A narrow interval for the universal multiplier step of two-block ADMM

Dev D. Goyal

30 October 2026; version 1.1, 31 October 2026; revised 1 November 2026

Abstract

We study classical two-block ADMM with a fixed penalty when only the multiplier update is multiplied by a constant step τ. Fortin and Glowinski proved convergence for $0 < \tau < (1+\sqrt{5})/2$ [1], and whether this range can be enlarged for general convex problems is Glowinski's open question [2]. We show that, for finite-dimensional convex problems whose proximal subproblems attain their minima and which have a saddle point, the iteration converges for every $0 < \tau \le 377359/200000 = 1.886795$: both primal blocks and the multiplier converge to a saddle point, and the objective values converge to the optimum. No smoothness, strong convexity, bounded iterates or summability is assumed. Conversely, one fixed pair of compact convex polytopes in $\mathbb{R}^{128}$, with zero objectives, has a nonconstant periodic ADMM orbit for every $\tau \ge \tau_{\mathrm{up}} \approx 1.8868840$. The largest step with universal convergence therefore lies in an interval of width less than 0.00009. In dimension one, by contrast, the iteration converges for every $0 < \tau < 2$, and this range is exact. The convergence proof combines quadratic energies of up to thirty-two consecutive states with exact rational certificates that are affine in τ. Affine-constraint versions are also proved. Lean 4 formalizes the convergence theorems for $0 < \tau \le 377359/200000$ with all seven certificates and the periodic orbit; the scalar theorem has a conventional proof.

Contents

1 Introduction **2**

2 Setting and results **3**

3 Proximal steps and supporting inequalities **4**

4 Finite-memory certificates **5**
4.1 Centred variables and supporting gaps 5
4.2 The certificate identity 5
4.3 The classical branch $0 < \tau < 3/2$ 6
4.4 A three-state certificate on $[3/2, 7/4]$ 6
4.5 Longer memories on $[7/4, 377359/200000]$ 7

5 From certificates to convergence **7**

6 Affine constraints **8**

7 The periodic obstruction **8**

8 Scalar problems **9**

9 An earlier published claim **10**

10 Formal verification **10**

11 Related work and novelty **11**

12 Limitations and open questions **12**

13 Research process and AI disclosure **13**

1 Introduction

The alternating direction method of multipliers (ADMM) for

$$\min\{f(x) + g(z) : x = z\}$$

alternates two proximal steps with a multiplier update. Fortin and Glowinski allowed the multiplier update to be scaled by a constant τ and proved convergence for $0 < \tau < (1 + \sqrt{5})/2$ [1]. Whether this golden-ratio range can be enlarged for general convex problems has been an open question; Tao and Yuan call it Glowinski's open question and answer it for quadratic objectives [2]. Gu and Yang showed that the conventional energy used in convergence proofs can increase beyond the golden ratio, so that proof technique cannot cross it [3]. Larger steps are known under extra assumptions, such as strong convexity [5], an added proximal term [6], or summability of multiplier increments [4]. The earlier result of this collection is an exact five-cycle at $\tau = 199/100$ for compact polytopes, valid from $\tau \geq 140111/70637 \approx 1.98353554$, which rules out universal convergence on all of $(0, 2)$ [15].

This paper brackets the universal threshold. The convergence theorem (Theorem 2.2) covers every $0 < \tau \leq 377359/200000 = 1.886795$. The periodic orbit (Theorem 2.4) exists for every $\tau \geq \tau_{\mathrm{up}}$, where

$$\tau_{\mathrm{up}} = \frac{3884783130026510326218902 7}{2058835202293311936890077 3} \approx 1.8868839651174105.$$

Hence the supremum τ_* of steps with universal convergence satisfies

$$1.886795 \leq \tau_* \leq \tau_{\mathrm{up}}, \qquad \tau_{\mathrm{up}} - 1.886795 < 0.00009. \tag{1.1}$$

The exact value of τ_* is not determined. Version 1.0 of this paper had the bracket [1.886, 1.8868954]; the present one is about ten times narrower.

The threshold depends on the dimension. For scalar problems the universal range is exactly $(0, 2)$ (Theorem 2.5): every scalar instance converges for $0 < \tau < 2$, and a trivial instance fails for every $\tau \geq 2$. The proof uses a coercive Lyapunov function built from the order structure of the real line rather than a quadratic certificate.

The method is not new. Convergence comes from a quadratic Lyapunov energy of several consecutive states, certified by nonnegative combinations of convex interpolation inequalities and positive-semidefinite matrices; this is the approach of automated Lyapunov analysis [8, 9]. The periodic orbit uses vertex normal-cone certificates for polytopes, as in the earlier five-cycle. What is new, as far as the literature we inspected shows, is the bracket (1.1) for the unmodified iteration in the general convex class, and the exact scalar range of Theorem 2.5.

Lean 4 checks the convergence theorem, the affine-constraint theorems, all seven longer-memory certificates and the periodic orbit (Section 10). The scalar theorem and a few routine reductions,

listed in that section, rest on the conventional proofs given here, and the exact certificates are also rechecked by independent rational-arithmetic programs.

Gao et al. asserted convergence of this iteration for every $1 < \tau \le 2$ [7]. Section 9 explains why that assertion cannot hold: its displayed inequality fails on a scalar example, the iteration has a two-cycle at $\tau = 2$, and Theorem 2.4 contradicts it for every $\tau \ge \tau_{\mathrm{up}}$.

Version history. Version 1.0 (30 October 2026) proved convergence for $0 < \tau \le 943/500$ and a periodic orbit for $\tau \ge 897744153596994489/475778439151265168 \approx 1.8868954$. Version 1.1 (31 October 2026) extends convergence to $0 < \tau \le 377359/200000$, lowers the orbit threshold to τ_{up} and adds Theorem 2.5. The statements of version 1.0 remain valid, and the Lean development now covers the extended range and the new orbit (Section 10).

2 Setting and results

Let E be a finite-dimensional real inner-product space. Let $C, D \subseteq E$ be convex sets and let $f, g : E \to \mathbb{R}$ be convex on C and D, respectively; values outside C and D play no role. Equivalently, one may add the convex indicators of C and D.

