) needs to use a time step Δt satisfying

$$\Delta t < \Delta x^2/\pi \equiv \pi/k_{\rm max}^2 \tag{32}$$

in order to guarantee numerical stability of high-k Fourier harmonics Lakoba (2016). In simulations reported in this study, we used the spatial domain of length $L = 40\pi$ and $N = 2^{13}$ equally spaced grid points; this corresponds to $\Delta x \approx 0.0153$ and $k_{\rm max} \approx 205$. According to (32), the threshold for high-k numerical stability is $\Delta t_{\rm thresh} = 7.5 \times 10^{-5}$, and we used $\Delta t = 5 \times 10^{-5}$ in all simulations.

APPENDIX B: LIST OF NOTATIONS AND ACRONYMS USED THROUGHOUT THE TEXT

$\alpha = (+)$ or $(-)$: labels positrons or electrons, respectively, in secondary plasma.

B_d: dipolar component of magnetic field.

B_s: magnetic field on surface of pulsar.

$b = B_s/B_d$.

c: speed of light in vacuum.

$\Delta\gamma = |\gamma_{s,(+)} - \gamma_{s,(-)}|$: difference between mean Lorentz factors of positrons and electrons in secondary plasma.

ΔR: distance along magnetic field lines where Langmuir turbulence is thought to be strong enough to lead to formation of coherent structures in charge density.

ΔV: potential drop across vacuum gap.

$E_\parallel$: envelope of Langmuir wave.

E_0: typical magnitude of $E_\parallel$ at a distance $R_{\rm onset}$.

e: charge of electron ($e > 0$).

f_α: momentum distribution function of type α particles in secondary plasma.

$\gamma_b \sim 2 \times 10^6$: average Lorentz factor of primary plasma beam particles.

$\gamma_p \sim 10^2$: average Lorentz factor of electrons and positrons in secondary plasma; $\gamma_2 = \gamma_p/10^2$.

γ_α, γ_T: mean and standard deviation of Lorentz factor γ in the secondary plasma of type α particles.

$h \gtrsim 10^3$ cm: height of the vacuum gap above the polar cap.

G: group velocity dispersion coefficient in NLSE (13).

G_d: related to G by (14).

$\kappa = n_p/n_b \sim 10^4$; $\kappa_4 = \kappa/10^4$.

λ_l: Langmuir wavelength in OFR.

λ_p: Langmuir wavelength in MFR at some reference location (around $R \sim 50R_s$).

L: nondimensional length of the computational domain used in section 6.

$l = c/\omega_p$: characteristic scale of Lamgmuir waves in MFR, used for normalization in section 5.

$l_{\rm corr}$: nondimensional correlation length of random field in initial condition (23a) for Langmuir wave.

m_e: mass of electron.

MFR: moving frame of reference.

n_GJ: Goldreich–Julian density.

n_b, n_p: density of electron-positron pairs in primary and secondary plasma, respectively.

NLSE: nonlinear Schrödinger equation.

$\Omega = 2\pi/P$: angular frequency of pulsar (nondimensional).

$\omega_l \sim 4 \times 10^{11}$ Hz: Langmuir frequency (in OFR).

$\omega_p \approx \omega_l/\gamma_p \sim 4 \times 10^9$ Hz: Langmuir frequency (in MFR) at some reference location (around $R \sim 50R_s$).

OFR: observer frame of reference.

$P \sim 1$: period of pulsar rotation (in seconds).

$\dot{P}$: rate of pulsar slow-down (nondimensional); $\dot{P}_{-15} = \dot{P}/10^{-15}$.

p_α, p_T: mean and standard deviation of the momentum distribution function f_α in the secondary plasma.

q: dimensional nonlinearity coefficient in NLSE (13).

q_d: related to q by (15).

Q: nondimensional nonlinearity coefficient in NLSE (20).

$\rho \sim 10^6$ cm: curvature radius of magnetic field lines in the vacuum gap; $\rho_6 = \rho/(10^6\,\mathrm{cm})$.

R_s: radius of pulsar (assumed to be 10 km in this paper).

R: distance from pulsar; $R_{50} = R/(50R_s) \sim 1$.

R_{em}: distance from the pulsar where coherent radio emission takes place.

$R_{\rm onset} \sim 200$ km: distance from the pulsar where Langmuir turbulence is thought to begin to develop.

r: relative part of random field in initial condition (23a) for Langmuir wave.

s: coefficient of nonlinear Landau damping in NLSE (13).

(s/q): magnitude of nonlinear Landau damping relative to pure cubic nonlinearity in NLSE.

s_d: related to s by (16).

θ: ratio of the characteristic scales of the envelope and carrier of the Langmuir wave, defined after (17); $\theta \sim 10^3 \ldots 10^4$.

τ, $\tau_\circ$: dimensional time variables in MFR and OFR, respectively; their relation is given after (13).

t: nondimensional time in MFR, related to τ by (19).

$t_{4\times}$: time that it takes amplitude of intense pulse to exceed four times average amplitude of initial state; see section 6.2.

u: nondimensional envelope of Langmuir wave; see (18).

χ, $\chi_\circ$: dimensional spatial variables along magnetic field lines in MFR and OFR, respectively; their relation is given after (13).

x: nondimensional space variable in MFR, related to χ by (17).