Nonlinear Functional Analysis and Applications Vol. 30, No. 4 (2025), pp. 1141-1152

ISSN: 1229-1595(print), 2466-0973(online)

https://doi.org/10.22771/nfaa.2025.30.04.10 http://nfaa.kyungnam.ac.kr/journal-nfaa



MODIFICATION OF THE ARMIJO LINE SEARCH TO SATISFY THE GLOBAL CONVERGENCE OF LIU-STOREY METHOD

Bouaziz khelifa¹, Iqbal M. Batiha^{2,3}, Nadir Djeddi⁴, Belal Batiha⁵ and Shaher Momani^{3,6}

 $^{1}\mathrm{Department}$ of Science and Technology, Larbi Tebessi University, Tebessa, Algeria

e-mail: khalifa.bouaziz@univ-tebessa.dz

²Department of Mathematics, Faculty of Science and Information Technology, Al Zaytoonah University of Jordan, Amman 11733, Jordan e-mail: i.batiha@zuj.edu.jo

³Nonlinear Dynamics Research Center (NDRC), Ajman University, Ajman, UAE e-mail: i.batiha@zuj.edu.jo

⁴Department of Mathematics, Larbi Tebessi University, Tebessa, Algeria e-mail: nadir.djeddi@univ-tebessa.dz

⁵Department of Mathematics, Jadara University, Irbid 21110, Jordan e-mail: b.bateha@jadara.edu.jo

⁶Department of Mathematics, University of Jordan, Jordan, Jordan e-mail: s.momani@ju.edu.jo

Abstract. In this paper, we propose a new Armijo-modified line search strategy that provides an efficient way to determine the step size at each iteration. The proposed modification ensures the global convergence of the original line search (LS) conjugate gradient method under some assumptions. By incorporating this modified line search, the algorithm benefits from stability when solving large-scale unconstrained optimization problems. To demonstrate the effectiveness of the proposed approach, we present a set of comprehensive numerical experiments. These tests compare the new scheme with existing classical line search techniques, highlighting its competitive performance in terms of accuracy, convergence speed, and computational efficiency. The results confirm that the proposed Armijo-modified line search is a valuable improvement to conjugate gradient methods.

⁰Received May 24, 2025. Revised October 10, 2025. Accepted October 18, 2025.

⁰2020 Mathematics Subject Classification: 65K05, 90C25, 90C26, 90C27.

 $^{^{0}\}mathrm{Keywords}$: Unconstrained optimization, conjugate gradient methods, global convergence, line search.

OCorresponding author: Iqbal M. Batiha(i.batiha@zuj.edu.jo).

1. Introduction

Unconstrained optimization is a central topic in numerical optimization, with applications in diverse fields such as machine learning, signal processing, and data science [3, 10, 11, 14, 15, 18]. Among iterative algorithms, conjugate gradient (CG) methods are particularly attractive for large-scale problems because of their low memory requirements and computational efficiency [7].

The Liu-Storey (LS) method [12] is a well-known variant of CG methods, but its global convergence is not always guaranteed without suitable line search strategies. Line searches play a crucial role in determining appropriate step sizes, directly affecting stability and convergence. The Armijo rule [2] is widely used for this purpose, as it ensures descent directions and is relatively simple to implement. However, when applied directly to the LS method in large-scale settings, it may fail to guarantee global convergence. To address this issue, we propose a modified Armijo line search that ensures the global convergence of the Liu-Storey method under mild assumptions.

Let us consider the following unconstrained minimization problem:

$$\min f(x), \quad x \in \mathbb{R}^n, \tag{1.1}$$

where f is a differentiable objective function, and has the following form

$$x_{k+1} = x_k + \alpha_k d_k, \tag{1.2}$$

where the step size α_k is obtained by carrying out some line search, and the direction d_k is defined by

$$d_k = \begin{cases} -g_k, & \text{for } k = 1, \\ -g_k + \beta_k d_{k-1}, & \text{for } k \ge 2, \end{cases}$$
 (1.3)

where $g_k = \nabla f(x_k)$ is the gradient of f at x_k , and β_k is a parameter that results in distinct conjugate gradient methods.