Definition 2.1. A point p is a *proximal point* of f on C at center c if $p \in C$ and

$$f(p) + \tfrac{1}{2}\|p - c\|^2 \le f(y) + \tfrac{1}{2}\|y - c\|^2 \qquad (y \in C).$$

We write $p = P_f(c)$. A *saddle point* is a pair (z_*, u_*) with $z_* \in C \cap D$ and

$$f(y) \ge f(z_*) - \langle u_*, y - z_* \rangle \ (y \in C), \qquad g(y) \ge g(z_*) + \langle u_*, y - z_* \rangle \ (y \in D).$$

Adding the two inequalities of a saddle point shows that z_* minimizes $f + g$ on $C \cap D$.

Standing hypotheses. For every $c \in E$, proximal points of f on C and of g on D exist, and a saddle point exists. Convexity makes proximal points unique (Lemma 3.1).

The algorithm, with penalty one in consensus form, is

$$x^{k+1} = P_f(z^k - u^k), \qquad z^{k+1} = P_g(x^{k+1} + u^k), \qquad u^{k+1} = u^k + \tau(x^{k+1} - z^{k+1}). \tag{2.1}$$

Only the last update contains τ. This is the multiplier-only relaxation of Fortin and Glowinski, not the over-relaxation that replaces x^{k+1} by $\alpha x^{k+1} + (1 - \alpha) z^k$ in the later steps.

Theorem 2.2 (Convergence)**.** *Assume the standing hypotheses and fix $0 < \tau \le 377359/200000$. For every initial $(z^0, u^0) \in E \times E$ there is a saddle point $(\bar{z}, \bar{u})$ such that*

$$x^{k+1} \to \bar{z}, \qquad z^k \to \bar{z}, \qquad u^k \to \bar{u},$$

and $f(x^{k+1}) \to f(\bar{z})$, $g(z^{k+1}) \to g(\bar{z})$. In particular $f(x^{k+1}) + g(z^{k+1})$ converges to the minimum of $f + g$ on $C \cap D$.

For a penalty $\rho > 0$ apply Theorem 2.2 to f/ρ and g/ρ; the variable u is then the multiplier divided by ρ. When f and g are proper, closed and convex with effective domains C and D, the proximal points exist for every center by the standard theory, so the standing hypotheses reduce to existence of a saddle point. The formal statement keeps attainment as an explicit hypothesis.

Theorem 2.3 (Affine constraints)**.** *Let U, V, E be finite-dimensional real inner-product spaces, let $A : U \to E$ and $B : V \to E$ be affine maps of arbitrary rank, let $C \subseteq U$ and $D \subseteq V$ be convex, and let f and g be convex on C and D. Fix $\rho > 0$ and $0 < \tau \leq 377359/200000$. Assume that for every center the subproblems*

$$\min_{x \in C} f(x) + \tfrac{\rho}{2}\|A(x) - c\|^2, \qquad \min_{y \in D} g(y) + \tfrac{\rho}{2}\|B(y) - c\|^2$$

attain their minima, and that the problem $\min\{f(x) + g(y) : x \in C,\ y \in D,\ A(x) = B(y)\}$ has a saddle point. Let (x_n, y_n, u_n) be any sequence produced by

$$x_{n+1} \in \arg\min_{x \in C} f(x) + \tfrac{\rho}{2}\|A(x) - B(y_n) + u_n\|^2, \quad y_{n+1} \in \arg\min_{y \in D} g(y) + \tfrac{\rho}{2}\|A(x_{n+1}) - B(y) + u_n\|^2,$$

and $u_{n+1} = u_n + \tau(A(x_{n+1}) - B(y_{n+1}))$, with arbitrary choices among minimizers. Then $A(x_n)$, $B(y_n)$, ρu_n and $f(x_n) + g(y_n)$ converge to the corresponding values at a saddle point of the original problem, and that limit value is the optimum. If A and B have continuous left inverses, then x_n and y_n also converge.

The usual form $A_0 x + B_0 y = b$ is the case $A = A_0$, $B(y) = b - B_0 y$. Without injectivity the raw iterates need not converge: if f does not depend on a coordinate of x that A ignores, an arbitrary choice of minimizers can move that coordinate at will.

Theorem 2.4 (Periodic orbit)**.** *There are compact convex polytopes $C, D \subset \mathbb{R}^{128}$, both containing 0, such that for every $\tau \geq \tau_{\mathrm{up}}$ the iteration (2.1) with $f = g = 0$ on C and D has a nonconstant periodic orbit whose period divides 128. Every proximal step on the orbit is the unique projection onto the respective polytope, $(0, 0)$ is a saddle point, $\|x^k - z^k\|^2 = 16383999999992680528112756654 08$ for every k, and z^k has no limit.*

The polytopes do not depend on τ; the initial multiplier is proportional to τ.

Theorem 2.5 (Scalar problems)**.** *Let $E = \mathbb{R}$ and assume the standing hypotheses. For every $0 < \tau < 2$ and every initial $(z^0, u^0) \in \mathbb{R}^2$ the conclusions of Theorem 2.2 hold. For every $\tau \geq 2$ the instance $C = D = \mathbb{R}$, $f = g = 0$, $z^0 = 0$, $u^0 = 1$ gives $z^k = 0$, $u^k = (1 - \tau)^k$ and $x^{k+1} = -(1 - \tau)^k$, which do not converge. Hence the set of steps $\tau > 0$ with universal convergence for scalar problems is exactly $(0, 2)$.*

Corollary 2.6. *The bracket (1.1) holds for the supremum τ_* of steps t such that every $0 < \tau \leq t$ gives convergence in the class of Theorem 2.2.*

The obstruction uses indicator objectives. Thresholds for smooth or strongly convex subclasses are not addressed; the earlier five-cycle has a strongly convex C^2 extension [15], but only from $\tau \geq 140111/70637$.

3 Proximal steps and supporting inequalities

For $x \in C$ write $a \in \partial_C f(x)$ when $f(y) \geq f(x) + \langle a, y - x\rangle$ for every $y \in C$. This does not assume differentiability.

Lemma 3.1. *$p = P_f(c)$ if and only if $p \in C$ and $c - p \in \partial_C f(p)$. Proximal points are unique and P_f is nonexpansive.*

Proof. If $c - p \in \partial_C f(p)$, add the supporting inequality to $\frac{1}{2}\|y-c\|^2 - \frac{1}{2}\|p-c\|^2 = \frac{1}{2}\|y-p\|^2 - \langle c-p, y-p\rangle$. Conversely, compare p with $p + t(y-p)$ for $0 < t \le 1$, use convexity, divide by t and let $t \downarrow 0$. For two centers, adding the supporting inequalities at $p = P_f(c)$ and $q = P_f(d)$ gives $\|p-q\|^2 \le \langle c-d, p-q\rangle$, which proves uniqueness and nonexpansiveness. □

Lemma 3.2. *Let $y_k \to y$ in C, let $a_k \in \partial_C f(y_k)$ converge, and let $b \in \partial_C f(y)$. Then $f(y_k) \to f(y)$.*

Proof. $f(y) + \langle b, y_k - y\rangle \le f(y_k) \le f(y) - \langle a_k, y - y_k\rangle$, and both outer terms tend to $f(y)$. □

Lemma 3.2 gives objective convergence without assuming continuity of f at boundary points of C, where convex functions can jump.

4 Finite-memory certificates

4.1 Centred variables and supporting gaps

Fix a saddle point (z_*, u_*) and replace f and g by

$$\widehat{f}(x) = f(x + z_*) + \langle u_*, x\rangle - f(z_*), \qquad \widehat{g}(z) = g(z + z_*) - \langle u_*, z\rangle - g(z_*).$$

Then $\widehat{f}, \widehat{g} \ge 0$ vanish at the origin with supporting normal zero there. Put $\zeta^k = z^k - z_*$ and $v^k = (u^k - u_*)/\tau$. The multiplier update gives $x^k - z_* = \zeta^k + v^k - v^{k-1}$, and Lemma 3.1 identifies the normals

$$a^k = \zeta^{k-1} - \tau v^{k-1} - (x^k - z_*) \in \partial\widehat{f}(x^k - z_*), \qquad b^k = (x^k - z_*) + \tau v^{k-1} - \zeta^k \in \partial\widehat{g}(\zeta^k). \quad (4.1)$$

The points are independent of τ and the normals are affine in τ.

In a window of m updates, label the saddle point as sample 0 and the updates as samples $1, \dots, m$. For every ordered pair of samples the supporting gaps

$$I_f(i,j) = \widehat{f}(X_j) - \widehat{f}(X_i) - \langle a_i, X_j - X_i\rangle \ge 0, \qquad I_g(i,j) = \widehat{g}(Z_j) - \widehat{g}(Z_i) - \langle b_i, Z_j - Z_i\rangle \ge 0$$

hold, where X_i, Z_i, a_i, b_i are the centred points and normals. A certificate assigns nonnegative weights $\lambda^f_{ij}, \lambda^g_{ij}$ whose row and column sums agree at every sample. Then all function values cancel in $\sum \lambda^f_{ij} I_f(i,j) + \sum \lambda^g_{ij} I_g(i,j)$, which is a quadratic form in the state window and is nonnegative on actual ADMM windows.

4.2 The certificate identity

Write a window of $m+1$ consecutive states as $w = (\zeta^0, v^0, \dots, \zeta^m, v^m) \in E^{2m+2}$ and, for a real symmetric matrix M, $[w]_M = \sum_{i,j} M_{ij}\langle w_i, w_j\rangle$. Let S_0 and S_1 select the first and last m states of the window, and let H_m represent $\|\zeta^m - \zeta^{m-1}\|^2 + \|v^m - v^{m-1}\|^2$. A certificate for the step interval $[\alpha, \beta]$ consists of rational symmetric matrices P_α, P_β and T_α, T_β and one set of weights such that, at both endpoints $\tau \in \{\alpha, \beta\}$,

$$[S_0 w]_{P(\tau)} - [S_1 w]_{P(\tau)} - [w]_{H_m} - \sum_{i,j} \lambda^f_{ij} I_f(i,j) - \sum_{i,j} \lambda^g_{ij} I_g(i,j) = [Dw]_{T(\tau)} \quad (4.2)$$

as an identity of quadratic forms, where D maps the window to differences of its states, so that $[Dw]_T$ vanishes on constant windows, and

$$P(\tau) - C_m \succeq 0, \qquad T(\tau) \succeq 0, \quad (4.3)$$

where C_m represents $\|\zeta^{m-1}\|^2 + \|v^{m-1}\|^2$ on the newest stored state. Because the normals (4.1) are affine in τ, so is the left side of (4.2). Defining $P(\tau)$ and $T(\tau)$ by linear interpolation between the endpoints, (4.2) holds for every $\tau \in [\alpha, \beta]$, and (4.3) holds by convexity of the positive-semidefinite cone. Scalar identities lift to vectors by bilinearity.

4.3 The classical branch $0 < \tau < 3/2$

For $k \geq 1$ let $r_k = v^k - v^{k-1} = x^k - z^k$, $q_{k+1} = \zeta^{k+1} - \zeta^k$ and $E_k = \|\zeta^k\|^2 + \tau\|v^k\|^2$. The supporting inequalities between the new proximal points and the saddle point give

$$E_k - E_{k+1} \geq (2-\tau)\|r_{k+1}\|^2 + \|q_{k+1}\|^2 + 2\langle r_{k+1}, q_{k+1}\rangle,$$

and monotonicity of the g-normals at consecutive points gives $\langle q_{k+1}, r_{k+1} + (\tau-1)r_k\rangle \geq 0$. With $W_k = E_k + |\tau-1|\|r_k\|^2$ and $2|t\langle a, b\rangle| \leq |t|(\|a\|^2 + \|b\|^2)$,

$$W_k - W_{k+1} \geq (2 - \tau - |\tau-1|)\|r_{k+1}\|^2 + (1 - |\tau-1|)\|q_{k+1}\|^2.$$

Both coefficients are positive exactly for $0 < \tau < 3/2$, so W_k is a coercive energy with strict decrease of the increments.