The Liu–Storey (LS) method [1] is a well-known conjugate gradient method. In the LS method, the parameter β_k is specified by

$$\beta_k^{LS} = \frac{-g_k^T y_{k-1}}{d_{k-1}^T g_{k-1}},\tag{1.4}$$

where $\|\cdot\|$ denotes the Euclidean norm, and $y_{k-1} = g_k - g_{k-1}$.

There are many approaches for finding an available step size. Among them, the exact line search is an ideal one, but it is cost-consuming or even impossible to use to find the step size. Some inexact line searches are sometimes useful and effective in practical computation, such as the Armijo line search [5, 8], and Goldstein and Wolfe line searches [4, 6, 9, 16, 17].

The Armijo line search is commonly used and easy to implement in practical computation.

Let s > 0 be a constant, $\rho \in (0,1)$ and $\mu \in (0,1)$. Choose α_k to be the largest α in $\{s, s\rho, s\rho^2, \ldots\}$ such that

$$f_k - f(x_k + \alpha d_k) \ge -\alpha \mu g_k^T d_k$$
.

The drawback of the Armijo line search is how to choose the initial step size s. If s is too large, then the procedure needs to call many more function evaluations. If s is too small, then the efficiency of the related algorithm will be decreased. Therefore, we should choose an adequate initial step size s at each iteration so as to find the step size α_k easily.

The remainder of this paper is organized as follows. Section 2 presents the proposed modification of the Armijo line search. Section 3 proves the convergence properties of the original Liu–Storey approach. Section 4 presents the global convergence analysis. Finally, numerical results and some discussions are given in Section 5.

2. Modification of the Armijo line search

To present the New Armijo-Modified analysis, we make the following fundamental assumptions about the objective function.

Assumption 2.1. The function f is LC^1 and strongly convex in \mathbb{R}^n , i.e., there exists a constant L > 0 such that

$$\|\nabla f(u) - \nabla f(v)\| \le L \|u - v\|, \quad \forall u, v \in \mathbb{R}^n.$$
 (2.1)

Note that Assumption 2.1 implies that the level set

$$\Gamma = \{x \in \mathbb{R}^n \mid f(x) \le f(x_0)\}$$
 is bounded.

Given $\mu \in \left(0, \frac{1}{2}\right)$, $\rho \in (0, 1)$ and $c \in \left[\frac{1}{2}, 1\right]$, the Armijo condition requires:

$$f(x_k + \alpha_k d_k) \le f(x_k) + c\alpha_k g_k^T d_k.$$

To satisfy this, we need a step size α_k that gives sufficient decrease. If ∇f is Lipschitz continuous with constant L_k , then we have the descent lemma:

$$f(x_k + \alpha_k d_k) \le f(x_k) + \alpha_k g_k^T d_k + \frac{L_k}{2} \alpha_k^2 \|d_k\|^2$$
.

This gives an upper bound on how much f can grow along the direction. To guarantee Armijo, we want

$$f(x_k + \alpha_k d_k) \le f(x_k) + c\alpha g_k^T d_k.$$

Using the descent lemma bound gives

$$f(x_k) + \alpha g_k^T d_k + \frac{L_k}{2} \alpha^2 \|d_k\|^2 \le f(x_k) + c\alpha g_k^T d_k.$$

Canceling $f(x_k)$ and rearranging yield

$$(1-c)(-g_k^T d_k) \ge \frac{L_k}{2} \alpha \|d_k\|^2.$$

So, we have

$$\alpha \le \frac{2(1-c)(-g_k^T d_k)}{L_k \|d_k\|^2}.$$

Therefore,

$$\alpha_k = \frac{(1-c)}{L_k} \frac{(-d_k^T g_k)}{\|d_k\|^2},\tag{2.2}$$

and α_k is the largest α in $\{s, s\rho, s\rho^2, \ldots\}$ such that

$$f_k - f(x_k + \alpha d_k) \ge -\alpha \mu g_k^T d_k \tag{2.3}$$

for $(-d_k^T g_k) > 0$, where L_k is an approximation to the Lipschitz constant L of g(u).

3. Convergent properties

In this section, we will reintroduce the convergence properties of the LS method. Now we give the following algorithm firstly.

Algorithm 3.1. (1) Step 0: Given $x_0 \in \mathbb{R}^n$, set $d_0 = -g_0$, k = 0.