4.4 A three-state certificate on $[3/2, 7/4]$

Here $m = 3$. The weights, with zero diagonals and samples ordered as saddle, first, second and third update, are

$$\Lambda_f = \begin{pmatrix} 0 & 3327 & 4525 & 3133 \\ 3011 & 0 & 25648 & 88936 \\ 3180 & 0 & 0 & 26993 \\ 4794 & 114268 & 0 & 0 \end{pmatrix}, \qquad \Lambda_g = \begin{pmatrix} 0 & 3957 & 3650 & 3378 \\ 2996 & 0 & 230607 & 0 \\ 3595 & 168638 & 0 & 150993 \\ 4394 & 61008 & 88969 & 0 \end{pmatrix}.$$

Each matrix has equal row and column sums at every index. The energy matrices on $(\zeta^0, v^0, \zeta^1, v^1, \zeta^2, v^2)$ are given by

$$8P_{3/2} = \begin{pmatrix} 532740 & -116092 & -189104 & 138756 & -331592 & -22664 \\ -116092 & 800006 & 306244 & -242242 & -190182 & -537802 \\ -189104 & 306244 & 488680 & -231192 & -286856 & -71272 \\ 138756 & -242242 & -231192 & 370768 & 93176 & -101376 \\ -331592 & -190182 & -286856 & 93176 & 637624 & 95536 \\ -22664 & -537802 & -71272 & -101376 & 95536 & 657976 \end{pmatrix},$$

$$8P_{7/4} = \begin{pmatrix} 599964 & -137604 & -298760 & 169140 & -289160 & -31536 \\ -137604 & 897627 & 378338 & -195741 & -240779 & -678597 \\ -298760 & 378338 & 697952 & -308936 & -386472 & -64992 \\ 169140 & -195741 & -308936 & 249496 & 140936 & -22080 \\ -289160 & -240779 & -386472 & 140936 & 694808 & 98128 \\ -31536 & -678597 & -64992 & -22080 & 98128 & 722608 \end{pmatrix},$$

and $P(\tau) = (7-4\tau)P_{3/2} + (4\tau-6)P_{7/4}$. The matrices $P_{3/2} - C_3$ and $P_{7/4} - C_3$ and the two dissipation matrices $T_{3/2}$, $T_{7/4}$ have exact rational LDL^{T} factorizations with positive pivots. The dissipation matrices are not independent input: they are the residual of (4.2) after subtracting the supporting gaps, and the independent checkers rebuild them from the iteration.

4.5 Longer memories on $[7/4, 377359/200000]$

Seven further certificates use longer windows. The first two date from version 1.0; the other five extend the range in version 1.1.

Step interval	States m	Energy matrix	Listed weights	Positive weights
[7/4, 47/25]	12	24×24	312	188
[47/25, 943/500]	20	40×40	840	353
[943/500, 47157/25000]	20	40×40	840	353
[47157/25000, 23579/12500]	20	40×40	840	219
[23579/12500, 3773/2000]	20	40×40	840	840
[3773/2000, 18867/10000]	20	40×40	840	840
[18867/10000, 377359/200000]	32	64×64	2112	2112

Their exact data are in the directory `certificates/`, one JSON file per interval, named by the decimal endpoints and the memory, for example `exact_188_1886_m20.json` and `exact_18867_1886795_m32.json`. Each interval uses one set of weights at both endpoints; different intervals use different weights. The rational checker verifies (4.2) at both endpoints, flow balance and nonnegativity of the weights, and exact LDL^{T} factorizations with positive pivots of $P_\alpha - C_m$, $P_\beta - C_m$, T_α and T_β. The Lean proof instead uses integer Gram certificates: for a rational matrix $A_{\mathbb{Z}}/d$ it checks $q^2 A_{\mathbb{Z}} = dR^{\mathsf{T}}R + E$ with $d, q > 0$ and an integer symmetric E that is diagonally dominant with nonnegative diagonal, which implies $A_{\mathbb{Z}}/d \succeq 0$. For the five certificates of version 1.1 the Lean data keep the matrices P, T and the weights of the JSON files, scaled to integers, and add such Gram factors; the kernel evaluates their checks as sequential sums over lists of integers, which a generic lemma identifies with the matrix entries.

5 From certificates to convergence

Fix τ and a certificate of memory m covering it; the classical branch is the case $m = 1$ with energy W_k. Relative to a saddle point define

$$\mathcal{E}_k = [(\zeta^k, v^k, \dots, \zeta^{k+m-1}, v^{k+m-1})]_{P(\tau)}.$$

By (4.3), $\mathcal{E}_k \geq \|\zeta^{k+m-1}\|^2 + \|v^{k+m-1}\|^2$. By (4.2) and nonnegativity of the gaps and of $[Dw]_{T(\tau)}$,

$$\mathcal{E}_k - \mathcal{E}_{k+1} \geq \|\zeta^{k+m} - \zeta^{k+m-1}\|^2 + \|v^{k+m} - v^{k+m-1}\|^2.$$

Hence the states are bounded and their increments are square summable; in particular $x^{k+1} - z^{k+1} = v^{k+1} - v^k \to 0$. These are conclusions, not hypotheses.

The state map $T_\tau : (z, u) \mapsto (z^+, u^+)$ is continuous by Lemma 3.1. A bounded sequence in finite dimension has a convergent subsequence $s_{k_j} \to \bar{s}$. Vanishing increments give $s_{k_j+1} \to \bar{s}$, and continuity gives $s_{k_j+1} = T_\tau(s_{k_j}) \to T_\tau(\bar{s})$, so $\bar{s}$ is a fixed point. Because $\tau > 0$, a fixed point has $x = z$, and Lemma 3.1 turns the two proximal equations into the saddle-point inequalities. Now centre the energy at this saddle point $\bar{s}$. Along the subsequence the finitely many states in the window converge to $\bar{s}$, so $\mathcal{E}_{k_j} \to 0$; monotonicity gives $\mathcal{E}_k \to 0$ and coercivity gives $s_k \to \bar{s}$. Continuity of P_f gives $x^{k+1} \to P_f(\bar{z} - \bar{u}) = \bar{z}$.

Finally $a^{k+1} = z^k - u^k - x^{k+1} \to -\bar{u}$ and $b^{k+1} = x^{k+1} + u^k - z^{k+1} \to \bar{u}$ are supporting normals converging to the saddle normals, so Lemma 3.2 gives $f(x^{k+1}) \to f(\bar{z})$ and $g(z^{k+1}) \to g(\bar{z})$. The intervals $(0, 3/2)$, $[3/2, 7/4]$, $[7/4, 47/25]$, $[47/25, 943/500]$ and the five intervals of version 1.1 cover $(0, 377359/200000]$, which proves Theorem 2.2.