- (2) Step 1: If $||g_k|| = 0$ then stop; else go to Step 2.
- (3) Step 2: Set $x_{k+1} = x_k + \alpha_k d_k$, where d_k is defined by (1.3), $\beta_k = \beta_k^{LS}$, and α_k is defined by the new Armijo-modified line search.
- (4) Step 3: Set k := k + 1, go to Step 1.

Some simple properties of the above algorithm are given as follows.

Lemma 3.2. ([3]) Let $\{x_k\}$ be a sequence generated by Algorithm 3.1. Then, under Assumption 2.1, the sequence $\{f(x_k)\}$ is non-increasing, and $\{x_k\}$ remains in the bounded level set Γ .

Lemma 3.3. Assume that Assumption 2.1 holds and that the LS method with the new Armijo-modified line search generates an infinite sequence $\{x_k\}$. If

$$\alpha_k \le \frac{(1-c)}{L} \frac{(-d_k^T g_k)}{\|d_k\|^2},$$
(3.1)

then

$$g_{k+1}^T d_{k+1} \le -c \|g_{k+1}\|^2. (3.2)$$

Proof. Under Assumption 2.1, by the Cauchy–Schwarz inequality and the LS method, we have

$$(1-c)(-d_k^T g_k) \ge \alpha_k L \|d_k\|^2$$

$$= \frac{\alpha_k L \|g_{k+1}\| \|d_k\|}{\|g_{k+1}\|^2} \|g_{k+1}\| \|d_k\|$$

$$\ge \frac{\|g_{k+1}\| \|g_{k+1} - g_k\|}{\|g_{k+1}\|^2} \|g_{k+1}\| \|d_k\|$$

$$\ge \frac{|g_{k+1}^T (g_{k+1} - g_k)|}{\|g_{k+1}\|^2} |g_{k+1}^T d_k|$$

$$\ge \frac{g_{k+1}^T (g_{k+1} - g_k)}{(-d_k^T g_k)} \frac{(-d_k^T g_k)}{\|g_{k+1}\|^2} (g_{k+1}^T d_k)$$

$$= \beta_{k+1}^{LS} \frac{(-d_k^T g_k)}{\|g_{k+1}\|^2} (g_{k+1}^T d_k).$$

Thus, we have

$$-c \|g_{k+1}\|^2 \ge -\|g_{k+1}\|^2 + \beta_{k+1}^{LS}(g_{k+1}^T d_k)$$
$$= g_{k+1}^T d_{k+1}.$$

Hence, the proof is completed.

Lemma 3.4. Assume that Assumption 2.1 holds. Then, the new Armijomodified line search is well defined.

Proof. On the one hand, since

$$\lim_{\alpha \to 0} \frac{f_k - f(x_k + \alpha d_k)}{\alpha} = -g_k^T d_k > -\mu g_k^T d_k,$$

there exists an $\alpha_k'>0$ such that

$$\frac{f_k - f(x_k + \alpha d_k)}{\alpha} \ge -\mu g_k^T d_k, \quad \forall \alpha \in (0, \alpha_k'].$$

Thus, letting $\alpha_k'' = \min\{s_k, \alpha_k'\}$ yields

$$\frac{f_k - f(x_k + \alpha d_k)}{\alpha} \ge -\mu g_k^T d_k, \quad \forall \alpha \in (0, \alpha_k''].$$

On the other hand, by Lemma 3.2, we can obtain

$$g(x_k + \alpha d_k)^T d(x_k + \alpha d_k) \le -c \|g(x_k + \alpha d_k)\|^2,$$

if
$$\alpha \leq \frac{(1-c)}{L} \frac{(-d_k^T g_k)}{\|d_k\|^2}$$
. Let

$$\bar{\alpha}_k = \min \left(\alpha_k'', \frac{(1-c)}{L} \frac{(-d_k^T g_k)}{\left\| d_k \right\|^2} \right).$$

We can prove that the new Armijo-modified line search is well defined when $\alpha \in (0, \bar{\alpha}_k]$. Hence, the proof is completed.

4. Global convergence

Lemma 4.1. Assume that Assumption 2.1 holds and that the LS method with the new Armijo-modified line search generates an infinite sequence $\{x_k\}$. Suppose there exist constants $m_0 > 0$ and $M_0 > 0$ such that $m_0 \leq L_k \leq M_0$. Then,

$$||d_k|| \le \left(1 + \frac{L(1-c)}{m_0}\right) ||g_k||, \quad \forall k.$$
 (4.1)

Proof. For k = 0, we have

$$||d_k|| = ||g_k|| \le \left(1 + \frac{L(1-c)}{m_0}\right) ||g_k||.$$

For k > 0, by the procedure of the new Armijo-modified line search, we have

$$\begin{split} \alpha_k & \leq s_k \\ & = \frac{(1-c)}{L_k} \frac{(-d_k^T g_k)}{\|d_k\|^2} \\ & \leq \frac{(1-c)}{m_0} \frac{(-d_k^T g_k)}{\|d_k\|^2}. \end{split}$$

By the Cauchy–Schwarz inequality, the above inequality, and noting the LS formula, we have

$$\begin{aligned} \|d_{k+1}\| &= \left\| -g_{k+1} + \beta_{k+1}^{LS} d_k \right\| \\ &\leq \|g_{k+1}\| + \left| \frac{g_{k+1}^T (g_{k+1} - g_k)}{(-d_k^T g_k)} \right| \|d_k\| \\ &\leq \left(1 + L\alpha_k \frac{\|d_k\|^2}{(-d_k^T g_k)} \right) \|g_{k+1}\| \\ &\leq \left(1 + \frac{L(1-c)}{m_0} \right) \|g_{k+1}\| \, . \end{aligned}$$

Hence, the proof is completed.

Theorem 4.2. Assume that Assumption 2.1 holds, the LS method with the new Armijo-modified line search generates an infinite sequence $\{x_k\}$, and there exist constants $m_0 > 0$ and $M_0 > 0$ such that $m_0 \le L_k \le M_0$. Then

$$\lim_{k \to \infty} \|g_k\| = 0. \tag{4.2}$$

Proof. Let $\eta_0 = \inf\{\alpha_k\}$. If $\eta_0 > 0$. Then we have

$$f_k - f_{k+1} \ge -\mu \alpha_k g_k^T d_k \tag{4.1}$$

$$\geq \mu \eta_0 c \|g_k\|^2. \tag{4.3}$$

Under Assumption 2.1, we have

$$\sum_{k=0}^{+\infty} \|g_k\|^2 < +\infty, \tag{4.4}$$

and thus,

$$\lim_{k \to \infty} \|g_k\| = 0. \tag{4.5}$$

In the following, we will prove that $\eta_0 > 0$.

For the contrary, assume that $\eta_0 = 0$. Then, there exists an infinite subset $K \subseteq \{0, 1, 2, \ldots\}$ such that

$$\lim_{k \in K, \, k \longrightarrow \infty} \alpha_k = 0. \tag{4.6}$$

Let $\alpha = \alpha_k/\rho$, there is a k' such that

$$\alpha_k/\rho \le s_k, \quad k \ge k'.$$

At least one the following two inequalities

$$f_k - f(x_k + \alpha d_k) \ge -\mu \alpha g_k^T d_k \tag{4.7}$$

and

$$g(x_k + \alpha d_k)^T d(x_k + \alpha d_k) \le -c \|g(x_k + \alpha d_k)\|^2$$
 (4.8)

does not hold.

If (4.7) does not hold, then we have

$$f_k - f(x_k + \alpha d_k) < -\mu \alpha g_k^T d_k.$$

Using the mean value theorem on the left-hand side of the above inequality, there exists $\theta_k \in [0, 1]$ such that

$$-\alpha g(x_k + \alpha \theta_k d_k)^T d_k < -\mu \alpha g_k^T d_k.$$

Thus

$$g(x_k + \alpha \theta_k d_k)^T d_k > \mu g_k^T d_k. \tag{4.9}$$

By the Cauchy–Schwarz inequality, (4.9) and Lemma 3.2, we have

$$L\alpha \|d_k\|^2 \ge \|g(x_k + \alpha \theta_k d_k) - g(x_k)\| \|d_k\|$$

$$\ge \left| (g(x_k + \alpha \theta_k d_k) - g(x_k))^T d_k \right|$$

$$\ge -(1 - \mu) g_k^T d_k$$

$$\ge c(1 - \mu) \|g_k\|^2.$$

We can obtain from Lemma 4.1 that

$$\alpha_k \ge \frac{c\rho(1-\mu)}{L} \frac{\|g_k\|^2}{\|d_k\|^2}$$

$$\ge \frac{c\rho(1-\mu)}{L} \left[1 + \frac{L(1-c)}{m_0} \right]^{-2},$$

 $k \ge k', k \in K$, which contradicts (4.6).