6 Affine constraints

For Theorem 2.3 define the infimal images

$$F(a) = \inf\{f(x) : x \in C,\ A(x) = a\}, \qquad G(b) = \inf\{g(y) : y \in D,\ B(y) = b\}$$

on the images $A(C)$ and $B(D)$. The original saddle point bounds every fibre from below, so F and G are finite and convex there. If x attains the original affine subproblem, it also attains the fibre infimum at $A(x)$; otherwise a better point in the same fibre would improve the subproblem. Hence $A(x)$ is the proximal point of F/ρ at the corresponding center, and every choice of original minimizers induces the same consensus iteration on the images, with multiplier u. Theorem 2.2 gives a limiting image saddle point (a, q). Attainment of the original subproblems at the centers $a - q$ and $a + q$, together with uniqueness of image proximal points, provides original points whose images both equal a; they attain their fibre infima and form an original saddle point with multiplier ρq. Fibre optimality identifies the original objective values along the trajectory with the image values, which proves convergence of $f(x_n) + g(y_n)$ without closedness of F or G. Continuous left inverses transfer image convergence to the raw iterates.

7 The periodic obstruction

The certificate file `certificates/exact_cycle_1886884_K128.json` supplies integer vectors $z, v \in \mathbb{Z}^{128}$. Let S be the cyclic left shift, so $S^{128} = I$, and set

$$x = z + v - S^{-1}v, \qquad C = \operatorname{conv}\{S^i x : 0 \le i < 128\}, \qquad D = \operatorname{conv}\{S^i z : 0 \le i < 128\}.$$

The coordinates of z and v sum to zero, hence so do those of x, and averaging all shifts shows $0 \in C \cap D$. With zero objectives, $(0, 0)$ is a saddle point.

For each offset j define the vertex margins

$$M_f(j, \tau) = \langle S^{-1}z - \tau S^{-1}v - x, x - S^j x\rangle, \qquad M_g(j, \tau) = \langle x + \tau S^{-1}v - z, z - S^j z\rangle.$$

Each is an integer plus τ times an integer. The certificate verifies that every slope is nonnegative and that every margin is nonnegative at $\tau = \tau_{\mathrm{up}}$; the binding inequality is the g-margin at offset $j = 1$. Hence all margins are nonnegative for every $\tau \ge \tau_{\mathrm{up}}$.

Nonnegative vertex margins extend by convexity to the whole polytope: if $p \in K$ and $\langle c - p, y - p\rangle \le 0$ for every $y \in K$, then $\|y - c\|^2 = \|p - c\|^2 + \|y - p\|^2 - 2\langle c - p, y - p\rangle \ge \|p - c\|^2 + \|y - p\|^2$, so p is the unique projection of c onto K. Define $x^i = S^i x$, $z^i = S^i z$ and $u^i = \tau S^i v$. Because S is orthogonal, the base inequalities give both proximal steps at every phase, and the multiplier update is the identity $u^i - u^{i-1} = \tau(S^i v - S^{i-1}v) = \tau(x^i - z^i)$. The sequences have period dividing 128, the residual $\|x - z\|^2 = 16383999999992680528112756654 08$ is constant, and $Sz \ne z$ gives $\|z^{k+1} - z^k\| = \|Sz - z\| > 0$ at every step, so z^k does not converge. This proves Theorem 2.4.

The binding margin has positive slope and vanishes at τ_{up}, so the projection characterization fails just below τ_{up} and the bound is tight for this orbit. A Fourier optimization located the integer vectors; it is not part of the proof. Version 1.0 used a different orbit of the same form, in `certificates/exact_cycle_18869_10000_K128.json`, valid for $\tau \ge 897744153596994489/475778439151265168 \approx 1.8868954$ with residual 16383999960516622876672; version 1.0 formalized that orbit, and the Lean development now formalizes the present one (Section 10).

8 Scalar problems

This section proves Theorem 2.5. The divergent instance is immediate: with $C = D = \mathbb{R}$ and $f = g = 0$ both proximal maps are the identity, $(0, 0)$ is a saddle point, and (2.1) gives $x^{k+1} = z^k - u^k$, $z^{k+1} = x^{k+1} + u^k = z^k$ and $u^{k+1} = (1 - \tau)u^k$. From $z^0 = 0$, $u^0 = 1$ this is the stated orbit; it alternates for $\tau = 2$ and diverges for $\tau > 2$.

Now fix $0 < \tau < 2$ and centre at a saddle point (z_*, u_*) as in Section 4. Let P and Q be the proximal maps of $\widehat{f}$ and $\widehat{g}$. In the variables $x - z_*$, $z - z_*$, $u - u_*$, again written x, z, u, the iteration is $x^{k+1} = P(z^k - u^k)$, $z^{k+1} = Q(x^{k+1} + u^k)$, $u^{k+1} = u^k + \tau(x^{k+1} - z^{k+1})$, and $P(0) = Q(0) = 0$. In dimension one the inequality $(p - q)^2 \le (c - d)(p - q)$ of Lemma 3.1 says that P and Q are nondecreasing and 1-Lipschitz, and so is $B = I - Q$. Thus $Q(t)$ and $B(t)$ have the weak sign of t, and $|t| = |Q(t)| + |B(t)|$.

A lifted recurrence. Put $h = \tau - 1$, $a = |h| < 1$ and $\varepsilon = 1 - a > 0$. For $k \ge 1$ let $t_k = x^k + u^{k-1}$. Then $z^k = Q(t_k)$ and $u^k = u^{k-1} + \tau(t_k - u^{k-1} - Q(t_k)) = \tau B(t_k) - hu^{k-1}$. Substituting $x^{k+1} = P(Q(t_k) - u^k)$ shows $(t_{k+1}, u^{k+1}) = \Phi(t_k, u^k)$ for the continuous map

$$\Phi(t, u) = (s,\ \tau B(s) - hu), \qquad s = u + P(Q(t) - u).$$

The state $(z^k, u^k) = (Q(t_k), u^k)$ is recovered from (t_k, u^k); no proximal map is inverted. Since P fixes 0, is nondecreasing and is 1-Lipschitz, $P(Q(t) - u)$ lies between 0 and $Q(t) - u$, so s lies in the closed interval with endpoints $q = Q(t)$ and u.

A Lyapunov function. Define

$$L(t, u) = \varepsilon|Q(t)| + |B(t)| + |u - B(t)|.$$

It is continuous, $L(t, u) \ge \varepsilon|t|$ because $Q(t)$ and $B(t)$ have the same sign, and $L(t, u) \ge |u|$ by the triangle inequality. Write $b = B(t)$ and $J(c) = |c| + |u - c|$, so that $L(t, u) = \varepsilon|q| + J(b)$. Because the new multiplier satisfies $\tau B(s) - hu - B(s) = h(B(s) - u)$,

$$L(\Phi(t, u)) = \psi(s), \qquad \psi(r) = \varepsilon|Q(r)| + |B(r)| + a|B(r) - u|.$$

For $0 \le r < r'$, monotonicity of Q and B and the reverse triangle inequality give $\psi(r') - \psi(r) \ge \varepsilon(Q(r') - Q(r)) + (1 - a)(B(r') - B(r)) = \varepsilon(r' - r) > 0$; symmetrically ψ is strictly decreasing on $(-\infty, 0]$. Hence on a closed interval ψ attains its maximum only at an endpoint, and every interior value is strictly smaller than the larger endpoint value. At the endpoint u, the common signs give $\psi(u) = |Q(u)| + |B(u)| = |u| = J(0) \le J(b) = L(t, u) - \varepsilon|q|$. At the endpoint q, monotonicity puts $B(q)$ between 0 and b (if $t \ge 0$ then $0 \le q \le t$, so $0 \le B(q) \le B(t)$; the case $t < 0$ is symmetric), and $|Q(q)| = |q| - |B(q)|$, so $\psi(q) = \varepsilon|q| + aJ(B(q))$. Convexity of J and $J(0) \le J(b)$ give $J(B(q)) \le J(b)$, hence $\psi(q) \le L(t, u) - \varepsilon J(b)$. Therefore

$$L(\Phi(t, u)) \le L(t, u) \qquad \text{for all } (t, u) \in \mathbb{R}^2. \tag{8.1}$$

Equality. Suppose $L(\Phi(w)) = L(w)$ for $w = (t, u)$. Then $s \in \{q, u\}$ and the corresponding bound is an equality. If $s = q$ then $J(b) = 0$, so $u = b = 0$, $t = q$, $Q(q) = q$ and $P(q) = q$; thus $\Phi(w) = w$. Otherwise $s = u$, $q = 0$ and $P(-u) = 0$, and $\Phi(w) = (u, u')$ with $u' = B(u) - hQ(u)$. If moreover $L(\Phi^2(w)) = L(\Phi(w))$, apply the same alternatives at $\Phi(w)$: either $\Phi(w)$ is fixed, or $Q(u) = 0$. In the latter case $B(u) = u$ and $u' = u$, and $\Phi(u, u) = (u + P(-u), \tau u - hu) = (u, u)$. In both cases

$$L(w) = L(\Phi(w)) = L(\Phi^2(w)) \implies \Phi(\Phi(w)) = \Phi(w). \tag{8.2}$$

One equality is not enough: for $P = 0$, Q the proximal map of $\max(0, z)$ on $\mathbb{R}$, $\tau = 19/10$ and $w = (1, 14/5)$ one finds $\Phi(w) = (14/5, -31/50) \neq w$ and $L(\Phi(w)) = L(w) = 14/5$.

Convergence. Let $w_k = (t_k, u^k)$. By (8.1) and coercivity the sequence is bounded and $L(w_k)$ decreases to some ℓ. Choose $w_{k_j} \to w'$. Continuity gives $L(w') = L(\Phi(w')) = L(\Phi^2(w')) = \ell$, so $\bar{w} = \Phi(w')$ is fixed by (8.2) and is a cluster point, the limit of w_{k_j+1}. Write $\bar{w} = (\bar{t}, \bar{u})$ and $\bar{q} = Q(\bar{t})$. The second coordinate gives $\tau\bar{u} = \tau B(\bar{t})$, so $\bar{u} = B(\bar{t})$ and $\bar{t} = \bar{q} + \bar{u}$, and the first gives $P(\bar{q} - \bar{u}) = \bar{q}$; also $Q(\bar{q} + \bar{u}) = \bar{q}$. By Lemma 3.1, $(\bar{q}, \bar{u})$ is a saddle point of the centred problem, with $\bar{q}$ in both domains. Recentring at it replaces P, Q by $c \mapsto P(\bar{q} - \bar{u} + c) - \bar{q}$ and $c \mapsto Q(\bar{q} + \bar{u} + c) - \bar{q}$, which have the same properties, and translates the lifted variables by $\bar{w}$; the recurrence keeps its form. The recentred function $L_{\bar{w}}(t, u) = \varepsilon|Q(t) - \bar{q}| + |B(t) - \bar{u}| + |u - B(t)|$ is therefore nonincreasing along w_k, and it tends to 0 along w_{k_j+1}. Hence $L_{\bar{w}}(w_k) \to 0$, and coercivity gives $t_k \to \bar{t}$ and $u^k \to \bar{u}$. Then $z^k = Q(t_k) \to \bar{q}$ and $x^{k+1} = P(z^k - u^k) \to P(\bar{q} - \bar{u}) = \bar{q}$. The normals a^{k+1} and b^{k+1} of Section 5 converge to the saddle normals, so Lemma 3.2 gives the objective limits. Undoing the centring proves Theorem 2.5.

The argument uses the order of the real line through the common signs of $Q(t)$ and $B(t)$. It does not extend to higher dimension, where Theorem 2.4 shows that the range is smaller.

9 An earlier published claim

Gao, Li, Ma, Li, Sun, Maihemuti and Jin study relaxed ADMM for model-predictive control and, according to our reading of their Sections 2.2.3 and 4.3, assert convergence of the multiplier-only iteration (2.1) for $1 < \tau \leq 2$ [7]. With penalty one in consensus form, a saddle point (z_*, u_*), $r_k = x^k - z^k$ and $q_k = (u^k - u_*)^2/\tau + (z^k - z_*)^2$, their inequalities (45) and (47) specialize to

$$q_k - q_{k+1} \geq \tau(z^{k+1} - z^k)^2 + (\tau - 1)r_k^2 + (2 - \tau)r_{k+1}^2.$$

Take $E = \mathbb{R}$, $f(x) = x$, $g(z) = -z$ and $z^0 = u^0 = 0$. The saddle point is $(0, -1)$, the updates are $x^{k+1} = z^k - u^k - 1$, $z^{k+1} = z^k = 0$ and $u^{k+1} = (1 - \tau)u^k - \tau$. At $\tau = 3/2$ one finds $q_1 - q_2 = 1/8$ while the right side is $5/8$; at $\tau = 7/4$ the values are $9/64$ and $57/64$. At $\tau = 2$ the iteration is an actual two-cycle, u^k alternating between 0 and -2 and x^k between -1 and 1. This scalar example converges for $\tau < 2$, as every scalar instance does by Theorem 2.5, so it refutes only the displayed inequality; Theorem 2.4 refutes the asserted range for every $\tau \geq \tau_{\mathrm{up}}$. We could not re-open the full text of [7] while preparing this release, so the identification of their equations rests on an earlier reading recorded in the research records of this package.