On the other hand, if (4.8) does not hold, then we have

$$g(x_k + \alpha d_k)^T d(x_k + \alpha d_k) > -c \|g(x_k + \alpha d_k)\|^2$$
 (4.10)

and

$$\begin{split} d(x_k + \alpha d_k) &= -g(x_k + \alpha d_k) \\ &+ \frac{g(x_k + \alpha d_k)^T (g(x_k + \alpha d_k) - g_k)}{-d_k^T g(x_k + \alpha d_k)} d_k. \end{split}$$

Thus, we have

$$\frac{g(x_k + \alpha d_k)^T (g(x_k + \alpha d_k) - g_k)}{-d_k^T g(x_k + \alpha d_k)} g(x_k + \alpha d_k)^T d_k > (1 - c) \|g(x_k + \alpha d_k)\|^2.$$

By using the Cauchy-Schwarz inequality on the left-hand side of the above inequality, we have

$$\alpha L \frac{\left\|d_k\right\|^2}{-d_k^T g(x_k + \alpha d_k)} > (1 - c).$$

Combining Lemma 3.2, we have

$$\alpha_k > \frac{\rho(1-c)}{L} \frac{-d_k^T g(x_k + \alpha d_k)}{\left[1 + \frac{L(1-\mu)}{m_0}\right]^2 \|g_k\|^2}$$
 $> 0,$

 $k \geq k', \ k \in K$ which also contradicts (4.6). This shows that $\eta_0 > 0$. The whole proof is completed.

Remark 4.3. The sequence $\{x_k\}$ generated by the LS method with the new Armijo-type line search converges to x^* , where $\nabla^2 f(x^*)$ is a symmetric positive definite matrix, and f(x) is twice continuously differentiable on

$$\Pi(x^*, \epsilon_0) = \{x : ||x - x^*|| < \epsilon_0\}.$$

5. Numerical experiments and discussion

In this section, we present several numerical experiments to demonstrate the efficiency of the new Armijo-modified line search used in the LS method. The test problems are taken from the CUTE library collection [7, 15], with dimension n=10,000. The parameters used in the experiments are set as follows: $\mu=0.25$, $\rho=0.75$, c=0.75, and $L_0=1$. Two conjugate gradient algorithms, denoted by LS and LS^+ , are compared in terms of numerical performance.

The tested methods are defined as follows:

- (1) **LS:** the standard LS method with the classical Armijo line search.
- (2) LS⁺: the LS method with the newly proposed Armijo-modified line search.

The inequality $||g_k|| \leq 10^{-7}$ is satisfied. All codes were implemented in MATLAB R2013.

The results are compared based on the number of iterations and CPU time. The performance of the algorithms is illustrated in Figures 1 and 2, showing the CPU time and number of iterations, respectively, using the performance profile approach introduced by Moré and Dolan [5].

In particular, Figure 1 compares the CPU time performance between the LS and LS^+ methods. For each method, the plot shows the fraction P of problems for which the method is within a factor t of the best performance. The left end of the curve indicates the percentage of problems for which the method was the fastest, while the right end indicates the percentage of problems that were successfully solved.

The experimental results indicate that the choice of initial points does not significantly influence the outcomes. Both the LS and LS^+ methods exhibit good performance. However, when comparing CPU time and the number of iterations, the LS^+ method consistently outperforms the LS method.

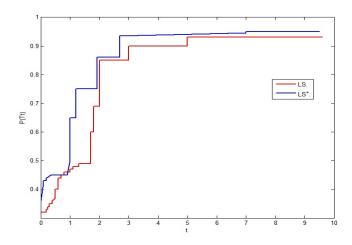


FIGURE 1. Performance profile based on CPU time.

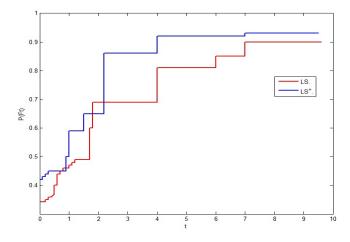


FIGURE 2. Performance profile based on function evaluations.