10 Formal verification

The Lean project in `lean/` pins Lean 4.19.0 (commit `6caaee842e94`) and Mathlib revision `c44e0c8ee63ca166450922a373c7409c5d26b00b`, with every transitive dependency fixed in `lake-manifest.json`. It contains 301 proof modules and one audit module. The eleven central declarations are:

- `ADMMNearSharpAllSteps.initialized_admm_objective_converges`: Theorem 2.2 for $0 < \tau \leq 377359/200000$, with the trajectory defined by iterating the actual map (2.1) from an arbitrary initial state, `ConvexOn` hypotheses, proximal points that are global minimizers for every center, and an existing saddle point. It concludes convergence to a saddle point, global optimality of the limit, and convergence of the summed objective values.

- `ADMMNearSharpAllSteps.rankDeficient_original_admm_converges` and `ADMMNearSharpAllSteps.fullRank_original_admm_converges`: Theorem 2.3 for $0 < \tau \le 377359/200000$, with arbitrary sequences satisfying the original affine minimization and multiplier relations; the second theorem assumes continuous left inverses.
- `ADMMNearSharpCertificate.C`K`.certificate` for $K = 188$, 1886, 188628, 188632, 18865, 18867 and 1886795: the seven longer-memory certificates on $[7/4, 377359/200000]$ as data, with their identities, flow balance and positivity proved.
- `ADMM128Counterexample.counterexample`: Theorem 2.4 for every $\tau \ge \tau_{\mathrm{up}}$, including unique proximal points for every center, compactness and convexity of both polytopes, the saddle point, the orbit of the actual map, period 128, the constant residual 16383999999992680528112756654 08, and nonconvergence of z^k.

The modules fall into groups: proximal model, generic memory convergence and the classical and three-state branches (inherited unchanged from the earlier convergence package); the generic finite-memory certificate theory (`MemoryCertificate`, `DelayedMemory`, `FiniteMemoryADMM`, `OriginalMemoryADMM`); the Gram-certificate bridge (`Quadratic`, `Integer*`, `StandardRows`), with `IntegerLists` for the list-based checks of the version 1.1 data; the generated certificate data (`NearSharpData*`, `NearSharpCertificate*`); the cycle (`Cycle*`); and the final assembly (`NearSharpAllSteps`, `NearSharpRankDeficient`, `NearSharpFullRank`).

The verifier `verify_lean.py` checks the dependency pins and clean checkouts, compiles every local module in dependency order into a new object directory with one thread, and then submits the stored bodies of the eleven central declarations to the Lean kernel again under fresh names. Every central declaration depends only on `propext`, `Classical.choice` and `Quot.sound`. Integer facts are decided by the kernel; no `native_decide` is used. A fresh replay on 31 October 2026, compiling all 301 modules from source into a new object directory against the pinned Mathlib checkout with its published build cache and resubmitting the eleven central declaration bodies to the kernel, passed; its record is in `verification/`. The standard-library programs `checks/check_exact_interval.py` and `checks/exact_circulant_cycle.py` independently rebuild the certificate identities and the cycle margins with rational arithmetic.

Formal scope: individual objective-value limits (rather than their sum), the reduction from proper closed convex functions to the attainment hypothesis, and penalty rescaling in consensus form are conventional arguments given above. Theorem 2.5 is not formalized; it has the conventional proof of Section 8. The orbit of version 1.0 is no longer formalized separately, since the present orbit covers every step it covered.

11 Related work and novelty

The comparisons below concern the exact iteration (2.1).

Source	Relation to this paper
Fortin–Glowinski [1]	Convergence for $0 < \tau < (1+\sqrt{5})/2$; the classical range.
Tao–Yuan [2]	The range $(0, 2)$ for quadratic objectives; general convex objectives are left open. Theorem 2.5 gives the range $(0, 2)$ for all convex scalar problems.

Source	Relation to this paper
Gu–Yang [3]	The conventional energy can increase beyond the golden ratio. This obstructs one proof technique, not convergence; the energies here use several states.
Chen–Sun–Toh [4]	Larger steps under an assumed summability of multiplier increments, which is derived here.
Ma–Shu–Li–Wu [5]; Ma [6]	Larger steps under strong convexity with a rank condition, or with an added proximal term.
Gao et al. [7]	Asserts $1 < \tau \leq 2$ for general convex problems; see Section 9.
Upadhyaya et al. [8]; AutoLyap [9]	Automated quadratic Lyapunov analysis, including iterate histories. This is the methodology used here.
Nilsson–Åkerman–Giselsson [10]	Sharp results for an extended family of splitting methods whose relaxation also changes the primal arguments; a different iteration.
Bai–Rao [11]; Zhang [12]	Larger dual steps with primal corrections or a proximal term, or with linear or strongly convex objectives and changing penalties.
Five-cycle [15]	The earlier obstruction of this collection, from $\tau \geq 140111/70637$, with a strongly convex extension.
Monotone and competitive planar maps [13, 14]	Convergence theorems for planar maps under order-preservation, invariant-box or injectivity hypotheses. The scalar ADMM state map need not be monotone in its native coordinates or injective, and we did not derive Theorem 2.5 from these results; a derivation after a change of variables is not excluded.

No earlier theorem establishing universal convergence of (2.1) beyond the golden ratio for general convex problems, and no convex obstruction below the five-cycle range of this collection, was found in the sources we inspected; Gao et al. is an earlier assertion whose stated range is false. Version 1.1 refines the bracket of version 1.0 with the same methods. For Theorem 2.5 we found no earlier statement covering arbitrary convex scalar objectives with nonsmooth or flat proximal maps, every initial state and possibly many saddle points; the search was bounded, and several original sources on planar monotone maps could only be read in part. This is a provisional assessment. Unindexed or inaccessible work, and equivalent reformulations of other splitting methods, remain priority risks.

12 Limitations and open questions

The exact threshold τ_* in (1.1) is open. Theorems 2.4 and 2.5 show that the universal threshold depends on the dimension: it is 2 for scalar problems and at most τ_{up} in dimension 128. Since an instance can be embedded in any larger dimension, the threshold does not increase with the dimension; its values in intermediate dimensions, and the smallest dimension in which it is below 2, are open. The results concern finite dimension, exact subproblems, fixed τ and penalty, and two blocks; weak convergence in infinite dimension, inexact or varying steps, more blocks and rates are not addressed. The obstruction uses indicator objectives, so the thresholds for smooth or strongly convex subclasses may be larger. No practical recommendation for τ follows from the theorems.

13 Research process and AI disclosure

This is fully autonomous agentic research. Dev D. Goyal is responsible for publication. The work has not been externally peer reviewed, no expert validation is claimed, and the novelty assessment of Section 11 is provisional.

The research, the exact certificates, the Lean formalization and earlier reports were produced by autonomous AI research agents in earlier sessions. For version 1.0 the records do not name the model or system used. Three uploaded packages contributed to it: a convergence dossier for $0 < \tau \leq 7/4$, an interval package for $[3/2, 7/4]$ that also treated general linear constraints and examined an earlier published claim, and a package with longer-memory certificates and the periodic orbit in $\mathbb{R}^{128}$. The first two are superseded by this paper; their Lean modules are included unchanged where the final proof uses them.

Opus 5.5 by Anthropic used with Claude Code assessed and consolidated the three packages for version 1.0. It coordinated independent AI referee checks, which included new exact-arithmetic rechecks of the cycle certificate and of both longer-memory certificates, ran a fresh Lean replay, added the comparison with Gao et al. (2022) and the scope notes, and wrote the manuscript.

A further uploaded package contributed version 1.1: five rational interval certificates extending the convergence range from 943/500 to 377359/200000, a new integer orbit in $\mathbb{R}^{128}$ with a lower threshold, and the scalar theorem of Section 8 with its proof. Its records state that it used an autonomous GPT-6 Astra workflow, in which separate instances reviewed the scalar and bracket proofs after first reconstructing them from the statements alone. Opus 5.5 by Anthropic used with Claude Code converted the new certificates to the file format of this package without changing any number, rechecked them and the new orbit with the package's exact checkers, checked the scalar proof, extended the Lean formalization to the new certificates and the new orbit, ran the fresh Lean replay described in Section 10, and revised this manuscript. On 1 November 2026 it also revised the presentation (front matter, abstract, introduction and this section) without changing the mathematics.

Neither these internal AI-assisted reviews nor the AI referee checks are human peer review. Lean verification establishes only the formal statements described in Section 10; it does not establish novelty or priority. The remaining limitations are listed in Section 12.

References

[1] M. Fortin and R. Glowinski, *Augmented Lagrangian Methods: Applications to the Numerical Solution of Boundary-Value Problems*, Studies in Mathematics and its Applications 15, North-Holland, 1983.

[2] M. Tao and X. Yuan, On Glowinski's open question on the alternating direction method of multipliers, *J. Optim. Theory Appl.* 179 (2018), 163–196. `https://doi.org/10.1007/s10957-018-1338-x`.

[3] G. Gu and J. Yang, On the dual step length of the alternating direction method of multipliers, arXiv:2006.08309, 2020.

[4] L. Chen, D. Sun and K.-C. Toh, A note on the convergence of ADMM for linearly constrained convex optimization problems, *Comput. Optim. Appl.* 66 (2017), 327–343. `https://doi.org/10.1007/s10589-016-9864-7`.

[5] F. Ma, J. Shu, Y. Li and J. Wu, The dual step size of the alternating direction method can be larger than 1.618 when one function is strongly convex, *J. Ind. Manag. Optim.* 17 (2021), 1173–1185. https://doi.org/10.3934/jimo.2020016.

[6] F. Ma, Convergence study on the proximal alternating direction method with larger step size, *Numer. Algorithms* 85 (2020), 399–425. https://doi.org/10.1007/s11075-019-00819-2.

[7] J. Gao, S. E. Li, F. Ma, W. Li, H. Sun, M. Maihemuti and C. Jin, Accelerated convergence of time-splitting algorithm by relaxation method, *IET Control Theory Appl.* 16 (2022), 776–788. https://doi.org/10.1049/cth2.12270.

[8] M. Upadhyaya, S. Banert, A. B. Taylor and P. Giselsson, Automated tight Lyapunov analysis for first-order methods, *Math. Program.* 209 (2025), 133–170. https://doi.org/10.1007/s10107-024-02061-8.

[9] M. Upadhyaya, S. Das Gupta, A. B. Taylor, S. Banert and P. Giselsson, The AutoLyap software suite for computer-assisted Lyapunov analyses of first-order methods, arXiv:2506.24076, 2025.

[10] M. Nilsson, A. Åkerman and P. Giselsson, Extending Douglas–Rachford splitting for convex optimization, arXiv:2511.19637, 2025.

[11] J. Bai and S. Rao, A new insight on the prediction-correction framework with applications to several first-order methods, *Oper. Res. Lett.* 68 (2026), 107487. https://doi.org/10.1016/j.orl.2026.107487.

[12] T. Zhang, On the $O(1/K^2)$ ergodic convergence of ADMM with dual step size from 0 to 2, *J. Oper. Res. Soc. China* 14 (2026), 521–532. https://doi.org/10.1007/s40305-024-00568-7.

[13] M. W. Hirsch and H. L. Smith, Monotone maps: a review, *J. Difference Equ. Appl.* 11 (2005), 379–398. https://doi.org/10.1080/10236190412331335445.

[14] M. R. S. Kulenović and O. Merino, A global attractivity result for maps with invariant boxes, *Discrete Contin. Dyn. Syst. Ser. B* 6 (2006), 97–110. https://doi.org/10.3934/dcdsb.2006.6.97.

[15] D. D. Goyal, A rational five-cycle for two-block ADMM with multiplier step below two, preprint, 2026. https://github.com/dvgyl/math-and-theory/tree/main/papers/mathematics/admm-counterexample.