Nine testing problems have been taken from [15]. Table 1 lists the numerical results. The meaning of each column is as follows:

- (1) Problem": the name of the test problem;
- (2) k": the number of iterations;
- (3) Time": the CPU time in seconds.

The following results show the effectiveness of the proposed new Armijomodified line search.

Table 1. Numerical results for test problems using the new Armijo-modified LS method.

No.	Problem	k (Iterations)	Time (s)
1	Booth 1	22	0.015
2	Branin	171	0.008
3	Sphere	23	0.030
4	SUMSQUARES	26	0.015
5	Himmelblau	6	0.138
6	Matyas	50	0.016
7	McCormick	48	0.204
8	Six-Hump Camel	6	0.095
9	Three-Hump Camel	25	0.017

6. Conclusion

A new Armijo-modified line search for the LiuStorey method was proposed and analyzed. Theoretical results proved its global convergence, and numerical tests confirmed that it outperforms the classical LS method in both CPU time and iterations. The proposed approach is simple, efficient, and suitable for large-scale optimization problems.

Acknowledgment: The authors would like to express their sincere gratitude to Ajman University for its generous support in covering the publication fee and facilitating the completion of this research.

References

- N. Andrei, An unconstrained optimization test functions collection, Adv. Model. Optim., 10(1) (2008), 147–161.
- [2] L. Armijo, Minimization of functions having Lipschitz continuous partial derivatives, Pacific J. Math., 16 (1966), 1–3.
- [3] D. Bertsekas, Nonlinear Programming, Athena Scientific, (1999).
- [4] K. Bouaziz and T. Bachawat, Convergence of a modified PRP conjugate gradient method with a new formula of step-size, Ital. J. Pure Appl. Math., 49 (2023), 45-53.
- [5] I. Bongartz, A.R. Conn, N.I.M. Gould and Ph.L. Toint, CUTE: Constrained and unconstrained testing environments, ACM Trans. Math. Softw., 21(1) (1995), 123–160.
- [6] K. Bouaziz and Y. Laskri, Convergence of a two-parameter family of conjugate gradient methods with a fixed formula of stepsize, Bol. Soc. Paran. Mat., 38(6) (2020), 127-140.
- [7] Y.-H. Dai and Y. Yuan, Nonlinear Conjugate Gradient Methods, Shanghai Scientific & Technical Publishers, (2000).

- [8] R. Fletcher, Practical Methods of Optimization, 2nd ed., Vol. I: Unconstrained Optimization, Wiley, New York, (1997).
- [9] A.A. Goldstein, On steepest descent, SIAM J. Control, 3 (1965), 147–151.
- [10] T. Hamadneh, A. Hioual, O. Alsayyed, Y.A. Al-Khassawneh, A. Al-Husban and A. Ouannas, Finite-time stability results for neural networks described by variable-order fractional difference equations, Fractal and Fractional, 7 (2023), 616.
- [11] A.F. Jameel, N.R. Anakira, A.K. Alomari, M. Al-Mahameed and A. Saaban, A new approximate solution of the fuzzy delay differential equations, Int. J. Math. Modell. Numer. Optim., 9(3) (2019), 221-240.
- [12] Y. Liu and C. Storey, Efficient generalized conjugate gradient algorithms. Part 1: Theory, J. Optim. Theory Appl., 69(1) (1991), 129–137.
- [13] J. Moré and E.D. Dolan, Benchmarking optimization software with performance profiles, Math. Program., 91(2) (2002), 201–213.
- [14] A. Al-Natoor and F. Alrimawi, Singular value inequalities for concave and convex functions of matrix sums and products, European J. Pure Appl. Math., 18(1) (2025), 1-12.
- [15] J. Nocedal and S.J. Wright, Numerical Optimization, Springer, (2006).
- [16] P. Wolfe, Convergence conditions for ascent methods, SIAM Rev., 11 (1969), 226–235.
- [17] Y. Yuan, Numerical Methods for Nonlinear Programming, Shanghai Scientific & Technical Publishers, (1993).
- [18] R. Abu-Zurayk, N. Alnairat, H. Waleed, A. Khalaf, D. Abu-Dalo, A. Bozeya and R. Afaneh, Dual-mode integration of a composite nanoparticle in PES membranes: enhanced performance and photocatalytic potential, Nanomaterials, 15(14) (2025), 1055.